



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 697 38 587 T2** 2009.04.30

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 935 665 B1**

(21) Deutsches Aktenzeichen: **697 38 587.6**

(86) PCT-Aktenzeichen: **PCT/US97/17555**

(96) Europäisches Aktenzeichen: **97 910 730.7**

(87) PCT-Veröffentlichungs-Nr.: **WO 1998/014601**

(86) PCT-Anmeldetag: **30.09.1997**

(87) Veröffentlichungstag
der PCT-Anmeldung: **09.04.1998**

(97) Erstveröffentlichung durch das EPA: **18.08.1999**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **19.03.2008**

(47) Veröffentlichungstag im Patentblatt: **30.04.2009**

(51) Int Cl.⁸: **C12N 15/82** (2006.01)

C12N 9/10 (2006.01)

C12N 15/62 (2006.01)

C12Q 1/68 (2006.01)

A01H 5/00 (2006.01)

(30) Unionspriorität:

26855 P 30.09.1996 US

(73) Patentinhaber:

**BASF Plant Science GmbH, 67063 Ludwigshafen,
DE**

(84) Benannte Vertragsstaaten:

**AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LI,
LU, MC, NL, PT, SE**

(72) Erfinder:

**KEELING, Peter, Ames, IA 50014, US; GUAN,
Hanping, Ames, IA 50010, US**

(54) Bezeichnung: **EINKAPSELUNG VON POLYPEPTIDEN IN DIE STÄRKEMATRIX**

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

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

Beschreibung

QUERVERWEIS AUF VERWANDTE ANMELDUNGEN

[0001] Diese Anmeldung nimmt die Priorität der am 30. September 1996 eingereichten Patentanmeldung mit der Seriennummer 60/026,855 in Anspruch.

ALLGEMEINER STAND DER TECHNIK

Polysaccharidenzyme

[0002] Sowohl prokaryontische als auch eukaryontische Zellen verwenden Polysaccharidenzyme als Speicherstoff. In der prokaryontischen Zelle ist Glycogen der wichtigste Speicherstoff. Obwohl Glykogen eine Ähnlichkeit mit der Stärke, die in den meisten Gefäßpflanzen auftritt, aufweist, verfügt es über unterschiedliche Kettenlängen und Polymerisationsgrade. In manchen Pflanzen wird Stärke als wichtigstes Speicherpolysaccharid verwendet. Stärke wird in verschiedenen Geweben der stärkehaltigen Pflanze gespeichert. Stärke besteht in den meisten Fällen aus zwei Komponenten, nämlich einerseits Amylose und andererseits Amylopektin. Amylose besteht aus linear angeordneten Glycanen und Amylopektin aus verzweigten Glycanketten. Typische Stärke weist ein Verhältnis von 25% Amylose zu 75% Amylopektin auf. Unterschiede im Amylose:Amylopektin-Verhältnis in einer Pflanze können die Eigenschaften der Stärke beeinflussen. Außerdem weisen Stärken aus unterschiedlichen Pflanzen häufig unterschiedliche Eigenschaften auf. Es scheint, daß sich Maisstärke und Kartoffelstärke aufgrund des Vorhandenseins bzw. Fehlens von Phosphatgruppen unterscheiden. Die Eigenschaften von Stärke unterschiedlicher Pflanzen unterscheiden sich auf Grund von Mutationen, die in das pflanzliche Genom eingeführt worden sind. Mutierte Stärken sind gut aus Mais, Reis und Erbsen und dergleichen bekannt.

[0003] Die Veränderungen bei den Verzweigungen der Stärke oder dem Verhältnis der Stärkekomponenten führen zur unterschiedlichen Stärkeeigenschaft. Eine Eigenschaft der Stärke ist die Bildung von Stärkekörnern, die insbesondere in Blättern, Wurzeln, Knollen und Samen gebildet werden. Diese Körner werden während der Stärkesynthese gebildet. Gewisse Stärkesynthasen, insbesondere die stärkekornggebundene Stärkesynthase, die löslichen Stärkesynthasen und die Verzweigungsenzyme, sind Proteine, die innerhalb des Stärkekorns bei dessen Bildung "eingekapselt" werden.

[0004] Die Verwendung von cDNA-Klonen von tierischen und bakteriellen Glycogensynthasen sind in der internationalen Patentanmeldung mit der Veröffentlichungsnummer WO94/09144 beschrieben. Die Nukleotid- und Aminosäuresequenzen der Glycogensynthase sind aus der Literatur bekannt. So kann man die Nukleotidsequenz für das glgA/Gen aus *E. coli*, das für die Glycogensynthase codiert, aus der Datei GenBank/EMBL (SWISSPORT) unter der Zugangsnummer J02616 (Kumar et al., 1986, J. Biol. Chem., 261: 16256–16259) entnehmen. Strukturgene für das Glycogen-Biosyntheseenzym aus *E. coli* wurden auch von Okita et al. (1981, J Biol. Chem. 256(13): 6944–6952 kloniert.

[0005] Das Glycogensynthesestrukturgen glgA wurde von Leung et al. (1987, J. Bacteriol., 169(9): 4349–4354) aus *Salmonella thyphimurium* LT2 kloniert. Die Sequenzen der Glycogensynthase aus Skelettmuskel des Kaninchens (Zhang et al., 1989, FASEB J., 3: 2532–2536) und menschlichem Muskel (Browner et al., 1989, Proc. Natl. Acad. Sci. 86: 1443–1447) sind ebenfalls bekannt.

[0006] Über die Verwendung von cDNA-Klonen von pflanzlichen löslichen Stärkesynthasen ist berichtet worden. Die Aminosäuresequenzen der löslichen Stärkesynthase aus der Erbse mit den Isoformen I und II wurden von Dry et al. (1991, Plant Journal, 2: 193202) veröffentlicht. Die Aminosäuresequenz der löslichen Stärkesynthase aus Reis (soluble starch synthase = SSTs) wurde von Baba et al. (1993, Plant Physiology) beschrieben. Die letztgenannte Sequenz (Reis-SSTs) zitiert fälschlich die N-terminale Sequenz und ist daher irreführend. Dies beruht vermutlich auf einem Extraktionsfehler, bei dem ein Proteaseabbau stattgefunden hat, oder einer anderen inhärenten Labilität in dem extrahierten Enzym. Die richtige N-terminale Sequenz (die mit AELSR beginnt) liegt in der sogenannten Transitpeptidsequenz der Reis-SSTs vor.

[0007] Die Sequenz des Verzweigungszyms I aus Mais wurde von Baba et al., 1991, BBRC, 181: 8794, untersucht. Das Stärkeverzweigungsenzym II aus Maisendosperm wurde von Fisher und Shrabie (1993, Plant Physiol., 102: 10451046) untersucht. Über die Verwendung von cDNA-Klonen von pflanzlichen, bakteriellen und tierischen Verzweigungsenzymen ist berichtet worden. Die Nukleotid- und Aminosäuresequenzen für bakterielle Verzweigungsenzyme (BE) sind aus der Literatur bekannt. So haben zum Beispiel Kiel et al. das Ver-

zweigungsenzym glgB aus *Cyanobacterium synechococcus* PPC7942 (1989, *Gene* (Amst), 78(1): 918) und aus *Bacillus stearothermophilus* (Kiel et al., 1991, *Mol. Gen. Genet.*, 230(12): 136–144) kloniert. Die Gene *glc3* und *ghal* aus *S. cerevisiae* sind Allele und codieren für das Glycogenverzweigungsenzym (Rowen et al., 1992, *Mol. Cell Biol.* 12(1): 22–29). Matsumoto et al. haben an dem Glycogenverzweigungsenzym aus *Neurospora crassa* (1990, *J. Biochem.*, 170: 118–122) geforscht. Die GenBank/EMBL enthält auch Sequenzen für das glgB-Gen aus *E. coli*, das für Verzweigungsenzym codiert.

[0008] Die Stärkesynthase (EC 2.4.1.11) verlängert Stärkemoleküle und soll sowohl auf Amylose als auch auf Amylopektin einwirken. Eine Stärkesynthase (STS)-Aktivität findet sich in Assoziation sowohl mit den Grana als auch im Stroma des Plastiden. Die Kapazität des Enzyms gebundene Stärkesynthase für die Assoziation von Stärke ist gut bekannt. Es ist nun bekannt, daß verschiedene Enzyme, die an der Stärkebiosynthese beteiligt sind, unterschiedliche Bindungsaffinitäten, wie dies von Mu-Forster et al. (1996, *Plant Phys.* 111: 821–829) beschrieben wurde, aufweisen. Die Aktivität der stärkekorngelundenen Stärkesynthase (granule-bound starch synthase, GBSTS) ist stark mit dem Produkt des waxy-Gens korreliert (Shure et al., 1983, *Cell* 35: 225–233). Es wurde gezeigt, daß die Amylosesynthese in verschiedenen Arten wie Mais, Reis und Kartoffel von der Expression dieses Gens abhängt (Tsai, 1974, *Biochem Gen* 11: 83–96; Hovenkamp-Hermelink et al., 1987, *Theor. Appl. Gen* 75: 217–221). Visser et al. beschrieben die molekulare Klonierung und teilweise Charakterisierung des Gens für die stärkekorngelundene Stärkesynthase aus der Kartoffel (1989, *Plant Sci.* 64(2): 185–192). Visser et al. haben auch die Hemmung der Expression des Gens für die stärkekorngelundene Stärkesynthase in der Kartoffel durch Antisense-Konstrukte beschrieben (1991, *Mol. Gen. Genet.* 225(2): 289–296).

[0009] Nach den bahnbrechenden Arbeiten von Frydman und Cardini (Frydman und Cardini, 1964, *Biochem. Biophys. Res. Communications* 17: 407–411) sind die anderen STS-Enzyme aus löslicher Stärkesynthase bekannt geworden. In jüngster Zeit ist jedoch der Begriff "löslich" angesichts von Entdeckungen, daß diese Enzyme mit dem Korn assoziiert sind, jedoch auch in der löslichen Phase vorliegen, in Frage gestellt worden (Denyer et al., 1993, *Plant J.* 4: 191–198; Denyer et al., 1995, *Planta* 97: 57–62; Mu-Forster et al., 1996, *Plant Physiol.* 111: 821–829). Es wird allgemein angenommen, daß die Amylopektinbiosynthese die Wechselwirkung von löslichen Stärkesynthasen und Stärkeverzweigungsenzymen beinhaltet. Unterschiedliche Isoformen der löslichen Stärkesynthase sind in der Erbse (Denyer und Smith, 1992, *Planta* 186: 609–617; Dry et al., 1992, *Plant Journal*, 2: 193–202), in der Kartoffel (Edwards et al., 1995, *Plant Physiol.* 112: 89–97; Marshall et al., 1996, *Plant Cell* 8, 1121–1135) sowie in Reis (Baba et al., 1993, *Plant Physiol.* 103: 565–573) identifiziert und kloniert worden, während es scheint, daß die Gerste multiple Isoformen enthält, von denen einige mit dem Stärkeverzweigungsenzym assoziiert sind (Tynnela und Schulman, 1994, *Physiol. Plantarum* 89: 835–841). Eine Eigenschaft, die den STS-Klonen gemeinsam ist, ist das Vorhandensein einer KXGGLGDV-Konsensus Sequenz, von der angenommen wird, daß sie die ADP-Glc-Bindungsstelle des Enzyms darstellt (Furukawa et al., 1990, *J. Biol. Chem.* 265: 2086–2090; Furukawa et al., 1993, *J. Biol. Chem.* 268: 23837–23842).

[0010] In Mais sind zwei lösliche Formen der STS, die als Isoform I und Isoform II bekannt sind, identifiziert worden (Macdonald und Preiss, 1983, *Plant Physiol.* 73: 175–178; Boyer und Preiss, 1978, *Carb. Res.* 61: 331–334; Pollock und Preiss, 1980, *Arch Biochem. Biophys.* 204: 578–588; Macdonald und Preiss, 1985 *Plant Physiol.* 78: 849–852; Dang und Boyer, 1988, *Phytochemistry* 27: 1255–1259; Mu et al., 1994, *Plant J.* 6: 151–159), es sind jedoch keine dieser beiden Isoformen kloniert worden. In jüngster Zeit wurde die STSI-Aktivität des Maisendosperms mit einem 76 kDa großen Polypeptid, das sowohl in löslichen als auch in stärkekorngelundenen Fraktionen auftritt, korreliert (Mu et al., 1994, *Plant J.* 6: 151–159). Die Polypeptidentität der STSII ist nach wie vor unbekannt. Die STSI und II weisen unterschiedliche Enzymcharakteristika auf. Die STSI weist eine primerunabhängige Aktivität auf, während die STSII einen Glycogen-Primer erfordert, um den Glycosyltransfer zu katalysieren. Es wurde berichtet, daß die löslichen Stärkesynthasen einen hohen Stromkontrollkoeffizienten für die Stärkeablagerung (Jenner et al., 1993, *Aust. J. Plant Physiol.* 22: 703–709; Keeling et al., 1993, *Planta* 191: 342–348) und bei erhöhten Temperaturen ungewöhnliche kinetische Eigenschaften aufweisen (Keeling et al., 1995, *Aust. J. Plant Physiol.* 21: 807–827). Die jeweiligen Isoformen in Mais weisen sowohl bezüglich der Temperaturoptima als auch der Temperaturstabilität wesentliche Unterschiede auf.

[0011] Sequenzen der pflanzlichen Stärkesynthase (und der Glycogensynthase aus *E. coli*) beinhalten die Sequenz KTGGL, von der bekannt ist, daß sie die ADPG-Bindungsdomäne darstellt. Die Gene für ein beliebiges Stärkesynthaseprotein dieser Art können in erfindungsgemäßen Konstrukten eingesetzt werden.

[0012] Das Verzweigungsenzym [α 1,4DGlycan: α 1,4DGlycan 6D(α 1,4DGlycano)transferase (E.C.2.4.1.18)], das manchmal Q-Enzym genannt wird, wandelt Amylose in Amylopektin um. Ein Abschnitt einer α 1,4DGlycankette über bakterielle Verzweigungsenzymgene und pflanzliche Sequenzen ist berichtet worden (Reisendosperm: Nakamura et al., 1992, *Physiologia Plantarum*, 84: 329–335 und Nakamura und Yamanou-

chi, 1992, Plant Physiol., 99: 1265–1266; Erbse: Smith, 1998, Planta, 175: 270–279 und Bhattacharyya et al., 1989, J. Cell Biochem., Suppl. 13D: 331; Maisendosperm: Singh und Preiss, 1985, Plant Physiology, 79: 34–40; VosScherperkeuter et al., 1989 Plant Physiology, 90: 75–84, Kartoffel: Kossmann et al., 1991, Mol. Gen. Genet., 230(12): 39–44; Cassava: Salehuzzaman und Vissar, 1992, Plant Mol Biol, 20: 809–819).

[0013] Auf dem Gebiet der Polysaccharidenzyme existieren Berichte über Vektoren für die Beeinflussung der Modifikation im Stärkebiosyntheseweg von Pflanzen dadurch, daß man verschiedene Stärkesynthesegene in verschiedenen Pflanzenarten einsetzt. Es ist gut bekannt, daß einige dieser Polysaccharidenzyme an Cellulose oder Stärke bzw. Glycogen binden. Ein spezifisches typisches Beispiel für die Verwendung eines Polysaccharidenzyms zeigt die Verwendung von Glycogenbiosyntheseenzymen für die Modifikation von pflanzlicher Stärke. In dem an Shewmaker ausgegebene US-Patent Nr. 5,349,123 wird ein Vektor gelehrt, der DNA für die Bildung von Glycogenbiosyntheseenzyme innerhalb von pflanzlichen Zellen enthält. Dieses Patent bezieht sich spezifisch auf die Veränderungen in Kartoffelstärke aufgrund der Einführung dieser Enzyme. Über andere Stärkesynthesegene und ihre Verwendung ist ebenfalls berichtet worden.

Hybridpeptide (Fusionspeptide)

[0014] Hybridproteine (auch "Fusionsproteine" genannt) sind Polypeptidketten, die aus zwei oder mehr Proteinen, die miteinander zu einem einzigen Polypeptid fusioniert sind, bestehen. Häufig ist eines der Proteine ein Ligand, der an eine spezifische Rezeptorzelle bindet. Vektoren, die für Fusionspeptide codieren, werden in ersten Linie dazu verwendet, um heterologe Proteine durch Fermentation von Mikroorganismen zu produzieren. Die produzierten Fusionsproteine können dann mittels Affinitätschromatographie aufgereinigt werden. Der Bindungsteil von einem der Polypeptide wird dazu verwendet, um das Hybridpolypeptid an eine Affinitätsmatrix anzuheften. So können zum Beispiel Fusionsproteine mit beta-Galactosidase gebildet werden, die an eine Säule gebunden werden können. Dieses Verfahren ist für die Bildung von Virus-Antigenen eingesetzt worden.

[0015] Eine weitere Verwendung besteht darin, eines der Polypeptide des Hybridpolypeptids zu gewinnen. Für die Spaltung des fusionierten Peptids sind chemische und biologische Methoden bekannt. Für die Abspaltung der Peptide kann ein niedriger pH-Wert eingesetzt werden, wenn eine säurelabile Aspartyl-Prolin-Bindung zwischen den Peptiden eingesetzt werden kann und die Peptide nicht von der Säure angegriffen werden. Hormone sind mittels Cyanobromid gespalten worden. Weiterhin wurde über Spaltung durch ortsspezifische Proteolyse berichtet. Weitere Verfahren der Proteinaufreinigung wie die Ionenchromatographie sind durch die Verwendung von Polyarginin-Schwänzen verbessert worden, die die Gesamtbasizität des Proteins erhöhen, wodurch die Bindung an Ionenaustauschssäulen gefördert wird.

[0016] Verschiedene Patente haben Verbesserungen bei Verfahren zur Herstellung von Hybridpeptiden oder spezifischen Hybridpeptiden, die spezifischen Verwendungszwecken dienen, umrissen. Das an Pastan et al. ausgegebene US-Patent umreißt eine Verbesserung von Hybridproteinen. In diesem Patent wird über einen ringförmig permutierten Liganden als Teil des Hybridpeptids berichtet. Dieser Ligand weist eine Spezifität und gute Bindungsaffinität auf. Über eine weitere Verbesserung bei Hybridproteinen wird in dem an Kuliopulos ausgegebenen US-Patent Nr. 5,648,244 berichtet. Dieses Patent beschreibt ein Verfahren zur Herstellung eines Hybridpeptids mit einem Trägerpeptid. Diese Nukleinsäureregion erzeugt, wenn sie von einer Restriktionsendonuklease erkannt wird, ein nichtpalindromartiges überstehendes Ende mit 3 Basen. Dieses gestattet die Spaltung des Vektors.

[0017] Über ein Beispiel für ein spezifisch adressiertes Hybridprotein wird in dem US-Patent 5,643,756 berichtet. Dieses Patent berichtet über einen Vektor für die Expression von glycosylierten Proteinen in Zellen. Dieses Hybridprotein ist für eine Verwendung in der richtigen Immunreaktivität von HIV gp120 adaptiert. Die Isolation von gp120-Domänen, die stark glycosyliert sind, wird durch diesen Vektor, über den berichtet wurde, verbessert.

[0018] In den US-Patenten 5,202,247 und 5,137,819 werden Hybridproteine mit Polysaccharidbindungsdomänen sowie Verfahren und Zusammensetzungen für die Herstellung von Hybridproteinen, die zur Bindung an eine Polysaccharidmatrix fähig sind, diskutiert. Das US-Patent 5,202,247 lehrt spezifisch ein Hybridprotein, das eine Cellulasebindungsregion mit einem interessierenden Peptid verknüpft. In dem Patent wird beschrieben, daß das Hybridprotein nach Expression in einem bakteriellen Wirt durch Affinitätschromatographie an Cellulose aufgereinigt werden kann.

[0019] Die Entwicklung von gentechnischen Verfahren hat es ermöglicht, Gene von verschiedenen Organismen und Pflanzen in andere Organismen oder Pflanzen einzuführen. Obwohl Stärke früher mittels Transfor-

mation und Mutagenese verändert worden ist, besteht noch immer ein Bedarf an einer weiteren Modifikation von Stärke. Zu diesem Zweck sind Vektoren erwünscht, die die Verkapselung von gewünschten Aminosäuren oder Peptiden innerhalb der Stärke, spezifisch innerhalb der Stärkekörner, ermöglichen. Die entstandene Stärke ist modifiziert, und das Gewebe der Pflanze mit dem Vektor ist modifiziert.

DARSTELLUNG DER ERFINDUNG

[0020] Die vorliegende Erfindung stellt ein rekombinantes Nukleinsäurekonstrukt, das einen Promoter, der dahingehend adaptiert ist, daß er die Expression eines interessierenden Polypeptids in einem stärkehaltigen Gewebe einer Pflanze während der Stärkesynthese beeinflusst, eine Nukleinsäure, die für ein Transitpeptid codiert, das das interessierende Polypeptid zu einem Amyloplasten zu transportieren vermag, eine Nukleinsäure, die für eine stärkebindende Region codiert, eine Nukleinsäure, die für das interessierende Polypeptid codiert; sowie eine Terminatorsequenz umfaßt, bereit, wobei das Konstrukt die Expression eines Hybridpolypeptids, das die stärkebindende Region und das interessierende Polypeptid umfaßt, steuert, wobei das Hybridpolypeptid innerhalb von Stärke der Pflanze eingekapselt ist.

[0021] Die Erfindung stellt außerdem einen Expressionsvektor, der das rekombinante Nukleinsäurekonstrukt umfaßt; eine Pflanzenzelle, die mit dem Konstrukt transformiert ist und die das in einem Amyloplasten verkapselte Hybridpolypeptid zu exprimieren vermag; eine aus der Pflanzenzelle regenerierte Pflanze; einen Samen der Pflanze, der das Konstrukt umfaßt und das in Stärke verkapselte Hybridpolypeptid zu exprimieren vermag; sowie eine modifizierte Stärke, die das von dem Samen abstammende Hybridpolypeptid umfaßt, bereit.

[0022] Die vorliegende Erfindung stellt auch ein Hybridpolypeptid bereit, das von den erfindungsgemäßen rekombinanten Nukleinsäurekonstrukten codiert wird und das eine stärkebindende Region (SBR) von einem stärkebindenden Enzym in Fusion mit einem "Payload"-Polypeptid, das bezüglich dieser stärkebindenden Region nicht endogen ist, d. h. das nicht natürlich an die stärkebindende Region gebunden vorkommt, umfaßt. Das Hybridpolypeptid eignet sich zur Herstellung von modifizierten Stärken, die das "Payload"-Polypeptid umfaßt. Solche modifizierten Stärken können dazu eingesetzt werden, um mit gewissen Aminosäuren angereicherte Körnerfutter bereitzustellen. Solche modifizierten Stärken eignen sich auch für die Bereitstellung von Polypeptiden wie Hormonen und sonstigen Medikamenten (z. B. Insulin) in stärke verkapselter Form zwecks Resistenz gegenüber Abbau durch Magensäuren. Die Hybridpolypeptide eignen sich auch für die Herstellung der "Payload"-Polypeptide in leicht aufgereinigter Form. So können zum Beispiel solche Hybridpolypeptide, die in Körnern hergestellt werden, aus den modifizierten Stärken, mit denen sie assoziiert sind, durch fachbekannte Techniken isoliert und aufgereinigt werden.

[0023] Der Begriff "Polypeptid" bedeutet im vorliegenden Zusammenhang eine Vielzahl von identischen oder unterschiedlichen Aminosäuren und umfaßt auch die Proteine.

[0024] Der Begriff "Hybridpolypeptid" bedeutet ein Polypeptid, das aus Peptiden oder Polypeptiden von mindestens zwei verschiedenen Ausgangsmaterialien aufgebaut ist, z. B. eine stärkebindende Region eines stärkebindenden Enzyms in Fusion mit einem anderen Polypeptid wie einem Hormon, wobei mindestens zwei Komponententeile des Hybridpolypeptids nicht in der Natur miteinander fusioniert vorkommen.

[0025] Der Begriff "interessierendes Polypeptid", auch "Payload-Polypeptid" genannt, bedeutet ein Polypeptid, das bezüglich der stärkeverkapselnden Region, die in Kombination mit dieser Region exprimiert werden soll, nicht endogen ist, um eine modifizierte Stärke, die das "Payload"-Polypeptid enthält, zu exprimieren. Wird das "Payload"-Polypeptid zur Erhöhung des Aminosäuregehalts von bestimmten Aminosäuren in der modifizierten Stärke eingesetzt, so besteht es vorzugsweise aus nicht mehr als drei unterschiedlichen Arten von Aminosäuren aus der Gruppe bestehend aus: Ala, Arg, Asn, Asp, Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr und Val.

[0026] Das "Payload"-Polypeptid kann ein Hormon, z. B. Insulin, ein Wachstumsfaktor, z. B. Somatotropin, ein Antikörper, ein Enzym, ein Immunglobulin oder ein Farbstoff sein oder ein biologisch aktives Fragment der genannten Substanzen, wie dies in der Fachwelt bekannt ist. Solange das Polypeptid eine biologische Wirksamkeit aufweist, muß es sich dabei nicht um ein natürlich vorkommendes Polypeptid handeln, sondern es kann mutiert, verkürzt oder auf sonstige Art und Weise modifiziert sein. Solche biologisch aktiven Polypeptide können modifizierte Polypeptide, die nur biologisch aktive Abschnitte von biologisch aktiven Polypeptiden enthalten, sein. Sie können auch Aminosäuresequenzen mit Homologie zu natürlich vorkommenden biologisch aktiven Aminosäuresequenzen (vorzugsweise mindestens ungefähr 75% Homologie), die die biologische Wirksamkeit beibehalten, sein.

[0027] Die stärkebindende Region des Hybridpolypeptids kann eine stärkebindende Region eines beliebigen fachbekannten stärkebindenden Enzyms sein, z. B. eines Enzyms, das aus der Gruppe bestehend aus der löslichen Stärkesynthase I, der löslichen Stärkesynthase II, der löslichen Stärkesynthase III, der stärkekorngelinkten Stärkesynthase, des Verzweigungsenzyms I, des Verzweigungsenzyms IIa, des Verzweigungsenzyms IIb und von Glycoamylasepolypeptiden ausgewählt ist.

[0028] Soll das Hybridpolypeptid eingesetzt werden, um das "Payload"-Polypeptid in reiner oder teilweiser aufgereinigter Form herzustellen, so umfaßt das Hybridpolypeptid vorzugsweise eine Spaltstelle zwischen der stärkebindenden Region und dem "Payload"-Polypeptid. Das Verfahren zur Isolation des aufgereinigten "Payload"-Polypeptids beinhaltet in diesem Fall den Schritt, daß man das Hybridpolypeptid mit einem für diese Spaltstelle spezifischen Spaltungsmittel in Kontakt bringt.

[0029] Die vorliegende Erfindung stellt auch rekombinante Nukleinsäure (RNA oder DNA)-Moleküle, die für die Hybridpolypeptide codieren, bereit. Solche rekombinanten Nukleinsäuremoleküle umfassen vorzugsweise Kontrollsequenzen, die an die Expression des Hybridpolypeptids in dem ausgewählten Wirt adaptiert sind. Der Begriff "Kontrollsequenzen" beinhaltet Promoter, Introns, bevorzugte Codonsequenzen für den jeweiligen Wirtsorganismus sowie andere fachbekannte Sequenzen, um die Expression von DNA oder RNA in bestimmten Wirten zu beeinflussen. Die Nukleinsäuresequenz, die für die Stärkeverkapselung der Region und das "Payload"-Polypeptid codieren, können natürlich vorkommende Nukleinsäuresequenzen oder biologisch aktive Fragmente davon oder auch biologisch aktive Sequenzen mit Homologie zu solchen Sequenzen, vorzugsweise mindestens ungefähr 75% Homologie zu solchen Sequenzen, sein.

[0030] Zu den Wirtsorganismen zählen Bakterien, Pflanzen und Tiere. Bevorzugte Wirte sind Pflanzen. Sowohl einkeimblättrige Pflanzen (monokotyle Pflanzen) als auch zweikeimblättrige Pflanzen (dikotyle Pflanzen) sind geeignete Wirte für die Expression der erfindungsgemäßen Hybridpolypeptide.

[0031] Die vorliegende Erfindung stellt auch Expressionsvektoren bereit, die die Nukleinsäuren, die für die erfindungsgemäßen Hybridproteine codieren, umfassen. Diese Expressionsvektoren werden für das Hineintransformieren der Nukleinsäuren in Wirtsorganismen eingesetzt und können auch Sequenzen umfassen, die die Expression der Nukleinsäure in dem Wirtsorganismus unterstützen. Die Expressionsvektoren können Plasmide, modifizierte Viren oder DNA- oder RNA-Moleküle oder sonstige Vektoren, die sich für die fachbekannten Transformationssysteme eignen, sein.

[0032] Gemäß dem erfindungsgemäßen Verfahren werden transformierte Pflanzenzellen erzeugt, die die rekombinanten Nukleinsäuremoleküle, die die erfindungsgemäßen Hybridpolypeptide zu exprimieren vermögen, umfassen. Es kann sich dabei um Pflanzenzellen handeln, die in Pflanzen regeneriert werden können, von denen das Hybridpolypeptid geerntet werden kann, oder es können solche Pflanzenzellen zu fruchtbaren Pflanzen mit Samen, die die für das Hybridpolypeptid codierenden Nukleinsäuren enthalten, regeneriert werden. In einer bevorzugten Ausführungsform enthalten solche Samen modifizierte Stärke, die das "Payload"-Polypeptid umfaßt.

[0033] Der Begriff "modifizierte Stärke" bedeutet, daß die natürlich vorkommende Stärke dahingehend modifiziert wurde, daß sie das "Payload"-Polypeptid umfaßt.

[0034] Bevorzugte rekombinante erfindungsgemäße Nukleinsäuremoleküle umfassen DNA, die für stärkeverkapselte Regionen codieren, die aus den in den Tabellen angegebenen stärke-synthetisierenden Gensequenzen ausgewählt sind.

[0035] Bevorzugte erfindungsgemäße Plasmide sind an die Verwendung mit spezifischen Wirten angepaßt. Plasmide, die einen Promoter, eine Plastidenzielsteuerungssequenz, eine für eine stärkeverkapselnde Region codierende Nukleinsäuresequenz sowie eine Terminatorsequenz umfassen, werden hier bereitgestellt. Solche Plasmide eignen sich für die Insertion von DNA-Sequenzen, die für "Payload"-Polypeptide und stärkeverkapselnde Regionen codieren, zwecks Expression in ausgewählten Wirten.

[0036] Plasmide der vorliegenden Erfindung können gewünschtenfalls einen Spacer bzw. eine Linker-Einheit in der Nähe der Fusionsstelle zwischen den für die SBR codierenden Nukleinsäuren und den für das "Payload"-Polypeptid codierenden Nukleinsäuren beinhalten. Die vorliegende Erfindung beinhaltet Plasmide, die Promoter umfassen, die für pflanzliche Wirte adaptiert sind. Solche Promoter können auch spezifisch für die Expression in einkeimblättrigen oder in zweikeimblättrigen Pflanzen adaptiert sein.

[0037] Ein Verfahren zur Bildung von erfindungsgemäßer peptidmodifizierter Stärke beinhaltet die folgenden Schritte: Bereitstellen eines Plasmids mit einem Promoter, der mit einer Nukleinsäuresequenz, die für eine stärkebindende Region codiert, assoziiert ist, wobei die für die stärkebindende Region codierende Nukleinsäuresequenz mit einer Nukleinsäureregion, die für ein "Payload"-Polypeptid codiert, verbunden ist, sowie Transformieren eines Wirts mit dem Plasmid, wodurch der Wirt peptidmodifizierte Stärke exprimiert.

[0038] Die vorliegende Erfindung umfaßt weiterhin stärkehaltige Körner umfassend: einen Embryo, Nährgewebe, sowie modifizierte Stärkekörner, in denen ein Protein verkapselt ist, das bezüglich der Stärkekörner des Korns, die nicht modifiziert sind, nicht endogen ist. Solche stärkehaltigen Körner können Körner sein, in denen der Embryo ein Maisembryo, ein Reisembryo oder ein Weizenembryo ist.

KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0039] [Fig. 1a](#) zeigt das Plasmid pEXS114, das das synthetische GFP (Green Fluorescent Protein) in pBSK der Fa. Stratagene subkloniert enthält.

[0040] [Fig. 1b](#) zeigt das Plasmid pEXS115.

[0041] [Fig. 2a](#) zeigt das Waxy-Gen mit Restriktionsstellen, das in ein im Handel erhältliches Plasmid subkloniert wurde.

[0042] [Fig. 2b](#) zeigt das Plasmid pET-21 A, das im Handel von der Fa. Novagen erhältlich ist, und in das das GFP-Fragment aus pEXS115 subkloniert wurde.

[0043] [Fig. 3a](#) zeigt das in pEXSWX subklonierte Plasmid pEXS114, sowie die GFP-FLWX-Karte.

[0044] [Fig. 3b](#) zeigt das Plasmid GFP-Bam HIWX.

[0045] [Fig. 4](#) zeigt das SGFP-Fragment von pEXS115, das in pEXSWX subkloniert wurde, sowie die GFP-Nco-WX-Karte.

[0046] [Fig. 5](#) zeigt eine lineare Darstellung eines für einkeimblättrige Pflanzen adoptierten Plasmids.

[0047] [Fig. 6](#) zeigt das Plasmid pEXS52.

[0048] [Fig. 7](#) zeigt die sechs Einbringungsplasmide, die für die Bildung von pEXS51 und pEX560 eingesetzt wurden.

[0049] [Fig. 7a](#) zeigt pEXS adh1. [Fig. 7b](#) zeigt pEXS adh1-nos3'.

[0050] [Fig. 7c](#) zeigt pEXS33. [Fig. 7d](#) zeigt pEXS10zp.

[0051] [Fig. 7e](#) zeigt pEXS10zp-adh1. [Fig. 7f](#) zeigt pEXS10zp-adh1-nos3'.

[0052] [Fig. 8a](#) und [Fig. 8b](#) zeigen die Plasmide pEXS50 bzw. pEXS51, die das MS-SIII-Gen, bei dem es sich um ein Gen für stärkelösliche Synthase handelt, enthalten.

[0053] [Fig. 9a](#) zeigt das Plasmid pEXS60, das das in pEXS50 gezeigte Intron nicht enthält, und

[0054] [Fig. 9b](#) zeigt das Plasmid pEXS61, das das in pEXS60 gezeigte Intron nicht enthält.

DETAILLIERTE BESCHREIBUNG

[0055] Die vorliegende Erfindung stellt, grob gesagt, ein Hybridpolypeptid, ein Verfahren zur Herstellung eines Hybridpolypeptids sowie Nukleinsäuren, die für das Hybridpolypeptid codieren, bereit. Ein Hybridpolypeptid besteht aus zwei oder mehr Untereinheiten, die gemeinsam zu einer Peptidkette fusioniert sind. Die Untereinheiten können Aminosäuren oder Peptide oder Polypeptide sein. Eine der Untereinheiten ist eine stärkebindende Region. Hybridpolypeptide können also an Stärkekörner, die von Pflanzen, die die Hybridpolypeptide exprimieren, produziert werden, adressiert werden.

[0056] Ein Verfahren zur Erzeugung der Hybridpolypeptide innerhalb von Zellen beinhaltet die Herstellung eines DNA-Konstrukts, das mindestens ein DNA-Fragment umfaßt, das für eine Sequenz codiert, die dahingehend fungiert, daß sie das Expressionsprodukt der verknüpften DNA in ein Stärkekorn bindet, wobei dieses Fragment mit einer DNA-Sequenz, die für das interessierende Polypeptid codiert, ligiert ist. Dieses Konstrukt wird innerhalb einer pflanzlichen Zelle exprimiert. Das Hybridpolypeptid kann zur Herstellung von aufgereinigtem Protein oder zur Immobilisierung eines interessierenden Proteins innerhalb des Schutzes eines Stärkekorns oder zur Herstellung von einem Samenkorn, das heterologe Aminosäuren oder Peptide enthält, eingesetzt werden.

[0057] Das erfindungsgemäße Hybridpolypeptid weist drei Regionen auf.

| | | |
|---------------------------------|-------------------------|-----------------------------|
| Interessierendes Polypeptid (X) | Zentrale Position (ZP)* | Stärkebindende Region (SBR) |
|---------------------------------|-------------------------|-----------------------------|

X bedeutete eine beliebige Aminosäure oder ein beliebiges interessierendes Peptid.

* Optionale Komponente

[0058] Das Gen für X kann in 5'- oder 3'-Stellung innerhalb des unten beschriebenen DNA-Konstrukts angeordnet werden.

[0059] Die ZP ist eine zentrale Position, bei der es sich um eine Abgangsstelle, eine Spaltstelle oder einen Spacer handeln kann, wie dies fachbekannt ist. Eine Spaltstelle wird von einem Spaltungsenzym erkannt. Ein Spaltungsenzym ist ein Enzym, das Peptide an einer bestimmten Stelle spaltet. Zu Beispielen von Chemikalien und Enzymen, die für die Spaltung von Peptiden eingesetzt worden sind, zählen Thrombin, Trypsin, Cyanobromid, Ameisensäure, Hydroxylamin, Collagenase und Alasubtilisin. Ein Spacer ist ein Peptid, das die Peptide, die das Hybridpolypeptid umfassen, verbindet. Üblicherweise hat es keine spezifische Aktivität außer daß es die Peptide miteinander verbindet, oder eine gewisse Minimaldistanz garantiert oder die Faltung, Ladung oder Wasseraufnahme des Proteins beeinflusst. Spacer können beliebige Peptidsequenzen sein, die die biologische Aktivität des Hybridpolypeptids nicht stören.

[0060] Die stärkebindende Region (SBR) ist diejenige Region des vorliegenden Polypeptids, die eine Bindungsaffinität für Stärke aufweist. Üblicherweise ist die SPR aus der Gruppe bestehend aus Peptiden umfassend stärkebindende Regionen von Stärkesynthasen und Verzweigungsenzyme von Pflanzen ausgewählt, kann jedoch stärker bindende Domänen aus anderem Ausgangsmaterial wie Glycoamylase und dergleichen beinhalten. In den bevorzugten Ausführungsformen der Erfindung beinhaltet die SBR Peptidprodukte von Genen, die natürlich im Stärkebiosyntheseweg vorkommen. Diese Untergruppe von bevorzugten SBRs wird als stärkebildende Bindungsregionen (SFBR) definiert. Eine weitere Untergruppe von im vorliegenden Zusammenhang bevorzugten SBRs ist die spezifische stärkebindende Region (SSBR) von den spezifischen Enzymen Stärkesynthase (STS), stärkekorngelundene Stärkesynthase (GBSTS) und Verzweigungsenzyme (BE) von stärkehaltigen Pflanzen. Das am stärksten bevorzugte Genprodukt dieser Gruppe ist die GBSTS. Zusätzlich sind die Stärkesynthase I und das Verzweigungsenzym II nützliche Genprodukte. Vorzugsweise sind die SBR (sowie alle oben diskutierten Untergruppen) verkürzte Versionen des Vollängen-Stärkesyntheseenzymgens so daß der verkürzte Abschnitt die stärkebindende Region beinhaltet.

[0061] Das DNA-Konstrukt für die Expression des Hybridpolypeptids innerhalb des Wirts lautet grob ausgedrückt folgendermaßen:

| | | | | | |
|----------|---------|----------------------------|---|-----|------------|
| Promoter | Intron* | Transitpeptid-Codierregion | X | SBR | Terminator |
|----------|---------|----------------------------|---|-----|------------|

* Optionale Komponente. Andere optionale Komponenten können ebenfalls eingesetzt werden.

[0062] Wie in der Fachwelt bekannt ist, ist ein Promoter eine DNA-Region, die die Transkription kontrolliert. Unterschiedliche Arten von Promotern werden für unterschiedliche Werte ausgewählt. Lac- und T7-Promoter funktionieren gut in Prokaryonten, der 35S-CaMV-Promoter funktioniert gut in zweikeimblättrigen Pflanzen, und der Polyubiquitin-Promoter funktioniert gut in vielen einkeimblättrigen Pflanzen. Es können im Rahmen der Erfindung beliebig viele verschiedene fachbekannte Promoter eingesetzt werden.

[0063] Wie ebenfalls fachbekannt ist, ist ein Intron eine Nukleotidsequenz in einem Gen, die nicht für das Genprodukt codiert. Ein Beispiel für ein Intron, das häufig die Expression in einkeimblättrigen Pflanzen ver-

stärkt, ist das Adh1-Intron. Diese Komponente des Konstrukts ist optional.

[0064] Die Transitpeptid-Codierregion ist eine Nukleotidsequenz, die für den Transport des Proteins in Organellen wie Plastiden codiert. Vorzugsweise wird ein Transitpeptid gewählt, das mit dem Wirt, in dem das Transitpeptid eingesetzt wird, kompatibel ist und erkannt wird. In der vorliegenden Erfindung handelt es sich bei den Plastiden um den Amyloplasten.

[0065] Das Hybridpolypeptid ist im Amyloplasten in Zellen wie Pflanzenzellen, die Stärke in Amyloplasten synthetisieren und speichern, lokalisiert.

[0066] Ein Terminator ist eine DNA-Sequenz, die die Transkription abbricht.

[0067] X ist die Codierregion für das interessierende Polypeptid, das ein beliebiges Polypeptid oder beliebige Ketten von Aminosäuren sein kann. Sie kann ein nützliches Fragment einer bzw. bis zu die gesamte Sequenz eines bekannten Polypeptids umfassen. Das interessierende Polypeptid kann ein Polypeptid, ein Fragment davon oder ein biologisch aktives Protein, bei dem es sich um ein Enzym, ein Hormon, einen Wachstumsfaktor, ein Immunglobulin, einen Farbstoff usw. handelt, sein. Zu Beispielen für einige der interessierenden Polypeptide, die in der vorliegenden Erfindung eingesetzt werden können, zählen Prolactin (PRL), Serumalbumin, Wachstumsfaktoren und Wachstumshormone, d. h. Somatotropin, was jedoch keine Einschränkung darstellen soll. Zu Serumalbuminen zählen Rinder-, Schaf-, Pferde-, Vogel- und Human-Serumalbumin. Zu Wachstumsfaktoren zählen der "Epidermal Growth Factor (EGF)", der "Insulin-like Growth Factor" I (IGF-I), der "Insulin-like Growth Factor" II (IGF-II), der Fibroblasten-Wachstumsfaktor (FGF), der transformierende Wachstumsfaktor alpha (TGF-Alpha), der transformierende Wachstumsfaktor beta (TGF-Beta), der Nerven-Wachstumsfaktor (NGF), der "plateletderived" Wachstumsfaktor (PDGF) sowie die rekombinanten menschlichen "Insulin-like Growth Factors" I (rHuIGF-I) und II (rHuIGF-II). Zu Somatotropinen, die in der Durchführung der vorliegenden Erfindung eingesetzt werden können, zählen Rinder-, Schweine-, Schaf-, Pferde-, Vogel- und Human-Somatotropin, was jedoch keine Einschränkung darstellen soll. Schweine-Somatotropin beinhaltet delta-7 rekombinantes Schweine-Somatotropin, wie es in der europäischen Patentanmeldung mit der Veröffentlichungsnummer 104,920 (Biogen) beschrieben und beansprucht wird. Bevorzugte interessierende Polypeptide sind Somatotropin, Insulin-A und -B-Ketten, Calcitonin, beta-Endorphin, Urogastron, beta-Globin, Myoglobin, das menschliche Wachstumshormon, Angiotensin, Proline, Proteasen, beta-Galactosidasen und Cellulasen.

[0068] Das Hybridpolypeptid, die SBR-Region und die interessierenden Polypeptide können auch fachbekannte posttranslationelle Modifikationen wie Glycosylierung, Acylierung und sonstige Modifikationen, die die gewünschte Aktivität des Polypeptids nicht stören, beinhalten.

Entwicklung eines Hybridpolypeptids

[0069] Die SBR-Region liegt in Genen, die an der Stärkesynthese beteiligt sind, vor. Zu Verfahren zur Isolation solcher Gene zählen das Screening von genomischen DNA-Bibliotheken und von cDNA-Bibliotheken. Gene können durch Ligation, mutagene Mittel, Verdau, Restriktion und sonstige Vorgänge dieser Art wie sie zum Beispiel in Maniatis et al., Molecular Cloning, Cold Spring Harbor Labs, Cold Springs Harbor, N. Y. umrissen sind, geschnitten und verändert werden. Ausgezeichnete Ausgangsmaterialien für die Gewinnung der SBR-Region beinhalten zum Beispiel die folgenden, was jedoch keine Einschränkung darstellen soll: Die Stärkesynthasen I, II, III, IV; die Verzweigungsenzyme I, IIA und B sowie die stärkekorngelundene Stärkesynthase (GBSTS). Diese Gene liegen in stärkehaltigen Pflanzen wie Reis, Mais, Erbsen, Kartoffeln, Weizen und dergleichen vor. Mit einer SBR-Sonde aus genomischer DNA oder cDNA oder mRNA oder Antikörpern gegen die SBR können nützliche Gene für die Klonierung isoliert und identifiziert werden. Die für Stärkeenzym codierenden Sequenzen können modifiziert werden, solange die Modifikationen nicht die Fähigkeit der SBR-Region, assoziierte Polypeptide zu verkapseln, stören.

[0070] Werden Gene, die für Proteine codieren, die in das Stärkekorn verkapselt werden, lokalisiert, so können, wie dies in der Fachwelt bekannt ist, verschiedene Ansätze für die Isolation der SBR eingesetzt werden. Ein Verfahren besteht darin, das Gen an unterschiedlichen Stellen mit Restriktionsenzymen zu schneiden, wobei Abschnitte vom N-terminalen Ende deletiert werden und das erhaltene Protein exprimieren gelassen wird. Das exprimierte verkürzte Protein wird dann an einem Stärkegel aufgetrennt, um die Assoziations- und Dissoziationskonstante des verbleibenden Proteins auszuwerten. Fachbekannte Markergene, z. B. das "Green Fluorescent Protein"-Gen, können an das verkürzte Protein angefügt werden und zur Bestimmung des Vorliegens des Markergens im Stärkekorn verwendet werden.

[0071] Sobald die SBR-Gensequenzregion isoliert worden ist, kann sie in der Herstellung der Genfragmentsequenz, die das in Stärke verkapselte interessierende Polypeptid exprimieren wird, eingesetzt werden. Die SBR-Gensequenz und die Gensequenz, die für das Polypeptid codiert, können miteinander verknüpft werden. Die erhaltene Fusions-DNA kann dann in verschiedene Vektorkonstrukte zwecks Expression in verschiedenen Wirten eingeführt werden. Die erfindungsgemäßen Wirte bilden Stärkekörner in Amyloplasten, die SBR kann jedoch leicht in bakteriellen Wirten wie *E. coli* ausgetestet werden.

[0072] Die für das interessierende Polypeptid codierende Nukleinsäuresequenz kann von DNA, RNA, genomischer DNA, cDNA, mRNA abgeleitet werden oder kann ganz oder teilweise synthetisch hergestellt werden. Die Sequenz des interessierenden Polypeptids kann dahingehend manipuliert werden, daß sie Mutationen enthält, so daß es sich bei dem erzeugten Protein um ein neues, mutiertes Protein handelt, solange die biologische Funktion beibehalten wird.

[0073] Wenn das interessierende Polypeptid, das für die Nukleinsäuresequenz codiert, mit der für SBR codierenden Sequenz ligiert ist, dann ist die Gensequenz für das interessierende Polypeptid vorzugsweise mit dem Ende der SBR-Sequenz, die für den N-Terminus codiert, verbunden. Obwohl das N-terminale Ende bevorzugt ist, scheint es, daß es für die Erfindung nicht kritisch ist, ob das interessierende Polypeptid mit dem N-terminalen Ende oder dem C-terminalen Ende der SBR ligiert ist. Es ist klar, daß das Verfahren zur Bildung der erfindungsgemäßen rekombinanten Nukleinsäuremoleküle, egal, ob synthetisch oder mittels Klonierung und Ligation, für die vorliegende Erfindung nicht kritisch ist.

[0074] Die Zentralregion des Hybridpolypeptids ist optional. Für manche Anwendungen der vorliegenden Erfindung kann es sehr nützlich sein, DNA, die für eine geeignete Proteasespaltstelle in dieser Region codiert, in das rekombinante Nukleinsäuremolekül, das für die Expression des Hybridpolypeptids verwendet wird, einzubrennen. Es kann jedoch auch sehr nützlich sein, DNA, die für eine Aminosäuresequenz, die pH-sensitiv ist, codiert, einzuführen, um die Zentralregion zu bilden.

[0075] Besteht der Einsatz der vorliegenden Erfindung darin, ein reines Protein zu entwickeln, das aus dem Stärkekorn durch eine Protease oder dergleichen extrahiert und freigesetzt werden kann, so ist eine Proteasespaltstelle nützlich. Weiterhin kann, wenn das Protein in einem Tier verdaut werden soll, eine Proteasespaltstelle nützlich sein, um die Enzyme im Verdauungstrakt des Tiers bei der Freisetzung des Proteins aus der Stärke zu unterstützen. Bei anderen Anwendungen und bei vielen Einsätzen im Zusammenhang mit der Verdauung wäre die Spaltstelle überflüssig.

[0076] Die Zentralregionsstelle kann einen Spacer umfassen. Ein "Spacer" bezieht sich auf ein Peptid, das das Protein, welches ein Hybridpolypeptid umfaßt, verbindet. Üblicherweise hat es keine spezifische Aktivität, außer, die Proteine miteinander zu verbinden, um eine Minimaldistanz zu gewährleisten oder um die Faltung, Ladung bzw. Hydrophobie oder Hydrophilie des Hybridpolypeptids zu beeinflussen.

Konstruktentwicklung

[0077] Sobald die ligierte DNA, die für das Hybridpolypeptid codiert, gebildet ist, werden Klonierungsvektoren oder Klonierungsplasmide hergestellt, die zum Übertragen der DNA auf einen Wirt zwecks Expression der Hybridpolypeptide fähig sind. Die rekombinante Nukleinsäuresequenz der vorliegenden Erfindung wird in einen geeigneten Klonierungsvektor bzw. ein geeignetes Klonierungsplamid insertiert. Für die vorliegende Erfindung ist der bevorzugte Wirt ein Wirt, der Stärkekörner erzeugt. Es können jedoch auch bakterielle Wirte eingesetzt werden. Besonders eignen sich bakterielle Wirte, die dahingehend transformiert worden sind, daß sie einige oder alle der Stärkesynthesegene einer Pflanze enthalten. Dem Durchschnittsfachmann ist klar, daß das Plasmid an den Wirt angepaßt ist. So beinhalten zum Beispiel in einem bakteriellen Wirt Promoter, die die Transkription regulieren, *lac*, *TAC*, *trp* und dergleichen. Zusätzlich würde DNA, die für ein Transitpeptid codiert, höchstwahrscheinlich nicht verwendet werden, und ein Sekretions-Leader, der stromaufwärts des Strukturgens liegt, kann dazu verwendet werden, das Polypeptid ins Medium zu befördern. Alternativ dazu wird das Produkt im Wirt zurückgehalten und der Wirt wird lysiert und das Produkt mittels Stärkeextraktionsverfahren oder dadurch, daß man das Material an eine Stärkematrix (oder eine stärkeartige Matrix wie Amylose oder Amylopektin, Glycogen oder dergleichen) bindet, um das Produkt zu extrahieren, isoliert und aufgereinigt.

[0078] Der erfindungsgemäße Wirt ist eine Pflanze, und das Plasmid ist daher dahingehend adoptiert, daß es sich für eine Pflanze eignet. Das Plasmid sollte einen Promoter enthalten, vorzugsweise einen Promoter, der dahingehend adaptiert ist, daß er die Expression des Proteins in das stärkehaltige Gewebe in der Pflanze adressiert. Der Promoter kann für verschiedene Gewebe wie Samen, Wurzeln, Knollen und dergleichen spe-

zifisch sein, oder es kann sich um einen konstitutiven Promoter für die Genexpression in allen Geweben der Pflanze handeln. Zu gut bekannten Promotern zählen der 10 kD große Zein(Mais)-Promoter, der CAB-Promoter, Patastin-, 35S und 19S-Blumenkohlmosaikvirus-Promoter (die bei zweikeimblättrigen Pflanzen sehr nützlich sind), der Polyubiquitin-Promoter (der sich für einkeimblättrige Pflanzen eignet), sowie deren Verbesserungen und Modifikationen, wie sie in der Fachwelt bekannt sind.

[0079] Der Klonierungsvektor enthält Codiersequenzen für ein Transitpeptid, um das Plasmid an die richtige Stelle zu steuern. Beispiele für Sequenzen, die für Transitpeptide codieren, sind in den Sequenztabelle dargestellt. Es können Codiersequenzen für andere Transitpeptide eingesetzt werden. Bevorzugt sind Transitpeptide, die natürlich in dem zu verwendenden Wirt vorkommen. Für Mais bevorzugte Regionen, die für Transitpeptide codieren, sind in den Tabellen und ihren Abbildungen dargestellt. Zweck des Transitpeptids ist die Zielsteuerung des Vektors an die korrekte Zone innerhalb der Zelle.

[0080] An die Sequenz, die für das Transitpeptid codiert, ist die DNA-Sequenz angehängt, die für das N-terminale Ende des interessierenden Polypeptids codiert. Die Richtung der Sequenz, die für das interessierende Polypeptid codiert, hängt davon ab, ob eine sense- oder antisense-Transkription erwünscht ist. Erfindungsgemäße DNA-Konstrukte, die spezifisch im vorliegenden Text beschrieben sind, weisen die Sequenz auf, die für das interessierende Polypeptid codiert, am N-terminalen Ende, aber die für SBR codierende Region kann auch am N-terminalen Ende liegen und die Sequenz des interessierenden Polypeptids danach. Am Ende des DNA-Konstrukts befindet sich die Terminatorsequenz. Solche Sequenzen sind in der Fachwelt gut bekannt.

[0081] Der Klonierungsvektor wird in einen Wirt hineintransformiert. Das Einbringen des Klonierungsvektors, vorzugsweise eines Plasmids, kann mit verschiedenen fachbekannten Transformationstechniken erfolgen. Diese Techniken können je nach dem Wirt schwanken, es zählen jedoch der Beschuß mit der Genkanone, die Mikroinjektion, die Transformation mit Agrobacterium, die "Whiskers"-Technology (US-Patent Nr. 5,302,523 und Nr. 5,464,765), die Elektroporation und dergleichen dazu. Ist der Wirt eine Pflanze, so können Zellen regeneriert werden, um Pflanzen zu bilden. Verfahren zur Regeneration von Pflanzen sind in der Fachwelt bekannt. Sobald der Wirt transformiert ist und die Proteine in ihm exprimiert werden, ist das Vorliegen der DNA, die für das interessierende "Payload"-Polypeptid codiert, in dem Wirt bestätigbar. Das Vorliegen von exprimierten Proteinen kann mittels Western-Blot oder ELISA oder auch auf Grund einer Veränderung in der Pflanze oder Zelle bestätigt werden.

Verwendung von verkapseltem Protein

[0082] Die vorliegende Erfindung hat verschiedene Anwendungen. Das Hybridpolypeptid kann in reiner Form von der Stärke abgespaltet werden (es können Spaltstellen eingebaut werden) und das reine Protein kann gewonnen werden. Es kann jedoch auch das verkapselte interessierende Polypeptid innerhalb der Stärke in Rohform eingesetzt werden, um Protein an verschiedene Teile des Verdauungstrakts des aufnehmenden Tiers zu liefern (der Begriff "Tier" soll Säugetiere, Vögel und Fische umfassen). Ist zum Beispiel die Stärke, in die das Material verkapselt ist, verdauungsresistent, so wird das Protein langsam in den Darm des Tiers freigesetzt, wodurch ein Abbau des wertvollen Proteins im Magen vermieden wird. Aminosäuren wie Methionin und Lysin können eingekapselt werden, um direkt in das Samenkorn, das dem Tier gefüttert wird, eingebaut zu werden, wodurch es sich erübrigt, die Ration mit diesen Aminosäuren in anderen Formen anzureichern.

[0083] Die vorliegende Erfindung ermöglicht es, daß Hormone, Enzyme, Proteine, proteinartige Nährstoffe und proteinartige Arzneimittel an bestimmte Verdauungsareale im Verdauungstrakt von Tieren adressiert werden. Proteine, die normalerweise im oberen Verdauungstrakt verdaut werden und die in Stärke verkapselt sind, vermögen unverdaut den Magen zu passieren und können ganz oder teilweise vom Darm absorbiert werden. Wenn die interessierenden Polypeptide die Darmwand zu passieren vermögen, können sie für die Verabreichung von Arzneimittel an ein Tier oder für die Bereitstellung von Hormonen wie Wachstumsfaktoren, z. B. Somatotropin, für die Impfung eines Tiers oder für die Verbesserung der Nährstoffe, die für ein Tier verfügbar sind, eingesetzt werden.

[0084] Ist die eingesetzte Stärke nicht gegenüber Verdauung im Magen resistent (zum Beispiel ist die "sugary 2"-Stärke hochverdaulich), so kann das zugesetzte Protein dahingehend adressiert werden, daß es im oberen Verdauungstrakt des Tiers absorbiert wird. Dies würde es erforderlich machen, daß der für die Produktion der modifizierten Stärke eingesetzte Wirt dahingehend mutiert oder transformiert wird, daß eine Stärke des "sugary 2"-Typs erzeugt wird. Die vorliegende Erfindung umfaßt die Verwendung von mutierten Organismen, die als Wirte modifizierte Stärke bilden. Zu einigen Beispielen für diese mutierten Wirte zählen Reis und Mais und dergleichen, die "sugary 1"-, "sugary 2"-, "brittle"-, "shrunk"-, "waxy"-, "amylose extender"-, "dull"-, "opaque"-

und "floury"-Mutationen und dergleichen aufweisen. Diese mutierten Stärken und Stärken aus unterschiedlichem pflanzlichen Ausgangsmaterial weisen unterschiedliche Verdaulichkeitsniveaus auf. So kann dadurch, daß man den Wirt auf Expression der DNA und des Tiers, dem die modifizierte Stärke verfüttert wird, auswählt, das Hybridpolypeptid dort verdaut werden, wohin es adressiert wird. Unterschiedliche Proteine werden am wirksamsten von unterschiedlichen Teilen des Körpers absorbiert. Dadurch, daß man das Protein in Stärke mit der gewählten Verdaulichkeit verkapselt, kann das Protein an irgendeinen Ort im Verdauungstrakt sowie zu spezifischen Zeitpunkten während des Verdauungsvorgangs geliefert werden.

[0085] Ein weiterer Vorteil der vorliegenden Erfindung ist die Fähigkeit, unterschiedliche Glycosylierungsniveaus des gewünschten Polypeptids zu hemmen oder zu exprimieren. Das Verkapselungsverfahren kann es dem Protein ermöglichen, innerhalb des Korns in einem unterschiedlichen Glycosylierungszustand exprimiert werden als wenn es von anderen DNA-Molekülen exprimiert wird. Die Glycosylierung hängt vom Grad der Verkapselung, dem eingesetzten Wirt und der Sequenz des Polypeptids ab.

[0086] Verbesserte Kulturpflanzen mit den oben beschriebenen Eigenschaften können dadurch erzeugt werden, daß man Pflanzen, von denen bekannt ist, daß sie andere günstige Eigenschaften aufweisen, genetisch manipuliert. Dadurch, daß man die Nukleotidsequenz eines Stärkesyntheseenzymgens manipuliert, ist es möglich, die Menge an Schlüsselaminosäuren-Proteinen oder -Peptiden, die in einer Pflanze erzeugt werden, zu verändern. Ein oder mehrere genetisch veränderte(s) Genkonstrukt(e), das/die pflanzlichen, pilzlichen, bakteriellen oder tierischen Ursprungs sein kann/können, kann/können in das pflanzliche Genom durch generatives Kreuzen oder durch Transformation eingeführt werden. Veränderte Gene können zusätzliche Kopien von Wildtyp-Genen umfassen oder können für modifizierte oder allele oder alternative Enzyme mit neuen Eigenschaften codieren. Der Einbau solch eines Genkonstrukts bzw. von solchen Genkonstrukten kann je nach Art und Menge des eingeführten Gens, bzw. der eingeführten Gene (in sense- oder antisense-Orientierung) unterschiedliche Auswirkungen ausüben. Er kann die Fähigkeit der Pflanze, ein bestimmtes Protein oder Peptid zur Produktion oder eine verbesserte Aminosäurezusammensetzung bereitzustellen, erhöhen.

Klonierung der an der Stärkebiosynthese beteiligten Enzyme

[0087] Um die erfindungsgemäßen DNA-Konstrukte bereitzustellen, kann man sich bekannter Klonierungstechniken bedienen. Das Ausgangsmaterial der speziellen Formen der SSTS, der GBSTS, des BE, der Glycogensynthase (GS), des Amylopektins oder von anderen Genen, die bzw. das im vorliegenden Zusammenhang verwendet werden, kann ein beliebiger Organismus sein, der Stärke oder Glycogen erzeugen kann. Kandidaten für Donatororganismen werden gescrēent und identifiziert. Danach kann man zwei Ansätze verfolgen: (a) mittels Enzymaufreinigung und Erzeugung von Antikörpern/Sequenzen nach den im vorliegenden Text beschriebenen Vorschriften; (b) mit SSTS, GBSTS, BE, GS, Amylopectin oder sonstigen cDNAs als heterologe Sonden zur Identifikation der genomischen DNAs für SSTS, GBSTS, BE, GS, Amylopectin oder sonstige stärkeverkapselnde Enzyme in Bibliotheken des jeweiligen Organismus. Gentransformation, Pflanzenregeneration und Testvorschriften sind in der Fachwelt bekannt. In diesem Fall ist es erforderlich, Genkonstrukte für die Transformation herzustellen, die Regulationssequenzen enthalten, welche die Expression während der Stärkebildung gewährleisten. Diese Regulationssequenzen sind in vielen kleinen Samenkörnern sowie in Knollen und Wurzeln vorhanden. So sind diese Regulationssequenzen leicht im Mais-Endosperm in DNA, die für die stärkekorngewundene Stärkesynthese (GBSTS), die lösliche Stärkesynthese (SSTS) oder Verzweigungsenzyme (BE) codiert, oder in anderen Enzymen des Stärkebiosynthesewegs des Maisendosperms verfügbar. Diese Regulationssequenzen aus dem Endosperm gewährleisten eine Proteinexpression zum korrekten Entwicklungszeitpunkt (z. B. ADPG-Pyrophosphorylase).

[0088] Bei diesem Verfahren messen wir die Stärkebindungskonstanten von stärkebindenden Proteinen mittels Elektrophorese von nativem Protein in Gegenwart von geeigneten Konzentrationen an Kohlenhydraten wie Glycogen oder Amylopectin. Stärkeverkapselnde Regionen können dadurch aufgeklärt werden, daß man die ortsgerichtete Mutagenese sowie sonstige Verfahren der Gentechnik, mit denen der Fachmann vertraut ist, einsetzt. Neue kinetisch veränderte Proteine, die neue Peptide oder Aminosäurekombinationen enthalten können, können mit den im vorliegenden Text beschriebenen Verfahren ausgewertet werden.

BEISPIELE

Beispiel Eins:

Verfahren zur Identifikation von stärkeverkapselnden Proteinen

Isolation von Stärkekornprotein:

[0089] 12,5 g Samenkörner in 25 ml Extraktionspuffer homogenisieren (50 mM Tris-Acetat, pH 7,5, 1 mM EDTA, 1 mM DTT, 3 × 20 Sekunden im Waring-Blender mit 1 minütigen Abständen zwischen dem Mixen). Proben auf Eis aufbewahren. Durch Miracloth filtrieren und 30 min bei 6000 U/min zentrifugieren. Überstand verwerfen und verfärbte Feststoffe, die als Schicht auf dem weißen Stärkepellet liegen, abschaben. Pellet erneut in 25 ml Puffer suspendieren und erneut zentrifugieren. Waschschrte noch zweimal wiederholen. Gewaschenes Pellet erneut in Aceton bei -20°C suspendieren, Pellet bei -20°C absetzen lassen. Wiederholen. Stärke im Luftstrom trocknen. Bei -20°C aufbewahren.

Proteinextraktion:

[0090] 50 mg Stärke mit 1 ml 2% SDS in Eppendorf-Hütchen vermischen. Im Vortex-Mixer mischen, bei 18.000 U/min 5 Minuten lang bei 4°C zentrifugieren. Überstand abgießen. Zweimal wiederholen. Mit 1 ml Probenpuffer (4 ml destilliertes Wasser, 1 ml 0,5 Tris-HCl, pH 6,8, 0,8 ml Glycerin, 1,6 ml 10% SDS, 0,4 ml B-Mer-captoethanol, 0,2 ml 0,5% Bromphenolblau) versetzen. Eppendorf-Hütchen 10 Minuten lang kochen, dabei Deckel mit Loch versehen. Abkühlung, 10 Min lang bei 10.000 U/min zentrifugieren. Überstand in frisches Eppendorf-Hütchen abdekantieren. 4 Minuten lang gemeinsam mit Standards kochen. Abkühlung.

SDS-PAGE-Gele: (nicht denaturierend)

| | 10%iges Trenngel | 4%iges Sammelgel |
|---------------------------------|------------------|------------------|
| Acryl/Bis 40%ige Stammlösung | 2,5 ml | 1,0 ml |
| 1,5 M Tris pH 8,8 | 2,5 ml | - |
| 0,5 M Tris pH 8,8 | - | 2,5 ml |
| 10% SDS | 100 µl | 100 µl |
| Wasser | 4,845 ml | 6,34 ml |
| 15 min entgasen, frisch zugeben | | |
| 10% Ammoniumpersulfat | 50 µl | 50 µl |
| TEMED | 5 µl | 10 µl |

[0091] Mini-Protean II Doppelgel-Zelle; 3,5 ml Trennpuffer pro Gel. Mit 4%igem Sammelgel überschichten. Das Gel wird bei 200 V Gleichspannung laufen gelassen. 10 × Laufpuffer (250 mM Tris, 1,92 M Glycin, 1% SDS, pH 8,3).

Meßverfahren von stärkebindenden Regionen:

[0092]

Lösungen:

| | |
|-------------------------------------|--|
| Extraktionspuffer: | 50 mM Tris-Acetat pH 7,5, 10 mM EDTA, 10% Saccharose, 2,5 mM DTT (frisch) |
| Sammelpuffer: | 0,5 M Tris-HCl, pH 6,8 |
| Trennpuffer: | 1,5 M Tris-HCl, pH 8,8 |
| 10 X Puffer für untere Elektrode: | 30,3 Tris + 144 g Glycin qs auf 1 l (pH beträgt –8,3, wird nicht eingestellt). Für Gebrauch verdünnen. |
| Puffer für obere Elektrode: | Wie für untere Elektrode |
| Saccharose Lösung: | 18,66 g Saccharose + 100 ml dH ₂ O |
| 30% Acryl/Bis-Stammlösung (2,67%C): | 146 g Acrylamid + 4 g Bis + 350 ml dH ₂ O. Auf 500 ml Auffüllen. Filtrieren und bis zu 1 Monat lang im Dunkeln bei 4 C aufbewahren. |
| 15% Acryl/Bis-Stammlösung (20%C): | 6 g Acrylamid + 1,5 g Bis + 25 ml dH ₂ O. Auf 50 ml Auffüllen. Filtrieren und bis zu 1 Monat lang im Dunkeln bei 4 C aufbewahren. |
| Riboflavinlösung: | 1,4 g Riboflavin + 100 ml dH ₂ O. Bis zu 1 Monat lang im Dunkeln aufbewahren. |
| SS-Assay-Mix: | 25 mM Natriumcitrat, 25 mM Bicin-NaOH (pH 8,0), 2 mM EDTA, 1 mM DTT (frisch), 1 mM Adenosin-5'-diphosphoglucose (frisch), 10 mg/ml Kaninchenleberglycogen Typ III (frisch) |
| Iodlösung: | 2 g Iod + 20 g KI, 0,1 N HCl auf 1 Liter |

Extrakt:

Extrahieren:

- 4 ml Extraktionspuffer + 12 g Endosperm. Homogenisieren.
- Durch Miracloth oder 4 Lagen Mull filtrieren, 20 min bei 4°C bei 20.000 g (14.500 U/min, SM-24-Rotor) zentrifugieren.
- Überstand mit Glaspipette entfernen
- 0,85 ml Extrakt + 0,1 ml Glycerin + 0,05 ml 0,5%iges Bromphenol Blau.
- An Vortex-Mixer mischen und 5 min in der Mikrozentrifuge bei Höchstgeschwindigkeit zentrifugieren. Direkt verwenden oder in flüssigem Stickstoff einfrieren und bis zu 2 Wochen lang bei –80°C aufbewahren.

Gießen der Gele:

[0093] Gelbond PAG Film (FMC Industries, Rockland, ME) auf die (Innenseite der) äußeren Glasplatte mit doppelseitigem Klebeband mit der hydrophilen Seite nach oben anbringen. Klebeband und Film werden so knapp und gleichmäßig wie möglich an den unteren Teil der Platte angeglichen. Der Film ist etwas kleiner als die Platte. Zwischen Film und Platte Wasser spritzen, um den Film haften zu lassen. Überschüssiges Wasser mit einem Tuch herausdrücken. Platten wie üblich anordnen, dann Unterseite der Platten mit Klebstoff verschließen. Die Kassette wird in die Gießanlage passen, wenn der graue Gummi von der Gießanlage entfernt wird. Das Gel polymerisiert mit dem Film und bleibt während aller folgenden Arbeitsschritte daran haften.

[0094] Gießen von 4,5%igem Mini-Trenngel (0,75 mm):

- 2,25 ml dH₂O
- + 3,75 ml Saccharoselösung
- + 2,5 ml Trennpuffer
- + 1,5 ml 30% Acryl/Bis-Stammlösung
- + verschiedene Mengen Glycogen für jedes Gel (d. h. 0–1,0%)

15 MIN LANG ENTGASEN

+ 50 µl 10% APS

+ 5 µl TEMED

30 MIN ODER ÜBER NACHT POLYMERISIEREN

3,125%iges Sammelgelgießen:

1,59 ml dH₂O

+ 3,75 ml Saccharoselösung
 + 2,5 ml Sammelpuffer
 + 2,083 ml 15% Acryl/Bis-Stammlösung

NICHT ENTGASEN

15 µl 10% APS

+ 35 µl Riboflavinlösung

+ 30 µl TEMED

2,5 STUNDEN IN DER NÄHE EINER GLÜHBIRNE POLYMERISIEREN

[0095] In 4°C Abkühlen, bevor die Kämme herausgezogen werden. Man kann auch keine Kämme verwenden und einfach einen Zentimeter Sammelgel gießen.

[0096] Das obengenannte Verfahren:

- kann bei unterschiedlichen Temperaturen laufen gelassen werden; Gele und Lösungen vorinkubieren
- 15 min bei 200 V vorlaufen lassen
- Gel laden: 7 µl pro Schlitz oder 115 µl falls kein Kamm verwendet wurde
- Bei 140 V laufen lassen, bis die Farbstofffront in der Nähe des unteren Endes ist. Verschiedene Lauftemperaturen werden dadurch erzielt, daß man die gesamte Gelanordnung in ein Wasserbad gibt. Elektrophorese kann gelegentlich gestoppt werden, um einen Temperaturfühler in das Gel einzuführen.
- Enzym-Assay: Gele bei Farbstofffront abschneiden. In SS-Assaymix über Nacht bei Raumtemperatur unter leichtem Bewegen inkubieren. Gele mit Wasser spülen. Mit I2/KI-Lösung überschichten.
- Aufnahmen der Gele auf Leuchtkasten machen und Aufnahmen messen. Rm = mm vom oberen Rand des Gels bis zu der Aktivitätsbande/mm vom oberen Rand des Gels bis zum unteren Rand des Gels, wo es abgeschnitten wurde (wo die Farbstofffront lag). % Glycogen gegen 1/Rm auftragen. Der Punkt, wo die Linie die x-Achse schneidet, ist -K (wo y = 0).

Test- und Auswertungsvorschrift für die Länge der SBR-Region:

[0097] Die Vorgehensweise der Vorschrift oben für die Auswahl der SBR-Region erfordert vier grundlegende Schritte. Zuerst muß DNA, die für ein Protein mit einer Stärkebindungsregion codiert, ausgewählt werden. Diese kann unter bekannten Stärkesynthesegenen oder Stärkebindungs-genen wie Genen für z. B. Amylasen ausgewählt werden. Das Protein muß extrahiert werden. Es sind verschiedene Proteinextraktionstechniken in der Fachwelt bekannt. Das Protein kann mit Proteasen behandelt werden, um zu Proteinfragmenten mit unterschiedlichen Längen zu gelangen. Die bevorzugten Fragmente weisen Deletionen in erster Linie von der N-terminalen Region des Proteins auf. Die SBR-Region befindet sich näher zum C-terminalen Ende als zum N-terminalen Ende. Das Protein wird auf den oben beschriebenen Gelen laufengelassen und die Affinität für die Gelmatrix wird ausgewertet. Eine höhere Affinität zeigt eine stärkere Bevorzugung dieser Region des Proteins für die Matrix an. Mit diesem Verfahren lassen sich unterschiedliche Proteine vergleichen, um die stärkebindenden Regionen in natürlichen bzw. synthetischen Proteinen zu vergleichen.

Beispiel Zwei:

SBR-Fusionsvektor:

[0098] Die folgenden Fusionsvektoren sind an die Verwendung in E. coli adaptiert. Das Fusionsgen, das mit der mutmaßlichen SBR in diesen Vektoren verbunden war, hat für das "Green Fluorescent Protein" (GFP) codiert. Eine beliebige Anzahl von unterschiedlichen Genen, die für Proteine und Polypeptide codieren, konnten in die Vektoren ligiert werden. Es wurde ein Fusionsvektor konstruiert, bei dem die SBR von waxy-Mais mit einem zweiten Gen oder Genfragment, im vorliegenden Fall GFP, fusioniert war.

[0099] pEXS114 (siehe [Fig. 1a](#)): Synthetisches GFP (SGFP) wurde ausgehend von dem Plasmid HBT-SGFP (von Jen Sheen; Dept. of Molecular Biology; Wellman 1 I, MGH; Boston, MA 02114) mit den Primern EXS73 (5'-GACTAGTCATATG GTG AGC AAG GGC GAG GAG-3') [SEQ ID NO: 1] und EXS74 (5'-CTAGATCTT-CATATG CTT GTA CAG CTC GTC CAT GCC-3') [SEQ ID NO: 2] mittels PCR amplifiziert. Die Enden des PCR-Produkts wurden mit T-DNA-Polymerase abgeglichen, um stumpfe Enden zu erzeugen; anschließend wurde das PCR-Produkt mit SpeI verdaut. Dieses SGFP-Fragment wurde in die EcoRV-SpeI Stellen von pBSK (Stratagene, 11011 North Torrey Pines Rd. La Jolla, CA.) subkloniert, wodurch man zu pEXS114 gelangte.

[0100] pEXS115 [siehe [Fig. 1b](#)]: Synthetisches GFP (SGFP) wurde aus dem Plasmid HBT-SGFP (von Jen Sheen) mit den Primern EXS73 (siehe oben) und EXS75 (5'-CTAGATCTTGCCATGGC CTT GTA CAG CTC

GTC CAT GCC-3') [SEQ ID NO: 3] mittels PCR amplifiziert. Die Enden des PCR-Produkts wurden mit T-DNA-Polymerase abgeglichen, um stumpfe Enden zu erzeugen; anschließend wurde das PCR-Produkt mit SpeI verdaut. Dieses SGFP-Fragment wurde in die EcoRV-SpeI Stellen von pBSK (Stratagene) subkloniert, wodurch man zu pEXS115 gelangte.

[0101] Das pEXSWX (siehe [Fig. 2a](#)): Mais-WX subkloniert NdeI-NotI in pEP-21a (see [Fig. 2b](#)). Die genomische DNA-Sequenz und assoziierte Aminosäuren, von denen ausgehend die mRNA-Sequenz erzeugt werden kann, sind in den Tabellen 1a und 1b unten gezeigt; alternativ dazu könnte die in den folgenden Tabellen angeführte DNA eingesetzt werden.

Tabelle 1a

DNA-Sequenz und abgeleitete Aminosäuresequenz des waxy-Gens in Mais
[SEQ ID NO: 4 und SEQ ID NO: 5]

| | |
|-----------------|---|
| LOCUS | ZMWAXY 4800 Bp DNA PLN |
| DEFINITION | waxy(wx+)-Locus von Zea mays für UDP-Glucose-Stärkeglycosyltransferase |
| ZUGANGSNUMMER | X03935 M24258 |
| SCHLÜSSELWÖRTER | Glycosyltransferase; Transitpeptid UDP-Glucosestärke, Glycosyltransferase; waxy-Locus |
| HERKUNFT | Mais |
| ORGANISMUS | Zea mays Eukaryota; Plantae; Embryobionta; Magnoliophyta; Liliopsida; Commelinidae, Cyperales; Poaceae |
| REFERENZ | 1 (Basen 1 bis 4800) |
| AUTOREN | Kloesgen, R. B., Gierl, A., Schwarz-Sommer, Z., und Saedler, H. |
| TITEL | Molecular analysis of the waxy locus of Zea mays |
| ZEITSCHRIFT | Mol. Gen. Genet. 203, 237-244 (1986) |
| STANDARD | Vollautomatisch |
| KOMMENTAR | NCBI gi: 22509 |
| MERKMALE | Lage/Kennzeichnungen |

```

source          1..4800
                /organism="Zea mays"
repeat_region   283..287
                /note="direct repeat 1"
repeat_region   288..292
                /note="direct repeat 1"
repeat_region   293..297
                /note="direct repeat 1"
repeat_region   298..302
                /note="direct repeat 1"
misc_feature     372..385
                /note="GC stretch (pot. regulatory factor binding
site)"
misc_feature     442..468
                /note="GC stretch (pot. regulatory factor binding
site)"
misc_feature     768..782
                /note="GC stretch (pot. regulatory factor binding
site)"
misc_feature     810..822
                /note="GC stretch (pot. regulatory factor binding
site)"
misc_feature     821..828
                /note="target duplication site (Ac7)"
CAAT_signal     821..828
TATA_signal     867..873
misc_feature     887..900
                /note="GC stretch (pot. regulatory factor binding
site)"
misc_feature     901
                /note="transcriptional start site"
exon            901..1080
                /number=1

```



```

intron      1081..1219
            /number=1
exon        1220..1553
            /number=2
transit_peptide 1233..1448
CDS         join(1449..1553,1685..1765,1860..1958,2055..2144,
3226..2289,2413..2513,2651..2760,2858..3101,3212..3394,
3490..3681,3793..3879,3977..4105,4227..4343)
            /note="NCBI gi: 22510"
            /codon_start=1
            /product="glucosyl transferase"

/translation="ASAGHNVPVFGAEMAPWSKTGGLGDVLGGLPPAHAANGHRVMVV
SPRYDQYKDAWDTSVVSEIKMGDDQYETVRFPHCYKRGVDRVFDHPLFLERVWGKTEE
KIYGPVAGTBYRDNQLRFSELLCQAALAPRILSLNNNPYFSGPYGEDVVFVCDNWHTC
PLSCYLKSNYQSHCIYRDAKTAFCIHNIYQGRFAPSDYPELNLPERFKSSFDFIGY
EKPEVEGRKINWMKAGILDADNVLTSPYYABELISGIARGCELDNIHRLTGITGIVNG
MDVSEWDPSRDKYIAVKYDUSTAVEAKALNKEALQAEVGLPVDNRNIPLVAFIGRLEEQ
KGPOVMAAAIPQIMCNVEDVQIVLLGTGKKKFERMLMSADEKFPGRKPAVVKENAAALA
NHIMAGADVLAVTSRPEPCGLIQLOQHRVGTPCACASTGGGLVDTTIECKTGFHRCRLS
VDCNVVEPADVKKVATTLQRAIKVVGTTPAYEEMVRNCHIQQLSWKGPAKNWENVLLSL
GVAGGEPGVEGEEIAPLAKENVAAP"
intron      1554..1684
            /number=2
exon        1685..1765
            /number=3
intron      1766..1859
            /number=3
exon        1860..1958
            /number=4
intron      1959..2054
            /number=4
exon        2055..2144
            /number=5
intron      2145..2225
            /number=5
exon        2226..2289
            /number=6
intron      2290..2412
            /number=6
exon        2413..2513
            /number=7
intron      2514..2650
            /number=7
exon        2651..2760
            /number=8
intron      2761..2857
            /number=8
exon        2858..3101
            /number=9
intron      3102..3211
            /number=9
exon        3212..3394
            /number=10
misc_feature 3358..3365
            /note="target duplication site (Ac9)"
intron      3395..3489
            /number=10
exon        3490..3681

```

```

/misc_feature      /number=11
                   3570..3572
                   /note="target duplication site (Spm 18)"
intron             3682..3792
                   /number=11
exon               3793..3879
                   /number=12
intron             3880..3976
                   /number=12
exon               3977..4105
                   /number=13
intron             4106..4226
                   /number=13
exon               4227..4595
                   /number=14
polyA_signal       4570..4575
polyA_signal       4593..4598
polyA_site         4595
polyA_signal       4597..4602
polyA_site         4618
polyA_site         4625

```

BASENZAHL 935 A 1413 C 1447 G 1005 T

ORIGIN

```

1  CAGCGACCTA TTACACAGCC CGCTCGGGCC CGCGACGTCG GGACACATCT TCTTCCCCCT
61  TTTGGTGAAG CTCTGCTCGC AGCTGTCCCG CTCCTTGSAC GTTCGTGTGG CAGATTCATC
121 TGTGTCTCTG TCTCTGTGTC TTCTTGGGTA GCTTGTGTAG TGGAGCTGAC ATGGTCTGAG
181 CAGGCTTAAA ATTTGCTCCT AGACGAGGAG TACCAGCACA GCACGTTGGG GATTTCTCTG
241 CCTGTGAAGT GCAACGCTTA GGATTGTGAC ACGCCTTGCT CGCGTCCCGT CGCGTCCCGT
301 CGATGCGGTG GTSAGCAGAG CAGCAACAGC TGGGCGGGCC AACCTTGGCT TCCGTCTCTT
361 CGTCTACCTT ACGCGCGCGC CGCGGACAGC CAGCAGAGAG CGGACAGCCA GCGGTGACCG
421 GCGAGCTGCT GTGGAAGTGG ACGCGCGCGC CGCGCGGGCC GCGCGCGGTG GGCACCCCAA
481 AAGTACCCAC GACAAAGCAA GGCGCCAAAG CGATCCAGC TCCGGAACGC AACAGCATGC
541 GTCCGCTCGG AGAGCCAGCC ACAAGCAGCC GAGAACCGAA CCGGTGGGCG ACGCGTCATG
601 GGACGGAACG CGGCGACGCT TCCAAACGGG CCACGTACGC CGGCGTGTGC GTCCGTGCGG
661 ACGACAAGCC AAGCGGAGGC AGCCCCCGAT CGGGAAGAGG TTTTGGGCGC GAGCGCTGGC
721 GTGCGGGTCA GTCCCTGGTG CGCAGTGCCG GGGGGAACGG GTATCGTGGG GGGCGCGGGC
781 GGAGGAGAGC GTGGCGAGGG CCGAGAGCAG CGGCGGGCCG GGTCAACGAA CGCGCCCCAC
841 GTACTGCCCT CCGCCTCCGC GCGCGCTAGA AATACCGAGG CTTGGACCGG GGGGCGGGCC
901 CGTCACATCC ATCCATCGAC CGATCGATCG CCACAGCCAA CACCACCGGC CGAGGCGACG
961 CGACAGCCGC CAGGAGGAAG GAATAAATC ACTGCCAGCC AGTGAAGGGG GAGAAGTGTA
1021 TGTCTCCGTC GACCAGTGGG CGCACCGCCC GGCAGGGCTG CTCATCTCGT CGACGACCAG
1081 GTTCTGTTCC GTTCCGATCC GATCCGATCC TGTCTTGA GTTCTGTCAG ATCCTGGGCG
1141 GTATCTGCGT GTTTGATGAT CCAGTTCTT CGAACCTAAA TCTGTCCGTG CACAGCTCTT
1201 TTCTCTCTCT CCTACGCACT GATTAAATCG GCATGCGCGC TCTGGCCACG TCGCAGCTCG
1261 TCGCAACCGC CGCGGGCGTG GGCCTCCCGG ACGCCTCCAC GTTCCGCGCG GCGCGCGCGC

```

1321 AUGGCCCTGAG GGGGGCCCCG GCGTCGGCCG CGGCGGACAC GCTCAGCATG CCGACCCAGCG
 1381 CGCCCGGGGC GCCCAGGCAC CAGCAGCAGG CGCGCCGCGG GGGCAGGTTC CCGTCGCTCG
 1441 TCGTGTGCGC CAGCGCGCGC ATGAACGTCG TCTTCGTCGG CGCCGAGATG GCGCCGTGGA
 1501 GCAAGACCGG CCGGCTCGGC GACGTCCTCG GCGGCTGCGC GCCGGCCATG GCCGTAAAGCG
 1561 CGCGCACCGA GACATGCATC CGTTGGATCG CGTCTTCTTC GTGCTCTTGC CGCGTGCAATG
 1621 ATGCATGTGT TTCTCTCTGG CTTGTGTTCC TGTATGTGAC GTGTTTGTTC GGGCATGCAT
 1681 GCAGGCGAAC GGGCACCGTG TCATGGTCGT CTCTCCCCGC TACGACCACT ACAAGGACCG
 1741 CTGGGACACC AGCGTCCTGT CCGAGGTACG GCCACCGAGA CCAGATTGAG ATCACAGTCA
 1801 CACACACCGT CATATGAACC TTCTCTGCT CTGATGCCTG CAACTGCAAA TGCATGCAGA
 1861 TCAAGATGCG AGACGGGTAC GAGACGGTCA GGTCTTCCA CTGCTACAAG CGCGGAGTGG
 1921 ACCGCGTGT CTGTGACCAC CCACGTGTTC TGGAGAGCGT GAGACGAGAT CTGATCACTC
 1981 GATACGCAAT TACCACCCCA TTGTAAGCAG TTACAGTGAG CTTTTTTTCC CCCCAGCGTG
 2041 GTCCCTGGTT TCAGGTTTGG GGAAACACCG AGGAGAAGAT CTACGGGCCT GTCCCTCGAA
 2101 CGGACTACAG GGACACCCAG CTGCGGTTCA GCCTGCTATG CCAGGTCAGG ATGGCTTGGT
 2161 ACTACAACTT CATATCATCT GTATGCAGCA GTATACACTG ATCAGAAATG CATGCTGTTC
 2221 TCGAGCCAGC ACTTGAAGCT CCAAGGATCC TGAGCCTCAA CAACACCCCA TACTCTCCG
 2281 GACCATACGG TAAGAGTTGC AGTCTTCGTA TATATATCTG TTGAGCTCGA GAATCTTCAC
 2341 AGGAAGCGCG CCATCAGACG GACTGTCAAT TTACACTGAC TACTGCTGCT GCTCTTCGTC
 2401 CATCCATACA AGGGGAGGAC GTGCTGTTCC TGTGCAACGA CTGGCACACC GGGCCTCTCT
 2461 CGTGCTACCT CAAGAGCAAC TACCAGTCCC ACGGCATCTA CAGGGACGCA AAGGTTGCCT
 2521 TCTCTGAAT GAACAACGCC GTTTTCGTTT TCCATGCTCG TATATACCTC GTCTGGTAGT
 2581 GGTGCTGCTT CTCTGAGAAA CTAAGTGAAG CTGACTGCAT GTCTGTCTGA CCATCTTCAC
 2641 GTACTACCAG ACCGCTTTCT GCATCCACAA CATCTCCTAC CAGGGCGCGT TCGCCTTCTC
 2701 CGACTACCGG GAGCTGAACC TCCCCGAGAG ATTCAAGTCG TCCTTCGATT TCATCGACGG
 2761 GTCTGTTTTT CTGGGTGCAT GTGAACATTC ATGAATGGTA ACCCACAACCT GTTCGCGTCC
 2821 TGCTGGTTCA TTATCTGACC TGATTGCATT ATTGCAGCTA CGAGAGCCCG GTGGAGGGCC
 2881 GGAAGATCAA CTGGATGAAG GCCGGGATCC TCGAGGCCGA CAGGGTCCTC ACCGTCAGCC
 2941 CCTACTACGC CGAGGAGCTC ATCTCCGGCA TCGCCAGGGG CTGCGAGCTC GACAACATCA
 3001 TGCGCTCAC CGGCATCACC GGCATCGTCA ACGGCATGSA CGTCAGCGAG TGGGACCCCA
 3061 GCAGGGACAA GTACATCGCC GTGAAGTACG ACGTGTGAC GGTGAGCTGG CTAGCTCTGA
 3121 TTCTGCTGCC TGGTCTCCT GCTCATCATG CTGGTTCGCT ACTGACCGCG CAAGTGTACG
 3181 TACGTGCGTG CGAGGCTGGT GTCCGGTTCA GGGCCTGCGG GCCAAGGCGC TGAACAAGGA
 3241 GCGGCTGCAG GCGGAGCTCG GGTCCCGGT GGACCGGAAC ATCCCGCTGG TGGCGTTCAT
 3301 CGGCAGGCTG GAAGAGCAGA AGGGCCCGA CGTCATCGCG GCCGCGATCC CGCAGCTCAT

```

3341 GGAGATCGTG CAGGACGTGC AGATCGTTCT GCTGCTACGT GTGCGCGCGC CGCCACCCCG
3421 CTACTACATG CGTGATCGT TCGTTCTACT GGAACATGG TGTAGCAAC GCGTGGATA
3481 ATGCTGAGG GCACGGGCAA GAAGAAGTC GACCGCATGG TCATGAGGCG CGAGGAGAG
3541 TTCCACAGCA AGGTGCGCGC CGTGGTCAAG TTCACGCGCG CGCTGCGGCA CACATCATG
3601 GCGGCGCGCG ACCTGCTCGC GGTACACAGC CGCTTCGAGC CCGCGCGCGT CATCCAGCTG
3661 CAGGGGATGC GATACGGAAC GGTACGACAG AAAAAAAAAA TCCTGAATCC TGACGAGAGG
3721 GACAGAGACA GATTATGAAT GCTTCATCGA TTGAATTGA TTGATCGATC TCTCCCGCTG
3781 CGACTCTTGC AGCCCTGCGC CTGCGCGTCC ACCGGTGGAC TGTGCGACAC CATCATGAA
3841 GCGAAGACCG GGTTCACATC GGGCGCGCTC AGCTTGAGC TAAGCCTAGC TGTGCGATG
3901 TCTTTCTTCT TTCTTTCTGT ATGTATGTAT GANTCAGCAC CGCCGTTCTT GTTTCGTCTG
3961 GGTCTCTCTT TCCAGTGTAT ACCTCGTGGC GCGGCGCGAC GTCAAGAAGG TGGCCACCAC
4021 ATTCCAGCGC GCCATCAAGG TGCTCGGCAC CCGGCGGTAC GAGGAGATGG TGAGGAAGTG
4081 CATGATCCAG GATCTCTCTT GGAAGGTACG TACGCGCGCC CCGCGCGCGC CCGCCAGAGC
4141 AGAGCGCCAA GATCGACGCA TCGACCGACC ACACGTAGGC GCTCGCTCC TGTGCTGAC
4201 CGTGGTTTAA TTTCGAAAT GCGCAGGCGC CTGCCAAGAA CTGGGACACG GTGCTGCTCA
4261 GCTTCGGGGT CGCGCGCGGC GAGCGAGGGG TCGAAGGCGA GGAGATCGCG CCGCTCGGCA
4321 AGGAGAACGT GCGCGCGCGC TGAAGAGTTC GCGCTGACAG GCGCTGATC TCGCGCTGG
4381 TCGAAGATG TTGGGACATC TTCTTATATA TGCTGTTTCC TTATGTGAT ATGGACAAGT
4441 ATGTGTAGCT GCTTCTCTGT GCTAGTGTAA TGTAGTGTAG TGCTGGCCAG TGGCACAACC
4501 TAATAAGCGC ATGAACATAT TGCTTCCGTG TGTAGTGAAG TACCGATCGG TAATTTTATA
4561 TTCCGAGTAA ATAAATGGAC CTCTAGTGGT CGAGTAAATA ATCGCTGCTG TTGCTGTCTC
4621 TTATCGCTCC TCGTATAGAT ATTATATAGA GTACATTTT CTCTCTCTGA ATCGTACGT
4681 TGTGAAATTT CTATATCATT ACTGTAAAT TTCTGCGTTC CAAAAGAGAC CATAGGCTAT
4741 CTTTGGCCCT GTTGTGTTTC GCTTCTGGCA GCTTCTGGCC ACCAAAGCT GGTGCGGACT

```

Tabelle 1b

DNA-Sequenz und abgeleitete Aminosäuresequenz des waxy-Gens in Reis
[SEQ ID NO: 6 und SEQ ID NO: 7]

| | |
|-----------------|--|
| LOCUS | OSWX 2542 Bp RNA PLN |
| DEFINITION | O. sativa Waxy mRNA |
| ZUGANGSNUMMER | X62134 S39554 |
| SCHLÜSSELWÖRTER | Glycosyltransferase; Stärkebiosynthese; waxy-Gen |
| HERKUNFT | Reis |
| ORGANISMUS | Oryza sativa Eukaryota; Plantae; Embryobionta; Magnoliophyta; Liliopsida; Commelinidae, Cyperales; Poaceae |
| REFERENZ | 1 (Basen 1 bis 2542) |
| AUTOREN | Okayaki, R. J |
| TITEL | Direkte Einreichung |
| ZEITSCHRIFT | Eingereicht (12-SEP-1991) bei den EMBL/GenBank/DDBJ-Datenbanken |
| R. J. | |
| | Okayaki, Universität Florida, Institut für Gemüsekulturen, 1255 Fifield Hall, 514 IFAS, Gainesville, Florida 32611-0514, USA |
| STANDARD | Vollautomatisch |
| REFERENZ | 2 (Basen 1 bis 2542) |
| AUTOREN | Okayaki, R. J |
| TITEL | Nukleotidsequenz einer langen cDNA des waxy-Gens aus Reis |
| ZEITSCHRIFT | Plant Mol. Biol. 19, 513-516 (1992) |
| STANDARD | Vollautomatisch |
| KOMMENTAR | NCBI gi: 20402 |
| MERKMALE | Lage/Kennzeichnungen |

```

source      1..2542
            /organism="Oryza sativa"
            /dev_stage="immature seed"
            /tissue_type="seed"
CDS         453..2282
            /gene="Wx"
            /standard_name="Waxy gene"
            /EC_number="2.4.1.21"
            /note="NCBI gi: 20403"
            /codon_start=1
            /function="starch biosynthesis"
            /product="starch (bacterial glycogen) synthase"

/tranelation="MSALTTSQLATSATGFGIADRSAPSSLLRHGFQGLKPRSPAGGD
ATSLSVTTSARATPKQQRSVQGRSRRFPVSVVYATGAGMNVVFGAEMAPWSKTGGLG
DVLGGLPPAMAANGHRVMVISPRYDQYKDAWDTSVVAEIKVADRYERVRRFFHCYKRCV
DRVFDHPSFLEKVVWCKTGEKIYCPDTGVVDYKDNQMRFSLLCQAALAPRILNLNNNP
YFKGTYGEDVVVFCNDWHTGPLASYLKNNYQPNGIYRNAKVAFCIHNISYQGRFAPED
YPELNLSEFRSSFDFIGYDTPVEGRKINWMKAGILEADRVLTVSPYYAEELISGIA
RCCELDNIMRLTGITGIVNGMDVSEWDPSKDKYITAKYDATTATIEAKALNKEALQAEA
GLPVDRKIPLIAFIGRLEEQKGPDMVMAAIPELMQEDVQIVLLGTGKKKFEKLLKSME
EKYPGKVRVAVVKFNAPLAHLIMAGADVLA VPSRFEPGLIQLGMRVGTFCACASTGG
LVDTVIEGKTGFHMGRLSVDCKVVEPSDVKKVAATLKRAIKVVGTPAYEEMVRNCHNQ
            DLSWKGFAKNWENVLLGLGVAGSAPGIEGDEIAPLAKENVAAP"
            2283..2535
            3'UTR
            polyA_site 2535
BASEZAHLE 610 A 665 C 693 G 574 T
ORIGIN
1 GAATTCAGTC TGAAGGAATA GATTCTCTTC AAAACAATTT AATCATT CAT CTCATCTGCT

```

61 CAAAGCTCTG TGCATCTCCG GGTGCAAGCG CCAGCATATT TATTGTGCAG TAAAAAATC
 121 TCATATCCCC TAGCCACCCA AGAACTGGT CTTAAGTCC TTATAAGCAG ATATGCCATT
 181 GTAATATATA TGTGTAGTT TTAGCAGCAA TTTTITTTAA AACCTTTGGT CTTTTTATG
 241 AACGTTTTAA GTTCACTGT CTTTTTTTTT CGAATTTTAA ATGTAGCTTC AAATTCTAAT
 301 CCCCAGTCCA AATTGTATA AACTTCATT CTCCTAATA ACATCTTAAT TCATTATTT
 361 GAAAACCACT TCAAATCTT TTTAGGCTCA CCAACCTTA AACATTCAA TTCAGTGCAG
 421 AGATCTTCCA CAGCAACAGC TAGACAACCA CCATGTCGGC TCTCACCAGC TCCCAGCTCG
 481 CCACCTCGGC CACCGGCTTC GGCATGCGCG ACAGGTGGGC GCGGTGCTCG CTGCTCGGCC
 541 ACGGGTTCCA GGGCTCTAAG CCCCAGAGCC CCGCGGCGCG CGACGCGACG TCGCTCAGCG
 601 TGACGACCAG CGGCGGCGCG ACGCCCAAGC AGCAGCGGTC GGTGCAGCGT GGCAGCGGGA
 661 GGTTCGCTC CGTCTGCTG TAGGCCACCG GCGCGGCGAT GAACCTGCTG TTGCTGCGCG
 721 CCGAGATGGC CCGCTGGAGC AAGACCGCG GCTCTGCTGA CTCTCTGGT GGCCTCCCCC
 781 CTGCCATGGC TCGAATGGC CACAGGGTCA TGGTATCTC TCCTCGGTAC GACCAGTACA
 841 AGGACGCTG GATACCCAGC GTTGTGCTG AGATCAAGT TCGAGACAGG TACGAGAGGG
 901 TGAGGTTTTT CCATTGCTAC AAGCGTGGAG TCGACCTGT GTTCATCGAC CATCGCTCAT
 961 TCCTCGAGAA GGTTCGGGA AAGACCGCTG AGAAGATCTA CGACCTGAC ACTGGAGTTG
 1021 ATTACAAGA CAGCCAGATG CGTTTCAGCC TTCTTTGCCA GGCAGCACTC GAGGCTCCTA
 1081 GGATCTTAA CCTCAACAC AACCCATACT TCAAGGAAC TTATGCTCAG GATCTTGTGT
 1141 TCCTCTGCAA CGACTGGCAG ACTGCCCCAG TGGCGAGCTA CTTGAAGAAC AACTACCGC
 1201 CCAATGGCAT CTACAGGAT GCAAGGTTG CTTTCTGCAT CCACACATC TCCTACCGG
 1261 GCGCTTTCCG TTTCCAGGAT TACCTGAGC TGAACCTCTC CGAGAGGTTG AGGTCTCTCT
 1321 TCGATTTCAT CGACGGGTAT GACACGCCCG TGGAGGGCAG GAAGATCAAC TCGATGAGG
 1381 CCGGAATCCT CGAAGCCGAC AGGTGCTCA CGGTGAGCCG GTACTACCGG GAGGAGCTCA
 1441 TCTCGGGCAT CGCCAGGGGA TCGAGCTCG ACAACATCAT GCGGCTCACC GGCCTCAGG
 1501 GCATGCTCAA CGGCATGGAC GTGAGCGAGT GGGATCTAG CAGGAGCAAG TACATCAGG
 1561 CCAAGTACGA CGCAACCCAG GCAATCGAGG CGAAGGCGCT GAACAGAGAG GCGTTGCAGG
 1621 CGGAGGCGGG TCTTCGGTC GACAGGAAA TCGCACTGAT CCGCTTCATC GGCAGGCTGG
 1681 AGGAACAGAA GGGCCCTGAC GTCATGCGCG CCGCCATCCC GGAGCTCATG CAGGAGGACG
 1741 TCCAGATCGT TCTTCTGGT ACTGGAAAG AGAAGTTGGA GAAGCTGCTC AAGAGCATGG
 1801 AGGAGAAGTA TCCGGGCAAG GTGAGGGCGG TGGTGAAGT CAACGCGCGG CTGCTCATC
 1861 TCATCATGGC CGGACCGGAC GTGCTGCGCG TCCCCAGCG CTTCAGGCGG TGTGACTCA
 1921 TCCAGCTGCA GCGGATGAGA TACGCAACGC CTTGTGCTT GCGTCCAGC GGTGGGCTCG
 1981 TGGACACGGT CATCGAAGC AAGACTGCT TCCACATGG CCGTCTCAGC GTGACTGCA
 2041 AGGTGCTGGA GCGAAGCGAC GTGAAGAAG TGGCGGCGAC CCGAAGCGG GCCATCAAG

 2101 TCTGTGGCAG CCGCGCTGAC GAGGAGATGG TCAGGAACTG CATGAGCAG GACCTCTCT
 2161 CGAAGCGGCC TCGAAGAAC TCGGAGATG TCGTCTGCG CCGGCGCTG CCGCGCAGG
 2221 CCGCGCGGAT CGAAGCGGAC GAGATGCGC CGCTCGGCAA GGAGAAGCTG CCGCTCTCT
 2281 GAAGAGCTG AGATCTACAT ATGGAGTGT TAAITTAATAT AGCAGTATAT GGATGAGAGA
 2341 CGAATGAGC ACTGCTTGT TTGTTGTAGT GAATTTGTAG CTATAGCCAA TTATATAGGC
 2401 TAATAAGTTT GATGTTGTAG TCTCTGCGT CTGCTTAAGT ATCTTATCGG ACCCTGAAT
 2461 TATGCTGCTG GCTTATTCG AATAATATTA AGTAATAAG CCTCTATTA ATTATTATAT
 2521 ATGTTATATT ATACTAAAA AA

Tabelle 2

DNA-Sequenz und abgeleitete Aminosäuresequenz des Gens für lösliche Stärkesynthase IIa in Mais
[SEQ ID NO: 8 und SEQ ID NO: 9]

| | | | |
|------------------------|-----------|----------------|---------|
| AKTENBEZEICH- NUNG: | MSS2C.SEQ | SEQUENZ:NORMAL | 2007 BP |
| CODONTABELLE: | UNIV.TCN | | |
| SEQUENZREGION | | 1-2007 | |
| TRANSLATIONSREGION | | 1-2007 | |

*** DNA-TRANSLATION ***

```

1  GCT GAG GGT GAG GCC GCG GGC AAG GAC GCG CCG CCG GAG AGG AGC GGC      48
1  A  E  A  E  A  G  G  K  D  A  P  P  E  R  S  G      16

49  GAC GCC GCC AGG TTG CCC CCG GGT CCG CGC AAT GCG CTC TCC AAA CCG      96
17  D  A  A  H  L  P  R  A  R  R  N  A  V  S  K  R      32

97  AGG GAT CCT CTT CAG CCG GTC GCG CCG TAC GCG TCC CCG ACG GGA AAC      144
13  R  D  P  L  Q  P  V  G  R  Y  G  S  A  T  S  N      48

145  ACG GCC AGG ACC GGC GCC GCG TCC TCC CAG AAC GCC GCA TTG CCG GAC      192
49  T  A  R  T  G  A  A  S  C  Q  N  A  A  L  A  D      64

193  GTT CAG ATC GTT GAG ATC AAG TCC ATC GTC GCC GCG CCG CCG ACG AGC      240
65  V  E  I  V  E  I  K  S  I  V  A  A  P  P  T  S      80

241  ATA GTG AAG TTC CCA GGG CCG GGG CTA CAG GAT GAT CCT TCC CTC TCG      288
81  I  V  K  F  P  G  R  G  L  Q  D  D  P  S  L  W      96

289  GAC ATA GCA CCG GAG ACT GTC CTC CCA GCC CCG AAG CCA CTC CAT GAA      336
97  D  I  A  P  E  T  V  L  P  A  P  K  P  L  H  E      112

337  TCG CCT GCG GTT GAC GCA GAT TCA AAT GGA ATT GCA CCT CCT ACA GTT      384
113  S  P  A  V  D  G  D  S  H  G  I  A  P  P  T  V      128

385  GAG CCA TTA GTA CAG GAG GCC ACT TGG GAT TTC AAG AAA TAC ATC GGT      432
129  E  P  L  V  Q  E  A  T  W  D  F  K  K  Y  I  G      144

433  TTT GAC GAG CCT GAC GAA CCG AAG GAT GAT TCC AGC GTT GGT CCA GAT      480

```

| | | |
|------|---|------|
| 145 | F D E P D E A K D D S R V G A D | 160 |
| 481 | GAT GCT GGT TCT TTT GAA CAT TAT GGG ACA ATG ATT CTG GGC CTT TGT | 528 |
| 161 | D A G S F E H Y G T M I L G L C | 176 |
| 529 | GGG GAC AAT GTT ATG AAC GTG ATC GTG GCT GGT GAA TGT TCT CCA | 576 |
| 177 | G E N V M N V I V V A A E C S P | 192 |
| 577 | TGG TGC AAA ACA GGT GGT CTT GGA GAT GTT GTG GGA GCT TTA CCC AAG | 624 |
| 193 | W C K T G G L G D V V G A L P K | 208 |
| 625 | GCT TTA GCG AGA AGA GGA CAT CGT GTT ATG GTT GTG GTA CCA AGG TAT | 672 |
| 209 | A L A R R G H R V M V V V P R Y | 224 |
| 673 | GGG GAC TAT GTG GAA GCC TTT GAT ATG GGA ATC CGG AAA TAC TAC AAA | 720 |
| 225 | G D Y V E A F D M G I R K Y Y K | 240 |
| 721 | GCT GCA GGA CAG GAC CTA GAA GTG AAC TAT TTC CAT GCA TTT ATT GAT | 768 |
| 241 | A A G Q D L E V N Y F H A F I D | 256 |
| 769 | GGA GTC GAC TTT GTG TTC ATT GAT GCC TCT TTC GGG CAC CGT CAA GAT | 816 |
| 257 | G V D F V F I D A S F R H R Q D | 272 |
| 817 | GAC ATA TAT GGG GGA AGT AGG CAG GAA ATC ATG AAG CGC ATG ATT TTG | 864 |
| 273 | D I Y G S R Q E I M K R M I L | 288 |
| 865 | TTT TGC AAG CTT CCT GTT GAG GTT CCT TGG CAC GTT CCA TGC GGT GGT | 912 |
| 289 | F C K V A V E V P W H V P C G G | 304 |
| 913 | GTG TGC TAC GGA GAT GGA AAT TTG GTG TTC ATT GCC ATG AAT TGG CAC | 960 |
| 305 | V C Y G D G N L V F I A M N W H | 320 |
| 961 | ACT GCA CTC CTG CCT GTT TAT CTC AAG GCA TAT TAC AGA GAC CAT GGG | 1008 |
| 321 | T A L L P V Y L K A Y Y R D H G | 336 |
| 1009 | TTA ATG CAG TAC ACT CGC TGC GTC CTC GTC ATA CAT AAC ATC GGC CAC | 1056 |
| 337 | L M Q Y T R S V L V I H N I G H | 352 |
| 1057 | CAG GGC CGT GGT CCT GTA CAT GAA TTC CCG TAC ATG GAC TTG CTG AAC | 1104 |
| 353 | Q G R G P V H E F P Y M D L L N | 368 |
| 1105 | ACT AAC CTT CAA CAT TTC GAG CTG TAC GAT CCC GTC GGT GGC GAG CAC | 1152 |
| 369 | T N L Q H F E L Y D P V G G E H | 384 |
| 1153 | GCC AAC ATC TTT GCC GCG TGT GTT CTG AAG ATG GCA GAC CGG GTG GTG | 1200 |
| 385 | A N I F A A C V L K M A D R V V | 400 |
| 1201 | ACT CTC AGC CGC GGC TAC CTG TGG GAG CTG AAG ACA GTG GAA GGC GCC | 1248 |
| 401 | T V S R G Y L W E L K T V E G G | 416 |
| 1249 | TGG GGC CTC CAC GAC ATC ATC CGT TCT AAC GAC TGG AAG ATC AAT GGC | 1296 |
| 417 | W G L H D I I R S N D W K I N G | 432 |
| 1297 | ATT CGT GAA CGC ATC GAC CAC CAG GAG TGG AAC CCC AAG GTG GAC GTG | 1344 |
| 433 | I R E R I D H Q E W N P K V D V | 448 |
| 1345 | CAC CTG CGG TCG GAC GGC TAC ACC AAC TAC TCC CTC GAG ACA CTC GAC | 1392 |
| 449 | H L R S D G Y T N Y S L E T L D | 464 |
| 1393 | GCT GGA AAG CGG CAG TGC AAG GCG GCC CTG CAG CGG GAC CTG GGC CTG | 1440 |
| 465 | A G K R Q C K A A L Q R D V G L | 480 |
| 1441 | GAA GTG CGC GAC GAC GTG CCG CTG CTC GGC TTC ATC GGC CGT CTG GAT | 1488 |
| 481 | E V R D D V P L L G F I G R L D | 496 |
| 1489 | GGA CAG AAG GGC GTC GAC ATC ATC GCG GAC GCG ATG CCG TGG ATC GCG | 1536 |
| 497 | G Q K G V D I I G D A M P W I A | 512 |
| 1537 | GGG CAG GAC CTG CAG CTG GTG ATG CTG GGC ACC GGC CCA CCT GAC CTG | 1584 |
| 513 | G Q D V Q L V M L G T G P P D L | 528 |
| 1585 | GAA CGA ATG CTG CAG CAC TTG GAG CGG GAG CAT CCC AAC AAG GTG CGC | 1632 |
| 529 | E R M L Q H L E R E H P N K V R | 544 |
| 1633 | GGG TGG GTC GGG TTC TCG CTC CTA ATG GTG CAT CGC ATC ACG CCG GGC | 1680 |
| 545 | G W V G F S V L M V H R I T P G | 560 |
| 1681 | GCC AGC GTG CTG GTG ATG CCC TCC CGC TTC GCC GGC GGG CTG AAC CAG | 1728 |
| 561 | A S V L V M P S R F A G G L N Q | 576 |
| 1729 | CTC TAC CCG ATG GCA TAC GGC ACC GTC CCT GTG GTG CAC GCC GTG GGC | 1776 |
| 577 | L Y A M A Y G T V P V V H A V G | 592 |
| 1777 | GGG CTC AGG GAC ACC GTG GCG CCG TTC GAC CCG TTC GGC GAC GCC GGC | 1824 |
| 593 | G L R D T V A P F D P P G D A G | 608 |
| 1825 | CTC GGG TGG ACT TTT GAC CGC GCC GAG GCC AAC AAG CTG ATC GAG GTG | 1872 |
| 609 | L G W T F D R A E A N K L I E V | 624 |
| 1873 | CTC AGC CAC TGC CTC GAC ACG TAC CGA AAC TAC GAG GAG AGC TGG AAG | 1920 |
| 625 | L S H C L D T Y R N Y E E S W K | 640 |
| 1921 | AGT CTC CAG GCG CGC GGC ATG TCG CAG AAC CTC AGC TGG GAC CAC GCG | 1968 |
| 641 | S L Q A R G M S Q N L S W D H A | 656 |
| 1969 | GCT GAG CTC TAC GAG GAC GTC CTT CTC AAG TAC CAG TGG | 2007 |
| 657 | A E L Y E D V L V K Y Q W | 669 |

Tabelle 3

DNA-Sequenz und abgeleitete Aminosäuresequenz des Gens für lösliche Stärkesynthase IIb in Mais
[SEQ ID NO: 10 und SEQ ID NO: 11]

| | | | |
|------------------------|-------------|----------------|---------|
| AKTENBEZEICH- NUNG: | MSS3FULL.DN | SEQUENZ:NORMAL | 2097 BP |
| CODONTABELLE: | UNIV.TCN | | |
| SEQUENZREGION | | 1-2097 | |
| TRANSLATIONSREGION | | 1-2097 | |

*** DNA-TRANSLATION ***

```

1  ATG CCG GGG GCA ATC TCT TCC TCG TCG TCG GCT TTT CTC CTC CCC GTC 48
1  K  P  G  A  I  S  S  S  S  S  A  F  L  L  P  V 16

49  GCG TCC TCC TCG CCG CCG CCG AGG CCG GGC AGT GTG GGT GCT GCT CTC 96
17  A  S  S  S  P  R  R  R  R  G  S  V  G  A  A  L 32

97  CCG TCG TAC GGC TAC AGC GGC GCG CAG CTG CCG TTG CAT TGG GCG CCG 144
33  R  S  Y  G  Y  S  G  A  E  L  R  L  H  W  A  R 48

145  CCG GGC CCG CCT CAG GAT GGA GCG GCG TCG GTA CCG GCC GCA GCG GCA 192
49  R  G  P  P  Q  D  G  A  A  S  V  R  A  A  A  A 64

193  CCG GCC GGG GGC GAA AGC GAG GAG GCA GCG AAG AGC TCC TCC TCG TCC 240
65  P  A  G  G  E  S  E  E  A  A  K  S  S  S  S  S 80

241  CAG GCG GGC GCT GTT CAG GGC AGC ACG GCG AAG GCT GTG GAT TCT GCT 288

```

| | | | | | | | | | | | | | | | | | |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 81 | Q | A | G | A | V | Q | G | S | T | A | K | A | V | D | S | A | 96 |
| 289 | TCA | CCT | CCC | AAT | CCT | TTG | ACA | TCT | GCT | CCG | AAG | CAA | AGT | CAG | AGC | GCT | 336 |
| 97 | S | P | P | N | P | L | T | S | A | P | K | Q | S | Q | S | A | 112 |
| 337 | GCA | ATG | CAA | AAC | GGA | ACG | AGT | GGG | GCC | AGC | AGC | GCG | AGC | ACC | GCC | GCG | 384 |
| 113 | A | M | Q | N | G | T | S | G | G | S | S | A | S | T | A | A | 126 |
| 385 | CCG | GTG | TCC | GGA | CCC | AAA | GCT | GAT | CAT | CCA | TCA | GCT | CCT | GTG | ACC | AAG | 432 |
| 129 | P | V | S | G | P | K | A | D | H | P | S | A | P | V | T | K | 144 |
| 433 | AGA | GAA | ATC | GAT | GCC | AGT | GCG | GTG | AAG | CCA | GAG | CCC | GCA | GGT | GAT | GAT | 480 |
| 145 | R | E | I | D | A | S | A | V | K | P | E | P | A | G | D | D | 160 |
| 481 | GCT | AGA | CCG | GTG | GAA | AGC | ATA | GGC | ATC | GCT | GAA | CCG | GTG | GAT | GCT | AAG | 528 |
| 161 | A | R | P | V | E | S | I | G | I | A | E | P | V | D | A | K | 176 |
| 529 | GCT | GAT | GCA | GCT | CCG | GCT | ACA | GAT | GCG | GCG | GCG | AGT | GCT | CCT | TAT | GAC | 576 |
| 177 | A | D | A | A | P | A | T | D | A | A | A | S | A | P | Y | D | 192 |
| 577 | AGG | GAG | GAT | AAT | GAA | CCT | GGC | CCT | TTG | GCT | GGG | CCT | AAT | GTG | ATG | AAC | 624 |
| 193 | R | E | D | N | E | F | G | P | L | A | G | P | N | V | M | N | 208 |
| 625 | GTG | GTG | GTG | GTG | GCT | TCT | GAA | TGT | GCT | CCT | TTC | TGC | AAG | ACA | GGT | GGC | 672 |
| 209 | V | V | V | V | A | S | E | C | A | P | F | C | K | T | G | G | 224 |
| 673 | CTT | GGA | GAT | GTG | GTG | GGT | GCT | TTG | CCT | AAG | GCT | CTG | GCG | AGG | AGA | GGA | 720 |
| 225 | L | G | D | V | V | G | A | L | P | K | A | L | A | R | R | G | 240 |
| 721 | CAC | CGT | GTT | ATG | GTG | ATA | CCA | AGA | TAT | GCA | GAG | TAT | CCG | GAA | GCC | | 768 |
| 241 | H | R | V | M | V | V | I | P | R | Y | G | E | Y | A | E | A | 256 |
| 769 | CGG | GAT | TTA | GGT | GTA | ACG | AGA | CGT | TAC | AAG | GTA | GCT | GGA | CAG | GAT | TCA | 816 |
| 257 | R | D | L | G | V | R | R | R | Y | K | V | A | G | Q | D | S | 272 |
| 817 | GAA | GTT | ACT | TAT | TTT | CAC | TCT | TAC | ATT | GAT | GGA | GTT | GAT | TTT | GTA | TTC | 864 |
| 273 | E | V | T | Y | F | H | S | Y | I | D | G | V | D | F | V | F | 288 |
| 865 | GTA | GAA | CCC | CCT | CCC | TTC | CGC | CAC | CGG | CAC | AAT | AAT | ATT | TAT | GGG | GGA | 912 |
| 289 | V | E | A | F | P | F | R | R | R | H | N | N | I | Y | G | G | 304 |
| 913 | GAA | AGA | TTG | GAT | ATT | TTG | AAG | CGC | ATG | ATT | TTG | TTC | TGC | AAG | GCC | GCT | 960 |
| 305 | E | R | L | D | I | L | K | R | M | I | L | F | C | K | A | A | 320 |
| 961 | GTG | GAG | GTT | CCA | TGG | TAT | GCT | CCA | TGT | GGC | GGT | ACT | GTG | TAT | GGT | GAT | 1008 |
| 321 | V | E | V | P | W | Y | A | P | C | G | G | T | V | Y | G | D | 336 |
| 1009 | GGC | AAC | TTA | GTT | TTC | ATT | GCT | AAT | GAT | TGG | CAT | ACC | GCA | CTT | CTG | CCT | 1056 |
| 337 | G | N | L | V | F | I | A | N | D | W | H | T | A | L | L | P | 352 |
| 1057 | GTG | TAT | CTA | AAG | GCC | TAT | TAC | CGG | GAC | AAT | GGT | TTG | ATG | CAG | TAT | GCT | 1104 |
| 353 | V | Y | L | K | A | Y | Y | R | D | N | G | L | M | Q | Y | A | 368 |
| 1105 | CGC | TCT | GTG | CTT | GTG | ATA | CAC | AAC | ATT | GCT | CAT | CAG | GGT | CGT | GGC | CCT | 1152 |
| 369 | R | S | V | L | V | I | H | N | I | A | H | Q | G | R | G | P | 384 |
| 1153 | GTA | GAC | GAC | TTC | GTG | ATT | TTT | GAC | TTG | CCT | GAA | CAC | TAC | ATC | GAC | CAC | 1200 |
| 385 | V | D | D | F | V | N | F | D | L | P | E | H | Y | I | D | H | 400 |
| 1201 | TTC | AAA | CTG | TAT | GAC | AAC | ATT | GGT | GGG | GAT | CAC | AGC | AAC | GTT | TTT | GCT | 1248 |
| 401 | F | K | L | Y | D | N | I | G | G | D | H | S | N | V | P | A | 416 |
| 1249 | CGG | GGG | CTG | AAG | ACG | GCA | GAC | CGG | GTG | GTG | ACC | GTT | AGC | AAT | GGC | TAC | 1296 |
| 417 | A | G | L | K | T | A | D | R | V | V | T | V | S | N | G | Y | 432 |
| 1297 | ATG | TGG | GAG | CTG | AAG | ACT | TGG | GAA | GGC | GGG | TGG | GGC | CTG | CAC | GAC | ATC | 1344 |
| 433 | M | W | E | L | K | T | S | E | G | G | W | G | L | H | D | I | 448 |

```

1345 ATA AAC CAG AAC GAC TGG AAG CTG CAG GGC ATC GTG AAC GGC ATC GAC 1392
449 I N Q N D W K L Q G I V N G I D 464

1393 ATG AGC GAG TGG AAC CCC GCT GTG GAC GTG CAC CTC CAC TCC GAC GAC 1440
465 M S E W N P A V D V H L H S D D 480

1441 TAC ACC AAC TAC ACG TTC GAG ACC CTG GAC ACC GGC AAG CCG CAG TGC 1488
481 Y T N Y T F E T L D T G K R Q C 496

1489 AAG GCC GCC CTG CAG CGG CAG CTG GGC CTG CAG GTC CGC GAC GAC GTG 1536
497 K A A L Q R Q L G L Q V R D D V 512

1537 CCA CTG ATC GGG TTC ATC GGG CGG CTG GAC CAC CAG AAG GGC GTG GAC 1584
513 P L I G F I G R L D H Q K G V D 528

1585 ATC ATC GCC GAC CGG ATC CAC TGG ATC CGG GGG CAG GAC GTG CAG CTC 632
529 I I A D A I H W I A G Q D V Q L 544

1633 GTG ATG CTG GGC ACC GGG CGG GCC GAC CTG GAG GAC ATG CTG CGG CGG 1680
545 V M L G T G R A D L E D M L R R 560

1681 TTC GAG TCG GAG CAC ACC GAC AAG GTG CGC GCG TGG GTG GGG TTC TCG 1728
561 P E S E H S D K V R A W V G F S 576

1729 GTG CCC CTG GCG CAC CGC ATC ACC CGC GGC GCG GAC ATC CTG CTG ATG 1776
577 V P L A H R I T A G A D I L L M 592

1777 CCG TCG CGG TTC GAG CCG TCG GGG CTG AAC CAG CTC TAC GCC ATG GCG 1824
593 P S R F E P C G L N Q L Y A M A 608

1825 TAC GGG ACC GTG CCC GTG GTG CAC GCC CTG GGG GGG CTC CGG GAC ACG 1872
609 Y G T V P V V H A V G G L R D T 624

1873 GTG GCG CCG TTC GAC CCG TTC AAC GAC ACC GGG CTC GGG TGG ACG TTC 1920
625 V A P F D P F N D T G L G W T F 640

1921 GAC CGC GCG GAG GCG AAC CGG ATG ATC GAC GCG CTC TCG CAC TGC CTC 1968
641 D R A E A N R M I D A L S H C L 556

1969 ACC ACG TAC CGG AAC TAC AAG GAG AGC TGG CGC GCC TGC AGG GCG CGC 2016
657 T T Y R N Y K E S W R A C R A R 672

2017 GGC ATG GCC GAG GAC CTC AGC TGG GAC CAC GCC GCC GTG CTG TAT GAG 2064
673 G M A E D L S W D H A A V L Y E 688

2065 GAC GTG CTC GTC AAG GCG AAG TAC CAG TGG TGA 2097
689 D V L V K A K Y Q W * 699

```

Tabelle 4

DNA-Sequenz und abgeleitete Aminosäuresequenz des Gens für lösliche Stärkesynthase I in Mais
[SEQ ID NO: 12 und SEQ ID NO: 13]

| | | | |
|------------------------|-------------|----------------|---------|
| AKTENBEZEICH- NUNG: | MSS1FULL.DN | SEQUENZ:NORMAL | 1752 BP |
| CODONTABELLE: | UNIV.TCN | | |
| SEQUENZREGION | | 1-1752 | |
| TRANSLATIONSREGION | | 1-1752 | |

| | |
|---|-----|
| TGC GTC GCG GAG CTG AGC AGG GAG GGG CCC GCG CCG CGC CCG CTG CCA Cys Val Ala Glu Leu Ser Arg Glu Gly Pro Ala Pro Arg Pro Leu Pro 700 705 710 715 | 48 |
| CCC GCG CTG CTG GCG CCC CCG CTC GTG CCC GGC TTC CTC GCG CCG CCG Pro Ala Leu Leu Ala Pro Pro Leu Val Pro Gly Phe Leu Ala Pro Pro 720 725 730 | 96 |
| GCC GAG CCC ACC GGT GAG CCG GCA TCG ACG CCG CCG CCC GTG CCC GAC Ala Glu Pro Thr Gly Glu Pro Ala Ser Thr Pro Pro Pro Val Pro Asp 735 740 745 | 144 |
| GCC GGC CTG GGG GAC CTC GGT CTC GAA CCT GAA GGG ATT GCT GAA GGT Ala Gly Leu Gly Asp Leu Gly Leu Glu Pro Glu Gly Ile Ala Glu Gly 750 755 760 | 192 |
| TCC ATC GAT AAC ACA GTA GTT GTG GCA AGT GAG CAA GAT TCT GAG ATT Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln Asp Ser Glu Ile 765 770 775 | 240 |
| GTG GTT GGA AAG GAG CAA GCT CGA GCT AAA GTA ACA CAA AGC ATT GTC Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr Gln Ser Ile Val 780 785 790 795 | 288 |
| TTT GTA ACC GGC GAA GCT TCT CCT TAT GCA AAG TCT GGC GCT CTA GGA Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser Gly Gly Leu Gly 800 805 810 | 336 |
| GAT GTT TGT GGT TCA TTG CCA GTT GCT CTT GCT GCT CGT GGT CAC CGT Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala Arg Gly His Arg 815 820 825 | 384 |
| GTG ATG GTT GTA ATG CCC AGA TAT TTA AAT GGT ACC TCC GAT AAG AAT Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr Ser Asp Lys Asn 830 835 840 | 432 |
| TAT GCA AAT GCA TTT TAC ACA GAA AAA CAC ATT CCG ATT CCA TGC TTT Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg Ile Pro Cys Phe 845 850 855 | 480 |
| GGC GGT GAA CAT GAA GTT ACC TTC TTC CAT GAG TAT AGA GAT TCA GTT Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr Arg Asp Ser Val 860 865 870 875 | 528 |
| GAC TGG GTG TTT GTT GAT CAT CCC TCA TAT CAC ACA CCT GGA AAT TTA Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg Pro Gly Asn Leu 880 885 890 | 576 |
| TAT CGA GAT AAG TTT GGT GCT TTT GGT GAT AAT CAG TTC AGA TAC ACA Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln Phe Arg Tyr Thr 895 900 905 | 624 |
| CTC CTT TGC TAT GCT GCA TGT GAG GCT CCT TTG ATC CTT GAA TTG GGA Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile Leu Glu Leu Gly 910 915 920 | 672 |
| GGA TAT ATT TAT GGA CAG AAT TGC ATG TTT GTT GTC AAT GAT TGG CAT Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val Asn Asp Trp His 925 930 935 | 720 |
| GCC AGT CTA GTG CCA GTC CTT CTT GCT GCA AAA TAT AGA CCA TAT GGT Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly 940 945 950 955 | 768 |
| GTT TAT AAA GAC TCC CGC AGC ATT CTT GTA ATA CAT AAT TTA GCA CAT Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His Asn Leu Ala His 960 965 970 | 816 |

| | |
|---|------|
| CAG GGT GTA GAG CCT GCA AGC ACA TAT CCT GAC CTT GGC TTG CCA CCT Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro 975 980 985 | 864 |
| GAA TGG TAT GGA GCT CTG GAG TGG GTA TTC CCT GAA TGG GCG AGC AGG Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg 990 995 1000 | 912 |
| CAT GCC CTT GAC AAG GGT GAG GCA GTT AAT TTT TTG AAA GGT GCA GTT His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val 1005 1010 1015 | 960 |
| CTG ACA GCA GAT CGA ATC GTG ACT GTC AGT AAG GGT TAT TCG TGG GAG Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly Tyr Ser Trp Glu 1020 1025 1030 1035 | 1008 |
| GTC ACA ACT GCT GAA GGT GGA CAG GGC CTC AAT GAG CTC TTA AGC TCC Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser 1040 1045 1050 | 1056 |
| AGA AAG ACT GTA TTA AAC GGA ATT GTA AAT GCA ATT GAC ATT AAT GAT Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp 1055 1060 1065 | 1104 |
| TGG AAC CCT GCC ACA GAC AAA TGT ATC CCC TGT CAT TAT TCT GTT GAT Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His Tyr Ser Val Asp 1070 1075 1080 | 1152 |
| GAC CTC TCT GGA AAG GCC AAA TGT AAA GGT GCA TTG CAG AAG GAG CTG Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu Gln Lys Glu Leu 1085 1090 1095 | 1200 |
| GGT TTA CCT ATA AGG CCT GAT GTT CCT CTG ATT GGC TTT ATT GGA AGG Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly Phe Ile Gly Arg 1100 1105 1110 1115 | 1248 |
| TTG GAT TAT CAG AAA GGC ATT GAT CTC ATT CAA CTT ATC ATA CCA GAT Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu Ile Ile Pro Asp 1120 1125 1130 | 1296 |
| CTC ATG CCG GAA GAT GTT CAA TTT GTC ATG CTT GGA TCT GGT GAC CCA Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro 1135 1140 1145 | 1344 |
| GAG CTT GAA GAT TGG ATG AGA TCT ACA GAC TCG ATC TTC AAG GAT AAA Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile Phe Lys Asp Lys 1150 1155 1160 | 1392 |
| TTT CGT GGA TGG GTT GGA TTT AGT GTT CCA GTT TCC CAC CGA ATA ACT Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr 1165 1170 1175 | 1440 |
| GCC GGC TGC GAT ATA TTG TTA ATG CCA TCC AGA TTC GAA CCT TGT GGT Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly 1180 1185 1190 1195 | 1488 |
| CTC AAT CAG CTA TAT GCT ATG CAG TAT GGC ACA GTT CCT GTT GTC CAT Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His 1200 1205 1210 | 1536 |
| GCA ACT GGC GGC CTT AGA GAT ACC GTG GAG AAC TTC AAC CCT TTC GGT Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe Asn Pro Phe Gly 1215 1220 1225 | 1584 |
| GAG AAT GGA GAG CAG GGT ACA GGG TGG GCA TTC GCA CCC CTA ACC ACA Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala Pro Leu Thr Thr 1230 1235 1240 | 1632 |

```

GAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC TAC ATA CAG GGA      1680
Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly
1245      1250      1255

ACA CAA GTC CTC CTG GGA AGG GCT AAT GAA GCG AGG CAT GTC AAA AGA      1728
Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg
1260      1265      1270      1275

CTT CAC GTG GGA CCA TGC CGC TGA      1752
Leu His Val Gly Pro Cys Arg *
1280

```

(2) INFORMATION ZU SEQ ID NO: 13:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 584 Aminosäuren
 (B) ART: Aminosäure
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

```

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

```

Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly
 245 250 255
 Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His Asn Leu Ala His
 260 265 270
 Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro
 275 280 285
 Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg
 290 295 300
 His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val
 305 310 315 320
 Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly Tyr Ser Trp Glu
 325 330 335
 Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser
 340 345 350
 Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp
 355 360 365
 Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His Tyr Ser Val Asp
 370 375 380
 Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu Gln Lys Glu Leu
 385 390 395 400
 Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly Phe Ile Gly Arg
 405 410 415
 Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu Ile Ile Pro Asp
 420 425 430
 Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro
 435 440 445
 Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile Phe Lys Asp Lys
 450 455 460
 Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr
 465 470 475 480
 Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly
 485 490 495
 Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His
 500 505 510
 Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe Asn Pro Phe Gly
 515 520 525
 Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala Pro Leu Thr Thr
 530 535 540
 Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly
 545 550 555 560
 Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg
 565 570 575
 Leu His Val Gly Pro Cys Arg *
 580

Tabelle 5

mRNA-Sequenz und abgeleitete Aminosäuresequenz des Verzweigungsenzyms-II-Gens aus Mais und des
Transitpeptids
[SEQ ID NO: 14 und SEQ ID NO: 15]

| | |
|-----------------|--|
| LOCUS | MZEGLUTRN 2725 Bp ss-mRNA PLN |
| DEFINITION | Mais-Stärkeverzweigungsenzym-II-mRNA, Vollst. CDS. |
| ZUGANGSNUMMER | L08065 |
| SCHLÜSSELWÖRTER | 1,4-alpha-Glucan Verzweigungsenzym; Amylo-Transglykosylase; Glucanotransferase; Stärkeverzweigungsenzym II |
| HERKUNFT | Zea mays-cDNA bis mRNA |
| ORGANISMUS | Zea mays Eukaryota; Plantae; Embryobionta; Magnoliophyta; Liliopsida; Commelinidae, Cyperales; Poaceae |
| REFERENZ | 1 (Basen 1 bis 2725) |
| AUTOREN | Fisher, D. K., Boyer, C. D. und Hannah, L. C. |
| TITEL | Starch branching enzyme II from maize endosperm |
| ZEITSCHRIFT | Plant Physiol. 102, 1045–1046 (1993) |
| STANDARD | Vollautomatisch |
| KOMMENTAR | NCBI gi: 168482 |
| MERKMALE | Lage/Kennzeichnungen |


```

source      1..2725
             /cultivar="W64Ax182E"
             /dev_stage="29 days post pollination"
             /tissue_type="endosperm"
             /organism="Zea mays"
sig_peptide 91..264
             /codon_start=1
CDS          91..2487
             /EC_number="2.4.1.18"
             /note="NCBI gi: 168483"
             /codon_start=1
             /product="starch branching enzyme II"

```

```

/translation="MAFRVSCAVLGGA VRAPRLTGGGEGSLVFRHTGLFLTRGARVGC
SGTHGAMRAAAAARKAVMVPEGEN DGLASRADSAQFQSDLEVPDISEETTCCAGVAD
AQALNRVRVVPFPFSDGQKIFQIDPMLQGYKYHLEYRYSLYRRIRSDIDEHEGGLEAFS
RSYEKFGFNASAEGIT YREWAPGAFSAALVGDVNNWDPNADRMSKNEFGVWEIFLPNN
ADGTSPIPHGSRVKVRMDTPSGIKDSIPAWIKYSVQAPGEIPYDGIYDPPEEVKYVF
RHAQPKRPKSLRIYETHVGMSSPEPKINTYVNF RDEVLPRIKKLCYNAVQIMAIQHS
YYGSFGYHVTNFFAPSSRFGT PEDLKSLIDRAHELGLLVLM DVVHSHASSNTLDGLNG
FDGTDTHYFHSGPRGHHWMD SRLFN YGNWEVLRFLLSNARWWLEBYKFDGFRFDGVT
SMYTHMGLQVTF TGNFNEYPGFATDVDAVVYLM LVNDLIHGLYPEAVTIGEDVSGMP
TFALPVHDGGVGFDYRMHMAVADK WIDLLKQSD ETWKMGDIVHTLTNRRWLEKCVTYA
ESHDAQALVGDKTIAFWLMDKDMYDFMALDRPSTPTIDRGIALHKMIRLITMGLGGEY
LNFMGNEFGHP EWIDFPRGPQRLPSGKFIPGNNNSYDKCRRRFDLGDADYLR YHGMQE
FDQAMQHLEQKYEFMTSDHQYISRKHEEDKVIVFEKGDLVFVFN FHCNNSYFDYRIGC
RKPGVYKVVLDSDAGLFGGFSRIHHA AEHFTADCSH DNRPYSF SVYTPSRTC VVYAPV

```

```

E"
mat_peptide 265..2487
             /codon_start=1
             /product="starch branching enzyme II"

```

BASENZAHL 727 A 534 C 715 G 749 T

ORIGIN

```

1  GGGCCAGAGC AGACCCGGAT TTCGCTCTTG CGGTCGCTGG GGTTTTTCGA TTGGCTGATC
61  AGTTCGATCC GATCCGGGCTG CGAAGGGGAG ATGGCGTTCC GCGGTTCTGG GCGGGTGCTC
121 GGTGGGGCCG TAAGGGCTCC CCGACTCACC GCGGGCGGGG AGGGTAGTCT AGTCTTCCGG
181 CAGACCGGCC TCTTCTTAAC TCGGGCTGCT GSAGTTGGAT GTTCGGGGAC GCACGGGGCC
241 ATGCGCGGCG CCGGGCGGGG CAGGAAGGGG STCATGGTTC CTCAGGGCCA GAATGATGGC
301 CTGCGATCRA GGGCTTACTC GGCTCAATTC CAGTGGGATC AACTGGAGGT ACCAGACATT
361 TCTGAAGAGA CAACGTTGCG TGCTGGTGTC GCTGATGCTC AAGCCTTCAA CAGAGTTGGA
421 GTGGTCCGCC CACCAAGCGA TGGACAAAAA ATATTCCACA TTGACCCCAT CTTCGAAGGC
481 TATAAGTACC ATCTTGAGTA TCGGTACAGC CTCTATAGAA GAATCCGTC AGACATTGAT
541 GAACATGAGG GAGGCTTGGG AGCCTTCTCC CGTAGTTATG AGAGTTTGG ATTTAATGCC
601 AGCGCGGAGG GTATCAGATA TCGAGAGTGG GCTCCTGGAG CATTTTCTGC AGCATTGGTG
661 GGTGCGGTCA ACACCTGGGA TCCAAATCCA GATCCTATGA GCAAAATGA GTTTGCTGTT
721 TCGGAATTTT TTCTGGCTAA CAATGCGAGT GGTACATCAC CTATTCTCA TGCATCTCGT
781 GTAAAGGTCA GAATGGATAC TCCATCAGGG ATAAACGATT CAATGCCACC CTGGATCAAG
841 TACTCAGTCC AGCGCCGAGG AGAATACCA TATGATGGGA TTTATTATCA TCCTCCTGAA
901 GAGGTAAAGT ATGTGTTTCA GCATGCGCAA CCTAAACGAC CAAATCATT CCGGATATAT
961 GAACACATG TCGGAATGAG TAGCCCGGAA CCGAGATTA ACACATATGT AAACTTTAGG
1021 GATGAAGTCC TCCCAAGAAT AAAAAAAGT GGATACAATC CAGTGCAAT AATGGCAATC
1081 CAAGAGCACT CATATTATGG AAGCTTTGGA TACCATGTAA CTAATTTTCT TCGGCCAAGT
1141 AGTCGTTTTG GTACCCGAGA AGATTGAGG TCTTTGATTG ATAGAGCACA TGACCTTGGT
1201 TTGCTAGTTC TCATGGATGT GGTTCATAGT CATGCGTCAA GTAATCTCT GGATGGGTTG
1261 AATGCTTTTG ATGGTACAGA TACACATTAC TTTCACAGTG GTCCACGTTG CCATCACTGG
1321 ATGTGGGATTT ATCTGAGGCT TAACCTATGG AACTGGGAAG TTTTAAGATT TCTTCTCTCC
1381 AATGCTAGAT GGTGGCTCGA GGAATATAAG TTGATGCTT TCGGTTTGA TGGTGTGACC
1441 TCCATGATGT ACACCTACCA CGGATTACAA GTACATTTA CCGGGAACCT CAATGAGTAT
1501 TTTGCTTTTG CCACCGATGT AGATGCAGTG GTTACTTGA TGCTGGTAA TGAATCTAAT
1561 CATGGACTTT ATCTTGAGGC TGTACCAATT GGTGAAGATG TTACTGCAAT GCTACATT
1621 GCGCTTCTTG TTCACGATGG TGGGCTAGGT TTTGACTATC GGATGCTAT GGTGTGGCT
1681 GACAAATGGA TTGACCTTCT CRAACAAAGT GATGAACTT GGAAGATGG TGTATTGTG
1741 CACACACTGA CAATAGGAG GTGCTTAGAG AAGTGTGTA CTTATGCTGA AAGTCATGAT
1801 CAAGCATTAG TCGGCGACAA GACTATTGCG TTTGGTTGA TGGACAAGGA TATGTATGAT
1861 TTCTGGCCCT TCGATAGACC TTCACTCCT ACCATTGATC GTGGGATAGC ATTACATAAG
1921 ATGATTAGAC TTATCACAAT GGGTTTAGGA GGACAGGCT ATCTTAATT CATGGGHAAT
1981 GAGTTTGGAC ATCTTCAATG GATAGATTTT CCAGAGGCT CCGAAGACT TCCAAGTGGT
2041 AAGTTTATTC CAGGGAATTA CAACAGTTAT GACAAATGTC GTCGAGGATT TGACCTGGGT
2101 GATGCAGACT ATCTTAGGTA TCATGGTATG CAGAGTTTG ATCAGGCAAT GCRACATCTT
2161 GAGCAAAAT ATGAATTCAAT GRCATCTGAT CACCAATATA TTTCCCGAA ACATGAGGAG
2221 CATAAGGTGA TTGTGTTCCA ARAGGGAGAT TTGGTATTTG TCTTCAACTT CCACTGCAAC
2281 AACAGCTATT TTGACTACCG TATTGGTTGT CGAAGGCTG GGTGTATAA GGTGGTCTTG
2341 GACTCCGAGC CTGCACTATT TGGTGGATTT AGCAGGATCC ATCAGCGAGC CGAGCACTTC
2401 ACCGCGGACT GTTCGCTATG TAATAGGCCA TATTCAATCT CGGTTTATAC ACCAAGCAGA
2461 ACATGTGTCG TCTATGCTCC ACTGGAGTGA TAGCGGGSTA CTCCTTGTG CCGGGCATGT
2521 GTGGGCTGT CGATGTGAGG AAAACCTTC TTCAAACCC GGCAGATGCA TGCATGCATG
2581 CTACAATAAG GTTCTGATAC TTTAATCGAT GCTGGAAAGC CCATGCATCT CGCTCGCTTG
2641 TCCTCTCTAT ATATATAAGA CCTTCAAGGT GTCAATTAAA CATAGAGTTT TCCTTTTTCG
2701 CTTTCTTAAA AAAAAAAAAA AAAAA

```

Tabelle 6

mRNA-Sequenz und abgeleitete Aminosäuresequenz des Verzweigungsenzyms-I-Gens aus Mais und des
Transitpeptids
[SEQ ID NO: 16 und SEQ ID NO: 17]

| | |
|-----------------|---|
| LOCUS | MZEBEI 2763 Bp ss-mRNA PLN |
| DEFINITION | Mais-Stärkeverzweigungsenzym-I-mRNA (BE-I). |
| ZUGANGSNUMMER | D11081 |
| SCHLÜSSELWÖRTER | Verzweigungsenzym I; |
| HERKUNFT | Zea mays L (Inzuchtlinie Oh43), cDNA bis mRNA |
| ORGANISMUS | Zea mays Eukaryota; Plantae; Embryobionta; Magnoliophyta; Liliopsida; Commelinidae; Poaceae |
| REFERENZ | 1 (Basen 1 bis 2763) |
| AUTOREN | Baba, T., Kimura, K., Mizuno, K., Etoh, H. Ishida, Y., Shida, O. und Arai, Y. |
| TITEL | Sequence conservation of the catalytic regions of Amylolytic enzymes in maize branching enzyme-I |
| ZEITSCHRIFT | Biochem. Biophys. Res. Commun. 181, 87-94 (1991) |
| STANDARD | Vollautomatisch |
| KOMMENTAR | Eingereicht (30-APR-1992) beim DDBJ durch: Tadashi Baba Institut für angewandte Biochemie Universität Tsukuba Tsukuba, Ibaraki 305 Japan Tel: 0298-53-6632 Fax: 0298-53-6632 NCBI gi: 217959 |
| MERKMALE | Lage/Kennzeichnungen |

source 1..2763
 /organism="Zea mays"
 CDS <1..2470
 /note="NCBI gi: 217960"
 /codon_start=2
 /product="branching enzyme-1 precursor"

/translation="LCLVSPSSSPTPLPPPRSRSHADRAAPPGIAGGQNVRLSVLSV
 QCKARRSGVRKVKSKFATAATVQEDKTMATAKGDVDHLPFYDLPKLEIFKDHFRYRM
 KRFLQKGSIEENEGSLESFSKGYLKFGINTNEDGTVYREWAPAAQEALIGDFNDWN
 GANHKKMEKDKFCGVWSIKIDHVKGKPAIPHNSKVKFRFLHGGVWVDRI PALIRYATVDA
 SKFGAPYDGVHWDPPASERYTFKHFRPSKPAAPRIYEAHVMSGGEKPAVSTYREFADN
 VLPRIIRANNYNTVQLMAVMENSYASFCYHVTNFFAVSSHSGTFEDLKYLVDKAHSLG
 LEVLMDEVHSHASHNNVTDLNGYDVGQSTQESYFHAGDRGYHKLWDSRLFMYAHWEVL
 RFLLSNLRYWLDEFMFQGFDFDGVTSMLYHHHGINVGFTGNYQEYFSLDTAVDAVVYM
 MLANHLMHKLLEATVVAEDVSGMPVLCRPVDEGGVGFYRLAMAI PDRWIDYLKKNKD
 DSEWSMGEIAHTLTNRRRYTEKCIAYAESHQQSIVGDKTIAFLMDKEMYTGMSDLQPA
 SPTIDRGIALQKMIHFITMALGGDGYLNFMCNEFGHPWIDFPREGNNWSYDKCRRQW
 SLVDTDLHLYKYMNAFOQAMNALDERFSFLSSSKQIVSDMNDEEKVIVFERGDLVFVF
 NFHPKKTYEGYKVGCDLPQKYRVALDSDALVFGGHRVGHVDVHFTSPEGVPGVPETN
 FNNRPNSPKVLSPPRTCVAYYRVDEAGAGRLHAKAETGKTPAESIDVKASRASSKE
 DKEATAGGKKGWKFARQPSDQDTK"

transit_peptide 2..190
 mat_peptide 191..2467
 /EC_number="2.4.1.18"
 /codon_start=1
 /product="branching enzyme-1 precursor"
 polyA signal 2734..2739

BASENZAHL 719 A 585 C 737 G 722 T

ORIGIN

```

1  GCTGTGCTC GTGTGCGCCT CTTCCTCGCC GACTCCGCTT CCGCCGCGGC GCGGCTCTCG
61  CTGGCATGCT GATCGGGCGG CACCGCCGGG GATCGCGGCT GCGGCGCAATG TCGGCTCTGAG
121 TGTCTTGTCT GTCCAGTGCA AGGCTCGCCG GTCAGGGGTG CCGAAGCTCA AGAGCAAATT
181 CGCCACTGCA GCTACTGTGC AAGAAGATAA AACTATGGCA ACTGCCAAAG GCGATGTCTGA
241 CCATCTCCCC ATATACGACC TGGACCCCAA GCTGGAGATA TTCAAGGACC ATTTCTAGGTA
301 CCGGATGAAA AGATTCTTAG AGCAGAAAGG ATCAATTGAA GAAAATGAGG GAAGTCTTGA
361 ATCTTTTCTT AAAGGCTATT TCAAAATTGG GATTAATACA AATGAGGATG GAAGTGTATA
421 TCGTGAAATG GCACCTGCTG CGCAGSAGGC AGAGCTTATT GCTGACTTCA ATGACTGGAA
481 TGGTGCAAAC CATAAGATGG AGAAGGATAA ATTTGGTGTG TGGTCGATCA AAATTGACCA
541 TCTCAAAGGG AAACCTGCCA TCCCTCACAA TTCCAAAGGT AAATTTCGCT TTCTACATGG
601 TGCAGTATGG GTTGATCGTA TTCCAGCATT GATTCGTTAT GCGACTGTTG ATGCCTCTAA

```

```

661 ATTTGGAGCT CCTATGATG GTGTTGATTG GGATCCTCCT GCTTCTGAAA GGTACACATT
721 TAAGCATCCT CGGCTTCAE AGGCTGCTGC TCCAGGTATC TATGAAGCCC ATGTAGGTAT
781 GAGTGGTGAA AAGCCAGCAG TAAGCACATA TAGGGAATTT GCAGACAATG TGTGGCCAGC
841 CATACGAGCA AATAACTACA ACACAGTTCA GTTGATGGCA GTTATGGAGC ATTCTGACTA
901 TGCTTCTTTC GGGTACCATG TGACAAATTT CTTTGGGCTT AGCAGCAGAT CAGGACACCC
961 AGAGGACCTC AATATCTTG TCGATAAGGC ACACAGTTTG GCTTTCGAG TTCTGATGGA
1021 TGTGTCCAT AGCCATCCAA CTAATAATGT CACAGATGGT TTAATGGCT ATGATGTTCG
1081 ACAAGCCACC CAGAGCTCCT ATTTTCATGC GGGAGCTAGA GCTTATCATA AACTTTGGGA
1141 TAGTCCGCTG TCAACTATG CTAAGTGGG GGTATTAGG TTTCTCTTT CTAAGCTGAG
1201 ATATTGGTTC GATGAATTCA TGTTCGATGC CTTCGGATTT GATCAGTTA CATCAATGCT
1261 GTATCATCAC CATEGTATCA ATGTGGGCTT TACTGGAAAC TACCAGGAAT ATTTCAGTTT
1321 GGACACAGCT GTGATGCCG TTGTTTACAT GATGCTTCCA AACCTTTAA TGCACAACT
1381 CTGGCCAGAA CCAACTCTTG TCGCTGAAGA TGTTCAGGC ATGCGGCTC TTGCGCGGCC
1441 AGTTGATCAA CGTGGGCTTG GOTTGACTA TCGGCTGGCA ATGCGTATCC CTGATGATG
1501 GATTGACTAC CTGAACAATA AAGATGACTC TCAGTGGTCC ATGCGTGAAA TAGCCGATAC
1561 TTGACTAAC AGGAGATATA CTGAAAAATG CATCGCATAT GCTGAGAGCC ATGATCAGTC
1621 TATTGTTGGC GACAAACTA TTGCATTTCT CCTGATGCC AAGGATATGT AACTGGCAT
1681 GTCAGACTTG CAGCCTGCTT CACCTACAAT TGATCGAGG ATTGCACTCC AAAAGATGAT
1741 TCACTTCATC ACAATGGCCC TTGGAGGTGA TGGCTACTTG AATTTTATGG GAAATCAGTT
1801 TGGTCACCCA GAATGGATTG ACTTTCCAAG AGAAGGGAAC AACTGGAGCT ATGATAAATG
1861 CAGACGACAG TCGAGCCTTG TCGACACTGA TCACTTGGCG TACAAGTACA TGAATGGCTT
1921 TGACCAAGCG ATCAATGCGC TCGATGAGAG ATTTTCTTTC CTTTCCTCTT CAAAGCAGAT
1981 CGTCAGCCAG ATCAACGATG AGGAAAGST TATTGCTTTT GAAGCTGGAG ATTTAGTTTT
2041 TGTTTTCAAT TTCCATCCCA AGMAACTTA CAGAGGCTAC AAGGTTGGAT GCGATTTGCC
2101 TCGSAAATAC AGGTAGCCCC TCGACTCTGA TGCTCTGCTC TTGCTGGAC ATGGAAGAGT
2161 TCGCCACGAC GTGATCACT TCAAGTCCGC TGAAGCGGTG CCAAGCTGC CCGAAGCAA
2221 GTTCAACAC CGGCGGAAT CGTTCAAGT CTTTCTGCTC CCGGACCTT GTGTGGCTTA
2281 TTAACCTGTA GACGAGCCAG GGGCTGGAGC AGCTCTTCA CCGAAAGCAG AGACAGGAAA
2341 GAGTCTTCCA CCAAGAGCCA TCGAGTCAA AGCTTCAGA GCTACTGCA AAGAGACAA
2401 GAGGCGACG GCTGCTGCA AGAAGGATG GAAGTTGCG CCGACGCTAT CCGATCAGA
2461 TACCAATGA AGCCACGAT CTTGCTGAG GACTGAGTG GCTGCGGCG CCGTGTAGT
2521 AGTCTGCTC TACTGAGTA GCGGCGGCTG GCGGCTTGG AACGCTCCTT TCGTGTAGCT
2581 TCGACGCGAC TGCTGCTCA TCAACGAGCA GGCAGGCACT GCTTGTATAG CTTTCTAGA
2641 ATAATAATCA GGGATGGATG GATGCTGCTT ATTGCTATC TGCTTAGAG TGCATGTGCC
2701 CAGTTTGTAT GTACAGGAGC AGTTCGCTC CAGATAAAA AAAACTTCT TGGGGGGTCT
2761 TTC

```

Tabelle 7

Codiersequenz und abgeleitete Aminosäuresequenz für die Transitpeptidregion des Gens für lösliche Stärke-synthase I aus Mais

[SEQ ID NO: 18 und SEQ ID NO: 19]

| | | | |
|------------------------|-------------|----------------|--------|
| AKTENBEZEICH- NUNG: | MSS1TRTP.DN | SEQUENZ:NORMAL | 153 BP |
| CODONTABELLE: | UNIV.TCN | | |
| SEQUENZREGION | | 1-153 | |
| TRANSLATIONSREGION | | 1-153 | |

*** DNA-TRANSLATION ***

```

1  ATG  GCG  ACG  CCC  TGG  GCG  GTG  GGC  GCG  GCG  TGC  CTC  CTC  CTC  GCG  CCG      48
1  M   A   T   P   S   A   V   G   A   A   C   L   L   L   A   R      16

49  GCC  GGC  TGG  CCG  GCC  GGC  GTC  GGC  GAC  GGG  GCG  GCG  CCS  CCG  AGG  CTC      96
17  A   A   W   P   A   A   V   G   D   R   A   R   P   R   R   L      32

97  CAG  GGC  GTG  CTC  CGC  CGC  GGG  TGC  CTC  GCG  GAG  CTG  AGC  AGC  GAG  GGG      144
33  Q   R   V   L   R   R   R   C   V   A   E   L   S   R   E   G      48

145  CCC  CAT  ATG      153
49  P   H   M      51

```

GFP-Konstrukte:

1. GFP ausschließlich in pET-21a:

pEXS115 wird mit NdeI und XhoI verdaut, und das 740 Bp Fragment, das die SGFP-Codiersequenz enthält, wird in die NdeI- und XhoI-Spaltstellen von pET-21a (Novagen 601 Science Dr. Madison WI) subkloniert. (Siehe Karte in [Fig. 2b](#) GFP-21a).

2. GFP, das leserastergerecht am 5'-Ende des reifen Vollängen-WX subkloniert wurde:

Das 740 Bp große NdeI-Fragment, das SGFP enthielt, aus pEXS114 wird in die NdeI-Spaltstelle von

pEXSWX subkloniert (siehe Karte in [Fig. 3b](#) GFG-FLWX).

3. GFP, das leserastergerecht am 5'-Ende des N-terminal verkürzten WX subkloniert wurde: WX, das um 700 Bp am N-Terminus verkürzt wurde. Das 1 kB große BamHI-Fragment, das für den C-Terminus von WX codiert, aus pEXSWX, wird in die BglII-Stelle von pEXS115 subkloniert. Anschließend wird das gesamte um SGFP verkürzte WX-Fragment als NdeI-HindIII-Fragment in PET21a subkloniert (siehe Karte in [Fig. 3b](#) GFP-BamHIWX).

4. GFP, das leserastergerecht am 5'-Ende des N-terminal verkürzten WX subkloniert wurde: WX, das um 100 Bp am N-Terminus verkürzt wurde. Das 740 Bp große NdeI-NcoI-Fragment, das SGFP enthält, aus pEXS115, wird an den NdeI- und NcoI-Spaltstellen in pEXSWX subkloniert (siehe Karte in [Fig. 4](#) GFP-NcoWX).

Beispiel Drei:

Plasmid-Transformation in Bakterien:

[0102] Präparation von kompetenten Escherichia coli-Zellen:

1. 2,5 ml LB-Medium mit einer Einzelkolonie des gewünschten E. coli-Stamms inokulieren: Der gewählte Stamm war XLIBLUE DL2IDE3 von (Statagene); inklusive geeignete Antibiotika. Bei 37°C, 250 U/min über Nacht züchten.
2. 100 LB-Medium mit einer 1:50-Verdünnung der Übernachtskultur inokulieren, inklusive geeignete Antibiotika. Bei 37°C, 250 U/min solange züchten, bis die $OD_{600} = 0,3-0,5$ beträgt.
3. Kultur in sterile Zentrifugenflasche überführen und 15 Minuten lang auf Eis kühlen.
4. 5 Minuten bei $3000 \times g$ (4°C) zentrifugieren.
5. Pellet erneut in 8 ml eiskaltem Transformationspuffer suspendieren. 15 Minuten auf Eis inkubieren.
6. 5 Minuten bei $3000 \times g$ (4°C) zentrifugieren.
7. Pellet erneut in 8 ml eiskaltem Transformationspuffer 2 suspendieren. In aliquote Teile teilen, in flüssigem Stickstoff schockgefrieren, bei -70°C aufbewahrt.

[0103]

| Transformationspuffer 1 | | Transformationspuffer 2 | |
|--------------------------------------|---------|--------------------------------------|---------|
| RbCl | 1,2 g | MOPS (10 mM) | 0,209 g |
| MnCl ₂ ·4H ₂ O | 0,99 g | RbCl | 0,12 g |
| Kaliumacetat | 0,294 g | CaCl ₂ ·2H ₂ O | 1,1 g |
| CaCl ₂ ·2H ₂ O | 0,15 g | Glycerin | 15 g |
| Glycerin | 15 g | dH ₂ O | 100 ml |
| dH ₂ O | 100 ml | pH auf 6,8 mit NaO | H |
| pH auf 5,8 mit 0,2 M Essigsäure | | Sterilfiltrieren | |
| Sterilfiltrieren | | | |

[0104] Transformation von Escherichia coli mittels der Rubidiumchlorid-Hitzeschockmethode: (Hanahan, D. (1985) in DNA cloning: a practical approach (Hrsg.: D. M. Glover), S. 109–135, IRL Press.

1. 1–5 µl DNA 30 Minuten mit 150 µl kompetenten E. coli-Zellen auf Eis inkubieren.
2. 45 Sekunden Hitzeschock bei 42°C.
3. Sofort 2 Minuten lang auf Eis geben.
4. Mit 600 µl LB-Medium versetzen und 1 Stunde lang bei 37°C inkubieren.
5. Auf LB-Agar, der die entsprechenden Antibiotika enthält, ausplattieren.

[0105] Dieses Plasmid wird das Hybridpolypeptid, das das "Green Fluorescent Protein" enthält, innerhalb der Bakterien exprimieren.

Beispiel Vier:

Expression des Konstrukts in E. coli:

1. 3 ml LB mit E. coli, der das interessierende Plasmid enthält, inokulieren. Geeignete Antibiotika mitverwenden. 37°C, 250 U/min, über Nacht.

2. 100 ml LB mit 2 ml Übernachtskultur inokulieren.
- Geeignete Antibiotika mitverwenden. Bei 37°C, 250 U/min züchten.
3. Bei einer OD₆₀₀ von ungefähr 0,4–0,5 bei Raumtemperatur aufbewahren, 200 U/min.
4. Bei einer OD₆₀₀ von ungefähr 0,6–0,8 mit 100 µl 1M IPTG induzieren. IPTG-Endkonzentration beträgt 1 mM.
5. Bei Raumtemperatur, 200 U/min, 4–5 Stunden züchten.
6. Zellen abzentrifugieren.
7. In flüssigem Stickstoff schockgefrieren und bei –70°C bis zur Verwendung aufbewahren.

[0106] Die Zellen können erneut in dH₂O suspendiert und unter UV-Licht ($\lambda_{\text{max}} = 395 \text{ nm}$) auf Eigenfluoreszenz beobachtet werden. Die Zellen können jedoch auch mit Ultraschall behandelt werden und ein aliquoter Teil des Zellextrakts kann mittels SDS-PAGE aufgetrennt und unter UV-Licht zum Nachweis der GFP-Fluoreszenz beobachtet werden. Handelt es sich bei dem verwendeten Protein um ein "Green Fluorescent Protein", so kann das Vorhandensein des Proteins in dem lysierten Material unter UV bei 395 nm im Leuchtkasten ausgewertet werden und das typische grüne Leuchten kann identifiziert werden.

Beispiel Fünf:

Plasmidextraktion aus Bakterien:

[0107] Es folgt nun eines der vielen allgemein verwendeten Plasmidaufreinigungsprotokolle mittels alkalischer Lyse, das sich für die Durchführung der vorliegenden Erfindung eignet.

1. 100–200 ml LB-Medium mit einer Einzelkolonie von mit einem der oben beschriebenen Plasmide transformierten E. coli inokulieren. Geeignete Antibiotika mitverwenden. Bei 37°C, 250 U/min über Nacht züchten.
2. 10 Minuten bei 5.000 × g (4°C) zentrifugieren.
3. Zellen erneut in 10 ml Wasser suspendieren, in ein 15-ml-Zentrifugenröhrchen überführen und die Zentrifugation wiederholen.
4. Pellet erneut in 5 ml 0,1 M NaOH (0,5% SDS) suspendieren. 10 Minuten auf Eis inkubieren.
5. Mit 2,5 ml 3 M Natriumacetat (pH 5,2) versetzen, vorsichtig umdrehen und 10 Minuten auf Eis inkubieren.
6. 5 Minuten bei 15.000–20.000 × g (4°C) zentrifugieren.
7. Überstand mit einem gleichen Volumen Phenol:Chloroform:Isoamylalkohol (25:24:1) extrahieren.
8. 10 Minuten bei 6.000–20.000 × g (4°C) zentrifugieren.
9. Wässrige Phase in sauberes Röhrchen überführen und mit 1 Volumenteil Isopropanol fällen.
10. 15 Minuten bei 12.000 × g (4°C) zentrifugieren.
11. Pellet in 0,5 ml TE lösen, mit 20 µl 10 mg/ml RNase versetzen und 1 Stunde bei 37°C inkubieren.
12. Zweimal mit Phenol:Chloroform:Isoamylalkohol (25:24:1) extrahieren.
13. Einmal mit Chloroform extrahieren.
14. Wässrige Phase mit 1 Volumenteil Isopropanol und 0,1 Volumenteil 3 M Natriumacetat fällen.
15. Pellet einmal mit 705%igem Ethanol waschen.
16. Pellet im SpeedVac trocknen und Pellet in TE suspendieren.

[0108] Dieses Plasmid kann nun in andere Wirte inseriert werden.

Tabelle 8

DNA-Sequenz und abgeleitete Aminosäuresequenz der Stärkesynthase-Codierregion aus pEXS52
[SEQ ID No: 20 und SEQ ID NO: 21]

| | | | |
|--------------------|-------------|----------------|---------|
| AKTENBEZEICHNUNG: | MSS1DELN.DN | SEQUENZ:NORMAL | 1626 BP |
| CODONTABELLE: | UNIV.TCN | | |
| SEQUENZREGION | | 1–1626 | |
| TRANSLATIONSREGION | | 1–1626 | |

(xi) SEQUENZBESCHREIBUNG : SEQ ID NO: 20

| | |
|---|-----|
| TGC GTC GCG GAG CTG AGC AGC GAG GAC CTC GGT CTC GAA CCT GAA GGG | 48 |
| Cys Val Ala Glu Leu Ser Arg Glu Asp Leu Gly Leu Glu Pro Glu Gly | |
| 55 60 65 | |
| ATT GCT GAA GGT TCC ATC GAT AAC ACA GTA GTT GTG CCA AGT GAG CAA | 96 |
| Ile Ala Glu Gly Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln | |
| 70 75 80 | |
| GAT TCT GAG ATT GTG GTT GGA AAG GAG CAA GCT CGA GCT AAA GTA ACA | 144 |
| Asp Ser Glu Ile Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr | |
| 85 90 95 | |
| CAA AGC ATT GTC TTT GTA ACC GGC GAA GCT TCT CCT TAT GCA AAG TCT | 192 |
| Gln Ser Ile Val Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser | |
| 100 105 110 115 | |
| GGG GGT CTA GGA GAT GTT TGT GGT TCA TTG CCA GTT GCT GTT GCT GCT | 240 |
| Gly Gly Leu Gly Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala | |
| 120 125 130 | |
| CGT GGT CAC CGT GTG ATG GTT GTA ATG CCC ACA TAT TTA AAT GGT ACC | 288 |
| Arg Gly His Arg Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr | |

| 135 | 140 | 145 | |
|---|-----|-----|------|
| TCC GAT AAG AAT TAT GCA AAT GCA TTT TAC ACA GAA AAA CAC ATT CGG Ser Asp Lys Asn Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg 150 | | | 336 |
| ATT CCA TGC TTT GGC GGT GAA CAT GAA GTT ACC TTC TTC CAT GAG TAT Ile Pro Cys Phe Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr 165 | | | 384 |
| AGA GAT TCA GTT GAC TGG CTG TTT GTT GAT CAT CCC TCA TAT CAC AGA Arg Asp Ser Val Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg 180 | | | 432 |
| CCT GGA AAT TTA TAT GCA GAT AAG TTT GGT GCT TTT GGT GAT AAT CAG Pro Gly Asn Leu Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln 200 | | | 480 |
| TTC AGA TAC ACA CTC CTT TGC TAT GCT GCA TGT GAG GCT CCT TTG ATC Phe Arg Tyr Thr Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile 215 | | | 528 |
| CTT GAA TTG GGA GGA TAT ATT TAT GGA CAG AAT TGC ATG TTT GTT GTC Leu Glu Leu Gly Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val 230 | | | 576 |
| AAT GAT TGG CAT GCC AGT CTA GTC CCA GTC CTT CTT GCT GCA AAA TAT Asn Asp Trp His Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr 245 | | | 624 |
| AGA CCA TAT GGT GTT TAT AAA GAC TCC CGC AGC ATT CTT GTA ATA CAT Arg Pro Tyr Gly Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His 260 | | | 672 |
| AAT TTA GCA CAT CAG GGT GTA GAG CCT GCA AGC ACA TAT CCT GAC CTT Asn Leu Ala His Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu 280 | | | 720 |
| GGG TTG CCA CCT GAA TGG TAT GCA GCT CTG GAG TGG GTA TTC CCT GAA Gly Leu Pro Pro Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu 295 | | | 768 |
| TGG GCG AGG AGG CAT GGC CTT GAC AAG GGT CAG GCA GTT AAT TTT TTG Trp Ala Arg Arg His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu 310 | | | 816 |
| AAA GCT CCA GTT GTG ACA GCA GAT CGA ATC GTG ACT GTC ACT AAG GGT Lys Gly Ala Val Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly 325 | | | 864 |
| TAT TCG TGG GAG GTC ACA ACT GCT GAA GGT GGA CAG GGC CTC AAT GAG Tyr Ser Trp Glu Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu 340 | | | 912 |
| CTC TTA AGC TCC AGA AAG AGT GTA TTA AAC GGA ATT GTA AAT GGA ATT Leu Leu Ser Ser Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile 360 | | | 960 |
| GAC ATT AAT GAT TGG AAC CCT GCC ACA GAC AAA TGT ATC CCC TGT CAT Asp Ile Asn Asp Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His 375 | | | 1008 |
| TAT TCT GTT GAT GAC CTC TCT GGA AAG GCC AAA TGT AAA GGT GCA TTG Tyr Ser Val Asp Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu 390 | | | 1056 |

```

CAG AAG GAG CTG GGT TTA CCT ATA AGG CCT GAT GTT CCT CTG ATT GGC      1104
Gln Lys Glu Leu Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly
405                               410                               415

TTT ATT GGA AGG TTG GAT TAT CAG AAA GGC ATT GAT CTC ATT CAA CTT      1152
Phe Ile Gly Arg Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu
420                               425                               430                               435

ATC ATA CCA GAT CTC ATG CCG GAA GAT GTT CAA TTT GTC ATG CTT GCA      1200
Ile Ile Pro Asp Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly
440                               445                               450

TCT GGT GAC CCA GAG CTT GAA GAT TGG ATG AGA TCT ACA GAG TCG ATC      1248
Ser Gly Asp Pro Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile
455                               460                               465

TTC AAG GAT AAA TTT CCT GGA TGG GTT GGA TTT AGT GTT CCA GTT TCC      1296
Phe Lys Asp Lys Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser
470                               475                               480

CAC CGA ATA ACT GCC GGC TGC GAT ATA TTG TTA ATG CCA TCC AGA TTC      1344
His Arg Ile Thr Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe
485                               490                               495

GAA CCT TGT GGT CTC AAT CAG CTA TAT GCT ATG CAG TAT GGC ACA GTT      1392
Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val
500                               505                               510                               515

CCT GTT GTC CAT GCA ACT GCG GGC CTT ACA GAT ACC CTG GAG AAC TTC      1440
Pro Val Val His Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe
520                               525                               530

AAC CCT TTC GGT GAG AAT GGA GAG CAG GGT ACA GGG TGC GCA TTC GCA      1488
Asn Pro Phe Gly Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala
535                               540                               545

CCC CTA ACC ACA GAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC      1536
Pro Leu Thr Thr Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile
550                               555                               560

TAC ATA CAG GGA ACA CAA GTC CTC CTG GGA AGG GCT AAT GAA GCG AGG      1584
Tyr Ile Gln Gly Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg
565                               570                               575

CAT GTC AAA ACA CTT CAC GTG GGA CCA TGC CGC TGA      1620
His Val Lys Arg Leu His Val Gly Pro Cys Arg *
580                               585                               590

```

(2) INFORMATION ZU SEQ ID NO: 21:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 540 Aminosäuren
 (B) ART: Aminosäure
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

```

Cys Val Ala Glu Leu Ser Arg Glu Asp Leu Gly Leu Glu Pro Glu Gly
1                               5                               10                               15

Ile Ala Glu Gly Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln
20                               25                               30

Asp Ser Glu Ile Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr
35                               40                               45

```

Gln Ser Ile Val Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser
 50 55 60
 Gly Gly Leu Gly Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala
 65 70 75 80
 Arg Gly His Arg Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr
 85 90 95
 Ser Asp Lys Asn Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg
 100 105 110
 Ile Pro Cys Phe Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr
 115 120 125
 Arg Asp Ser Val Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg
 130 135 140
 Pro Gly Asn Leu Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln
 145 150 155 160
 Phe Arg Tyr Thr Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile
 165 170 175
 Leu Glu Leu Gly Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val
 180 185 190
 Asn Asp Trp His Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr
 195 200 205
 Arg Pro Tyr Gly Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His
 210 215 220
 Asn Leu Ala His Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu
 225 230 235 240
 Gly Leu Pro Pro Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu
 245 250 255
 Trp Ala Arg Arg His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu
 260 265 270
 Lys Gly Ala Val Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly
 275 280 285
 Tyr Ser Trp Glu Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu
 290 295 300
 Leu Leu Ser Ser Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile
 305 310 315 320
 Asp Ile Asn Asp Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His
 325 330 335
 Tyr Ser Val Asp Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu
 340 345 350
 Gln Lys Glu Leu Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly
 355 360 365
 Phe Ile Gly Arg Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu
 370 375 380
 Ile Ile Pro Asp Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly
 385 390 395 400
 Ser Gly Asp Pro Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile
 405 410 415

```

Phe Lys Asp Lys Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser
      420                      425                      430
His Arg Ile Thr Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe
      435                      440                      445
Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val
      450                      455                      460
Pro Val Val His Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe
      465                      470                      475                      480
Asn Pro Phe Gly Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala
      485                      490                      495
Pro Leu Thr Thr Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile
      500                      505                      510
Tyr Ile Gln Gly Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg
      515                      520                      525
His Val Lys Arg Leu His Val Gly Pro Cys Arg *
      530                      535                      540

```

Beispiel Sechs:

[0109] Bei diesem Versuch wird ein Plasmid eingesetzt, das über einen Maispromoter, ein Maistransitpeptid, eine Stärkebindungsregion des Gens für die Stärkesynthase I und ein daran gebundenes ligiertes Genfragment verfügt. Das in [Fig. 6](#) gezeigte Plasmid enthält die in Tabelle 8 aufgelistete DNA-Sequenz.

[0110] Das Plasmid pEXS52 wurde nach der folgenden Vorschrift konstruiert:

Bei dem für die Konstruktion von transgenen Plasmiden verwendeten Materialien handelt es sich um:

Plasmid pBluescript SK-

Plasmid pMF6 (enthalten den nos3'-Terminator)

Plasmid pHKH1 (enthalten das Mais-adh1-Intron)

Plasmid Mstsl(6-4) (enthalten das Mais-STs-Transitpeptid, Verwendung als Matrice für PCT von stsl-Transitpeptid weg)

Plasmid MstslIII in pBluescript SK-

Primers EXS29 (GTGGATCCATGGCGACGCCCTCGGCCGTGG) [SEQ ID NO:22]

Primer EXS35 (CTGAATTCCATATGGGGCCCCCTCCCTGCTCAGCTC) [SEQ ID NO:23]

beide für PCT des stsl-Transitpeptids verwendet

Primer EXS31 (CTCTGAGCTCAAGCTTGCTACTTTCTTTCTTAATG) [SEQ ID NO:24]

Primer EXS32 (GTCTCCGCGGTGGTGTCTTGCTTCCTAG) [SEQ ID NO:25]

beide für PCR des Mais-10KD-Zein-Promoters verwendet (Journal: Gene 71: 359-370 [1988])
genomische DNA aus Mais A632 (als Matrice für die PCR des Mais-10KD-Zein-Promoters verwendet)

Schritt 1: Mais-10KD-Zein-Promoter in pBluescript SK- klonieren (Bezeichnung: pEXS10zp).

1. PCR des 1,1 kB Mais-10KD-Zein-Promoters

Primer: EXS31, EXS32

Matrice: genomische DNA aus Mais A632

2. 1,1 kB Mais, 10KD Zein-Promoter PCR Produkt in pBluescript SK- Plasmid an der SacI- und SacII-Spaltstelle klonieren (siehe **Fig. 7**)

Schritt 2: NdeI-Spaltstelle in pEXS10zp deletieren (Bezeichnung pEXS10zp-NdeI).

[0111] NdeI wird durch Auffüllen und Stumpfendiges Ligieren des Mais-10KD-Zein-Promoters in pBluescript SK- entfernt.

Schritt 3: Mais-Adh-1-Intron in pBluescriptSK-klonieren (Bezeichnung pEXSadh1).

[0112] Das Mais-Adh1-Intron wird aus dem Plasmid pHKH1 an der XbaI und der BamHI-Stelle freigesetzt. Das Mais-ADH1-Intron (XbaI/BamHI-Fragment) wird in pBluescriptSK- an der XbaI- und BamHI-Stelle kloniert (siehe **Fig. 7**).

Schritt 4: Mais-10KD-Zein-Promoter und Mais-Adh-1-Intron in pBluescriptSK- klonieren (Bezeichnung pEXS10zp-adh1).

[0113] Der Mais-10KD-Zein-Promoter wird aus dem Plasmid pEXS 10zp-NdeI an der SacI- und der SacII-Stelle freigesetzt. Der Mais-10KD-Zein-Promoter (SacI/SacII-Fragment) wird in das Plasmid pEXSadh 1 (das das Mais-and1-Intron enthält) an der SacI- und der SacII-Stelle kloniert (siehe **Fig. 7**).

Schritt 5: Mais-nos3'-Terminator in das Plasmid pEXSadh1 klonieren (Bezeichnung pEXSadh1-nos3').

[0114] Der Mais-nos3'-Terminator wird aus dem Plasmid pMF6 an der EcoRI- und HindIII-Spaltstelle freigesetzt. Der Mais-nos3'-Terminator (EcoRI/HindIII-Fragment) wird in das Plasmid pEXSadh1 an EcoRI und HindIII kloniert (siehe **Fig. 7**).

Schritt 6: Mais-nos3'-Terminator in das Plasmid pEXS10zp-adh1 klonieren (Bezeichnung pEXS10zp-adh1-nos3').

[0115] Der Mais-nos3'-Terminator wird aus dem Plasmid pEXSadh1-nos3' an der EcoRI- und Apal-Spaltstelle freigesetzt. Der Mais-nos3'-Terminator (EcoRI/Apal-Fragment) wird in das Plasmid pEXS10zp-adh1 an EcoRI- und Apal-Spaltstelle kloniert (siehe **Fig. 7**).

Schritt 7: Mais-STSI-Transitpeptid in das Plasmid pEXS10zp-adh1-nos3' klonieren (Bezeichnung: pEXS33).

1. PCR des 150 Bp Mais-STSI-Transitpeptids

Primer: EXS29, EXS35

Matrize: STSI(6-4) Plasmid

2. Das 150 Bp Mais-STSI-Transitpeptid-PCR-Produkt an der EcoRI- und der BamHI-Spaltstelle in das Plasmid pEXS10zp-adh1-nos3' klonieren (siehe **Fig. 7**).

Schritt 8: Ortsgerichtete Mutagenese an dem Mais-STSI-Transitpeptid in pEXS33 (Bezeichnung pEXS33(m)).

[0116] Es liegt eine Mutation (Stop-Codon) auf dem Mais-STSI-Transitpeptid in dem Plasmid pEXS33 vor. Die ortsgerichtete Mutagenese wird durchgeführt, um das Stop-Codon zu einem Nicht-Stop-Codon zu verändern. Das neue Plasmid (enthaltend den Mais 10 kd-Promoter das Mais-STSI-Transitpeptid, das Mais-adh1-Intron, den Mais-nos3'-Terminator) wird mit pEXS33(m) bezeichnet.

Schritt 9: NotI-Spaltstelle in pEXS33(m) deletiert (Bezeichnung pEXS50).

[0117] Die NotI-Spaltstelle wird als pEXS33 durch NotI-Auffüllen, stumpfendiges Ligieren unter Bildung von pEXS50 entfernt (siehe **Fig. 8**).

Schritt 10: Mais-adh1-Intron in pEXS33(m) deletiert (Bezeichnung pEXS60)

[0118] Das Mais-adh1-Intron wird durch Verdauen mit NotI/BamHI entfernt, mit Klenow-Fragment aufgefüllt, stumpfendiges Ligieren unter Bildung von pEXS60 (siehe **Fig. 9**).

Schritt 11: Mais-STSI in pEXS50, pEXS60 klonieren

[0119] Die Mais-STSI wird aus dem Plasmid Mais-STSI in pBluescriptSK- an der NdeI- und der EcoRI-Stelle freigesetzt. Die Mais-STSI (NdeI-EcoRI-Fragment) wird getrennt in pEXS50 und pEXS60, kloniert, Bezeichnung pEXS51 bzw. pEXS61 (siehe **Fig. 8** bzw. 9).

Schritt 12:

[0120] Gen in Tabelle 8 an der NdeI/NotI-Spaltsstelle in pEXS51 klonieren, wodurch man zu pEXS52 gelangt. Weitere ähnliche Plasmide können dadurch hergestellt werden, daß man andere Gene (STSI, II, WX, glgA, glgB, glgB, BEI BEII usw.) an der NdeI/NotI-Spaltsstelle in pEXS51, pEXS61 kloniert.

[0121] Das Plasmid EXS52 wurde in Reis hineintransformiert. Die regenerierten Reispflanzen, die mit pEXS52 transformiert waren, wurden gekennzeichnet und in eine Magenta-Box gesetzt.

[0122] Zwei Geschwister jeder Linie wurden aus der Magenta-Box ausgewählt und in 2,5-Inch-Töpfe, die mit einer Erdmischung (Mischung aus Oberboden mit Torf-Vermiculit 50/50) gefüllt waren, überführt. Die Töpfe wurden in ein Aquarium (Fischtank) mit Wasser in einer Höhe von 1/2 Inch gestellt. Oben wurde abgedeckt, um eine hohe Feuchtigkeit aufrechtzuerhalten (es wurden einige Löcher gemacht, so daß Wärme entweichen konnte). Die Temperatur wurde mit einem Thermometer überwacht. Der Fischtank wurde unter Fluoreszenzlampen gestellt. In der ersten Woche wurde kein Dünger bei den Pflanzen verwendet. Die Photoperiode dauerte von 6 Uhr vormittags – 8 Uhr nachmittags, Minimum 14 Stunden Licht. Die Temperatur betrug mindestens 68°F in der Nacht und 80–90°F während des Tages. Unter dem Fischtank wurde eine Heizmatte verwendet, um wenn erforderlich das Wurzelwachstum zu unterstützen. Die Pflanzen blieben ungefähr eine Woche unter der genannten Bedingung. (Anmerkung: Die Keimpflanzen begannen sich aufgrund der niedrigen Lichtintensität zu strecken).

[0123] Nach der ersten Woche wurde das Aquarium oben geöffnet und die Reistransformanten wurden drei Wochen lang in Wachstumskammern mit hoher Feuchtigkeit und hoher Lichtintensität umgestellt.

[0124] Alternativ dazu kann die Wassermischung im Gewächshaus dazu verwendet werden, eine hohe Feuchtigkeit aufrechtzuerhalten. Die Pflanzen wuchsen drei Wochen lang. Anschließend wurden die Pflanzen in 6-Inch-Töpfe (Minimum: 5-Inch-Töpfe) mit Erdmischung (Oberboden und Torf-Vet, 50/50) umgesetzt. Die Töpfe standen in einer Schale, die mit ½ Inch Wasser gefüllt war. Die Pflanzen wurden einmal pro Woche oder nach dem Bedarf der Pflanzen gemäß ihres Aussehens mit 15-16-17 (N-K-P) gedüngt (250 ppm). Die Pflanzen verblieben in 14 Stunden Licht (Minimum) von 6 Uhr vormittags – 8 Uhr nachmittags bei hoher Lichtintensität, Temperatur 85–90°/70°F Tag/Nacht.

[0125] Die Pflanzen bildeten Reiskörner und die Reiskörner wurden geerntet. Von diesen geernteten Samen kann man die Stärke extrahieren und auch das Vorliegen der ligierten Aminosäuren C, V, A, E, L, S, R, E [SEQ ID NO: 27] in der Stärke in dem Samen analysieren.

Beispiel Sieben:

SBR-Vektor für Pflanzen:

[0126] Das in [Abb. 6](#) gezeigte Plasmid wird an die Verwendung bei einkeimblättrigen Pflanzen, d. h. Mais, adaptiert. Das Plasmid pEXS52 ([Fig. 6](#)) verfügt über einen Promoter, ein Transitpeptid (aus Mais) und ein ligiertes Genfragment (TGC GTC GCG GAG CTG AGC AGG GAG) [SEQ ID NO: 26], das für die Aminosäuresequenz C V A E L S R E [SEQ ID NO: 27] kodiert.

[0127] Dieses Genfragment tritt natürlich in der Nähe des N-terminalen Endes des Gens für die lösliche Stärkesynthase aus Mais (MSTSI) auf. Wie in Tabelle 8 dargestellt, beginnt die SBR der Stärkesynthase ungefähr bei Aminosäure 292. Dieser Vektor wird vorzugsweise in einen Maiswirt hineintransformiert. Das Transitpeptid ist derartig an Mais adaptiert, daß dieser der bevorzugte Wirt ist. Natürlich können Transitpeptid und Promoter gegebenenfalls dahingehend verändert werden, daß sie sich für die gewünschte Wirtspflanze eignen. Nach Transformation mittels "Whiskers"-Technologie (US-Patente Nr. 5,302,523 und 5,464,765) werden die transformierten Wirtszellen nach fachbekannten Verfahren regeneriert, die Transformante wird bestäubt und die erhaltenen Körner können gewonnen und auf das Vorhandensein des Peptids in der Stärke und in den Stärkekörnern analysiert werden.

[0128] Es können die folgenden bevorzugten Gene in Mais zur Verbesserung von Futtermitteln eingesetzt werden: Phytasegen, Somatotropingen, die folgenden verketteten Aminosäuren: AUG AUG AUG AUG AUG AUG AUG AUG [SEQ ID NO: 28]; und/oder AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG [SEQ ID NO: 29] und/oder AAA AAA AAA AAA AAA AAA [SEQ ID NO: 30]; oder eine Kombination der Codons, die für die Aminosäure Lysin in einer Kette oder ein Kombination der Codons, die sowohl für das Lysin- als auch das MethioninCodon kodieren, oder eine beliebige Kombination von zwei oder drei dieser Aminosäuren. Die Länge der Ketten sollte nicht allzulang sein, es scheint jedoch, daß die Kettenlänge nicht kritisch ist. So werden die Aminosäuren innerhalb des Stärkekorns verkapselt oder innerhalb der in dem stärkehaltigen Teil der Wirtspflanze gebildeten Stärke gebunden werden.

[0129] Dieses Plasmid kann in andere Getreide wie Reis, Weizen, Gerste, Hafer, Sorgumhirse oder kleinkörnige Hirse mit wenig oder gar keiner Modifikation des Plasmids hineintransformiert werden. Bei dem Promoter kann es sich um den Promoter des waxy-Gens, dessen Sequenz veröffentlicht worden ist, oder um andere fachbekannte Zein-Promoter handeln.

[0130] Außerdem können diese Plasmide ohne unzumutbare Versuchstätigkeit in zweikeimblättrige Pflanzen wie Kartoffeln, Süßkartoffel, Taro, Yam, Lotus, Cassava, Erdnüsse, Erbsen, Sojabohne, Bohnen oder Kichererbsen hineintransformiert werden. Der Promoter kann dahingehend ausgewählt werden, daß die Stärkespeicherzone von bestimmten zweikeimblättrigen Pflanzen oder Knollen adressiert wird, zum Beispiel kann der Patatin-Promoter für Kartoffelknollen eingesetzt werden.

[0131] Es sind in diesem Fachgebiet verschiedene Verfahren zur Transformation von ein- und zweikeimblättrigen Pflanzen bekannt, und das Verfahren, mit dem die Gene transformiert werden, ist für die vorliegende Erfindung nicht kritisch. Das Plasmid kann in *Agrobacterium tumefaciens* mittels der Gefrier-Tau-Methode von An et al (1988) Binary Vektors, in *Plant Molecular Biology Manual A3*, S. B. Gelvin und R. A. Schilperroot, Hrsg. (Dordrecht, Niederlande: Kluwer Academic Publishers), S. 1–19, eingeführt. Die Vorbereitung des *Agrobacterium*-Inokulums mit dem Konstrukt sowie die Inokulation des Pflanzenmaterials, die Regeneration von Sprossen und die Bewurzelung von Sprossen sind in Edwards et al., "Biochemical and molecular characterization of a novel starch synthase from potatoes," *Plant J.* 8, 283–294 (1995), beschrieben.

[0132] Es liegen in mehreren unterschiedlichen Genen mehrere Bindungsregionen vor. Obwohl bevorzugt ist, daß das Protein innerhalb des Stärkekorns eingekapselt wird (Stärkekornverkapselung), erstreckt sich die Erfindung mit dem Begriff "Einkapselung" auch auf eine Einkapselung in Stärke, die nicht in Stärkekornform vorliegt. Für diesen Zweck eignen sich die folgenden Arten von Genen.

Verwendung von Stärkebindungsregionen der Glycogensynthase:

[01333] Die Glykogensynthase aus *E. coli* ist kein großes Protein: das Strukturgen ist 1431 Basenpaare lang, wobei ein Protein mit 477 Aminosäuren mit einem geschätzten Molekulargewicht von 49.000 spezifiziert wird. Es ist bekannt, daß bei bakteriellen Genen, die in pflanzliche Genome insertiert werden, Probleme mit dem "Codon-Usage" auftreten können, dies ist jedoch bei *E. coli*-Genen nicht so ein großes Problem wie bei Genen von anderen Bakterien wie *Bacillus*. Die Glykogensynthase aus *E. coli* weist ein "Codon-Usage"-Profil auf, das mit Maisgenen viel gemeinsam hat, es wird jedoch bevorzugt, die Sequenz am Translationsstart mit bekannten Verfahren so zu verändern, daß sie mit einer pflanzlichen Konsensussequenz besser kompatibel ist:

g1qA G A T A A T G C A G [SEQ ID NO:31]

cons A A C A A T G G C T [SEO ID NO:32]

Verwendung der Stärkebindungsregionen der löslichen Stärkesynthese:

[0134] cDNA-Klone von pflanzlichen löslichen Stärkesynthasen sind im Abschnitt oben "Allgemeiner Stand der Technik" beschrieben und können in der vorliegenden Erfindung eingesetzt werden. Die Gene für beliebige solche SSTS-Proteine können in erfindungsgemäßen Konstrukten eingesetzt werden.

Verwendung der Stärkebindungsregionen des Verzweigungsenzyms:

[0135] cDNA-Klone von pflanzlichen, bakteriellen und tierischen Verzweigungsenzymen sind im Abschnitt oben "Allgemeiner Stand der Technik" beschrieben und können in der vorliegenden Erfindung eingesetzt werden. Das Verzweigungsenzym [1,4-D-Glucan: 1,4-D-Glucan-6D(1,4-D-Glucano)transferase (E.C.2.4.1.18)] wandelt Amylose in Amylopektin um, (ein Abschnitt einer 1,4-D-Glucankette wird auf eine primäre Hydroxy-

gruppe in einer ähnlichen Glucankette übertragen), manchmal Q-Enzym genannt.

[0136] Die Sequenz des Verzweigungsenzyms I des Maises wurde von Baba et al. (1991) BBRC, 181: 87–94 untersucht. Das Stärkeverzweigungsenzym II aus dem Maisendosperm wurde von Fisher et al. (1993) Plant Physiol, 102: 1045–1046 untersucht. Das BE-Genkonstrukt kann das Vorhandensein eines Amyloplasten-Transitpeptids erforderlich machen, um seine korrekte Lokalisierung im Amyloplasten zu gewährleisten. Die Gene für ein beliebiges solches Verzweigungsenzym des GBSTS-Proteins können in erfindungsgemäßen Konstrukten eingesetzt werden.

Verwendung von Stärkebindungsdomänen der stärkekorngelassenen Stärkesynthase:

[0137] Die Verwendung von cDNA-Klonen von pflanzlichen stärkekorngelassenen Stärkesynthasen wird bei Shure et al. (1983) Cell 35: 225–233, und Visser et al. (1989) Plant Sci. 64(2): 185–192 beschrieben. Visser et al. haben auch die Hemmung der Expression des Gens für die stärkekorngelassene Stärkesynthase in der Kartoffel mittels antisense-Konstrukten beschrieben (1991) Mol. Gen. Genetic 225(2): 289–296; (1994) The Plant Cell 6: 43–52.) Shimada et al. zeigen Antisense bei Reis (1993) Theor. Appl. Genet. 86: 665–672. Van der Leij et al. zeigen die Wiederherstellung der Amylosesynthese in "low-amylose"-Kartoffel nach Transformation mit dem Wildtyp-waxy-Gen der Kartoffel (1991) Theor. Appl. Genet. 82: 289–295.

[0138] Die Aminosäuresequenzen und Nukleotidsequenzen der Stärkekorngelassenen Stärkesynthasen aus z. B. Mais, Reis, Weizen, Kartoffel, Cassava, Erbsen oder Gerste sind gut bekannt. Die Gene von beliebigem solchen GBSTS-Protein können in erfindungsgemäßen Konstrukten eingesetzt werden.

Konstruktion von Pflanzentransformationsvektoren:

[0139] Pflanzentransformationsvektoren für das erfindungsgemäße Verfahren können nach Standard-Techniken konstruiert werden.

Verwendung von Transitpeptidsequenzen

[0140] Manche Genkonstrukte machen das Vorhandensein eines Amyloplasten-Transitpeptids erforderlich, um eine korrekte Lokalisierung im Amyloplasten zu gewährleisten. Es wird angenommen, daß Chloroplasten-Transitpeptide ähnliche Sequenzen aufweisen (Heijne et al. beschreiben eine Datenbank von Chloroplasten-Transitpeptiden in (1991) Plant Mol. Biol. Reporter, 9(2): 104–126). Andere Transitpeptide, die sich für die vorliegende Erfindung eignen, sind diejenigen der ADPG-Pyrophosphorylase (1991) Plant Mol. Biol. Reporter, 9(2): 104–126), der kleinen RUBISCO-Untereinheit der Acetolactatsynthase, der Glyceraldehyd-3P-dehydrogenase und der Nitritreduktase.

[0141] Die Consensus-Sequenz des Transitpeptids der kleinen RUBISCO-Untereinheit von vielen Genotypen weist die folgende Sequenz auf:

MASSMLSSAAVATRTNPAQASM VAPFTGLKSAAFPVSEKQNLDI

TSIASNGGRVQC [SEQ ID: NO 33]

[0142] Die kleine RUBISCO-Untereinheit weist die folgende Sequenz auf:

MAPTVMMASSATATRTNPAQAS AVAPFQGLKSTASLPVARRSSR SLGNVASNGGRIRC
[SEQ ID NO:34]

[0143] Das Transitpeptid der Blatt-Glyceraldehyd-3P-dehydrogenase des Maises weist die folgende Sequenz auf:

MAQILAPSTQWQMRITKTSPCA TPITSKMWSSLVMKQTKKVAHS
AKFRVMAVNSENGT [SEQ ID NO:35]

[0144] Die Transitpeptidsequenz der endospermgebundenen Stärkesynthase des Maises weist die folgende Sequenz auf:

MAALATSQLVATRAGHGVPDASTFRRGAAQGLRGARASAAADTLMSMRTSARAAPRHQ
QQARRGGRFPFSLVVC [SEQ ID NO:36]

[0145] Die Transitpeptidsequenz der löslichen Stärkesynthase des Maisendosperms weist die folgende Sequenz auf:

MATPSAVGAACLLLRXAWPAAVGDRARPRRLQRVLRRL [SEQ ID NO:37]

Gentechnisches Einführen von neuen Aminosäuren oder Peptiden in stärkeverkapselnde Proteine:

[0146] Die in der vorliegenden Erfindung eingesetzten stärkebildenden Proteine können nach Methoden, mit denen der Fachmann vertraut ist, dahingehend modifiziert werden, daß sie neue Aminosäurekombinationen beinhalten. So können zum Beispiel Sequenzen von stärkebildenden Proteinen dahingehend modifiziert werden, daß sie höhere Lysin-, Methionin- oder Tryptophanniveaus als normal exprimieren. Solche Niveaus können nützlicherweise über die normalen Niveaus hinweg angehoben werden, und solche Proteine stellen eine nährstoffmäßige Verbesserung in Kulturen wie Getreiden bereit.

[0147] Zusätzlich zu einer Veränderung der Aminosäurezusammensetzung ist es möglich, die stärkebindenden Proteine dahingehend genetisch zu verändern, daß wertvolle Peptide in das stärkebindende Protein eingebaut werden können. Ein Anhängen des interessierenden Polypeptids an das stärkebindende Protein am N-terminalen Ende des Proteins stellt ein bekanntes Mittel zum Hinzufügen von Peptidfragmenten unter Beibehaltung der Stärkebindungsfähigkeit bereit. Weitere Verbesserungen können dadurch erfolgen, daß man spezifische Proteasespaltstellen in die Stelle, an der das interessierende Polypeptid an die Stärkebindungsregion angeheftet wird, einführt. Es ist dem Fachmann gut bekannt, daß Proteasen bevorzugte Spezifitäten für unterschiedliche Aminosäurebindungen aufweisen. Solche Spezifitäten können dazu eingesetzt werden, um ein Mittel für die Abgabe von wertvollen Peptiden an unterschiedliche Regionen des Verdauungstrakts von Mensch und Tier bereitzustellen.

[0148] In einer weiteren Ausführungsform der vorliegenden Erfindung kann das interessierende Polypeptid im Anschluß an die Aufreinigung und Verarbeitung der Stärkekörner freigesetzt werden. Mittels Amylolyse- und/oder Verkleisterungsverfahren ist bekannt, daß Proteine, die an das Stärkekorn gebunden sind, freigesetzt werden können bzw. für die Proteolyse verfügbar gemacht werden können. Dies ermöglicht die Gewinnung von wirtschaftlich bedeutenden Mengen an Proteinen und Peptiden aus der Stärkekornmatrix.

[0149] In einer weiteren Ausführungsform der Erfindung ist es möglich, die Stärkekörner auf verschiedene unterschiedliche Art und Weise aufzuarbeiten, um ein Mittel zur Veränderung der Verdaulichkeit der Stärke bereitzustellen. Mit dieser Methodik ist es möglich, die biologische Verfügbarkeit der innerhalb der Stärkekörner festgelegten Proteine, Peptide oder Aminosäuren zu verändern.

SEQUENZBESCHREIBUNG

(1) ALLGEMEINE INFORMATIONEN:

(i) ANMELDER: Keeling, Peter
Guan, Hanping

(ii) ANMELDETITEL: STÄRKEVERKAPSELUNG

(iii) ANZAHL DER SEQUENZEN: 37

(iv) KORRESPONDENZADRESSE:

(A) ADRESSAT: Greenlee, Winner und Sullivan,
P.C.

(B) STRASSE: 5370 Manhattan Circle

(C) STADT: Boulder

(D) STAAT: CO

(E) LAND: USA

(F) POSTLEITZAHL: 80303

(v) COMPUTERLESBARE FORM:

(A) DATENTRÄGER: Diskette

(B) COMPUTER: IBM PC compatible

(C) BETRIEBSSYSTEM:

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

(vi) VORLIEGENDE PATENTANMELDUNG:

(A) ANMELDENUMMER: US

(B) ANMELDETAG: 30. Sept. 1997

(C) KLASSIFIKATION:

(vii) FRÜHERE PATENTANMELDUNG

(A) ANMELDENUMMER: US 60/026,855

(B) ANMELDETAG: 30. Sept. 1996

(viii) ANGABEN ZUM PATENANWALT

(A) NAME: Winner, Ellen P

(B) REGISTRIERUNGSNUMMER: 28,547

(C) REFERENZ/LISTENNUMMER: 89-97

(ix) ANGABEN ZUR TELEKOMMUNIKATION:

(A) TELEFON: (303) 499-8080

(B) TELEFAX: (303) 499-8089

(2) INFORMATION ZU SEQ ID NO: 1:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 31 Basenpaare

(B) ART: Nukleinsäure

(C) Strangform: Einzel

(D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Sonstige Nukleinsäure

(A) BESCHREIBUNG: / desc = "Oligonukleotide"

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

GACTAGTCAT ATGGTGAGCA AGGCCGAGCA G

31

(2) INFORMATION ZU SEQ ID NO: 2:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 36 Basenpaare

- (B) ART: Nukleinsäure
- (C) Strangform: Einzel
- (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Sonstige Nukleinsäure

(A) BESCHREIBUNG: / desc = "Oligonukleotide"

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

CTAGATCITC ATATGCTTGT ACAGCTCGTC CATGCC

36

(2) INFORMATION ZU SEQ ID NO: 3:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 39 Basenpaare
- (B) ART: Nukleinsäure
- (C) Strangform: Einzel
- (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Sonstige Nukleinsäure

(A) BESCHREIBUNG: / desc = "Oligonukleotide"

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

CTAGATCTTG GCCATGGCCT TGTACAGCTC GTCCATGCC

39

(2) INFORMATION ZU SEQ ID NO: 4:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 4800 Basenpaare
- (B) ART: Nukleinsäure
- (C) Strangform: Doppel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: DNA (genomisch)

(iii) HYPOTHETISCH: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Zea mays

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS

(B) LAGE: join(1449..1553, 1685..1765,
1860..1958, 2055

..2144, 2226..2289, 2413..2513, 2651..2760, 2858
..3101, 3212..3394, 3490..3681, 3793..3879, 3977
..4105, 4227..4343)

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:4

| | |
|---|-----|
| CAGCGACCTA TTACACAGCC CGCTCGGGCC CGCGACGTGG GGACACATCT TCTTCCCCCT | 60 |
| TTTGGTGAAG CTCTGCTCGC AGCTGTCCGG CTCCTTGGAC GTTCGTGTGG CAGATTGATC | 120 |
| TGTTGTCTCG TCTGCTGTGC TTCTGGGTA GCTTGTGTAG TGGAGCTGAC ATGCTCTGAG | 180 |
| CAGGCTTAAA ATTTGCTCGT AGACGAGGAG TACCAGCACA GCACGTTGCG GATTTCTCTG | 240 |

| | |
|---|------|
| CCTGTGAAGT GCAACGTCTA GGATTGTCA CACGCTTGGT CGGTCGGCT CGCGTCGGCT | 300 |
| CGATGCCGTC GTGAGCAGAG CAGCAACAGC TGGCGCGCCC AACGTTGGCT TCGGTGTCTT | 360 |
| CGTCGTACGT ACGCGCGCGC CGGCGACAGC CAGCAGAGAG CGGAGAGCGA GCGGTGCACG | 420 |
| GGGAGGTGGT GTGGAACTGG AGCGCGCGCG CCGCGCGCCC GCGCGCGGTG GGCAACCCAA | 480 |
| AAGTACCCAC GACAAGCGAA GCGCGCAAG CGATCCAGC TCGGAACGC AACAGCATGC | 540 |
| GTCCGCTCGG AGAGCCAGCC ACAAGCAGCC GAGAACCGAA CCGGTGGGCG ACGGTGCATG | 600 |
| GGACGGACGC GGGCGACGCT TCCAAACGGG CCACGTACGC CGGCGTGTGC GTCCGTGCAG | 660 |
| ACGACAAGCC AAGGCGAGGC AGCCCCGAT CGGGAAGCC TTTTGGGCG GAGCGCTGGC | 720 |
| GTCCGCTCA GTGCTGTG CGCAGTCCG GGGGAACCG GTATCGTGG GGGCGCGGCG | 780 |
| GGAGGAGAGC GTGCGAGGG CCGAGAGCAG CCGCGGCGG GTCAACGCA CGCGCCCCAC | 840 |
| GTACTGCCCT CCCCCTCGC GCGCGTAGA AATACGAGG CCTGGACCG GGGGGGCGCC | 900 |
| CCTCACATCC ATCCATCGAC CGATCGATCG CCACAGCCAA CACCACCGCG CGAGGCGAGC | 960 |
| CGACAGCCGC CAGGAGGAAG GAATAACTC ACTGCCAGCC AGTGAAGGG GAGAGTGTA | 1020 |
| CTGCTCGGTC GACCACTGCG CGCACCCCG GCGAGGCTG CTGATCTCGT CGACGACCAG | 1080 |
| CTTCTGTTCC GTTCGGATCC GATCCGATCC TGTCTTCAG TTTCTCCAG ATCTGGCGC | 1140 |
| GTATCTCGCT GTTTGATGAT CCAGTTCTT CGAACCTAAA TCTGTCCTG CACACGTCTT | 1200 |
| TTCTCTCTCT CCTACGCACT GGATTAATCG GCATGGCGGC TCTGCCCACG TCGCAGCTCG | 1260 |
| TCCCAACCGC CGCCGGCGTG GCGTCCCGG ACGCTCCAC GTTCCGCGC GCGCGCGCGC | 1320 |
| AGGGCCTGAG GGGGCCCCG GCGTCCGCG CCGCGACAC GCTCAGCATG CGGACCGCG | 1380 |
| CSCGCGCGGC GCGCAGGCAC CAGCAGCAGG CGCGCGCGG GGGCAGGTT CCGTCCGCTG | 1440 |
| TCGTGTGC GCC ACC GGC GGC ATG AAC GTC GTC TTC GTC GGC GCC GAG ATG | 1490 |
| Ala Ser Ala Gly Met Asn Val Val Phe Val Gly Ala Glu Met | |
| 1 5 10 | |
| GCG CGG TGG AGC AAG ACC GGC GGC CTC GGC GAC GTC CTC GGC GGC CTG | 1538 |

Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val Leu Gly Gly Leu
15 20 25 30

CCG CCG GCC ATG GCC GTAAGCGCGC GCACCGAGAC ATGCATCCGT TGGATCGCGT 1593
Pro Pro Ala Met Ala
35

CTTCTTCGTG CTCTTSCCGC GTGCATGATG CATGTGTTTC CTCCTGGCCT GTGTTCTGT 1653

ATGTGACGTG TTGTTCGGG CATGCATSCA G CCG AAC GGG CAC CCT GTC ATG 1705
Ala Asn Gly His Arg Val Met
40

GTC GTC TCT CCC CGC TAC GAC CAG TAC AAG GAC GCC TCG GAC ACC AGC 1753
Val Val Ser Pro Arg Tyr Asp Gln Tyr Lys Asp Ala Trp Asp Thr Ser
45 50 55

GTC GTG TCC GAG GTACGGGCAC CGAGACCAGA TTCAGATCAC AGTCACACAC 1805
Val Val Ser Glu
60

ACCGTCATAT GAACCTTTCT CTGCTCTGAT GCCTGCNACT GCAAATGCAT GCAG ATC 1862
Ile

AAG ATG GGA GAC GGG TAC GAG ACG GTC AAG TTC TTC CAC TGC TAC AAG 1910
Lys Met Gly Asp Gly Tyr Glu Thr Val Arg Phe Phe His Cys Tyr Lys
65 70 75

CGC GGA GTG GAC CGC GTG TTC GTT GAC CAC CCA CTG TTC CTG GAG AGG 1958
Arg Gly Val Asp Arg Val Phe Val Asp His Pro Leu Phe Leu Glu Arg
80 85 90 95

GTGAGACGAG ATCTGATCAC TCGATACGCA ATTACACCCC CATTGTAAGC AGTTACAGTG 2018

AGCTTTT TTTT CCCCCCGGCC TGGTCGCTGG TTTCAG GTT TGG GGA AAG ACC GAG 2072
Val Trp Gly Lys Thr Glu
100

GAG AAG ATC TAC GGG CCT GTC GCT GGA ACG GAC TAC AGG GAC AAC CAG 2120
Glu Lys Ile Tyr Gly Pro Val Ala Gly Thr Asp Tyr Arg Asp Asn Gln
105 110 115

CTG CGG TTC AGC CTG CTA TGC CAG GTCAGGATCG CTTGCTACTA CAACTTCATA 2174
Leu Arg Phe Ser Leu Leu Cys Gln

120

125

TCATCTGTAT GCAGCAGTAT ACACTGATGA GAAATGCATG CTGTCTGCA G GCA GCA 2231
Ala Ala

CTT GAA GCT CCA AGG ATC CTG AGC CTC AAC AAC AAC CCA TAC TTC TCC 2279
Leu Glu Ala Pro Arg Ile Leu Ser Leu Asn Asn Asn Pro Tyr Phe Ser
130 135 140

GGA CCA TAC G GTAAGAGTTG CAGTCTTCGT ATATATATCT GTTGAGCTCG 2329
Gly Pro Tyr
145

AGAATCTTCA CAGGAACCGG CCCATCAGAC GGAATGTCCT TTTACACTGA CTACTGCTGC 2389
TCCTCTTCGT CCATCCATAC AAG GG GAG GAC GTC GTG TTC GTC TGC AAC 2438
Gly Glu Asp Val Val Phe Val Cys Asn
150 155

GAC TGG CAC ACC GGC CCT CTC TCG TGC TAC CTC AAG AGC AAC TAC CAG 2486
Asp Trp His Thr Gly Pro Leu Ser Cys Tyr Leu Lys Ser Asn Tyr Gln
160 165 170

TCC CAC GGC ATC TAC AGG GAC GCA AAG GTTGCCTTCT CTGAAC TGAA 2533
Ser His Gly Ile Tyr Arg Asp Ala Lys
175 180

CAACGCCGTT TTCCTTCTCC ATGCTCGTAT ATACCTCGTC TGGTAGTGGT GGTGCTTCTC 2593

TGAGAACTA ACTGAACTG ACTGCATGTC TGTCTGACCA TCTTCACTGA CTACCAG 2650

ACC GCT TTC TGC ATC CAC AAC ATC TCC TAC CAG GGC CCG TTC GCC TTC 2698
Thr Ala Phe Cys Ile His Asn Ile Ser Tyr Gln Gly Arg Phe Ala Phe
185 190 195

TCC GAC TAC CCG GAG CTG AAC CTC CCG GAG AGA TTC AAG TCG TCC TTC 2746
Ser Asp Tyr Pro Glu Leu Asn Leu Pro Glu Arg Phe Lys Ser Ser Phe
200 205 210

GAT TTC ATC GAC GG GTCTGTTTC CTGCGTGCAT GTGAACATTC ATGAATGGTA 2800
Asp Phe Ile Asp Gly
215

ACCCACAACCT GTTCGCGTCC TGCTGGTTCA TTATCTGACC TGATTGCATT ATTGCAG C 2858

| | |
|---|------|
| TAC GAG AAG CCC GTC GAA GGC CGG AAG ATC AAC TGG ATG AAG GCC GGG Tyr Glu Lys Pro Val Glu Gly Arg Lys Ile Asn Trp Met Lys Ala Gly 220 225 230 | 2906 |
| ATC CTC GAG GCC GAC AGG GTC CTC ACC GTC AGC CCC TAC TAC GCC GAG Ile Leu Glu Ala Asp Arg Val Leu Thr Val Ser Pro Tyr Tyr Ala Glu 235 240 245 | 2954 |
| GAG CTC ATC TCC GCC ATC GCC AGG GGC TGC GAG CTC GAC AAC ATC ATG Glu Leu Ile Ser Gly Ile Ala Arg Gly Cys Glu Leu Asp Asn Ile Met 250 255 260 265 | 3002 |
| CGC CTC ACC GGC ATC ACC GGC ATC GTC AAC GGC ATG GAC GTC AGC GAG Arg Leu Thr Gly Ile Thr Gly Ile Val Asn Gly Met Asp Val Ser Glu 270 275 280 | 3050 |
| TGG GAC CCC AGC AGG GAC AAG TAC ATC GCC GTG AAG TAC GAC GTG TCG Trp Asp Pro Ser Arg Asp Lys Tyr Ile Ala Val Lys Tyr Asp Val Ser 285 290 295 | 3098 |
| ACG GTGAGCTGGC TAGCTCTGAT TCTGCTGGCT GGTCTCTGGT CTCATCATGC Thr | 3151 |
| TGCTTCGGTA CTGACGGGGC AAGTGTACGT ACCTGGGTGC GACGGTGCTG TCGGGTTCAG | 3211 |
| GCC GTG GAG CCC AAG GCG CTC AAC AAG GAG GCG CTC CAG GCG GAG GTC Ala Val Glu Ala Lys Ala Leu Asn Lys Glu Ala Leu Gln Ala Glu Val 300 305 310 | 3259 |
| GGG CTC CCG GTG GAC CCG AAC ATC CCG CTC GTG GCG TTC ATC GCG AGG Gly Leu Pro Val Asp Arg Asn Ile Pro Leu Val Ala Phe Ile Gly Arg 315 320 325 330 | 3307 |
| CTG GAA GAG CAG AAG GGC CCC GAC GTC ATG GCG GCC GCC ATC CCG CAG Leu Glu Glu Gln Lys Gly Pro Asp Val Met Ala Ala Ala Ile Pro Gln 335 340 345 | 3355 |
| CTC ATG GAG ATG GTG GAG GAC GTG CAG ATC GTT CTC CTG GTACGTGTGC Leu Met Glu Met Val Glu Asp Val Gln Ile Val Leu Leu 350 355 | 3404 |
| GCCGGGCGCC ACCCGGCTAC TACATGGGTG TATCGTGGT TCTACTGCAA CATCGGTGTC | 3464 |
| AGCAACGGCA TGGRTAATGC TGCAG GGC ACG GGC AAG AAG AAG TTC GAG GCG | 3516 |

Gly Thr Gly Lys Lys Lys Phe Glu Arg
 360 365

ATG CTC ATG AGC GCC GAG GAG AAG TTC CCA GGC AAG GTG CGC GCC GTG 3564
 Met Leu Met Ser Ala Glu Glu Lys Phe Pro Gly Lys Val Arg Ala Val
 370 375 380

GTC AAG TTC AAC GCG GCG CTG GCG CAC CAC ATC ATG GCC GCG GCC GAC 3612
 Val Lys Phe Asn Ala Ala Leu Ala His His Ile Met Ala Gly Ala Asp
 385 390 395 400

GTG CTC GCC GTC ACC AGC CGC TTC GAG CCC TGC GGC CTC ATC CAG CTG 3660
 Val Leu Ala Val Thr Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu
 405 410 415

CAG GCG ATG CGA TAC GGA ACG CTACGAGAGA AAAAAAAAAAT CCTGAATCCT 3711
 Gln Gly Met Arg Tyr Gly Thr
 420

GACGACAGGG ACAGACACAG ATTATGAATG CTCATCGAT TTGAATTGAT TGATCGATGT 3771

CTCGCGCTGC GACTCTTGCA G CCC TGC GCC TGC GCG TCC ACC GGT GCA CTC 3822
 Pro Cys Ala Cys Ala Ser Thr Gly Gly Leu
 425 430

GTC GAC ACC ATC ATC GAA GGC AAG ACC GGC TTC CAC ATG GGC CGC CTC 3870
 Val Asp Thr Ile Ile Glu Gly Lys Thr Gly Phe His Met Gly Arg Leu
 435 440 445

AGC GTC GAC GTAAGCCTAG CTCTGCCATG TTCCTTCTTC TTTCTTTCTG 3919
 Ser Val Asp
 450

TATGTATGTA TGAATCAGCA CCGCGGTCT TCTTTCTGTC TCGTCTCTTC TTCCCAG 3976

TGT AAC GTC GTG GAG CCG GCG GAC GTC AAG AAG GTG GCC ACC ACA TTG 4024
 Cys Asn Val Val Glu Pro Ala Asp Val Lys Lys Val Ala Thr Thr Leu
 455 460 465

CAG CGC GCC ATC AAG GTC GTC GCC ACG CCG GCG TAC GAG GAG ATG GTG 4072
 Gln Arg Ala Ile Lys Val Val Gly Thr Pro Ala Tyr Glu Glu Met Val
 470 475 480

AGG AAC TGC ATG ATC CAG GAT CTC TCC TGG AAG GTACCTACGC CCGCCCCGCC 4125
 Arg Asn Cys Met Ile Gln Asp Leu Ser Trp Lys

485

490

495

CGGCCCCCGC AGAGCAGAGC GCCAAGATCG ACGATCGAC CGACCACACG TAGCGCGCTC 4185

GCTCCTGTCTG CTGACCGTGG TTTAATTGCG GAAATGCGCA G GGC CCT GCC AAG 4238
Gly Pro Ala Lys

AAC TGG GAG AAC GTG CTG CTC AGC CTC GGG CTC GCC GGC GGC GAG CCA 4286
Asn Trp Glu Asn Val Leu Leu Ser Leu Gly Val Ala Gly Gly Glu Pro
500 505 510 515

GGG GTC GAA GGC GAG GAG ATC GCG CCG CTC GCC AAG GAG AAC GTG GCC 4334
Gly Val Glu Gly Glu Glu Ile Ala Pro Leu Ala Lys Glu Asn Val Ala
520 525 530

GCG CCC TGA AGAGTTCGGC CTGCAGGGCC CCTGATCTCG CGCGTGGTGC 4383
Ala Pro *

AAAGATGTTG GGACATCTTC TTATATATGC TGTTCGTTT ATGTGATATG GACAACTATG 4443

TGTAGCTGCT TGCTTGCTCT AGTGTAACTGT AGTGTAGTGG TGGCCAGTGG CACAACCTAA 4503

TARGCGCATC AACTAATTGC TTGCGTGTGT AGTTAAGTAC CGATCGGTAA TTTTATATTG 4563

CCAGTAAATA AATGCACCTG TAGTGGTGGG GTAAATAATC CCTGCTGTTG GGTGTTCTTA 4623

TGGCTCCTCG TTAGATATT ATATACAGTA CATTTTTCTC TCTCTGATC CTACGTTTGT 4683

GAAATTTCTA TATCATTACT GTAAATTTTC TCGGTTCCAA AAGAGACCAT AGCCTATCTT 4743

TGGCCCTGTT TGTTCGGCT TCTGGCAGCT TCTGGCCACC AAAAGCTGCT GCGGACT 4800

(2) INFORMATION ZU SEQ ID NO: 5:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 534 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

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

Leu Thr Val Ser Pro Tyr Tyr Ala Glu Glu Leu Ile Ser Gly Ile Ala
 245 250 255

Arg Gly Cys Glu Leu Asp Asn Ile Met Arg Leu Thr Gly Ile Thr Gly
 260 265 270

Ile Val Asn Gly Met Asp Val Ser Glu Trp Asp Pro Ser Arg Asp Lys
 275 280 285

Tyr Ile Ala Val Lys Tyr Asp Val Ser Thr Ala Val Glu Ala Lys Ala
 290 295 300

Leu Asn Lys Glu Ala Leu Gln Ala Glu Val Gly Leu Pro Val Asp Arg
 305 310 315 320

Asn Ile Pro Leu Val Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly
 325 330 335

Pro Asp Val Met Ala Ala Ala Ile Pro Gln Leu Met Glu Met Val Glu
 340 345 350

Asp Val Gln Ile Val Leu Leu Gly Thr Gly Lys Lys Lys Phe Glu Arg
 355 360 365

Met Leu Met Ser Ala Glu Glu Lys Phe Pro Gly Lys Val Arg Ala Val
 370 375 380

Val Lys Phe Asn Ala Ala Leu Ala His His Ile Met Ala Gly Ala Asp
 385 390 395 400

Val Leu Ala Val Thr Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu
 405 410 415

Gln Gly Met Arg Tyr Gly Thr Pro Cys Ala Cys Ala Ser Thr Gly Gly
 420 425 430

Leu Val Asp Thr Ile Ile Glu Gly Lys Thr Gly Phe His Met Gly Arg
 435 440 445

Leu Ser Val Asp Cys Asn Val Val Glu Pro Ala Asp Val Lys Lys Val
 450 455 460

Ala Thr Thr Leu Gln Arg Ala Ile Lys Val Val Gly Thr Pro Ala Tyr
 465 470 475 480

```

Glu Glu Met Val Arg Asn Cys Met Ile Gln Asp Leu Ser Trp Lys Gly
      485              490              495

Pro Ala Lys Asn Trp Glu Asn Val Leu Leu Ser Leu Gly Val Ala Gly
      500              505              510

Gly Glu Pro Gly Val Glu Gly Glu Glu Ile Ala Pro Leu Ala Lys Glu
      515              520              525

Asn Val Ala Ala Pro *
      530

```

(2) INFORMATION ZU SEQ ID NO: 6:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 2542 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppelt
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA oder mRNA

(iii) HYPOTHETISCH: NEIN

(vi) URSPRÜNGLICHE HERKUNFT

- (A) ORGANISMUS: Oryza sativa

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) Lage: 453 .. 2282

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

| | |
|--|-----|
| GAATTCAGTG TGAAGGAATA GATTCTCTTC AAAACAATTT AATCATTCAAT CTGATCTGCT | 50 |
| CAAAGGCTCTG TGCATCTCCG GGTGCAACGG CCAGGATATT TATTGTGCAG TAAAAAATG | 120 |
| TCATATCCCC TAGCCACCCA AGAACTGCT CCTTAAGTCC TTATAAGCAC ATATGGCATT | 180 |
| GTAATATATA TGTTCAGTT TTAGCGACAA TTTTTTAAA AACTTTTGGT CCTTTTATG | 240 |
| AACGTTTTAA GTTCACTCT CTTTTTTTTT CGAATTTTAA ATGTAGCTTC AAATTCTAAT | 300 |
| CCCCAATCCA AATTGTAATA AACTTCAATT CTCCTAATTA ACATCTTAAT TCATTTATTT | 360 |
| GAAAACCAGT TCAATTCTT TTTAGGCTCA CCACACCTTA AACAATTCHA TTCAGTGCAG | 420 |
| AGATCTTCCA CAGCAACAGC TAGACAACCA CC ATG TCG GCT CTC ACC ACG TCC | 473 |
| Met Ser Ala Leu Thr Thr Ser | |
| 535 540 | |
| CAG CTC GCC ACC TCG GCC ACC GGC TTC GGC ATC GCG GAC AGG TCG GCG | 521 |
| Gln Leu Ala Thr Ser Ala Thr Gly Phe Gly Ile Ala Asp Arg Ser Ala | |
| 545 550 555 | |
| CCG TCG TCG CTG CTC CGC CAC GGG TTC CAG GGC CTC AAG CCC CGC AGC | 569 |
| Pro Ser Ser Leu Leu Arg His Gly Phe Gln Gly Leu Lys Pro Arg Ser | |
| 560 565 570 | |
| CCG GCG GCG GCG GAC GCG ACG TCG CTC ACC GTG ACG ACC AGC GCG CGC | 617 |
| Pro Ala Gly Gly Asp Ala Thr Ser Leu Ser Val Thr Thr Ser Ala Arg | |
| 575 580 585 | |
| GCG ACG CCC AAG CAG CAG CGG TCG GTG CAG CGT GGC ACC CGG AGG TTC | 665 |
| Ala Thr Pro Lys Gln Gln Arg Ser Val Gln Arg Gly Ser Arg Arg Phe | |
| 590 595 600 605 | |
| CCC TCC GTC GTC GTG TAC GCC ACC GGC GCG GGC ATG AAC GTC GTG TTC | 713 |
| Pro Ser Val Val Val Tyr Ala Thr Gly Ala Gly Met Asn Val Val Phe | |
| 610 615 620 | |
| GTG GGC GCG GAG ATG GCC CCC TGG AGC AAG ACC GGC GCG CTC GGT GAC | 761 |
| Val Gly Ala Glu Met Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp | |
| 625 630 635 | |
| GTG CTC GGT GGC CTC CCC CCT GCC ATG GCT GCG AAT GGC CAC AGG GTC | 809 |
| Val Leu Gly Gly Leu Pro Pro Ala Met Ala Ala Asn Gly His Arg Val | |
| 640 645 650 | |
| ATG GTG ATC TCT CCT CGG TAC GAC CAG TAC AAG GAC GCT TGG GAT ACC | 857 |
| Met Val Ile Ser Pro Arg Tyr Asp Gln Tyr Lys Asp Ala Trp Asp Thr | |
| 655 660 665 | |
| AGC GTT GTG GCT GAG ATC AAG GTT GCA GAC AGG TAC GAG AGG GTG AGG | 905 |
| Ser Val Val Ala Glu Ile Lys Val Ala Asp Arg Tyr Glu Arg Val Arg | |
| 670 675 680 685 | |
| TTT TTC CAT TGC TAC AAG CGT GGA GTC GAC CGT GTG TTC ATC GAC CAT | 953 |
| Phe Phe His Cys Tyr Lys Arg Gly Val Asp Arg Val Phe Ile Asp His | |
| 690 695 700 | |

| | |
|---|------|
| CCG TCA TTC CTG GAG AAG GTT TGG GGA AAG ACC CGT GAG AAG ATC TAC Pro Ser Phe Leu Glu Lys Val Trp Gly Lys Thr Gly Glu Lys Ile Tyr 705 710 715 | 1001 |
| CGA CCT GAC ACT GGA GTT GAT TAC AAA GAC AAC CAG ATG CGT TTC AGC Gly Pro Asp Thr Gly Val Asp Tyr Lys Asp Asn Gln Met Arg Phe Ser 720 725 730 | 1049 |
| CTT CTT TGC CAG GCA GCA CTC GAG GCT CCT AGG ATC CTA AAC CTC AAC Leu Leu Cys Gln Ala Ala Leu Glu Ala Pro Arg Ile Leu Asn Leu Asn 735 740 745 | 1097 |
| AAC AAC CCA TAC TTC AAA GGA ACT TAT GGT GAG GAT GTT GTG TTC GTC Asn Asn Pro Tyr Phe Lys Gly Thr Tyr Gly Glu Asp Val Val Phe Val 750 755 760 765 | 1145 |
| TGC AAC GAC TGG CAC ACT GGC CCA CTG GCG AGC TAC CTG AAG AAC AAC Cys Asn Asp Trp His Thr Gly Pro Leu Ala Ser Tyr Leu Lys Asn Asn 770 775 780 | 1193 |
| TAC CAG CCC AAT GGC ATC TAC AGG AAT GCA AAG GTT GCT TTC TGC ATC Tyr Gln Pro Asn Gly Ile Tyr Arg Asn Ala Lys Val Ala Phe Cys Ile 785 790 795 | 1241 |
| CAC AAC ATC TCC TAC CAG GGC CGT TTC GCT TTC GAG GAT TAC CCT GAG His Asn Ile Ser Tyr Gln Gly Arg Phe Ala Phe Glu Asp Tyr Pro Glu 800 805 810 | 1289 |
| CTG AAC CTC TCC GAG AGG TTC AGG TCA TCC TTC GAT TTC ATC GAC GGC Leu Asn Leu Ser Glu Arg Phe Arg Ser Ser Phe Asp Phe Ile Asp Gly 815 820 825 | 1337 |
| TAT GAC ACG CCG GTG GAG GGC AGG AAG ATC AAC TGG ATG AAG GCC GGA Tyr Asp Thr Pro Val Glu Gly Arg Lys Ile Asn Trp Met Lys Ala Gly 830 835 840 845 | 1385 |
| ATC CTG GAA GCC GAC AGG GTG CTC ACC GTG AGC CCG TAC TAC GCC GAG Ile Leu Glu Ala Asp Arg Val Leu Thr Val Ser Pro Tyr Tyr Ala Glu 850 855 860 | 1433 |
| GAG CTC ATC TCC GGC ATC GCC AGG GGA TGC GAG CTC GAC AAC ATC ATG Glu Leu Ile Ser Gly Ile Ala Arg Gly Cys Glu Leu Asp Asn Ile Met 865 870 875 | 1481 |
| CGG CTC ACC GGC ATC ACC GGC ATC GTC AAC GGC ATG GAC GTC AGC GAG | 1529 |

| | |
|---|------|
| Arg Leu Thr Gly Ile Thr Gly Ile Val Asn Gly Met Asp Val Ser Glu 880 885 890 | |
| TGG GAT CCT AGC AAG GAC AAG TAC ATC ACC GGC AAG TAC GAC GCA ACC Trp Asp Pro Ser Lys Asp Lys Tyr Ile Thr Ala Lys Tyr Asp Ala Thr 895 900 905 | 1577 |
| ACG GCA ATC GAG GCG AAG GCG CTC AAC AAG GAG GCG TTG CAG GCG GAG Thr Ala Ile Glu Ala Lys Ala Leu Asn Lys Glu Ala Leu Gln Ala Glu 910 915 920 925 | 1625 |
| GGG GGT CTT CCG CTC GAC AGG AAA ATC CCA CTG ATC GCG TTC ATC GGC Ala Gly Leu Pro Val Asp Arg Lys Ile Pro Leu Ile Ala Phe Ile Gly 930 935 940 | 1673 |
| ACG CTG GAG GAA CAG AAG GGC CCT GAC GTC ATG GCG GCG GCG ATC CCG Arg Leu Glu Glu Gln Lys Gly Pro Asp Val Met Ala Ala Ala Ile Pro 945 950 955 | 1721 |
| GAG CTC ATG CAG CAG GAG GTC CAG ATC GTT CTT CTG GGT ACT GGA AAG Glu Leu Met Gln Glu Asp Val Gln Ile Val Leu Leu Gly Thr Gly Lys 960 965 970 | 1769 |
| AAG AAG TTC GAG AAG CTG CTC AAG AGC ATC GAG GAG AAG TAT CCG GCG Lys Lys Phe Glu Lys Leu Leu Lys Ser Met Glu Glu Lys Tyr Pro Gly 975 980 985 | 1817 |
| AAG GTC AGG GCG CTG GTC AAG TTC AAC GCG CCG CTT GCT CTT CTC ATC Lys Val Arg Ala Val Val Lys Phe Asn Ala Pro Leu Ala His Leu Ile 990 995 1000 1005 | 1865 |
| ATG GCG GGA GCG GAC GTC CTC GCG GTC CCG AGC GCG TTC GAG CCG TGT Met Ala Gly Ala Asp Val Leu Ala Val Pro Ser Arg Phe Glu Pro Cys 1010 1015 1020 | 1913 |
| GGA CTC ATC CAG CTG CAG GGG ATG ACA TAC GGA ACG CCG TGT GCT TGC Gly Leu Ile Gln Leu Gln Gly Met Arg Tyr Gly Thr Pro Cys Ala Cys 1025 1030 1035 | 1961 |
| GCG TCC ACC GGT GCG CTC GTG GAC AGC GTC ATC GAA GGC AAG ACT GGT Ala Ser Thr Gly Gly Leu Val Asp Thr Val Ile Glu Gly Lys Thr Gly 1040 1045 1050 | 2009 |
| TTC CAC ATG GCG GGT CTC AGC GTC GAC TCC AAG CTG GTG GAG CCA AGC Phe His Met Gly Arg Leu Ser Val Asp Cys Lys Val Val Glu Pro Ser | 2057 |

| 1055 | 1060 | 1065 | |
|--|------|------|------|
| GAC GTG AAG AAG GTG GCG GCC ACC CTG AAG CGC GCC ATC AAG GTC GTC | | | 2105 |
| Asp Val Lys Lys Val Ala Ala Thr Leu Lys Arg Ala Ile Lys Val Val | | | |
| 1070 | 1075 | 1080 | 1085 |
| GCG ACG CCG GCG TAC GAG GAG ATG GTC AGG AAC TGC ATG AAC CAG GAC | | | 2153 |
| Gly Thr Pro Ala Tyr Glu Glu Met Val Arg Asn Cys Met Asn Gln Asp | | | |
| 1090 | 1095 | 1100 | |
| CTC TCC TCG AAG GCG CCT GCG AAG AAC TGG GAG AAT GTG CTC CTC GCG | | | 2201 |
| Leu Ser Trp Lys Gly Pro Ala Lys Asn Trp Glu Asn Val Leu Leu Gly | | | |
| 1105 | 1110 | 1115 | |
| CTG GCG GTC GCC GGC AGC CCG CCG GGG ATC GAA GGC GAC GAG ATC CCG | | | 2249 |
| Leu Gly Val Ala Gly Ser Ala Pro Gly Ile Glu Gly Asp Glu Ile Ala | | | |
| 1120 | 1125 | 1130 | |
| CCG CTC GCG AAG GAG AAC GTG GCT CCT CCT TGA AGAGCCTGAG ATCTACATAT | | | 2302 |
| Pro Leu Ala Lys Glu Asn Val Ala Ala Pro * | | | |
| 1135 | 1140 | | |
| GGAGTGATTA ATTAATATAG CAGTATATCG ATCAGAGACG AATGAACCAG TGGTTTGT TT | | | 2362 |
| GTGTAGTGA ATTGTAGCT ATAGCCAATT ATATAGGCTA ATAAGTTTGA TGTGTACTC | | | 2422 |
| TTCTGGGTGT GCTTAAGTAT CTATCGGAC CCTGAATTTA TGTGTGTGGC TTATTGCCAA | | | 2482 |
| TAATATTAAG TAATAAAGGG TTTATTATAT TATTATATAT GTTATATTAT ACTAAAAAAA | | | 2542 |

(2) INFORMATION ZU SEQ ID NO: 7:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 610 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Ala | Leu | Thr | Thr | Ser | Gln | Leu | Ala | Thr | Ser | Ala | Thr | Gly | Phe |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |

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

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

83

Tyr Gly Thr Pro Cys Ala Cys Ala Ser Thr Gly Gly Leu Val Asp Thr
 500 505 510

Val Ile Glu Gly Lys Thr Gly Phe His Met Gly Arg Leu Ser Val Asp
 515 520 525

Cys Lys Val Val Glu Pro Ser Asp Val Lys Lys Val Ala Ala Thr Leu
 530 535 540

Lys Arg Ala Ile Lys Val Val Gly Thr Pro Ala Tyr Glu Glu Met Val
 545 550 555 560

Arg Asn Cys Met Asn Gln Asp Leu Ser Trp Lys Gly Pro Ala Lys Asn
 565 570 575

Trp Glu Asn Val Leu Leu Gly Leu Gly Val Ala Gly Ser Ala Pro Gly
 580 585 590

Ile Glu Gly Asp Glu Ile Ala Pro Leu Ala Lys Glu Asn Val Ala Ala
 595 600 605

Pro *
 610

(2) INFORMATION ZU SEQ ID NO: 8:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 2007 Basenpaare
- (B) ART: Nukleinsäure
- (C) Strangform: Doppel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA oder mRNA

(iii) HYPOTHETISCH: NEIN

(vi) URSPRÜNGLICHE HERKUNFT

- (A) ORGANISMUS: Zea mays

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) Lage: 1 .. 2007

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

84

| | |
|---|-----|
| GCT GAG GCT GAG GCC GGG GGC AAG GAC GCG CCG CCG GAG AGG AGC GGC Ala Glu Ala Glu Ala Gly Gly Lys Asp Ala Pro Pro Glu Arg Ser Gly 615 620 625 | 48 |
| GAC GCC GCC AGG TTG CCC CCG GCT CGG CGC AAT GCG GTC TCC AAA CCG Asp Ala Ala Arg Leu Pro Arg Ala Arg Arg Asn Ala Val Ser Lys Arg 630 635 640 | 96 |
| AGG GAT CCT CTT CAG CCG GTC GGC CGG TAC GGC TCC GCG ACG GGA AAC Arg Asp Pro Leu Gln Pro Val Gly Arg Tyr Gly Ser Ala Thr Gly Asn 645 650 655 | 144 |
| ACG GCC AGG ACC GGC GCC GCG TCC TGC CAG AAC GCC GCA TTG GCG GAC Thr Ala Arg Thr Gly Ala Ala Ser Cys Gln Asn Ala Ala Leu Ala Asp 660 665 670 | 192 |
| GTT GAG ATC GTT GAG ATC AAG TCC ATC GTC GCC GCG CCG CCG ACG AGC Val Glu Ile Val Glu Ile Lys Ser Ile Val Ala Ala Pro Pro Thr Ser 675 680 685 690 | 240 |
| ATA GTG AAG TTC CCA GGG CGC GGG CTA CAG GAT GAT CCT TCC CTC TGG Ile Val Lys Phe Pro Gly Arg Gly Leu Gln Asp Asp Pro Ser Leu Trp 695 700 705 | 288 |
| GAC ATA GCA CCG GAG ACT GTC CTC CCA GCC CCG AAG CCA CTG CAT GAA Asp Ile Ala Pro Glu Thr Val Leu Pro Ala Pro-Lys Pro Leu His Glu 710 715 720 | 336 |
| TCG CCT GCG GTT GAG GGA GAT TCA AAT GGA ATT GCA CCT CCT ACA GTT Ser Pro Ala Val Asp Gly Asp Ser Asn Gly Ile Ala Pro Pro Thr Val 725 730 735 | 384 |
| GAG CCA TTA GTA CAG GAG GCC ACT TCG GAT TTC AAG AAA TAC ATC GGT Glu Pro Leu Val Gln Glu Ala Thr Trp Asp Phe Lys Lys Tyr Ile Gly 740 745 750 | 432 |
| TTT GAC GAG CCT GAC GAA GCG AAG GAT GAT TCC AGG GTT GGT GCA GAT Phe Asp Glu Pro Asp Glu Ala Lys Asp Asp Ser Arg Val Gly Ala Asp 755 760 765 770 | 480 |
| GAT GCT GGT TCT TTT GAA CAT TAT GGG ACA ATG ATT CTG GGC CTT TGT Asp Ala Gly Ser Phe Glu His Tyr Gly Thr Met Ile Leu Gly Leu Cys 775 780 785 | 528 |
| GGG GAG AAT GTT ATG AAC GTG ATC GTG GTG GCT GCT GAA TGT TCT CCA | 576 |

| | |
|---|------|
| Gly Glu Asn Val Met Asn Val Ile Val Val Ala Ala Glu Cys Ser Pro | |
| 790 795 800 | |
| TGG TGC AAA ACA GGT GGT CTT GGA GAT GTT GTG GGA GGT TTA CCC AAG | 624 |
| Trp Cys Lys Thr Gly Gly Leu Gly Asp Val Val Gly Ala Leu Pro Lys | |
| 805 810 815 | |
| GCT TTA GCG AGA AGA GGA CAT CGT GTT ATG GTT GTG GTA CCA AGG TAT | 672 |
| Ala Leu Ala Arg Arg Gly His Arg Val Met Val Val Val Pro Arg Tyr | |
| 820 825 830 | |
| GGG GAC TAT GTG GAA GCC TTT GAT ATG GGA ATC CGG AAA TAC TAC AAA | 720 |
| Gly Asp Tyr Val Glu Ala Phe Asp Met Gly Ile Arg Lys Tyr Tyr Lys | |
| 835 840 845 850 | |
| GCT GCA GGA CAG GAC CTA GAA GTG AAC TAT TTC CAT GCA TTT ATT GAT | 768 |
| Ala Ala Gly Gln Asp Leu Glu Val Asn Tyr Phe His Ala Phe Ile Asp | |
| 855 860 865 | |
| GGA CTC GAC TTT GTG TTC ATT GAT GCC TCT TTC CGG CAC CGT CAA GAT | 816 |
| Gly Val Asp Phe Val Phe Ile Asp Ala Ser Phe Arg His Arg Gln Asp | |
| 870 875 880 | |
| GAC ATA TAT GGG GGA AGT AGG CAG GAA ATC ATG AAG CGC ATG ATT TTG | 864 |
| Asp Ile Tyr Gly Gly Ser Arg Gln Glu Ile Met Lys Arg Met Ile Leu | |
| 885 890 895 | |
| TTT TGC AAG GTT GCT GTT CAG GTT CCT TGG CAC GTT CCA TGC GGT GGT | 912 |
| Phe Cys Lys Val Ala Val Glu Val Pro Trp His Val Pro Cys Gly Gly | |
| 900 905 910 | |
| GTG TGC TAC GGA GAT GGA AAT TTG GTG TTC ATT GCC ATG AAT TGG CAC | 960 |
| Val Cys Tyr Gly Asp Gly Asn Leu Val Phe Ile Ala Met Asn Trp His | |
| 915 920 925 930 | |
| ACT GCA CTC CTG CCT GTT TAT CTG AAG GCA TAT TAC AGA GAC CAT GGC | 1008 |
| Thr Ala Leu Leu Pro Val Tyr Leu Lys Ala Tyr Tyr Arg Asp His Gly | |
| 935 940 945 | |
| TTA ATG CAG TAC ACT CGC TCC GTC CTC GTC ATA CAT AAC ATC GGC CAC | 1056 |
| Leu Met Gln Tyr Thr Arg Ser Val Leu Val Ile His Asn Ile Gly His | |
| 950 955 960 | |
| CAG GGC CGT GGT CCT GTA CAT GAA TTC CCG TAC ATG GAC TTG CTG AAC | 1104 |
| Gln Gly Arg Gly Pro Val His Glu Phe Pro Tyr Met Asp Leu Leu Asn | |

| 965 | 970 | 975 | |
|---|------|------|------|
| ACT AAC CTT CAA CAT TTC GAG CTG TAC GAT CCC GTC GGT GGC GAG CAC | | | 1152 |
| Thr Asn Leu Gln His Phe Glu Leu Tyr Asp Pro Val Gly Gly Glu His | | | |
| 980 | 985 | 990 | |
| GGC AAC ATC TTT GCC GCG TGT GTT CTG AAG ATG GCA GAC CGG GTG GTG | | | 1200 |
| Ala Asn Ile Phe Ala Ala Cys Val Leu Lys Met Ala Asp Arg Val Val | | | |
| 995 | 1000 | 1005 | 1010 |
| ACT GTC AGC CCC GGC TAC CTG TGG GAG CTG AAG ACA GTG GAA GGC GGC | | | 1248 |
| Thr Val Ser Arg Gly Tyr Leu Trp Glu Leu Lys Thr Val Glu Gly Gly | | | |
| | 1015 | 1020 | 1025 |
| TGG GGC CTC CAC GAC ATC ATC CGT TCT AAC GAC TGG AAG ATC AAT GGC | | | 1296 |
| Trp Gly Leu His Asp Ile Ile Arg Ser Asn Asp Trp Lys Ile Asn Gly | | | |
| | 1030 | 1035 | 1040 |
| ATT CGT GAA CGC ATC GAC CAC CAG GAG TGG AAG CCC AAG GTG GAC GTG | | | 1344 |
| Ile Arg Glu Arg Ile Asp His Gln Glu Trp Asn Pro Lys Val Asp Val | | | |
| | 1045 | 1050 | 1055 |
| CAC CTG CGG TCG GAC GGC TAC ACC AAC TAC TCC CTC GAG ACA CTC GAC | | | 1392 |
| His Leu Arg Ser Asp Gly Tyr Thr Asn Tyr Ser Leu Glu Thr Leu Asp | | | |
| | 1060 | 1065 | 1070 |
| GCT GGA AAG CGG CAG TGC AAG GCG CCC CTG CAG CGG GAC GTG GGC CTG | | | 1440 |
| Ala Gly Lys Arg Gln Cys Lys Ala Ala Leu Gln Arg Asp Val Gly Leu | | | |
| | 1075 | 1080 | 1085 |
| GAA GTG CCC GAC GAC GTG CCG CTG CTC GGC TTC ATC GGG CGT CTG GAT | | | 1488 |
| Glu Val Arg Asp Asp Val Pro Leu Leu Gly Phe Ile Gly Arg Leu Asp | | | |
| | 1095 | 1100 | 1105 |
| GGA CAG AAG GGC GTG GAC ATC ATC GGG CAC GCG ATC CCG TCG ATC GCG | | | 1536 |
| Gly Gln Lys Gly Val Asp Ile Ile Gly Asp Ala Met Pro Trp Ile Ala | | | |
| | 1110 | 1115 | 1120 |
| GGG CAG GAC GTG CAG CTG GTG ATG CTG GGC ACC GGC CCA CCT GAC CTG | | | 1584 |
| Gly Gln Asp Val Gln Leu Val Met Leu Gly Thr Gly Pro Pro Asp Leu | | | |
| | 1125 | 1130 | 1135 |
| GAA CGA ATG CTG CAG CAC TTG GAG CCG GAC CAT CCC AAC AAG GTG CGC | | | 1632 |
| Glu Arg Met Leu Gln His Leu Glu Arg Glu His Pro Asn Lys Val Arg | | | |
| | 1140 | 1145 | 1150 |

| | |
|---|------|
| GGG TGG GTC GGG TTC TCG GTC CTA ATG GTG CAT GCG ATC ACG CCG GCG | 1680 |
| Gly Trp Val Gly Phe Ser Val Leu Met Val His Arg Ile Thr Pro Gly | |
| 1155 1160 1165 1170 | |
| GCC AGC GTG CTG GTG ATG CCC TCC CGC TTC GGC GGC GGG CTG AAC CAG | 1728 |
| Ala Ser Val Leu Val Met Pro Ser Arg Phe Ala Gly Gly Leu Asn Gln | |
| 1175 1180 1185 | |
| CTC TAC GCG ATG GCA TAC GGC ACC GTC CCT GTG GTG CAC GCG GTG GGC | 1776 |
| Leu Tyr Ala Met Ala Tyr Gly Thr Val Pro Val Val His Ala Val Gly | |
| 1190 1195 1200 | |
| GGG CTC AGG GAC ACC GTG GCG CCG TTC GAC CCG TTC GGC GAC GCG GCG | 1824 |
| Gly Leu Arg Asp Thr Val Ala Pro Phe Asp Pro Phe Gly Asp Ala Gly | |
| 1205 1210 1215 | |
| CTC GGG TGG ACT TTT GAC CCG GCC GAG GCC AAC AAG CTG ATC GAG GTG | 1872 |
| Leu Gly Trp Thr Phe Asp Arg Ala Glu Ala Asn Lys Leu Ile Glu Val | |
| 1220 1225 1230 | |
| CTC AGC CAC TGC CTC GAC ACG TAC CGA AAC TAC GAG GAG AGC TGG AAG | 1920 |
| Leu Ser His Cys Leu Asp Thr Tyr Arg Asn Tyr Glu Glu Ser Trp Lys | |
| 1235 1240 1245 1250 | |
| AGT CTC CAG GCG CGC GGC ATG TCG CAG AAC CTC AGC TGG GAC CAC GCG | 1968 |
| Ser Leu Gln Ala Arg Gly Met Ser Gln Asn Leu Ser Trp Asp His Ala | |
| 1255 1260 1265 | |
| GCT GAG CTC TAC GAG GAC GTC CTT GTC AAG TAC CAG TGG | 2007 |
| Ala Glu Leu Tyr Glu Asp Val Leu Val Lys Tyr Gln Trp | |
| 1270 1275 | |

(2) INFORMATION ZU SEQ ID NO: 9:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 669 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

Ala Glu Ala Glu Ala Gly Gly Lys Asp Ala Pro Pro Glu Arg Ser Gly

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

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

| 485 | 490 | 495 |
|--|-----|---------|
| Gly Gln Lys Gly Val Asp Ile Ile Gly Asp Ala Met Pro Trp Ile Ala 500 | 505 | 510 |
| Gly Gln Asp Val Gln Leu Val Met Leu Gly Thr Gly Pro Pro Asp Leu 515 | 520 | 525 |
| Glu Arg Met Leu Gln His Leu Glu Arg Glu His Pro Asn Lys Val Arg 530 | 535 | 540 |
| Gly Trp Val Gly Phe Ser Val Leu Met Val His Arg Ile Thr Pro Gly 545 | 550 | 555 560 |
| Ala Ser Val Leu Val Met Pro Ser Arg Phe Ala Gly Gly Leu Asn Gln 565 | 570 | 575 |
| Leu Tyr Ala Met Ala Tyr Gly Thr Val Pro Val Val His Ala Val Gly 580 | 585 | 590 |
| Gly Leu Arg Asp Thr Val Ala Pro Phe Asp Pro Phe Gly Asp Ala Gly 595 | 600 | 605 |
| Leu Gly Trp Thr Phe Asp Arg Ala Glu Ala Asn Lys Leu Ile Glu Val 610 | 615 | 620 |
| Leu Ser His Cys Leu Asp Thr Tyr Arg Asn Tyr Glu Glu Ser Trp Lys 625 | 630 | 635 640 |
| Ser Leu Gln Ala Arg Gly Met Ser Gln Asn Leu Ser Trp Asp His Ala 645 | 650 | 655 |
| Ala Glu Leu Tyr Glu Asp Val Leu Val Lys Tyr Gln Trp 660 | 665 | |

(2) INFORMATION ZU SEQ ID NO: 10:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 2097 Basenpaare
- (B) ART: Nukleinsäure
- (C) Strangform: Doppel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA oder mRNA

(iii) HYPOTHETISCH: NEIN

(vi) URSPRÜNGLICHE HERKUNFT

(A) ORGANISMUS: Zea mays

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) Lage: 1 .. 2097

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

| | |
|---|-----|
| ATG CCG GCG GCA ATC TCT TCC TCG TCG TCG GCT TTT CTC CTC CCC GTC | 48 |
| Met Pro Gly Ala Ile Ser Ser Ser Ser Ala Phe Leu Leu Pro Val | |
| 670 675 680 685 | |
| GCG TCC TCC TCG CCG CGG CGC AGG CGG GCG AGT GTG GGT GCT GCT CTG | 96 |
| Ala Ser Ser Ser Pro Arg Arg Arg Arg Gly Ser Val Gly Ala Ala Leu | |
| 690 695 700 | |
| CGC TCG TAC GGC TAC AGC GGC GCG GAG CTG CCG TTG CAT TGG GCG GGG | 144 |
| Arg Ser Tyr Gly Tyr Ser Gly Ala Glu Leu Arg Leu His Trp Ala Arg | |
| 705 710 715 | |
| CGG GGC CCG CCT CAG GAT GGA GCG GCG TCG GTA CCG GCC GCA GCG GCA | 192 |
| Arg Gly Pro Pro Gln Asp Gly Ala Ala Ser Val Arg Ala Ala Ala Ala | |
| 720 725 730 | |
| CCG GCC GGC GGC GAA AGC GAG GAG GCA GCG AAG AGC TCC TCC TCG TCC | 240 |
| Pro Ala Gly Gly Glu Ser Glu Glu Ala Ala Lys Ser Ser Ser Ser Ser | |
| 735 740 745 | |
| CAG CCG GGC GCT GTT CAG GGC ACC ACC GCC AAG GCT GTC GAT TCT GCT | 288 |
| Gln Ala Gly Ala Val Gln Gly Ser Thr Ala Lys Ala Val Asp Ser Ala | |
| 750 755 760 765 | |
| TCA CCT CCC AAT CCT TTG ACA TCT GCT CCG AAG CAA AGT CAG ACC GCT | 336 |
| Ser Pro Pro Asn Pro Leu Thr Ser Ala Pro Lys Gln Ser Gln Ser Ala | |
| 770 775 780 | |
| GCA ATG CAA AAC GGA ACG AGT GGG GGC AGC AGC GCG AGC ACC GCC GCG | 384 |
| Ala Met Gln Asn Gly Thr Ser Gly Gly Ser Ser Ala Ser Thr Ala Ala | |
| 785 790 795 | |
| CCG GTG TCC GGA CCC AAA GCT GAT CAT CCA TCA GCT CCT CTC ACC AAG | 432 |

| | |
|---|-----|
| Pro Val Ser Gly Pro Lys Ala Asp His Pro Ser Ala Pro Val Thr Lys | |
| 800 805 810 | |
| AGA GAA ATC GAT GCC AGT CCG GTG AAG CCA GAG CCC GCA GGT GAT GAT | 480 |
| Arg Glu Ile Asp Ala Ser Ala Val Lys Pro Glu Pro Ala Gly Asp Asp | |
| 815 820 825 | |
| GCT AGA CCG GTG GAA AGC ATA GGC ATC GCT GAA CCG GTG GAT GCT AAG | 528 |
| Ala Arg Pro Val Glu Ser Ile Gly Ile Ala Glu Pro Val Asp Ala Lys | |
| 830 835 840 845 | |
| GCT GAT GCA GCT CCG GCT ACA GAT CCG CCG CCG AGT GCT CCT TAT GAC | 576 |
| Ala Asp Ala Ala Pro Ala Thr Asp Ala Ala Ala Ser Ala Pro Tyr Asp | |
| 850 855 860 | |
| AGG GAG GAT AAT GAA CCT GGC CCT TTG GCT GGG CCT AAT GTG ATG AAC | 624 |
| Arg Glu Asp Asn Glu Pro Gly Pro Leu Ala Gly Pro Asn Val Met Asn | |
| 865 870 875 | |
| GTC GTC GTG GTG GCT TCT GAA TGT GCT CCT TTC TCC AAG ACA GGT GGC | 672 |
| Val Val Val Val Ala Ser Glu Cys Ala Pro Phe Cys Lys Thr Gly Gly | |
| 880 885 890 | |
| CTT GGA GAT GTC GTG GGT GCT TTG CCT AAG GCT CTG GCG AGG AGA GGA | 720 |
| Leu Gly Asp Val Val Gly Ala Leu Pro Lys Ala Leu Ala Arg Arg Gly | |
| 895 900 905 | |
| CAC CGT GTT ATG GTC GTG ATA CCA AGA TAT GGA GAG TAT GCC GAA GCC | 768 |
| His Arg Val Met Val Val Ile Pro Arg Tyr Gly Glu Tyr Ala Glu Ala | |
| 910 915 920 925 | |
| CGG GAT TTA GGT GTA AGG AGA CGT TAC AAG GTA GCT GGA CAG GAT TCA | 816 |
| Arg Asp Leu Gly Val Arg Arg Arg Tyr Lys Val Ala Gly Gln Asp Ser | |
| 930 935 940 | |
| GAA GTT ACT TAT TTT CAC TCT TAC ATT GAT GGA GTT GAT TTT GTA TTC | 864 |
| Glu Val Thr Tyr Phe His Ser Tyr Ile Asp Gly Val Asp Phe Val Phe | |
| 945 950 955 | |
| GTA GAA GCC CCT CCC TTC CCG CAC CCG CAC AAT AAT ATT TAT GGG GGA | 912 |
| Val Glu Ala Pro Pro Phe Arg His Arg His Asn Asn Ile Tyr Gly Gly | |
| 960 965 970 | |
| GAA AGA TTG GAT ATT TTG AAG CCG ATG ATT TTG TTC TGC AAG GCC GCT | 960 |
| Glu Arg Leu Asp Ile Leu Lys Arg Met Ile Leu Phe Cys Lys Ala Ala | |

| 975 | 980 | 985 | |
|---|------|------|------|
| GTT GAG GTT CCA TGG TAT GGT CCA TGT GGC GGT ACT GTC TAT GGT GAT | | | 1008 |
| Val Glu Val Pro Trp Tyr Ala Pro Cys Gly Gly Thr Val Tyr Gly Asp | | | |
| 990 | 995 | 1000 | 1005 |
| GGC AAC TTA GTT TTC ATT GGT AAT GAT TGG CAT ACC GCA CTT CTG CCT | | | 1056 |
| Gly Asn Leu Val Phe Ile Ala Asn Asp Trp His Thr Ala Leu Leu Pro | | | |
| | 1010 | 1015 | 1020 |
| GTC TAT CTA AAG GGC TAT TAC CGG GAC AAT GGT TTG ATG CAG TAT GCT | | | 1104 |
| Val Tyr Leu Lys Ala Tyr Tyr Arg Asp Asn Gly Leu Met Gln Tyr Ala | | | |
| | 1025 | 1030 | 1035 |
| CGC TCT GTG CTT GTG ATA CAC AAC ATT GCT CAT CAG GGT CGT GGC CCT | | | 1152 |
| Arg Ser Val Leu Val Ile His Asn Ile Ala His Gln Gly Arg Gly Pro | | | |
| | 1040 | 1045 | 1050 |
| GTA GAC GAC TTC GTC AAT TTT GAC TTG CCT GAA CAC TAC ATC GAC CAC | | | 1200 |
| Val Asp Asp Phe Val Asn Phe Asp Leu Pro Glu His Tyr Ile Asp His | | | |
| | 1055 | 1060 | 1065 |
| TTC AAA CTC TAT GAC AAC ATT GGT CGG GAT CAC AGC AAC GTT TTT GCT | | | 1248 |
| Phe Lys Leu Tyr Asp Asn Ile Gly Gly Asp His Ser Asn Val Phe Ala | | | |
| | 1070 | 1075 | 1080 |
| CGC GGC CTC AAG ACG GCA GAC CGG GTG GTG ACC GTT AGC AAT GGC TAC | | | 1296 |
| Ala Gly Leu Lys Thr Ala Asp Arg Val Val Thr Val Ser Asn Gly Tyr | | | |
| | 1090 | 1095 | 1100 |
| ATG TGG GAG CTG AAG ACT TCG GAA GGC GGG TGG GGC CTC CAC GAC ATC | | | 1344 |
| Met Trp Glu Leu Lys Thr Ser Glu Gly Gly Trp Gly Leu His Asp Ile | | | |
| | 1105 | 1110 | 1115 |
| ATA AAC CAG AAC GAC TGG AAG CTG CAG GGC ATC GTG AAC GGC ATC GAC | | | 1392 |
| Ile Asn Gln Asn Asp Trp Lys Leu Gln Gly Ile Val Asn Gly Ile Asp | | | |
| | 1120 | 1125 | 1130 |
| ATG AGC GAG TGG AAC CCC GGT GTG GAC GTG CAC CTC CAC TCC GAC GAC | | | 1440 |
| Met Ser Glu Trp Asn Pro Ala Val Asp Val His Leu His Ser Asp Asp | | | |
| | 1135 | 1140 | 1145 |
| TAC ACC AAC TAC ACG TTC GAG ACG CTG GAC ACC GGC AAC CGG CAG TGC | | | 1488 |
| Tyr Thr Asn Tyr Thr Phe Glu Thr Leu Asp Thr Gly Lys Arg Gln Cys | | | |
| | 1150 | 1155 | 1160 |
| | | | 1165 |

| | |
|---|------|
| AAG GCC GCC CTG CAG CGG CAG CTG GGC CTG CAG GTC CGC GAC GAC GTG Lys Ala Ala Leu Gln Arg Gln Leu Gly Leu Gln Val Arg Asp Asp Val 1170 1175 1180 | 1536 |
| CCA CTG ATC GCG TTC ATC GGG CGG CTG GAC CAC CAG AAG GGC GTG GAC Pro Leu Ile Gly Phe Ile Gly Arg Leu Asp His Gln Lys Gly Val Asp 1185 1190 1195 | 1584 |
| ATC ATC GCC GAC GCG ATC CAC TGG ATC GCG GGG CAG GAC GTG CAG CTC Ile Ile Ala Asp Ala Ile His Trp Ile Ala Gly Gln Asp Val Gln Leu 1200 1205 1210 | 1632 |
| GTG ATG CTG GCG ACC GGG CGG GCC GAC CTG GAG GAC ATG CTG CGG CGG Val Met Leu Gly Thr Gly Arg Ala Asp Leu Glu Asp Met Leu Arg Arg 1215 1220 1225 | 1680 |
| TTC GAG TCG GAG CAC ACC GAC AAG GTG CGC GCG TGG GTG GCG TTC TCG Phe Glu Ser Glu His Ser Asp Lys Val Arg Ala Trp Val Gly Phe Ser 1230 1235 1240 1245 | 1728 |
| GTG CCG CTG GCG CAC CGC ATC ACG GCG GCG GCG GAC ATC CTG CTG ATG Val Pro Leu Ala His Arg Ile Thr Ala Gly Ala Asp Ile Leu Leu Met 1250 1255 1260 | 1776 |
| CCG TCG CGC TTC GAC CCG TCG GGG CTG AAC CAG CTC TAC GCC ATG GCG Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Ala 1265 1270 1275 | 1824 |
| TAC GCG ACC CTG CCC GTG GTG CAC GCC CTG GGG GCG CTC CCG GAC ACG Tyr Gly Thr Val Pro Val Val His Ala Val Gly Gly Leu Arg Asp Thr 1280 1285 1290 | 1872 |
| GTG GCG CCG TTC GAC CCG TTC AAC GAC ACC CGG CTC GCG TGG ACC TTC Val Ala Pro Phe Asp Pro Phe Asn Asp Thr Gly Leu Gly Trp Thr Phe 1295 1300 1305 | 1920 |
| GAC CGC GCG GAG CGG AAC CGG ATG ATC GAC GCG CTC TCG CAC TGC CTC Asp Arg Ala Glu Ala Asn Arg Met Ile Asp Ala Leu Ser His Cys Leu 1310 1315 1320 1325 | 1968 |
| ACC ACG TAC CGG AAC TAC AAG GAG AGC TGG CGC GCC TGC AGG GCG CGC Thr Thr Tyr Arg Asn Tyr Lys Glu Ser Trp Arg Ala Cys Arg Ala Arg 1330 1335 1340 | 2016 |
| GGC ATG GCC GAG GAC CTC ACC TGG GAC CAC GCC GCC GTG CTC TAT GAG | 2064 |

Gly Met Ala Glu Asp Leu Ser Trp Asp His Ala Ala Val Leu Tyr Glu
 1345 1350 1355

GAC GTC CTC GTC AAG GCG AAG TAC CAG TGG TGA
 Asp Val Leu Val Lys Ala Lys Tyr Gln Trp °
 1360 1365

(2) INFORMATION ZU SEQ ID NO: 11:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 669 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

Met Pro Gly Ala Ile Ser Ser Ser Ser Ala Phe Leu Leu Pro Val
 1 5 10 15

Ala Ser Ser Ser Pro Arg Arg Arg Arg Gly Ser Val Gly Ala Ala Leu
 20 25 30

Arg Ser Tyr Gly Tyr Ser Gly Ala Glu Leu Arg Leu His Trp Ala Arg
 35 40 45

Arg Gly Pro Pro Gln Asp Gly Ala Ala Ser Val Arg Ala Ala Ala Ala
 50 55 60

Pro Ala Gly Gly Glu Ser Glu Glu Ala Ala Lys Ser Ser Ser Ser Ser
 65 70 75 80

Gln Ala Gly Ala Val Gln Gly Ser Thr Ala Lys Ala Val Asp Ser Ala
 85 90 95

Ser Pro Pro Asn Pro Leu Thr Ser Ala Pro Lys Gln Ser Gln Ser Ala
 100 105 110

Ala Met Gln Asn Gly Thr Ser Gly Gly Ser Ser Ala Ser Thr Ala Ala
 115 120 125

Pro Val Ser Gly Pro Lys Ala Asp His Pro Ser Ala Pro Val Thr Lys
 130 135 140

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

Val Asp Asp Phe Val Asn Phe Asp Leu Pro Glu His Tyr Ile Asp His
 385 390 395 400
 Phe Lys Leu Tyr Asp Asn Ile Gly Gly Asp His Ser Asn Val Phe Ala
 405 410 415
 Ala Gly Leu Lys Thr Ala Asp Arg Val Val Thr Val Ser Asn Gly Tyr
 420 425 430
 Met Trp Glu Leu Lys Thr Ser Glu Gly Gly Trp Gly Leu His Asp Ile
 435 440 445
 Ile Asn Gln Asn Asp Trp Lys Leu Gln Gly Ile Val Asn Gly Ile Asp
 450 455 460
 Met Ser Glu Trp Asn Pro Ala Val Asp Val His Leu His Ser Asp Asp
 465 470 475 480
 Tyr Thr Asn Tyr Thr Phe Glu Thr Leu Asp Thr Gly Lys Arg Gln Cys
 485 490 495
 Lys Ala Ala Leu Gln Arg Gln Leu Gly Leu Gln Val Arg Asp Asp Val
 500 505 510
 Pro Leu Ile Gly Phe Ile Gly Arg Leu Asp His Gln Lys Gly Val Asp
 515 520 525
 Ile Ile Ala Asp Ala Ile His Trp Ile Ala Gly Gln Asp Val Gln Leu
 530 535 540
 Val Met Leu Gly Thr Gly Arg Ala Asp Leu Glu Asp Met Leu Arg Arg
 545 550 555 560
 Phe Glu Ser Glu His Ser Asp Lys Val Arg Ala Trp Val Gly Phe Ser
 565 570 575
 Val Pro Leu Ala His Arg Ile Thr Ala Gly Ala Asp Ile Leu Leu Met
 580 585 590
 Pro Ser Arg Phe Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Ala
 595 600 605
 Tyr Gly Thr Val Pro Val Val His Ala Val Gly Gly Leu Arg Asp Thr
 610 615 620

```

Val Ala Pro Phe Asp Pro Phe Asn Asp Thr Gly Leu Gly Trp Thr Phe
625                630                635                640

Asp Arg Ala Glu Ala Asn Arg Met Ile Asp Ala Leu Ser His Cys Leu
                645                650                655

Thr Thr Tyr Arg Asn Tyr Lys Glu Ser Trp Arg Ala Cys Arg Ala Arg
                660                665                670

Gly Met Ala Glu Asp Leu Ser Trp Asp His Ala Ala Val Leu Tyr Glu
        675                680                685

Asp Val Leu Val Lys Ala Lys Tyr Gln Trp *
        690                695

```

(2) INFORMATION ZU SEQ ID NO: 12:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 1752 Basenpaare
- (B) ART: Nukleinsäure
- (C) Strangform: Doppel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA oder mRNA

(iii) HYPOTHETISCH: NEIN

(vi) URSPRÜNGLICHE HERKUNFT

- (A) ORGANISMUS: Zea mays

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) Lage: 1 .. 1752

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

| | |
|---|-----|
| TGC CTC GCG GAG CTG ACC AGG GAG GGG CCC GCG CCG CGC CCG CTG CCA Cys Val Ala Glu Leu Ser Arg Ser Gly Pro Ala Pro Arg Pro Leu Pro 700 705 710 715 | 68 |
| CCC GCG CTC CTG GCG CCC CCG CTC GTG CCC GGC TTC ACC GCG Pro Ala Leu Leu Ala Pro Pro Leu Val Pro Gly Phe Leu Ala Pro 720 725 730 | 96 |
| GCC GAG CCC ACC GGT GAG CCG GCA TCG ACC CCG CCG CCC GTG CCC GAC Ala Glu Pro Thr Gly Glu Pro Ala Ser Thr Pro Pro Pro Val Pro Asp 735 740 745 | 144 |
| GCC GGC CTC GCG GAC CTC GGT CTC GAA CCT GAA GGG ATT GCT GAA GGT Ala Gly Leu Gly Asp Leu Gly Leu Glu Pro Glu Gly Ile Ala Glu Gly 750 755 760 | 192 |
| TCC ATC GAT AAC ACA GTA GTT GTG GCA AGT GAG CAA GAT TCT GAG ATT Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln Asp Ser Glu Ile 765 770 775 | 240 |
| GTG GTT GGA AAG GAG CAA GCT CGA GCT AAA GTA ACA CAA AGC ATT GTC Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr Gln Ser Ile Val 780 785 790 795 | 288 |
| TTT GTA ACC GGC GAA GCT TCT CCT TAT GCA AAG TCT GGG GGT CTA GGA Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser Gly Gly Leu Gly 800 805 810 | 336 |
| GAT GTT TGT GGT TCA TTG CCA GTT GCT CTT GCT GCT CCG GGT CAC CGT Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala Arg Gly His Arg 815 820 825 | 384 |
| GTG ATG GTT GTA ATG CCC AGA TAT TTA AAT GGT ACC TCC GAT AAG AAT Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr Ser Asp Lys Asn 830 835 840 | 432 |
| TAT GCA AAT GCA TTT TAC ACA GAA AAA CAC ATT CCG ATT CCA TGC TTT Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg Ile Pro Cys Phe 845 850 855 | 480 |
| GGC GGT GAA CAT GAA GTT ACC TTC TTC CAT GAG TAT AGA GAT TCA GTT Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr Arg Asp Ser Val 860 865 870 875 | 528 |
| GAC TGG CTG TTT GTT GAT CAT CCC TCA TAT CAC AGA CCT GGA AAT TTA Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg Pro Gly Asn Leu 880 885 890 | 576 |
| TAT GGA GAT AAG TTT GGT GCT TTT GGT GAT AAT CAG TTC AGA TAC ACA Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln Phe Arg Tyr Thr 895 900 905 | 624 |
| CTC CTT TGC TAT GCT GCA TGT GAG GCT CCT TTG ATC CTT GAA TTG GGA | 672 |

| | |
|---|------|
| Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile Leu Glu Leu Gly | |
| 910 915 920 | |
| GCA TAT ATT TAT CGA CAG AAT TGC ATG TTT GTT GTC AAT GAT TGG CAT | 720 |
| Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val Asn Asp Trp His | |
| 925 930 935 | |
| GCC AGT CTA GTG CCA CTC CTT CTT GCT GCA AAA TAT AGA CCA TAT GGT | 768 |
| Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly | |
| 940 945 950 955 | |
| GTT TAT AAA GAC TCC CGC AGC ATT CTT GTA ATA CAT AAT TTA GCA CAT | 816 |
| Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His Asn Leu Ala His | |
| 960 965 970 | |
| CAG GGT GTA GAG CCT GCA AGC ACA TAT CCT GAC CTT GCG TTG CCA CCT | 864 |
| Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro | |
| 975 980 985 | |
| GAA TGG TAT GGA GCT CTG GAG TGG GTA TTC CCT GAA TGG GCG AGG AGG | 912 |
| Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg | |
| 990 995 1000 | |
| CAT GCC CTT GAC AAG GGT GAG GCA GTT AAT TTT TTG AAA GGT GCA GTT | 960 |
| His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val | |
| 1005 1010 1015 | |
| GTG ACA GCA GAT CGA ATC GTG ACT GTC AGT AAG GGT TAT TCG TGG GAG | 1008 |
| Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly Tyr Ser Trp Glu | |
| 1020 1025 1030 1035 | |
| GTG ACA ACT GCT GAA GGT GGA CAG GGC CTC AAT GAG CTC TTA AGC TCC | 1056 |
| Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser | |
| 1040 1045 1050 | |
| AGA AAG AGT GTA TTA AAC GGA ATT GTA AAT GGA ATT GAC ATT AAT GAT | 1104 |
| Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp | |
| 1055 1060 1065 | |
| TGG AAC CCT GCC ACA GAC AAA TGT ATC CCC TGT CAT TAT TCT GTT GAT | 1152 |
| Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His Tyr Ser Val Asp | |
| 1070 1075 1080 | |
| GAC CTC TCT GGA ARG GCC AAA TGT AAA GGT GCA TTG CAG AAG GAG CTG | 1200 |
| Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu Gln Lys Glu Leu | |

| 1085 | 1090 | 1095 | |
|---|------|------|------|
| GGT TTA CCT ATA AGG CCT GAT GTT CCT CTG ATT GGC TTT ATT GGA AGG | | | 1248 |
| Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly Phe Ile Gly Arg | | | |
| 1100 | 1105 | 1110 | 1115 |
| TTG GAT TAT CAG AAA GCC ATT GAT CTC ATT CAA CTT ATC ATA CCA GAT | | | 1296 |
| Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu Ile Ile Pro Asp | | | |
| 1120 | 1125 | | 1130 |
| CTC ATG CGG GAA GAT GTT CAA TTT GTC ATG CTT GGA TCT GGT GAC CCA | | | 1344 |
| Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro | | | |
| 1135 | 1140 | | 1145 |
| GAG CTT GAA GAT TGG ATG AGA TCT ACA GAG TCG ATC TTC AAG GAT AAA | | | 1392 |
| Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile Phe Lys Asp Lys | | | |
| 1150 | 1155 | | 1160 |
| TTT COT GGA TGG GTT GGA TTT AGT GTT CCA GTT TCC CAC CGA ATA ACT | | | 1440 |
| Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr | | | |
| 1165 | 1170 | | 1175 |
| GCC GGC TGC GAT ATA TTG TTA ATG CCA TCC AGA TTC GAA CCT TGT GGT | | | 1488 |
| Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly | | | |
| 1180 | 1185 | 1190 | 1195 |
| CTC AAT CAG CTA TAT GCT ATG CAG TAT GCC ACA GTT CCT GTT GTC CAT | | | 1536 |
| Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His | | | |
| 1200 | 1205 | | 1210 |
| GCA ACT GGG GGC CTT AGA GAT ACC GTG GAG AAC TTC AAC CCT TTC GGT | | | 1584 |
| Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe Asn Pro Phe Gly | | | |
| 1215 | 1220 | | 1225 |
| GAG AAT GGA GAG CAG GGT ACA GGG TGG GCA TTC GCA CCC CTA ACC ACA | | | 1632 |
| Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala Pro Leu Thr Thr | | | |
| 1230 | 1235 | | 1240 |
| GAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC TAC ATA CAG GGA | | | 1680 |
| Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly | | | |
| 1245 | 1250 | | 1255 |
| ACA CAA GTC CTC CTG GGA AGC GCT AAT GAA GCG AGG CAT GTC AAA AGA | | | 1728 |
| Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg | | | |
| 1260 | 1265 | 1270 | 1275 |
| CTT CAC GTG GGA CCA TGC CGC TGA | | | 1752 |
| Leu His Val Gly Pro Cys Arg " | | | |
| 1280 | | | |

(2) INFORMATION ZU SEQ ID NO: 13:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 584 Aminosäuren
 (B) ART: Aminosäure
 (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

```

Cys Val Ala Glu Leu Ser Arg Glu Gly Pro Ala Pro Arg Pro Leu Pro
 1              5              10              15

Pro Ala Leu Leu Ala Pro Pro Leu Val Pro Gly Phe Leu Ala Pro Pro
      20              25              30

Ala Glu Pro Thr Gly Glu Pro Ala Ser Thr Pro Pro Pro Val Pro Asp
      35              40              45

Ala Gly Leu Gly Asp Leu Gly Leu Glu Pro Glu Gly Ile Ala Glu Gly
      50              55              60

Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln Asp Ser Glu Ile
      65              70              75              80

Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr Gln Ser Ile Val
      85              90              95

Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser Gly Gly Leu Gly
      100             105             110

Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala Arg Gly His Arg
      115             120             125

Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr Ser Asp Lys Asn
      130             135             140

Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg Ile Pro Cys Phe
      145             150             155             160

```


Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr Arg Asp Ser Val
165 170 175

Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg Pro Gly Asn Leu
180 185 190

Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln Phe Arg Tyr Thr
195 200 205

Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile Leu Glu Leu Gly
210 215 220

Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val Asn Asp Trp His
225 230 235 240

Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr Arg Pro Tyr Gly
245 250 255

Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His Asn Leu Ala His
260 265 270

Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu Gly Leu Pro Pro
275 280 285

Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu Trp Ala Arg Arg
290 295 300

His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu Lys Gly Ala Val
305 310 315 320

Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly Tyr Ser Trp Glu
325 330 335

Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu Leu Leu Ser Ser
340 345 350

Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile Asp Ile Asn Asp
355 360 365

Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His Tyr Ser Val Asp
370 375 380

Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu Gln Lys Glu Leu
385 390 395 400

Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly Phe Ile Gly Arg
 405 410 415
 Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu Ile Ile Pro Asp
 420 425 430
 Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly Ser Gly Asp Pro
 435 440 445
 Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile Phe Lys Asp Lys
 450 455 460
 Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser His Arg Ile Thr
 465 470 475 480
 Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe Glu Pro Cys Gly
 485 490 495
 Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val Pro Val Val His
 500 505 510
 Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe Asn Pro Phe Gly
 515 520 525
 Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala Pro Leu Thr Thr
 530 535 540
 Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile Tyr Ile Gln Gly
 545 550 555 560
 Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg His Val Lys Arg
 565 570 575
 Leu His Val Gly Pro Cys Arg *
 580

(2) INFORMATION ZU SEQ ID NO: 14:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 2725 Basenpaare
- (B) ART: Nukleinsäure
- (C) Strangform: Einzel

(D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: mRNA

(iii) HYPOTHETISCH: NEIN

(vi) URSPRÜNGLICHE HERKUNFT

(A) ORGANISMUS: Zea mays

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: sig_peptide

(B) Lage: 91 .. 264

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: mat_peptide

(B) Lage: 265 .. 2487

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: sig_peptide

(B) Lage: 91 .. 2490

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

| | |
|--|-----|
| GGCCCCAGAGC AGACCCGGAT TTCGCTCTTG CGGTGCGTGG GGTCTTAGCA TTGGGTGATC | 60 |
| AGTTCGATCC GATCCGGCTG CGAAGGCGAG ATG GCG TTC CGG GTT TCT GGG GCG | 114 |
| Met Ala Phe Arg Val Ser Gly Ala | |
| -58 -55 | |
| GTG CTC GGT GGG GCG GTA AGG GCT CCC CGA CTC ACC GGC GCG GCG GAC | 162 |
| Val Leu Gly Gly Ala Val Arg Ala Pro Arg Leu Thr Gly Gly Gly Glu | |
| -50 -45 -40 -35 | |
| GCT AGT CTA CTC TTC CGG CAC ACC GGC CTC TTC TTA ACT CGG GGT GCT | 210 |
| Gly Ser Leu Val Phe Arg His Thr Gly Leu Phe Leu Thr Arg Gly Ala | |
| -30 -25 -20 | |
| CGA GTT GGA TGT TCG GGG ACG CAC GGG GCC ATG CGC GCG GCG GCG GCG | 258 |
| Arg Val Gly Cys Ser Gly Thr His Gly Ala Met Arg Ala Ala Ala Ala | |
| -15 -10 -5 | |
| GCC AGG AAG GCG GTC ATG GTT CCT GAG GGC GAG AAT GAT GGC CTC GCA | 306 |
| Ala Arg Lys Ala Val Met Val Pro Glu Gly Glu Asn Asp Gly Leu Ala | |
| 1 5 10 | |
| TCA AGG GCT GAC TCG GCT CAA TTC CAG TCG GAT GAA CTG GAG GTA CCA | 354 |
| Ser Arg Ala Asp Ser Ala Gln Phe Gln Ser Asp Glu Leu Glu Val Pro | |
| 15 20 25 30 | |

| | |
|---|-----|
| GAC ATT TCT GAA GAG ACA ACG TGC GGT GCT GGT GTG GCT GAT GCT CAA Asp Ile Ser Glu Glu Thr Thr Cys Gly Ala Gly Val Ala Asp Ala Gln 35 40 45 | 402 |
| GCC TTG AAG ACA GTT CGA GTG GTC CCC CCA CCA AGC GAT GGA CAA AAA Ala Leu Asn Arg Val Arg Val Val Pro Pro Pro Ser Asp Gly Gln Lys 50 55 60 | 450 |
| ATA TTC CAG ATT GAC CCC ATG TTG CAA GGC TAT AAG TAC CAT CTT GAG Ile Phe Gln Ile Asp Pro Met Leu Glu Gly Tyr Lys Tyr His Leu Glu 65 70 75 | 498 |
| TAT CGG TAC AGC CTC TAT AGA ACA ATC CGT TCA GAC ATT GAT GAA CAT Tyr Arg Tyr Ser Leu Tyr Arg Arg Ile Arg Ser Asp Ile Asp Glu His 80 85 90 | 546 |
| GAA GGA GGC TTG GAA GCC TTC TCC CGT ACT TAT CAG AAG TTT GGA TTT Glu Gly Gly Leu Glu Ala Phe Ser Arg Ser Tyr Glu Lys Phe Gly Phe 95 100 105 110 | 594 |
| AAT GCC AGC GCG GAA GGT ATC ACA TAT CGA GAA TGG GCT CCT GGA GCA Asn Ala Ser Ala Glu Gly Ile Thr Tyr Arg Glu Trp Ala Pro Gly Ala 115 120 125 | 642 |
| TTT TCT GCA GCA TTG GTG GGT GAC CTC AAC AAC TGG GAT CCA AAT GCA Phe Ser Ala Ala Leu Val Gly Asp Val Asn Asn Trp Asp Pro Asn Ala 130 135 140 | 690 |
| GAT CGT ATG AGC AAA AAT GAG TTT GGT GTT TGG CAA ATT TTT CTG CCT Asp Arg Met Ser Lys Asn Glu Phe Gly Val Trp Glu Ile Phe Leu Pro 145 150 155 | 738 |
| AAC AAT GCA GAT GGT ACA TCA CCT ATT CCT CAT GGA TCT CGT GTA AAG Asn Asn Ala Asp Gly Thr Ser Pro Ile Pro His Gly Ser Arg Val Lys 160 165 170 | 786 |
| GTG AGA ATG GAT ACT CCA TCA GCG ATA AAG GAT TCA ATT CCA GCC TGG Val Arg Met Asp Thr Pro Ser Gly Ile Lys Asp Ser Ile Pro Ala Trp 175 180 185 190 | 834 |
| ATC AAG TAC TCA GTG CAG GCC CCA GGA GAA ATA CCA TAT GAT GCG ATT Ile Lys Tyr Ser Val Gln Ala Pro Gly Glu Ile Pro Tyr Asp Gly Ile 195 200 205 | 882 |
| TAT TAT GAT CCT CCT GAA GAG GTA AAG TAT GTG TTC AGG CAT GCG CAA | 930 |

| | |
|---|------|
| Tyr Tyr Asp Pro Pro Glu Glu Val Lys Tyr Val Phe Arg His Ala Gln | |
| 210 215 220 | |
| CCT AAA CGA CCA AAA TCA TTG CGG ATA TAT GAA ACA CAT GTC GGA ATG | 978 |
| Pro Lys Arg Pro Lys Ser Leu Arg Ile Tyr Glu Thr His Val Gly Met | |
| 225 230 235 | |
| AGT AGC CCG GAA CCG AAG ATA AAC ACA TAT GTA AAC TTT AGG CAT GAA | 1026 |
| Ser Ser Pro Glu Pro Lys Ile Asn Thr Tyr Val Asn Phe Arg Asp Glu | |
| 240 245 250 | |
| GTC CTC CCA AGA ATA AAA AAA CTT GGA TAC AAT GCA GTG CAA ATA ATG | 1074 |
| Val Leu Pro Arg Ile Lys Lys Leu Gly Tyr Asn Ala Val Gln Ile Met | |
| 255 260 265 270 | |
| CCA ATC CAA GAG CAC TCA TAT TAT GCA AGC TTT GGA TAC CAT GTA ACT | 1122 |
| Ala Ile Gln Glu His Ser Tyr Tyr Gly Ser Phe Gly Tyr His Val Thr | |
| 275 280 285 | |
| AAT TTT TTT GCG CCA AGT AGT CGT TTT GGT ACC CCA GAA CAT TTG AAG | 1170 |
| Asn Phe Phe Ala Pro Ser Ser Arg Phe Gly Thr Pro Glu Asp Leu Lys | |
| 290 295 300 | |
| TCT TTG ATT GAT AGA GCA CAT GAG CTT GGT TTG CTA GTT CTC ATG GAT | 1218 |
| Ser Leu Ile Asp Arg Ala His Glu Leu Gly Leu Leu Val Leu Met Asp | |
| 305 310 315 | |
| GTG GTT CAT AGT CAT GCG TCA AGT AAT ACT CTG CAT GGG TTG AAT GGT | 1266 |
| Val Val His Ser His Ala Ser Ser Asn Thr Leu Asp Gly Leu Asn Gly | |
| 320 325 330 | |
| TTT GAT GGT ACA GAT ACA CAT TAC TTT CAC AGT GGT CCA CGT GGC CAT | 1314 |
| Phe Asp Gly Thr Asp Thr His Tyr Phe His Ser Gly Pro Arg Gly His | |
| 335 340 345 350 | |
| CAC TGG ATG TGG GAT TCT CGC CTA TTT AAC TAT GGG AAC TGG GAA GTT | 1362 |
| His Trp Met Trp Asp Ser Arg Leu Phe Asn Tyr Gly Asn Trp Glu Val | |
| 355 360 365 | |
| TTA AGA TTT CTT CTC TCC AAT GCT AGA TGG TGG CTC GAG CAA TAT AAG | 1410 |
| Leu Arg Phe Leu Leu Ser Asn Ala Arg Trp Trp Leu Glu Glu Tyr Lys | |
| 370 375 380 | |
| TTT GAT GGT TTC CGT TTT GAT GGT GTG ACC TCC ATG ATG TAC ACT CAC | 1458 |
| Phe Asp Gly Phe Arg Phe Asp Gly Val Thr Ser Met Met Tyr Thr His | |

| 385 | 390 | 395 | |
|---|-----|-----|------|
| CAC GGA TTA CAA GTA ACA TTT ACG GGG AAC TTC AAT GAG TAT TTT GGC | | | 1506 |
| His Gly Leu Gln Val Thr Phe Thr Gly Asn Phe Asn Glu Tyr Phe Gly | | | |
| 400 | 405 | 410 | |
| TTT GCC ACC GAT GTA GAT GCA GTG GTT TAC TTG ATG CTG GTA AAT GAT | | | 1554 |
| Phe Ala Thr Asp Val Asp Ala Val Val Tyr Leu Met Leu Val Asn Asp | | | |
| 415 | 420 | 425 | 430 |
| CTA ATT CAT GGA CTT TAT CCT GAG GCT GTA ACC ATT GGT GAA GAT GTT | | | 1602 |
| Leu Ile His Gly Leu Tyr Pro Glu Ala Val Thr Ile Gly Glu Asp Val | | | |
| | 435 | 440 | 445 |
| AGT GGA ATG CCT ACA TTT GCC CTT CCT GTT CAC GAT GGT GGG GTA GGT | | | 1650 |
| Ser Gly Met Pro Thr Phe Ala Leu Pro Val His Asp Gly Gly Val Gly | | | |
| | 450 | 455 | 460 |
| TTT GAC TAT CGG ATG CAT ATG GCT GTG GCT GAC AAA TGG ATT GAC CTT | | | 1698 |
| Phe Asp Tyr Arg Met His Met Ala Val Ala Asp Lys Trp Ile Asp Leu | | | |
| | 465 | 470 | 475 |
| CTC AAG CAA AGT GAT GAA ACT TGG AAG ATG GGT GAT ATT GTG CAC ACA | | | 1746 |
| Leu Lys Gln Ser Asp Glu Thr Trp Lys Met Gly Asp Ile Val His Thr | | | |
| | 480 | 485 | 490 |
| CTG ACA AAT AGG AGG TCG TTA GAG AAG TGT GTA ACT TAT GCT GAA AGT | | | 1794 |
| Leu Thr Asn Arg Arg Trp Leu Glu Lys Cys Val Thr Tyr Ala Glu Ser | | | |
| | 495 | 500 | 505 |
| CAT GAT CAA GCA TTA GTC GGC GAC AAG ACT ATT GCG TTT TGG TTG ATG | | | 1842 |
| His Asp Gln Ala Leu Val Gly Asp Lys Thr Ile Ala Phe Trp Leu Met | | | |
| | 515 | 520 | 525 |
| GAC AAG GAT ATG TAT CAT TTC ATG GCC CTC GAT AGA CCT TCA ACT CCT | | | 1890 |
| Asp Lys Asp Met Tyr Asp Phe Met Ala Leu Asp Arg Pro Ser Thr Pro | | | |
| | 530 | 535 | 540 |
| ACC ATT CAT CGT GGG ATA GCA TTA CAT AAG ATG ATT AGA CTT ATC ACA | | | 1938 |
| Thr Ile Asp Arg Gly Ile Ala Leu His Lys Met Ile Arg Leu Ile Thr | | | |
| | 545 | 550 | 555 |
| ATG GGT TTA GGA GGA GAG GGC TAT CTT AAT TTC ATG GGA AAT GAG TTT | | | 1986 |
| Met Gly Leu Gly Gly Glu Gly Tyr Leu Asn Phe Met Gly Asn Glu Phe | | | |
| | 560 | 565 | 570 |

| | |
|---|------|
| GGA CAT CCT GAA TGG ATA GAT TTT CCA AGA GGT CCG CAA AGA CTT CCA Gly His Pro Glu Trp Ile Asp Phe Pro Arg Gly Pro Gln Arg Leu Pro 575 580 585 590 | 2034 |
| AGT GGT AAG TTT ATT CCA GGG AAT AAC AAC AGT TAT GAC AAA TGT CGT Ser Gly Lys Phe Ile Pro Gly Asn Asn Asn Ser Tyr Asp Lys Cys Arg 595 600 605 | 2082 |
| CGA AGA TTT GAC CTG GGT GAT GCA GAC TAT CTT AGG TAT CAT GGT ATG Arg Arg Phe Asp Leu Gly Asp Ala Asp Tyr Leu Arg Tyr His Gly Met 610 615 620 | 2130 |
| CAA GAG TTT GAT CAG GCA ATG CAA CAT CTT GAG CAA AAA TAT GAA TTC Gln Glu Phe Asp Gln Ala Met Gln His Leu Glu Gln Lys Tyr Glu Phe 625 630 635 | 2178 |
| ATG ACA TCT GAT CAC CAG TAT ATT TCC CGG AAA CAT GAG GAG GAT AAG Met Thr Ser Asp His Gln Tyr Ile Ser Arg Lys His Glu Glu Asp Lys 640 645 650 | 2226 |
| GTG ATT GTG TTC GAA AAG GGA GAT TTG GTA TTT GTG TTC AAC TTC CAC Val Ile Val Phe Glu Lys Gly Asp Leu Val Phe Val Phe Asn Phe His 655 660 665 670 | 2274 |
| TGC AAC AAC AGC TAT TTT GAC TAC CGT ATT GGT TGT CGA AAG CCT GGG Cys Asn Asn Ser Tyr Phe Asp Tyr Arg Ile Gly Cys Arg Lys Pro Gly 675 680 685 | 2322 |
| GTG TAT AAG GTG CTC TTG GAC TCC GAC GCT GGA CTA TTT GGT GGA TTT Val Tyr Lys Val Val Leu Asp Ser Asp Ala Gly Leu Phe Gly Gly Phe 690 695 700 | 2370 |
| ACC AGG ATC CAT CAC GCA GCC GAG CAC TTC ACC GCC GAC TGT TCG CAT Ser Arg Ile His His Ala Ala Glu His Phe Thr Ala Asp Cys Ser His 705 710 715 | 2418 |
| GAT AAT AGG CCA TAT TCA TTC TCG GTT TAT ACA CCA AGC AGA ACA TGT Asp Asn Arg Pro Tyr Ser Phe Ser Val Tyr Thr Pro Ser Arg Thr Cys 720 725 730 | 2466 |
| GTC GTC TAT GCT CCA GTG GAG TGA TAGCGGGGCTA CTCGTTGCTG CGCGGCATGT Val Val Tyr Ala Pro Val Glu * 735 740 | 2520 |
| GTGGGGCTGT CGATGTGAGG AAAAACTTC TTCAAAGCC GGCAGATGCA TGCATGCATG | 2580 |
| CTACAATAAG GTTCTGATAC TTTAATCGAT GGTGGAAGC CCATGCATCT CGCTGCGTTG | 2640 |
| TCCTCTCTAT ATATATAAGA CCTTCAAGGT GTCAATTAAA CATAGAGTTT TCGTTTTTCG | 2700 |
| CTTTCCTAAA AAAAAA AAAA | 2725 |

(2) INFORMATION ZU SEQ ID NO: 15:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 800 Aminosäuren
 (B) ART: Aminosäure
 (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

```

Met Ala Phe Arg Val Ser Gly Ala Val Leu Gly Gly Ala Val Arg Ala
-58          -55          -50          -45

Pro Arg Leu Thr Gly Gly Gly Glu Gly Ser Leu Val Phe Arg His Thr
      -40          -35          -30

Gly Leu Phe Leu Thr Arg Gly Ala Arg Val Gly Cys Ser Gly Thr His
-25          -20          -15

Gly Ala Met Arg Ala Ala Ala Ala Arg Lys Ala Val Met Val Pro
-10          -5          1          5

Glu Gly Glu Asn Asp Gly Leu Ala Ser Arg Ala Asp Ser Ala Gln Phe
      10          15          20

Gln Ser Asp Glu Leu Glu Val Pro Asp Ile Ser Glu Glu Thr Thr Cys
      25          30          35

Gly Ala Gly Val Ala Asp Ala Gln Ala Leu Asn Arg Val Arg Val Val
      40          45          50

Pro Pro Pro Ser Asp Gly Gln Lys Ile Phe Gln Ile Asp Pro Met Leu
      55          60          65          70

Gln Gly Tyr Lys Tyr His Leu Glu Tyr Arg Tyr Ser Leu Tyr Arg Arg
      75          80          85

```

Ile Arg Ser Asp Ile Asp Glu His Glu Gly Gly Leu Glu Ala Phe Ser
90 95 100

Arg Ser Tyr Glu Lys Phe Gly Phe Asn Ala Ser Ala Glu Gly Ile Thr
105 110 115

Tyr Arg Glu Trp Ala Pro Gly Ala Phe Ser Ala Ala Leu Val Gly Asp
120 125 130

Val Asn Asn Trp Asp Pro Asn Ala Asp Arg Met Ser Lys Asn Glu Phe
135 140 145 150

Gly Val Trp Glu Ile Phe Leu Pro Asn Asn Ala Asp Gly Thr Ser Pro
155 160 165

Ile Pro His Gly Ser Arg Val Lys Val Arg Met Asp Thr Pro Ser Gly
170 175 180

Ile Lys Asp Ser Ile Pro Ala Trp Ile Lys Tyr Ser Val Gln Ala Pro
185 190 195

Gly Glu Ile Pro Tyr Asp Gly Ile Tyr Tyr Asp Pro Pro Glu Glu Val
200 205 210

Lys Tyr Val Phe Arg His Ala Gln Pro Lys Arg Pro Lys Ser Leu Arg
215 220 225 230

Ile Tyr Glu Thr His Val Gly Met Ser Ser Pro Glu Pro Lys Ile Asn
235 240 245

Thr Tyr Val Asn Phe Arg Asp Glu Val Leu Pro Arg Ile Lys Lys Leu
250 255 260

Gly Tyr Asn Ala Val Gln Ile Met Ala Ile Gln Glu His Ser Tyr Tyr
265 270 275

Gly Ser Phe Gly Tyr His Val Thr Asn Phe Phe Ala Pro Ser Ser Arg
280 285 290

Phe Gly Thr Pro Glu Asp Leu Lys Ser Leu Ile Asp Arg Ala His Glu
295 300 305 310

Leu Gly Leu Leu Val Leu Met Asp Val Val His Ser His Ala Ser Ser
315 320 325

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

Leu Asn Phe Met Gly Asn Glu Phe Gly His Pro Glu Trp Ile Asp Phe
 570 575 580

Pro Arg Gly Pro Gln Arg Leu Pro Ser Gly Lys Phe Ile Pro Gly Asn
 585 590 595

Asn Asn Ser Tyr Asp Lys Cys Arg Arg Arg Phe Asp Leu Gly Asp Ala
 600 605 610

Asp Tyr Leu Arg Tyr His Gly Met Gln Glu Phe Asp Gln Ala Met Gln
 615 620 625 630

His Leu Glu Gln Lys Tyr Glu Phe Met Thr Ser Asp His Gln Tyr Ile
 635 640 645

Ser Arg Lys His Glu Glu Asp Lys Val Ile Val Phe Glu Lys Gly Asp
 650 655 660

Leu Val Phe Val Phe Asn Phe His Cys Asn Asn Ser Tyr Phe Asp Tyr
 665 670 675

Arg Ile Gly Cys Arg Lys Pro Gly Val Tyr Lys Val Val Leu Asp Ser
 680 685 690

Asp Ala Gly Leu Phe Gly Gly Phe Ser Arg Ile His His Ala Ala Glu
 695 700 705 710

His Phe Thr Ala Asp Cys Ser His Asp Asn Arg Pro Tyr Ser Phe Ser
 715 720 725

Val Tyr Thr Pro Ser Arg Thr Cys Val Val Tyr Ala Pro Val Glu *
 730 735 740

(2) INFORMATION ZU SEQ ID NO: 16:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 2763 Basenpaare
- (B) ART: Nukleinsäure
- (C) Strangform: Einzel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: mRNA

(iii) HYPOTHETISCH: NEIN

(vi) URSPRÜNGLICHE HERKUNFT

(A) ORGANISMUS: Zea mays

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: transit_peptide

(B) Lage: 2 .. 190

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: mat_peptide

(B) Lage: 191 .. 2467

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) Lage: 2 .. 2470

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

| | |
|--|-----|
| G CTG TCG CTC GTC TCG CCC TCT TCG TCG CCG ACT CCG CTT CCG CCG | 46 |
| Leu Cys Leu Val Ser Pro Ser Ser Ser Pro Thr Pro Leu Pro Pro | |
| -63 -60 -55 -50 | |
| CCG CCG CCG TCT CCG TCG CAT GGT GAT CCG GCG GCA CCG CCG GGG ATC | 94 |
| Pro Arg Arg Ser Arg Ser His Ala Asp Arg Ala Ala Pro Pro Gly Ile | |
| -45 -40 -35 | |
| GCG GGT CCG GGC AAT GTG CCG CTG AGT GTG TTC TCT GTC CAG TGC AAG | 142 |
| Ala Gly Gly Gly Asn Val Arg Leu Ser Val Leu Ser Val Gln Cys Lys | |
| -30 -25 -20 | |
| GCT CCG CCG TCA GGC GTG CCG AAG GTC AAG AGC AAA TTC GCC ACT GCA | 190 |
| Ala Arg Arg Ser Gly Val Arg Lys Val Lys Ser Lys Phe Ala Thr Ala | |
| -15 -10 -5 | |
| GCT ACT GTG GAA GAA CAT AAA ACT ATG GCA ACT GCC AAA GGC GAT GTC | 238 |
| Ala Thr Val Gln Glu Asp Lys Thr Met Ala Thr Ala Lys Gly Asp Val | |
| 1 5 10 15 | |
| GAC CAT CTC CCC ATA TAC GAC CTG GAC CCC AAG CTG GAG ATA TTC AAG | 286 |
| Asp His Leu Pro Ile Tyr Asp Leu Asp Pro Lys Leu Glu Ile Phe Lys | |
| 20 25 30 | |
| GAC CAT TTC AGG TAC CCG ATG AAA AGA TTC CTA GAG CAG AAA GGA TCA | 334 |
| Asp His Phe Arg Tyr Arg Met Lys Arg Phe Leu Glu Gln Lys Gly Ser | |
| 35 40 45 | |

| | |
|---|-----|
| ATT GAA GAA AAT GAG GGA AGT CTT GAA TCT TTT TCT AAA GGC TAT TTG Ile Glu Glu Asn Glu Gly Ser Leu Glu Ser Phe Ser Lys Gly Tyr Leu 50 55 60 | 382 |
| AAA TTT GGG ATT AAT ACA AAT GAG GAT GGA ACT GTA TAT CGT GAA TGG Lys Phe Gly Ile Asn Thr Asn Glu Asp Gly Thr Val Tyr Arg Glu Trp 65 70 75 80 | 430 |
| GCA CCT GCT GCG CAG GAG GCA GAG CTT ATT GGT GAC TTC AAT GAC TGG Ala Pro Ala Ala Gln Glu Ala Glu Leu Ile Gly Asp Phe Asn Asp Trp 85 90 95 | 478 |
| AAT GGT GCA AAC CAT AAG ATG GAG AAG GAT AAA TTT GGT GTT TGG TCG Asn Gly Ala Asn His Lys Met Glu Lys Asp Lys Phe Gly Val Trp Ser 100 105 110 | 526 |
| ATC AAA ATT GAC CAT GTC AAA GGG AAA COT GCC ATC COT CAC AAT TCC Ile Lys Ile Asp His Val Lys Gly Lys Pro Ala Ile Pro His Asn Ser 115 120 125 | 574 |
| AAG GTT AAA TTT CGC TTT CTA CAT GGT GGA GTA TGG GTT GAT CGT ATT Lys Val Lys Phe Arg Phe Leu His Gly Gly Val Trp Val Asp Arg Ile 130 135 140 | 622 |
| GCA GCA TTG ATT CGT TAT GCG ACT GTT GAT GCC TCT AAA TTT GCA GCT Pro Ala Leu Ile Arg Tyr Ala Thr Val Asp Ala Ser Lys Phe Gly Ala 145 150 155 160 | 670 |
| CCC TAT GAT GGT GTT CAT TGG GAT COT COT GCT TCT GAA AGG TAC ACA Pro Tyr Asp Gly Val His Trp Asp Pro Pro Ala Ser Glu Arg Tyr Thr 165 170 175 | 718 |
| TTT AAG CAT COT CGG CCT TCA AAG COT GCT GCT CCA CGT ATC TAT GAA Phe Lys His Pro Arg Pro Ser Lys Pro Ala Ala Pro Arg Ile Tyr Glu 180 185 190 | 766 |
| GCC CAT GTA GGT ATG AGT GGT GAA AAG CCA GCA GTA AGC ACA TAT AGG Ala His Val Gly Met Ser Gly Glu Lys Pro Ala Val Ser Thr Tyr Arg 195 200 205 | 814 |
| GAA TTT GCA GAC AAT GTG TTG CCA CGC ATA CGA GCA AAT AAC TAC AAC Glu Phe Ala Asp Asn Val Leu Pro Arg Ile Arg Ala Asn Asn Tyr Asn 210 215 220 | 862 |
| ACA GTT CAG TTG ATG GCA GTT ATG GAG CAT TCG TAC TAT GCT TCT TTC | 910 |

| | | | |
|---|-----|-----|------|
| Thr Val Gln Leu Met Ala Val Met Glu His Ser Tyr Tyr Ala Ser Phe | | | |
| 225 | 230 | 235 | 240 |
| GGG TAC CAT GTG ACA AAT TTC TTT GCC GTT ACC AGC ACA TCA GGC ACA | | | 958 |
| Gly Tyr His Val Thr Asn Phe Phe Ala Val Ser Ser Arg Ser Gly Thr | | | |
| 245 | 250 | 255 | |
| CCA GAG GAC CTC AAA TAT CTT GTT GAT AAG CCA CAC AGT TTG GGT TTG | | | 1006 |
| Pro Glu Asp Leu Lys Tyr Leu Val Asp Lys Ala His Ser Leu Gly Leu | | | |
| 260 | 265 | 270 | |
| CGA GTT CTG ATG GAT GTT GTC CAT AGC CAT GCA ACT AAT AAT GTC ACA | | | 1054 |
| Arg Val Leu Met Asp Val Val His Ser His Ala Ser Asn Asn Val Thr | | | |
| 275 | 280 | 285 | |
| GAT GGT TTA AAT GGC TAT GAT GTT GGA CAA AGC ACC CAA GAG TCC TAT | | | 1102 |
| Asp Gly Leu Asn Gly Tyr Asp Val Gly Gln Ser Thr Gln Glu Ser Tyr | | | |
| 290 | 295 | 300 | |
| TTT CAT GCG GCA GAT AGA GGT TAT CAT AAA CTT TGG GAT AGT CGG CTG | | | 1150 |
| Phe His Ala Gly Asp Arg Gly Tyr His Lys Leu Trp Asp Ser Arg Leu | | | |
| 305 | 310 | 315 | 320 |
| TTC AAC TAT GCT AAC TGG GAG GTA TTA AGC TTT CTT CTT TCT AAC CTG | | | 1198 |
| Phe Asn Tyr Ala Asn Trp Glu Val Leu Arg Phe Leu Leu Ser Asn Leu | | | |
| 325 | 330 | 335 | |
| AGA TAT TGG TTG GAT GAA TTC ATG TTT GAT GGC TTC CGA TTT GAT GGA | | | 1246 |
| Arg Tyr Trp Leu Asp Glu Phe Met Phe Asp Gly Phe Arg Phe Asp Gly | | | |
| 340 | 345 | 350 | |
| GTT ACA TCA ATG CTG TAT CAT CAC CAT GGT ATC AAT GTG GGG TTT ACT | | | 1294 |
| Val Thr Ser Met Leu Tyr His His His Gly Ile Asn Val Gly Phe Thr | | | |
| 355 | 360 | 365 | |
| GGA AAC TAC CAG GAA TAT TTC AGT TTG GAC ACA GCT GTG GAT GCA GTT | | | 1342 |
| Gly Asn Tyr Gln Glu Tyr Phe Ser Leu Asp Thr Ala Val Asp Ala Val | | | |
| 370 | 375 | 380 | |
| GTT TAC ATG ATG CTT GCA AAC CAT TTA ATC CAC AAA CTC TTG CCA GAA | | | 1390 |
| Val Tyr Met Met Leu Ala Asn His Leu Met His Lys Leu Leu Pro Glu | | | |
| 385 | 390 | 395 | 400 |
| GCA ACT GTT GTT GCT GAA GAT GTT TCA GGC ATG CCG GTC CTT TGC CGG | | | 1438 |
| Ala Thr Val Val Ala Glu Asp Val Ser Gly Met Pro Val Leu Cys Arg | | | |

| 405 | 410 | 415 | |
|---|-----|-----|------|
| CCA GTT GAT GAA GGT GGG GTT GGG TTT GAC TAT CGC CTG GCA ATG GCT | | | 1486 |
| Pro Val Asp Glu Gly Gly Val Gly Phe Asp Tyr Arg Leu Ala Met Ala | | | |
| 420 | 425 | 430 | |
| ATC CCT GAT AGA TGG ATT GAC TAC CTG AAG AAT AAA GAT GAC TCT CAG | | | 1534 |
| Ile Pro Asp Arg Trp Ile Asp Tyr Leu Lys Asn Lys Asp Asp Ser Glu | | | |
| 435 | 440 | 445 | |
| TGG TCG ATG GGT GAA ATA GCG CAT ACT TTG ACT AAC AGG AGA TAT ACT | | | 1582 |
| Trp Ser Met Gly Glu Ile Ala His Thr Leu Thr Asn Arg Arg Tyr Thr | | | |
| 450 | 455 | 460 | |
| GAA AAA TGC ATC GCA TAT GCT GAG AGC CAT GAT CAG TCT ATT GTT GGC | | | 1630 |
| Glu Lys Cys Ile Ala Tyr Ala Glu Ser His Asp Gln Ser Ile Val Gly | | | |
| 465 | 470 | 475 | 480 |
| GAC AAA ACT ATT GCA TTT CTC CTG ATG GAC AAG GAA ATG TAC ACT GGC | | | 1678 |
| Asp Lys Thr Ile Ala Phe Leu Leu Met Asp Lys Glu Met Tyr Thr Gly | | | |
| 485 | 490 | 495 | |
| ATG TCA GAC TTG CAG CCT GCT TCA CCT ACA ATT GAT CGA GGG ATT GCA | | | 1726 |
| Met Ser Asp Leu Gln Pro Ala Ser Pro Thr Ile Asp Arg Gly Ile Ala | | | |
| 500 | 505 | 510 | |
| CTC CAA AAG ATG ATT CAC TTC ATC ACA ATG GCC CTT GGA GGT GAT GGC | | | 1774 |
| Leu Gln Lys Met Ile His Phe Ile Thr Met Ala Leu Gly Gly Asp Gly | | | |
| 515 | 520 | 525 | |
| TAC TTG AAT TTT ATG GCA AAT GAG TTT GGT CAC CCA GAA TGG ATT GAC | | | 1822 |
| Tyr Leu Asn Phe Met Gly Asn Glu Phe Gly His Pro Glu Trp Ile Asp | | | |
| 530 | 535 | 540 | |
| TTT CCA AGA GAA GGG AAC AAC TGG AGC TAT GAT AAA TGC AGA CGA CAG | | | 1870 |
| Phe Pro Arg Glu Gly Asn Asn Trp Ser Tyr Asp Lys Cys Arg Arg Gln | | | |
| 545 | 550 | 555 | 560 |
| TGG AGC CTT GTG GAC ACT GAT CAC TTG CGG TAC AAG TAC ATG AAT GCG | | | 1918 |
| Trp Ser Leu Val Asp Thr Asp His Leu Arg Tyr Lys Tyr Met Asn Ala | | | |
| 565 | 570 | 575 | |
| TTT GAC CAA GCG ATG AAT GCG CTC GAT GAG AGA TTT TCC TTC CTT TCG | | | 1966 |
| Phe Asp Gln Ala Met Asn Ala Leu Asp Glu Arg Phe Ser Phe Leu Ser | | | |
| 580 | 585 | 590 | |

| | |
|--|------|
| TCG TCA AAG CAG ATC GTC AGC GAC ATG AAC GAT GAG GAA AAG GTT ATT | 2014 |
| Ser Ser Lys Gln Ile Val Ser Asp Met Asn Asp Glu Glu Lys Val Ile | |
| 595 600 605 | |
| GTC TTT GAA CGT GGA GAT TTA GTT TTT GTT TTC AAT TTC CAT CCG AAG | 2062 |
| Val Phe Glu Arg Gly Asp Leu Val Phe Val Phe Asn Phe His Pro Lys | |
| 610 615 620 | |
| AAA ACT TAC GAG GGC TAC AAA GTG GGA TGC GAT TTG CCT GGG AAA TAC | 2110 |
| Lys Thr Tyr Glu Gly Tyr Lys Val Gly Cys Asp Leu Pro Gly Lys Tyr | |
| 625 630 635 640 | |
| AGA GTA GCC CTG GAC TGT GAT GCT CTG CTC TTC GGT GGA CAT GGA AGA | 2158 |
| Arg Val Ala Leu Asp Ser Asp Ala Leu Val Phe Gly Gly His Gly Arg | |
| 645 650 655 | |
| GTT GGC CAC GAC GTG GAT CAC TTC ACG TCG CCT GAA GGG GTG CCA GGG | 2206 |
| Val Gly His Asp Val Asp His Phe Thr Ser Pro Glu Gly Val Pro Gly | |
| 660 665 670 | |
| GTG CCC GAA ACG AAC TTC AAC AAC CGG CCG AAC TCG TTC AAA GTC CTT | 2254 |
| Val Pro Glu Thr Asn Phe Asn Asn Arg Pro Asn Ser Phe Lys Val Leu | |
| 675 680 685 | |
| TCT CCG CCC CGC ACC TGT GTG GCT TAT TAC CGT GTA GAC GAA GCA GCG | 2302 |
| Ser Pro Pro Arg Thr Cys Val Ala Tyr Tyr Arg Val Asp Glu Ala Gly | |
| 690 695 700 | |
| GCT GGA CGA CGT CTT CAC GCG AAA GCA GAG ACA GGA AAG ACG TCT CCA | 2350 |
| Ala Gly Arg Arg Leu His Ala Lys Ala Glu Thr Gly Lys Thr Ser Pro | |
| 705 710 715 720 | |
| GCA GAG AGC ATC GAC GTC AAA GCT TCC AGA CCT AGT AGC AAA GAA GAC | 2398 |
| Ala Glu Ser Ile Asp Val Lys Ala Ser Arg Ala Ser Ser Lys Glu Asp | |
| 725 730 735 | |
| AAG GAG GCA ACG GCT GCT GGC AAG AAG GGA TGG AAG TTT CCG CGG CAG | 2446 |
| Lys Glu Ala Thr Ala Gly Gly Lys Lys Gly Trp Lys Phe Ala Arg Gln | |
| 740 745 750 | |
| CCA TCC GAT CAA GAT ACC AAA TGA AGCCACGAGT CATTGGTGAG GACTGGACTG | 2500 |
| Pro Ser Asp Gln Asp Thr Lys * | |
| 755 760 | |
| GCTGCCGGCG CCCTGTTAGT AGTCTGCTC TACTGGAATA GCCGCCGCTG GCCGCCCTGG | 2560 |
| AACGGTCCTT TCCTGTAGCT TGCAGGCGAC TGGTGTCTCA TCACCGAGCA GGCAGGCACT | 2620 |
| GCTTGATATAG CTTTCTAGTA ATAATAATCA GGGATGGATG GATGGTGTGT ATTGGCTATC | 2680 |
| TGGCTAGACG TGCATGTGCC CAGTTTGTAT GTACAGGAGC AGTTCCCGTC CAGAATAAAA | 2740 |
| AAAAACTTGT TGGCGGGTTT TTC | 2763 |

2) INFORMATION ZU SEQ ID NO: 17:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 823 Aminosäuren
 (B) ART: Aminosäure
 (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 17:

```

Leu Cys Leu Val Ser Pro Ser Ser Ser Pro Thr Pro Leu Pro Pro Pro
-63          -60          -55          -50

Arg Arg Ser Arg Ser His Ala Asp Arg Ala Ala Pro Pro Gly Ile Ala
      -45          -40          -35

Gly Gly Gly Asn Val Arg Leu Ser Val Leu Ser Val Gln Cys Lys Ala
-30          -25          -20

Arg Arg Ser Gly Val Arg Lys Val Lys Ser Lys Phe Ala Thr Ala Ala
-15          -10          -5          1

Thr Val Gln Glu Asp Lys Thr Met Ala Thr Ala Lys Gly Asp Val Asp
      5          10          15

His Leu Pro Ile Tyr Asp Leu Asp Pro Lys Leu Glu Ile Phe Lys Asp
      20          25          30

His Phe Arg Tyr Arg Met Lys Arg Phe Leu Glu Gln Lys Gly Ser Ile
      35          40          45

Glu Glu Asn Glu Gly Ser Leu Glu Ser Phe Ser Lys Gly Tyr Leu Lys
      50          55          60          65

Phe Gly Ile Asn Thr Asn Glu Asp Gly Thr Val Tyr Arg Glu Trp Ala

```

70

75

80

Pro Ala Ala Gln Glu Ala Glu Leu Ile Gly Asp Phe Asn Asp Trp Asn
85 90 95

Gly Ala Asn His Lys Met Glu Lys Asp Lys Phe Gly Val Trp Ser Ile
100 105 110

Lys Ile Asp His Val Lys Gly Lys Pro Ala Ile Pro His Asn Ser Lys
115 120 125

Val Lys Phe Arg Phe Leu His Gly Gly Val Trp Val Asp Arg Ile Pro
130 135 140 145

Ala Leu Ile Arg Tyr Ala Thr Val Asp Ala Ser Lys Phe Gly Ala Pro
150 155 160

Tyr Asp Gly Val His Trp Asp Pro Pro Ala Ser Glu Arg Tyr Thr Phe
165 170 175

Lys His Pro Arg Pro Ser Lys Pro Ala Ala Pro Arg Ile Tyr Glu Ala
180 185 190

His Val Gly Met Ser Gly Glu Lys Pro Ala Val Ser Thr Tyr Arg Glu
195 200 205

Phe Ala Asp Asn Val Leu Pro Arg Ile Arg Ala Asn Asn Tyr Asn Thr
210 215 220 225

Val Gln Leu Met Ala Val Met Glu His Ser Tyr Tyr Ala Ser Phe Gly
230 235 240

Tyr His Val Thr Asn Phe Phe Ala Val Ser Ser Arg Ser Gly Thr Pro
245 250 255

Glu Asp Leu Lys Tyr Leu Val Asp Lys Ala His Ser Leu Gly Leu Arg
260 265 270

Val Leu Met Asp Val Val His Ser His Ala Ser Asn Asn Val Thr Asp
275 280 285

Gly Leu Asn Gly Tyr Asp Val Gly Gln Ser Thr Gln Glu Ser Tyr Phe
290 295 300 305

His Ala Gly Asp Arg Gly Tyr His Lys Leu Trp Asp Ser Arg Leu Phe

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

| | | |
|---|-----|---------|
| 550 | 555 | 560 |
| Ser Leu Val Asp Thr Asp His Leu Arg Tyr Lys Tyr Met Asn Ala Phe | | |
| 565 | 570 | 575 |
| Asp Gln Ala Met Asn Ala Leu Asp Glu Arg Phe Ser Phe Leu Ser Ser | | |
| 580 | 585 | 590 |
| Ser Lys Gln Ile Val Ser Asp Met Asn Asp Glu Glu Lys Val Ile Val | | |
| 595 | 600 | 605 |
| Phe Glu Arg Gly Asp Leu Val Phe Val Phe Asn Phe His Pro Lys Lys | | |
| 610 | 615 | 620 625 |
| Thr Tyr Glu Gly Tyr Lys Val Gly Cys Asp Leu Pro Gly Lys Tyr Arg | | |
| 630 | 635 | 640 |
| Val Ala Leu Asp Ser Asp Ala Leu Val Phe Gly Gly His Gly Arg Val | | |
| 645 | 650 | 655 |
| Gly His Asp Val Asp His Phe Thr Ser Pro Glu Gly Val Pro Gly Val | | |
| 660 | 665 | 670 |
| Pro Glu Thr Asn Phe Asn Asn Arg Pro Asn Ser Phe Lys Val Leu Ser | | |
| 675 | 680 | 685 |
| Pro Pro Arg Thr Cys Val Ala Tyr Tyr Arg Val Asp Glu Ala Gly Ala | | |
| 690 | 695 | 700 705 |
| Gly Arg Arg Leu His Ala Lys Ala Glu Thr Gly Lys Thr Ser Pro Ala | | |
| 710 | 715 | 720 |
| Glu Ser Ile Asp Val Lys Ala Ser Arg Ala Ser Ser Lys Glu Asp Lys | | |
| 725 | 730 | 735 |
| Glu Ala Thr Ala Gly Gly Lys Lys Gly Trp Lys Phe Ala Arg Gln Pro | | |
| 740 | 745 | 750 |
| Ser Asp Gln Asp Thr Lys * | | |
| 755 | 760 | |

(2) INFORMATION ZU SEQ ID NO: 18:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 153 Basenpaare
- (B) ART: Nukleinsäure
- (C) Strangform: Einzel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA bis mRNA

(iii) HYPOTHETISCH: NEIN

(vi) URSPRÜNGLICHE HERKUNFT

- (A) ORGANISMUS: Zea mays

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) Lage: 1 .. 153

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 18:

| | |
|---|-----|
| ATG GCG ACG CCC TCG GCG GTG GCG GCG GCG TGC CTC CTC CTC GCG CGG | 48 |
| Met Ala Thr Pro Ser Ala Val Gly Ala Ala Cys Leu Leu Leu Ala Arg | |
| 765 770 775 | |
| GCC GCG TGG CCG GCG GCG GTC GCG GAG CCG GCG CCG CCG CCG AGG CTC | 96 |
| Ala Ala Trp Pro Ala Ala Val Gly Asp Arg Ala Arg Pro Arg Arg Leu | |
| 780 785 790 | |
| CAG GCG GTG CTG CCG GCG CCG TGC GTC CCG GAG CTC ACC AGG GAG GGG | 144 |
| Gln Arg Val Leu Arg Arg Arg Cys Val Ala Glu Leu Ser Arg Glu Gly | |
| 795 800 805 | |
| CCC CAT ATG | 153 |
| Pro His Met | |
| 810 | |

2) INFORMATION ZU SEQ ID NO: 19:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 51 Aminosäuren

(B) ART: Aminosäure

(D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 19:

```
Met Ala Thr Pro Ser Ala Val Gly Ala Ala Cys Leu Leu Leu Ala Arg
 1           5           10           15
Ala Ala Trp Pro Ala Ala Val Gly Asp Arg Ala Arg Pro Arg Arg Leu
          20          25          30
Gln Arg Val Leu Arg Arg Arg Cys Val Ala Glu Leu Ser Arg Glu Gly
          35          40          45
Pro His Met
      50
```

(2) INFORMATION ZU SEQ ID NO: 20:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 1620 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Doppel

(D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA bis mRNA

(iii) HYPOTHETISCH: NEIN

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) Lage: 1 .. 1620

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 20:

| | |
|---|-----|
| TGC GTC GCG GAG CTC AGC AGG GAG GAC CTC GGT CTC GAA CCT GAA GGG | 48 |
| Cys Val Ala Glu Leu Ser Arg Glu Asp Leu Gly Leu Glu Pro Glu Gly | |
| 55 60 65 | |
| ATT GCT GAA GGT TCC ATC GAT AAC ACA GTA GTT CTC GCA AGT GAG CAA | 96 |
| Ile Ala Glu Gly Ser Ile Asp Asn Thr Val Val Val Ala Ser Glu Gln | |
| 70 75 80 | |
| GAT TCT GAG ATT CTC GTT GGA AAG GAG CAA GCT CGA GCT AAA GTA ACA | 144 |
| Asp Ser Glu Ile Val Val Gly Lys Glu Gln Ala Arg Ala Lys Val Thr | |
| 85 90 95 | |
| CAA AGC ATT GTC TTT GTA ACC GGC GAA GCT TCT CCT TAT GCA AAG TCT | 192 |

Gln Ser Ile Val Phe Val Thr Gly Glu Ala Ser Pro Tyr Ala Lys Ser
100 105 110 115

GGG GGT CTA GGA GAT GTT TGT GGT TCA TTG CCA GPT GGT CTT GCT GCT 240
Gly Gly Leu Gly Asp Val Cys Gly Ser Leu Pro Val Ala Leu Ala Ala
120 125 130

CGT GGT CAC CGT GTG ATG GTT GTA ATG CCC AGA TAT TTA AAT GGT ACC 288
Arg Gly His Arg Val Met Val Val Met Pro Arg Tyr Leu Asn Gly Thr
135 140 145

TCC GAT AAG AAT TAT GCA AAT GCA TTT TAC ACA GAA AAA CAC ATT CGG 336
Ser Asp Lys Asn Tyr Ala Asn Ala Phe Tyr Thr Glu Lys His Ile Arg
150 155 160

ATT CCA TGC TTT GGC GGT GAA CAT GAA GTT ACC TTC TTC CAT GAC TAT 384
Ile Pro Cys Phe Gly Gly Glu His Glu Val Thr Phe Phe His Glu Tyr
165 170 175

AGA GAT TCA GTT GAC TGG GTG TTT GTT GAT CAT CCC TCA TAT CAC AGA 432
Arg Asp Ser Val Asp Trp Val Phe Val Asp His Pro Ser Tyr His Arg
180 185 190 195

CCT GGA AAT TTA TAT GGA GAT AAG TTT GGT GCT TTT GGT GAT AAT CAG 480
Pro Gly Asn Leu Tyr Gly Asp Lys Phe Gly Ala Phe Gly Asp Asn Gln
200 205 210

TTC AGA TAC ACA CTC CTT TGC TAT GCT GCA TGT GAG GCT CCT TTG ATC 528
Phe Arg Tyr Thr Leu Leu Cys Tyr Ala Ala Cys Glu Ala Pro Leu Ile
215 220 225

CTT GAA TTG GGA GGA TAT ATT TAT GGA CAG AAT TGC ATG TTT GTT CTC 576
Leu Glu Leu Gly Gly Tyr Ile Tyr Gly Gln Asn Cys Met Phe Val Val
230 235 240

AAT GAT TGG CAT GCC AGT CTA GTG CCA GTC CTT CTT GCT CCA AAA TAT 624
 Asn Asp Trp His Ala Ser Leu Val Pro Val Leu Leu Ala Ala Lys Tyr
 245 250 255

AGA CCA TAT GGT GTT TAT AAA GAC TCC CGC AGC ATT CTT GTA ATA CAT 672
Arg Pro Tyr Gly Val Tyr Lys Asp Ser Arg Ser Ile Leu Val Ile His
260 265 270 275

AAT TTA GCA CAT CAG GGT GTA GAG CCT GCA AGC ACA TAT CCT GAC CTT 720
Asn Leu Ala His Gln Gly Val Glu Pro Ala Ser Thr Tyr Pro Asp Leu

| 280 | 285 | 290 | |
|---|-----|-----|------|
| GGG TTG CCA CCT GAA TGG TAT GGA GCT CTG GAG TGG GTA TTC CCT GAA | | | 768 |
| Gly Leu Pro Pro Glu Trp Tyr Gly Ala Leu Glu Trp Val Phe Pro Glu | | | |
| 295 | 300 | 305 | |
| TGG GCG AGG AGG CAT GCC CTT GAC AAG GGT GAG GCA GTT AAT TTT TTC | | | 816 |
| Trp Ala Arg Arg His Ala Leu Asp Lys Gly Glu Ala Val Asn Phe Leu | | | |
| 310 | 315 | 320 | |
| AAA GGT GCA GTT GTG ACA GCA GAT CGA ATC GTG ACT GTC AGT AAG GGT | | | 864 |
| Lys Gly Ala Val Val Thr Ala Asp Arg Ile Val Thr Val Ser Lys Gly | | | |
| 325 | 330 | 335 | |
| TAT TCG TGG GAG GTC ACA ACT GCT GAA GGT GGA CAG GGC CTC AAT GAG | | | 912 |
| Tyr Ser Trp Glu Val Thr Thr Ala Glu Gly Gly Gln Gly Leu Asn Glu | | | |
| 340 | 345 | 350 | 355 |
| CTC TTA AGC TCC ACA AAG AGT GTA TTA AAC GGA ATT GTA AAT GGA ATT | | | 960 |
| Leu Leu Ser Ser Arg Lys Ser Val Leu Asn Gly Ile Val Asn Gly Ile | | | |
| 360 | 365 | 370 | |
| GAC ATT AAT GAT TGG AAC CCT GCC ACA GAC AAA TGT ATC CCC TGT CAT | | | 1008 |
| Asp Ile Asn Asp Trp Asn Pro Ala Thr Asp Lys Cys Ile Pro Cys His | | | |
| 375 | 380 | 385 | |
| TAT TCT GTT GAT GAC CTC TCT GGA AAG GCC AAA TGT AAA GGT GCA TTG | | | 1056 |
| Tyr Ser Val Asp Asp Leu Ser Gly Lys Ala Lys Cys Lys Gly Ala Leu | | | |
| 390 | 395 | 400 | |
| CAG AAG GAG CTG GGT TTA CCT ATA AGG CCT GAT GTT CCT CTG ATT GGC | | | 1104 |
| Gln Lys Glu Leu Gly Leu Pro Ile Arg Pro Asp Val Pro Leu Ile Gly | | | |
| 405 | 410 | 415 | |
| TTT ATT GGA AGG TTG GAT TAT CAG AAA GGC ATT GAT CTC ATT CAA CTT | | | 1152 |
| Phe Ile Gly Arg Leu Asp Tyr Gln Lys Gly Ile Asp Leu Ile Gln Leu | | | |
| 420 | 425 | 430 | 435 |
| ATC ATA CCA GAT CTC ATG CGG GAA GAT GTT CAA TTT GTC ATG CTT GGA | | | 1200 |
| Ile Ile Pro Asp Leu Met Arg Glu Asp Val Gln Phe Val Met Leu Gly | | | |
| 440 | 445 | 450 | |
| TCT GGT GAC CCA GAG CTT GAA GAT TGG ATG AGA TCT ACA GAG TCG ATC | | | 1248 |
| Ser Gly Asp Pro Glu Leu Glu Asp Trp Met Arg Ser Thr Glu Ser Ile | | | |
| 455 | 460 | 465 | |

| | |
|---|------|
| TTC AAG GAT AAA TTT CGT GGA TGG GTT GGA TTT AGT GTT CCA GTT TCC Phe Lys Asp Lys Phe Arg Gly Trp Val Gly Phe Ser Val Pro Val Ser 470 475 480 | 1296 |
| CAC CGA ATA ACT GCC GGC TGC GAT ATA TTG TTA ATG CCA TCC AGA TTC His Arg Ile Thr Ala Gly Cys Asp Ile Leu Leu Met Pro Ser Arg Phe 485 490 495 | 1344 |
| GAA CCT TGT GGT CTC AAT CAG CTA TAT GCT ATG CAG TAT GGC ACA GTT Glu Pro Cys Gly Leu Asn Gln Leu Tyr Ala Met Gln Tyr Gly Thr Val 500 505 510 515 | 1392 |
| CCT GTT GTC CAT GCA ACT GGG GGC CTT AGA GAT ACC GTG GAG AAC TTC Pro Val Val His Ala Thr Gly Gly Leu Arg Asp Thr Val Glu Asn Phe 520 525 530 | 1440 |
| AAC CCT TTC GGT CAG AAT GGA GAG CAG GGT ACA GGG TGG GCA TTC GCA Asn Pro Phe Gly Glu Asn Gly Glu Gln Gly Thr Gly Trp Ala Phe Ala 535 540 545 | 1488 |
| CCC CTA ACC ACA CAA AAC ATG TTT GTG GAC ATT GCG AAC TGC AAT ATC Pro Leu Thr Thr Glu Asn Met Phe Val Asp Ile Ala Asn Cys Asn Ile 550 555 560 | 1536 |
| TAC ATA CAG GGA ACA CAA GTC CTC CTG GGA AGG GCT AAT GAA GCG AGG Tyr Ile Gln Gly Thr Gln Val Leu Leu Gly Arg Ala Asn Glu Ala Arg 565 570 575 | 1584 |
| CAT GTC AAA ACA CTT CAC GTG GGA CCA TGC CGC TGA His Val Lys Arg Leu His Val Gly Pro Cys Arg * 580 585 590 | 1620 |

2) INFORMATION ZU SEQ ID NO: 21:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 540 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

Cys Val Ala Glu Leu Ser Arg Glu Asp Leu Gly Leu Glu Pro Glu Gly

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

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

| | | | |
|-----|-----|-----|-----|
| | 485 | 490 | 495 |
| Pro | Leu | Thr | Thr |
| Glu | Asn | Met | Phe |
| Val | Asp | Ile | Ala |
| Asn | Cys | Asn | Ile |
| 500 | 505 | 510 | |
| Tyr | Ile | Gln | Gly |
| Thr | Gln | Val | Leu |
| Leu | Gly | Arg | Ala |
| Asn | Glu | Ala | Arg |
| 515 | 520 | 525 | |
| His | Val | Lys | Arg |
| Leu | His | Val | Gly |
| Pro | Cys | Arg | * |
| 530 | 535 | 540 | |

(2) INFORMATION ZU SEQ ID NO: 22:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 30 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: sonstige Nukleinsäure

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

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 22:

GTGGATCCAT GCGGACGCCC TCGGCGGTGG

30

(2) INFORMATION ZU SEQ ID NO: 23:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 35 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: sonstige Nukleinsäure

(A) BESCHREIBUNG: / desc = "Oligonukleotide"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 23:

CTGAATTCCA TATGGGGCCC CTCCTGCTC AGCTC

35

(2) INFORMATION ZU SEQ ID NO: 24:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 36 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: sonstige Nukleinsäure

(A) BESCHREIBUNG: / desc = "Oligonukleotide"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 24:

CTCTGAGCTC AAGCTTGCTA CTTTCTTCC TTAATC

36

(2) INFORMATION ZU SEQ ID NO: 25:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 29 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: Linear

(ii) ART DES MOLEKÜLS: sonstige Nukleinsäure

(A) BESCHREIBUNG: / desc = "Oligonukleotide"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 25:

GTCTCGCGGG TGGTGTCCTT GCTTCCTAG

29

(2) INFORMATION ZU SEQ ID NO: 26:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 53 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Doppel

(D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA bis mRNA

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 26:

TGCGTGGCGG AGCTGAGCAG GGAGGTCTCC GCGGTGGTGT CCTTGCTTCC TAG

53

2) INFORMATION ZU SEQ ID NO: 27:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 8 Aminosäuren

(B) ART: Aminosäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 27:

Cys Val Ala Glu Leu Ser Arg Glu
1 5

(2) INFORMATION ZU SEQ ID NO: 28:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 16 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA bis mRNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 28:

AGAGAGAGAG AGAGAG

16

(2) INFORMATION ZU SEQ ID NO: 29:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 36 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA bis mRNA

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 29:

AAGAAGAAGA AGAAGAAGAA GAAGAAGAAG AAGAAG

36

(2) INFORMATION ZU SEQ ID NO: 30:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 18 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Doppel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: cDNA bis mRNA

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 30:

AAAAAAAAAA AAAAAAAAA

18

(2) INFORMATION ZU SEQ ID NO: 31:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 11 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: sonstige Nukleinsäure

(A) BESCHREIBUNG: / desc = "Oligonukleotide"

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 31:

AGATAATGCA G

11

(2) INFORMATION ZU SEQ ID NO: 32:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 10 Basenpaare

(B) ART: Nukleinsäure

(C) STRANGFORM: Einzel

(D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: sonstige Nukleinsäure

(A) BESCHREIBUNG: / desc = "Oligonukleotide"

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 32:

AACAATGGCT

10

2) INFORMATION ZU SEQ ID NO: 23:

(i) SEQUENZCHARAKTERISTIKA:

(A) LÄNGE: 56 Aminosäuren

- (B) ART: Aminosäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: Peptid

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 33:

```

Met Ala Ser Ser Met Leu Ser Ser Ala Ala Val Ala Thr Arg Thr Asn
1           5           10           15

Pro Ala Gln Ala Ser Met Val Ala Pro Phe Thr Gly Leu Lys Ser Ala
          20           25           30

Ala Phe Pro Val Ser Arg Lys Gln Asn Leu Asp Ile Thr Ser Ile Ala
          35           40           45

Ser Asn Gly Gly Arg Val Gln Cys
50           55
  
```

2) INFORMATION ZU SEQ ID NO: 34:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 58 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: Peptid

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 34

```

Met Ala Pro Thr Val Met Met Ala Ser Ser Ala Thr Ala Thr Arg Thr
1           5           10           15

Asn Pro Ala Gln Ala Ser Ala Val Ala Pro Phe Gln Gly Leu Lys Ser
          20          25          30

Thr Ala Ser Leu Pro Val Ala Arg Arg Ser Ser Arg Ser Leu Gly Asn
          35          40          45

Val Ala Ser Asn Gly Gly Arg Ile Arg Cys
          50          55

```

2) INFORMATION ZU SEQ ID NO: 35:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 58 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: Peptid

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 35

```

Met Ala Gln Ile Leu Ala Pro Ser Thr Gln Trp Gln Met Arg Ile Thr
1           5           10           15

Lys Thr Ser Pro Cys Ala Thr Pro Ile Thr Ser Lys Met Trp Ser Ser
          20          25          30

Leu Val Met Lys Gln Thr Lys Lys Val Ala His Ser Ala Lys Phe Arg
          35          40          45

Val Met Ala Val Asn Ser Glu Asn Gly Thr
          50          55

```

2) INFORMATION ZU SEQ ID NO: 36:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 74 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: Peptid

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 36

```

Met Ala Ala Leu Ala Thr Ser Gln Leu Val Ala Thr Arg Ala Gly His
1           5           10           15

Gly Val Pro Asp Ala Ser Thr Phe Arg Arg Gly Ala Ala Gln Gly Leu
          20           25           30

Arg Gly Ala Arg Ala Ser Ala Ala Ala Asp Thr Leu Ser Met Arg Thr
          35           40           45

Ser Ala Arg Ala Ala Pro Arg His Gln Gln Gln Ala Arg Arg Gly Gly
          50           55           60

Arg Phe Pro Phe Pro Ser Leu Val Val Cys
65           70

```

2) INFORMATION ZU SEQ ID NO: 37:

(i) SEQUENZCHARAKTERISTIKA:

- (A) LÄNGE: 39 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: irrelevant

(ii) ART DES MOLEKÜLS: Peptid

(iii) HYPOTHETISCH: NEIN

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 37

```
Met Ala Thr Pro Ser Ala Val Gly Ala Ala Cys Leu Leu Leu Ala Arg
1           5           10           15

Xua Ala Trp Pro Ala Ala Val Gly Asp Arg Ala Arg Pro Arg Arg Leu
          20          25          30

Gln Arg Val Leu Arg Arg Arg
          35
```

Patentansprüche

1. Rekombinantes Nukleinsäurekonstrukt, das folgendes umfaßt:
 - a) einen Promoter, der dahingehend adaptiert ist, daß er die Expression eines interessierenden Polypeptids in einem stärkehaltigen Gewebe einer Pflanze während der Stärkesynthese beeinflusst.
 - b) eine Nukleinsäure, die für ein Transitpeptid codiert, das das interessierende Polypeptid zu einem Amyloplasten zu transportieren vermag;
 - c) eine Nukleinsäure, die für eine stärkebindende Region codiert;
 - d) eine Nukleinsäure, die für das interessierende Polypeptid codiert; sowie
 - e) eine Terminatorsequenz;
 wobei das Konstrukt die Expression eines Hybridpolypeptids, das die stärkebindende Region und das interessierende Polypeptid umfaßt, steuert, wobei das Hybridpolypeptid innerhalb von Stärke der Pflanze eingekapselt ist.
2. Konstrukt nach Anspruch 1, wobei Promoter und Terminator an die Expression des Hybridpolypeptids innerhalb von Stärke einer monokotylen Pflanze adaptiert sind.
3. Konstrukt nach Anspruch 1, wobei Promoter und Terminator an die Expression des Hybridpolypeptids innerhalb von Stärke einer dikotylen Pflanze adaptiert sind.
4. Expressionsvektor, der das Konstrukt nach Anspruch 1 umfaßt.
5. Pflanzenzelle, die dahingehend transformiert ist, daß sie das Konstrukt nach Anspruch 1 umfaßt, wobei die Zelle das in Amyloplasten der Zelle eingekapselte Hybridpolypeptid zu exprimieren vermag.
6. Pflanze, die aus der Zelle nach Anspruch 5 regeneriert worden ist.
7. Same der Pflanze nach Anspruch 6, wobei der Same das Konstrukt nach einem der Ansprüche 1 bis 3 umfaßt und das in Stärke dieses Samens eingekapselte Hybridpolypeptid zu exprimieren vermag.
8. Modifizierte Stärke aus dem Samen nach Anspruch 7, wobei die Stärke das Hybridpolypeptid umfaßt.
9. Modifizierte Stärke nach Anspruch 8 zur Verwendung als Arzneimittel.
10. Verwendung der modifizierten Stärke nach Anspruch 8 in der Herstellung eines Nahrungsmittels oder eines Tierfutters.
11. Nahrungsmittel oder Tierfutter, das die modifizierte Stärke nach Anspruch 8 umfaßt.
12. Verfahren zur Herstellung eines reinen interessierenden Polypeptids, wobei das Verfahren die folgen-

den Schritte umfaßt:

- a) Transformieren einer Pflanzenzelle mit dem Konstrukt nach Anspruch 1;
- b) Exprimierenlassen des Hybridpolypeptids in der Pflanzenzelle;
- c) Isolieren des Hybridpolypeptids aus der Zelle; sowie
- d) Aufreinigen des interessierenden Polypeptids aus dem Hybridpolypeptid.

13. Hybridpolypeptid, das von dem Konstrukt nach Anspruch 1, 2 oder 3 gebildet wird.

14. Hybridpolypeptid nach Anspruch 13, das weiterhin zwischen der stärkebindenden Region und dem interessierenden Polypeptid eine Spaltstelle umfaßt.

15. Hybridpolypeptid nach Anspruch 13 oder 14, wobei das interessierende Polypeptid aus der Gruppe bestehend aus Hormonen, Wachstumsfaktoren, Antikörpern, Peptiden, Polypeptiden, Enzymen und Immunglobulinen stammt.

16. Hybridpolypeptid nach Anspruch 15, wobei das interessierende Polypeptid aus der Gruppe bestehend aus Somatotropin, Insulin A, Insulin B, Calcitonin, beta-Endorphin, Urogatron, beta-Globin, Myoglobin, menschliches Wachstumshormon, Angiotensin, Prolactin, einer Protease, einer beta-Galactosidase sowie einer Cellulase stammt.

17. Hybridpolypeptid nach einem der Ansprüche 13 bis 16, wobei die stärkebindende Region von einem stärkebindenden Enzym abstammt.

18. Hybridpolypeptid nach Anspruch 17, wobei das stärkebindende Enzym aus der Gruppe bestehend aus einer Stärkesynthase, einem Stärkeverzweigungsenzym und einer Glycogensynthase stammt.

19. Hybridpolypeptid nach Anspruch 18, wobei das stärkebindende Enzym aus der Gruppe bestehend aus der löslichen Stärkesynthase I, der löslichen Stärkesynthase II, der löslichen Stärkesynthase III, der stärke-korngebundenen Stärkesynthase, dem Verzweigungsenzym I, dem Verzweigungsenzym IIa und dem Verzweigungsenzym IIb stammt.

20. Hybridpolypeptid nach einem der Ansprüche 13 bis 19, wobei es sich bei der stärkebindenden Region um einen carboxyterminalen Abschnitt des stärkebindenden Enzyms handelt.

21. Stärkekorn, das das Hybridpolypeptid nach einem der Ansprüche 13 bis 20 enthält.

22. Stärkekorn, das ein von dem Konstrukt nach einem der Ansprüche 1 bis 3 codiertes Hybridpolypeptid enthält.

23. Stärkehaltiges Samenkorn, das das Stärkekorn nach Anspruch 21 oder 22 umfaßt.

24. Stärkehaltiges Samenkorn nach Anspruch 23, das weiterhin einen Embryo und Nährgewebe umfaßt.

Es folgen 12 Blatt Zeichnungen

Anhängende Zeichnungen

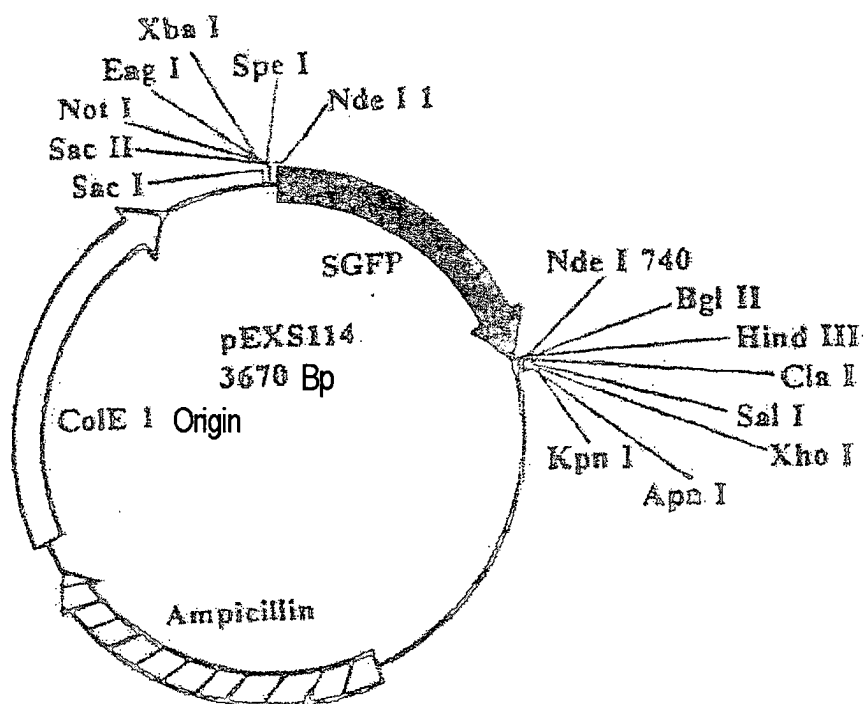


FIG. 1A

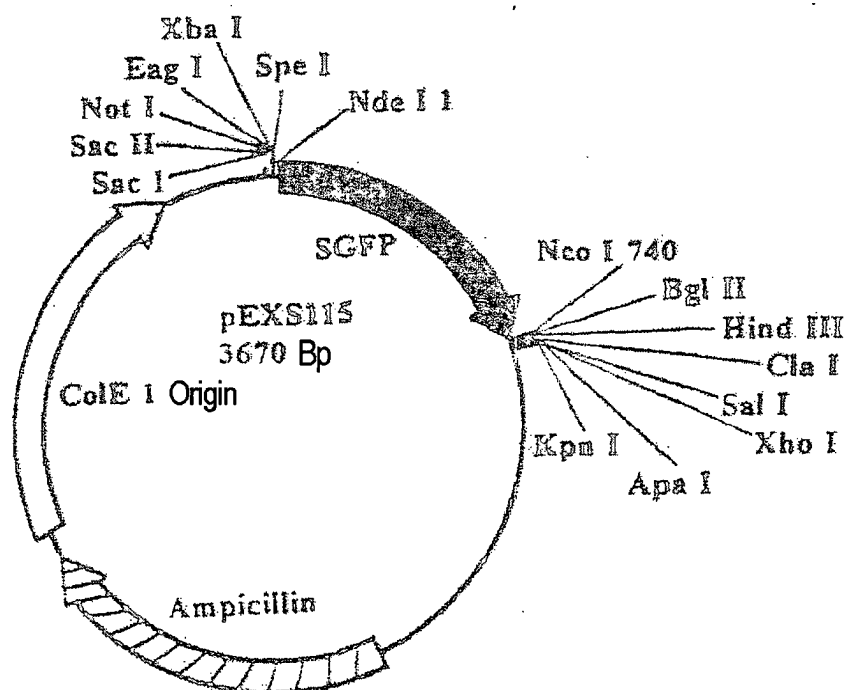


FIG. 1B

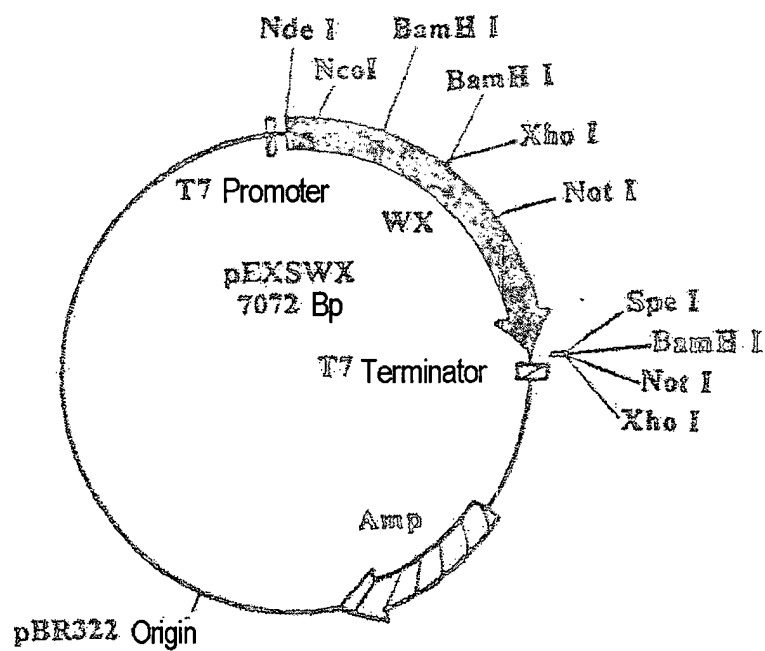


FIG. 2A

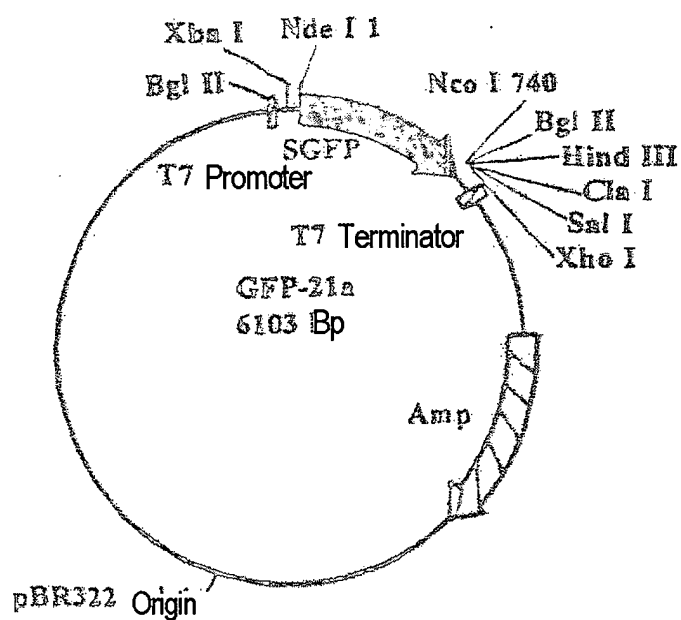


FIG. 2B

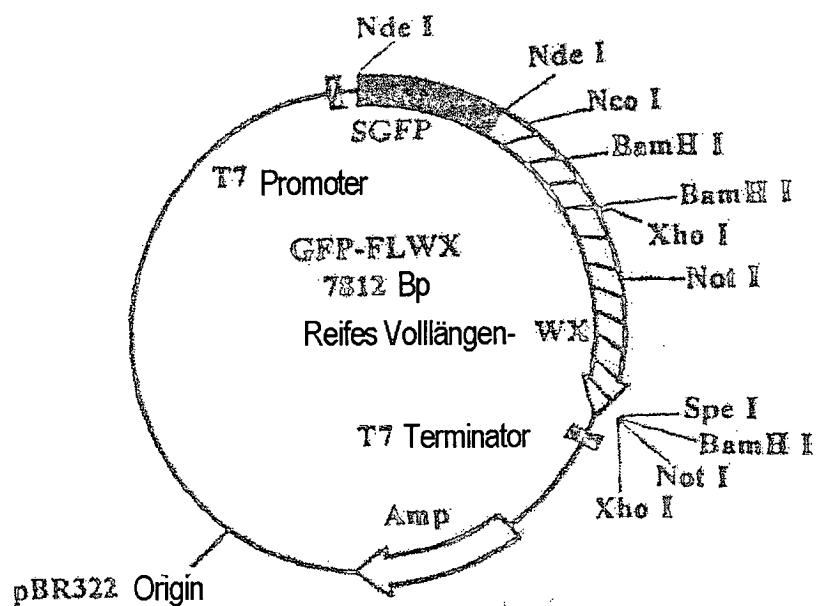


FIG. 3A

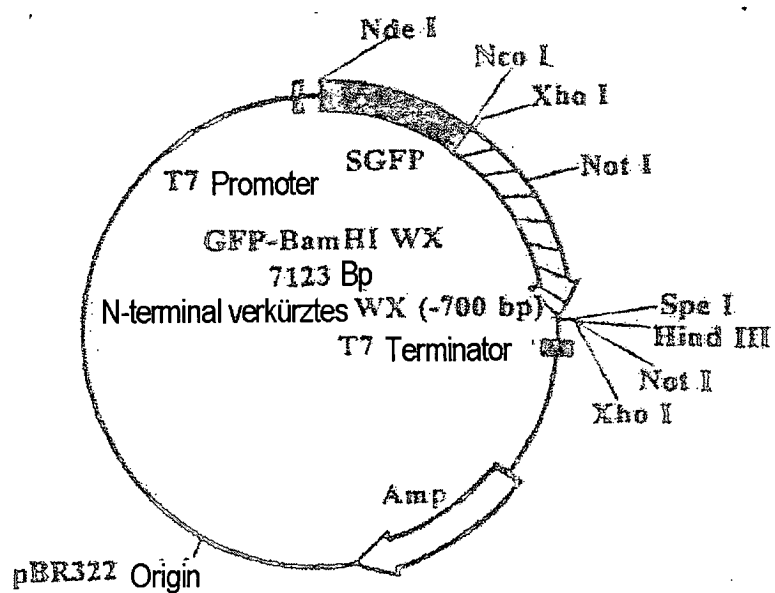


FIG. 3B

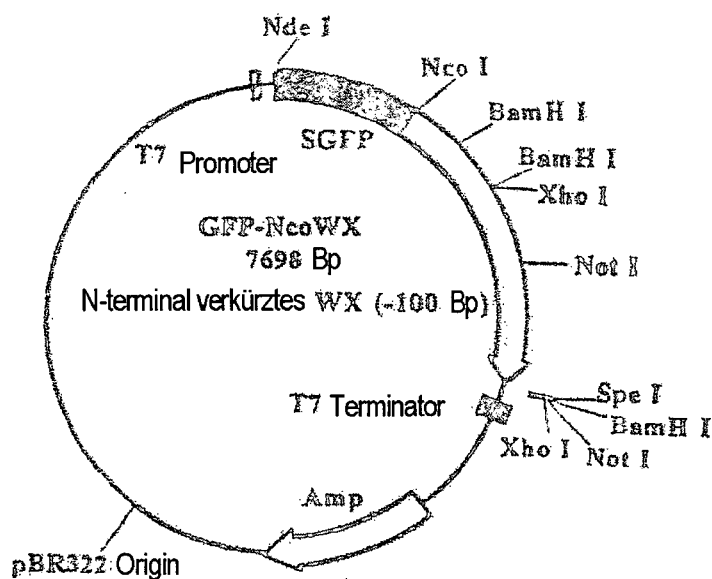


FIG. 4

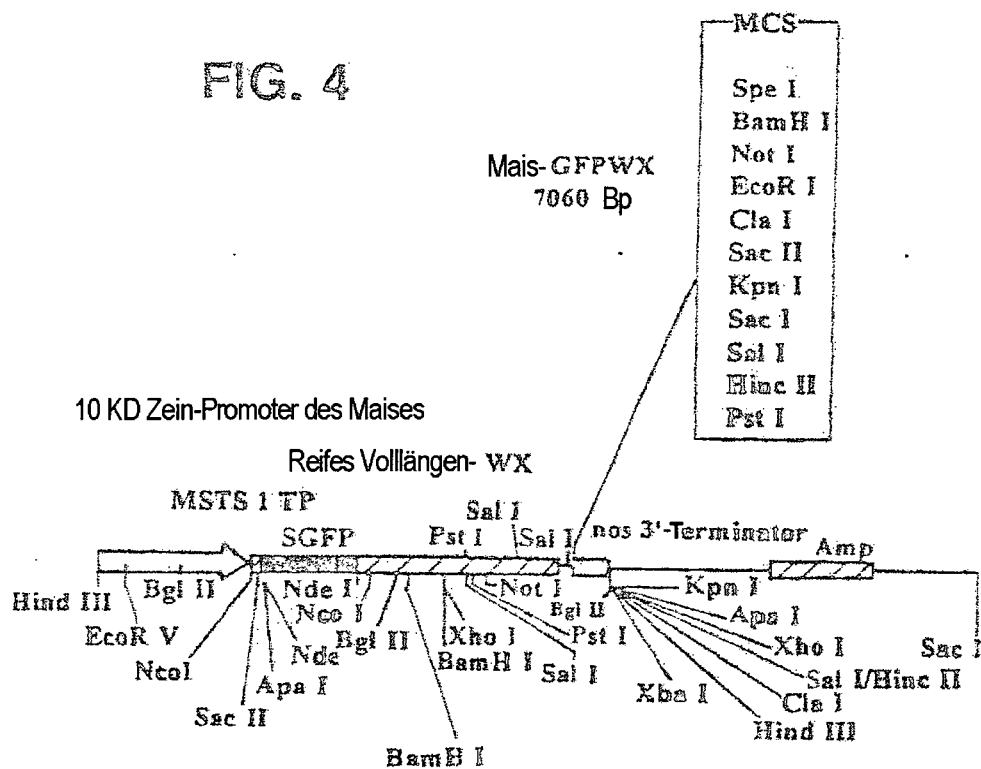


FIG. 5

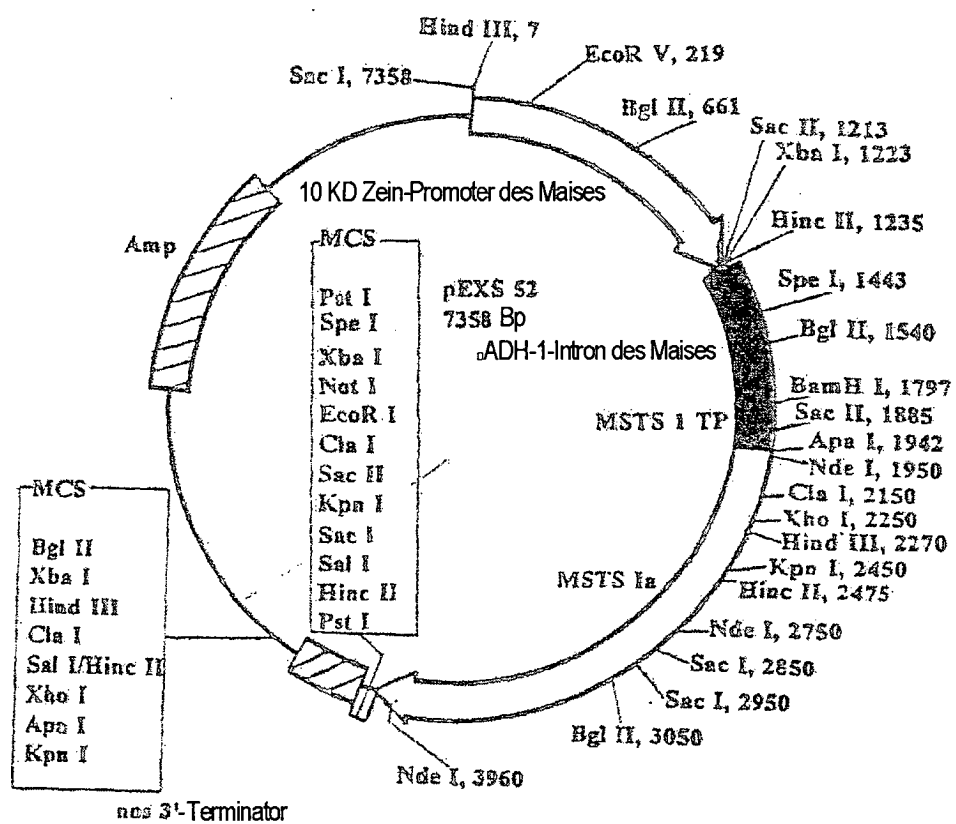


FIG. 6

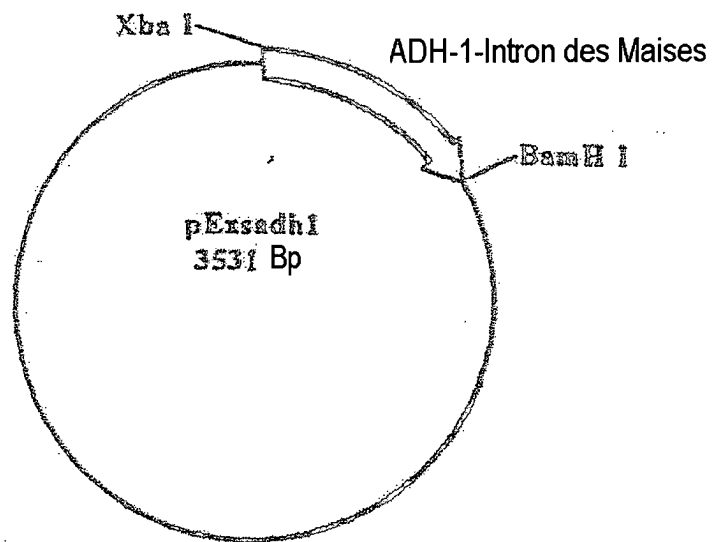


FIG. 7A

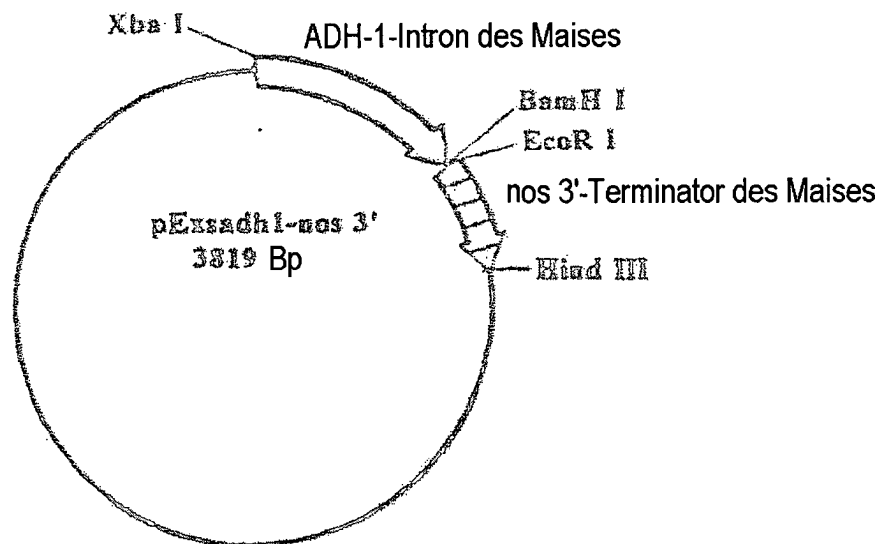


FIG. 7B

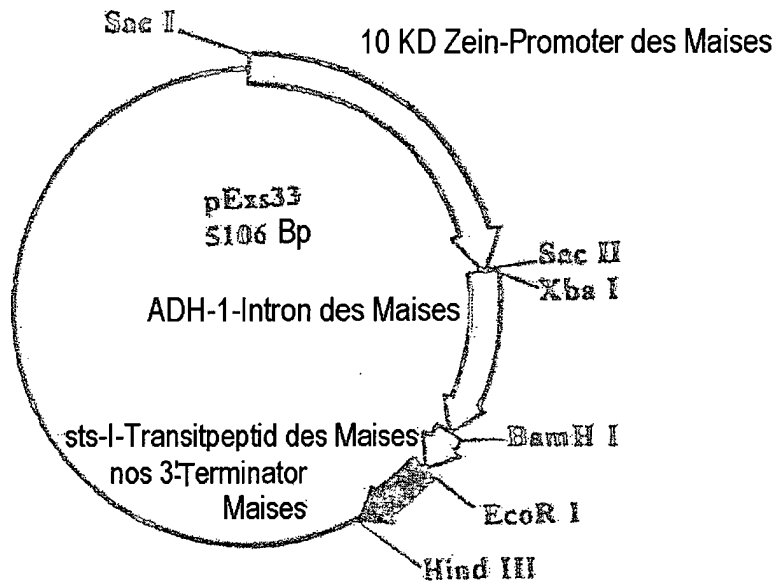


FIG. 7C

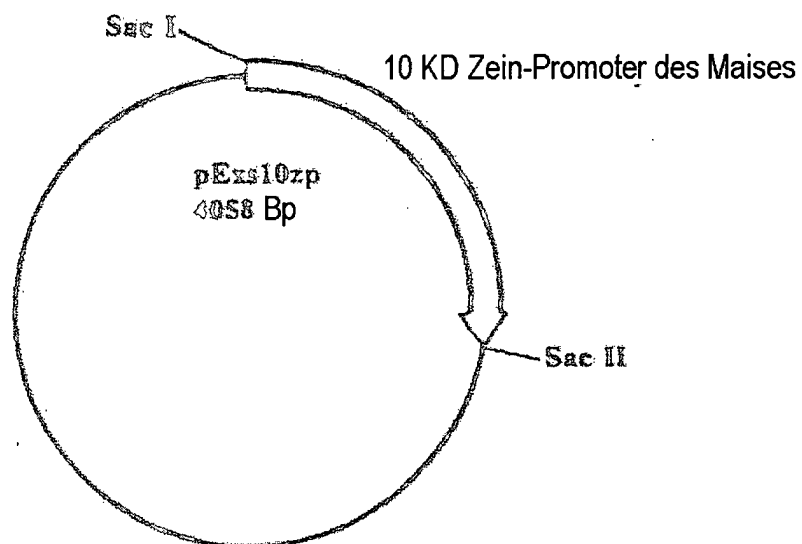


FIG. 7D

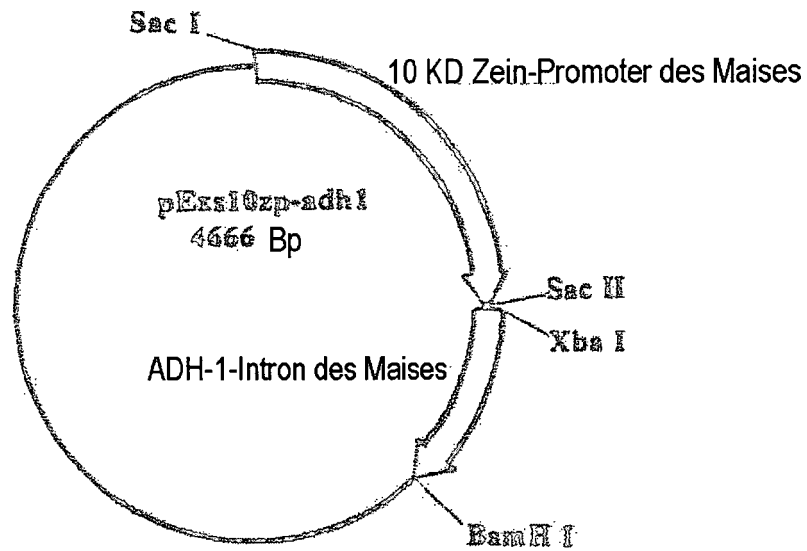


FIG. 7E

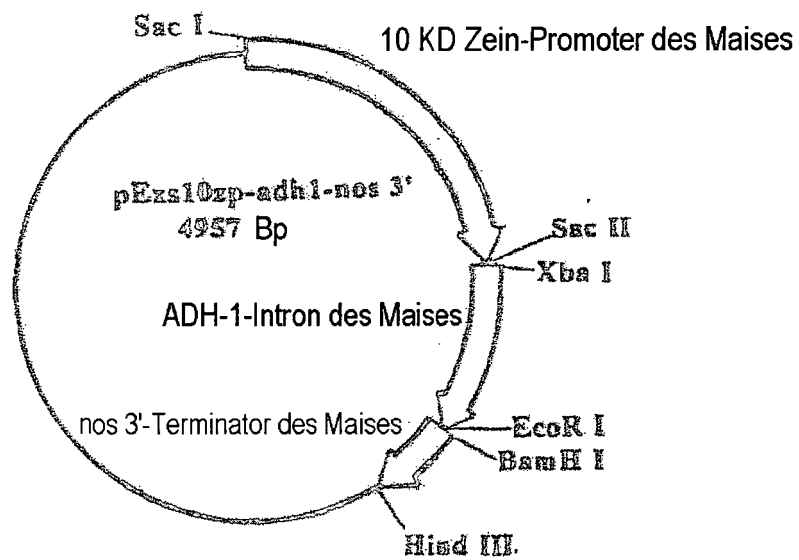


FIG. 7F

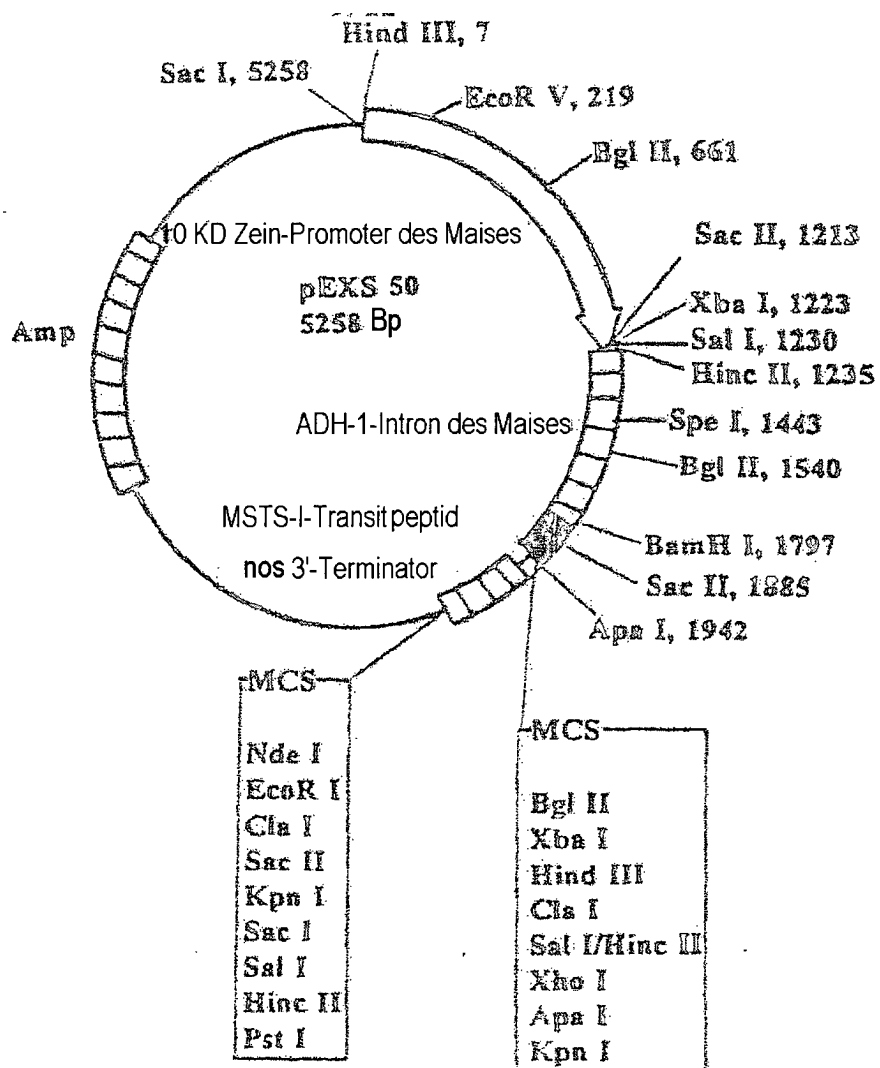


FIG. 8A

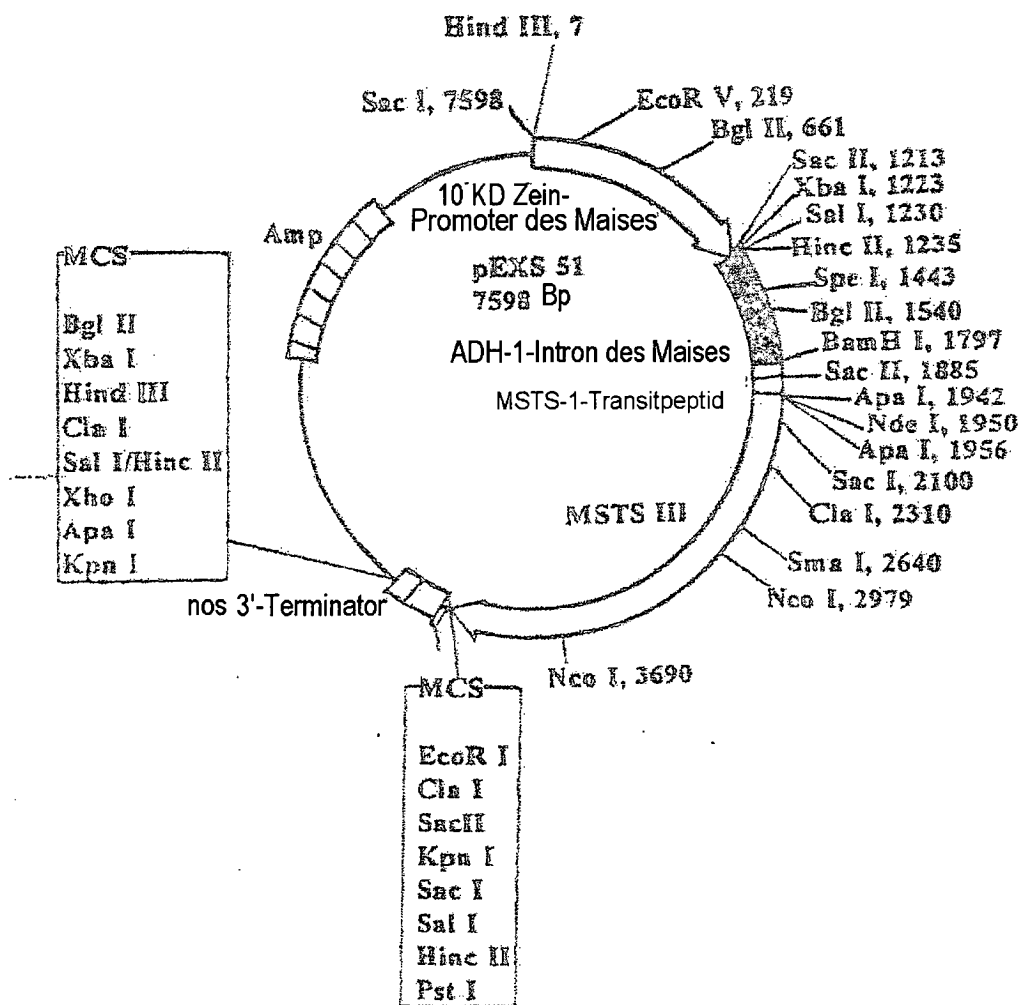


FIG. 8B

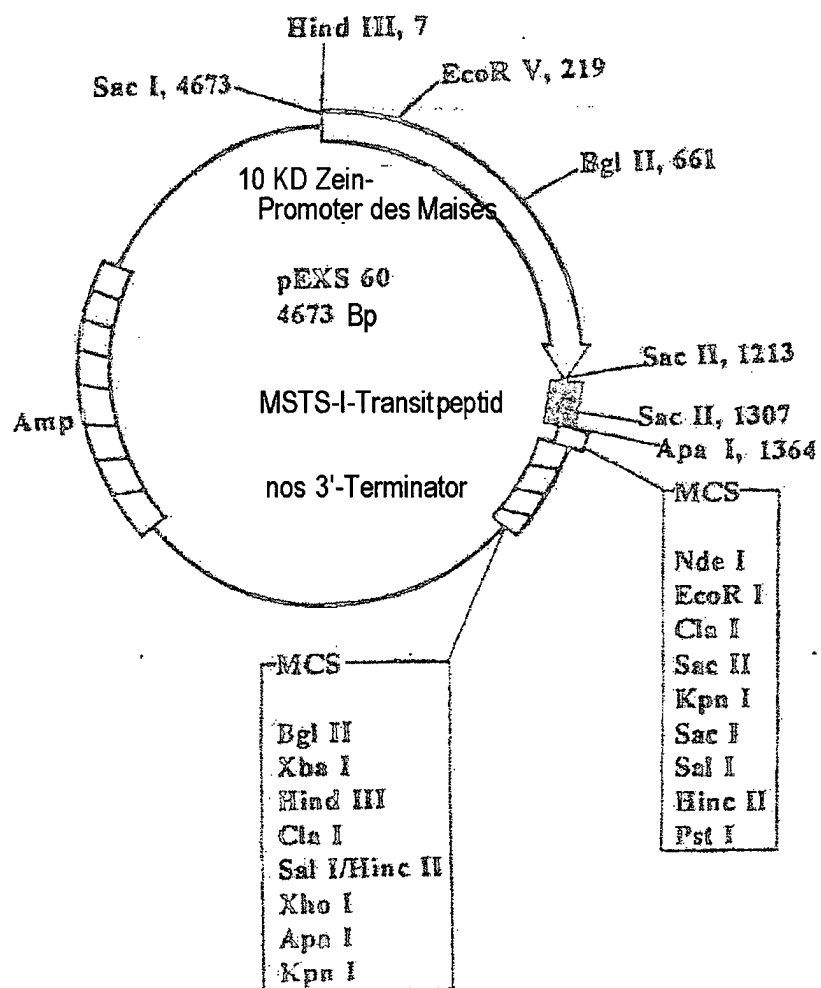


FIG. 9A

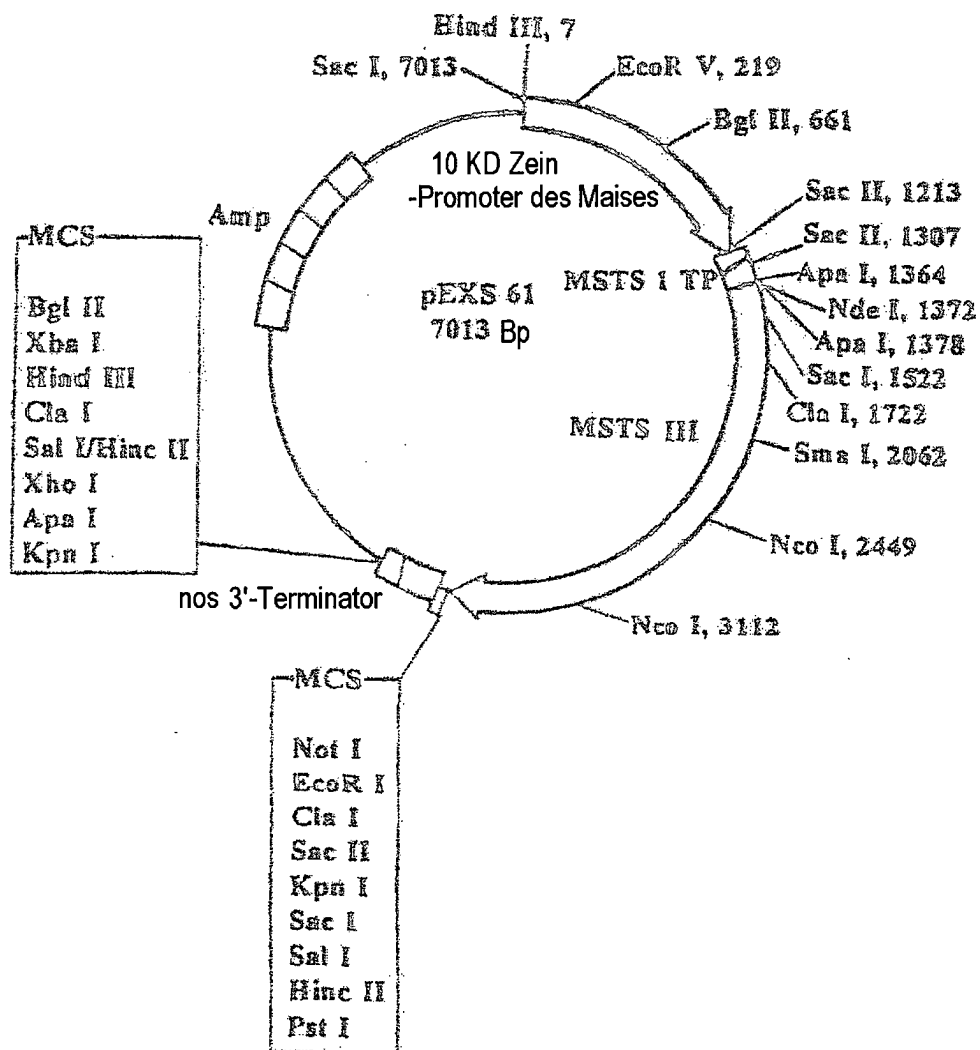


FIG. 9B