



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2020년03월23일

(11) 등록번호 10-2092429

(24) 등록일자 2020년03월17일

(51) 국제특허분류(Int. Cl.)  
*C12P 19/02* (2006.01) *C12P 19/14* (2006.01)  
*C12P 7/10* (2006.01) *C12P 7/14* (2017.01)  
(52) CPC특허분류  
*C12N 9/2445* (2013.01)  
*C12N 9/2437* (2013.01)  
(21) 출원번호 10-2015-7008688  
(22) 출원일자(국제) 2013년09월04일  
심사청구일자 2018년08월21일  
(85) 번역문제출일자 2015년04월03일  
(65) 공개번호 10-2015-0058274  
(43) 공개일자 2015년05월28일  
(86) 국제출원번호 PCT/FR2013/052036  
(87) 국제공개번호 WO 2014/037667  
국제공개일자 2014년03월13일  
(30) 우선권주장  
12 58260 2012년09월05일 프랑스(FR)  
(56) 선행기술조사문헌  
US04472504 A  
W01992010581 A1  
W02010029259 A1  
W02010148148 A2

(73) 특허권자  
**아이에프피 에너지스 누벨**  
프랑스 루이-말메종 세데 92852 아브뉴 드 브와  
쁘레오 1 & 4  
**상뜨르 나쇼날 드 라 리셰르쉬 상띠피끄(씨엔알에스)**  
프랑스 에프-75794 파리 세텍스 튀 미셀 앙즈 3  
(72) 발명자  
**마르조 앙뚜안**  
프랑스 에프-75018 파리 튀 마르가데 212  
**마띠 위그**  
프랑스 에프-77600 뷔시 생 조르쥬 뿔라스 뿔강스  
비앵베뉴 30  
(뒷면에 계속)  
(74) 대리인  
**특허법인코리아나**

전체 청구항 수 : 총 21 항

심사관 : 김정아

(54) 발명의 명칭 **저온에서 강화된 베타-글루코시다아제 활성을 갖는 폴리펩티드**

### (57) 요약

본 발명은 약 30℃ 내지 약 35℃ 의 온도에서 향상된 베타-글루코시다아제 활성을 갖는 폴리펩티드에 관한 것이다.

(52) CPC특허분류

**C12P 19/02** (2013.01)

**C12P 19/14** (2013.01)

**C12P 7/10** (2013.01)

**C12P 7/14** (2013.01)

**C12Y 302/01021** (2013.01)

**C12P 2203/00** (2013.01)

**Y02E 50/16** (2013.01)

**Y02P 20/52** (2015.11)

(72) 발명자

**아이리나 셸런**

프랑스 에프-30350 도메사르쥬 로띠스명 르 꼬또 7

**윌망 크리스토프**

프랑스 에프-30000 님 뤼 데파르시외 5

**페르실롱 세실**

프랑스 에프-30000 님 뤼 드 라 콩페스 3

**포르 세바스티앙**

프랑스 에프-38410 위리아쥬 슈맹 뒤 파르끄 55

**아르망 실비**

프랑스 에프-38100 그르노블 뤼 마르퀴앙 23

**프띠 모**

프랑스 에프-38700 라 트롱쉬 뤼 마리 블레 4

## 명세서

### 청구범위

#### 청구항 1

아미노산 서열 SEQ ID No. 1 의 베타-글루코시다아제 활성을 갖는 폴리펩티드.

#### 청구항 2

제 1 항에 있어서, 서열 SEQ ID No. 3 의 야생형 BGL1 단백질의 베타-글루코시다아제 활성과 비교해, 30℃ 내지 35℃ 를 포함하는 온도에서 향상된 베타-글루코시다아제 활성을 갖는 것을 특징으로 하는 폴리펩티드.

#### 청구항 3

제 1 항에 있어서, 아미노산 서열 SEQ ID No. 5 의 100B11 폴리펩티드의 베타-글루코시다아제 활성과 비교해, 30℃ 내지 35℃ 를 포함하는 온도에서 향상된 베타-글루코시다아제 활성을 갖는 것을 특징으로 하는 폴리펩티드.

#### 청구항 4

제 3 항에 있어서, 아미노산 서열 SEQ ID No. 5 의 100B11 폴리펩티드의 베타-글루코시다아제 활성과 비교해, 30℃ 내지 35℃ 를 포함하는 온도에서 10% 이상, 또는 20% 이상, 또는 30% 이상, 또는 40% 이상 향상된 베타-글루코시다아제 활성을 갖는 것을 특징으로 하는 폴리펩티드.

#### 청구항 5

제 1 항에 정의된 바와 같은 폴리펩티드를 엔코딩하는 것을 특징으로 하는, 정제된 또는 분리된 핵산.

#### 청구항 6

제 5 항에 있어서, 핵산 서열 SEQ ID No. 2 를 포함하는 것을 특징으로 하는 핵산.

#### 청구항 7

제 5 항에 따른 핵산을 포함하는 것을 특징으로 하는 벡터.

#### 청구항 8

제 1 항 내지 제 4 항 중 어느 한 항에 따른 폴리펩티드, 또는 제 5 항에 따른 핵산 또는 제 7 항에 따른 벡터를 포함하는 것을 특징으로 하는 분리된 숙주 세포.

#### 청구항 9

제 8 항에 있어서, 트리코데르마 (*Trichoderma*), 아스페르길루스 (*Aspergillus*), 뉴로스포라 (*Neurospora*), 휴미콜라 (*Humicola*), 페니실리움 (*Penicillium*), 푸사리움 (*Fusarium*), 테르모모노스포라 (*Thermomonospora*), 마이셀리오프토라 (*Myceliophthora*), 크리소스포리움 (*Chrysosporium*), 바실러스 (*Bacillus*), 슈도모나스 (*Pseudomonas*), 에스케리차 (*Escherichia*), 클로스트리듐 (*Clostridium*), 셀룰로모나스 (*Cellulomonas*), 스트렙토마이세스 (*Streptomyces*), 야로위아 (*Yarrowia*), 피치아 (*Pichia*) 및 사카로마이세스 (*Saccharomyces*) 중에서 선택되는 것을 특징으로 하는 분리된 숙주 세포.

#### 청구항 10

제 8 항에 있어서, 트리코데르마 레에세이 (*Trichoderma reesei*), 트리코데르마 비리다에 (*Trichoderma viridae*), 트리코데르마 코닝기이 (*Trichoderma koningii*), 아스페르길루스 니게르 (*Aspergillus niger*), 아스페르길루스 니둘란스 (*Aspergillus nidulans*), 아스페르길루스 웬티이 (*Aspergillus wentii*), 아스페르길루스 오리자에 (*Aspergillus oryzae*), 아스페르길루스 포에니시스 (*Aspergillus phoenicis*), 뉴로스포라 크라사 (*Neurospora crassa*), 휴미콜라 그리사에 (*Humicola grisae*), 마이셀리오프토라 테르모필라 (*Myceliophthora thermopila*), 크리소스포리움 룩노웬세 (*Chrysosporium lucknowense*), 페니실리움 피노필름 (*Penicillium*

pinophilum), 페니실리움 옥살리쿰 (Penicillium oxalicum), 에스케리차 콜라이 (Escherichia coli), 클로스트리듐 아세토부틸리쿰 (Clostridium acetobutylicum), 클로스트리듐 사카롤라이티쿰 (Clostridium saccharolyticum), 클로스트리듐 벤제린키이 (Clostridium benjerinckii), 클로스트리듐 부틸리쿰 (Clostridium butylicum), 피치아 파스토리스 (Pichia pastoris), 야로위아 리폴리티카 (Yarrowia lipolityca), 사카로마이세스 세레비지아에 (Saccharomyces cerevisiae), 및 이의 혼합물 중에서 선택되는 것을 특징으로 하는 단리된 숙주 세포.

#### 청구항 11

제 9 항에 있어서, 종 트리코데르마 레에세이 (Trichoderma reesei) 의 것인 것을 특징으로 하는 단리된 숙주 세포.

#### 청구항 12

제 9 항에 있어서, 종 사카로마이세스 세레비지아에 (Saccharomyces cerevisiae) 의 것인 것을 특징으로 하는 단리된 숙주 세포.

#### 청구항 13

제 1 항 내지 제 4 항 중 어느 한 항에 따른 폴리펩티드의 존재 하에서의 가수분해 단계를 포함하는, 베타-올리고당의 가수분해를 위한 공정.

#### 청구항 14

제 8 항에 따른 세포의 존재 하에서의 가수분해 단계를 포함하는, 베타-올리고당의 가수분해를 위한 공정.

#### 청구항 15

제 1 항 내지 제 4 항 중 어느 한 항에 따른 폴리펩티드의 존재 하에서의 가수분해 단계를 포함하는, 셀로비오스의 글루코오스로의 가수분해를 위한 공정.

#### 청구항 16

제 8 항에 따른 세포의 존재 하에서의 가수분해 단계를 포함하는, 셀로비오스의 글루코오스로의 가수분해를 위한 공정.

#### 청구항 17

제 1 항 내지 제 4 항 중 어느 한 항에 따른 폴리펩티드의 존재 하에서의 가수분해 단계를 포함하는, 바이오연료의 제조를 위한 공정.

#### 청구항 18

제 8 항에 따른 세포의 존재 하에서의 가수분해 단계를 포함하는, 바이오연료의 제조를 위한 공정.

#### 청구항 19

리그노셀룰로오스 바이오매스 상에 작용하는 효소 조성물로서, 사상 진균에 의해 제조되고 제 1 항 내지 제 4 항 중 어느 한 항에 따른 하나 이상의 폴리펩티드를 포함하는 효소 조성물.

#### 청구항 20

하기 단계:

- 수성상 내에, 가수분해하고자 하는 물질의 현탁;
- 글루코오스를 함유하는 가수분해물을 생성하도록, 제 19 항에 따른 효소 조성물의 존재 하에, 리그노셀룰로오스 바이오매스의 가수분해;
- 발효 포도액을 생성하도록, 가수분해물의 글루코오스의 발효;
- 발효 포도액으로부터 바이오연료의 분리

를 포함하고,

가수분해 및 발효 단계가 동시에 실시되는 것을 특징으로 하는, 바이오매스로부터 바이오연료의 제조 방법.

## 청구항 21

하기 단계:

- 수성상 내에, 가수분해하고자 하는 물질의 현탁;
- 글루코오스를 함유하는 가수분해물을 생성하도록, 제 8 항에 따른 세포의 존재 하에, 리그노셀룰로오스 바이오매스의 가수분해;
- 발효 포도액을 생성하도록, 가수분해물의 글루코오스의 발효;
- 발효 포도액으로부터 바이오연료의 분리

를 포함하고,

가수분해 및 발효 단계가 동시에 실시되는 것을 특징으로 하는, 바이오매스로부터 바이오연료의 제조 방법.

## 발명의 설명

### 기술 분야

- [0001] 셀룰로오스로부터 에탄올을 제조하는 가능성은 다량의 원료의 이용능 및 또한 연료로서의 에탄올에 대한 관심 덕분에 큰 주목을 받아왔다. 이러한 방법을 위한 셀룰로오스-기반 천연 원료는 "바이오매스" 로 표시된다. 수많은 유형의 바이오매스, 예를 들어 목재, 농경 잔류물, 초본 작물 및 고체 도시 폐기물은 바이오연료를 제조하기 위한 잠재적인 원료로서 고려되어 왔다. 상기 물질은 주로 셀룰로오스, 헤미셀룰로오스 및 리그닌으로 이루어진다.

### 배경 기술

- [0002] 셀룰로오스는 베타-1,4 결합에 의해 연결된 글루코오스 분자로 이루어진 중합체이며, 이들은 분해 또는 해중합에 대해 매우 저항성이 있다. 일단 셀룰로오스가 글루코오스로 전환되면, 글루코오스는 효모를 사용하여 바이오연료, 예를 들어 에탄올로 쉽게 발효된다.
- [0003] 셀룰로오스를 글루코오스로 전환하기 위해 연구된 가장 오래된 방법은 산 가수분해에 기반하는 것이다. 상기 방법은 농축 또는 희석 산의 존재 하에 실시될 수 있다. 그러나, 여러가지 단점, 예컨대 농축 산이 사용되는 경우 산의 열악한 회수 및 희석 산을 사용하는 경우 글루코오스의 적은 제조가, 산 가수분해 방법의 경제학에는 악영향을 준다.
- [0004] 산 가수분해 방법의 단점을 극복하기 위해, 셀룰로오스 전환 방법은 더욱 최근에는 셀룰라아제 유형의 효소를 사용하는, 효소 가수분해와 관련된다. 리그노셀룰로오스 바이오매스 (예를 들어, 셀룰로오스) 의 상기 효소 가수분해는 그러나 값비싼 산업적 공정이라는 단점을 갖는다. 따라서, 점점 더 효과적인 셀룰라아제-분비 미생물 균주를 사용하는 것이 필수적이다. 이와 관련하여, 많은 미생물, 예컨대 진균 트리코데르마 (Trichoderma), 아스페르길루스 (Aspergillus), 휴미콜라 (Humicola) 또는 푸사리움 (Fusarium), 및 또한 박테리아, 예컨대 테르모모노스포라 (Thermomonospora), 바실러스 (Bacillus), 셀룰로모나스 (Cellulomonas) 및 스트렙토마이세스 (Streptomyces) 는 셀룰로오스를 가수분해하는 효소를 포함한다. 상기 미생물에 의해 분비되는 효소는 셀룰로오스의 글루코오스로의 전환에 사용되는 3 가지 유형의 활성을 갖고, 하기 3 가지 그룹으로 나뉜다: 셀룰로오스 섬유를 내부에서 랜덤으로 공격하는 엔도글루카나아제, 섬유의 말단을 공격하여 셀로비오스를 방출시킬 엑소글루카나아제, 및 상기 셀로비오스를 글루코오스로 가수분해할 베타-글루코시다아제. 베타-글루코시다아제는 셀룰로오스 전환 공정의 제한 단계를 구성한다. 게다가, 공정의 첫번째 어려움은, 공정의 종료시 가수분해되지 않는 임의의 셀로비오스가 바이오연료의 제조 동안 수율 상실을 나타내므로, 셀로비오스의 글루코오스로의 전환에 놓여 있다.
- [0005] 셀로비오스의 상기 축적은, 트리코데르마 (Trichoderma) 를 포함하는 여러 가지 셀룰라아제-제조 미생물이, 매우 적은 베타-글루코시다아제를 생성한다는 점을 고려하면, 효소 가수분해에서의 주요 문제점이다. 사실상, 상업적 트리코데르마 (Trichoderma) 균주에 의해 분비되는 총 단백질의 1% 미만은 베타-글루코시다아제 유형이

다. 상기 적은 양의 베타-글루코시다아제는 따라서 셀로비오스를 글루코오스로 가수분해하는 낮은 능력을 산출하고, 따라서 시스템 내의 이의 축적을 야기한다. 이것이 일어나면서, 고 농도의 셀로비오스는 셀로비오스가 반응의 최종 생성물인 기타 셀룰라아제 및 특히 엑소글루카나아제의 활성을 억제한다. 상기 단점을 극복하기 위해, 본 출원인은 이들의 특허 출원 WO 2010/029259 에서, 증가된 특이적 활성을 가진 효소를 수득하는 것을 가능하게 만드는 베타-글루코시다아제 유전자를 개발하여, 이에 의해 리그노셀룰로오스 바이오매스를 바이오연료로 전환하기 위한 공정을 실질적으로 개선하였다.

[0006] 가수분해 및 발효는 다양한 기작에 따라 실시될 수 있다. 가장 흔한 것은 분리된 가수분해 및 발효 (SHF: Separate Hydrolysis and Fermentation) 로 이루어진다. 상기 방법은 최적 반응 조건을 유지함으로써 각 단계를 최적화시키는 것을 가능하게 한다. 상기 발효는 약 28℃ 내지 약 30℃ 의 온도에서 즉흥적으로 실시되는 반면, 가수분해는 일반적으로 45℃ 이상의 온도에서 실시된다. 그러나, SHF 에서, 반응 마지막에 방출되는 당은 농도가 매우 높고 효소의 억제를 야기하여, 공정의 효율을 감소시킨다.

[0007] 상기 결점을 피하기 위해, 또다른 유형의 공정 (SSF - Simultaneous Saccharification and Fermentation: 동시 당화 및 발효) 이 구상될 수 있다. SSF 에서, 2 개 단계 (핵소오스의 가수분해 및 발효) 가 동시에 일어나서, 효소에 대해 억제적인 농도에서의 당의 축적을 방지한다. 투자 비용은 또한 단일 반응기의 사용 후에 감소된다. 가수분해 정도는, 분비되는 당이 에탄올로의 발효에 즉시 사용되므로 억제제의 부재에 후속하여 더 높다.

[0008] 상기 방법에서, 반응기의 온도는 필수적으로 가수분해 및 발효에 대한 최적 온도 사이의 절충, 전형적으로는 약 30℃ 내지 약 35℃ 를 구성한다. 그러나, 이러한 온도에서, 베타-글루코시다아제를 비롯한 셀룰로오스분해 효소의 활성은 약 30% 감소한다.

## 발명의 내용

### 해결하려는 과제

[0009] 따라서 SSF 공정의 최적 가수분해 및 발효 온도에서, 특히 약 30℃ 내지 약 35℃ 의 온도에서 효율적인 베타-글루코시다아제 활성을 유지할 수 있는 효소에 대한 필요성이 있다.

### 과제의 해결 수단

[0010] 본 출원인은 약 30℃ 내지 약 35℃ 의 온도에서, 특히 서열 SEQ ID No. 3 의 야생형 BGL1 단백질의 베타-글루코시다아제 활성과 비교하여, 향상된 베타-글루코시다아제 활성을 갖는 폴리펩티드를 개발하였다. BGL1 은 트리코테르마 레에세이 (Trichoderma reesei) 로부터의 베타-글루코시다아제에 상응한다.

[0011] 본 출원인은 야생형 BGL1 단백질의 베타-글루코시다아제 활성과 비교하여 향상된 특이적 베타-글루코시다아제 활성을 갖는 여러 개의 클론을 이전에 확인하였다. 이러한 결과는 이들의 특허 출원 WO 2010/029259 에 제시된다. 더욱 구체적으로는, 이들은 강한 프로모터의 조절 하에서 트리코테르마 레에세이 (Trichoderma reesei) 에서의 이의 발현이 상기 효소를 발현하지 않는 균주에 의해 생성되는 것과 비교하여, 생성된 효소 각 테일의 베타-글루코시다아제 활성에서 26.2 배 증가를 야기하는 (특허 출원 WO 2010/029259 의 표 6), SEQ ID No. 5 의 폴리펩티드 (100B11 이라고 불림) 를 엔코딩하는 특정 클론을 입증하였다.

[0012] 이제 본 출원인은 놀랍고도 예상치 못하게, 이전에 확인된 클론 100B11 과 비교하여 향상된 활성을 갖는 (이것은 약 30℃ 내지 약 35℃ 의 온도에서임) 효소를 엔코딩하는 새로운 클론을 입증하였다.

[0013] 본 발명은 따라서 아미노산 서열 SEQ ID No. 1 의, 베타-글루코시다아제 활성을 갖는 폴리펩티드에 관한 것이다.

[0014] 본 발명의 폴리펩티드의 아미노산 서열은 다음과 같다:

[0015] MRYRTAAALALATGPFARADSHSTSGASAEAVVPPAGTPWGTAYDKAKAALAKLNLQDKVGIVSGVGWNGGPCVGNTPSPASKIGYPQLCLQDGPLGIRFGGS  
VTAFTPGIQAASSTWDTLMRQRGEYLGAEAKGCGIHVLLGPVAGPLGKTPQGGRNWEGFGVDPYLTGIAAETIEGLQSAGVQACAKHYIVNEQELNRETIS  
SNPDDRTLHELHYLWPFADAVHANVASVMCSYNKINGSWACEDQYTLTQTLVKDQLGFGPYVMTDWNQAHTTVQSANSGLDMSMPGTDGNGNRLWGPALTNV  
NSNQVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNI SRNVQGNHKTNVRAIARDGIVLLKNDANILPLKPPASIAVVGSAIIGNHARNSPSCNDKGCDDGA  
LGMGWGSGAVNYPYFVAPYDAINTRASSQGTQVTLSTNTSSGASAAARGKDVAIVFITADSGEGYITVEGNAGDRNLDLPWHNGNALVQAVAGANSNVIVV  
VHSVGAILEQILALPQVKAVVWAGLPQSQESGNALVDVLWGDVSPSGKLVTYIAKSPNDYNTRIVSGGSDSFSEGLFIDYKHFDANIIPRYEFGYGLSYTK

FNYSRLSVLSTAKSGPATGAVVPGGSDLFQNVATVTVDIANSGQVTGA EVAQLYITYPSSAPRTTPPKQLRGFAKLNLTGQSGTATFNIRRRDLSYWD TAS  
QKVVVPSGSFGISVGASSRDIRLTSTLSVA.

- [0016] 상기 폴리펩티드는 핵산 서열 SEQ ID No. 2 에 의해 엔코딩된다.
- [0017] 우선적으로, 아미노산 서열 SEQ ID No. 1 의 상기 폴리펩티드는 약 30℃ 내지 약 35℃ 의 온도에서, 특히 서열 SEQ ID No. 3 의 야생형 BGL1 단백질의 베타-글루코시다아제 활성과 비교하여 상기 동일한 온도에서 향상된 베타-글루코시다아제 활성을 갖는다. BGL1 단백질은 핵산 서열 SEQ ID No. 4 에 의해 엔코딩된다.
- [0018] 더욱 우선적으로, 아미노산 서열 SEQ ID No. 1 의 상기 폴리펩티드는 약 30℃ 내지 약 35℃ 의 온도에서 아미노산 서열 SEQ ID No. 5 의 100B11 폴리펩티드의 베타-글루코시다아제 활성과 비교하여 상기 동일한 온도에서 향상된 베타-글루코시다아제 활성을 갖는다. 100B11 폴리펩티드는 핵산 서열 SEQ ID No. 6 에 의해 엔코딩된다.
- [0019] 게다가, 본 발명에 따른 폴리펩티드는 글루코오스에 의한 억제에 대해 덜 민감하다는 장점을 갖고, 그 결과 높은 글루코오스 농도의 존재 하에 더욱 양호한 베타-글루코시다아제 활성을 보유한다.
- [0020] 하나의 구현예에서, 이전에 기재된 바와 같은 폴리펩티드는 글루코오스의 부재 하에서 측정된 야생형 단백질 BGL1 (SEQ ID No. 3) 의 베타-글루코시다아제 활성과 비교하여 향상된, 글루코오스의 존재하에서 측정된 베타-글루코시다아제 활성을 갖는다.
- [0021] 하나의 바람직한 구현예에서, 본 발명의 폴리펩티드는 아미노산 서열 SEQ ID No. 5 의 100B11 폴리펩티드의 베타-글루코시다아제 활성과 비교하여, 약 30℃ 내지 약 35℃ 의 온도에서 10% 이상, 우선적으로 20% 이상, 우선적으로 30% 이상, 더 더욱 우선적으로 40% 이상 향상된 베타-글루코시다아제 활성을 갖는다.
- [0022] 당업자는, 예를 들어, 기질 파라-니트로페닐 베타-D-글루코피라노시드 (pNPG) 를 사용하는 효소 활성 시험에 의해 본 발명에 따른 폴리펩티드의 효소 활성의 증가 또는 다른 말로는 개선을 측정하는 것이 가능할 것이다. 베타-글루코시다아제의 작용 후 수득되는 파라-니트로페놀의 양은, 예를 들어, 414 nm 에서 광학 밀도를 관측함으로써 측정될 수 있다.
- [0023] 당업자가 본 발명에 따른 폴리펩티드가 야생형 BGL1 단백질의 것과 비교하여 향상된 효소 활성을 갖는지를 측정하기 위해 사용할 수 있는 프로토콜의 예는 하기와 같다:
- [0024] - 37℃ 에서 밤새, 본 발명에 따른 폴리펩티드를 발현하는 대장균 (E. coli) 의 저장 배양액의 제조;
- [0025] - 24h 동안 20℃ 에서 LB 배양 배지의 1% 의 저장 배양액으로의 접종;
- [0026] - 2 분 동안 13 000 rpm 에서 원심분리;
- [0027] - pH 5 (최종 OD<sub>600</sub> = 100) 에서 100 mM 숙시네이트 완충액으로의 세포 펠렛의 재현탁;
- [0028] - 50  $\mu$ l 의 세포의 15 mM 의 파라-니트로페닐 베타-D-글루코피라노시드 (pNPG) 를 함유하는 100  $\mu$ l 의 100 mM 숙시네이트 완충액 (pH 5) 으로, 50℃ 에서 1 시간 30 분 이후, 빙상에서 5 분 동안의 인큐베이션;
- [0029] - 150  $\mu$ l 의 0.2 M Na<sub>2</sub>CO<sub>3</sub> 의 첨가;
- [0030] - 2 분 동안 13 000 rpm 에서의 원심분리;
- [0031] - 150  $\mu$ l 의 상청액에 대한 414 nm 에서의 광학 밀도의 관측.
- [0032] 게다가, 당업자는 본 발명에 따른 폴리펩티드가 야생형 BGL1 단백질보다 글루코오스 억제에 대해 덜 민감한지를 측정하기 위해, 50℃ 에서 1 시간 30 분 동안, 50  $\mu$ l 의 세포의 15 mM 의 pNPG 및 60 g/l 의 글루코오스를 함유하는 100  $\mu$ l 의 100 mM 숙시네이트 완충액 (pH 5) 으로 인큐베이션에 의해 상기 기재된 프로토콜을 사용할 수 있을 것이다.
- [0033] 상기 프로토콜은 약 30℃ 내지 약 35℃ 의 온도 조건 하에, 특히 아미노산 서열 SEQ ID No. 5 의 100B11 폴리펩티드와 비교하여, 베타-글루코시다아제 활성의 향상을 측정하기 위해 쉽게 적용가능하다.
- [0034] 본 발명은 또한 아미노산 서열 SEQ ID No. 1 의 폴리펩티드를 엔코딩하는 핵산에 관한 것이다. 우선적으로, 상기 핵산은 핵산 서열 SEQ ID No. 2 를 포함한다.
- [0035] 본 발명은 또한 이전에 기재된 핵산을 포함하는 벡터에 관한 것이다.



- [0036] 본 발명에 따르면, 용어 "백터" 는 외래 핵산 단편을 삽입하는 것을 가능하게 하는 임의의 DNA 서열을 의미하는 것으로 의도되며, 백터는 외래 DNA 를 숙주 세포 내로 도입하는 것을 가능하게 한다. 백터의 예는 플라스미드, 코스미드, 효모 인공 염색체 (YAC), 박테리아 인공 염색체 (BAC) 및 P1 박테리오파지-유래 인공 염색체 (PAC), 및 바이러스-유래 백터이다.
- [0037] 본 발명에 따르면, 이전에 기재된 핵산은 숙주 세포의 발현에 필요한 프로모터, 터미네이터 또는 임의의 기타 서열에 기능적으로 연결될 수 있다.
- [0038] 본 발명에 따른 백터는 또한 선별가능 마커를 가질 수 있다. 용어 "선별가능 마커" 는 유전자의 발현이 상기 마커를 함유하는 세포에게 이들을 선별할 수 있게 하는 특징을 부여하게 하는 그러한 유전자를 의미하는 것으로 의도된다. 이것은, 예를 들어, 항생제에 대한 내성에 대한 유전자이다.
- [0039] 본 발명의 주제는 또한 이전에 기재된 본 발명의 폴리펩티드를 제조할 수 있는, 또는 본 발명의 상기 폴리펩티드를 엔코딩하는 핵산을 포함하는 단리된 숙주 세포이다.
- [0040] 당업자는 적어도 이전에 기재된 폴리펩티드, 핵산 또는 백터를 잘-공지된 통상의 방법에 의해 숙주 세포 내로 도입할 수 있을 것이다. 예를 들어, 염화칼슘 처리, 전기천공, 또는 입자 총의 사용을 언급할 수 있다.
- [0041] 하나의 구현예에 따르면, 당업자는 숙주 세포 내로, 그리고 통상의 방법에 의해, 본 발명에 따른 향상된 베타-글루코시다아제 활성을 갖는 폴리펩티드를 엔코딩하는 핵산의 여러 카피를 도입할 수 있을 것이다.
- [0042] 하나의 구현예에 따르면, 이전에 기재된 단리된 숙주 세포는 트리코데르마 (Trichoderma), 아스페르길루스 (Aspergillus), 뉴로스포라 (Neurospora), 휴미콜라 (Humicola), 마이셀리오프토라 (Myceliophthora), 크리소스포리움 (Chrysosporium), 페니실리움 (Penicillium), 푸사리움 (Fusarium), 테르모모노스포라 (Thermomonospora), 바실러스 (Bacillus), 슈도모나스 (Pseudomonas), 에스케리차 (Escherichia), 클로스트리듐 (Clostridium), 셀룰로모나스 (Cellulomonas), 스트렙토마이세스 (Streptomyces), 야로위아 (Yarrowia), 피치아 (Pichia) 및 사카로마이세스 (Saccharomyces) 로부터 선택된다.
- [0043] 하나의 바람직한 구현예에 따르면, 이전에 기재된 단리된 숙주 세포는 트리코데르마 레에세이 (Trichoderma reesei), 트리코데르마 비리다에 (Trichoderma viridae), 트리코데르마 코닝기이 (Trichoderma koningii), 아스페르길루스 니게르 (Aspergillus niger), 아스페르길루스 니돌란스 (Aspergillus nidulans), 마이셀리오프토라 테르모필라 (Myceliophthora thermopila), 크리소스포리움 룩노웬세 (Chrysosporium lucknowense), 아스페르길루스 웬티이 (Aspergillus wentii), 아스페르길루스 오리자에 (Aspergillus oryzae), 아스페르길루스 포에니시스 (Aspergillus phoenicis), 뉴로스포라 크라사 (Neurospora crassa), 휴미콜라 그리사에 (Humicola grisae), 페니실리움 피노필룸 (Penicillium pinophilum), 페니실리움 옥살리쿰 (Penicillium oxalicum), 에스케리차 콜라이 (Escherichia coli), 클로스트리듐 아세트부틸리쿰 (Clostridium acetobutylicum), 클로스트리듐 사카롤라이티쿰 (Clostridium saccharolyticum), 클로스트리듐 벤제린키이 (Clostridium benjerinckii), 클로스트리듐 부틸리쿰 (Clostridium butylicum), 피치아 파스토리스 (Pichia pastoris), 야로위아 리폴리티카 (Yarrowia lipolityca), 사카로마이세스 세레비지아에 (Saccharomyces cerevisiae), 및 이의 혼합물로부터 선택된다.
- [0044] 하나의 바람직한 구현예에 따르면, 이전에 기재된 단리된 숙주 세포는 트리코데르마 레에세이 (Trichoderma reesei) 및 사카로마이세스 세레비지아에 (Saccharomyces cerevisiae) 로부터 선택된다.
- [0045] 본 발명은 또한 베타-올리고당의 가수분해를 위한, 이전에 기재된 폴리펩티드 또는 이전에 기재된 세포 중 임의의 하나의 용도에 관한 것이다.
- [0046] 본 발명은 또한 셀로비오스의 글루코오스로의 가수분해를 위한, 이전에 기재된 폴리펩티드 또는 이전에 기재된 세포 중 임의의 하나의 용도에 관한 것이다.
- [0047] 본 발명의 주제는 또한 바이오연료의 생성을 위한, 이전에 기재된 폴리펩티드 또는 이전에 기재된 세포 중 임의의 하나의 용도이다.
- [0048] 본 발명에 따르면, 용어 "바이오연료" 는 바이오매스의 전환으로부터 산출되고 에너지 목적을 위해 사용될 수 있는 임의의 생성물로서 정의될 수 있다. 처음으로, 그리고 제한 없이, 바이오가스는 예를 들어, (임의로 후속 전환 후) 연료 내로 도입될 수 있거나 또는 그 자체로 연료일 수 있는 생성물, 예컨대 알코올 (사용되는 발효성 유기체의 유형에 따라 에탄올, 부탄올 및/또는 이소프로판올), 용매 (아세톤), 산 (부티르산), 지질 및



이의 유도체 (단쇄 또는 장쇄 지방산, 지방산 에스테르), 및 또한 수소가 언급될 수 있다.

- [0049] 바람직하게는, 본 발명에 따른 바이오연료는 알코올, 예를 들어 에탄올, 부탄올 및/또는 이소프로판올이다. 더욱 우선적으로, 본 발명에 따른 바이오연료는 에탄올이다.
- [0050] 또다른 구현예에서, 바이오연료는 바이오가스이다.
- [0051] 또다른 구현예에서, 생성물은 화학 산업에서 유리한 분자, 예를 들어, 기타 알코올, 예컨대 1,2-프로판디올, 1,3-프로판디올, 1,4-부탄디올, 2,3-부탄디올, 유기 산, 예컨대 아세트산, 프로피온산, 아크릴산, 부티르산, 숙신산, 말산, 푸마르산, 시트르산 또는 이타콘산, 또는 히드록시산, 예컨대 글리콜산, 히드록시프로피온산 또는 락트산이다.
- [0052] 바이오연료의 제조에 더해, 30℃ 내지 35℃ 의 온도에서 향상된 베타-글루코시다아제 활성을 갖는 폴리펩티드는 또한 다양한 기질의 가수분해를 촉매화하고, 따라서 다양한 아로마/풍미의 방출을 가능하게 함으로써 다른 유형의 적용에 사용될 수 있다. 예를 들어, 이것은 상기 과일 내에 존재하는 여러 글루코시드를 가수분해함으로써 과일 풍미를 방출하기 위해 사용될 수 있거나, 또는 그 밖에 이것은 포도 모노테르페닐 베타-글루코시드를 가수분해하여, 와인에 대한 풍미의 중요한 공급원을 나타낼 수 있다. 따라서, 이전에 기재된 폴리펩티드는 여러 분야, 특히 향수산업, 식품 산업, 와인 양조학 등에서 사용될 수 있다.
- [0053] 본 발명에 따른 폴리펩티드를 발현할 수 있는 사상 진균의 균주, 바람직하게는 트리코데르마 (Trichoderma), 더욱 우선적으로 T. 레에세이 (T. reesei) 는 미생물의 성장을 위해 선택된, 탄소-기재 배지, 예컨대 락토오스 또는 글루코오스의 존재 하에, 발효조에서 배양된다. 하나의 구현예에서, 상기 탄소-기재 기질은, 그의 특성에 따라, 살균 전에 발효조 내에 도입되거나 또는 개별적으로 살균되고 20 내지 35 g/l 의 초기 농도를 수득하도록 이의 살균 후에 발효조 내에 도입된다.
- [0054] 효소의 생성으로부터 선택되는 기질을 함유하는 수용액을 이후 첨가한다. 진균에 의해 생성된, 리그노셀룰로오스 바이오매스 상에 작용하는 효소 조성물은 배양 배지의 여과에 의해 최종적으로 회수된다. 상기 조성물은 특히 엔도글루카나아제, 엑소글루카나아제 및 본 발명에 따른 베타-글루코시다아제를 함유한다. 하나의 구현예에서, 효소의 생성으로부터 선택되는 기질을 함유하는 수용액을 200-250 g/l 의 농도로 제조하고; 상기 용액은 락토오스와 같은 유도체 기질을 함유해야만 한다. 상기 수용액은 35 내지 45 mg/g 의 셀 (공급 배치) 의 최적 양을 제공하도록 초기 탄소-기재 기질의 소모 후에 주입된다. 상기 공급 배치 상 동안, 배양 배지 내의 당의 잔류 농도는 1 g/l 미만이고, 리그노셀룰로오스 바이오매스 상에 작용하는 효소는 진균에 의해 분비된다. 상기 효소는 배양 배지의 여과에 의해 회수될 수 있다.
- [0055] 본 발명의 주제는 리그노셀룰로오스 바이오매스 상에 작용하는 효소 조성물로서, 상기 효소 조성물은 사상 진균 또는 효모에 의해 생성되고 이전에 기재된 폴리펩티드를 포함하는 효소 조성물이다.
- [0056] 마지막으로, 본 발명의 주제는 하기 단계:
- [0057] - 수성상 내에, 가수분해하고자 하는 물질의 현탁;
- [0058] - 글루코오스를 함유하는 가수분해물을 생성하도록, 효소 조성물의 존재 하에, 이전에 기재된 리그노셀룰로오스 바이오매스의 가수분해;
- [0059] - 가수분해물의 글루코오스의 발효;
- [0060] - 발효 포도액으로부터 바이오연료의 분리
- [0061] 를 포함하고,
- [0062] 가수분해 및 발효 단계가 동시에 실시되는 것을 특징으로 하는, 바이오매스로부터 바이오연료의 제조 방법이다.
- [0063] 본 발명의 또다른 주제는 하기 단계:
- [0064] - 수성상 내에, 가수분해하고자 하는 바이오매스의 현탁;
- [0065] - 이전에 정의된 리그노셀룰로오스 바이오매스 상에 작용하는 효소 조성물 및 발효성 유기체의 동시 첨가 및 인큐베이션;
- [0066] - 발효 포도액으로부터 바이오연료의 분리

- [0067] 를 포함하는 것을 특징으로 하는, 바이오매스로부터 바이오연료를 제조하기 위한 방법이다.
- [0068] 본 발명의 또다른 주제는 하기 연속 단계:
- [0069] - 수성상 내에, 가수분해하고자 하는 바이오매스의 현탁;
- [0070] - 발효 포도액을 생성하도록, 30℃ 내지 35℃ 의 온도에서 이전에 정의된 하나 이상의 셀룰로오스분해성 및/또는 발효성 유기체의 첨가;
- [0071] - 발효 포도액으로부터 바이오연료의 분리
- [0072] 를 포함하는 것을 특징으로 하는, 바이오매스로부터 바이오연료의 제조 방법이다.
- [0073] 상기 구현예에 따르면, 바이오매스 내에 존재하는 셀룰로오스는 글루코오스로 전환되고, 동시에, 동일한 반응기 내에서, 발효성 유기체 (예를 들어 효모) 는 당업자에게 공지된 SSF (동시 당화 및 발효) 공정에 따라 글루코오스를 최종 생성물로 전환시킨다. 발효성 유기체의 대사성 및 가수분해성 능력에 따라, 작업의 올바른 수행은 더 많은 또는 더 적은 양의 외생 셀룰로오스분해성 혼합물의 첨가를 필요로 할 수 있다.
- [0074] 또다른 구현예에서, 발효성 유기체는 사상 진균에 의해 생성된 효소에 대한 필요성을 제한하거나 제거하도록 리그노셀룰로오스 바이오매스 상에 작용하는 다른 효소와 임의로 함께, 이의 세포 표면 상에 또는 분비에 의해 본 발명의 주제인 폴리펩티드를 생성한다.
- [0075] 본 발명에 따라 약 30℃ 내지 약 35℃ 의 온도에서 더욱 양호한 베타-글루코시다아제 활성을 나타내는 폴리펩티드의 사용은 따라서 더욱 양호한 글루코오스 생성 수율을 수득하는 장점을 갖는다. 따라서, 본 발명은 이전보다 효소를 적게 사용하는 것을 가능하게 하며, 이것은 바이오연료 생성 비용이 예를 들어, 줄어드는 경제적인 장점을 갖는다.
- [0076] 본 발명의 기타 양상, 주제, 장점 및 특성은 하기에 이어지고 예를 들어 제시된, 본 발명의 바람직한 구현예를 기술하는 비-제한적인 설명의 판독 시 제시될 것이다.

## 발명을 실시하기 위한 구체적인 내용

- [0077] 실시예
- [0078] 실시예 1 : 첫번째 라운드 서플링
- [0079] 트리코데르마 레에세이 (Trichoderma reesei) 베타-글루코시다아제 유전자 (모체 유전자 BGL1, SEQ ID No. 4) 의 서열을 EP 1 104 457 B1 에 기재된 특허 공정에 따라 BGL1 모체 유전자와 70% 일치성을 갖는 카에토뮴 글로보숨 (Chaetomium globosum) (유전자 A) (SEQ ID No. 7, 핵산 서열 SEQ ID No. 8 에 의해 엔코딩됨) 의 추정적 글루코시다아제 유전자와의 첫번째 라운드 서플링에 적용하였다.
- [0080] 1- 고효율 스크리닝
- [0081] 고효율 스크리닝 시험은 상기 2 개의 서열의 서플링으로부터 산출되는 최고의 클론, 즉, T. 레에세이 (T. reesei) 로부터의 BGL1 모체 유전자와 비교하였을 때 베타-글루코시다아제 활성 수준에서 2 보다 큰 향상 인자를 갖는 것들을 선별하는 것을 가능하게 하였다.
- [0082] 첫번째 라운드 서플링의 라이브러리 스크리닝 시험을 하기 단계에 따라 실시하였다:
- [0083] - 본 발명에 따른 제조합 효소의 서플링 변이체를 발현하는 대장균의 다양한 콜로니의 아가 상의 단리 및 37℃ 에서 밤새 LB 배지에서 상기 콜로니의 예비배양;
- [0084] - 3% 예비배양물로 LB 배지의 접종, 이후 4h 동안 37℃ 에서 인큐베이션;
- [0085] - 100  $\mu$ M 이소프로필-베타-티오-갈락토사이드 (IPTG) 의 첨가에 의한 변이체의 발현 유도, 이후 20℃ 에서 밤새 인큐베이션;
- [0086] - 13 000 rpm 에서 2 분 동안 원심분리;
- [0087] - 2.2 mM 의 파라-니트로페닐 베타-D-글루코피라노시드 (pNPG) 를 함유하는 100  $\mu$ l 의 0.1 M 숙시네이트 완충액 중에 세포 펠렛의 재현탁;
- [0088] - 주위 온도에서 3h 동안 인큐베이션;

- [0089] - 알칼리화 후 414 nm 에서의 광학 밀도의 관독.
- [0090] 상기 스크리닝 조건 하에서, BGL1 참조 효소와 비교하여 베타-글루코시다아제 활성의 향상을 나타내는 여러 개의 클론을 확인하였다.
- [0091] **2-베타-글루코시다아제 활성의 향상 측정**
- [0092] **2-1/ pNPG 기질 상에서**
- [0093] 첫번째 라운드 서플링에서 선택되는 변이체의 상대적 kcat 를 측정하기 위해, 하기 절차를 실시하였다:
- [0094] - 37℃ 에서 밤새, 본 발명에 따른 재조합 효소를 발현하는 대장균의 저장 배양액의 형성;
- [0095] - IPTG (250  $\mu$ M) 유도과 함께 24h 동안 20℃ 에서 LB 배양 배지의 1% 의 저장 배양액으로의 접종;
- [0096] - 2 분 동안 13 000 rpm 에서 원심분리;
- [0097] - pH 5 (최종 OD<sub>600</sub> = 100) 에서 100 mM 숙시네이트 완충액으로의 세포 펠렛의 재현탁;
- [0098] - 50  $\mu$ l 의 세포의 15 mM 의 파라-니트로페닐 베타-D-글루코피라노시드 (pNPG) 를 함유하는 100  $\mu$ l 의 100 mM 숙시네이트 완충액 (pH 5) 으로, 50℃ 에서 1 시간 30 분 이후, 빙상에서 5 분 동안의 인큐베이션;
- [0099] - 150  $\mu$ l 의 0.2 M Na<sub>2</sub>CO<sub>3</sub> 의 첨가;
- [0100] - 2 분 동안 13 000 rpm 에서의 원심분리;
- [0101] - 150  $\mu$ l 의 상청액에 대한 414 nm 에서의 광학 밀도의 관독.

[0102] 표 2 는 상기 실험 조건 하에서 3 개의 이전에 확인된 클론 (10H7, 59B8 및 164A2 로 불림) 에 대해 수득된 kcat 값 및 또한 향상 인자를 제공한다.

[0103] **표 2 : 베타-글루코시다아제 활성의 향상**

[0104] (유도된 배양물의 결과)

	클론	K <sub>cat</sub> (분 <sup>-1</sup> )	향상 인자
1 <sup>st</sup> -라운드 클론	<b>10H7</b>	590.0	8
	<b>59B8</b>	518.6	7
	<b>164A2</b>	1437.3	20
참조 단백질	<b>BGL1</b>	71.0	1

[0105]

[0106] 결과는 3 개의 클론 10H7, 59B8 및 164A2 에 대해 야생형 효소 (BGL1) 와 비교한 효소 활성의 매우 상당한 향상을 보여준다.

[0107] **2-2/ 셀로비오스 상에서**

- [0108] 10H7, 59B8 및 164A2 클론의 활성의 향상을 이후 두번째 기질: 셀로비오스에 대해 확인하였다.
- [0109] 상기 시험은 본 발명에 따른 재조합 효소를 발현하는 대장균의 배양물 상에서 실시하였다. 시험 단계는 하기와 같다:
- [0110] - IPTG 로 유도되는 1% 의 저장 배양액으로의 LB 배양 배지의 접종, 이후 37℃ 에서 밤새 인큐베이션;
- [0111] - 상기 세포를 37℃ 에서 0.4 의 600 nm 에서의 광학 밀도가 수득될 때까지 배양;
- [0112] - 상기 세포를 250  $\mu$ M IPTG 로 20℃ 에서 20 시간 동안 유도;
- [0113] - 배양 배지 글루코오스를 제거하기 위해, 100 mM 숙시네이트 완충액 (pH 5) 에서 세포 펠렛을 3 회 세정;
- [0114] - 마이크로플레이트에서 12 시간 동안 50℃ 에서 10  $\mu$ l 의 상기 세포 및 190  $\mu$ l 의 셀로비오스 (263.2 mM) 로 이루어지는 반응 믹스 (RM1) (최종 농도 250 mM) 를 제조함;

- [0115] - 마이크로플레이트에서 12 시간 동안 50℃ 에서 인큐베이션.
- [0116] 발현:
- [0117] - 하기로 이루어지는 반응 믹스 (RM2) 를 제조한다:
- [0118] - 10  $\mu$ l 의 RM1,
- [0119] - 90  $\mu$ l 의 100 mM 숙시네이트 완충액 (pH 5),
- [0120] - 5  $\mu$ l 의 글루코오스 옥시다아제 (44 U/ml).
- [0121] - 주위 온도에서 1 시간 동안 인큐베이션한다.
- [0122] - 하기를 30 분 동안 주위 온도에서 혼합 및 인큐베이션한다:
- [0123] - 10  $\mu$ l 의 RM2,
- [0124] - 2  $\mu$ l 의 호스래디쉬 퍼록시다아제 (10 U/ml),
- [0125] - 5  $\mu$ l 의 100 mM ABTS,
- [0126] - 83  $\mu$ l 의 50 mM 포스페이트 완충액, pH 7.4.
- [0127] - 420 nm 에서 광학 밀도를 판독한다.

[0128] **표 3 : 베타-글루코시다아제 활성의 향상**  
 [0129] (유도된 배양액의 결과)

	클론	K <sub>cat</sub> (분 <sup>-1</sup> )	향상 인자
1 <sup>st</sup> -라운드 클론	10H7	69.1	13
	59B8	37.7	7
	164A2	213.2	41
참조 단백질	BGL1	5.2	1

- [0130]
- [0131] 마찬가지로, 결과는 셀로비오스가 기질로서 사용되었을 때 10H7, 59B8 및 164A2 클론에 대해 야생형 효소 (BGL1) 와 비교한 효소 활성의 매우 상당한 향상을 보여준다.

[0132] 실시예 2 : 두번째 라운드 서플링

- [0133] 첫번째 라운드 서플링에서 수득된 향상된 유전자의 서열을 이어서 두번째 라운드 서플링 (여전히 EP 1 104 457 B1 에 기재된 특허받은 방법에 따름) 에 적용하였다. 유전적 다양성을 증가시키기 위해, 야생형 BGL1 효소와 70% 일치성을 갖는 베타-글루코시다아제를 엔코딩하는 하나 이상의 유전자를 첨가하였다.

- [0134] 더욱 구체적으로는, 뉴로스포라 크라사 (*Neurospora crassa*) 의 추정적 글루코시다아제 유전자 (유전자 C) (핵산 서열 SEQ ID No. 10 에 의해 엔코딩되는 SEQ ID No. 9) 가 사용되었다.

[0135] 1- 고효율력 스크리닝

- [0136] 이전에 기재된 고효율력 스크리닝 시험 (IPTG 유도 단계를 제외함, 첫번째 라운드 서플링에서 제공된 향상이 오직 프로모터 누출에 근거한 베타-글루코시다아제 활성의 검출을 가능하게 하므로) 을 최고의 클론, 즉, 164A2 클론과 비교하였을 때 베타-글루코시다아제 활성 수준에서 2 보다 큰 향상 인자를 나타내는 것들을 선별하기 위해, 본 두번째 라운드 서플링에 따라 수득되는 클론 상에서 실시하였다.

- [0137] 상기 스크리닝 조건 하에서, 참조 효소 (164A2) 와 비교한 베타-글루코시다아제 활성의 향상은, 특히 100B11 (핵산 서열 SEQ ID No. 6 에 의해 엔코딩된 SEQ ID No. 5) 및 115E1 (핵산 서열 SEQ ID No. 12 에 의해 엔코딩된 SEQ ID No. 11) 클론을 비롯한 여러 클론에서 발견되었다.

[0138] 2-베타-글루코시다아제 활성의 향상의 측정

[0139] 2-1/ pNPG 상에서

[0140] 상대적 kcat 를 측정하기 위해, 100B11 및 115E1 클론의 활성을 이전에 기재된 활성 시험에 의해 측정하였다.

[0141] 표 4 는 상기 실험 조건 하에서 100B11 및 115E1 클론에 대해 수득된 kcat 값 및 또한 향상 인자를 제시한다.

[0142] 표 4 : 베타-글루코시다아제 활성의 향상

[0143] (유도된 배양액의 결과)

	클론	K <sub>cat</sub> (분 <sup>-1</sup> )	향상 인자
2 <sup>nd</sup> -라운드 클론	100B11	4342.8	3.0
	115E1	3989.2	2.8
참조 단백질	164A2	1437.3	1

[0144]

[0145] 결과는 100B11 및 115E1 클론에 대해 참조 효소 (164A2) 및 BGL1 (X60) 과 비교한 효소 활성의 매우 상당한 향상을 보여준다.

[0146] 2-2/ 셀로비오스 상에서

[0147] 100B11 및 115E1 클론의 활성의 향상을 이후 두번째 기질: 셀로비오스 상에서 확인하였다.

[0148] 상대적 kcat 를 측정하기 위해, 100B11 및 115E1 클론의 활성을, 실시예 1 의 2-2 지점에서 기재된 바와 같은 기질로서 셀로비오스를 사용하는 이전에 기재된 50℃ 에서의 활성 시험에 의해 측정하였다.

[0149] 표 5 : 베타-글루코시다아제 활성의 향상

[0150] (유도된 배양액의 결과)

	클론	K <sub>cat</sub> (분 <sup>-1</sup> )	향상 인자
2 <sup>nd</sup> -라운드 클론	100B11	387.2	1.8
	115E1	406.4	1.9
참조 단백질	164A2	213.2	1

[0151]

[0152] 마찬가지로, 결과는 셀로비오스가 기질로서 사용되었을 때 100B11 및 115E1 클론에 대해 참조 효소 (164A2) 와 비교한 효소 활성의 매우 상당한 향상을 보여준다.

[0153] 실시예 3 : 세번째 라운드 서플링

[0154] 두번째 라운드 서플링에서 수득된 14 개의 향상된 유전자 (138E12, 134G2, 100B11, 115E1, 99G11, 127B12, 91F6, 135F9, 116D9, 212D11, 210A6, 124F5, 129D2 및 141F7) 의 서열을 이어서 세번째 라운드 서플링 (여전히 EP 1 104 457 B1 에 기재된 특허받은 방법에 따름) 에 적용하였다. 유전적 다양성을 증가시키기 위해, 상기 유전자와 70% 일치성을 갖는 베타-글루코시다아제를 엔코딩하는 하나 이상의 유전자를 첨가하였다. 본 정확한 예에서, 뉴로스포라 크라사 (Neurospora crassa) 의 추정적 베타-글루코시다아제 유전자 (유전자 C) (핵산 서열 SEQ ID No. 10 에 의해 엔코딩되는 SEQ ID No. 9) 및 카에토뮴 글로보숨 (Chaetomium globosum) 의 추정적 베타-글루코시다아제 유전자 (유전자 A) (핵산 서열 SEQ ID No. 8 에 의해 엔코딩되는 SEQ ID No. 7) 가 사용되었다.

[0155] 1- 고출력 스크리닝

[0156] 이전에 기재된 고출력 스크리닝 시험 (IPTG 유도 단계를 제외함, 첫번째 라운드 서플링에서 제공된 향상이 오직 프로모터 누출에 근거한 베타-글루코시다아제 활성의 검출을 가능하게 하므로) 을 본 세번째 라운드 서플링에 따라 수득되는 클론 상에서 실시하였다. 상기 클론의 활성을 30℃ 및 50℃ 에서 측정하였다.

[0157] 상기 스크리닝 조건 하에, 17E5 클론 (핵산 서열 SEQ ID No. 2 에 의해 엔코딩되는, 아미노산 서열 SEQ ID No. 1) 을 선택하였는데, 이것이 유리한 30℃/50℃ 활성 비를 가지고 있기 때문이다.

[0158] 표 6 은 17E5 클론 및 100B11 클론 (두번째 라운드 서플링으로부터 산출되는 참조 클론) 에 대해 50℃ 및 30℃

에서 수득된 상대적 활성을 제공한다.

**표 6 : 30℃ 에서의 상대적 활성**

	50℃	30℃
<b>17E5</b>	100%	80%
<b>100B11</b>	100%	53%

결과는 17E5 클론이, 100B11 클론의 경우 53% 에 대비하여, 50℃ 에서의 이의 활성과 비교하여 30℃ 에서 80% 활성을 보유한다는 것을 보여준다.

게다가, 이의 특이적 활성은 100B11 효소의 것보다 2 의 인자가 더 크다.

#### **2-베타-글루코시다아제 활성의 측정**

상대적 kcat 를 측정하기 위해, 17E5 클론의 활성을 이전에 기재된 활성 시험에 의해 30℃ 및 50℃ 에서 측정하였다.

표 7 은 상기 실험 조건 하에서 17E5 클론에 대해 수득된 kcat 값 및 또한 향상 인자를 제시한다.

**표 7 : 30℃ 에서 베타-글루코시다아제 활성의 향상**

(비유도된 배양물의 결과)

	kcat (분 <sup>-1</sup> )		향상	
	30℃	50℃	30℃	50℃
<b>17E5</b>	4.2	10.94	2.32	2.17
<b>100B11</b>	1.81	5.03		

결과는 참조 클론과 비교하여 2 의 인자의 17E5 클론의 효소 활성의 향상을 보여준다 (이것은 두가지 온도 모두에서임).

#### **실시예 4 : 트리코데르마 레에세이 (Trichoderma reesei) 중의 베타-글루코시다아제의 향상된 변이체의 발현**

17E5 유전자를 하이그로마이신을 사용하는 선별 (스트렙토마이세스 하이그로스코피쿠스 (Streptomyces hygroscopicus) Hph 유전자) 로 RUT C30 (ATCC 56765), CL847 (Durand et al., Enzyme Microb. Technol., 1988; 10:341-346) 로부터 유래된 트리코데르마 레에세이 (Trichoderma reesei) 균주에서의 발현을 허용하는 벡터 내로 클로닝하였다. 17E5 유전자는 다른 T. 레에세이 (T. reesei) 셀룰라아제와 동시에 유도가능한 cbh1 강한 프로모터의 통제 하에 위치하였다.

트리코데르마 레에세이 (Trichoderma reesei) 의 형질전환을 당업자에게 공지된 통상의 방법에 따라 실시하였다 (갈슘 쇼크에 의한 원형질체의 형질전환 및 50 µg/ml 하이그로마이신으로의 선별). 포자형성에 의해 형질전환체를 정제하고, 이후 불안정한 클론을 제거하기 위해 선별 배지에서 2 회 서브배양하였다.

이후 30 개의 클론을 24-웰 플레이트 중의 셀룰라아제 생성에 대해 평가하였다. 각각의 클론의 몇몇 포자를 하기 조성을 갖는 2 ml 의 배지를 접종하는데 사용하였다: 20 g/l 락토오스, 20 g/l Solka floc 셀룰로오스, 5 g/l 펙톤, 15 g/l KH<sub>2</sub>PO<sub>4</sub>, 5 g/l (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.6 g/l CaCl<sub>2</sub>, 0.6 g MgSO<sub>4</sub>, 0.005 g/l FeSO<sub>4</sub>, 0.0014 g/l MnSO<sub>4</sub>, 0.0014 g/l ZnSO<sub>4</sub>, 0.0037 g/l CoCl<sub>2</sub>, 11.6 g/l 의 말레산, 12.1 g/l 의 트리스 및 2.08 g/l 의 NaOH. 플라스크를 150 rpm 로 교반하면서 30℃ 에서 인큐베이션하였다.

5 일 후, 배양물을 원심분리하고, 상청액의 단백질 농도를 폴린 (Folin) 방법을 사용하여 측정하였다. 상청액의 베타-글루코시다아제 활성을 하기 조건 하에서 파라-니트로페닐 베타-D-글루코피라노시드 (pNPG) 발색단 기질의 가수분해에 의해 측정하였다:

- 50 mM 의 시트레이트 완충액 (pH 4.8)

- 5 mM 의 pNPG



- [0177] - 10  $\mu$ l 의 샘플
- [0178] - 30°C 에서 30 분 동안 인큐베이션.
- [0179] 반응은 100  $\mu$ l 의 2% 탄산나트륨을 첨가함으로써 중지되었다. pNPG 의 가수분해에 의해 방출된 파라-니트로페놀의 양을 410 nm 에서의 흡광도를 측정함으로써 측정하고, 파라-니트로페놀 범위와 비교하였다. 반응은 25 에서 400  $\mu$ M 의 파라-니트로페놀로 선형이었다. 샘플을 임의로 희석하여, 측정된 흡광도가 선형성의 범위에 남아있게 하였다. 베타-글루코시다아제 활성을 비교를 위해, 상기와 동일한 조건 하에서 또한 50°C 에서 측정하였다. 가장 높은 베타-글루코시다아제 활성 (기원 균주와 비교하여 적어도 5 의 인자가 큼) 을 나타내는 클론을 선별하였다.
- [0180] 표 8 은 야생형 CL847 균주로부터, 변이체 100B11 을 발현하는 균주로부터 및 상기 기재된 방법에 따라 수득된, 변이체 17E5 를 발현하는 클론 중 하나로부터 각각 유래된, 상청액에 대해  $\mu$ mol/min/mg 의 효소로 측정된 30°C /50°C pNPase 베타-글루코시다아제 활성을 보여준다.

[0181] 표 8: 야생형 CL847, 100B11 폴리펩티드 및 17E5 폴리펩티드의 베타-글루코시다아제 활성

	30°C/50°C 활성 비	30°C 에서의 특이적 활성	50°C 에서의 특이적 활성
CL847	0.2	0.06	0.3
100B11	0.3	3.7	12.5
17E5	0.5	4.7	9.5

- [0182]
- [0183] 30°C/50°C 비의 증가는 30°C 의 온도에서 100B11 변이체의 것보다 큰 특이적 활성을 갖고, 17E5 클론에서 명시된다.

[0184] 실시예 5 : 사카로마이세스 세레비지아에 (*Saccharomyces cerevisiae*) 내의 야생형 베타-글루코시다아제 (BGL1) 및 향상된 변이체 100B11 및 17E5 의 재조합 발현

[0185] 1- 효모 세포질 내의 BGL1, 100B11 및 17E5 단백질의 제조:

- [0186] 트리코데르마 레에세이 (*Trichoderma reesei*) (BGL1) 의 야생형 베타-글루코시다아제 유전자 및 또한 100B11 및 17E5 변이체의 것을 신호 펩티드 없이 pESC-Leu 벡터 (Agilent Technologies) 내로 클로닝하였다. 상기 구축물은 류신 및 트립토판에 대해 영양요구성인, 사카로마이세스 세레비지아에 (*Saccharomyces cerevisiae*) EBY100 균주의 세포질에서 단백질의 발현을 허용한다 (Boder ET and Wittrup KD, Biotechnol Prog, 1998, 14:55-62). 상기 플라스미드는 유전자 발현을 갈락토오스-유도가능 GAL1 프로모터의 조절 하에 두는 것을 가능하게 하고, 형질전환체의 선별을 허용하는 선별가능 영양요구성 마커 유전자 (Leu2) 를 갖는다. 생성된 단백질은 마지막으로 N-말단 c-myc 태그에 융합되어, 친화성 크로마토그래피에 의해 생성된 효소의 검출 및 정제를 허용한다.

- [0187] S. 세레비지아에 (*S. cerevisiae*) EBY100 의 형질전환을 당업자에게 공지된 통상의 방법에 따라 실시하였다 (열 쇼크 및 리튬 아세테이트에 의한 효모의 형질전환). 형질전환체를 0.67% 의 Yeast Nitrogen Base (YNB), 2% 의 글루코오스 및 0.01% 의 트립토판을 함유하는 YNB-Glc-Trp 배지 상에서 선별하였다.

- [0188] 각각의 유전자에 대해 (Sc-BGL1, Sc-100B11 및 Sc-17E5) 하나의 형질전환체를 0.67% 의 YNB, 0.5% 의 카사미노산 (CAA), 0.01% 의 트립토판 및 2% 의 글루코오스를 함유하는 15 ml 의 YNB-Glc-CAA-Trp 최소 배지를 접종하는데 사용하였다. 220 rpm 에서 교반하면서 30°C 에서 24 시간의 예비배양 후, 3 개의 Sc-BGL1, Sc-100B11 및 Sc-17E5 균주를 0.67% 의 YNB, 0.5% 의 CAA, 0.01% 의 트립토판 및 2% 의 갈락토오스를 함유하는 150 ml 의 YNB-Gal-CAA-Trp 배지 (0.5 의 OD<sub>600</sub> 에서) 를 접종하는데 사용하였다. 배양물을 220 rpm 에서 교반하면서 25°C 에서 인큐베이션하였다.

- [0189] 4 일 인큐베이션 후, 20 ml 의 배양물을 3000 g 로, 4°C 에서 5 분 동안 원심분리하였다. 효모 펠렛을 3 ml 의 50 mM 시트레이트 완충액, pH 5 에 취하고, 2.5 kbar 의 압력으로 기계적으로 용리하였다. 세포질 추출물을 4°C 에서 30 분 동안 50 000 g 에서의 원심분리 후 수득하였다.

[0190] 2- 베타-글루코시다아제 활성의 측정



- [0191] 세포질 추출물 중의 총 단백질 농도를 Bradford 어세이 (Bradford MM., Anal Biochem, 1976, 72:248-54) 에 의해, 평균적으로 1.7 mg/ml 에서 추정하였다.
- [0192] 세포질 추출물의 베타-글루코시다아제 활성을 하기 조건 하에서 600  $\mu$ l 의 부피로 파라-니트로페닐 베타-D-글루코피라노시드 (pNPG) 기질의 가수분해에 의해 측정하였다:
- [0193] - 50 mM 의 시트레이트 완충액 (pH 5)
- [0194] - 5 mM 의 pNPG
- [0195] - 6.1  $\mu$ g 의 총 단백질을 함유하는 3.6  $\mu$ l 의 세포질 추출물
- [0196] - 30°C 또는 50°C 에서 30 분 동안 인큐베이션.
- [0197] 100  $\mu$ l 의 1M 탄산나트륨을 100  $\mu$ l 의 가수분해 반응에 첨가함으로써 반응을 중지시켰다. pNPG 의 가수분해에 의해 방출된 파라-니트로페놀 (pNP) 의 농도를 415 nm 에서의 흡광도를 측정함으로써 결정하고, 파라-니트로페놀의 표준 범위 (0.36  $\mu$ M 에서 360  $\mu$ M 로 선형) 와 비교하였다. 세포질 추출물을 초기 반응 속도 조건 하에 두도록 하기 위해 임의로 희석하였다.
- [0198] 표 9 는 야생형 효소 (Sc-BGL1) 를 발현하는 균주, 100B11 (Sc 100B11) 을 발현하는 균주 및 17E5 (Sc-17E5) 를 발현하는 균주로부터 각각 유래된 세포질 추출물에 대해  $\mu$ mol.min<sup>-1</sup>.mg<sup>-1</sup> 의 총 단백질로 측정된 30°C/50°C 베타-글루코시다아제 활성 비를 보여준다.

[0199] 표 9: Sc-BGL1, Sc-100B11 및 Sc-17E5 의 베타-글루코시다아제 활성

	30°C 에서의 특이적 활성	50°C 에서의 특이적 활성	30°C/50°C 활성 비	야생형 BGL1 과 비교한 30°C 에서의 특이적 활성의 향상
Sc-BGL1	0.15	0.41	0.4	-
Sc-100B11	0.18	0.64	0.3	1.2
Sc-17E5	0.46	1.12	0.4	3.1

- [0200]
- [0201] 상기 결과는 Sc-17E5 균주의 30°C 에서의 특이적 활성은 Sc-BGL1 균주와 비교하여 3 의 인자 및 Sc-100B11 과 비교하여 2.5 의 인자가 더 크다는 것을 보여준다.
- [0202] **실시예 6 : 야생형 베타-글루코시다아제 (BGL1) 및 S. 세레비지아에 (S. cerevisiae) 에서 생성된 향상된 변이체 100B11 및 17E5 의 정제 및 특징분석**
- [0203] 1- **베타-글루코시다아제 정제:**
- [0204] 실시예 5 의 Sc-BGL1 및 Sc-100B11 및 Sc-17E5 변이체의 세포질 추출물을 하기 프로토콜에 따라 상응하는 효소, BGL1, 100B11 및 17E5 를 정제하기 위해 사용하였다:
- [0205] 500  $\mu$ l 의 세포질 추출물을 축 교반으로 4°C 에서 1h 동안 20  $\mu$ l 의 "Anti-c-Myc tag Gel" 수지 (MBL) 와 함께 인큐베이션하였다. 13 000 rpm 에서 10 초의 원심분리 후에, 수지를 1X PBS 로 3 회 세정하였다. c-myc 펩티드 (EQKLISEEDL) (1 mg.ml<sup>-1</sup>) 로 구성된 용리 용액 중의 5 분 동안 4°C 에서의 수지의 인큐베이션 후, 단백질의 용리를 13 000 rpm 에서 10 초 동안의 원심분리에 의해 실시하였다.
- [0206] 2- **베타-글루코시다아제 활성의 측정**
- [0207] 정제된 효소의 농도를 고유의 BGL1 에 대해 120 125 M<sup>-1</sup>.cm<sup>-1</sup> 및 100B11 및 17E5 에 대해 120 250 M<sup>-1</sup>.cm<sup>-1</sup> 과 동일한 몰 흡광 계수를 사용하여, 나노드롭으로 280 nm 에서의 흡광도를 측정함으로써 수득하였다. 상기 농도는 평균 0.19 mg/ml 과 동일하다.
- [0208] 각각의 효소의 순도를 코마시 블루 (Coomassie blue) 를 사용하는 단백질 염색으로 SDS 의 존재 하에 10% 폴리 아크릴아미드 겔 상에서 전기영동에 의해 입증하였다.
- [0209] BGL1 및 정제된 100B11 및 17E5 변이체의 활성을 30°C 및 50°C 에서 이전에 기재된 바와 같이 측정하였다.

[0210] 표 10 은 30℃ 및 50℃ 에서 pNPG 의 가수분해 동안 측정된 각각의 효소 ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$  의 효소) 의 특이적 활성을 보여준다.

[0211] 표 10: 정제된 BGL1, 100B11 및 17E5 의 베타-글루코시다아제 활성

	30℃ 에서의 특이적 활성	50℃ 에서의 특이적 활성	30℃/50℃ 활성 비	야생형 BGL1 과 비교한 30℃ 에서의 특이적 활성의 향상
<b>BGL1</b>	5.1	8.9	0.57	-
<b>100B11</b>	7.1	17.2	0.41	1.4
<b>17E5</b>	10.2	23.6	0.43	2.0

[0212]

[0213] 결과는 야생형 BGL1 과 비교하여 2 및 100B11 변이체와 비교하여 1.4 의 인자에 의해 17E5 변이체의 특이적 활성의 30℃ 에서의 향상을 보여준다.

## 서열 목록

<110> IFP Energies nouvelles

<120> POLYPEPTIDE WITH REINFORCED BETA-GLUCOSIDASE ACTIVITY AT LOW TEMPERATURE

<130> BFF120334

<160> 12

<170> BiSSAP 1.0

<210> 1

<211> 744

<212> PRT

<213> Trichoderma reesei

<400> 1

Met Arg Tyr Arg Thr Ala Ala Ala Leu Ala Leu Ala Thr Gly Pro Phe

1 5 10 15

Ala Arg Ala Asp Ser His Ser Thr Ser Gly Ala Ser Ala Glu Ala Val

20 25 30

Val Pro Pro Ala Gly Thr Pro Trp Gly Thr Ala Tyr Asp Lys Ala Lys

35 40 45

Ala Ala Leu Ala Lys Leu Asn Leu Gln Asp Lys Val Gly Ile Val Ser

50 55 60

Gly Val Gly Trp Asn Gly Gly Pro Cys Val Gly Asn Thr Ser Pro Ala

65 70 75 80

Ser Lys Ile Gly Tyr Pro Gln Leu Cys Leu Gln Asp Gly Pro Leu Gly

85 90 95

Ile Arg Phe Gly Gly Ser Val Thr Ala Phe Thr Pro Gly Ile Gln Ala

100 105 110

Ala Ser Thr Trp Asp Thr Glu Leu Met Arg Gln Arg Gly Glu Tyr Leu

115 120 125

Gly Ala Glu Ala Lys Gly Cys Gly Ile His Val Leu Leu Gly Pro Val

130 135 140

Ala Gly Pro Leu Gly Lys Thr Pro Gln Gly Gly Arg Asn Trp Glu Gly

145 150 155 160

Phe Gly Val Asp Pro Tyr Leu Thr Gly Ile Ala Met Ala Glu Thr Ile

165 170 175

Glu Gly Leu Gln Ser Ala Gly Val Gln Ala Cys Ala Lys His Tyr Ile

180 185 190

Val Asn Glu Gln Glu Leu Asn Arg Glu Thr Ile Ser Ser Asn Pro Asp

195 200 205

Asp Arg Thr Leu His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val

210 215 220

His Ala Asn Val Ala Ser Val Met Cys Ser Tyr Asn Lys Ile Asn Gly

225 230 235 240

Ser Trp Ala Cys Glu Asp Gln Tyr Thr Leu Gln Thr Val Leu Lys Asp

245 250 255

Gln Leu Gly Phe Pro Gly Tyr Val Met Thr Asp Trp Asn Ala Gln His

260 265 270

Thr Thr Val Gln Ser Ala Asn Ser Gly Leu Asp Met Ser Met Pro Gly

275 280 285

Thr Asp Phe Asn Gly Asn Asn Arg Leu Trp Gly Pro Ala Leu Thr Asn

290 295 300

Ala Val Asn Ser Asn Gln Val Pro Thr Ser Arg Val Asp Asp Met Val

305 310 315 320

Thr Arg Ile Leu Ala Ala Trp Tyr Leu Thr Gly Gln Asp Gln Ala Gly

325 330 335

Tyr Pro Ser Phe Asn Ile Ser Arg Asn Val Gln Gly Asn His Lys Thr

340                      345                      350  
 Asn Val Arg Ala Ile Ala Arg Asp Gly Ile Val Leu Leu Lys Asn Asp  
 355                      360                      365  
 Ala Asn Ile Leu Pro Leu Lys Lys Pro Ala Ser Ile Ala Val Val Gly  
 370                      375                      380  
 Ser Ala Ala Ile Ile Gly Asn His Ala Arg Asn Ser Pro Ser Cys Asn  
  
 385                      390                      395                      400  
 Asp Lys Gly Cys Asp Asp Gly Ala Leu Gly Met Gly Trp Gly Ser Gly  
 405                      410                      415  
 Ala Val Asn Tyr Pro Tyr Phe Val Ala Pro Tyr Asp Ala Ile Asn Thr  
 420                      425                      430  
 Arg Ala Ser Ser Gln Gly Thr Gln Val Thr Leu Ser Asn Thr Asp Asn  
 435                      440                      445  
 Thr Ser Ser Gly Ala Ser Ala Ala Arg Gly Lys Asp Val Ala Ile Val  
 450                      455                      460  
  
 Phe Ile Thr Ala Asp Ser Gly Glu Gly Tyr Ile Thr Val Glu Gly Asn  
 465                      470                      475                      480  
 Ala Gly Asp Arg Asn Asn Leu Asp Pro Trp His Asn Gly Asn Ala Leu  
 485                      490                      495  
 Val Gln Ala Val Ala Gly Ala Asn Ser Asn Val Ile Val Val Val His  
 500                      505                      510  
 Ser Val Gly Ala Ile Ile Leu Glu Gln Ile Leu Ala Leu Pro Gln Val  
 515                      520                      525  
 Lys Ala Val Val Trp Ala Gly Leu Pro Ser Gln Glu Ser Gly Asn Ala  
  
 530                      535                      540  
 Leu Val Asp Val Leu Trp Gly Asp Val Ser Pro Ser Gly Lys Leu Val  
 545                      550                      555                      560  
 Tyr Thr Ile Ala Lys Ser Pro Asn Asp Tyr Asn Thr Arg Ile Val Ser  
 565                      570                      575  
 Gly Gly Ser Asp Ser Phe Ser Glu Gly Leu Phe Ile Asp Tyr Lys His  
 580                      585                      590  
 Phe Asp Asp Ala Asn Ile Thr Pro Arg Tyr Glu Phe Gly Tyr Gly Leu

595	600	605
Ser Tyr Thr Lys Phe Asn Tyr Ser Arg Leu Ser Val Leu Ser Thr Ala		
610	615	620
Lys Ser Gly Pro Ala Thr Gly Ala Val Val Pro Gly Gly Pro Ser Asp		
625	630	635
640		
Leu Phe Gln Asn Val Ala Thr Val Thr Val Asp Ile Ala Asn Ser Gly		
645	650	655
Gln Val Thr Gly Ala Glu Val Ala Gln Leu Tyr Ile Thr Tyr Pro Ser		
660	665	670
Ser Ala Pro Arg Thr Pro Pro Lys Gln Leu Arg Gly Phe Ala Lys Leu		
675	680	685
Asn Leu Thr Pro Gly Gln Ser Gly Thr Ala Thr Phe Asn Ile Arg Arg		
690	695	700
Arg Asp Leu Ser Tyr Trp Asp Thr Ala Ser Gln Lys Trp Val Val Pro		
705	710	715
720		
Ser Gly Ser Phe Gly Ile Ser Val Gly Ala Ser Ser Arg Asp Ile Arg		
725	730	735
Leu Thr Ser Thr Leu Ser Val Ala		
740		
<210>	2	
<211		
>	2235	
<212>	DNA	
<213>	Trichoderma reesei	
<400>	2	
atgcgttacc gaacagcagc tgcgtggca ctgcccactg ggccttttgc tagggcagac	60	
agtcactcaa catcgggggc ctcggtgag gcagttgtac ctctgcagg gactccatgg	120	
ggaaccgcgt acgacaagcg gaagccgca ttggcaaagc tcaatctcca agataaggtc	180	
ggcatcgtga gcggtgtcgg ctggaacggc ggtccttgcg ttggaaacac atctccggcc	240	
tccaagatcg gctatccaca gctatgcctt caagacggac ccctcggtat ccgattcggc	300	
ggcagcgtca cagcctttac gccgggcac caagcggcct cgacgtggga taccgagttg	360	
atgcgccagc gtggagagta cctgggtgcc gaggccaagg gctgcgggat tcatgtcctg	420	

cttggctcctg tggctgggcc gctgggaaaag actccgcagg gcggtcgcaa ctgggagggc	480
ttcgggtgtcg atccatatct caggggcatt gccatggccg agacaatcga gggcctgcag	540
tggccggcg tgcaggcgtg cgcaagcac tatatcgta acgagcagga gctcaatcga	600
gaaaccattt cgagcaacc agatgaccga actctccatg agctgtatct gtggccattt	660
gccgacgcgg ttacgcca tgctgcttct gtcattgtct cgtacaaca gatcaatggc	720
agctgggcct gcgaggatca gtacacgtg cagactgtgc tgaaagacca gctggggttc	780
ccaggctatg tcatgacgga ctggaacgca cagcacacga ctgtccaaag cgcaattct	840
gggcttgaca tgtcaatgcc tggcacagac ttcaacgta acaatcggt ctgggtcca	900
gctctacca atgcggtaaa tagcaatcag gtccccacga gcagagtcga cgatatggtg	960
actcgtatcc tgcgcgatg gtacttgaca ggccaggacc aggcaggcta tccgtcgttc	1020
aacatcagca gaaatgtta aggaaaccac aagaccaatg tcagggcaat tgccagggac	1080
ggcatcgttc tgtcaagaa tgacccaac atcctgccgc tcaagaagcc cgctagcatt	1140
gccgtcgttg gatctgccgc aatcattggt aaccaccca gaaactgcc ctctgcaac	1200
gacaaaggct gcgacgacgg ggccttgggc atgggttggg gttccggcgc cgtcaactat	1260
ccgtacttcg tgcgcaccta cgatgccatc aataccagag cgtcttcga gggcacccag	1320
gttaccttga gcaacaccga caacacgtcc tcaggcgcgt ctgcagcaag aggaaaggac	1380
gtcgccatcg tcttcacac cgccgactcg ggtgaaggct acatcacct ggagggcaac	1440
gcgggcgac gcaacaacct ggatccgtgg cacaacgca atgccctggt ccaggcggtg	1500
gccggtgcca acagcaacgt cattgttgtt gtccactccg ttggcgccat cattctggag	1560
cagattcttg ctcttcgca ggtcaaggcc gttgtctggg cgggtcttcc ttctcaggag	1620
agcggcaatg cgtcgtcga cgtgtgtggt ggagatgtca gcccttctgg caagctggtg	1680
tacaccattg cgaagagccc caatgactat aacactcga tcgtttccgg cggcagtgc	1740
agcttcagcg agggactgtt catcgactat aagcacttcg acgacgcaa tatcacccg	1800
cggtagcagt tcggctatgg actgtcttac accaagtca actactcac cctctccgtc	1860
ttgtcgaccg ccaagtctgg tcctcgact ggggccgttg tgccgggagg cccgagtgat	1920
ctgttcaga atgtcgcgac agtcaccgtt gacatcgaa actctggcca agtgactggt	1980
gccgagtag cccagctgta catcacctac ccattctcag caccaggac ccctccgaag	2040
cagctcgag gctttgcaa gctgaacct acgcctggtc agagcggaac agcaacgttc	2100
aacatccgac gacgagatct cagctactgg gacacgctt cgagaaatg ggtggtgccg	2160
tcggggtcgt ttggcatcag cgtgggagcg agcagccggg atatcaggt gacgagcact	2220
ctgtcggtag cgtag	2235

<210> 3

<211> 744

<212> PRT

<213> *Trichoderma reesei*

<400> 3

Met Arg Tyr Arg Thr Ala Ala Ala Leu Ala Leu Ala Thr Gly Pro Phe

1 5 10 15

Ala Arg Ala Asp Ser His Ser Thr Ser Gly Ala Ser Ala Glu Ala Val

20 25 30

Val Pro Pro Ala Gly Thr Pro Trp Gly Thr Ala Tyr Asp Lys Ala Lys

35 40 45

Ala Ala Leu Ala Lys Leu Asn Leu Gln Asp Lys Val Gly Ile Val Ser

50 55 60

Gly Val Gly Trp Asn Gly Gly Pro Cys Val Gly Asn Thr Ser Pro Ala

65 70 75 80

Ser Lys Ile Ser Tyr Pro Ser Leu Cys Leu Gln Asp Gly Pro Leu Gly

85 90 95

Val Arg Tyr Ser Thr Gly Ser Thr Ala Phe Thr Pro Gly Val Gln Ala

100 105 110

Ala Ser Thr Trp Asp Val Asn Leu Ile Arg Glu Arg Gly Gln Phe Ile

115 120 125

Gly Glu Glu Val Lys Ala Ser Gly Ile His Val Ile Leu Gly Pro Val

130 135 140

Ala Gly Pro Leu Gly Lys Thr Pro Gln Gly Gly Arg Asn Trp Glu Gly

145 150 155 160

Phe Gly Val Asp Pro Tyr Leu Thr Gly Ile Ala Met Gly Gln Thr Ile

165 170 175

Asn Gly Ile Gln Ser Val Gly Val Gln Ala Thr Ala Lys His Tyr Ile

180 185 190

Leu Asn Glu Gln Glu Leu Asn Arg Glu Thr Ile Ser Ser Asn Pro Asp

195 200 205

Asp Arg Thr Leu His Glu Leu Tyr Thr Trp Pro Phe Ala Asp Ala Val



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

Phe Ile Thr Ala Asp Ser Gly Glu Gly Tyr Ile Thr Val Glu Gly Asn  
 465 470 475 480  
 Ala Gly Asp Arg Asn Asn Leu Asp Pro Trp His Asn Gly Asn Ala Leu  
 485 490 495  
 Val Gln Ala Val Ala Gly Ala Asn Ser Asn Val Ile Val Val Val His  
 500 505 510  
 Ser Val Gly Ala Ile Ile Leu Glu Gln Ile Leu Ala Leu Pro Gln Val  
 515 520 525  
 Lys Ala Val Val Trp Ala Gly Leu Pro Ser Gln Glu Ser Gly Asn Ala  
 530 535 540  
 Leu Val Asp Val Leu Trp Gly Asp Val Ser Pro Ser Gly Lys Leu Val  
 545 550 555 560  
 Tyr Thr Ile Ala Lys Ser Pro Asn Asp Tyr Asn Thr Arg Ile Val Ser  
 565 570 575  
 Gly Gly Ser Asp Ser Phe Ser Glu Gly Leu Phe Ile Asp Tyr Lys His  
 580 585 590  
 Phe Asp Asp Ala Asn Ile Thr Pro Arg Tyr Glu Phe Gly Tyr Gly Leu  
 595 600 605  
 Ser Tyr Thr Lys Phe Asn Tyr Ser Arg Leu Ser Val Leu Ser Thr Ala  
 610 615 620  
 Lys Ser Gly Pro Ala Thr Gly Ala Val Val Pro Gly Gly Pro Ser Asp  
 625 630 635 640  
 Leu Phe Gln Asn Val Ala Thr Val Thr Val Asp Ile Ala Asn Ser Gly  
 645 650 655  
 Gln Val Thr Gly Ala Glu Val Ala Gln Leu Tyr Ile Thr Tyr Pro Ser  
 660 665 670  
 Ser Ala Pro Arg Thr Pro Pro Lys Gln Leu Arg Gly Phe Ala Lys Leu  
 675 680 685  
 Asn Leu Thr Pro Gly Gln Ser Gly Thr Ala Thr Phe Asn Ile Arg Arg  
 690 695 700  
 Arg Asp Leu Ser Tyr Trp Asp Thr Ala Ser Gln Lys Trp Val Val Pro  
 705 710 715 720

Ser Gly Ser Phe Gly Ile Ser Val Gly Ala Ser Ser Arg Asp Ile Arg

725

730

735

Leu Thr Ser Thr Leu Ser Val Ala

740

<210> 4

<211> 2235

<212> DNA

<213> Trichoderma reesei

<400> 4

atgcgttacc gaacagcagc tgcgctggca cttgccactg ggccctttgc tagggcagac 60

agtcactcaa catcgggggc ctcggctgag gcagttgtac ctctgcagg gactccatgg 120

ggaaccgcgt acgacaaggc gaaggccgca ttggcaaagc tcaatctcca agataaggtc 180

ggcatcgtga gcggtgtcgg ctggaacggc ggtccttgcg ttggaaacac atctccggcc 240

tccaagatca gctatccatc gctatgcctt caagacggac cctcgggtgt tcgatactcg 300

acaggcagca cagcctttac gccgggcgtt caagcggcct cgacgtggga tgtcaatttg 360

atccgcgaac gtggacagtt catcggtagag gaggtgaagg cctcggggat tcatgtcata 420

cttggctcgt tggtctgggc gctgggaaag actccgcagg gcggtcgcaa ctgggagggc 480

ttcgggtgtc atccatatct cacgggcatt gccatgggtc aaaccatcaa cggcatccag 540

tcggtaggcg tgcaggcgac agcgaagcac tatatcctca acgagcagga gctcaatcga 600

gaaaccattt cgagcaaccc agatgaccga actctccatg agctgtatac ttggccattt 660

gccgacgcgg ttcaggccaa tgctgcttct gtcattgtct cgtacaacaa ggtcaatacc 720

acctgggcct gcgaggatca gtacacgtg cagactgtgc tgaaagacca gctggggttc 780

ccaggctatg tcatgacgga ctggaacgca cagcacacga ctgtccaaag cgcaattct 840

gggcttgaca tgtcaatgcc tggcacagac ttcaacggta acaatcggct ctgggggtcca 900

gtcttcacca atgcggtaaa tagcaatcag gtccccacga gcagagtcga cgatatggtg 960

actcgtatcc tcgccgatg gtacttgaca ggccaggacc aggcaggcta tccgtcgttc 1020

aacatcagca gaaatgttca aggaaaccac aagaccaatg tcagggcaat tgccaggggc 1080

ggcatcgttc tgetcaagaa tgacccaac atcctgccgc tcaagaagcc cgctagcatt 1140

gccgtcgttg gatctgccgc aatcattggt aaccacgcca gaaactcgcc ctctgcaac 1200

gacaaaggct gcgacgacgg ggccttgggc atgggttggg gttccggcgc cgtcaactat 1260

ccgtacttcg tcgccccta cgatgccatc aataccagag cgtcttcgca gggcacccag 1320

gttaccttga gcaacaccga caacacgtcc tcaggcgcat ctgcagcaag aggaaaggac 1380  
gtcgccatcg tcttcatcac cgccgactcg ggtgaaggct acatcacctg ggagggaac 1440  
gcggcgatc gcaacaacct ggatccgtgg cacaacggca atgccctggt ccaggcggtg 1500  
gccggtgcc aacagcaact cattgttgtt gtccactccg ttggcgccat cattctggag 1560

cagattcttg ctcttcgca ggtcaaggcc gttgtctggg cgggtcttcc ttctcaggag 1620  
agcggcaatg cgctcgtaga cgtgctgtgg ggagatgtca gcccttctgg caagctggtg 1680  
tacaccattg cgaagagccc caatgactat aacactcgca tcgtttccgg cggcagtgac 1740  
agcttcagcg agggactgtt catcgactat aagcacttcg acgacgcaa tatcacgccg 1800  
cggtagcagt tcggctatgg actgtcttac accaagtcca actactcac cctctccgtc 1860  
ttgtcgaccg ccaagtctgg tcttcgcact ggggccgttg tgccgggagg cccgagtgat 1920  
ctgttcaga atgtcgagc agtcaccgtt gacatcgcaa actctggcca agtgactggt 1980

gccgaggtag cccagctgta catcacctac ccatcttcag caccaggac ccctccgaag 2040  
cagctgcgag gctttgcaa gctgaacctc acgctgggc agagcggaac agcaacgttc 2100  
aacatccgac gacgagatct cagctactgg gacacggctt cgcagaaatg ggtggtgccg 2160  
tcggggtcgt ttggcatcag cgtgggagcg agcagccggg atatcaggct gacgagcact 2220  
ctgtcggtag cgtag 2235

<210> 5

<211> 744

<212> PRT

<213> Trichoderma reesei

<400> 5

Met Arg Tyr Arg Thr Ala Ala Ala Leu Ala Leu Ala Thr Gly Pro Phe

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

Ser Lys Ile Ser Tyr Pro Ser Leu Cys Leu Gln Asp Gly Pro Leu Gly  
 85 90 95  
 Ile Arg Phe Gly Thr Gly Ser Thr Ala Phe Thr Pro Gly Val Gln Ala  
 100 105 110  
 Ala Ser Thr Trp Asp Thr Glu Leu Met Arg Gln Arg Gly Glu Tyr Leu  
 115 120 125  
 Gly Ala Glu Ala Lys Gly Cys Gly Ile His Val Leu Leu Gly Pro Val  
 130 135 140  
 Ala Gly Pro Leu Gly Lys Thr Pro Gln Gly Gly Arg Asn Trp Glu Gly  
 145 150 155 160  
 Phe Gly Val Asp Pro Tyr Leu Thr Gly Ile Ala Met Ala Glu Thr Ile  
 165 170 175  
 Glu Gly Leu Gln Ser Ala Gly Val Gln Ala Cys Ala Lys His Tyr Ile  
 180 185 190  
 Val Asn Glu Gln Glu Leu Asn Arg Glu Thr Ile Ser Ser Asp Val Asp  
 195 200 205  
 Asp Arg Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val  
 210 215 220  
 Gln Ala Asn Val Ala Ser Val Met Cys Ser Tyr Asn Lys Ile Asn Gly  
 225 230 235 240  
 Ser Trp Ala Cys Glu Asp Gln Tyr Thr Leu Gln Thr Val Leu Lys Asp  
 245 250 255  
 Gln Leu Gly Phe Pro Gly Tyr Val Met Thr Asp Trp Asn Ala Gln His  
 260 265 270  
 Thr Thr Val Gln Ser Ala Asn Ser Gly Leu Asp Met Ser Met Pro Gly  
 275 280 285  
 Thr Asp Phe Asn Gly Asn Asn Arg Leu Trp Gly Pro Ala Leu Thr Asn  
 290 295 300  
 Ala Val Asn Ser Asn Gln Val Pro Thr Ser Arg Val Asp Asp Met Val  
 305 310 315 320  
 Thr Arg Ile Leu Ala Ala Trp Tyr Leu Thr Gly Gln Asp Gln Ala Gly  
 325 330 335

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

580 585 590  
 Phe Asp Asp Ala Asn Ile Thr Pro Arg Tyr Glu Phe Gly Tyr Gly Leu  
 595 600 605  
 Ser Tyr Thr Lys Phe Asn Tyr Ser Arg Leu Ser Val Phe Ser Thr Ala  
 610 615 620  
 Lys Ser Gly Pro Ala Thr Gly Ala Val Val Pro Gly Gly Pro Ser Asp  
 625 630 635 640  
 Leu Phe Gln Asn Val Ala Thr Val Thr Val Asp Ile Ala Asn Ser Gly  
 645 650 655  
  
 Gln Val Thr Gly Ala Glu Val Ala Gln Leu Tyr Ile Thr Tyr Pro Ser  
 660 665 670  
 Ser Ala Pro Arg Thr Pro Pro Lys Gln Leu Arg Gly Phe Ala Lys Leu  
 675 680 685  
 Asn Leu Thr Pro Gly Gln Ser Gly Thr Ala Thr Phe Asn Ile Arg Arg  
 690 695 700  
 Arg Asp Leu Ser Tyr Trp Asp Thr Ala Ser Gln Lys Trp Val Val Pro  
 705 710 715 720  
 Ser Gly Ser Phe Gly Ile Ser Val Gly Ala Ser Ser Arg Asp Ile Arg

725 730 735  
 Leu Thr Ser Thr Leu Ser Val Ala  
 740

<210> 6

<211> 2235

<212> DNA

<213> Trichoderma reesei

<400> 6

atgcgttacc gaacagcagc tgcgctggca ctgcccactg ggccctttgc tagggcagac 60

agtcactcaa catcgggggc ctcggtgag gcagttgtac ctctgcagg gactccatgg 120

ggaaccgcgt acgacaaggc gaaggccgca ttggcaaagc tcaatctcca agataaggtc 180

ggcatcgtga gcggtgtcgg ctggaacggc ggtccttgcg ttggaaacac atctccggcc 240

tccaagatca gctatccatc gctatgcctt caagacggac ccctcggtat ccgattcggc 300

acaggcagca cagcctttac gccgggcgtt caagcggcct cgacgtggga taccgagttg 360



atgcgccagc gtggagagta cctgggtgcc gaggccaagg gctgcgggat tcatgtcctg	420
cttggctcctg tggctgggcc gctgggaaag actccgcagg gcggtcgcaa ctgggagggc	480
ttcgggtgctg atccatatct caccggcatt gccatggccg agacaatcga gggcctgcag	540
tcggccggcg tgcaggcgtg cgcaagcac tatatctca acgagcagga gctcaatcga	600
gaaaccattt cgagcgacgt cgatgaccga actatgcatg agctgtatct gtggccattt	660
gccgacgagg ttcaggccaa tgtcgtttct gtcattgtct cgtacaaca gatcaatggc	720
agctgggcct gcgaggatca gtacacgtg cagactgtgc tgaaagacca gctggggttc	780
ccaggctatg tcatgacgga ctggaacgca cagcacacga ctgtccaaag cgcaattct	840
gggcttgaca tgtaaatgcc tggcacagac ttcaacggta acaatcggct ctggggcca	900
gctctacca atgcggtaaa tagcaatcag gtccccacga gcagagtcga cgatatggtg	960
actcgtatcc tcgcccatg gtacttgaca ggccaggacc aggcaggcta tccgtcgttc	1020
aacatcagca gaaatgttca aggaaccac aagaccaatg tcagggcaat tgccagggac	1080
ggcatcgttc tgcacaagaa tgacccaac atcctgccgc tcaagaagcc cgctagcatt	1140
gccgtcgttg gatctgccgc aatcatttgt aaccacgcca gaaactgcc ctcgtgcaac	1200
gacaaaggct gcgacgacgg ggccttgggc atgggttggg gttccggcgc cgtcaactat	1260
ccgtacttcg tcgcgccta cgatgccatc aataccagag cgtcttcga gggcaccag	1320
gttaccttga gcaacacga caacacgtcc tcaggcgcat ctgcagcaag aggaaaggac	1380
gtcgccatcg tcttcacac cgccgactcg ggtgaaggct acatcacctg ggagggcaac	1440
gcggcgcatc gcaacaacct ggatccgtgg cacaacggca atgccctggt ccaggcggtg	1500
gccggtgcca acagcaacgt cattgttgtt gtccactccg ttggcgccat cattctggag	1560
cagattcttg ctcttcgca ggtcaaggcc gttgtctggg cgggtcttcc ttctcaggag	1620
agcggcaatg cgctcgtga cgtgctgtgg ggagatgtca gcccttctgg caagctggtg	1680
tacaccattg cgaagagccc caatgactat aacactcga tcgtttccgg cggcagtgac	1740
agcttcagcg agggactgtt catcgactat aagcacttcg acgacgcaa tatcacccg	1800
cggtagcagt tcggctatgg actgtcttac accaagtca actactcag cctctccgtc	1860
ttttcgaccg ccaagtctgg tcttcgact ggggccgttg tgccgggagg cccgagtgat	1920
ctgttcaga atgtcgcgac agtcaccgtt gacatcgcaa actctggcca agtgactggt	1980
gccgaggtag cccagctgta catcacctac ccatcttcag caccaggac ccctccaag	2040
cagctgcgag gctttgcaa gctgaacctc acgctggtc agagcgaac agcaacgttc	2100
aacatccgac gacgagatct cagctactgg gacacggctt cgcaaaaatg ggtggtgccg	2160
tcggggtcgt ttggcatcag cgtgggagcg agcagccggg atatcaggct gacgagcact	2220

ctgtcggtag cgtag

2235

<210> 7

<211> 726

<212> PRT

<213> Chaetomium globosum

<400> 7

Met Thr Thr Leu Arg Asn Phe Ala Leu Leu Ala Ala Ala Val Leu Ala

1 5 10 15

Arg Val Glu Ala Leu Glu Ala Ala Asp Trp Ala Ala Ala Glu Ala Ser

20 25 30

Ala Lys Thr Ala Leu Ala Lys Met Ser Gln Gln Asp Lys Ile Ser Ile

35 40 45

Val Thr Gly Ile Gly Trp Asp Lys Gly Pro Cys Val Gly Asn Thr Ala

50 55 60

Ala Ile Asn Ser Ile Asn Tyr Pro Gln Leu Cys Leu Gln Asp Gly Pro

65 70 75 80

Leu Gly Ile Arg Phe Gly Thr Gly Ser Thr Ala Phe Thr Pro Gly Val

85 90 95

Gln Ala Ala Ser Thr Trp Asp Thr Glu Leu Met Arg Gln Arg Gly Glu

100 105 110

Tyr Leu Gly Ala Glu Ala Lys Gly Cys Gly Ile His Val Leu Leu Gly

115 120 125

Pro Val Ala Gly Ala Leu Gly Lys Ile Pro His Gly Gly Arg Asn Trp

130 135 140

Glu Gly Phe Gly Thr Asp Pro Tyr Leu Ala Gly Ile Ala Met Ala Glu

145 150 155 160

Thr Ile Glu Gly Leu Gln Ser Ala Gly Val Gln Ala Cys Ala Lys His

165 170 175

Tyr Ile Val Asn Glu Gln Glu Leu Asn Arg Glu Thr Ile Ser Ser Asp

180 185 190

Val Asp Asp Arg Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp

195	200	205
Ala Val His Ala Asn Val Ala Ser Val Met Cys Ser Tyr Asn Lys Ile		
210	215	220
Asn Gly Ser Trp Gly Cys Glu Asn Asp His Ala Gln Asn Gly Leu Leu		
225	230	235
Lys Lys Glu Leu Gly Phe Lys Gly Tyr Val Val Ser Asp Trp Asn Ala		
245	250	255
Gln His Thr Thr Asp Gly Ala Ala Asn Asn Gly Met Asp Met Thr Met		
260	265	270
Pro Gly Ser Asp Tyr Asn Gly Asn Asn Val Leu Trp Gly Pro Gln Leu		
275	280	285
Ser Asn Ala Val Asn Ser Asn Arg Val Ser Arg Asp Arg Leu Asp Asp		
290	295	300
Met Ala Lys Arg Ile Leu Thr Ser Trp Tyr Leu Leu Gly Gln Asn Ser		
305	310	315
Gly Tyr Pro Asn Ile Asn Ile Asn Ala Asn Val Gln Gly Asn His Lys		
325	330	335
Glu Asn Val Arg Ala Val Ala Arg Asp Gly Ile Val Leu Leu Lys Asn		
340	345	350
Asp Glu Gly Val Leu Pro Leu Lys Lys Pro Gly Lys Val Ala Leu Val		
355	360	365
Gly Ser Ala Ala Ser Val Asn Ser Ala Gly Pro Asn Ala Cys Val Asp		
370	375	380
Lys Gly Cys Asn Thr Gly Ala Leu Gly Met Gly Trp Gly Ser Gly Ser		
385	390	395
Val Asn Tyr Pro Tyr Phe Val Ala Pro Tyr Asp Ala Leu Lys Thr Arg		
405	410	415
Ala Gln Ala Asp Gly Thr Thr Leu Ser Leu His Asn Ser Asp Ser Thr		
420	425	430
Asn Gly Val Ser Gly Val Val Ser Gly Ala Asp Val Ala Ile Val Val		
435	440	445

Ile Thr Ala Asp Ser Gly Glu Gly Tyr Ile Thr Val Glu Gly His Ala  
450 455 460

Gly Asp Arg Asn His Leu Asp Pro Trp His Asp Gly Asn Ala Leu Val  
465 470 475 480

Lys Ala Val Ala Ala Ala Asn Lys Asn Thr Ile Val Val Val His Ser  
485 490 495

Thr Gly Pro Ile Ile Leu Glu Thr Ile Leu Ala Thr Glu Gly Val Lys  
500 505 510

Ala Val Val Trp Ala Gly Leu Pro Ser Gln Glu Asn Gly Asn Ala Leu  
515 520 525

Val Asp Val Leu Tyr Gly Leu Thr Ser Pro Ser Gly Lys Leu Val Tyr  
530 535 540

Ser Ile Ala Lys Arg Pro Glu Asp Tyr Gly Thr Ala Pro Ser Lys Gly  
545 550 555 560

Ser Asn Asp Lys Phe Thr Glu Gly Leu Phe Val Asp Tyr Arg His Phe  
565 570 575

Asp Asn Ala Lys Ile Glu Pro Arg Tyr Glu Phe Gly Phe Gly Leu Ser  
580 585 590

Tyr Thr Glu Phe Thr Tyr Ala Asp Leu Ser Val Thr Ser Thr Val Thr  
595 600 605

Ala Gly Pro Ala Ser Gly Glu Thr Ile Pro Gly Gly Ala Ala Asp Leu  
610 615 620

Trp Glu Thr Val Ala Thr Val Thr Ala Ser Ile Thr Asn Ser Gly Glu  
625 630 635 640

Val Glu Gly Ala Glu Val Ala Gln Leu Tyr Ile Thr Leu Pro Ser Ala  
645 650 655

Ala Pro Ser Thr Pro Pro Lys Gln Leu Arg Gly Phe Ala Lys Leu Lys  
660 665 670

Leu Glu Pro Gly Ala Ser Gly Val Ala Thr Phe Asn Leu Arg Arg Arg  
675 680 685

Asp Leu Ser Tyr Trp Asp Ala Gly Arg Gly Gln Trp Val Val Pro Ala  
690 695 700

Gly Glu Phe Thr Val Ser Val Gly Ala Ser Ser Arg Asp Val Arg Leu

705 710 715 720

Thr Gly Ser Leu Thr Ala

725

<210> 8

<211> 2181

<212> DNA

<213> Chaetomium globosum

<400> 8

atgacgacgc tccgcaactt tgcgtgctc gcagcggcgg tgcttgcgcg ggtcgaggcc	60
ctcgaggccg ccgactgggc tgcggtgag gcctcagcca aaaccgact ggcaaagatg	120
tcacaacaag acaaaatcag cattgtgacg ggcatcggct gggacaaggg tccctgtgtc	180
ggcaacacgg ccgccatcaa ctgatcaac taccgcagc tctgcctaca ggacggccc	240
ctcgggatcc gcttcggcac cggctcgacg gccttcaccc cgggcgtcca agccgcctcg	300
acatgggata ccgagctgat gcgccagcgc ggcgagtacc tcggggccga ggccaagggg	360
tgcggcatcc acgtgttgtc gggccccgtg gccggggcac tgggcaagat cccgcacggc	420
gggcgcaact gggaaggatt cgggacggac ccgtacctgg cgggcatcgc catggccgag	480
acgatcgagg ggctgcagtc ggcgggggtg caggcgtgcg ccaagcacta catcgtcaac	540
gagcaggagc tcaaccgcga gaccatcagc agcgacgtcg acgaccgcac catgcacgag	600
ctgtacctgt ggcccttcgc cgacgccgtg cagccaacg tggccagcgt catgtgcagc	660
tacaacaaga tcaacggctc gtggggctgc gagaacgacc acgccccaaa cggcctgctc	720
aagaaggagc tcggcttcaa gggttacgtc gtcagcgact ggaacgcgca gcacacgacc	780
gacggcgccg ccaacaacgg catggacatg accatgccgg gcagcgacta caacggcaac	840
aacgtgctct ggggcccgcg gtcagcaac gccgtcaaca gcaaccgggt ctgcgcgac	900
cggctcgacg acatggccaa acgcatcctc acctcatggt acctcctggg ccagaactcg	960
ggctacccca acatcaacat caacccaac gtgcagggca accacaagga gaacgtgcgg	1020
gcggtggcgc gcgacggcat cgtgtgtctc aagaacgacg agggcgtgct cccgctgaag	1080
aagccaggca aggtggctct cgtcggatcg gcggcctcgg tcaacagcgc gggccccaac	1140
gcgtgcgtcg acaagggtc caacacgggc gcgtcggca tgggtgggg gtccgggtcc	1200
gtcaactacc cctactttgt ggccccctac gacgcgtca agacgcgcgc ccaggccgac	1260
ggcaccacgc tcagcctgca caactcggac tcgaccaacg gcgtatcggg cgtggtgtcg	1320

ggcgccgacg tggccatcgt ggtgatcacg gcggactcgg gcgagggcta catcacggtc 1380  
gagggccacg ccggcgaccg caaccacctg gacccgtggc acgacggcaa cgcgctggtt 1440  
aaggcgggtg ccgcgcccaa caagaacacc atcgtggtag tgcacagcac agggcccatc 1500

atcctcgaga ccatcctggc gacggagggt gtcaaggcgg ttgtgtgggc cggcctgccg 1560  
agtcaggaga acggcaacgc gctagttagc gttttgtacg gcctgacttc gccctcaggc 1620  
aaactgggtc actccatcgc caagcgcccc gaggactatg gcacggcccc ctccaagggc 1680  
agtaacgaca agttcaccga aggcctgttt gtcgactacc ggcactttga caacgccaag 1740  
attgagccgc ggtacagatt tggctttggt ttgtcctaca ccgaattcac ctacgccgac 1800  
ctctccgtca ctccaccgt aacggccggc cccgcctcag gcgagacat acccggcggc 1860  
gcggccgacc tcgggagac tgtcgcaacg gtcacggcgt ccatcacgaa cagcggcgag 1920

gtggagggcg ccgaggtggc gcagctgtac atcacgtgc cgtcggcggc cccctcgacg 1980  
ccgccaagc agctgcgcgg gtccccaag ctcaagctcg agccgggggc gtcgggcgtc 2040  
gcgacctca acctgcgcg tcgcgatctg agttattggg atgccgggcg cggccagtgg 2100  
gtggtgccgg cgggcgagtt tacggtttcg gttggtcga gttcgaggga tgtgcgttg 2160  
acggggagct tgactgctta g 2181

<210> 9

<211> 735

<212> PRT

<213> Neurospora crassa

<400> 9

Met His Leu Arg Ile Phe Ala Val Leu Ala Ala Thr Ser Leu Ala Trp

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

Gln Leu Cys Leu Gln Asp Gly Pro Leu Gly Ile Arg Phe Gly Gly Ser

	85	90	95
Val Thr Ala Phe Thr Pro Gly Ile Gln Ala Ala Ser Thr Trp Asp Val			
	100	105	110
Glu Leu Ile Arg Gln Arg Gly Val Tyr Leu Gly Ala Glu Ala Arg Gly			
	115	120	125
Val Gly Val His Val Leu Leu Gly Pro Val Ala Gly Ala Leu Gly Lys			
	130	135	140
Ile Pro Asn Gly Gly Arg Asn Trp Glu Gly Phe Gly Pro Asp Pro Tyr			
145	150	155	160
Leu Thr Gly Ile Ala Met Ser Glu Thr Ile Glu Gly Ile Gln Ser Asn			
	165	170	175
Gly Val Gln Ala Cys Ala Lys His Phe Ile Leu Asn Glu Gln Glu Thr			
	180	185	190
Asn Arg Asp Thr Ile Ser Ser Val Val Asp Asp Arg Thr Met His Glu			
	195	200	205
Leu Tyr Leu Phe Pro Phe Ala Asp Ala Val His Ser Asn Val Ala Ser			
	210	215	220
Val Met Cys Ser Tyr Asn Lys Val Asn Gly Thr Trp Ala Cys Glu Asn			
225	230	235	240
Asp Lys Ile Gln Asn Gly Leu Leu Lys Lys Glu Leu Gly Phe Lys Gly			
	245	250	255
Tyr Val Met Ser Asp Trp Asn Ala Gln His Thr Thr Asn Gly Ala Ala			
	260	265	270
Asn Ser Gly Met Asp Met Thr Met Pro Gly Ser Asp Phe Asn Gly Lys			
	275	280	285
Thr Ile Leu Trp Gly Pro Gln Leu Asn Thr Ala Val Asn Asn Gly Gln			
290	295	300	
Val Ser Lys Ala Arg Leu Asp Asp Met Ala Lys Arg Ile Leu Ala Ser			
305	310	315	320
Trp Tyr Leu Leu Glu Gln Asn Ser Gly Tyr Pro Ala Thr Asn Leu Lys			
	325	330	335
Ala Asn Val Gln Gly Asn His Lys Glu Asn Val Arg Ala Val Ala Arg			



340	345	350
Asp Gly Ile Val Leu Leu Lys Asn Asp Asp Asn Ile Leu Pro Leu Lys		
355	360	365
Lys Pro Ser Lys Leu Ala Ile Ile Gly Ser Ser Ser Val Val Asn Pro		
370	375	380
Ala Gly Arg Asn Ala Cys Thr Asp Arg Gly Cys Asn Thr Gly Ala Leu		
385	390	395
Gly Met Gly Trp Gly Ser Gly Thr Ala Asp Tyr Pro Tyr Phe Val Ala		
405	410	415
Pro Tyr Asp Ala Leu Lys Thr Arg Ala Gln Ser Asp Gly Thr Thr Val		
420	425	430
Asn Leu Leu Ser Ser Asp Ser Thr Ser Gly Val Ala Asn Ala Ala Ser		
435	440	445
Gly Ala Asp Ala Ala Leu Val Phe Ile Thr Ala Asp Ser Gly Glu Gly		
450	455	460
Tyr Ile Thr Val Glu Gly Val Thr Gly Asp Arg Pro Asn Leu Asp Pro		
465	470	475
Trp His Asn Gly Asn Gln Leu Val Gln Ala Val Ala Gln Ala Asn Lys		
485	490	495
Asn Thr Ile Val Val Val His Ser Thr Gly Pro Ile Ile Leu Glu Thr		
500	505	510
Ile Leu Ala Gln Pro Gly Val Lys Ala Val Val Trp Ala Gly Leu Pro		
515	520	525
Ser Gln Glu Asn Gly Asn Ala Leu Val Asp Val Leu Tyr Gly Leu Val		
530	535	540
Ser Pro Ser Gly Lys Leu Pro Tyr Thr Ile Ala Lys Ser Glu Ser Asp		
545	550	555
Tyr Gly Thr Ala Val Gln Arg Gly Gly Thr Asp Leu Phe Thr Glu Gly		
565	570	575
Leu Phe Ile Asp Tyr Arg His Phe Asp Lys Asn Gly Ile Ala Pro Arg		
580	585	590

Tyr Glu Phe Gly Phe Gly Leu Ser Tyr Thr Asn Phe Thr Tyr Ser Ser  
595 600 605  
Leu Ser Ile Thr Ser Thr Ala Ser Ser Gly Pro Ala Ser Gly Asp Thr  
610 615 620  
Ile Pro Gly Gly Arg Ala Asp Leu Trp Glu Thr Val Ala Thr Val Thr  
625 630 635 640  
Ala Val Val Lys Asn Thr Gly Gly Val Gln Gly Ala Glu Ala Pro Gln  
645 650 655  
  
Leu Tyr Ile Thr Leu Pro Ser Ser Ala Pro Ser Ser Pro Pro Lys Gln  
660 665 670  
Leu Arg Gly Phe Ala Lys Leu Lys Leu Ala Pro Gly Glu Ser Lys Thr  
675 680 685  
Ala Thr Phe Ile Leu Arg Arg Arg Asp Leu Ser Tyr Trp Asp Thr Gly  
690 695 700  
Ser Gln Asn Trp Val Val Pro Ser Gly Ser Phe Gly Val Val Val Gly  
705 710 715 720  
Ala Ser Ser Arg Asp Leu Arg Leu Asn Gly Lys Phe Asp Val Tyr

725 730 735

<210> 10

<211> 2208

<212> DNA

<213> Neurospora crassa

<400> 10

atgcaccttc gaatatttgc ggtgttggcc gcgacttccc tcgcctgggc cgagactagc	60
gagaaacaag ctcgtcaagc tggctcaggt tttgcggcgt gggacgcagc ctattctcag	120
gcaagcactg ctctctccaa gctttcacag caagacaagg tcaacatcgt caccggagtc	180
ggctggaata agggcccatg tgttggcaac accccagcta ttgcatcaat cggttatccc	240
cagctctgtt tacaagacgg ccctctcggc attcggtttg gaggaagtgt caccgcgttc	300
acgcctggta tccaggcggc ttcaacatgg gacgtcgaac tgattcgaca gcgcggcgtc	360
tacctcgggtg cagaagccag aggggttggc gtacatgtcc ttcttggacc cgtggccgga	420
gcgcttggca agatccccaa tgggtggacgt aactgggagg gcttttggtcc ggatccctac	480
ctcacaggta ttgcatgag cgaaacaatt gaagggatcc agagcaatgg tgtacaagct	540

tgcgccaagc acttcattct caacgaacag gagacaaacc gcgatactat cagcagtgtc 600  
 gtcgacgacc gcaccatgca tgaactatac ctcttccttt ttgccgatgc cgtacactca 660  
 aatgttgcaa gtgtgatgtg cagctacaac aaggtcaacg gtacgtgggc atgtgagaat 720  
  
 gacaaaatcc agaatggcct tctcaagaaa gagctaggct tcaaaggata tgtcatgagt 780  
 gattggaacg cccagcacac cacgaacggc gctgcaaaca gtggtatgga tatgacgatg 840  
 ccaggcagtg actttaatgg caagacgata ctgtggggac cacagctcaa caccgccgtc 900  
 aacaatggcc aggtctccaa agcaagactg gacgacatgg ccaagcgcac tctcgcatcg 960  
 tggatatctac tcgagcaaaa ctcaggctac cctgcgacta acctcaaggc caatgttcaa 1020  
 ggaaaccaca aggagaacgt tcgcgcagtg gcaagagacg gcattgttct gctgaagaac 1080  
 gacgataaca tcctcccgct caagaagcct agcaagctgg caatcattgg gtcacgtcc 1140  
  
 gtgtgcaacc ctgcgggaag gaacgcctgc accgatcgag gatgcaacac cgggtgcgtc 1200  
 ggcatgggtt ggggctccgg cacggccgat taccctact tcgtagcacc ctatgatgct 1260  
 ctcaagacgc gggctcagtc cgacggaaca actgtcaacc tactcagctc tgacagcacc 1320  
 agcggcgtag ccaacgtgc ctccggagcc gacgcggcac tagtcttcat cacagccgat 1380  
 tccggcgaag gctacatcac ggtcgagggc gtgaccggcg accgtcccaa cctcgatccc 1440  
 tggcacaacg gcaaccagct agtccaagcc gtggctcaag ccaacaagaa caccattgtc 1500  
 gtcgtccaca gtaccggccc catcattctg gagactatcc tcgcgcagcc gggcgtaag 1560  
  
 gcggtcgtgt gggccggtct cccagccaa gagaacggca acgcccttgt cgatgtccta 1620  
 tacggttgg tctctccctc gggtaagctg ccgtatacta tcgccaagag cgaaagcgac 1680  
 tacggcactg ccgtgcaaag gggagggacg gatctgttca ctgagggtct gtcatcgat 1740  
 taccgccact ttgacaagaa cggatatcgt ccccggtatg agttcggttt cggctcttcc 1800  
 tacacgaact tcacctact ctccctctcc atcacctcca ccgcctctc cggteccgcc 1860  
 tcgggtgaca ccatccctgg cgcccgccgc gacctctggg aaaccgtggc aaccgtcact 1920  
 gccgtcgta aaaacacggg tgggtgtcag ggcccgagg caccagct atacatcacc 1980  
  
 ttgccctctt ccgcgccgtc gagcccgccg aaacagctca gagggtttgc aaagctgaag 2040  
 ctggcgccc gggagagcaa gacagctacg ttcatcttgc ggaggaggga tttagttat 2100  
 tgggatacgg gcagccagaa ttgggtggtg cctagtggca gctttggggg gtagtgggt 2160  
 gctagtctga gggattttag gttgaatggg aagtttgatg tttattga 2208

<210> 11  
 <211> 744  
 <212> PRT

<213> Trichoderma reesei

<400> 11

Met Arg Tyr Arg Thr Ala Ala Ala Leu Ala Leu Ala Thr Gly Pro Phe  
1 5 10 15

Ala Arg Ala Asp Ser His Ser Thr Ser Gly Ala Ser Ala Glu Ala Val  
20 25 30

Val Pro Pro Ala Gly Thr Pro Trp Gly Thr Ala Tyr Asp Lys Ala Lys  
35 40 45

Ala Ala Leu Ala Lys Leu Asn Leu Gln Asp Lys Val Gly Ile Val Ser  
50 55 60

Gly Val Gly Trp Asn Gly Gly Pro Cys Val Gly Asn Thr Ser Pro Ala  
65 70 75 80

Ser Lys Ile Ser Tyr Pro Ser Leu Cys Leu Gln Asp Gly Pro Leu Gly  
85 90 95

Ile Arg Phe Gly Thr Gly Ser Thr Ala Phe Thr Pro Gly Val Gln Ala  
100 105 110

Ala Ser Thr Trp Asp Val Asn Leu Ile Arg Glu Arg Gly Gln Phe Ile  
115 120 125

Gly Glu Glu Val Lys Ala Ser Gly Ile His Val Ile Leu Gly Pro Val  
130 135 140

Ala Gly Pro Leu Gly Lys Ile Pro His Gly Gly Arg Asn Trp Glu Gly  
145 150 155 160

Phe Gly Val Asp Pro Tyr Leu Thr Gly Ile Ala Met Ala Glu Thr Ile  
165 170 175

Glu Gly Leu Gln Ser Ala Gly Val Gln Ala Cys Ala Lys His Tyr Ile  
180 185 190

Leu Asn Glu Gln Glu Leu Asn Arg Glu Thr Ile Ser Ser Asn Pro Asp  
195 200 205

Asp Arg Thr Leu His Glu Leu Tyr Thr Trp Pro Phe Ala Asp Ala Val  
210 215 220

His Ala Asn Val Ala Ser Val Met Cys Ser Tyr Asn Lys Ile Asn Gly

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

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

Leu Thr Ser Thr Leu Ser Val Ala

740

<210> 12

<211> 2235

<212> DNA

<213> Trichoderma reesei

<400> 12

atgcgttacc gaacagcagc tgcgtggca ctgccactg ggccctttgc tagggcagac	60
agtcactcaa catcgggggc ctcggtgag gcagttgtac ctctgcagg gactccatgg	120
ggaaccgcgt acgacaaggc gaagccgca ttggcaaagc tcaatctcca agataaggtc	180
ggcatcgtga gcggtgtcgg ctggaacggc ggtccttgcg ttggaacac atctccggcc	240
tccaagatca gctatccatc gctatgcctt caagacggac ccctcggtat ccgattcggc	300
acaggcagca cagcctttac gccgggcgtt caagcggcct cgacgtggga tgtcaatttg	360
atccgcgaac gtggacagtt catcggtgag gaggtgaagg cctcggggat tcatgtcata	420
cttggtcctg tggctgggcc gctgggaaag atccgcacg gcggtcgcaa ctgggagggc	480
ttcggtgtcg atccatact caccggcatt gccatggccg agacaatcga gggcctgcag	540
tcggccggcg tgcaggcgtg cggaagcac tatatctca acgagcagga gctcaatcga	600
gaaaccattt cgagcaacc agatgaccga actctccatg agctgtatac ttggccattt	660
gccgacgcgg ttcacgcaa tgctcgttct gtcattgtct cgtacaaca gatcaatggc	720
agctgggcct gcgaggatca gtacacgtg cagactgtgc tgaaagacca gctggggttc	780
ccaggctatg tcatgacgga ctggaacgca cagcacagca ctgtccaaag cgcgaattct	840
gggcttgaca tgtcaatgcc tggcacagac ttcaacggta acaatcggt ctggggtcca	900
gtcttcacca atgcggtaaa tagcaatcag gtccccacga gcagagtcga cgatatggtg	960
actcgtatcc tcgccgatg gtacttgaca ggccaggacc aggcaggcta tccgtcgttc	1020
aacatcagca gaaatgttca aggaaccac aagaccaatg tcagggcaat tgccaggggac	1080
ggcatcgttc tgctcaagaa tgacgccaac atcctgccgc tcaagaagcc cgctagcatt	1140
gccgtcgttg gatctgccgc aatcattggt aaccacgcca gaaactcgcc ctctgcaac	1200
gacaaaggct gcgacgacgg ggccttgggc atgggttggg gttccggcgc cgtcaactat	1260
ccgtacttcg tcgcgccta cgatgccatc aataccagag cgtcttcgca gggcacccag	1320
gttaccttga gcaaaccca caacacgtcc tcaggcgcgt ctgcagcaag aggaaaggac	1380
gtcgccatcg tcttcatcac cgccgactcg ggtgaaggct acatcacctg ggagggcaac	1440
gcgggcgatc gcaacaacct ggatccgtgg cacaacggca atgccctggt ccaggcggtg	1500

gccggtgccacacgaacgt cattgttggt gtccactccg ttggcgccat cattctggag	1560
cagattcttg ctcttcgca ggtcaaggcc gttgtctggg cgggtcttcc ttctcaggag	1620
agcggcaatg cgctcgtcga cgtgctgtgg ggagatgtca gcccttctgg caagctggtg	1680
tacaccattg cgaagagccc caatgactat aacactcgca tcgtttccgg cggcagtgac	1740
agcttcagcg agggactgtt catcgactat aagcacttcg acgacgcaa taccacgccg	1800
cggtagcagt tcggctatgg actgtcttac accaagtcca actactcag cctctccgtc	1860
ttgtcgaccg ccaagtctgg tctgcgact ggggccgttg tgccgggagg cccgagtgat	1920
ctgttcaga atgtcgcgac agtcaccgtt gacatcgcaa actctggcca agtgactggt	1980
gccgaggtag ccagctgta catcacctac ccatcttcag caccaggac ccctccgaag	2040
cagctgcgag gctttgcaa gctgaacctc acgctggtc agagcggaac agcaacgttc	2100
aacatccgac gacgagatct cagctactgg gacacggctt cgcagaaatg ggtggtgccg	2160
tcggggtcgt ttggcatcag cgtgggagcg agcagccggg atatcaggct gacgagcact	2220
ctgtcggtag cgtag	2235