



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0079676
(43) 공개일자 2015년07월08일

(51) 국제특허분류(Int. Cl.)
C12N 9/42 (2006.01) C12N 15/63 (2006.01)
(52) CPC특허분류
C12N 9/2445 (2013.01)
C12N 15/63 (2013.01)
(21) 출원번호 10-2015-7011633
(22) 출원일자(국제) 2013년10월30일
심사청구일자 없음
(85) 번역문제출일자 2015년05월01일
(86) 국제출원번호 PCT/US2013/067413
(87) 국제공개번호 WO 2014/070837
국제공개일자 2014년05월08일
(30) 우선권주장
61/720,697 2012년10월31일 미국(US)

(71) 출원인
다니스코 유에스 인크.
미합중국 캘리포니아 (우편번호 94304) 팔로 알토
페이지 밀 로드 925
(72) 발명자
바우어 벤자민 에스
미국 94304 캘리포니아주 팔로 알토 페이지 밀 로
드 925 다니스코 유에스 인크.
푸즈달라 메리디스 케이
미국 94304 캘리포니아주 팔로 알토 페이지 밀 로
드 925 다니스코 유에스 인크.
(74) 대리인
특허법인코리아나

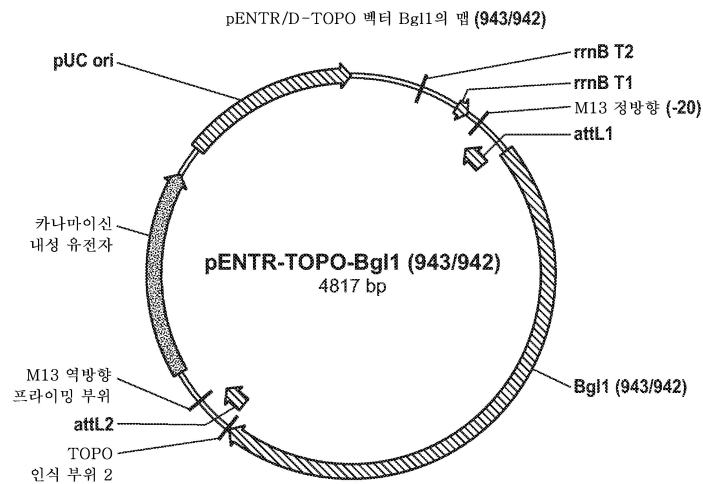
전체 청구항 수 : 총 22 항

(54) 발명의 명칭 마그나포르테 그리세아로부터의 베타-글루코시다제

(57) 요약

본 조성물 및 방법은 마그나포르테 그리세아로부터의 베타-글루코시다제, 상기 베타-글루코시다제를 암호화하는 폴리뉴클레오타이드, 및 그의 제조 및/또는 사용 방법에 관한 것이다. 베타-글루코시다제를 함유하는 제제는 리그노셀룰로오스계 바이오매스 기질을 가수분해하는데 사용하기에 적합하다.

대표도 - 도1



(52) CPC특허분류

C12Y 302/01004 (2013.01)

C12Y 302/01008 (2013.01)

C12Y 302/01021 (2013.01)

C12Y 302/01037 (2013.01)

C12Y 302/01055 (2013.01)

명세서

청구범위

청구항 1

서열 번호 2 또는 서열 번호 3의 아미노산 서열과 75% 이상 동일한 아미노산 서열을 포함하는 재조합 폴리펩티드로서, 베타-글루코시다제 활성을 갖는 재조합 폴리펩티드.

청구항 2

제1항에 있어서, 재조합 폴리펩티드 및 *트리코테르마 레에세이* Bgl1을 사용하여 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 경우 *트리코테르마 레에세이* Bgl1과 비교하여 개선된 베타-글루코시다제 활성을 갖는 재조합 폴리펩티드.

청구항 3

제1항 또는 제2항에 있어서, 개선된 베타-글루코시다제 활성이 증가된 셀로비아제 활성인, 재조합 폴리펩티드.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 개선된 베타-글루코시다제 활성이 동일한 당화 조건하에 리그노셀룰로오스계 바이오매스로부터 글루코오스의 증가된 수율인, 재조합 폴리펩티드.

청구항 5

제4항에 있어서, 리그노셀룰로오스계 바이오매스가 당화전에 전처리에 적용되는, 재조합 폴리펩티드.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 서열 번호 2 또는 서열 번호 3의 아미노산 서열과 80% 이상 동일한 아미노산 서열을 포함하는 재조합 폴리펩티드.

청구항 7

제1항 내지 제5항 중 어느 한 항에 있어서, 서열 번호 2 또는 서열 번호 3의 아미노산 서열과 90% 이상 동일한 아미노산 서열을 포함하는 재조합 폴리펩티드.

청구항 8

하나 이상의 다른 셀룰라제를 추가로 포함하는, 제1항 내지 제7항 중 어느 한 항의 재조합 폴리펩티드를 포함하는 조성물.

청구항 9

제8항에 있어서, 하나 이상의 다른 셀룰라제가, 전혀 없거나 하나 이상의 다른 베타-글루코시다제, 하나 이상의 셀로비오히드롤라제, 및 하나 이상의 엔도글루카나제로부터 선택되는 조성물.

청구항 10

하나 이상의 다른 헤미셀룰라제를 추가로 포함하는, 제1항 내지 제9항 중 어느 한 항의 재조합 폴리펩티드를 포함하는 조성물.

청구항 11

제8항 또는 제9항에 있어서, 하나 이상의 헤미셀룰라제를 추가로 포함하는 조성물.

청구항 12

제10항 또는 제11항에 있어서, 하나 이상의 헤미셀룰라제가 하나 이상의 자일라나제, 하나 이상의 베타-자일로

시다제, 및 하나 이상의 L-아라비노푸라노시다제로부터 선택되는 조성물.

청구항 13

제1항 내지 제7항 중 어느 한 항의 재조합 폴리펩티드를 암호화하는 핵산.

청구항 14

제13항에 있어서, 폴리펩티드가 신호 펩티드 서열을 추가로 포함하는 핵산.

청구항 15

제14항에 있어서, 신호 펩티드 서열이 서열 번호 13-42로 이루어진 군으로부터 선택되는 핵산.

청구항 16

조절 서열과 작동가능하게 조합된 제13항 내지 제15항 중 어느 한 항의 핵산을 포함하는 발현 벡터.

청구항 17

제16항의 발현 벡터를 포함하는 숙주 세포.

청구항 18

제17항에 있어서, 세균 세포 또는 진균 세포인 숙주 세포.

청구항 19

제17항 또는 제18항의 숙주 세포 및 배양 배지를 포함하는 조성물.

청구항 20

베타-글루코시다제를 생성하기에 적합한 조건하에, 배양 배지에서 제17항 또는 제18항의 숙주 세포를 배양하는 것을 포함하는, 베타-글루코시다제의 생성 방법.

청구항 21

배양 배지의 상청액 중 제20항의 방법에 따라 생성된 베타-글루코시다제를 포함하는 조성물.

청구항 22

리그노셀룰로오스계 바이오매스 기질을 제1항 내지 제7항 중 어느 한 항의 폴리펩티드, 또는 제21항의 조성물과 접촉시켜, 글루코오스 또는 다른 당을 수득하는 것을 포함하는, 리그노셀룰로오스계 바이오매스 기질의 가수분해 방법.

발명의 설명

기술 분야

[0001]

우선권

[0002]

본 출원은 2012년 10월 31일에 출원된 미국 가출원 제61/720,697호의 우선권을 주장하며, 이 가출원은 그 전문이 본원에 참조로 포함된다.

[0003]

본 조성물 및 방법은 *마그나포르테 그리세아*(*Magnaporthe grisea*)로부터 수득가능한 베타-글루코시다제 폴리펩티드, 상기 베타-글루코시다제 폴리펩티드를 암호화하는 폴리뉴클레오티드, 및 그의 제조 및 사용 방법에 관한 것이다. 상기 베타-글루코시다제 폴리펩티드를 포함하는 제제 및 조성물은 리그노셀룰로오스계 바이오매스를 분해 또는 가수분해하는데 유용하다.

배경 기술

[0004]

셀룰로오스 및 헤미셀룰로오스는 광합성에 의해 생성되는 가장 풍부한 식물성 물질이다. 이들은 중합체성 기질을 단량체성 당으로 가수분해할 수 있는 세포의 효소를 생성하는 수많은 미생물 (예를 들어, 세균, 효모 및 진

균)에 의해 에너지원으로서 분해되고 사용될 수 있다 (문헌 [Aro *et al.*, (2001) J. Biol. Chem., 276: 24309-24314]). 재생산가능하지 않은 자원 접근법에 대한 한계로 인해, 셀룰로오스가 주요한 재생가능 에너지 자원이 될 가능성은 막대하다 (문헌 [Krishna *et al.*, (2001) Bioresource Tech., 77: 193-196]). 생물학적 공정을 통한 셀룰로오스의 효과적 이용은 식품, 사료 및 연료의 부족을 극복하기 위한 하나의 접근법이다 (문헌 [Ohmiya *et al.*, (1997) Biotechnol. Gen. Engineer Rev., 14: 365-414]).

[0005]

셀룰라제는 셀룰로오스 (베타-1,4-글루칸 또는 베타 D-글루코시드 결합 포함)를 가수분해하여 글루코오스, 셀로비오스, 셀로올리고당 등의 형성을 야기하는 효소이다. 셀룰라제는 통상적으로 3가지 주요 부류로 나뉘어져 있다: 엔도글루카나제 (EC 3.2.1.4) ("EG"), 엑소글루카나제 또는 셀로비오히드롤라제 (EC 3.2.1.91) ("CBH") 및 베타-글루코시다제 ([베타]-D-글루코시드 글루코히드롤라제; EC 3.2.1.21) ("BG") (문헌 [Knowles *et al.*, (1987) TIBTECH 5: 255-261]; 및 [Schulein, (1998) Methods Enzymol., 160: 234-243]). 엔도글루카나제는 셀룰로오스 섬유의 비정질 부분에 주로 작용하는 한편, 셀로비오히드롤라제는 또한 결정질 셀룰로오스를 분해할 수 있다 (문헌 [Nevalainen and Penttila, (1995) Mycota, 303-319]). 따라서, 셀룰라제 시스템 내의 셀로비오히드롤라제의 존재는 결정질 셀룰로오스의 효율적인 가용화를 위해 필요하다 (문헌 [Suurnakki *et al.*, (2000) 셀룰로오스, 7: 189-209]). 베타-글루코시다제는 셀로비오스, 셀로-올리고당 및 다른 글루코시드로부터 D-글루코오스 단위를 유리시키는 작용을 한다 (문헌 [Freer, (1993) J. Biol. Chem., 268: 9337-9342]).

[0006]

셀룰라제는 수많은 세균, 효모 및 진균에 의해 생성되는 것으로 알려져 있다. 특정 진균은 셀룰로오스의 결정질 형태를 분해할 수 있는 완전한 셀룰라제 시스템을 생성한다. 이들 진균은 발효되어 일련의 셀룰라제 또는 셀룰라제 혼합물을 생성할 수 있다. 동일 진균 및 다른 진균은 또한 조작되어(engineered) 특정 셀룰라제를 생성하거나 과잉생성하여, 셀룰라제의 상이한 유형 또는 분율을 포함하는 셀룰라제의 혼합물을 야기할 수 있다. 진균은 또한, 이들이 발효를 통해 대량으로 다양한 셀룰라제를 생성할 수 있도록 조작될 수 있다. *사카로마이세스 세레비시아에(Saccharomyces cerevisiae)*와 같은 많은 효모가 그의 천연 상태에서 셀룰로오스를 가수분해하는 능력이 결여되어 있기 때문에, 사상균(filamentous fungi)은 특수 역할을 한다 (예를 들어, 문헌 [Wood *et al.*, (1998) Methods in Enzymology, 160: 87-116] 참조).

[0007]

CBH, EG 및 BG의 진균성 셀룰라제 분류는 각각의 분류 내에 다수의 성분이 포함되도록 추가로 확장될 수 있다. 예를 들어 다수의 CBH, EG 및 BG가, 2개의 CBH, 즉, CBH I ("CBH1") 및 CBH II ("CBH2"), 8개 이상의 EG, 즉, EG I, EG II, EG III, EGIV, EGV, EGVII 및 EGVIII, 및 5개 이상의 BG, 즉, BG1, BG2, BG3, BG4, BG5 및 BG7에 대해 알려진 유전자를 함유하는, *트리코데르마 레에세이(Trichoderma reesei)* (또한 *히포크레아 제코리나(Hypocrea jecorina)*)로서 지칭됨)를 포함한 다양한 진균 공급원으로부터 단리된 바 있다 (문헌 [Foreman *et al.* (2003), J. Biol. Chem. 278(34):31988-31997]). EGV, EGVII 및 EGVIII은 또한 자일로글루카나제 활성을 갖는다.

[0008]

결정질 셀룰로오스를 글루코오스로 효율적으로 전환하기 위해, 결정질 셀룰로오스를 가수분해하는데 있어서 덜 효과적인 단리된 성분과 함께, CBH, EG 및 BG 분류 각각으로부터의 성분을 포함하는 완전한 셀룰라제 시스템이 필요하다 (문헌 [Filho *et al.*, (1996) Can. J. Microbiol., 42:1-5]). 엔도-1,4-베타-글루카나제 (EG) 및 엑소-셀로비오히드롤라제 (CBH)는 셀룰로오스의 셀로올리고당 (주 생성물로서 셀로비오스)으로의 가수분해를 촉매하며, 한편 베타-글루코시다제 (BGL)는 상기 올리고당을 글루코오스로 전환시킨다. 상이한 분류로부터의 셀룰라제 성분 사이에서 상승 작용적인 관계가 관찰되었다. 특히, EG-유형 셀룰라제 및 CBH-유형 셀룰라제가 상승 작용적으로 상호 작용하여 효율적으로 셀룰로오스를 분해한다. 베타-글루코시다제는 셀로비오스 (이는 당을 에탄올로 발효시키는데 사용되는, 예를 들어, 효모 등과 같은 미생물에 유해하며; 또한 엔도글루카나제 및 셀로비오히드롤라제의 활성에 대해 억제성이어서, 이들이 결정질 셀룰로오스를 추가로 가수분해하는데 비효과적이 되게 한다)와 같은 셀로올리고당으로부터 글루코오스를 유리시키는 중요한 역할을 한다.

[0009]

셀룰로오스계 물질의 분해 또는 전환에서 베타-글루코시다제에 의해 행해진 중요한 역할의 관점에서, 셀룰로오스 원료(cellulosic feedstock)를 가수분해하는 개선된 효능 또는 능력을 갖는 베타-글루코시다제 상동체의 발견, 특성화, 제조, 및 적용이 바람직하며 유리하다.

[0010]

발명의 개요

[0011]

마그나포르테 그리세아로부터 수득가능한 베타-글루코시다제 및 그의 용도

[0012]

셀룰로오스의 효소적 가수분해는 물질의 리그노셀룰로오스계 바이오매스 원료 (이는 셀룰로오스계 당 및/또는 하류 제품(downstream product)일 수 있다)로부터 생물학적 생산의 주된 제한적 단계 중 하나로 남아 있다. 베

타-글루코시다제는 억제성 셀로비오스로부터 글루코오스를 방출하는, 그 공정의 마지막 단계를 촉매하는 중요한 역할을 하며, 따라서 그의 활성 및 효능은 효소적 리그노셀룰로오스계 바이오매스 전환의 전반적인 효능에 직접 기여하고, 결과적으로 효소 용액의 사용에서의 비용에 기여한다. 따라서 신규하고 더 효과적인 베타-글루코시다제를 발견하고, 제조하고 사용하는데 큰 관심이 존재한다.

[0013]

트리코테르마 레에세이 또는 히포크레아 제코리나 (문헌 [Korotkova O.G. et al., (2009) Biochemistry 74:569-577]; [Chauve, M. et al., (2010) Biotechnol. Biofuels 3:3-3])로부터의 베타-글루코시다제 Bgl1, Bgl3, Bgl5, Bgl7 등, 후미콜라 그리세아 변종 썬모이데아(*Humicola grisea* var. *thermoidea*) (문헌 [Nascimento, C.V. et al., (2010) J. Microbiol. 48, 53-62])로부터; 스포로트리쿰 풀베룰렌툼(*Sporotrichum pulverulentum*) (문헌 [Deshpande V. et al., (1988) Methods Enzymol., 160:415-424])으로부터; 아스페르길루스 오리자에(*Aspergillus oryzae*) (문헌 [Fukuda T. et al, (2007) Appl. Microbiol. Biotechnol. 76:1027-1033])로부터, 탈라로마이세스 썬모필루스(*Talaromyces thermophilus*) CBS 236.58 (문헌 [Nakkharat P. et al., (2006) J. Biotechnol., 123:304-313])로부터, 탈라로마이세스 에메르소니이(*Talaromyces emersonii*) (문헌 [Murray P., et al, (2004) Protein Expr. Purif. 38:248-257])로부터의 베타-글루코시다제를 포함한 다수의 베타-글루코시다제가 공지되어 있지만, 지금까지 트리코테르마 레에세이 베타-글루코시다제 Bgl1 및 아스페르길루스 니게르(*Aspergillus niger*) 베타-글루코시다제 SP188은 이에 대해 다른 베타-글루코시다제의 활성 및 성능이 평가되는 기준(benchmark) 베타-글루코시다제로 여겨진다. 트리코테르마 레에세이 Bgl1이 아스페르길루스 니게르 베타-글루코시다제 SP188보다 더 높은 비활성도를 갖지만, 전자는 분비가 저조할 수 있으며, 한편 후자는 글루코오스 억제에 더 민감한 것으로 보고되어 있다 (문헌 [Chauve, M. et al., (2010) Biotechnol. Biofuels, 3(1):3]).

[0014]

본 조성물 및 방법의 한 측면은 리그노셀룰로오스계 바이오매스 기질을 가수분해하기 위한, 진균 종 마그나포르테 그리세아로부터 단리된 고도 활성 베타-글루코시다제의 적용 또는 용도이다. 마그나포르테 그리세아, 벼도열병 진균(rice blast fungus)의 계통은, 2005년에 딘(Dean) 등에 의해, 문헌 [Nature, 434(7036):980-986 (2005)]에 주해되어 있다. 서열 번호 2의 본원에 기재된 서열은 미국 국립 생명공학 정보 센터(National Center for Biotechnology Information), 미국 국립 의학 도서관(U.S. National Library of Medicine: NCBI)에 수탁 번호 XP_003709907로 공개되고, 베타-글루코시다제 I로 지정되었다. 상기 효소는 리그노셀룰로오스계 바이오매스 기질을 가수분해하는데 유용한 효소 조성물에 포함되거나, 재조합 형태로 이전에 제조된 바가 없다. 상기 효소 또는 이러한 효소를 포함하는 조성물이 리그노셀룰로오스계 바이오매스 기질의 효소적 가수분해의 적합한 방법으로 이러한 기질에 적용된 바도 없다. 더욱이, 마그나포르테 그리세아의 베타-글루코시다제는 조작된 미생물에 의해 이전에 발현된 바도 없다. 이는 하나 이상의 셀룰라제 유전자 및/또는 하나 이상의 헤미셀룰라제 유전자와 공동-발현된 바도 없다 다년간의 발전을 통해, 유전적 툴(tool)의 비축을 활용하여, 이중 단백질 및 효소의 고도로 효과적이고 효율적인 생산자가 되어 온 적합한 미생물에서의 발현은, 이들이 미조작된(unengineered) 미생물에서 내생적으로 발현되는 경우, 또는 이들이 식물에서 발현되는 경우보다 실질적으로 더 대량으로 이러한 유용한 베타-글루코시다제를 발현할 수 있게 한다. 베타-글루코시다제로서 분류되는 효소는 그의 기원에서 뿐만 아니라 리그노셀룰로오스계 기질에 대한 그의 활성에서도 다양하지만, 전부는 아니더라도 대부분의 베타-글루코시다제는 적합한 조건 하에 셀로비오스 가수분해를 촉매할 수 있다. 예를 들어, 어떤 것은 셀로비오스에 대해서 뿐만 아니라 장쇄 올리고당에 대해서도 활성이며, 한편 다른 것은 셀로비오스에 대해서만 더 배타적으로 활성이다. 심지어 유사한 기질 선호도를 갖는 그러한 베타-글루코시다제의 경우, 어떤 것은 이들을 더 촉매적으로 활성이고 효율적으로 만들고, 따라서 효소적으로 촉매화된 가수분해가 최대 수일보다 더 오래걸릴 여유가 없는 산업상 이용에서 더 유용하게 만드는 효소 반응속도 프로파일을 갖는다. 더욱이, 리그노셀룰로오스계 바이오매스의 효소적 가수분해로부터 수득된 셀룰로오스계 당을 전환시킬 수 있는 어떠한 발효 또는 에타놀로젠(ethanlogen) 미생물도 조작되어 마그나포르테 그리세아로부터의 베타-글루코시다제, 예컨대 본원에서의 Mg3A 폴리펩티드를 발현시킨 바 없다 에타놀로젠 미생물에서의 베타-글루코시다제의 발현은, 효소 당화에 의해 완전히 전환되지는 않은 잔존 셀로비오스로부터 D-글루코오스를 추가로 유리시키는 중요한 기회를 제공하며, 여기서 이렇게 생성된 D-글루코오스는 에타놀로젠에 의해 적시에 즉시 소모되거나 발효될 수 있다.

[0015]

본 조성물 및 방법의 측면은 본원에서 "Mg3A" 또는 "Mg3A 폴리펩티드"로서 지칭되는, 마그나포르테 그리세아로부터 유래된 글리코실 히드롤라제 패밀리의 3의 베타-글루코시다제 폴리펩티드, 이를 암호화하는 핵산, 이를 포함하는 조성물, 및 리그노셀룰로오스계 바이오매스를 가용성 발효성 당으로 전환시키거나 가수분해하는데 있어서 베타-글루코시다제 폴리펩티드 및 이를 포함하는 조성물을 생성하고 적용하는 방법에 관한 것이다. 그 다음에 이러한 발효성 당은 셀룰로오스계 에탄올, 연료, 및 다른 생화학물질 및 유용한 제품으로 전환될 수 있다. 특정 실시양태에서, Mg3A 베타-글루코시다제 폴리펩티드는 공지된 고 충실도 베타-글루코시다제인 기준 트리코테

르마 레에세이 Bgl1과 비교하여, 더 높은 베타-글루코시다제 활성을 갖고/거나 소정의 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 증가된 능력을 나타낸다. (문헌 [Chauve, M. et al., (2010) Biotechnol. Biofuels, 3(1):3]).

[0016]

일부 실시양태에서, Mg3A 폴리펩티드는 적합한 바이오매스 기질을 가수분해하거나 분해하는 효소 조성물 중 하나 이상의 다른 셀룰라제와 함께, 또는 그의 존재하에 적용된다. 하나 이상의 다른 셀룰라제는, 예를 들어, 다른 베타-글루코시다제, 셀로비오히드롤라제, 및/또는 엔도글루카나제일 수 있다. 예를 들어, 효소 조성물은 Mg3A 폴리펩티드, 셀로비오히드롤라제, 및 엔도글루카나제를 포함할 수 있다. 일부 실시양태에서, Mg3A 폴리펩티드는 효소 조성물 중 하나 이상의 헤미셀룰라제와 함께, 또는 그의 존재하에 적용된다. 하나 이상의 헤미셀룰라제는, 예를 들어, 자일라나제, 베타-자일로시다제, 및/또는 L-아라비노푸라노시다제일 수 있다. 추가 실시양태에서, Mg3A 폴리펩티드는 효소 조성물 중 하나 이상의 셀룰라제 및 하나 이상의 헤미셀룰라제와 함께, 또는 이들의 존재하에 적용된다. 예를 들어, 효소 조성물은 Mg3A 폴리펩티드, 전혀 없거나 1개 또는 2개의 다른 베타-글루코시다제, 하나 이상의 셀로비오히드롤라제, 하나 이상의 엔도글루카나제; 임의로 전혀 없거나 하나 이상의 자일라나제, 전혀 없거나 하나 이상의 베타-자일로시다제, 및 전혀 없거나 하나 이상의 L-아라비노푸라노시다제를 포함한다.

[0017]

특정 실시양태에서, Mg3A 폴리펩티드, 또는 상기 Mg3A 폴리펩티드를 포함하는 조성물은, 리그노셀룰로오스계 바이오매스 기질의 효소적 가수분해에 의해 생성된 가용성 발효성 당을 대사시킬 수 있고 이러한 당을 에탄올, 생화학물질 또는 다른 유용한 물질로 전환시킬 수 있는 에타놀로겐 미생물의 존재하에 리그노셀룰로오스계 바이오매스 기질 또는 부분적으로 가수분해된 리그노셀룰로오스계 바이오매스 기질에 적용된다. 이러한 공정은 엄격한 순차 공정일 수 있고 그로 인해 가수분해 단계가 발효 단계 전에 일어난다. 이러한 공정은, 대안으로, 하이브리드 공정일 수 있고, 그로 인해 가수분해 단계가 먼저 시작되지만, 나중에 시작되는 발효 단계와 일정 기간 중복된다. 이러한 공정은, 추가 대안에서, 동시 가수분해 및 발효 공정일 수 있으며, 그로 인해 효소적 가수분해로부터 생성된 당이 에타놀로겐에 의해 발효되는 동안 바이오매스 기질의 효소적 가수분해가 일어난다.

[0018]

Mg3A 폴리펩티드는, 예를 들어, 효소적 가수분해 공정에 기여하며, 셀로비오스와 같은 올리고당으로부터 D-글루코오스의 유리에 기여하는, 효소 조성물의 일부일 수 있다. 특정 실시양태에서, Mg3A 폴리펩티드는, 에타놀로겐 미생물이 이러한 베타-글루코시다제 활성을 발현하고/거나 분비하도록, 유전자 조작되어 에타놀로겐에서 발현될 수 있다. 더욱이, Mg3A 폴리펩티드는 가수분해 효소 조성물의 일부일 수 있으며 한편 동시에 또한 에타놀로겐에 의해 발현되고/거나 분비될 수 있고, 그로 인해 가수분해 효소 조성물을 사용한 리그노셀룰로오스계 바이오매스 기질의 가수분해에 의해 생성된 가용성 발효성 당은 Mg3A 폴리펩티드를 또한 발현시키고/거나 분비하는 에타놀로겐 미생물에 의해 에탄올로 전환되고/거나 대사된다. 가수분해 효소 조성물은 하나 이상의 다른 셀룰라제 및/또는 하나 이상의 헤미셀룰라제 이외에도 Mg3A 폴리펩티드를 포함할 수 있다. 에타놀로겐은 Mg3A 폴리펩티드, 하나 이상의 다른 셀룰라제, 하나 이상의 다른 헤미셀룰라제, 또는 이들 효소의 조합을 발현시키도록 조작될 수 있다. 베타-글루코시다제 중 하나 이상은 가수분해 효소 조성물에 있을 수 있고 에타놀로겐에 의해 발현되고/거나 분비될 수 있다. 예를 들어, 리그노셀룰로오스계 바이오매스 기질의 가수분해는 Mg3A 폴리펩티드를 포함하는 효소 조성물을 사용하여 달성될 수 있고, 그 다음에 가수분해로부터 생성된 당은 Mg3A 폴리펩티드를 발현시키고/거나 분비하도록 조작된 미생물로 발효될 수 있다. 대안으로, 제1 베타-글루코시다제를 포함하는 효소 조성물은 가수분해 단계에 참여하고 제1 베타-글루코시다제와 상이한 제2 베타-글루코시다제는 에타놀로겐에 의해 발현되고/거나 분비된다. 예를 들어, 리그노셀룰로오스계 바이오매스 기질의 가수분해는 *트리코테르마 레에세이* Bgl1을 포함하는 가수분해 효소 조성물을 사용하여 달성될 수 있고, 가수분해로부터 생성된 발효성 당은 Mg3A 폴리펩티드를 발현시키고/거나 분비하는 에타놀로겐 미생물에 의해 발효되거나, 그 역도 마찬가지다.

[0019]

본원에서 실증된 바와 같이, Mg3A 폴리펩티드 및 Mg3A 폴리펩티드를 포함하는 조성물은 리그노셀룰로오스계 바이오매스의 당화 및 분해가 실시되는 조건하에서 개선된 효능을 갖는다. Mg3A 폴리펩티드를 포함하는 효소 조성물의 개선된 효능은, 소정의 바이오매스 기질을 가수분해하는 그의 성능이 *트리코테르마 레에세이*의 Bgl1을 포함하는 달리 필적하는 효소 조성물의 것과 비교되는 경우, 나타난다.

[0020]

특정 실시양태에서, 개선되거나 증가된 베타-글루코시다제 활성이 Mg3A 폴리펩티드의 개선되거나 증가된 셀로비아제 활성에서 반영되고, 이는, 예를 들어, 약 30℃ 내지 약 65℃ (예를 들어, 약 35℃ 내지 약 60℃, 약 40℃ 내지 약 55℃, 약 45℃ 내지 약 55℃, 약 48℃ 내지 약 52℃, 약 40℃, 약 45℃, 약 50℃, 약 55℃ 등)의 온도에서 기질로서 셀로비오스를 사용하여 측정된다. 일부 실시양태에서, *트리코테르마 레에세이* Bgl1의 것과 비교하여 Mg3A 폴리펩티드의 개선된 베타-글루코시다제 활성은, 베타-글루코시다제 폴리펩티드를 사용하여 인산 펄

윤된 셀룰로오스 (PASC), 예를 들어, 문헌 [Walseth, TAPPI 1971, 35:228] 및 [Wood, Biochem. J. 1971, 121:353-362]의 개조된 프로토콜을 사용하여 전처리된 이렇게 전처리된 아비셀(Avicel)을 가수분해하는 경우 관찰된다. 일부 실시양태에서, *트리코테르마 레에세이* Bgl1의 것과 비교하여 Mg3A 폴리펩티드의 개선된 베타-글루코시다제 활성은, 베타-글루코시다제 폴리펩티드를 사용하여 회식 암모니아로 전처리된 옥수수 대, 예를 들어, 국제 공개 특허 출원: W02006110891, W02006110899, W02006110900, W02006110901, 및 W02006110902; 미국 특허 제7,998,713호, 제7,932,063호에 기재된 것을 가수분해하는 경우 관찰된다.

[0021]

일부 실시양태에서, 개선된 베타-글루코시다제 활성은 *트리코테르마 레에세이* Bgl1의 것과 비교하여, 더 고온에서 수행된 당화 반응에서 바이오매스 기질의 전환 또는 가수분해를 촉매하는 Mg3A 폴리펩티드의 능력에서의 증가된 강건함(robustness)에서 반영된다. 예를 들어, *트리코테르마 레에세이* Bgl1로 관찰되는 것에 비해 Mg3A로 관찰되는 증가된 가수분해의 차이 또는 수준은, 저온에서보다는 고온에서 더 크며, 따라서 이는 본 개시내용의 Mg3A 폴리펩티드가 그의 더 높은 열안정성으로 인해 고온에서 바이오매스 가수분해에 특히 적합함을 나타내는 것이다. 일부 실시양태에서, 약 55℃의 가수분해 온도에서, Mg3A 폴리펩티드는, 동일 수준에서 투여된 *트리코테르마 레에세이* Bgl1과 비교하여 등량의 다른 셀룰라제 및 헤미셀룰라제의 존재하에, 동일 기질의 가수분해의 적어도 2%, 예를 들어, 적어도 5%, 적어도 7%, 적어도 10%, 또는 심지어는 적어도 15% 더 높은 수준을 달성할 수 있다. 특정 실시양태에서, Mg3A의 가수분해 수준 대 Bgl1의 가수분해 수준 간의 차이는 50℃에서보다 55℃에서 더 높다. Mg3A의 이러한 개선된 열안정성 또는 고온-성능은 오히려 예기치 못한 것인데 그 이유는 고온에서 폴리펩티드의 폴림(unfolding) 또는 분해(breakdown)를 측정하는데 전형적으로 사용되는 용해 온도 (Tm)가 실제로 Mg3A의 경우보다는 *트리코테르마 레에세이* Bgl1의 경우 훨씬 더 높은 Tm (예를 들어, 약 10℃ 더 높음)을 나타냈기 때문이라는 점을 주목한다. 이론에 의해 구속되는 것을 바라지 않으면서, 이러한 놀라운 관찰은 Tm이 바이오매스 기질의 존재하에 열안정성의 양호한 지표 또는 예측 변수(predictor)는 아니라는 것을 시사할 수 있다.

[0022]

일부 측면에서, Mg3A 폴리펩티드 및/또는 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 효소 조성물에서 또는 방법에서 적용되는 바와 같은 상기 폴리펩티드는 (a) *마그나포르테 그리세아*로부터 유래되거나, 그로부터 수득가능하거나, 그에 의해 생성되고; (b) 서열 번호 2의 아미노산 서열과 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%) 동일한 아미노산 서열을 포함하는 재조합 폴리펩티드; (c) 서열 번호 2의 촉매적 도메인, 즉 아미노산 잔기 19-873과 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%) 동일한 아미노산 서열을 포함하는 재조합 폴리펩티드; (d) 서열 번호 3의 아미노산 서열의 성숙 형태, 즉 서열 번호 2의 아미노산 잔기 19-873과 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%) 동일한 아미노산 서열을 포함하는 재조합 폴리펩티드; 또는 (e) 베타-글루코시다제 활성을 갖는 (a), (b), (c) 또는 (d)의 단편이다. 특정 실시양태에서, 서열 번호 2의 하나 이상의 아미노산 잔기의 치환, 결실 및/또는 삽입을 포함하는, 베타-글루코시다제 활성을 갖는 변이체 폴리펩티드가 제공된다.

[0023]

일부 측면에서, Mg3A 폴리펩티드 및/또는 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 효소 조성물에서 또는 방법에서 적용되는 바와 같은 상기 폴리펩티드는 (a) 서열 번호 1에 대하여 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%) 서열 동일성인 핵산 서열에 의해 암호화되는 폴리펩티드, 또는 (b) 중간 엄격성 조건하에, 높은 엄격성 조건하에 또는 매우 높은 엄격성 조건하에 서열 번호 1 또는 적어도 100개의 인접 뉴클레오티드의 서열 번호 1의 부분서열(subsequence), 또는 그의 상보적 서열과 혼성화할 수 있는 것이며, 여기서 폴리펩티드는 베타-글루코시다제 활성을 갖는다. 일부 실시양태에서, Mg3A 폴리펩티드 및/또는 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 조성물에서 또는 방법에서 적용되는 바와 같은 상기 폴리펩티드는, 유전 암호의 축중(degeneracy)으로 인해, 중간 엄격성 조건하에, 높은 엄격성 조건하에 또는 매우 높은 엄격성 조건하에 서열 번호 1 또는 적어도 100개의 인접 뉴클레오티드의 서열 번호 1의 부분서열과 혼성화하지 않으나, 그럼에도 불구하고 베타-글루코시다제 활성을 갖고 서열 번호 2의 것과 또는 서열 번호 3의 성숙 베타-글루코시다제 서열과 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%) 동일한 아미노산 서열을 포함하는 폴리펩티드를 암호화하는 것이다. 핵산 서열은 함성일 수 있고, 반드시 *마그나포르테 그리세아*로부터 유래될 필요는 없으나, 핵산 서열은 베타-글루코시다제 활성을 갖는 폴리펩티드를 암호화하고 서열 번호 2 또는 서열 번호 3과 75% 이상 동일한 아미노산 서열을 포함한다.

[0024]

일부 바람직한 실시양태에서, Mg3A 폴리펩티드 또는 상기 Mg3A 폴리펩티드를 포함하는 조성물은 (서열 번호 4의) 야생형 *트리코테르마 레에세이* Bgl1, 또는 상기 *트리코테르마 레에세이* Bgl1을 포함하는 효소 조성물의 것

과 비교하여, 개선된 베타-글루코시다제 활성을 갖는다. 특정 실시양태에서, 본원에서의 조성물 및 방법의 Mg3A 폴리펩티드의 셀룰라제 활성은, 클로로-니트로-페닐-글루코시드 (CNP) 가수분해 검정을 사용하여 측정된 바와 같이, *트리코테르마 레에세이* Bgl1의 것보다 적어도 약 5% 더 높거나, 약 10% 더 높거나, 또는 약 15% 더 높거나, 또는 약 20% 더 높다. CNP 검정은 본원에서 실시예 2B에 기재되어 있다. 일부 실시양태에서, 본원에서의 조성물 및 방법의 Mg3A 폴리펩티드의 셀룰라제 활성은, CNP 가수분해 검정을 사용하여 측정된 바와 같이, *아스페르길루스 니게르* B-glu의 것보다 적어도 약 2배 더 높거나, 적어도 약 5배 더 높거나, 적어도 약 6배 더 높거나, 적어도 약 7배 더 높거나, 적어도 약 8배 더 높거나, 적어도 약 9배 더 높거나, 또는 심지어는 적어도 약 10배 더 높다.

[0025]

예를 들어, 본원에서의 조성물 및 방법의 Mg3A 폴리펩티드의 베타-글루코시다제 활성은, 셀로비오스 가수분해 검정을 사용하여 측정된 바와 같이, *트리코테르마 레에세이* Bgl1의 것보다 적어도 약 20% 더 높다 (예를 들어, 적어도 약 20% 더 높거나, 적어도 약 30% 더 높거나, 적어도 약 40% 더 높거나, 적어도 약 50% 더 높거나, 적어도 약 60% 더 높거나, 적어도 약 70% 더 높거나, 적어도 약 80% 더 높거나, 적어도 약 85% 더 높거나, 예를 들어 적어도 약 87% 더 높다 등이다). 셀로비오스 가수분해 검정은 본원에서 실시예 2C에 기재되어 있다. 일부 실시양태에서, 본원에서의 조성물 및 방법의 Mg3A 폴리펩티드의 베타-글루코시다제 활성은, 셀로비오스 가수분해 검정을 사용하여 측정된 바와 같이, *아스페르길루스 니게르* B-glu의 것보다 약 2% 덜 하거나, 약 5% 덜 하거나, 약 7% 덜 하거나, 약 10% 덜 하거나, 예를 들어, 약 13% 덜하다 등이다.

[0026]

일부 실시양태에서, 본원에서의 조성물 및 방법의 Mg3A 폴리펩티드는 *트리코테르마 레에세이* Bgl1과 비교하여 실질적으로 증가된 (예를 들어, 적어도 약 20% 더 높거나, 적어도 약 30% 더 높거나, 적어도 약 40% 더 높거나, 적어도 약 50% 더 높거나, 적어도 약 60% 더 높거나, 적어도 약 70% 더 높거나, 적어도 약 80% 더 높거나, 적어도 약 85% 더 높거나, 예를 들어 적어도 약 87% 더 높은 등) 셀로비오스 가수분해 활성을 갖지만 클로로-니트로-페닐-글루코시드 (CNP)를 가수분해하는 다소 덜 극적으로 증가된 (예를 들어, 약 5% 높거나, 약 10% 더 높거나, 약 15% 더 높거나, 약 20% 더 높은) 능력을 갖는다. 예를 들어, 본원에서의 조성물 및 방법의 Mg3A 폴리펩티드는 *아스페르길루스 니게르* B-glu와 비교하여 실질적으로 증가된 CNP 가수분해 활성을 갖지만, *아스페르길루스 니게르* B-glu와 비교시 다소 유사하거나 약간 덜한 셀로비오스 가수분해 활성을 갖는다.

[0027]

일부 실시양태에서, *트리코테르마 레에세이* Bgl1과 비교하여, 재조합 Mg3A 폴리펩티드는, CNP/셀로비오스에 비해 약 10% 감소되거나, 약 20% 감소되거나, 약 30% 감소되거나, 또는 심지어는 약 40% 감소된 가수분해 활성 비를 갖는다. 일부 실시양태에서, Mg3A 폴리펩티드는, *아스페르길루스 니게르* B-glu와 비교하여, CNP/셀로비오스에 비해 약 2배, 약 5배, 약 7배, 약 10배, 약 15배, 또는 심지어는 약 20배 더 높은 상대 가수분해 활성 비를 갖는다.

[0028]

특정 측면에서, 본 발명의 Mg3A 폴리펩티드 및 상기 Mg3A 폴리펩티드를 포함하는 조성물은 (서열 번호 4의) 야생형 *트리코테르마 레에세이* Bgl1의 것과 비교하여, 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 개선된 성능을 갖는다. 일부 실시양태에서, Mg3A 폴리펩티드 또는 Mg3A 폴리펩티드를 포함하는 조성물의 개선된 가수분해 성능은, 동일한 당화 조건하에, 동일한 방식으로 전처리된 동일한 바이오매스로부터 *트리코테르마 레에세이* Bgl1 또는 *트리코테르마 레에세이* Bgl1을 포함하는 동일한 효소 조성물에 의해 생성된 글루코오스의 수준과 비교하여, 특정 방식으로 전처리된, 소정의 리그노셀룰로오스계 바이오매스 기질로부터 더 많은 양의 글루코오스의 생성에 의해 관찰가능하다. 예를 들어, Mg3A 폴리펩티드에 의해 또는 상기 Mg3A 폴리펩티드를 포함하는 효소 조성물에 의해 생성된 글루코오스의 양은, 0-10 mg (예를 들어, 약 1 mg, 약 2 mg, 약 3 mg, 약 4 mg, 약 5 mg, 약 6 mg, 약 7 mg, 약 8 mg, 약 9 mg, 약 10 mg)의 베타-글루코시다제 (Mg3A 폴리펩티드 또는 *트리코테르마 레에세이* Bgl1)를 사용하여 바이오매스 기질 중 1 g의 글루칸을 가수분해하는 경우, (Mg3A 폴리펩티드라기 보다는) *트리코테르마 레에세이* Bgl1 또는 *트리코테르마 레에세이* Bgl1을 포함하는 달리 동일한 효소 조성물에 의해 생성된 글루코오스의 양보다 적어도 약 5% (예를 들어, 적어도 약 5%, 적어도 약 10%, 적어도 약 15%, 적어도 약 20%, 또는 적어도 약 25) 더 많다.

[0029]

일부 측면에서, Mg3A 폴리펩티드 또는 Mg3A 폴리펩티드를 포함하는 조성물의 개선된 가수분해 성능은, 동일한 당화 조건하에, 동일한 방식으로 전처리된 동일한 바이오매스로부터 *트리코테르마 레에세이* Bgl1 또는 *트리코테르마 레에세이* Bgl1을 포함하는 달리 동일한 효소 조성물에 의한 글루칸 전환율 (%)의 수준과 비교하여, 특정 방식으로 전처리된 소정의 리그노셀룰로오스계 바이오매스 기질로부터 증가된 글루칸 전환율 (%)에 의해 관찰가능하다. 예를 들어, Mg3A 폴리펩티드 또는 상기 Mg3A 폴리펩티드를 포함하는 효소 조성물에 의한 글루칸 전환율 (%)은, 0-10 mg (예를 들어, 약 1 mg, 약 2 mg, 약 3 mg, 약 4 mg, 약 5 mg, 약 6 mg, 약 7 mg, 약 8 mg, 약 9 mg, 약 10 mg)의 베타-글루코시다제 (Mg3A 폴리펩티드 또는 *트리코테르마 레에세이* Bgl1)를 사용하여 바이

오메스 기질 중 1 g의 글루칸을 가수분해하는 경우, (Mg3A 폴리펩티드라기 보다는) *트리코테르마 레에세이* Bgl1 또는 *트리코테르마 레에세이* Bgl1 을 포함하는 달리 동일한 효소 조성물에 의한 글루칸 전환율 (%)보다 적어도 약 5% (예를 들어, 적어도 약 5%, 적어도 약 10%, 또는 적어도 약 15%) 더 크다.

[0030]

추가 측면에서, Mg3A 폴리펩티드 및 Mg3A 폴리펩티드를 포함하는 조성물의 개선된 가수분해 성능은, 동일한 당화 조건하에 *트리코테르마 레에세이* Bgl1 또는 *트리코테르마 레에세이* Bgl1을 포함하는 달리 동일한 조성물에 의해 동일한 바이오매스 기질을 가수분해하는 경우 셀로비오스의 잔존량과 비교하여, 특정 방식으로 전처리된 소정의 리그노셀룰로오스계 바이오매스 기질을 가수분해함으로써, 생성물 혼합물 중 잔존 셀로비오스의 감소된 양 및/또는 더 높은 셀로비아제 활성에 의해 관찰가능하다. 예를 들어, Mg3A 폴리펩티드 또는 상기 Mg3A 폴리펩티드를 포함하는 조성물에 의해, 특정 방식으로 전처리된 소정의 바이오매스 기질의 가수분해로부터 생성된 생성물 혼합물 중 잔존 셀로비오스의 양은, 동일한 당화 조건하에 *트리코테르마 레에세이* Bgl1 또는 *트리코테르마 레에세이* Bgl1을 포함하는 달리 동일한 효소 조성물에 의해 동일한 방식으로 전처리된 동일한 바이오매스 기질의 가수분해로부터 생성된 생성물 혼합물에서 생성된 잔존 셀로비오스의 양보다 적어도 약 5% (예를 들어, 적어도 약 5%, 적어도 약 10%, 적어도 약 15%, 또는 심지어는 적어도 약 20) 더 적다. 이는 0-10 mg (예를 들어, 약 1 mg, 약 2 mg, 약 3 mg, 약 4 mg, 약 5 mg, 약 6 mg, 약 7 mg, 약 8 mg, 약 9 mg, 약 10 mg)의 베타-글루코시다제 (예를 들어, Mg3A 폴리펩티드 또는 *트리코테르마 레에세이* Bgl1)를 사용하여 바이오매스 기질 중 1 g의 글루칸을 가수분해하는 경우이다.

[0031]

본 조성물 및 방법의 측면은 리그노셀룰로오스계 바이오매스 및 상기 상세히 설명된 바와 같은 재조합 Mg3A 폴리펩티드를 포함하는 조성물을 포함한다. 적합한 리그노셀룰로오스계 바이오매스는, 예를 들어, 농작물, 식품 또는 사료 생산의 부산물, 리그노셀룰로오스계 폐기물, 예를 들어, 목초 잔해(grass residue)를 포함한 식물 잔해(plant residue), 또는 폐지 또는 폐지 제품으로부터 유래될 수 있다. 특정 실시양태에서, 리그노셀룰로오스계 바이오매스는 하나 이상의 전처리 단계에 적용되어, 자일란, 헤미셀룰로오스, 셀룰로오스 및/또는 리그닌 물질이 효소에 더욱 접근가능하거나 민감하게 하며, 그에 따라 효소에 의해 더욱 가수분해될 수 있게 한다. 적합한 전처리 방법은, 예를 들어, 바이오매스 물질을 반응기 내에서 금속 염 및 강산의 희석 용액을 포함하는 촉매에 적용하는 것일 수 있다. 예를 들어, 미국 특허 제6,660,506호, 제6,423,145호를 참조한다. 대안으로, 적합한 전처리는, 예를 들어, 미국 특허 제5,536,325호에 기재된 바와 같은 다단계 공정일 수 있다. 특정 실시양태에서, 바이오매스 물질을 미국 특허 제6,409,841호의 개시내용에 따라, 약 0.4% 내지 약 2%의 강산을 사용한 묽은 산 가수분해의 하나 이상의 단계에 적용할 수 있다. 전처리 방법의 추가 실시양태는, 예를 들어, 미국 특허 제5,705,369호; 문헌 [Gould, (1984) Biotech. & Bioengr., 26:46-52]; [Teixeira et al., (1999) Appl. Biochem & Biotech., 77-79:19-34]; 국제 공개 특허 출원 W02004/081185; 또는 미국 특허 공개 제20070031918호, 또는 국제 공개 특허 출원 W006110901에 기재된 것들을 포함할 수 있다.

[0032]

본 발명은 또한 베타-글루코시다제 활성을 갖는 폴리펩티드를 암호화하는 단리된 폴리뉴클레오티드에 관한 것으로, 여기서 단리된 폴리뉴클레오티드는 하기로부터 선택된다:

[0033]

(1) 서열 번호 2 또는 서열 번호 3에 대하여 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%)의 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드를 암호화하는 폴리뉴클레오티드;

[0034]

(2) 서열 번호 1에 대하여 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%)의 동일성을 갖거나, 중간 엄격성 조건하에, 높은 엄격성 조건하에, 또는 매우 높은 엄격성 조건하에 서열 번호 1, 또는 그의 상보적 서열과 혼성화하는 폴리뉴클레오티드.

[0035]

본 조성물 및 방법의 측면은 서열 번호 2의 것, 또는 성숙 서열 서열 번호 3의 것과 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%) 동일한 아미노산 서열을 포함하는 재조합 폴리펩티드를 암호화하는 단리된 핵산 서열을 사용하는, 베타-글루코시다제 활성을 갖는 Mg3A 폴리펩티드의 제조 또는 생성 방법을 포함한다. 일부 실시양태에서, 폴리펩티드는 생성되는 Mg3A 폴리펩티드가 숙주 생물에 의해 분비되도록 천연 또는 비천연 신호 펩티드를 추가로 포함하며, 예를 들어, 신호 펩티드는 서열 번호 13 (*트리코테르마 레에세이* Bgl1의 신호 서열)과 90% 이상 동일한 서열을 포함한다. 특정 실시양태에서 단리된 핵산은 서열 번호 1과 적어도 75% (예를 들어, 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%) 동일한 서열을 포함한다. 특정 실시양태에서, 단리된 핵산은 신호 펩티드 서열을 암호화하는 핵산 서열을 추가로 포함한다. 특정 실시양태에서, 신호 펩티드 서열은 서열 번호 13-42로부터 선택된 것일 수 있다. 특정 특별한 실시양태에서, 서열 번호 13의 신호 펩티드 서열을 암호화하는 핵산

서열을 사용하여 *트리코테르마 레에세이* 중에서 Mg3A 폴리펩티드를 발현시킨다.

- [0036] 본 조성물 및 방법의 측면은 조절 서열과 작동가능하게 조합된 상기 기재된 바와 같은 단리된 핵산을 포함하는 발현 벡터를 포함한다.
- [0037] 본 조성물 및 방법의 측면은 발현 벡터를 포함하는 숙주 세포를 포함한다. 특정 실시양태에서, 숙주 세포는 세균 세포 또는 진균 세포이다. 특정 실시양태에서, 발현 벡터를 포함하는 숙주 세포는 리그노셀룰로오스계 바이오매스의 가수분해로부터 생성된 가용성 당을 대사시킬 수 있는 에타놀로겐 미생물이고, 여기서 가수분해는 화학적 및/또는 효소적 공정의 결과이다.
- [0038] 본 조성물 및 방법의 측면은 상기 기재된 숙주 세포 및 배양 배지를 포함하는 조성물을 포함한다. 본 조성물 및 방법의 측면은 베타-글루코시다제를 생성하기에 적합한 조건하에, 배양 배지에서 상기 기재된 숙주 세포를 배양하는 것을 포함하는, Mg3A 폴리펩티드의 생성 방법을 포함한다.
- [0039] 본 조성물 및 방법의 측면은 상기 기재된 바와 같은 베타-글루코시다제의 생성 방법에 따라서 생성된 배양 배지의 상청액 중 Mg3A 폴리펩티드를 포함하는 조성물을 포함한다.
- [0040] 일부 측면에서 본 발명은 상기 및 여기에서 기재된 바와 같은, 베타-글루코시다제 활성을 갖는 폴리펩티드를 암호화하는 폴리뉴클레오티드를 포함하는, 핵산 구축물, 재조합 발현 벡터, 조작된 숙주 세포에 관한 것이다. 추가 측면에서, 본 발명은 핵산 구축물, 재조합 발현 벡터, 및/또는 조작된 숙주 세포를 사용하는, 본 발명의 베타-글루코시다제 폴리펩티드 또는 이러한 베타-글루코시다제 폴리펩티드를 포함하는 조성물의 제조 또는 생성 방법에 관한 것이다. 특히, 본 발명은, 예를 들어, 서열 번호 2와 또는 서열 번호 3의 성숙 서열과 적어도 75% 동일하거나, 서열 번호 1과 75% 이상 동일한 폴리뉴클레오티드에 의해 암호화되는 베타-글루코시다제의 성숙 서열에 작동가능하게 연결된 적합한 신호 펩티드를 포함하는 핵산 구축물, 단리된 폴리뉴클레오티드, 핵산 구축물, 재조합 발현 벡터, 또는 이러한 핵산 구축물을 포함하는 조작된 숙주 세포에 관한 것이다. 일부 실시양태에서, 신호 펩티드 및 베타-글루코시다제 서열은 상이한 미생물로부터 유래된다.
- [0041] 또한, 조절 서열과 작동가능하게 조합된 단리된 핵산을 포함하는 발현 벡터가 제공된다. 게다가, 발현 벡터를 포함하는 숙주 세포가 제공된다. 여전히 추가 실시양태에서, 숙주 세포 및 배양 배지를 포함하는 조성물이 제공된다.
- [0042] 일부 실시양태에서, 숙주 세포는 세균 세포 또는 진균 세포이다. 특정 실시양태에서, 숙주 세포는 리그노셀룰로오스계 바이오매스 기질의 가수분해 (여기서, 가수분해는 화학적 가수분해 또는 효소적 가수분해 또는 이들 공정의 조합을 통해서일 수 있다)로부터 생성되는 가용성 당을 대사시킬 수 있으나, 또한 이중 효소를 발현시킬 수 있는 에타놀로겐 미생물이다. 일부 실시양태에서, 숙주 세포는 *사카로마이세스 세레비지에* 또는 *지모모나스 모빌리스*(*Zymomonas mobilis*) 세포이고, 이들은 이중 폴리펩티드, 예컨대 본 발명의 Mg3A 폴리펩티드를 발현시킬 수 있을 뿐만 아니라, 당을 에탄올 및/또는 하류 제품으로 발효시킬 수도 있다. 특정 특별한 실시양태에서, 베타-글루코시다제를 발현시키는 *사카로마이세스 세레비지에* 세포 또는 *지모모나스 모빌리스* 세포는, 하나 이상의 베타-글루코시다제를 포함하는 효소 조성물에 의해 리그노셀룰로오스계 바이오매스로부터 생성된 당을 발효시킬 수 있다. 하나 이상의 베타-글루코시다제를 포함하는 효소 조성물은 동일한 베타-글루코시다제를 포함할 수 있거나 하나 이상의 상이한 베타-글루코시다제를 포함할 수 있다. 특정 실시양태에서, 하나 이상의 베타-글루코시다제를 포함하는 효소 조성물은 세균 또는 진균 세포일 수 있는, 조작된 숙주 세포에 의해 생성된 효소 혼합물일 수 있다. *사카로마이세스 세레비지에* 또는 *지모모나스 모빌리스* 세포가 본 개시내용의 Mg3A 폴리펩티드를 발현시키는 경우, Mg3A 폴리펩티드는 발현될 수는 있지만, 분비되지는 않을 수 있다. 따라서 셀로비오스는 베타-글루코시다제 Mg3A 폴리펩티드가 D-글루코오스의 유리를 촉매하도록 이러한 숙주 세포에 도입되거나 또는 "수송되어"야 한다. 따라서 특정 실시양태에서, *사카로마이세스 세레비지에* 또는 *지모모나스 모빌리스* 세포는 Mg3A 폴리펩티드를 암호화하는 것 이외에도 셀로비오스 수송체 유전자로 형질전환된다. 셀로비오스 수송체 및 베타-글루코시다제는, 예를 들어, 문헌 [Ha *et al.*, (2011) PNAS, 108(2):504-509]에서, 생성 미생물이 셀로비오스를 발효시킬 수 있도록 *사카로마이세스 세레비지에*에서 발현되었다. 또 다른 셀로비오스 수송체는, 예를 들어, 공개된 미국 특허 출원 제20110262983호에서, *피치아*(*Pichia*) 효모에서 발현되었다. 셀로비오스 수송체는, 예를 들어, 문헌 [Sekar *et al.*, (2012) Applied Environmental Microbiology, 78(5):1611-1614]에서, *이. 콜라이*(*E. coli*)에 도입되었다.
- [0043] 추가 실시양태에서, Mg3A 폴리펩티드는 숙주 세포에 의해 이중 발현된다. 예를 들어, Mg3A 폴리펩티드는 *마그나포르테 그리세아*가 아닌 조작된 미생물에 의해 발현된다. 일부 실시양태에서, Mg3A 폴리펩티드는 하나 이상

의 상이한 셀룰라제 유전자와 공동-발현된다. 일부 실시양태에서, Mg3A 폴리펩티드는 하나 이상의 헤미셀룰라제 유전자와 공동-발현된다.

[0044]

일부 측면에서, 선행 단락의 재조합 Mg3A 폴리펩티드를 포함하는 조성물 및 이러한 조성물의 제조 방법이 제공된다. 일부 실시양태에서, 조성물은 하나 이상의 다른 셀룰라제를 추가로 포함하고, 그로 인해 하나 이상의 다른 셀룰라제는 숙주 세포에 의해 Mg3A 폴리펩티드와 공동-발현된다. 예를 들어, 하나 이상의 다른 셀룰라제는, 전혀 없거나 하나 이상의 다른 베타-글루코시다제, 하나 이상의 셀로비오히드롤라제, 및/또는 하나 이상의 엔도글루카나제로부터 선택될 수 있다. 이러한 다른 베타-글루코시다제, 셀로비오히드롤라제 및/또는 엔도글루카나제는, 존재할 경우, 단일 숙주 세포에 의해 Mg3A 폴리펩티드와 공동-발현될 수 있다. 2개 이상의 셀룰라제 중 적어도 2개는 서로 이종일 수 있거나 상이한 유기체로부터 유래될 수 있다. 예를 들어, 조성물은 2개의 베타-글루코시다제를 포함할 수 있고, 제1의 것은 Mg3A 폴리펩티드이고, 제2 베타-글루코시다제는 *마그나포르테 그리세아* 균주로부터 유래되지 않는다. 예를 들어, 조성물은 적어도 하나의 셀로비오히드롤라제, 하나의 엔도글루카나제, 또는 하나의 베타-글루코시다제 (이는 *마그나포르테 그리세아*로부터 유래되지 않음)를 포함할 수 있다. 일부 실시양태에서, 셀룰라제 중 하나 이상은 숙주 세포에 내인성(endogenous)이나, 달리 숙주 세포에서 자연 발생될 수 있는 것과 상이한 수준에서 과발현 또는 발현된다. 예를 들어, 셀룰라제 중 하나 이상은 *트리코테르마 레에세이* CBH1 및/또는 CBH2일 수 있고, 이는 *트리코테르마 레에세이* 숙주 세포에 고유하나, CBH1 및 CBH2 중 하나 또는 둘 다는, 이들이 *트리코테르마 레에세이* 숙주 세포에서 Mg3A 폴리펩티드와 공동-발현되는 경우, 과발현 또는 저발현된다.

[0045]

특정 실시양태에서, 재조합 Mg3A 폴리펩티드를 포함하는 조성물은 하나 이상의 헤미셀룰라제를 추가로 포함할 수 있고, 그로 인해 하나 이상의 헤미셀룰라제는 숙주 세포에 의해 Mg3A 폴리펩티드와 공동-발현된다. 예를 들어, 하나 이상의 헤미셀룰라제는 하나 이상의 자일라나제, 하나 이상의 베타-자일로시다제, 및/또는 하나 이상의 L-아라비노푸라노시다제로부터 선택될 수 있다. 이러한 다른 자일라나제, 베타-자일로시다제 및 L-아라비노푸라노시다제는, 존재할 경우, 단일 숙주 세포에 의해 Mg3A 폴리펩티드와 공동-발현될 수 있다. 일부 실시양태에서, 조성물은 적어도 하나의 베타-자일로시다제, 자일라나제 또는 아라비노푸라노시다제 (이는 *마그나포르테 그리세아*로부터 유래되지 않음)를 포함할 수 있다.

[0046]

추가 측면에서, 재조합 Mg3A 폴리펩티드를 포함하는 조성물은 하나 이상의 다른 셀룰라제 및 하나 이상의 헤미셀룰라제를 추가로 포함할 수 있고, 그로 인해 하나 이상의 셀룰라제 및/또는 하나 이상의 헤미셀룰라제가 숙주 세포에 의해 Mg3A 폴리펩티드와 공동-발현된다. 예를 들어, Mg3A 폴리펩티드는 동일한 숙주 세포에서 다른 비-셀룰라제 비-헤미셀룰라제 효소 또는 단백질 이외에도, 하나 이상의 다른 베타-글루코시다제, 하나 이상의 셀로비오히드롤라제, 하나 이상의 엔도글루카나제, 하나 이상의 엔도-자일라나제, 하나 이상의 베타-자일로시다제, 및 하나 이상의 L-아라비노푸라노시다제와 공동-발현될 수 있다. 따라서 본 조성물 및 방법의 측면은 Mg3A 폴리펩티드 이외에도 다수의 효소를 공동-발현하는 상기 기재된 숙주 세포 및 배양 배지를 포함하는 조성물을 포함한다. 따라서 본 조성물 및 방법의 측면은, Mg3A 폴리펩티드와 상기 기재된 바와 같은 다수의 효소를 공동-발현하는 숙주 세포를, Mg3A 및 다른 효소를 생성하는데 적합한 조건하에, 배양 배지에서 배양하는 것을 포함하는 Mg3A-함유 효소의 조성물의 제조 방법을 포함한다. 또한, 배양 배지의 상청액 중 본원에서의 방법에 따라 생성된 Mg3A 폴리펩티드 및 다른 효소를 포함하는 조성물이 제공된다. 배양 배지의 이러한 상청액은, 그대로, 또는 최소한의 생성-후 처리(post-production processing) (이는 전형적으로 세포 잔해(cell debris)를 제거하는 여과, 세포-사멸 절차, 및/또는 한외여과 또는 거기서 효소를 풍부하게 하거나 농축시키는 다른 단계를 포함할 수 있다) 또는 이러한 처리 없이 사용될 수 있다. 이러한 상청액은 본원에서 "전체 브로쓰(whole broth)" 또는 "전체 셀룰라제 브로쓰"로 칭해진다.

[0047]

추가 측면에서, 본 발명은 셀룰로오스계 물질을 분해 또는 전환시키고 셀룰로오스계 물질로부터 물질을 생성하기에 적합한 조건하에 상기 기재된 바와 같은 조성물을 적용 또는 사용하는 방법에 관한 것이다.

[0048]

추가 측면에서, 바람직하게는 하나 이상의 전처리 단계에 이미 적용된 셀룰로오스계 물질을 선행 단락 중 하나의 Mg3A 폴리펩티드 또는 이러한 폴리펩티드를 포함하는 조성물과 접촉시켜 발효성 당을 수득하는 것을 포함하는, 셀룰로오스계 물질을 발효성 당으로 분해 또는 전환시키는 방법이 제공된다.

[0049]

따라서 본 명세서는 하기 특별한 측면에 관한 것이다:

[0050]

제1 측면에서, 서열 번호 2 또는 성숙 서열 서열 번호 3과 75% 이상 동일한 아미노산 서열을 포함하는 재조합 폴리펩티드로서, 베타-글루코시다제 활성을 갖는 재조합 폴리펩티드.

- [0051] 제2 측면에서, 재조합 폴리펩티드 및 *트리코테르마 레에세이* Bgl1을 사용하여 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 경우, *트리코테르마 레에세이* Bgl1과 비교하여 개선된 베타-글루코시다제 활성 또는 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 능력을 갖는, 제1 측면의 재조합 폴리펩티드.
- [0052] 제3 측면에서, 개선된 베타-글루코시다제 활성이 증가된 셀로비아제 활성 또는 셀로비오스를 가수분해하고, 그로 인해 D-글루코오스를 유리시키는 개선된 능력인, 상기한 바와 같은, 제1 측면 또는 제2 측면의 재조합 폴리펩티드.
- [0053] 제4 측면에서, 개선된 베타-글루코시다제 활성이 소정의 세트의 당화 조건하에 소정의 리그노셀룰로오스계 바이오매스로부터 글루코오스의 증가된 수율인, 상기한 바와 같은, 제1 측면, 제2 측면 또는 제3 측면의 재조합 폴리펩티드.
- [0054] 제5 측면에서, 리그노셀룰로오스계 바이오매스가 당화 전에 전처리에 적용된 것인, 상기 제1 측면 내지 제4 측면 중 어느 한 측면의 재조합 폴리펩티드. 전처리는 적합하게는, 리그노셀룰로오스계 바이오매스 기질을, 예를 들어 본원에 기재된 그러한 전처리 방법을 포함할 수 있는, 효소적 접근 및 가수분해를 더욱 가능하게 하는 관련 기술분야에 공지된 것들일 수 있다.
- [0055] 제6 측면에서, 서열 번호 2와 또는 서열 번호 3의 성숙 서열과 80% 이상 동일한 아미노산 서열을 포함하는, 상기 제1 측면 내지 제5 측면 중 어느 한 측면의 재조합 폴리펩티드.
- [0056] 제7 측면에서, 서열 번호 2와 또는 서열 번호 3의 성숙 서열과 90% 이상 동일한 아미노산 서열을 포함하는, 상기 제1 측면 내지 제6 측면 중 어느 한 측면의 재조합 폴리펩티드.
- [0057] 제8 측면에서, 하나 이상의 다른 셀룰라제를 추가로 포함하는, 상기 제1 측면 내지 제7 측면 중 어느 한 측면의 재조합 폴리펩티드를 포함하는 조성물.
- [0058] 제9 측면에서, 하나 이상의 다른 셀룰라제가, 전혀 없거나 하나 이상의 다른 베타-글루코시다제, 하나 이상의 셀로비오히드롤라제 및 하나 이상의 엔도글루카나제로부터 선택되는, 제8 측면의 조성물.
- [0059] 제10 측면에서, 하나 이상의 헤미셀룰라제를 추가로 포함하는, 상기 제1 측면 내지 제7 측면 중 어느 한 측면의 재조합 폴리펩티드를 포함하는 조성물.
- [0060] 제11 측면에서, 하나 이상의 헤미셀룰라제를 추가로 포함하는, 상기한 바와 같은 제8 측면 또는 제9 측면의 조성물.
- [0061] 제12 측면에서, 하나 이상의 헤미셀룰라제가 하나 이상의 자일라나제, 하나 이상의 베타-자일로시다제, 및 하나 이상의 L-아라비노푸라노시다제로부터 선택되는, 상기한 바와 같은 제10 측면 또는 제11 측면의 조성물.
- [0062] 제13 측면에서, 제1 측면 내지 제7 측면 중 어느 한 측면의 재조합 폴리펩티드를 암호화하는 핵산.
- [0063] 제14 측면에서, 신호 서열을 추가로 포함하는 제13 측면의 핵산.
- [0064] 제15 측면에서, 신호 서열이 서열 번호 13-42로 이루어진 군으로부터 선택되는, 제14 측면의 핵산.
- [0065] 제16 측면에서, 조절 서열과 작동가능하게 조합된 제13 측면 내지 제15 측면 중 어느 한 측면의 핵산을 포함하는 발현 벡터.
- [0066] 제17 측면에서, 제16 측면의 발현 벡터를 포함하는 숙주 세포.
- [0067] 제18 측면에서, 세균 세포 또는 진균 세포인, 제17 측면의 숙주 세포. 다수의 세균 세포는 본원에 기재된 바와 같이 적합한 숙주 세포인 것으로 공지되어 있다. 다수의 진균 세포가 또한 적합하다. 일부 실시양태에서, 세균 또는 진균 숙주 세포는 특정 단량체성 당을 에탄올로 발효 또는 대사시킬 수 있는, 에타놀로젠인 것일 수 있다. 예를 들어, 세균 에타놀로젠 *지모모나스 모빌리스*는 본 개시내용의 베타-글루코시다제 폴리펩티드를 발현시키는 숙주 세포일 수 있다. 예를 들어, 진균 에타놀로젠 *사카로마이세스 세레비지에* 효모는 또한 본 개시내용의 베타-글루코시다제 폴리펩티드를 생성하는 숙주 세포로서 역할을 할 수 있다.
- [0068] 제19 측면에서, 제16 측면 또는 제17 측면의 숙주 세포 및 배양 배지를 포함하는 조성물.
- [0069] 제20 측면에서, 베타-글루코시다제를 생성하기에 적합한 조건하에, 배양 배지에서, 제17 측면 또는 제18 측면의 숙주 세포를 배양하는 것을 포함하는, 베타-글루코시다제의 생성 방법.
- [0070] 제21 측면에서, 배양 배지의 상청액 중, 상기 제20 측면의 방법에 따라 생성된 베타-글루코시다제를 포함하는

조성물.

[0071] 제22 측면에서, 리그노셀룰로오스계 바이오매스 기질을 제1 측면 내지 제7 측면 중 어느 한 측면의 폴리펩티드와, 또는 제21 측면의 조성물과 접촉시켜, 글루코오스 및/또는 다른 당을 수득하는 것을 포함하는, 리그노셀룰로오스계 바이오매스 기질의 가수분해 방법.

[0072] Mg3A 조성물 및 방법의 이들 및 다른 측면은 하기 설명으로부터 명백해질 것이다.

도면의 간단한 설명

[0073] 도 1은 pENTR/D-TOPO-Bg11(943/942) 벡터의 맵을 도시한다.

도 2는 pTrex3g 943/942 구축물의 맵을 도시한다.

도 3a-3f는 50℃, 및 1.5시간에서, 기질로서 인산 팽윤된 셀룰로오스 (PASC)를 사용하여 Mg3A 대 기준 *트리코테르마 레에세이* Bg11의 가수분해 성능을 비교하기 위해 다소 상이한 Mg3A 효소 제제를 사용한 2 세트의 실험을 제공하며, 여기서 Mg3A 및 Bg11을 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된 백그라운드(background) 전체 셀룰라제에 첨가하였고, 베타-글루코시다제+전체 셀룰라제 혼합물을 다양한 베타-글루코시다제 용량에서 PASC 기질과 혼합하였다. 도 3a는 본원에서 실시예 4-A에 기재된 조건에 따라서 다양한 베타-글루코시다제 용량에서 글루칸 전환율 (%)의 측정 및 비교를 도시한다. 도 3b는 본원에서 실시예 4-A에 기재된 조건에 따라서 다양한 베타-글루코시다제 용량에서 총 글루코오스 생성의 측정 및 비교를 도시한다. 도 3c는 본원에서 실시예 4-A에 기재된 조건에 따라서, 다양한 베타-글루코시다제 용량에서 PASC를 가수분해함으로써 생성된 셀로비오스의 측정 및 비교를 도시한다. 도 3d는 본원에서 실시예 4-B에 기재된 조건에 따라서 다양한 베타-글루코시다제 용량에서 글루칸 전환율 (%)의 측정 및 비교를 도시한다. 도 3e는 본원에서 실시예 4-B에 기재된 조건에 따라서 다양한 베타-글루코시다제 용량에서 총 글루코오스 생성의 측정 및 비교를 도시한다. 도 3f는 본원에서 실시예 4-B에 기재된 조건에 따라서, 다양한 베타-글루코시다제 용량에서 PASC를 가수분해함으로써 생성된 셀로비오스의 측정 및 비교를 도시한다.

도 4a-4b는 2일 동안 50℃에서, 기질로서 희석 암모니아로 전처리된 옥수수 대 (DACS)를 사용한 Mg3A 대 기준 *트리코테르마 레에세이* Bg11의 가수분해 성능의 비교를 제공하며, 여기서 Mg3A 및 Bg11을 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된 백그라운드 전체 셀룰라제에 첨가하였고, 베타-글루코시다제+셀룰라제 혼합물을 셀룰로오스 용량에 대해 다양한 총 단백질로 DACS 기질과 혼합하였다. 도 4a는 셀룰로오스 용량에 대해 다양한 총 단백질에서 총 글루칸 전환의 측정 및 비교를 도시한다. 도 4b는 셀룰로오스 용량에 대해 다양한 총 단백질에서 총 글루코오스 생성의 측정 및 비교를 도시한다.

도 5a-5b는 2일 동안 50℃에서, 기질로서 희석 암모니아로 전처리된 옥수수 대 (DACS)를 사용한 Mg3A 대 기준 *트리코테르마 레에세이* Bg11의 가수분해 성능의 비교를 제공하며, 여기서 다양한 용량의 Mg3A 및 Bg11을 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된 전체 셀룰라제 (이는 일정량 13.4 mg 단백질/g 셀룰로오스로 사용됨)에 첨가하였다. 도 5a는 다양한 베타-글루코시다제 로딩 (여기서 전체 셀룰라제 백그라운드는 일정량 13.4 mg 단백질/g 셀룰로오스로 유지됨)에서 총 글루칸 전환의 측정 및 비교를 도시한다. 도 5b는 다양한 베타-글루코시다제 로딩 (여기서 전체 셀룰라제 백그라운드는 일정량 13.4 mg 단백질/g 셀룰로오스로 유지됨)에서 총 글루코오스 생성의 측정 및 비교를 도시한다.

도 6a-6c는 2일 동안 55℃에서, 기질로서 희석 암모니아로 전처리된 옥수수 대 (DACS)를 사용한 Mg3A 대 기준 *트리코테르마 레에세이* Bg11의 가수분해 성능의 비교를 제공하며, 여기서 다양한 용량의 Mg3A 및 Bg11을 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된 전체 셀룰라제 백그라운드 (이는 10 mg 단백질/g 셀룰로오스를 함유함)에 첨가하였다. 도 6a는 다양한 베타-글루코시다제 로딩 (여기서 전체 셀룰라제 백그라운드는 일정량 10 mg 단백질/g 셀룰로오스로 유지됨)에서 총 글루칸 전환의 측정 및 비교를 도시한다. 도 6b는 다양한 베타-글루코시다제 로딩 (여기서 전체 셀룰라제 백그라운드는 일정량 10 mg 단백질/g 셀룰로오스로 유지됨)에서 총 글루코오스 생성의 측정 및 비교를 도시한다. 도 6c는 55℃에서 수행된 당화 반응에서 다양한 용량으로 Mg3A 및 Bg11의 존재에 의해 셀로비오스의 고갈의 용량 곡선을 도시한다.

도 7은 *사카로마이세스 세레비지에* 에타놀로겐에서 Mg3A 폴리펩티드의 발현을 위해 최적화되고 합성되는 Mg3A 유전자를 포함하는 효모 서틀 벡터 pSC11 구축물을 도시한다.

도 8은 *지모모나스 모빌리스* 에타놀로겐에서 Mg3A 폴리펩티드의 발현을 위해 최적화되고 합성되는 Mg3A 유전자

를 포함하는 지모모나스 모빌리스 통합 벡터 pZC11을 도시한다.

도 9a-9e는 본 개시내용의 서열 및 서열 식별자를 도시한다.

발명을 실시하기 위한 구체적인 내용

I. 개관

본원에는 마그나포르테 그리세아로부터 글리코실 히드롤라제 패밀리 3에 속하는 재조합 베타-글루코시다제 Mg3A와 관련된 조성물 및 방법이 기재되어 있다. 본 조성물 및 방법은, 부분적으로는, 효소 조성물을 사용하여 리그노셀룰로오스계 바이오매스 물질 또는 원료를 가수분해하는 경우, 예를 들어, *트리코테르마 레에세이*의 공지된 기준인 고 충실도 베타-글루코시다제 Bgl1보다, 재조합 Mg3A 폴리펩티드가 효소 조성물의 성분으로서 더 높은 셀룰라제 활성을 갖고 더 강건하다는 관찰을 기반으로 한다. Mg3A 폴리펩티드의 이들 특질은 이들, 또는 그의 변이체를, 예를 들어, 리그노셀룰로오스계 바이오매스 원료의 전환 또는 가수분해를 포함한 수많은 공정에서 사용하기에 적합하게 한다.

본 조성물 및 방법을 더 상세히 설명하기 전에, 본 조성물 및 방법은 기재된 특성의 실시양태로 제한되는 것이 아니고, 물론, 그 자체로 변화할 수 있다는 점이 이해되어야 한다. 본원에서 사용되는 전문용어는 단지 특성의 실시양태를 기재하는 목적을 위한 것이고, 제한하려고 의도된 것이 아닌데, 그 이유는 본 조성물 및 방법의 범위는 첨부된 청구범위에 의해서만 제한될 것이기 때문이라는 점이 또한 이해되어야 한다.

어떤 범위의 값들이 제공되는 경우, 문맥이 명확하게 달리 지시하지 않는 한, 하한치의 1/10 단위로, 그 범위의 상한치 및 하한치 사이의 각각의 개재 값 및 그 명시된 범위 내에서 임의의 다른 명시된 또는 개재된 값이 본 조성물 및 방법에 포함되는 것으로 이해되어야 한다. 명시된 범위에서 임의의 구체적으로 배제된 한계치에 따라, 이들 더 작은 범위의 상한치 및 하한치는 더 작은 범위에 독립적으로 포함될 수 있고, 또한 본 조성물 및 방법에 포함된다. 명시된 범위가 한계치 중 하나 또는 둘 다를 포함하는 경우, 그러한 포함된 한계치 중 하나 또는 둘 다를 배제하는 범위 또한 본 조성물 및 방법에 포함된다.

특정 범위가 본원에 제시되며 여기서 용어 "약"이 수치에 선행한다. 용어 "약"은 본원에서 사용되어 그것이 선행하는 정확한 수에 뿐만 아니라, 상기 용어가 선행하는 수에 가깝거나 또는 근사한 수에 문자 그대로의 지지를 제공한다. 수가 구체적으로 열거된 수에 가깝거나 근사한지를 결정하는데 있어서, 가깝거나 근사한 열거되지 않은 수는, 그것이 제시되는 맥락에서, 구체적으로 열거된 수의 실질적인 등가물을 제공할 수 있다. 예를 들어, 수치와 관련하여, 용어 "약"은, 맥락에서 그 용어가 달리 구체적으로 한정되지 않는 한, 수치의 -10% 내지 +10%의 범위를 지칭한다. 또 다른 예에서, 문구 "약 6의 pH 값"은, pH 값이 달리 구체적으로 한정되지 않는 한, 5.4 내지 6.6의 pH 값을 지칭한다.

본원에 제공된 표제는 전체로서 명세서를 참조로 할 수 있는 본 조성물 및 방법의 다양한 측면 또는 실시양태의 제한이 아니다. 따라서, 바로 아래에 정의된 용어들은 전체로서 명세서를 참조로 더 충분히 정의된다.

본 문서는 읽기 쉽게 하기 위해 다수의 섹션으로 편성되어 있으나; 독자는 한 섹션에서 이루어진 서술이 다른 섹션에 적용될 수 있음을 이해할 것이다. 이러한 방식으로, 개시내용의 상이한 섹션에 사용된 표제는 제한적인 것으로 해석되어서는 안된다.

달리 정의되지 않는 한, 본원에 사용된 모든 기술적 및 과학적 용어는 본 조성물 및 방법이 속하는 관련 기술 분야의 통상의 기술자가 흔히 이해하는 바와 동일한 의미를 갖는다. 비록 본원에 기재된 것과 유사하거나 균등한 임의의 방법 및 물질이 또한 본 조성물 및 방법의 실시 또는 시험에서 사용될 수 있긴 하지만, 대표적인 예시적 방법 및 물질이 이제 기재된다.

본 명세서에서 인용된 모든 간행물 및 특허는 각각의 개별 간행물 또는 특허가 구체적으로 그리고 개별적으로 참조로 포함되는 것으로 지시되는 경우 본원에 참조로 포함되며 간행물이 인용되는 것과 관련하여 방법 및/또는 물질을 개시하고 기재하기 위해 본원에 참조로 포함된다. 임의의 간행물의 인용은 출원일 이전에 그의 개시내용에 대한 것이며, 본 조성물 및 방법이 선행 발명에 의하여 이러한 간행물보다 선행할 자격이 없다는 인정으로서 해석되어서는 안된다. 추가로, 제공된 간행물의 날짜는 독립적으로 확인될 필요가 있을 수 있는 실제 공개 날짜와 상이할 수 있다.

이러한 상세한 설명에 따라서, 하기 약어 및 정의가 적용된다. 단수 형태 표현은 문맥이 달리 명확히 지시하지 않는 한 복수의 지시대상을 포함한다는 점을 주목한다. 따라서, 예를 들어, "효소"라는 언급은 복수의 이러한

효소를 포함하고, "투여"라는 언급은 하나 이상의 투여 및 통상의 기술자에게 공지된 그의 등가물 등을 포함한다.

[0084] 청구범위는 임의의 임의적 요소를 배제하도록 초안이 작성될 수 있다는 점을 추가로 주목한다. 그와 같이, 이러한 서술은 청구범위 요소의 열거와 관련하여 "오로지," "단지" 등과 같은 배타적인 전문용어의 사용, 또는 "부정적인" 제한의 사용의 선행 기준으로서의 역할을 하는 것으로 의도된다.

[0085] 용어 "재조합"은, 대상 세포, 핵산, 폴리펩티드/효소 또는 벡터와 관련하여 사용되는 경우, 대상이 그의 천연 상태에서부터 변형되었다는 것을 나타낸다. 따라서, 예를 들어, 재조합 세포는, 세포의 천연 (비-재조합) 형태 내에서 발견되지 않는 유전자를 발현시키거나, 자연에서 발견되는 것과는 상이한 수준에서 또는 상이한 조건하에 천연 유전자를 발현시킨다. 재조합 핵산은 하나 이상의 뉴클레오티드에 의해 천연 서열과 상이할 수 있고/거나 발현 벡터 내에서, 이중 서열, 예를 들어, 이중 프로모터, 분비를 가능하게 하는 신호 서열 등에 작동적으로 연결된다. 재조합 폴리펩티드/효소는 하나 이상의 아미노산에 의해 천연 서열과 상이할 수 있고/거나, 이중 서열과 융합된다. 베타-글루코시다제를 암호화하는 핵산을 포함하는 벡터는, 예를 들어, 재조합 벡터이다.

[0086] 본원에 사용된 바와 같은 용어 "로 본질적으로 이루어진"은, 용어 뒤의 성분(들)이 전체 조성물의 30 중량% 미만인 총량으로 다른 공지된 성분(들)의 존재하에 있으며 성분(들)의 작용 또는 활성에 기여하거나 지장을 주지 않는, 조성물을 지칭한다는 점을 추가로 주목한다.

[0087] 본원에 사용된 바와 같은 용어 "포함하는"은, 용어 "포함하는" 뒤의 성분(들)을 포함하지만 그에 제한되지는 않음을 의미한다는 점을 추가로 주목한다. 용어 "포함하는" 뒤의 성분(들)은 필요하거나 강제적이거나, 성분(들)을 포함하는 조성물은 다른 강제적인 아닌 또는 임의적 성분(들)을 추가로 포함할 수 있다.

[0088] 본원에 사용된 바와 같은 용어 "로 이루어진"은, 용어 "로 이루어진" 뒤의 성분(들)을 포함하고 그에 제한됨을 의미한다는 점을 또한 주목한다. 따라서 용어 "로 이루어진" 뒤의 성분(들)은 필요하거나 강제적이고, 어떤 다른 성분(들)도 조성물에 존재하지 않는다.

[0089] 본 개시내용을 읽으면 통상의 기술자에게 분명해지는 바와 같이, 본원에 기재되고 예시되는 개별 실시양태 각각은 본원에 기재된 본 조성물 및 방법의 범주 또는 취지로부터 벗어남이 없이 다른 몇몇 실시양태 중 임의의 것의 특징로부터 용이하게 분리되거나 그 특징과 조합될 수 있는 구별되는 성분 및 특징을 갖는다. 임의의 열거된 방법은 열거된 사상(event)의 순으로 또는 논리적으로 가능한 임의의 다른 순서로 수행될 수 있다.

[0090] II. 정의

[0091] "베타-글루코시다제"는 E.C. 3.2.1.21의 베타-D-글루코시드 글루코히드롤라제를 지칭한다. 따라서 용어 "베타-글루코시다제 활성"은 베타-D-글루코오스 또는 셀로비오스의 가수분해를 촉매하여 D-글루코오스를 방출하는 능력을 지칭한다. 베타-글루코시다제 활성은, 본 개시내용의 실시예 2C에 기재된 바와 같이, 예를 들어, 셀로비오스 기질의 가수분해를 촉매하여 D-글루코오스를 산출하는 효소의 능력을 측정하는 셀로비아제 검정을 사용하여 결정할 수 있다.

[0092] 본원에 사용된 바와 같이, "Mg3A" 또는 "Mg3A 폴리펩티드"는 *마그나포르테 그리세아*로부터 유래된 글리코실 히드롤라제 패밀리 3에 속하는 베타-글루코시다제 (예를 들어, 재조합 베타-글루코시다제) (및 그의 변이체)를 지칭하며, 이는 기준 베타-글루코시다제인, 서열 번호 4의 아미노산 서열을 갖는 야생형 *트리코테르마 레에세이* Bg11 폴리펩티드와 비교하는 경우 리그노셀룰로오스계 바이오매스 기질을 가수분해하는 개선된 성능을 갖는다. 본 조성물 및 방법의 측면에 따르면, Mg3A 폴리펩티드는 서열 번호 2에 도시된 아미노산 서열을 갖는 것뿐만 아니라, 서열 번호 2의 아미노산 서열, 또는 성숙 서열 서열 번호 2, 또는 서열 번호 2의 길이에 있어서 적어도 100 개 잔기의 단편에 대하여 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 적어도 99% 서열 동일성을 갖는 유도체 또는 변이체 폴리펩티드를 포함하며, 여기서 Mg3A 폴리펩티드는 베타-글루코시다제 활성을 갖고 셀로비오스의 D-글루코오스로의 전환을 촉매할 수 있을 뿐만 아니라, *트리코테르마 레에세이* Bg11보다 더 높은 베타-글루코시다제 활성을 갖고 셀로비오스의 D-글루코오스로의 전환을 촉매하는 더 큰 능력도 갖는다.

[0093] 본 개시내용의 조성물 및 방법에 사용되는 Mg3A 폴리펩티드는 적어도 5%, 적어도 10%, 바람직하게는 적어도 20%, 더 바람직하게는 적어도 30%, 더욱더 바람직하게는 적어도 40%, 더 바람직하게는 적어도 50%, 더욱더 바람직하게는 적어도 60%, 바람직하게는 적어도 70%, 더 바람직하게는 적어도 90%, 더욱더 바람직하게는 적어도 100% 또는 그 초과, 서열 번호 2의 아미노산 서열의 폴리펩티드; 또는 서열 번호 2의 잔기 19 내지 873으로

이루어진 폴리펩티드; 또는 성숙 서열 서열 번호 3의 베타-글루코시다제 활성을 가질 것이다.

[0094] "패밀리 3 글리코실 히드롤라제" 또는 "GH3"은 문헌 [Henrissat, Biochem. J. 280:309-316 (1991)], 및 [Henrissat &airoch, Biochem. J., 316:695-696 (1996)]에 의한 분류에 따른 글리코실 히드롤라제 패밀리 3의 정의 내에 해당하는 폴리펩티드를 지칭한다.

[0095] 본원에 기재된 본 조성물 및 방법에 따른 Mg3A 폴리펩티드는 단리 또는 정제될 수 있다. 정제 또는 단리된 Mg3A 폴리펩티드가 자연적으로 연관되어 있는 자연 발생 구성성분 중 일부 또는 모두로부터 Mg3A를 분리하는 것에 의해 그의 자연 상태에서부터 변경됨을 의미한다. 이러한 단리 또는 정제는 최종 조성물 중 바람직하지 않은 전체 세포, 세포 잔해, 불순물, 외래 단백질, 또는 효소를 제거하는, 관련 기술분야-인식된 분리 기법, 예컨대 이온 교환 크로마토그래피, 친화도(affinity) 크로마토그래피, 소수성 분리, 투석, 프로테아제 처리, 황산암모늄 침전 또는 다른 단백질 염 침전, 원심분리, 크기 배제 크로마토그래피, 여과, 미세여과, 겔 전기영동 또는 구배 상에서의 분리에 의해 완수될 수 있다. 그 다음에, Mg3A-함유 조성물에 추가의 이점을 제공하는 구성성분, 예를 들어, 활성화제, 항-억제제, 바람직한 이온, pH를 제어하는 화합물 또는 다른 효소 또는 화합물질을 첨가하는 것이 추가로 가능하다.

[0096] 본원에 사용된 바와 같이, "미생물"은 세균, 진균, 바이러스, 원생동물, 및 다른 미생물 (microbe 또는 microscopic organism)을 지칭한다.

[0097] 본원에 사용된 바와 같이, 폴리펩티드의 "유도체" 또는 "변이체"는, C- 및 N-말단 중 하나 또는 둘 다에 하나 이상의 아미노산의 부가, 아미노산 서열에서의 하나 또는 다수의 상이한 위치에서 하나 이상의 아미노산의 치환, 폴리펩티드의 한쪽 또는 양쪽 말단에서 또는 아미노산 서열에서의 하나 이상의 부위에서 하나 이상의 아미노산의 결실, 또는 아미노산 서열에서의 하나 이상의 부위에서 하나 이상의 아미노산의 삽입에 의해, 전구체 폴리펩티드 (예를 들어, 천연 폴리펩티드)로부터 유래되는 폴리펩티드를 의미한다. Mg3A 유도체 또는 변이체의 제조는 임의의 편리한 방식으로, 예를 들어, 천연 폴리펩티드를 암호화하는 DNA 서열을 변형하고, 그 DNA 서열을 적합한 숙주로 형질전환하고, 변형된 DNA 서열을 발현시켜 유도체/변이체 Mg3A를 형성함으로써 달성될 수 있다. 유도체 또는 변이체는, 예를 들어, 글리코실화 또는 달리 Mg3A 폴리펩티드의 특성을 변화시켜, 화학적으로 변형되는 Mg3A 폴리펩티드를 추가로 포함한다. Mg3A의 유도체 및 변이체는 본 조성물 및 방법에 의해 포함되지만, 이러한 유도체 및 변이체는 동일한 리그노셀룰로오스계 바이오매스 기질 가수분해 조건하에, 서열 번호 4의 야생형 *트리코테르마 레세이* Bgl1의 것과 비교하는 경우 개선된 베타-글루코시다제 활성을 나타낼 것이다.

[0098] 특정 측면에서, 본원에서의 조성물 및 방법의 Mg3A 폴리펩티드는 또한 베타-글루코시다제 활성을 갖는 폴리펩티드 단편 또는 폴리펩티드의 기능적 단편을 포함할 수 있고, 이는 모 폴리펩티드로부터 유래되며, 이는 서열 번호 3을 포함하거나 그로 이루어진 성숙 서열, 또는 서열 번호 2를 포함하거나 그로 이루어진 전장(full length) 폴리펩티드일 수 있다. 기능적 폴리펩티드는 N-말단 영역, 또는 C-말단 영역 중 하나, 또는 양 영역에서 절두되어 모 폴리펩티드의 단편을 생성할 수 있다. 본 개시내용의 목적상, 기능적 단편은 모 폴리펩티드의 것의 적어도 20%, 더 바람직하게는 적어도 30%, 40%, 50%, 또는 바람직하게는, 적어도 60%, 70%, 80%, 또는 더욱더 바람직하게는 적어도 90%의 베타-글루코시다제 활성을 가져야 한다.

[0099] 특정 측면에서, Mg3A 유도체/변이체는 서열 번호 2의 아미노산 서열, 또는 성숙 서열 서열 번호 3에 대하여 75% 내지 99% (또는 그 초과) 아미노산 서열 동일성의 어딘가, 예를 들어, 서열 번호 2의 아미노산 서열, 또는 성숙 서열 서열 번호 3에 대하여 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 가질 것이다. 일부 실시양태에서, 아미노산 치환은 L-아미노산을 사용하는 "보존적 아미노산 치환"이고, 여기서 하나의 아미노산이 또 다른 생물학적으로 유사한 아미노산에 의해 대체된다. 보존적 아미노산 치환은 치환되는 아미노산의 일반적인 전하, 소수성/친수성, 및/또는 입체 벌크를 보존하는 것이다. 보존적 치환의 예는 하기 군 사이의 것들이다: Gly/Ala, Val/Ile/Leu, Lys/Arg, Asn/Gln, Glu/Asp, Ser/Cys/Thr, 및 Phe/Trp/Tyr. 유도체는, 예를 들어, 겨우 1 내지 10개 아미노산 잔기, 예컨대 6-10개, 겨우 5개, 겨우 4개, 3개, 2개, 또는 심지어는 1개 아미노산 잔기만큼 상이할 수 있다. 일부 실시양태에서, Mg3A 유도체는 N-말단 및/또는 C-말단 결실을 가질 수 있고, 여기서 결실된 말단 부분(들)을 배제한 Mg3A 유도체는 서열 번호 2 또는 서열 번호 3에서의 인접 서브-영역과 동일하다.

[0100] 본원에 사용된 바와 같이, 본원에서 동정된 아미노산 또는 뉴클레오타이드 서열과 관련하여 "퍼센트 (%) 서열 동일성"은, 필요한 경우, 최대 퍼센트 서열 동일성을 달성하기 위해 서열을 정렬하고 겹을 도입한 후, 그리고, 서열 동일성의 일부로서 임의의 보존적 치환을 고려하지 않고, Mg3A 서열에서의 아미노산 잔기 또는 뉴클레오타이드

와 동일한 후보 서열에서의 아미노산 잔기 또는 뉴클레오타이드의 백분율로서 정의된다.

- [0101] "상동체"란 대상 아미노산 서열 및 대상 뉴클레오타이드 서열과의 특정된 정도의 동일성을 갖는 실체를 의미해야 한다. 상동 서열은 통상적인 서열 정렬 툴 (예를 들어, 클러스탈(Clustal), BLAST 등)을 사용하여, 대상 서열과 적어도 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 심지어는 99% 동일한 아미노산 서열을 포함하도록 취해진다. 전형적으로, 상동체는 달리 규정되지 않는 한, 대상 아미노산 서열과 동일한 활성 부위 잔기를 포함할 것이다.
- [0102] 서열 정렬을 수행하고 서열 동일성을 결정하는 방법은 통상의 기술자에게 공지되어 있어, 과도한 실험 없이 수행될 수 있으며, 동일성 값의 계산은 명확하게 수득될 수 있다. 예를 들어, 문헌 [Ausubel *et al.*, eds. (1995) *Current Protocols in Molecular Biology*, Chapter 19 (Greene Publishing and Wiley-Interscience, New York)]; 및 ALIGN 프로그램 (문헌 [Dayhoff (1978) in *Atlas of Protein Sequence and Structure* 5:Suppl. 3 (National Biomedical Research Foundation, Washington, D.C.)] 참조. 다수의 알고리즘이 서열을 정렬하고 서열 동일성을 결정하기 위해 이용가능하며, 예를 들어, 문헌 [Needleman *et al.* (1970) *J. Mol. Biol.* 48:443]의 상동성 정렬 알고리즘; 문헌 [Smith *et al.* (1981) *Adv. Appl. Math.* 2:482]의 국소 상동성 알고리즘; 문헌 [Pearson *et al.* (1988) *Proc. Natl. Acad. Sci.* 85:2444]의 유사성 검색 방법; 스미스-워터먼 (Smith-Waterman) 알고리즘 (문헌 [Meth. Mol. Biol. 70:173-187 (1997)]); 및 BLASTP, BLASTN, 및 BLASTX 알고리즘 (문헌 [Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410] 참조)를 포함한다.
- [0103] 이들 알고리즘을 사용하는 컴퓨터화된 프로그램이 또한 이용가능하며, 다음을 포함하지만 그에 제한되지는 않는다: ALIGN 또는 메갈라인(Megalign) (DNASTAR) 소프트웨어, 또는 WU-BLAST-2 (문헌 [Altschul *et al.*, (1996) *Meth. Enzym.*, 266:460-480]); 또는 GAP, BESTFIT, BLAST, FASTA, 및 TFASTA [지네틱스 컴퓨팅 그룹 (Genetics Computing Group: GCG, 미국 위스콘신주 매디슨) 패키지, 버전(Version) 8로 이용가능함]; 및 PC/유전자 프로그램에서의 CLUSTAL [인텔리지네틱스(Intelligenetics) (캘리포니아주 마운틴 뷰)에 의함]. 통상의 기술자는 비교될 서열의 길이에 걸쳐 최대 정렬을 달성하는데 필요한 알고리즘을 포함하여, 정렬을 측정하기 위한 적절한 파라미터를 결정할 수 있다. 바람직하게는, 서열 동일성은 프로그램에 의해 결정된 디폴트 파라미터를 사용하여 결정된다. 구체적으로, 서열 동일성은 하기 디폴트 파라미터로 클러스탈 W (문헌 [Thompson J.D. *et al.* (1994) *Nucleic Acids Res.* 22:4673-4680])를 사용하여 결정될 수 있다:
- [0104] 갭 오픈링 페널티(Gap opening penalty): 10.0
- [0105] 갭 익스텐션 페널티(Gap extension penalty): 0.05
- [0106] 단백질 중량 매트릭스: BLOSUM 시리즈
- [0107] DNA 중량 매트릭스: IUB
- [0108] 지연 발산(Delay divergent) 서열 %: 40
- [0109] 갭 분리 거리 8
- [0110] DNA 전이 중량 0.50
- [0111] 친수성 잔기 열거: GPSNDQEKR
- [0112] 네거티브 매트릭스 사용: OFF
- [0113] 잔기 특이적 페널티 토글(Toggle) ON
- [0114] 친수성 페널티 토글 ON
- [0115] 엔드 갭 분리 페널티 토글 OFF
- [0116] 본원에 사용된 바와 같이, "발현 벡터"는 적합한 숙주에서 DNA의 발현에 영향을 미칠 수 있는 적합한 제어 서열에 작동가능하게 연결되는 DNA 서열을 포함하는 DNA 구축물을 의미한다. 이러한 제어 서열은 전사에 영향을 미치는 프로모터, 전사를 제어하는 임의적 오퍼레이터(operator) 서열, mRNA 상에 적합한 리보솜-결합 부위를 암호화하는 서열, 및 전사 및 번역의 종결을 제어하는 서열을 포함할 수 있다. 상이한 세포 유형은 상이한 발현 벡터로 사용될 수 있다. *바실루스 서브틸리스(Bacillus subtilis)*에서 사용되는 벡터를 위한 예시적 프로모터는 AprE 프로모터이고; *스트렙토마이세스 리비단스(Streptomyces lividans)*에서 사용되는 예시적 프로모터는 A4 프로모터 (*아스페르길루스 니게르*로부터); *이. 콜라이*에서 사용되는 예시적 프로모터는 Lac 프로모터, *사카로마*

이세스 세레비지에에서 사용되는 예시적 프로모터는 *PGK1*, *아스페르길루스 니게르*에서 사용되는 예시적 프로모터는 *glcA*이고, *트리코테르마 레에세이*를 위한 예시적 프로모터는 *cbh1*이다. 벡터는 플라스미드, 파지 입자, 또는 단순히 잠재적 게놈 삽입물일 수 있다. 일단 적합한 숙주로 형질전환되면, 벡터는 숙주 게놈과 무관하게 기능하고 복제할 수 있거나, 적합한 조건하에, 게놈 자체에 통합될 수 있다. 본 명세서에서, 플라스미드와 벡터는 때때로 교체사용이 가능하다. 그러나, 본 조성물 및 방법은 균등한 기능을 하고 관련 기술 분야에 공지되어 있거나, 공지된 발현 벡터의 다른 형태를 포함하도록 의도된다. 따라서, 다종다양한 숙주/발현 벡터 조합이 본원에 기재된 DNA 서열을 발현시키는데 사용될 수 있다. 유용한 발현 벡터는, 예를 들어, 염색체성, 비염색체성 및 합성 DNA 서열의 절편, 예컨대 SV40의 다양한 공지된 유도체 및 공지된 세균성 플라스미드, 예를 들어, *col E1*, *pCR1*, *pBR322*, *pMb9*, *pUC 19* 및 그의 유도체를 포함한, *이. 콜라이*로부터의 플라스미드, 더 넓은 숙주 범위의 플라스미드, 예를 들어, *RP4*, 파지 DNA, 예를 들어, 파지 λ 의 수많은 유도체, 예를 들어, *NM989*, 및 다른 DNA 파지, 예를 들어, *M13* 및 사상 외가닥 DNA 파지, 효모 플라스미드, 예컨대 2μ 플라스미드 또는 그의 유도체, 진핵 세포에서 유용한 벡터, 예컨대 동물 세포에서 유용한 벡터 및 플라스미드와 파지 DNA의 조합으로부터 유래된 벡터, 예컨대 파지 DNA 또는 다른 발현 제어 서열을 사용하기 위해 변형된 플라스미드로 이루어질 수 있다. 본 조성물 및 방법의 발현 벡터를 사용한 발현 기법은 관련 기술분야에 공지되어 있고, 일반적으로, 예를 들어, 문헌 [Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Press (1989)]에 기재되어 있다. 종종, 본원에 기재된 DNA 서열을 포함하는 이러한 발현 벡터는 통합 현상(integration event)을 통해 특정의 종의 게놈에 직접 삽입에 의해 단세포 숙주로 형질전환된다 (예를 들어, 문헌 [Bennett & Lasure, *More Gene Manipulations in Fungi*, Academic Press, San Diego, pp. 70-76 (1991)] 및 진균 숙주에서 표적 게놈 삽입을 기재하는 거기에 인용된 논문 참조).

- [0117] 본원에 사용된 바와 같이, "숙주 균주" 또는 "숙주 세포"는 본 조성물 및 방법에 따른 DNA를 포함하는 발현 벡터에 적합한 숙주를 의미한다. 본 조성물 및 방법에 유용한 숙주 세포는 일반적으로, 발현이 달성될 수 있는 임의의 형질전환가능한 미생물을 포함한 원핵생물 또는 진핵생물 숙주이다. 구체적으로, 숙주 균주는 *바실러스 서브틸리스*, *스트렙토마이세스 리비단스*, *에스케리키아 콜라이*(*Escherichia coli*), *트리코테르마 레에세이*, *사카로마이세스 세레비지에* 또는 *아스페르길루스 니게르*일 수 있다. 특정 실시양태에서, 숙주 세포는 에타놀로겐 미생물일 수 있고, 이는, 예를 들어, 효모, 예컨대 *사카로마이세스 세레비지에* 또는 세균 에타놀로겐, 예컨대 *지모모나스 모빌리스*일 수 있다. *사카로마이세스 세레비지에* 또는 *지모모나스 모빌리스*가 숙주 세포로서 사용되는 경우, 그리고 베타-글루코시다제 유전자가 숙주 세포로부터 분비되도록 하는 것이 아니고 세포 내로 발현된다면, 셀로비오스 수송체 유전자는 숙주 세포에 도입되어 세포내에서 발현된 베타-글루코시다제가 셀로비오스 기질에 작용하여 글루코오스를 유리시키도록 할 수 있고, 이는 그 다음에, 미생물에 의해 후속적으로 또는 즉시 대사되어 에탄올로 전환될 것이다.
- [0118] 숙주 세포는 재조합 DNA 기법을 사용하여 구축된 벡터로 형질전환 또는 형질감염된다. 이러한 형질전환된 숙주 세포는 *Mg3A* (및 그의 유도체 또는 변이체 (돌연변이체))를 암호화하는 벡터를 복제하는 것과 원하는 펩티드 생성물을 발현시키는 것 중 하나 또는 둘 다 가능할 수 있다. 본 조성물 및 방법에 따른 특정 실시양태에서, "숙주 세포"는 *트리코테르마* 종의 세포로부터 추출된 세포 및 원형질체 둘 다를 의미한다.
- [0119] 세포와 관련하여 사용된, 용어 "형질전환된", "안정적으로 형질전환된" 또는 "유전자 도입된"은 세포가 그의 게놈에 통합되거나 다세대를 통해 유지되는 에피솜으로서 담지되는 비천연 (예를 들어, 이중) 핵산 서열을 함유하는 것을 의미한다.
- [0120] 핵산 서열을 세포에 삽입하는 맥락에서 용어 "도입된"은, 관련 기술분야에 공지된 바와 같이, "형질감염", 또는 "형질전환" 또는 "형질도입"을 의미한다.
- [0121] "숙주 균주" 또는 "숙주 세포"는 발현 벡터, 파지, 바이러스, 또는 다른 DNA 구축물 [관심 폴리펩티드 (예를 들어, 베타-글루코시다제)를 암호화하는 폴리뉴클레오티드 포함]이 도입된 유기체이다. 예시적인 숙주 균주는 관심 폴리펩티드를 발현시킬 수 있는 미생물 세포 (예를 들어, 세균, 사상균, 및 효모)이다. 용어 "숙주 세포"는 세포로부터 추출된 원형질체를 포함한다.
- [0122] 폴리뉴클레오티드 또는 폴리펩티드와 관련하여 용어 "이중"은 숙주 세포에서 자연 발생하지 않는 폴리뉴클레오티드 또는 폴리펩티드를 지칭한다.
- [0123] 폴리뉴클레오티드 또는 폴리펩티드와 관련하여 용어 "내인성"은 숙주 세포에서 자연 발생하는 폴리뉴클레오티드 또는 폴리펩티드를 지칭한다.

- [0124] 용어 "발현"은 핵산 서열을 기반으로 폴리펩티드가 생성되는 공정을 지칭한다. 상기 공정은 전사 및 번역 둘 다를 포함한다.
- [0125] 따라서 리그노셀룰로오스계 바이오매스 기질을 에탄올로 전환시키는 공정은, 일부 실시양태에서, 두가지 베타-글루코시다제 활성을 포함할 수 있다. 예를 들어, 제1 베타-글루코시다제 활성은 당화 또는 가수분해 단계 동안 리그노셀룰로오스계 바이오매스 기질에 적용될 수 있고, 제2 베타-글루코시다제 활성은 당화 또는 가수분해 단계로부터 생성된 단량체성 또는 발효성 당이 대사되는 발효 단계에서 에타놀로젠 미생물의 부분으로서 적용될 수 있다. 제1 및 제2 베타-글루코시다제 활성은, 일부 실시양태에서, 동일한 베타-글루코시다제 폴리펩티드의 존재로부터 생길 수 있다. 예를 들어, 당화에서의 제1 베타-글루코시다제 활성은 본 발명의 Mg3A 폴리펩티드의 존재로부터 생길 수 있으며, 한편 발효 단계에서의 제2 베타-글루코시다제 활성은 에타놀로젠 미생물에 의한 상이한 베타-글루코시다제의 발현으로부터 생길 수 있다. 또 다른 예에서, 제1 및 제2 베타-글루코시다제 활성은 당화 또는 가수분해 단계 및 발효 단계에서 동일한 폴리펩티드의 존재로부터 생길 수 있다. 예를 들어, 본 발명의 동일한 Mg3A 폴리펩티드는, 일부 실시양태에서, 가수분해 또는 당화 단계 및 발효 단계 둘 다에 베타-글루코시다제 활성을 제공할 수 있다.
- [0126] 특정 다른 실시양태에서, 리그노셀룰로오스계 바이오매스 기질을 에탄올로 전환시키는 공정은, 두가지 베타-글루코시다제 활성을 포함할 수 있으며, 한편 당화 또는 가수분해 단계 및 발효 단계는, 예를 들어, 동일 탱크에서 동시에 일어난다. 2개 이상의 베타-글루코시다제 폴리펩티드는 베타-글루코시다제 활성에 기여할 수 있고, 그 중 하나는 본 발명의 Mg3A 폴리펩티드일 수 있다.
- [0127] 특정 추가 실시양태에서, 리그노셀룰로오스계 바이오매스를 에탄올로 전환시키는 공정은 단일의 베타-글루코시다제 활성을 포함할 수 있으며, 한편 당화 또는 가수분해 단계 또는 발효 단계 중 어느 하나(그러나 단계 둘 모두는 아님)는 베타-글루코시다제의 참여를 수반한다. 예를 들어, 본 발명의 Mg3A 폴리펩티드 또는 상기 Mg3A 폴리펩티드를 포함하는 조성물은 당화 단계에서 사용될 수 있다. 또 다른 예에서, 리그노셀룰로오스계 바이오매스 기질을 가수분해하는데 사용되는 효소 조성물은 베타-글루코시다제 활성을 포함하지 않으며, 한편 에타놀로젠 미생물은 베타-글루코시다제 폴리펩티드, 예를 들어, 본 발명의 Mg3A 폴리펩티드를 발현시킨다.
- [0128] 본원에 사용된 바와 같이, "신호 서열"은 세포 외부로 폴리펩티드의 성숙 형태의 분비를 촉진하는 폴리펩티드의 N-말단 부분에 결합된 아미노산의 서열을 의미한다. 신호 서열의 이러한 정의는 기능적 정의이다. 세포의 폴리펩티드의 성숙 형태는 분비 공정 동안 절단되는 신호 서열이 결합되어 있다. Mg3A의 천연 신호 서열이 본 조성물 및 방법의 측면에서 사용될 수 있지만, 다른 비천연 신호 서열이 사용될 수도 있다 (예를 들어, 서열 번호 13). 용어 "성숙"은, 본원에서의 폴리펩티드를 지칭하는 경우, 번역 및 임의의 번역 후 변형에 이은 그의 최종 형태(들)의 폴리펩티드를 의미한다. 예를 들어, 본 발명의 Mg3A 폴리펩티드는 하나 이상의 성숙 형태를 갖고, 그 중 적어도 하나는 서열 번호 3의 아미노산을 서열을 갖는다.
- [0129] 본 발명의 베타-글루코시다제 폴리펩티드는 "전구체," "미성숙," 또는 "전장"으로서 지칭될 수 있거나 (이 경우에 이들은 신호 서열을 포함함), "성숙"으로서 지칭될 수 있다 (이 경우에 이들은 신호 서열이 결합되어 있다). 폴리펩티드의 성숙 형태가 일반적으로 가장 유용하다. 달리 언급되지 않는 한, 본원에서 사용된 아미노산 잔기의 번호 매김은 각각의 아미노산 폴리펩티드의 성숙 형태를 지칭한다. 본 발명의 베타-글루코시다제 폴리펩티드는 또한, 생성 폴리펩티드가 베타-글루코시다제 활성을 보유하는 한, 절단되어 N 또는 C-말단이 제거될 수 있다.
- [0130] 본 발명의 베타-글루코시다제 폴리펩티드는 또한, 이것이 제1 베타-글루코시다제 폴리펩티드의 적어도 일부, 및 제2 베타-글루코시다제 폴리펩티드의 적어도 일부를 포함한다는 점에서, "키메라" 또는 "하이브리드" 폴리펩티드일 수 있다 (이러한 키메라 베타-글루코시다제 폴리펩티드는, 예를 들어, 베타-글루코시다제 각각에 대한 도메인의 스와핑(swapping)을 수반하는 공지된 기술을 사용하여, 제1 및 제2 베타-글루코시다제로부터 유래될 수 있다). 본 베타-글루코시다제 폴리펩티드는 이중 신호 서열, 추적 또는 정제를 가능하게 하는 에피토프 등을 추가로 포함할 수 있다. 용어 "이중"을 사용하여, 관심 폴리펩티드를 발현시키기 위해 사용되는 신호 서열을 지칭하는 경우, 이는 신호 서열이, 예를 들어, 관심 폴리펩티드로서 상이한 미생물로부터 유래됨을 의미하는 것이다. 본원에서의 Mg3A 폴리펩티드를 발현시키기 위한 적합한 이중 신호 서열의 예는, 예를 들어, *트리코테르마 레에세이*로부터의 것일 수 있다.
- [0131] 본원에 사용된 바와 같이, "기능적으로 부착된" 또는 "작동가능하게 연결된"은 알려진 또는 원하는 활성을 갖는 조절 영역 또는 기능 도메인, 예컨대 프로모터, 터미네이터, 신호 서열 또는 인핸서 영역이 표적 (예를 들어, 유전자 또는 폴리펩티드)에, 조절 영역 또는 기능 도메인이 그의 알려진 또는 원하는 활성에 따라 그 표적의 발

현, 분비 또는 기능을 제어할 수 있도록 하는 방식으로 부착되거나 연결됨을 의미한다.

[0132] 본원에 사용된 바와 같이, 용어 "폴리펩티드" 및 "효소"는 펩티드 결합에 의해 연결된 아미노산 잔기를 포함하는 임의의 길이의 중합체를 지칭하는데 교체 사용이 가능하다. 아미노산 잔기를 위한 통상적인 1문자 또는 3문자 코드가 본원에서 사용된다. 중합체는 선형 또는 분지형일 수 있고, 이는 변형된 아미노산을 포함할 수 있고, 비-아미노산이 개재될 수 있다. 상기 용어는 또한 자연적으로 또는 개입에 의해; 예를 들어, 디설피드 결합 형성, 글리코실화, 지질화, 아세틸화, 인산화, 또는 임의의 다른 조작 또는 변형, 예컨대 표지 성분과의 컨쥬게이션에 의해 변형된 아미노산 중합체를 포함한다. 또한, 정의 내에, 예를 들어, 아미노산의 하나 이상의 유사체 (예를 들어, 비천연 아미노산 등 포함)를 함유하는 폴리펩티드뿐만 아니라, 관련 기술분야에 공지된 다른 변형이 포함된다.

[0133] 본원에 사용된 바와 같이, "야생형" 및 "천연" 유전자, 효소, 또는 균주는 자연에서 발견되는 것들이다.

[0134] 용어 "야생형", "모" 또는 "기준" 은, 폴리펩티드와 관련하여, 하나 이상의 아미노산 위치에서 인공 치환, 삽입, 또는 결실을 포함하지 않는 자연 발생 폴리펩티드를 지칭한다. 유사하게, 용어 "야생형", "모" 또는 "기준"은, 폴리뉴클레오타이드와 관련하여, 인공 뉴클레오시드(nucleoside) 변화를 포함하지 않는 자연 발생 폴리뉴클레오타이드를 지칭한다. 그러나, 야생형, 모, 또는 기준 폴리펩티드를 암호화하는 폴리뉴클레오타이드는 자연 발생 폴리뉴클레오타이드에 제한되는 것이 아니라, 야생형, 모, 또는 기준 폴리펩티드를 암호화하는 임의의 폴리뉴클레오타이드를 포함한다.

[0135] 본원에 사용된 바와 같이, "변이체 폴리펩티드"는, 전형적으로 재조합 DNA 기법에 의해, 하나 이상의 아미노산의 치환, 부가, 또는 결실에 의해 모 (또는 기준) 폴리펩티드로부터 유래되는 폴리펩티드를 지칭한다. 변이체 폴리펩티드는 작은 수의 아미노산 잔기에 의해 모 폴리펩티드와 상이할 수 있다. 이들은 모 폴리펩티드와의 1차 아미노산 서열 상동성/동일성의 그의 수준에 의해 정의될 수 있다. 적합하게는, 변이체 폴리펩티드는 모 폴리펩티드에 대하여 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 심지어는 적어도 99% 아미노산 서열 동일성을 갖는다.

[0136] 본원에 사용된 바와 같이, "변이체 폴리뉴클레오타이드"는 변이체 폴리펩티드를 암호화하거나, 모 폴리뉴클레오타이드와 특정된 정도의 상동성/동일성을 갖거나, 엄격한 조건하에 모 폴리뉴클레오타이드 또는 그의 상보체와 혼성화한다. 적합하게는, 변이체 폴리뉴클레오타이드는 모 폴리뉴클레오타이드 또는 모 폴리뉴클레오타이드의 상보체에 대하여 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 심지어는 적어도 99% 뉴클레오타이드 서열 동일성을 갖는다. 퍼센트 동일성을 결정하는 방법은 관련 기술분야에 공지되어 있고 상기 기재되어 있다.

[0137] 용어 "로부터 유래된"은 용어 "로부터 기원된", "로부터 수득된", "로부터 수득가능한", "로부터 단리된", 및 "로부터 창출된"을 포함하며, 일반적으로 한 특정된 물질이 또 다른 특정된 물질에서 그의 기원을 찾거나 또 다른 특정된 물질과 관련하여 기재될 수 있는 특질을 가짐을 나타낸다.

[0138] 본원에 사용된 바와 같이, 용어 "혼성화 조건"은 혼성화 반응이 수행되는 조건을 지칭한다. 이들 조건은 전형적으로, 혼성화가 측정되는 조건의 "엄격성"의 정도에 의해 분류된다. 엄격성의 정도는, 예를 들어, 핵산 결합 복합체 또는 프로브의 용해 온도 (T_m)를 기반으로 할 수 있다. 예를 들어, "최대 엄격성"은 전형적으로 약 $T_m - 5^\circ\text{C}$ (프로브의 T_m 보다 5°C 미만)에서 일어나고; "높은 엄격성"은 T_m 보다 약 $5-10^\circ\text{C}$ 미만에서 일어나고; "중간 엄격성"은 프로브의 T_m 보다 약 $10-20^\circ\text{C}$ 미만에서 일어나고; "낮은 엄격성"은 T_m 보다 약 $20-25^\circ\text{C}$ 미만에서 일어난다. 대안으로, 또는 게다가, 혼성화 조건은 혼성화의 염 또는 이온 강도 조건, 및/또는 하나 이상의 엄격성 세척, 예를 들어: 6X SSC = 매우 낮은 엄격성; 3X SSC = 낮은 내지 중간 엄격성; 1X SSC = 중간 엄격성; 및 0.5X SSC = 높은 엄격성을 기반으로 할 수 있다. 기능적으로, 최대 엄격성 조건은 혼성화 프로브와 엄밀 동일성 또는 근-엄밀 동일성을 갖는 핵산 서열을 동정하는데 사용될 수 있으며; 한편 높은 엄격성 조건은 프로브와 약 80% 이상의 서열 동일성을 갖는 핵산 서열을 동정하는데 사용된다. 높은 선택성을 필요로 하는 적용을 위해, 전형적으로, 비교적 엄격한 조건을 사용하여 하이브리드를 형성하는 것이 바람직하다 (예를 들어 비교적 낮은 염 및/또는 고온 조건이 사용된다).

[0139] 본원에 사용된 바와 같이, 용어 "혼성화"는 관련 기술분야에 공지된 바와 같이, 핵산의 한 가닥이 염기 쌍 형성을 통해 상보적 가닥과 결합하는 공정을 지칭한다. 더 구체적으로, "혼성화"는 블롯 혼성화 기법 및 PCR 기법 동안 일어나는 바와 같이, 핵산의 외가닥이, 상보적 가닥, 즉, 염기 쌍으로 이중 가닥을 형성하는 과정을 지칭

한다. 핵산 서열은 두 서열이 중정도 내지 높은 엄격성 혼성화 및 세척 조건하에 서로 특이적으로 혼성화하는 경우 참조 핵산 서열에 "선택적으로 혼성화가능한" 것으로 여겨진다. 혼성화 조건은 핵산 결합 복합체 또는 프로브의 용해 온도 (T_m)를 기반으로 한다. 예를 들어, "최대 엄격성"은 전형적으로 약 $T_m - 5^\circ\text{C}$ (프로브의 T_m 보다 5°C 미만)에서 일어나고; "높은 엄격성"은 T_m 보다 약 $5-10^\circ\text{C}$ 미만에서 일어나고; "중간 엄격성"은 프로브의 T_m 보다 약 $10-20^\circ\text{C}$ 미만에서 일어나고; "낮은 엄격성"은 T_m 보다 약 $20-25^\circ\text{C}$ 미만에서 일어난다. 기능적으로, 최대 엄격성 조건은 혼성화 프로브와 엄밀 동일성 또는 근-엄밀 동일성을 갖는 서열을 동정하는데 사용될 수 있으며; 한편 중간 또는 낮은 엄격성 혼성화는 폴리뉴클레오티드 서열 상동체를 동정 또는 검출하는데 사용될 수 있다.

[0140] 중간 및 높은 엄격성 혼성화 조건은 관련 기술분야에 주지되어 있다. 예를 들어, 중간 엄격성 혼성화는 20% 포름아미드, 5 x SSC (150 mM NaCl, 15 mM 시트르산삼나트륨), 50 mM 인산나트륨 (pH 7.6), 5 x 덴하르트 용액 (Denhardt's solution), 10% 덱스트란 술페이트 및 20 mg/ml 변성 전단된 연어 정자(denatured sheared salmon sperm) DNA를 포함하는 용액에서 37°C 에서 밤새 인큐베이션으로 수행된 후, 필터를 약 $37 - 50^\circ\text{C}$ 에서 1x SSC 중에서 세척할 수 있다. 높은 엄격성 혼성화 조건은 65°C 및 0.1X SSC (여기서 1X SSC = 0.15 M NaCl, 0.015 M Na_3 시트레이트, pH 7.0)에서의 혼성화일 수 있다. 대안으로, 높은 엄격성 혼성화 조건은 50% 포름아미드, 5X SSC, 5X 덴하르트 용액, 0.5% SDS 및 100 $\mu\text{g/ml}$ 변성 담체 DNA에서 약 42°C 에서 수행된 후, 실온에서 2X SSC 및 0.5% SDS에서 2회 세척 및 42°C 에서 0.1X SSC 및 0.5% SDS에서 추가로 2회 세척할 수 있다. 그리고 매우 높은 엄격성 혼성화 조건은 68°C 및 0.1X SSC에서의 혼성화일 수 있다 통상의 기술자는 프로브 길이 등과 같은 인자를 수용하기 위해 필요한 경우 온도, 이온 강도 등을 조정하는 법을 알고 있다.

[0141] 변이체 베타-글루코시다제를 암호화하는 핵산은 서열 번호 1의 뉴클레오티드 및 그의 동일한 상보체 사이에 형성된 이중가닥과 비교하여 $1^\circ\text{C} - 3^\circ\text{C}$ 이상 감소된 T_m 을 가질 수 있다.

[0142] 적어도 2개의 핵산 또는 폴리펩티드의 맥락에서, 문구 "실질적으로 유사한" 또는 "실질적으로 동일한"은, 폴리뉴클레오티드 또는 폴리펩티드가 모 또는 참조 서열과 적어도 약 90%, 적어도 약 91%, 적어도 약 92%, 적어도 약 93%, 적어도 약 94%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 또는 심지어는 적어도 약 99% 동일한 서열을 포함하거나, 관능기를 첨가함이 없이 단지 본 기재를 회피하기 위해 이루어진 아미노산 치환, 삽입, 결실, 또는 변형은 포함하지 않음을 의미한다.

[0143] 본원에 사용된 바와 같이, "발현 벡터"는 적합한 숙주에서 폴리펩티드의 발현을 수행할 수 있는 적합한 제어 서열에 작동가능하게 연결되고 특정된 폴리펩티드를 암호화하는 DNA 서열을 함유하는 DNA 구축물을 지칭한다. 이러한 제어 서열은 전사를 수행하는 프로모터, 이러한 전사를 제어하는 임의적 오퍼레이터 서열, 적합한 mRNA 리보솜 결합 부위를 암호화하는 서열 및/또는 전사 및 번역의 종결을 제어하는 서열을 포함할 수 있다. 벡터는 플라스미드, 파지 입자, 또는 잠재적 게놈 삽입물일 수 있다. 일단 적합한 숙주로 형질전환되면, 벡터는 숙주 게놈과 무관하게 기능하고 복제할 수 있거나, 일부 경우에, 게놈 자체에 통합될 수 있다.

[0144] 용어 "재조합"은, 예컨대 암호화 서열을 돌연변이시켜 변경된 폴리펩티드를 생성시키고, 암호화 서열을 또 다른 유전자의 것에 융합시키고, 상이한 프로모터의 제어하에 유전자를 위치시키고, 이중 유기체에서 유전자를 발현시키고, 감소된 또는 상승된 수준에서 유전자를 발현시키고, 그의 천연 발현 프로파일과 상이한 방식으로 조건적으로 또는 구성적으로 유전자를 발현시키는 것 등에 의해, 그의 서열 또는 발현 특성을 변경시키기 위해 변경된 유전 물질 (즉, 핵산, 이들이 암호화하는 폴리펩티드, 및 이러한 폴리뉴클레오티드를 포함하는 벡터 및 세포)을 지칭한다. 일반적으로 재조합 핵산, 폴리펩티드, 및 이를 기반으로 한 세포는, 이들이 관련 핵산, 폴리펩티드, 및 자연에서 발견되는 세포와 동일하지 않도록 인공적으로 조작되어 왔다.

[0145] "신호 서열"은 폴리펩티드의 N-말단 부분에 결합된 아미노산의 서열을 지칭하며, 이는 세포로부터 폴리펩티드의 성숙 형태의 분비를 촉진한다. 세포의 폴리펩티드의 성숙 형태는 분비 공정 동안 절단되는 신호 서열이 결여되어 있다.

[0146] 용어 "선택적 마커" 또는 "선택가능한 마커"는 도입된 핵산 또는 벡터를 함유하는 그러한 숙주를 용이하게 선택할 수 있게 하는 숙주 세포에서 발현할 수 있는 유전자를 지칭한다. 선택가능한 마커의 예는 항균 물질 (예를 들어, 히그로마이신, 블레오마이신, 또는 클로람페니콜) 및/또는 숙주 세포 상에, 영양적 이점과 같은 대사적 이점을 부여하는 유전자를 포함하나 그에 제한되지는 않는다.

[0147] 용어 "조절 요소"는 핵산 서열의 발현의 일부 측면을 제어하는 유전 요소를 지칭한다. 예를 들어, 프로모터는 작동가능하게 연결된 암호화 영역의 전사의 개시를 촉진하는 조절 요소이다. 추가의 조절 요소는 스플라이싱

(splicing) 신호, 폴리아데닐화 신호 및 종결 신호를 포함한다.

- [0148] 본원에 사용된 바와 같이, "숙주 세포"는 일반적으로, 관련 기술분야에 공지된 재조합 DNA 기법을 사용하여 구축된 벡터로 형질전환되거나 형질감염되는 원핵생물 또는 진핵생물 숙주의 세포이다. 형질전환된 숙주 세포는 폴리펩티드 변이체를 암호화하는 벡터를 복제할 수 있거나 원하는 폴리펩티드 변이체를 발현시킬 수 있다. 폴리펩티드 변이체의 예비(pre)- 또는 사전(pro)-형태를 암호화하는 벡터의 경우에, 이러한 변이체는, 발현시, 전형적으로 숙주세포로부터 숙주 세포 배지로 분리된다.
- [0149] 용어 "도입된"은, 핵산 서열을 세포에 삽입하는 맥락에서, 형질전환, 형질도입, 또는 형질감염을 의미한다. 형질전환의 수단은 관련 기술분야에 공지된 바와 같은 원형질체 형질전환, 염화칼슘 침전, 전기천공법, 네이키드(naked) DNA 등을 포함한다. (문헌 [Chang and Cohen *Mol. Gen. Genet.* 168:111-115, 1979]; [Smith *et al.* (1986) *Appl. Env. Microbiol.* 51:634]; 및 검토 논문 [Ferrari *et al.*, in Harwood, *Bacillus*, Plenum Publishing Corporation, pp. 57-72, 1989] 참조).
- [0150] "융합" 폴리펩티드 서열은 2개의 대상 폴리펩티드 서열 사이의 펩티드 결합을 통해, 연결된다, 즉, 작동가능하게 연결된다.
- [0151] 용어 "사상균"은 아문(subdivision) 유마이코티나(Eumycotina), 특히 페치조마이코티나(Pezizomycotina) 종의 모든 사상 형태를 지칭한다.
- [0152] "에타놀로젠성 미생물"은 당 또는 올리고당을 에탄올로 전환시키는 능력을 갖는 미생물을 지칭한다.
- [0153] 다른 기술적 및 과학적 용어는 본 개시내용이 속하는 관련 기술분야의 통상의 기술자에 의해 통상 이해되는 바와 같은 의미를 갖는다 (예를 들어, 문헌 [Singleton and Sainsbury, *Dictionary of Microbiology and Molecular Biology*, 2d Ed., John Wiley and Sons, NY 1994]; 및 [Hale and Marham, *The Harper Collins Dictionary of Biology*, Harper Perennial, NY 1991] 참조).
- [0154] **III. 베타-글루코시다제 폴리펩티드, 폴리뉴클레오티드, 벡터, 및 숙주 세포**
- [0155] **A. Mg3A 폴리펩티드**
- [0156] 한 측면에서, 본 조성물 및 방법은 베타-글루코시다제 활성을 갖는 재조합 Mg3A 베타-글루코시다제 폴리펩티드, 그의 단편, 또는 그의 변이체를 제공한다. 재조합 베타-글루코시다제 폴리펩티드의 예는 *마그나포르테 그리세아*로부터 단리되었다. 성숙 Mg3A 폴리펩티드는 서열 번호 3으로 제시된 아미노산 서열을 갖는다. 유사한, 실질적으로 유사한 Mg3A 폴리펩티드는, 예를 들어, *마그나포르테 그리세아*의 기타 균주 또는 단리물에서, 자연 발생할 수 있다. 이들 및 다른 재조합 Mg3A 폴리펩티드는 본 조성물 및 방법에 의해 포함된다.
- [0157] 일부 실시양태에서, 재조합 Mg3A 폴리펩티드는 예시된 Mg3A 폴리펩티드에 대하여 특정된 정도의 아미노산 서열 동일성, 예를 들어, 서열 번호 2의 아미노산 서열 또는 성숙 서열 서열 번호 3에 대하여 적어도 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 심지어는 적어도 99% 서열 동일성을 갖는 변이체 Mg3A 폴리펩티드이다. 서열 동일성은, 본원에 기재된 바와 같이, 예를 들어, BLAST, ALIGN, 또는 CLUSTAL과 같은 프로그램 사용하여, 아미노산 서열 정렬에 의해 결정될 수 있다.
- [0158] 특정 실시양태에서, 재조합 Mg3A 폴리펩티드는 미생물 중에서, 예를 들어, 세균 또는 진균 숙주 생물에서, 재조합적으로 생성되며, 한편 다른 것에서는 Mg3A 폴리펩티드는 합성적으로 생성되거나, 천연 공급원 (예를 들어, *마그나포르테 그리세아*)으로부터 정제된다.
- [0159] 특정 실시양태에서, 재조합 Mg3A 폴리펩티드는 폴리펩티드의 구조 및/또는 기능에 실질적으로 영향을 미치지 않는 치환을 포함한다. 이들 치환의 예는 표 1에 요약된 바와 같이 보존적 돌연변이이다.

[0160]

[표 1]

아미노산 치환

원래 잔기	코드	허용되는 치환
알라닌	A	D-Ala, Gly, 베타-Ala, L-Cys, D-Cys
아르기닌	R	D-Arg, Lys, D-Lys, 호모-Arg, D-호모-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
아스파라긴	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
아스파르트산	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
시스테인	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
글루타민	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
글루탐산	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
글리신	G	Ala, D-Ala, Pro, D-Pro, beta-Ala, Acp
이소류신	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
류신	L	D-Leu, Val, D-Val, Leu, D-Leu, Met, D-Met
리신	K	D-Lys, Arg, D-Arg, 호모-Arg, D-호모-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
메티오닌	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
페닐알라닌	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, 트랜스-3,4, 또는 5-페닐프롤린, 시스-3,4, 또는 5-페닐프롤린
프롤린	P	D-Pro, L-1-티아졸리딘-4-카복실산, D- 또는 L-1-옥사졸리딘-4-카복실산
세린	S	D-Ser, Thr, D-Thr, 알로-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
트레오닌	T	D-Thr, Ser, D-Ser, 알로-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
티로신	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
발린	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

[0161]

[0162]

자연 발생 아미노산을 수반하는 치환은 일반적으로 재조합 Mg3A 폴리펩티드를 암호화하는 핵산을 돌연변이시킨 다음, 변이체 폴리펩티드를 유기체에서 발현시킴으로써 이루어진다. 아미노산에 대해 화학적 변형 또는 비-자연 발생 아미노산을 수반하는 치환은 일반적으로 Mg3A 폴리펩티드가 유기체에 의해 합성된 후 이를 화학적으로 변형시킴으로써 이루어진다.

[0163]

일부 실시양태에서, 변이체 재조합 Mg3A 폴리펩티드는 서열 번호 2 또는 서열 번호 3과 실질적으로 동일하며, 이는 이들이 폴리펩티드의 구조, 기능, 또는 발현에 그다지 영향을 미치지 않는 아미노산 치환, 삽입, 또는 결실을 포함하지 않음을 의미한다. 이러한 변이체 재조합 Mg3A 폴리펩티드는 본 기재에 회피하도록 설계된 것들을 포함할 것이다. 일부 실시양태에서, 변이체 재조합 Mg3A 폴리펩티드, 이들 변이체를 포함하는 조성물 및 방법은 서열 번호 2 또는 서열 번호 3과 실질적으로 동일한 것이 아니라, 특정 상황에서, 실질적으로, 본원에서의 폴리펩티드의 구조, 기능, 또는 발현에 영향을 미치는 아미노산 치환, 삽입, 또는 결실을 포함하여, 결과적으로 서열 번호 2 또는 서열 번호 3의 폴리펩티드의 것과 비교하여, 예를 들어, 리그노셀룰로오스계 기질을 가수분해하는 개선된 비활성도, 바람직한 숙주 생물에서의 개선된 발현, 개선된 열안정성, pH 안정성 등을 포함하는 개선된 특성이 달성될 수 있게 된다.

[0164]

일부 실시양태에서, 재조합 Mg3A 폴리펩티드 (그의 변이체 포함)는 베타-글루코시다제 활성을 갖는다. 베타-글루코시다제 활성은 본원에 기재된 검정, 예를 들어, 실시예 2에 기재된 것을 사용하여, 또는 관련 기술분야에 공지된 다른 검정에 의해 결정되고 측정될 수 있다.

[0165]

재조합 Mg3A 폴리펩티드는 베타-글루코시다제 활성을 보유하는 "전장" Mg3A 폴리펩티드의 단편을 포함한다. 바람직하게는 그러한 기능적 단편 (즉, 베타-글루코시다제 활성을 보유하는 단편)은 길이가 적어도 100 아미노산 잔기 (예를 들어, 길이가 적어도 100 아미노산 잔기, 적어도 120 아미노산 잔기, 적어도 140 아미노산 잔기, 적어도 160 아미노산 잔기, 적어도 180 아미노산 잔기, 적어도 200 아미노산 잔기, 적어도 220 아미노산 잔기, 적어도 240 아미노산 잔기, 적어도 260 아미노산 잔기, 적어도 280 아미노산 잔기, 적어도 300 아미노산 잔기, 적어도 320 아미노산 잔기, 또는 적어도 350 아미노산 잔기 또는 그 초과)이다. 이러한 단편은 적합하게는 전장 전구체 폴리펩티드 또는 전장 성숙 폴리펩티드의 활성 부위를 보유하나 중요성이 낮은(non-critical) 아미노산 잔기의 결실을 가질 수 있다. 단편의 활성은 본원에 기재된 검정, 예를 들어 실시예 2에 기재된 것을

사용하여, 또는 관련 기술분야에 공지된 다른 검정에 의해 용이하게 결정될 수 있다.

[0166] 일부 실시양태에서, Mg3A 아미노산 서열 및 유도체는 N- 및/또는 C-말단 융합 단백질로서 생성되어, 예를 들어, 추출, 검출 및/또는 정제에 도움을 주고/거나 Mg3A 폴리펩티드에 기능적 특성을 부가한다. 융합 단백질 파트너의 예는 글루타티온-S-트랜스퍼라제 (GST), 6XHis, GAL4 (DNA 결합 및/또는 전사 활성 도메인), FLAG-, MYC-태그 또는 통상의 기술자에게 공지된 다른 태그를 포함하나 그에 제한되지는 않는다. 일부 실시양태에서, 관심 폴리펩티드 서열과 융합 단백질 파트너 사이에 단백질분해성 절단 부위가 제공되어 융합 서열을 제거할 수 있게 한다. 적합하게는, 융합 단백질은 재조합 Mg3A 폴리펩티드의 활성을 저해하지 않는다. 일부 실시양태에서, 재조합 Mg3A 폴리펩티드는 리더 펩티드(leader peptide), 프로펩티드(propeptide), 결합 도메인 및/또는 촉매적 도메인을 포함한 기능적 도메인에 융합된다. 융합 단백질은 임의로, 두 성분의 특성에 그다지 영향을 미치지 않으면서 Mg3A 폴리펩티드와 융합 도메인을 결합하는 링커 서열을 통해 재조합 Mg3A 폴리펩티드에 연결된다. 링커는 임의로, 의도된 적용에 기능적으로 기여한다.

[0167] 본 개시내용은 조작되어 본 개시내용의 하나 이상의 Mg3A 폴리펩티드를 발현시키는 숙주 세포를 제공한다. 적합한 숙주 세포는 임의의 미생물의 세포 (예를 들어, 세균, 원생생물, 조류, 진균 (예를 들어, 효모 또는 사상균), 또는 다른 미생물의 세포)를 포함하고, 바람직하게는 세균, 효모, 또는 사상균의 세포이다.

[0168] 세균 속의 적합한 숙주 세포는 *에스케리키아*, *바실루스*, *락토바실루스*, *슈도모나스(Pseudomonas)*, 및 *스트렙토마이세스*의 세포를 포함하나 그에 제한되지는 않는다. 세균 종의 적합한 세포는 *에스케리키아 콜라이*, *바실루스 서브틸리스*, *바실루스 리케니포르미스(Bacillus licheniformis)*, *락토바실루스 브레비스(Lactobacillus brevis)*, *슈도모나스 아에루기노사(Pseudomonas aeruginosa)*, 및 *스트렙토마이세스 리비단스*의 세포를 포함하나 그에 제한되지는 않는다.

[0169] 효모의 속의 적합한 숙주 세포는 *사카로마이세스(Saccharomyces)*, *스키조사카로마이세스(Schizosaccharomyces)*, *칸디다(Candida)*, *한센울라(Hansenula)*, *피치아(Pichia)*, *클루이베로마이세스(Kluyveromyces)*, 및 *파피아(Phaffia)*의 세포를 포함하나 그에 제한되지는 않는다. 효모 종의 적합한 세포는 *사카로마이세스 세레비지에*, *스키조사카로마이세스 폼베(Schizosaccharomyces pombe)*, *칸디다 알비칸스(Candida albicans)*, *한센울라 폴리모르파(Hansenula polymorpha)*, *피치아 파스토리스(Pichia pastoris)*, *피. 캐나다렌시스(P. canadensis)*, *클루이베로마이세스 마르시아누스(Kluyveromyces marxianus)*, 및 *파피아 로도지마(Phaffia rhodozyma)*의 세포를 포함하나 그에 제한되지는 않는다.

[0170] 사상균의 적합한 숙주 세포는 아문 유마이크로티나의 모든 사상 형태를 포함한다. 사상균 속의 적합한 세포는 *아크레모니움(Acremonium)*, *아스페르길루스*, *아우레오바시디움(Aureobasidium)*, *비에르칸테라(Bjerkandera)*, *세리포리옵시스(Ceriporiopsis)*, *크리소스포리움(Chrysoporium)*, *코프리누스(Coprinus)*, *코리올루스(Coriolus)*, *코리나스쿠스(Corynascus)*, *차에르토미움(Chaetomium)*, *크립토크커스(Cryptococcus)*, *필로바시디움(Filobasidium)*, *푸사리움*, *지베렐라(Gibberella)*, *후미콜라*, *마그나포르테*, *무코르(Mucor)*, *마이셀리오포토라(Myceliophthora)*, *무코르*, *네오칼리마스틱스(Neocallimastix)*, *뉴로스포라(Neurospora)*, *파에실로마이세스(Paecilomyces)*, *페니실리움(Penicillium)*, *파네로차에테(Phanerochaete)*, *플레비아(Phlebia)*, *피로마이세스(Piromyces)*, *플레우로투스(Pleurotus)*, *사이탈리움(Scytalidium)*, *스키조필룸(Schizophyllum)*, *스포로트리쿰(Sporotrichum)*, *탈라로마이세스(Talaromyces)*, *써모아스쿠스(Thermoascus)*, *티엘라비아(Thielavia)*, *톨리포클라디움(Tolypocladium)*, *트라메테스(Trametes)*, 및 *트리코테르마*의 세포를 포함하나 그에 제한되지는 않는다.

[0171] 사상균 종의 적합한 세포는 *아스페르길루스 아와모리(Aspergillus awamori)*, *아스페르길루스 푸미가투스(Aspergillus fumigatus)*, *아스페르길루스 포에티두스(Aspergillus foetidus)*, *아스페르길루스 야포니쿠스(Aspergillus japonicus)*, *아스페르길루스 니들란스*, *아스페르길루스 니게르*, *아스페르길루스 오리자에*, *크리소스포리움 룩크노웬스(Chrysosporium lucknowense)*, *푸사리움 박트리디오이데스(Fusarium bac-tridioides)*, *푸사리움 세레알리스(Fusarium cerealis)*, *푸사리움 크루크웰렌스(Fusarium crookwellense)*, *푸사리움 쿨모룸(Fusarium culmorum)*, *푸사리움 그라미네아룸(Fusarium graminearum)*, *푸사리움 그라미눔(Fusarium graminum)*, *푸사리움 헤테로스포룸(Fusarium heterosporum)*, *푸사리움 네군디(Fusarium negundi)*, *푸사리움 옥시스포룸(Fusarium oxysporum)*, *푸사리움 레티쿨라툼(Fusarium reticulatum)*, *푸사리움 로세움(Fusarium roseum)*, *푸사리움 삼부시눔(Fusarium sambucinum)*, *푸사리움 사르코크로움(Fusarium sarcocroum)*, *푸사리움 스포로트리키오이데스(Fusarium sporotrichioides)*, *푸사리움 숄푸레움(Fusarium sulphureum)*, *푸사리움 토룰로숨(Fusarium torulosum)*, *푸사리움 트리코테시오이데스(Fusarium trichothecioides)*, *푸사리움 베네나툼(Fusarium venenatum)*, *비에르칸테라 아두스타(Bjerkandera adusta)*, *세리포리옵시스 아네이리나(Ceriporiopsis*

aneirina), 세리포리옵시스 아네이리나, 세리포리옵시스 카레지애아(*Ceriporiopsis caregiea*), 세리포리옵시스 질베슨스(*Ceriporiopsis gilvescens*), 세리포리옵시스 파노신타(*Ceriporiopsis pannocinta*), 세리포리옵시스 리볼로사(*Ceriporiopsis rivulosa*), 세리포리옵시스 수브루파(*Ceriporiopsis subrufa*), 세리포리옵시스 수브베르미스포라(*Ceriporiopsis subvermisporea*), 코프린누스 시네레우스(*Coprinus cinereus*), 코리올루스 히르수투스(*Coriolus hirsutus*), 후미콜라 인솔렌스(*Humicola insolens*), 후미콜라 라누지노사(*Humicola lanuginosa*), 무코르 미에헤이(*Mucor miehei*), 마이셀리오프토라 써모필라(*Myceliophthora thermophila*), 뉴로스포라 크라사(*Neurospora crassa*), 뉴로스포라 인터메디아(*Neurospora intermedia*), 페니실리움 푸르푸로제눔(*Penicillium purpurogenum*), 페니실리움 카네스스(*Penicillium canescens*), 페니실리움 솔리툼(*Penicillium solitum*), 페니실리움 푸니쿨로숨, 파네로차에테 크리스소스포리움(*Phanerochaete chrysosporium*), 플레비아 라디아테(*Phlebia radiata*), 플레우로투스 에린지이(*Pleurotus eryngii*), 탈라로마이세스 플라부스(*Talaromyces flavus*), 티엘라비아 테레스트리스(*Thielavia terrestris*), 트라메테스 빌로사(*Trametes villosa*), 트라메테스 베르시콜로르(*Trametes versicolor*), 트리코데르마 하리아눔(*Trichoderma harzianum*), 트리코데르마 코닝지이(*Trichoderma koningii*), 트리코데르마 롱지브라키아툼(*Trichoderma longibrachiatum*), 트리코데르마 레에세이, 및 트리코데르마 비리데(*Trichoderma viride*)의 세포를 포함하나 그에 제한되지는 않는다.

[0172] 핵산을 이들 유기체로 형질전환하는 방법은 관련 기술분야에 공지되어 있다. 예를 들어, *아스페르길루스* 숙주 세포를 형질전환하기에 적합한 절차는 EP 238 023에 기재되어 있다.

[0173] 일부 실시양태에서, 재조합 Mg3A 폴리펩티드는 신호 펩티드에 융합되어, 예를 들어, 재조합 Mg3A 폴리펩티드의 세포의 분비를 촉진한다. 예를 들어, 특정 실시양태에서, 신호 펩티드는 서열 번호 13-42로부터 선택된 서열에 의해 암호화된다. 특정의 실시양태에서, 재조합 Mg3A 폴리펩티드는 분비된 폴리펩티드로서 이종 유기체에서 발현된다. 따라서 본원에서의 조성물 및 방법은 이종 유기체에서 분비된 폴리펩티드로서 Mg3A 폴리펩티드를 발현시키는 방법을 포함한다. 일부 실시양태에서 재조합 Mg3A 폴리펩티드는, 예를 들어, 이종 유기체가 에타놀로겐 미생물, 예컨대 *사카로마이세스 세레비지에* 또는 *지모모나스 모빌리스*인 경우, 이종 유기체에서 세포내에서 발현된다. 그러한 경우에, 셀로비오스 수송체 유전자는 유전 공학 툴을 사용하여 유기체에 도입되어, Mg3A 폴리펩티드가 유기체 내부에서 셀로비오스 기질에 대해 작용하여 셀로비오스를 D-글루코오스로 전환시키도록 한 다음, 이를 유기체에 의해 에탄올로 전환되거나 대사되도록 할 수 있다.

[0174] 본 개시내용은 또한 상기-기재된 핵산을 포함하는 발현 카세트 및/또는 벡터를 제공한다. 적합하게는, 본 개시내용의 Mg3A 폴리펩티드를 암호화하는 핵산은 프로모터에 작동가능하게 연결된다. 프로모터는 관련 기술분야에 주지되어 있다. 숙주 세포에서 기능하는 임의의 프로모터가 베타-글루코시다제 및/또는 본 개시내용의 다른 핵산 중 임의의 것을 발현시키는데 사용될 수 있다. 다양한 숙주 세포에서의 베타-글루코시다제 핵산 및/또는 본 개시내용의 다른 핵산 중 임의의 것의 발현을 구동하기 위해 유용한, 개시 제어 영역 또는 프로모터는 수많으며 통상의 기술자에게 익숙하다 (예를 들어, WO 2004/033646 및 거기에 인용된 참조문헌 참조). 사실상 이들 핵산을 구동할 수 있는 임의의 프로모터가 사용될 수 있다.

[0175] 구체적으로, 사상균 숙주에서의 재조합 발현을 원하는 경우, 프로모터는 사상균 프로모터일 수 있다. 핵산은, 예를 들어, 이종 프로모터의 제어하에 있을 수 있다. 핵산은 또한 구성적 또는 유도성 프로모터의 제어 하에 발현될 수 있다. 사용될 수 있는 프로모터의 예는 셀룰라제 프로모터, 자일라나제 프로모터, 1818 프로모터 (이전에, *트리코데르마*의 EST 맵핑(mapping)에 의해 고도로 발현된 단백질을 동정됨)를 포함하나 그에 제한되지는 않는다. 예를 들어, 프로모터는 적합하게는 셀로비오히드롤라제, 엔도글루카나제, 또는 베타-글루코시다제 프로모터일 수 있다. 특히 적합한 프로모터는, 예를 들어, *티. 레에세이*(*T. reesei*) 셀로비오히드롤라제, 엔도글루카나제, 또는 베타-글루코시다제 프로모터일 수 있다. 예를 들어, 프로모터는 셀로비오히드롤라제 I (*cbh1*) 프로모터이다. 프로모터의 비제한적 예는 *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *pki1*, *gpd1*, *xyn1*, 또는 *xyn2* 프로모터를 포함한다. 프로모터의 추가의 비제한적 예는 *티. 레에세이* *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl4*, *egl5*, *pki1*, *gpd1*, *xyn1*, 또는 *xyn2* 프로모터를 포함한다.

[0176] 본원에서 Mg3A 폴리펩티드를 암호화하는 핵산 서열은 벡터에 포함될 수 있다. 일부 측면에서, 벡터는 발현 제어 서열의 제어하에 Mg3A 폴리펩티드를 암호화하는 핵산 서열을 함유한다. 일부 측면에서, 발현 제어 서열은 천연 발현 제어 서열이다. 일부 측면에서, 발현 제어 서열은 비천연 발현 제어 서열이다. 일부 측면에서, 벡터는 선택적 마커 또는 선택가능한 마커를 함유한다. 일부 측면에서, Mg3A 폴리펩티드를 암호화하는 핵산 서열은 선택가능한 마커없이 숙주 세포의 염색체로 통합된다.

[0177] 적합한 벡터는 사용된 숙주 세포에 적합한 것들이다. 적합한 벡터는 예를 들어, 세균, 바이러스 (예컨대, 박테

리오파지 T7 또는 M-13 유래의 파지), 코스미드, 효모 또는 식물로부터 유래될 수 있다. 적합한 벡터는 숙주 세포에 저, 중, 또는 고 카피 수로 유지될 수 있다. 이러한 벡터를 얻어 사용하기 위한 프로토콜은 통상의 기술자에게 공지되어 있다 (예를 들어, 문헌 [Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor, 1989] 참조).

[0178] 일부 측면에서, 발현 벡터는 또한 종결 서열을 포함한다. 종결 제어 영역은 또한 숙주 세포 고유의 다양한 유전자로부터 유래될 수 있다. 일부 측면에서, 종결 서열 및 프로모터 서열은 동일한 공급원으로부터 유래된다.

[0179] Mg3A 폴리펩티드를 암호화하는 핵산 서열은 표준 기법을 이용하여, 벡터, 예컨대 발현 벡터에 혼입될 수 있다 (문헌 [Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1982]).

[0180] 일부 측면에서, 본 개시내용에 기재된 Mg3A 폴리펩티드 및/또는 하나 이상의 임의의 다른 핵산을 자연 발생 세포에서 현재 발견된 것보다 훨씬 높은 수준으로 과발현시키는 것이 바람직할 수 있다. 일부 실시양태에서, 본 개시내용에 기재된 내인성 베타-글루코시다제 및/또는 하나 이상의 임의의 다른 핵산을 자연 발생 세포에서 현재 발견된 것보다 훨씬 낮은 수준으로 저발현 (예를 들어, 돌연변이, 불활성화, 또는 결실)시키는 것이 바람직할 수 있다.

[0181] **B. Mg3a 폴리뉴클레오티드**

[0182] 본원에 기재된 조성물 및 방법의 또 다른 측면은 베타-글루코시다제 활성을 갖는 재조합 Mg3A 폴리펩티드 (그의 변이체 및 단편 포함)를 암호화하는 폴리뉴클레오티드 또는 핵산 서열이다. 일부 실시양태에서 폴리뉴클레오티드는 이중 유기체, 예컨대 본원에서 특정된 것에서 Mg3A 폴리펩티드의 발현을 지향하기 위한 발현 벡터의 맥락에서 제공된다. 재조합 Mg3A 폴리펩티드를 암호화하는 폴리뉴클레오티드는 조절 요소 (예를 들어, 프로모터, 터미네이터, 인핸서 등)에 작동가능하게 연결되어 암호화된 폴리펩티드를 발현시키는 것을 도울 수 있다.

[0183] 재조합 Mg3A 폴리펩티드를 암호화하는 폴리뉴클레오티드 서열의 예는 서열 번호 1의 뉴클레오티드 서열을 갖는다. 재조합 Mg3A 폴리펩티드 및 변이체를 암호화하는, 실질적으로 동일한 것을 포함한, 유사한 폴리뉴클레오티드는, 예를 들어, *마그나포르테 그리세아*, 또는 *마그나포르테* 종의 다른 균주 또는 단리물에서, 자연 발생할 수 있다. 유전 암호의 축중의 관점에서, 상이한 뉴클레오티드 서열을 갖는 폴리뉴클레오티드가 동일한 Mg3A 폴리펩티드, 변이체, 또는 단편을 암호화할 수 있음이 인식될 것이다.

[0184] 일부 실시양태에서, 재조합 Mg3A 폴리펩티드를 암호화하는 폴리뉴클레오티드는 Mg3A 폴리펩티드를 암호화하는 예시된 폴리뉴클레오티드에 대하여 특정된 정도의 아미노산 서열 동일성, 예를 들어, 서열 번호 2의 아미노산 서열에 대하여 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 또는 심지어는 적어도 99% 서열 동일성을 갖는다. 상동성은, 본원에 기재된 바와 같이, 예를 들어, BLAST, ALIGN, 또는 CLUSTAL과 같은 프로그램을 사용하여, 아미노산 서열 정렬에 의해 결정될 수 있다.

[0185] 일부 실시양태에서, 재조합 Mg3A 폴리펩티드를 암호화하는 폴리뉴클레오티드는 재조합 Mg3A 폴리펩티드의 세포의 분비를 지향하는 신호 펩티드를 위한 암호화 서열의 뒤 (즉, 서열의 하류) 프레임에서 융합된다. 본원에 기재된 바와 같이, 용어 "이중"은 관심 폴리펩티드를 발현시키는데 사용되는 신호 서열을 지칭하기 위해 사용되는 경우, 이는 신호 서열과 관심 폴리펩티드가 상이한 유기체로부터임을 의미하는 것이다. 이중 신호 서열은, 예를 들어, 다른 진균 셀룰라제 유전자로부터의 것, 예를 들어, 서열 번호 13의, *트리코테르마 레에세이* Bg11의 신호 서열 등을 포함한다. 발현 벡터는 재조합 Mg3A 폴리펩티드를 발현시키기에 적합하거나, 적합한 숙주 세포에 발현 벡터를 도입하기 이전에 발현 벡터를 증식시키기에 적합한 이중 숙주 세포에서 제공될 수 있다.

[0186] 일부 실시양태에서, 재조합 Mg3A 폴리펩티드를 암호화하는 폴리뉴클레오티드는 특정된 혼성화 조건하에 서열 번호 1의 폴리뉴클레오티드에 (또는 그의 상보체에) 혼성화한다. 조건의 예는 본원에 기재되는 바와 같이, 중간 엄격성, 높은 엄격성 및 극히 높은 엄격성 조건이다.

[0187] Mg3A 폴리뉴클레오티드는 자연 발생 또는 합성 (즉, 인공)일 수 있고, 상이한 숙주에서 발현시키기 위해 코돈-최적화되거나, 클로닝 부위를 도입하기 위해 돌연변이되거나, 또는 달리 기능성을 부가하기 위해 변경될 수 있다.

[0188] **C. Mg3A 벡터 및 숙주 세포**

[0189] 개시된 재조합 Mg3A 폴리펩티드를 생성하기 위해, 폴리펩티드를 암호화하는 DNA는 공개된 서열로부터 화학적으

로 합성될 수 있거나 유전자 함유 숙주 세포로부터 직접 수득할 수 있다 (예를 들어, cDNA 라이브러리 스크리닝 또는 PCR 증폭에 의해). 일부 실시양태에서, Mg3A 폴리뉴클레오티드는 발현 카세트에 포함되고/거나 표준 분자 클로닝 기법에 의해 적합한 발현 벡터로 클로닝된다. 이러한 발현 카세트 또는 벡터는 전사의 개시 및 종결을 돕는 서열 (예를 들어, 프로모터 및 터미네이터)을 함유하고, 전형적으로는 또한 하나 이상의 선택가능한 마커를 함유할 수 있다.

[0190]

발현 카세트 또는 벡터는 적합한 발현 숙주 세포에 도입된 다음에, 이는 상응하는 Mg3A 폴리뉴클레오티드를 발현시킨다. 적합한 발현 숙주는 세균 또는 진균 미생물일 수 있다. 세균 발현 숙주는, 예를 들어, *에스케리키아* (예를 들어, *에스케리키아 콜라이*), *슈도모나스* (예를 들어, *피. 플루오레센스*(*P. fluorescens*) 또는 *피. 스투츠제레이*(*P. stutzeri*)), *프로테우스*(*Proteus*) (예를 들어, *프로테우스 미라빌리스*(*Proteus mirabilis*)), *랄스토니아*(*Ralstonia*) (예를 들어, *랄스토니아 유티트로파*(*Ralstonia eutropha*)), *스트렙토마이세스*, *스타필로코커스*(*Staphylococcus*) (예를 들어, *에스. 카르노수스*(*S. carnosus*)), *락토코커스*(*Lactococcus*) (예를 들어, *엘. 락티스*(*L. lactis*)), 또는 *바실루스* (예를 들어, *바실루스 서브틸리스*, *바실루스 메가테리움*(*Bacillus megaterium*), *바실루스 리케니포르미스* 등)일 수 있다. 진균 발현 숙주는, 예를 들어, 에타놀로겐으로서도 역할을 할 수 있는 효모일 수 있다. 효모 발현 숙주는, 예를 들어, *사카로마이세스 세레비지에*, *스키조사카로마이세스 폼베*, *야로위아 리폴리티카*(*Yarrowia lipolytica*), *한세놀라 폴리모르파*, *클루이베로마이세스 락티스*(*Kluyveromyces lactis*) 또는 *피치아 파스토리스*일 수 있다. 진균 발현 숙주는 또한, 예를 들어, *아스페르길루스 니게르*, *크리소스포리움 토크노웬스*, *마이셀리오포토라 쉼모필라*, *아스페르길루스* (예를 들어, *에이. 오리자에*, *에이. 니게르*, *에이. 니들란스* 등) 또는 *트리코테르마 레에세이*를 포함하는 사상균 숙주일 수 있다. 또한, 포유동물 발현 숙주, 예컨대 마우스 (예를 들어, NS0), 차이니스 햄스터 난소 (CHO) 또는 새끼 햄스터 신장 (BHK) 세포주가 적합하다. 다른 진핵생물 숙주, 예컨대 곤충 세포 또는 바이러스 발현 시스템 (예를 들어, 박테리오파지, 예컨대 M13, T7 파지 또는 람다, 또는 바이러스, 예컨대 바칼로바이러스)이 또한 Mg3A 폴리펩티드를 생성하는데 적합하다.

[0191]

관심 특성의 숙주에서 분비된 단백질과 연관되어 있는 프로모터 및/또는 신호 서열은 그 숙주 또는 다른 숙주에서 Mg3A 폴리펩티드의 이중 생성 및 분비에서 사용하기 위한 후보이다. 예로서, 사상균계에서, 셀로비오히드롤라제 I (cbh1), 글루코아밀라제 A (glcA), TAKA-아밀라제 (amyA), 자일라나제 (exlA), gpd-프로모터 cbh1, cbh11, 엔도글루카나제 유전자 eg1-eg5, Cel61B, Cel74A, gpd 프로모터, Pgl1, pkl1, EF-1알파, tef1, cDNA1 및 hex1을 위한 유전자를 구동하는 프로모터가 적합하고 다수의 상이한 유기체로부터 유래될 수 있다 (예를 들어, *에이. 니게르*, *티. 레에세이*, *에이. 오리자에*, *에이. 아와모리*, *에이. 니들란스*).

[0192]

일부 실시양태에서, Mg3A 폴리뉴클레오티드는, 재조합 Mg3A 폴리펩티드의 세포외 (또는 원형질막주위) 공간으로의 분비를 야기하며, 그로 인해 세포 상층액 (또는 원형질막주위 공간 또는 용해물)에서 효소 활성을 직접 검출할 수 있게 하는, 적합한 상동 또는 이중 신호 서열을 암호화하는 폴리뉴클레오티드와 재조합적으로 연관되어 있다. *에스케리키아 콜라이*, 관련 기술분야에 공지된 다른 그램 음성균 및 다른 유기체에 적합한 신호 서열은 HlyA, DsbA, Pbp, PhoA, PelB, OmpA, OmpT 또는 M13 파지 Gill 유전자의 발현을 구동하는 것들을 포함한다. *바실루스 서브틸리스*, 관련 기술분야에 공지된 그램 양성균 및 다른 유기체의 경우, 적합한 신호 서열은 AprE, NprB, Mpr, AmyA, AmyE, Blac, SacB의 발현을 구동하는 것들을 추가로 포함하고, *에스. 세레비지에* 또는 다른 효모의 경우, 살독소, Bar1, Suc2, 교배 인자 알파, InulA 또는 Ggplp 신호 서열을 포함한다. 신호 서열은 다수의 신호 펩티다제에 의해 절단될 수 있고, 따라서 발현된 단백질의 나머지로부터 이들을 제거할 수 있다. 진균 발현 신호 서열은, 본원에서, 예를 들어, 서열 번호 13-37로부터 선택되는 것일 수 있다. 효모 발현 신호 서열은, 예를 들어, 서열 번호 38-40으로부터 선택되는 것일 수 있다. *지모모나스 모빌리스*에서 본 발명의 Mg3A 폴리펩티드를 발현시키는데 사용하기에 적합할 수 있는 신호 서열은, 예를 들어, 서열 번호 41-42로부터 선택된 것을 포함할 수 있다. (문헌 [Linger J.G. et al., (2010) Appl. Environ. Microbiol. 76(19):6360-6369]).

[0193]

일부 실시양태에서, 재조합 Mg3A 폴리펩티드는 단독으로 또는 N- 또는 C-말단에 위치된 다른 펩티드, 태그 또는 단백질 (예를 들어, 6XHis, HA 또는 FLAG 태그)과의 융합물로서 발현된다. 적합한 융합물은 친화도 정제 또는 검출을 촉진하는 태그, 펩티드 또는 단백질 (예를 들어, 6XHis, HA, 키틴 결합 단백질, 티오레독신 또는 FLAG 태그) 뿐만 아니라, 표적 베타-글루코시다제의 발현, 분비 또는 프로세싱을 촉진하는 것들을 포함한다. 적합한 프로세싱 부위는 엔테로키나제, STE13, Kex2 또는 생체내 또는 시험관내 절단을 위한 다른 프로테아제 절단 부위를 포함한다.

[0194]

Mg3A 폴리뉴클레오티드는 전기천공법, 지질-보조 형질전환 또는 형질감염 ("리포펙션"), 화학적으로 매개된 형

질감염 (예를 들어, CaCl 및/또는 CaP), 리튬 아세테이트-매개 형질전환 (예를 들어, 숙주 세포 원형질체의), 바이올리스틱(biolistic) "유전자총" 형질전환, PEG-매개 형질전환 (예를 들어, 숙주 세포 원형질체의), 원형질체 융합 (예를 들어, 세균 또는 진핵생물 원형질체), 리포솜-매개 형질전환, *아그로박테리움 투메파시엔스* (*Agrobacterium tumefaciens*), 아데노바이러스 또는 다른 바이러스 또는 파지 형질전환 또는 형질도입을 포함하나 그에 제한되지는 않는 다수의 형질전환 방법에 의해 발현 숙주 세포에 도입된다.

[0195] **D. 세포 배양 배지**

[0196] 일반적으로, 미생물은 본원에 기재된 Mg3A 폴리펩티드의 생성에 적합한 세포 배양 배지에서 배양된다. 배양은 관련 기술분야에 공지된 절차 및 변형을 이용하여, 탄소원 및 질소원 및 무기 염을 포함하는 적합한 영양 배지에서 실시된다. 성장 및 셀룰라제의 생성을 위한 적합한 배양 배지, 온도 범위 및 기타 조건이 관련 기술분야에 공지되어 있다. 비제한적인 예로서, *트리코테르마 레에세이*에 의한 셀룰라제 생성을 위한 전형적인 온도 범위는 24℃ 내지 37℃, 예를 들어, 25℃ 내지 30℃이다.

[0197] **1. 세포 배양 조건**

[0198] 진균 배양물의 유지 및 성장에 적합한 물질 및 방법은 관련 기술분야에 주지되어 있다. 일부 측면에서, 세포는 숙주 세포에 삽입된 핵산에 의해 암호화되는 하나 이상의 베타-글루코시다제 폴리펩티드의 발현을 가능하게 하는 조건하에 배양 배지에서 배양된다. 표준 세포 배양 조건은 세포를 배양하는데 사용될 수 있다. 일부 측면에서, 세포는 적절한 온도, 기체 혼합물, 및 pH 에서 성장되고 유지된다. 일부 측면에서, 세포는 적절한 세포 배지에서 성장된다.

[0199] **IV. Mg3A의 활성**

[0200] 본원에 개시된 재조합 Mg3A 폴리펩티드는 베타-글루코시다제 활성 또는 셀로비오스를 가수분해하고 그로부터 D-글루코오스를 유리시키는 능력을 갖는다. 본원에서의 Mg3A 폴리펩티드는 동일한 당화 조건하에, *트리코테르마 레에세이*의 기준 고 충실도 베타-글루코시다제 Bgl1보다 더 높은 베타-글루코시다제 활성 및 셀로비오스로부터 D-글루코오스를 유리시키는 개선되거나 증가된 능력을 갖는다. 일부 실시양태에서, 본원에서의 Mg3A 폴리펩티드는 *아스페르길루스 니게르*의 또 다른 기준 베타-글루코시다제 B-glu보다 더 높은 베타-글루코시다제 활성 및/또는 셀로비오스로부터 D-글루코오스를 유리시키는 개선되거나 증가된 능력을 가질 수 있다.

[0201] 실시예 3에 나타난 바와 같이, 재조합 Mg3A 폴리펩티드는, *트리코테르마 레에세이* Bgl1과 비교하여, 적어도 약 5%, 적어도 약 10%, 바람직하게는 적어도 약 15%, 더 바람직하게는 적어도 약 20% 더 높은, 클로로-니트로-페닐-글루코시드 (CNP) 기질을 가수분해하는 활성을 갖는다. 일부 실시양태에서, 재조합 Mg3A 폴리펩티드는, *아스페르길루스 니게르* B-glu와 비교하여, 적어도 약 2배, 적어도 약 5배, 적어도 약 7배, 바람직하게는 적어도 약 9배, 더 바람직하게는 적어도 약 10배 더 높은, CNP 기질을 가수분해하는 활성을 갖는다.

[0202] 재조합 Mg3A 폴리펩티드는, *트리코테르마 레에세이* Bgl1과 비교하여, 극적으로 개선되거나 증가된, 예를 들어, 적어도 약 30% 더 높은, 더 바람직하게는 적어도 약 40% 더 높은, 바람직하게는 적어도 약 50% 더 높은, 더 바람직하게는 적어도 55% 더 높은, 바람직하게는 적어도 60%, 더욱더 바람직하게는 적어도 약 65% 더 높은, 바람직하게는 적어도 약 70% 더 높은, 더 바람직하게는 적어도 약 75% 더 높은, 가장 바람직하게는 적어도 약 80% 더 높은 셀로비아제 활성을 갖고, 이는 D-글루코오스를 유리시키는, 셀로비오스의 가수분해를 촉매하는 효소의 촉매의 능력을 측정한다. 일부 실시양태에서, 재조합 Mg3A 폴리펩티드는, *아스페르길루스 니게르* B-glu와 비교하여, D-글루코오스를 유리시키는, 셀로비오스의 가수분해를 촉매하는 능력의 약 1/2, 약 1/3, 약 1/4, 약 1/5, 또는 심지어는 약 1/6을 갖는다.

[0203] 일부 실시양태에서, 재조합 Mg3A 폴리펩티드는, *트리코테르마 레에세이* Bgl1과 비교하여, CNP/셀로비오스에 비해 약 10% 감소되거나, 약 20% 감소되거나, 약 30% 감소되거나, 또는 심지어는 약 40% 감소된 가수분해 활성 비를 갖는다. 일부 실시양태에서, Mg3A 폴리펩티드는, *아스페르길루스 니게르* B-glu와 비교하여, CNP/셀로비오스에 비해 약 2배, 약 5배, 약 7배, 약 10배, 약 15배, 또는 심지어는 약 20배 더 높은 상대 가수분해 활성 비를 갖는다.

[0204] 실시예 4에 나타난 바와 같이, 재조합 Mg3A 폴리펩티드는, *트리코테르마 레에세이* Bgl1과 비교하여, 더 많은 글루코오스 그러나 인산 팽윤된 셀룰로오스 기질로부터의 총 당 이하의 양을 생성하였다.

[0205] 실시예 5에 나타난 바와 같이, 재조합 Mg3A 폴리펩티드는 또한, *트리코테르마 레에세이* Bgl1과 비교하여, 더 많은 글루코오스 그러나 희석 암모니아로 전처리된 옥수수 대 기질로부터의 총 당 이하의 양을 생성하였다.

[0206] 실시예 6에 나타난 바와 같이, 재조합 Mg3A 폴리펩티드는, *트리코테르마 레에세이* Bgl1과 비교하여, 더 높은 가수분해 반응 온도, 55℃에서 글루칸 전환에서 더 효과적이고 희석 암모니아로 전처리된 옥수수 대 기질로부터 더 많은 글루코오스를 생성한다. 50℃ 초과 온도에서, *트리코테르마 레에세이* Bgl1은 불량한 열안정성으로 인한 효소 활성의 상실에 기인한 불량한 베타-글루코시다제 성능을 부여하는 것으로 관찰되었다. 그에 반해서, 비록 *트리코테르마 레에세이* Bgl1의 측정된 용해 온도 (Tm)는 Mg3A의 것보다 약 10℃ 높긴 하지만, 놀랍게도 Mg3A 활성 및 성능은 55℃ 만큼 높은 당화 온도에서 온전한 상태로 남는 것으로 관찰되었다. Mg3A의 이러한 열안정성 및 고온 성능 이점은 이를 바이오매스 전환을 위한 특히 적합한 베타-글루코시다제가 되게 한다.

[0207] **V. 재조합 베타-글루코시다제 Mg3A 폴리펩티드를 포함하는 조성물**

[0208] 본 개시내용은 재조합 Mg3A 폴리펩티드가 풍부하게 된 조작된 효소 조성물 (예를 들어 셀룰라제 조성물) 또는 발효 브로쓰를 제공한다. 일부 측면에서, 조성물은 셀룰라제 조성물이다. 셀룰라제 조성물은, 예를 들어 사상균 셀룰라제 조성물, 예컨대 *트리코테르마* 셀룰라제 조성물일 수 있다. 일부 측면에서, 조성물은 하나 이상의 셀룰라제 폴리펩티드를 암호화하는 하나 이상의 핵산을 포함하는 세포이다. 일부 측면에서, 조성물은 셀룰라제 활성을 포함하는 발효 브로쓰이고, 여기서 브로쓰는 바이오매스 샘플에 존재하는 셀룰로오스를 약 50 중량% 초과하여 당으로 전환시킬 수 있다. 본원에 사용된 바와 같은 용어 "발효 브로쓰" 및 "전체 브로쓰"는 발효 후에 회수 및/또는 정제를 전혀 겪지 않거나 최소한 겪는 조작된 미생물의 발효에 의해 생성된 효소 제제를 지칭한다. 발효 브로쓰는 사상균의 발효 브로쓰, 예를 들어, *트리코테르마*, *후미콜라*, *푸사리움*, *아스페르길루스*, *뉴로스포라*, *페니실리움*, *세팔로스포리움(Cephalosporium)*, *아클리아(Achlya)*, *포도스포라(Podospora)*, *엔도티아(Endothia)*, *무코르*, *코클리오볼루스(Cochliobolus)*, *피리컬라리아(Pyricularia)*, *마이셀리오프토라* 또는 *크리소스포리움* 발효 브로쓰일 수 있다. 특히, 발효 브로쓰는, 예를 들어, *트리코테르마* 종, 예컨대 *트리코테르마 레에세이*, 또는 *페니실리움* 종, 예컨대 *페니실리움 푸니쿨로숨* 중 하나일 수 있다. 발효 브로쓰는 또한 적합하게는 무세포 발효 브로쓰일 수 있다. 한 측면에서, 본 발명의 셀룰라제, 세포 또는 발효 브로쓰 조성물 중 어느 것이든 하나 이상의 헤미셀룰라제를 추가로 포함할 수 있다.

[0209] 일부 측면에서, 전체 브로쓰 조성물은 *티. 레에세이* 또는 그의 조작된 균주에서 발현된다. 일부 측면에서 전체 브로쓰는 *티. 레에세이*의 통합된 균주에서 발현되고 여기서 Mg3A 폴리펩티드를 포함한 다수의 셀룰라제는 *티. 레에세이* 숙주 세포의 계통에 통합되어 있다. 일부 측면에서, 통합된 *티. 레에세이* 균주에서 발현된 폴리펩티드의 하나 이상의 성분은 결실되어 있다.

[0210] 일부 측면에서, 전체 브로쓰 조성물은 *에이. 니게르(A. niger)* 또는 그의 조작된 균주에서 발현된다.

[0211] 대안으로, 재조합 Mg3A 폴리펩티드는 세포내에서 발현될 수 있다. 임의로, 효소 변이체의 세포내 발현, 또는 상기 언급된 것들과 같은 신호 서열을 사용한 원형질막주위 공간으로의 분비 후, 투과화 또는 용해 단계를 사용하여 재조합 Mg3A 폴리펩티드를 상청액에 방출할 수 있다. 막 장벽의 파괴는 기계적 수단, 예컨대 초음파, 가압 처리 (프렌치 프레스(French press)), 캐비테이션(cavitation)의 사용에 의해, 또는 막-소화 효소, 예컨대 리소자임 또는 효소 혼합물의 사용에 의해 수행된다. 이러한 실시양태의 변형은 세포내에서 에타놀로겐 미생물에서의 재조합 Mg3A 폴리펩티드의 발현을 포함한다. 예를 들어, 셀로비오스 수송체는 동일 에타놀로겐 미생물에 유전 공학을 통해 도입될 수 있어, 리그노셀룰로오스계 바이오매스의 가수분해로부터 생성된 셀로비오스가 에타놀로겐 유기체에 수송될 수 있고, 거기서 가수분해되고 D-글루코오스로 될 수 있고, 결국 에타놀로겐에 의해 대사될 수 있도록 한다.

[0212] 일부 측면에서, 재조합 Mg3A 폴리펩티드를 암호화하는 폴리뉴클레오티드는 적합한 무세포 발현 시스템을 사용하여 발현된다. 무세포 시스템에서, 관심 폴리뉴클레오티드는 전형적으로 프로모터의 보조로 전사되지만, 원형 발현 벡터를 형성하는 결정은 임의적이다. 일부 실시양태에서, RNA는 전사 없이 외부에서 부가되거나 발생되고 무세포 시스템에서 번역된다.

[0213] **VI. 리그노셀룰로오스계 바이오매스 기질을 가수분해하기 위한 Mg3A 폴리펩티드의 사용**

[0214] 일부 측면에서, 리그노셀룰로오스 바이오매스 기질을 발효성 당으로 전환시키는데 효과적인 양으로 Mg3A 폴리펩티드를 포함하는 본원에 개시된 조성물과 상기 바이오매스 기질을 접촉시키는 것을 포함하는, 리그노셀룰로오스 바이오매스를 당으로 전환시키는 방법이 본원에 제공된다. 일부 측면에서, 상기 방법은 바이오매스를 산 및/또는 염기 및/또는 기계적 또는 다른 물리적 수단으로 전처리하는 것을 추가로 포함한다. 일부 측면에서 산은 인산을 포함한다. 일부 측면에서, 염기는 수산화나트륨 또는 암모니아를 포함한다. 일부 측면에서, 기계적 수단은, 예를 들어, 풀링(pulling), 프레스싱(pressing), 크러싱(crushing), 그라인딩(grinding), 및 리그노셀룰로오

스게 바이오매스를 더 작은 물리적 형태로 물리적으로 분해하는 다른 수단을 포함할 수 있다. 다른 물리적 수단은 또한, 예를 들어, 셀룰로오스 및 헤미셀룰로오스에 대한 효소에 의한 접근성을 증가시키기 위해, 리그노셀룰로오스계 바이오매스를 "느슨하게 하는(loosen)" 스팀 또는 다른 가압 흡(fume) 또는 증기(vapor)를 사용하는 것을 포함할 수 있다. 특정 실시양태에서, 전처리의 방법은 또한 리그노셀룰로오스계 바이오매스 기질의 리그닌을 분해할 수 있는 효소를 수반할 수 있어, 바이오매스의 셀룰로오스 및 헤미셀룰로오스에 대한 효소 조성물을 가수분해하는 바이오매스의 효소의 접근성이 증가되도록 한다.

[0215] **바이오매스:** 본 개시내용은 본 개시내용의 효소 조성물을 사용하는, 바이오매스 당화의 방법 및 공정을 제공한다. 본원에 사용된 바와 같은 용어 "바이오매스"는 셀룰로오스 및/또는 헤미셀룰로오스 (임의로 또한 리그노셀룰로오스계 바이오매스 물질 중의 리그닌)를 포함하는 임의의 조성물을 지칭한다. 본원에 사용된 바와 같이, 바이오매스는 제한 없이, 종자, 난알, 덩이줄기, 식물 폐기물 (예를 들어, 야자 나무의 빈 과실 송이, 또는 야자 섬유 폐기물), 또는 식품 가공 또는 산업적 가공의 부산물 (예를 들어, 줄기), 옥수수 (예를 들어, 옥수수 속대, 옥수수 대 등 포함), 목초 (예를 들어, 인디안 그래스(Indian grass), 예컨대 *소르카스트럼 누탄스* (*Sorghastrum nutans*); 또는, 스위치그래스(switchgrass), 예를 들어, *파니쿰(Panicum)* 종, 예컨대 *파니쿰 비르가툼(Panicum virgatum)* 포함), 다년생 줄기 (예를 들어, 물대), 목재 (예를 들어, 목재 칩, 가공 폐기물 포함), 종이, 펄프, 및 재생지 (예를 들어, 신문지, 인쇄 용지 등 포함)를 포함한다. 다른 바이오매스 물질은 제한 없이, 감자, 대두 (예를 들어, 평지씨), 보리, 호밀, 귀리, 밀, 비트(beet) 및 사탕수수 바가스(sugar cane bagasse)를 포함한다.

[0216] 따라서 본 개시내용은 바이오매스 물질, 예를 들어, 자일란, 헤미셀룰로오스, 셀룰로오스 및/또는 발효성 당을 포함하는 물질을 포함하는 조성물을 본 개시내용의 Mg3A 폴리펩티드, 또는 본 개시내용의 핵산 또는 폴리뉴클레오티드에 의해 암호화된 Mg3A 폴리펩티드, 또는 본 개시내용의 Mg3A 폴리펩티드를 포함하는 셀룰라제 또는 비-자연 발생 헤미셀룰라제 조성물 또는 본 개시내용의 제조 제품 중 임의의 것과 접촉시키는 것을 포함하는 당화 방법을 제공한다.

[0217] 당화된 바이오매스 (예를 들어, 본 개시내용의 효소에 의해 처리되는 리그노셀룰로오스계 물질)는, 예를 들어, 미생물 발효 및/또는 화학적 합성과 같은 공정을 통하여, 다수의 바이오-기반의 제품으로 제조될 수 있다. 본원에 사용된 바와 같이, "미생물 발효"는 적합한 조건하에 발효 미생물을 성장시키고 수집하는 공정을 말한다. 발효 미생물은 바이오-기반의 제품의 제조를 위한 원하는 발효 공정에서 사용하기에 적합한 임의의 미생물일 수 있다. 적합한 발효 미생물은 제한 없이, 사상균, 효모 및 세균을 포함한다. 당화된 바이오매스는, 예를 들어, 발효 및/또는 화학적 합성을 통해 그것이 연료 (바이오연료(biofuel), 예를 들어, 바이오에탄올, 바이오부탄올, 바이오메탄올, 바이오프로판올, 바이오디젤, 제트 연료 등)로 제조될 수 있다. 또한, 당화된 바이오매스는, 예를 들어, 발효 및/또는 화학적 합성을 통해, 범용 화학물질 (예를 들어, 아스코르브산, 아이소프렌, 1,3-프로판다이올), 지질, 아미노산, 폴리펩티드 및 효소로 제조될 수 있다.

[0218] **전처리:** 당화 또는 효소적 가수분해 및/또는 당화로부터 생긴 발효성 당의 발효 전에, 바이오매스 (예를 들어, 리그노셀룰로오스계 물질)는 바람직하게는 하나 이상의 전처리 단계(들)에 적용되어, 자일란, 헤미셀룰로오스, 셀룰로오스 및/또는 리그닌 물질이 효소 조성물 (예를 들어, Mg3A 폴리펩티드를 포함하는 본 발명의 효소 조성물) 중 효소에 더욱 접근가능하거나 민감하게 하며, 따라서 효소(들) 및/또는 효소 조성물에 의해 더욱 가수분해될 수 있게 한다.

[0219] 일부 측면에서, 적합한 전처리 방법은 바이오매스 물질을 반응기 내에서 금속 염 및 강산의 희석 용액을 포함하는 촉매에 적용하는 것을 수반할 수 있다. 바이오매스 물질은, 예를 들어, 원료 또는 건조 물질일 수 있다. 이러한 전처리는 셀룰로오스 가수분해의, 활성화 에너지 또는 온도를 낮추어, 궁극적으로 발효성 당의 수율이 더 높아지게 할 수 있다. 예를 들어, 미국 특허 제6,660,506호; 제6,423,145호 참조.

[0220] 일부 측면에서, 적합한 전처리 방법은 바이오매스 물질을 셀룰로오스의 글루코오스로의 상당한 해중합을 달성함이 없이 주로 헤미셀룰로오스의 해중합을 유발하기 위해 선택된 온도 및 압력에서 수성 매질 중에서 제1 가수분해 단계에 적용하는 것을 수반할 수 있다. 이러한 단계에 의해, 슬러리가 산출되며, 여기서, 액체 수상은 헤미셀룰로오스의 해중합으로부터 생기는 용해된 단당류를 함유하며, 고체상은 셀룰로오스와 리그닌을 함유한다. 그 다음에, 슬러리를 상당 부분의 셀룰로오스가 해중합되게 하는 조건하에 제2 가수분해 단계에 적용하여, 셀룰로오스의 용해된/가용성 해중합 생성물을 함유하는 액체 수상을 산출한다. 예를 들어, 미국 특허 제5,536,325호 참조.

[0221] 추가 측면에서, 적합한 전처리 방법은 바이오매스 물질을 약 0.4% 내지 약 2%의 강산을 사용한 묽은 산 가수분

해의 하나 이상의 단계에 의해 처리하고; 이어서, 산에 의해 가수분해된 물질의 미반응된 고체 리그노셀룰로오스계 성분을 알칼리 탈리그닌화(alkaline delignification)로 처리하는 것을 수반할 수 있다. 예를 들어, 미국 특허 제6,409,841호 참조.

[0222] 더 추가 측면에서, 적합한 전처리 방법은 전가수분해 반응기에서 바이오매스 (예를 들어, 리그노셀룰로오스계 물질)를 전가수분해하고; 산성 액체를 고체 리그노셀룰로오스계 물질에 첨가하여 혼합물을 제조하고; 혼합물을 반응 온도로 가열하고; 리그노셀룰로오스계 물질을 리그노셀룰로오스계 물질로부터 적어도 약 20%의 리그닌을 함유하는 가용화된 부분, 및 셀룰로오스를 함유하는 고체 분획으로 분별시키기에 충분한 기간 동안 반응 온도를 유지하고; 반응 온도에서 또는 반응 온도 근처에 있는 동안, 고체 분획으로부터 가용화된 부분을 분리하고, 가용화된 부분을 제거하고; 가용화된 부분을 회수하는 것을 수반할 수 있다. 고체 분획 중의 셀룰로오스는 효소에 의해 더욱 분해될 수 있게 된다. 예를 들어, 미국 특허 제5,705,369호 참조. 이러한 측면의 변형에서, 전가수분해는, 예를 들어, 리그노셀룰로오스계 바이오매스 물질의 리그닌을 분해할 수 있는 효소를 사용한 전가수분해를 대안으로 또는 추가로 수반할 수 있다.

[0223] 더 추가 측면에서, 적합한 전처리는 과산화수소 H_2O_2 의 사용을 수반할 수 있다. 문헌 [Gould, 1984, Biotech, and Bioengr. 26:46-52] 참조.

[0224] 다른 측면에서, 또한, 전처리는 바이오매스 물질을 매우 낮은 농도에서 화학량론적 양의 수산화나트륨 및 수산화암모늄과 접촉시키는 것을 포함할 수 있다. 문헌 [Teixeira *et al.*, (1999) Appl. Biochem. and Biotech. 77-79:19-34] 참조.

[0225] 일부 실시양태에서, 전처리는 리그노셀룰로오스를 적당한 온도, 압력 및 pH에서, 약 9 내지 약 14의 pH에서의 화학물질 (예를 들어, 염기, 예컨대 탄산나트륨 또는 수산화칼륨)과 접촉시키는 것을 포함할 수 있다. 공개된 국제 출원 WO2004/081185 참조. 예를 들어, 바람직한 전처리 방법에서 암모니아가 사용된다. 이러한 전처리 방법은 바이오매스 물질을 높은 고형분의 조건하에 낮은 암모니아 농도에 적용하는 것을 포함한다.

[0226] 예를 들어, 미국 특허 공개 제20070031918호 및 공개된 국제 출원 WO 06110901 참조.

[0227] A. 당화 공정

[0228] 일부 측면에서, 바이오매스를 폴리펩티드를 포함하는 효소 조성물로 처리하는 것을 포함하는 당화 공정이 본원에 제공되며, 여기서 폴리펩티드는 베타-글루코시다제 활성을 갖고, 상기 공정은 바이오매스를 적어도 약 50 wt.% (예를 들어, 적어도 약 55 wt.%, 60 wt.%, 65 wt.%, 70 wt.%, 75 wt.%, 또는 80 wt.%)로 발효성 당으로 전환시킨다. 일부 측면에서, 바이오매스는 리그닌을 포함한다. 일부 측면에서 바이오매스는 셀룰로오스를 포함한다. 일부 측면에서 바이오매스는 헤미셀룰로오스를 포함한다. 일부 측면에서, 셀룰로오스를 포함하는 바이오매스는 자일란, 갈락탄, 또는 아라비난 중 하나 이상을 추가로 포함한다. 일부 측면에서, 바이오매스는 제한 없이, 종자, 낱알, 덩이줄기, 식물 폐기물 (예를 들어, 야자 나무의 빈 과일 송이, 또는 야자 섬유 폐기물) 또는 식품 가공 또는 산업적 가공의 부산물 (예를 들어, 줄기), 옥수수 (예를 들어, 옥수수 속대, 옥수수 대 등 포함), 목초 (예를 들어, 인디안 그래스, 예컨대 *소르카스트럼 누탄스*; 또는, 스위치그래스, 예를 들어, *파니쿨* 종, 예컨대 *파니쿨 비르가툼* 포함), 다년생 줄기 (예를 들어, 물대), 목재 (예를 들어, 목재 칩, 가공 폐기물 포함), 종이, 펄프, 및 재생지 (예를 들어, 신문지, 인쇄 용지 등 포함), 감자, 대두 (예를 들어, 평지씨), 보리, 호밀, 귀리, 밀, 비트, 및 사탕수수 바가스일 수 있다. 일부 측면에서, 바이오매스를 포함하는 물질은 폴리펩티드로 처리하기 이전에 하나 이상의 전처리 방법/단계에 적용된다. 일부 측면에서, 당화 또는 효소적 가수분해는 바이오매스를 본 발명의 Mg3A 폴리펩티드를 포함하는 효소 조성물로 처리하는 것을 추가로 포함한다. 효소 조성물은, 예를 들어, Mg3A 폴리펩티드 이외에도, 하나 이상의 다른 셀룰라제를 포함할 수 있다. 대안으로, 효소 조성물은 하나 이상의 다른 헤미셀룰라제를 포함할 수 있다. 특정 실시양태에서, 효소 조성물은 본 발명의 Mg3A 폴리펩티드, 하나 이상의 다른 셀룰라제, 하나 이상의 헤미셀룰라제를 포함한다. 일부 실시양태에서, 효소 조성물은 전체 브로쓰 조성물이다.

[0229] 일부 측면에서, 리그노셀룰로오스계 바이오매스 물질을 폴리펩티드를 포함하는 조성물로 처리하는 것을 포함하는 당화 공정이 제공되며, 여기서, 폴리펩티드는 서열 번호 2에 대하여 적어도 약 75% (예를 들어, 적어도 약 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) 서열 동일성을 갖고, 상기 공정은 바이오매스를 적어도 약 50% (예를 들어, 적어도 약 55%, 60%, 65%, 70%, 75%, 80%, 85%, 또는 90%) (중량 기준)로 발효성 당으로 전환시킨다. 일부 측면에서, 리그노셀룰로오스계 바이오매스 물질은 본원에 기재된 바와 같은 하나 이상의 전처리 방법/단계에 적용되었다.

- [0230] 본 조성물 및 방법의 다른 측면 및 실시양태는 전술한 설명 및 하기 실시예에 의해 명백해질 것이다.
- [0231] **실시예**
- [0232] 하기 실시예는 본 개시내용의 특정 바람직한 실시양태 및 측면을 실증하고 예시하기 위해 제공되며 한정하는 것으로 해석되어서는 안된다.
- [0233] **실시예 1**
- [0234] **1-A. *Mg3A* 및 기준 *티. 레에세이 Bg11*의 유전자 발현의 클로닝 및 발현**
- [0235] **1-A-a. *티. 레에세이 bg11* 발현 벡터의 구축**
- [0236] 천연 *티. 레에세이* β -글루코시다제 유전자 *bg11*의 N-말단 부분을 코돈 최적화시켰다 (DNA 2.0, 미국 캘리포니아주 멘로 파크). 이러한 합성된 부분은 이러한 효소의 암호화 영역의 최초 447개 염기를 포함하였다. 그 다음에, 이러한 단편을 프라이머 SK943 및 SK941 (하기)을 사용하여 PCR에 의해 증폭시켰다. 천연 *bg11* 유전자의 나머지 영역을 프라이머 SK940 및 SK942 (하기)를 사용하여, *티. 레에세이* 균주 RL-P37로부터 추출된 게놈 DNA 샘플로부터 PCR 증폭시켰다 (문헌 [Sheir-Neiss, G *et al.* (1984) Appl. Microbiol. Biotechnol. 20:46-53]). *bg11* 유전자의 이들 2개의 PCR 단편을 프라이머 SK943 및 SK942를 사용하여, 융합 PCR 반응에서 함께 융합시켰다:
- [0237] 정방향 프라이머 SK943: (5'- CACCATGAGATATAGAACAGCTGCCGCT-3') (서열 번호 5)
- [0238] 역방향 프라이머 SK941: (5'-CGACCGCCCTGCGGAGTCTTGCCAGTGGTCCGCGACAG-3') (서열 번호 6)
- [0239] 정방향 프라이머 (SK940): (5'-CTGTGCGGGGACCAGTGGGCAAGACTCCGAGGGCGGTCG-3') (서열 번호 7)
- [0240] 역방향 프라이머 (SK942): (5'- CCTACGCTACCGACAGAGTG-3') (서열 번호 8)
- [0241] 생성된 융합 PCR 단편을 게이트웨이(Gateway)® 엔트리(Entry) 벡터 pENTR™/D-TOPO® 로 클로닝하고 (도 1), *이. 콜라이 원 샷(One Shot)® TOP10* 화학적 적격(Chemically Competent) 세포 (Invitrogen(인비트로젠))로 형질전환시켜, 중간 벡터, pENTR TOPO-Bg11(943/942) (도 1)을 생성시켰다. 삽입된 DNA의 뉴클레오티드 서열을 결정하였다. 정확한 *bg11* 서열이 있는 pENTR-943/942 벡터를 LR 클로나제(clonase)® 반응을 사용하여 pTrex3g와 재조합시켰다 (인비트로젠에 의해 약속된 프로토콜 참조). LR 클로나제 반응 혼합물을 *이. 콜라이 원 샷®TOP10* 화학적 적격 세포 (인비트로젠)로 형질전환시켜, 발현 벡터, pTrex3g 943/942 (도 2)를 생성시켰다. 또한, 벡터는 *티. 레에세이*의 형질전환을 위한 선택가능한 마커로서, 아세트아미다제를 암호화하는 *아스페르길루스 니들란스 amdS* 유전자를 함유하였다. 발현 카세트를 프라이머 SK745 및 SK771 (하기)로 PCR 증폭시켜, 형질전환을 위한 생성물을 생성하였다.
- [0242] 정방향 프라이머 SK771: (5' - GTCTAGACTGGAAACGCAAC -3') (서열 번호 9)
- [0243] 역방향 프라이머 SK745: (5' - GAGTTGTGAAGTCGGTAATCC -3') (서열 번호 10)
- [0244] **1-A-b. *mg3a* 발현 벡터의 구축**
- [0245] 베타-글루코시다제 유전자의 오픈 리딩 프레임을 주형으로서 *마그나포르테 그리세아*로부터 추출된 게놈 DNA를 사용하여 PCR에 의해 증폭시켰다. 오픈 리딩 프레임을 천연 신호 서열로 증폭시켰다. 사용된 PCR 써모사이클러(thermocycler)는 DNA 엔진 테트라드(Engine Tetrad) 2 펠티에르 써멀 사이클러(Peltier Thermal Cycler) (바이오-래드 레보러터리즈(Bio-Rad Laboratories))이었다. 사용된 DNA 중합효소는 PfuUltra II 퓨전(Fusion) HS DNA 중합효소 (스트라타진(Stratagene)) 또는 유사한 품질의 프루프리딩(proofreading) DNA 중합효소이었다. 오픈 리딩 프레임을 증폭시키기 위해 사용된 프라이머는 하기와 같았다:
- [0246] Mg3A-F: 5'-CAC CAT GCG TTT CTC CGG GAT CGT -3' (서열 번호 11)
- [0247] Mg3A-R: 5'-TCA GTT CAG GTC AGC ACT CAG ATG GAG C-3' (서열 번호 12)
- [0248] Mg3A-F 정방향 프라이머는 5'-말단에 4개의 추가의 뉴클레오티드 (서열 - CACC)를 포함시켜, pENTR/D-TOPO로의 방향성 클로닝을 촉진하였다. 오픈 리딩 프레임의 PCR 산물을 퀴아퀵(Qiaquick) PCR 정제 키트 (퀴아젠(Qiagen), 캘리포니아주 발렌시아)를 사용하여 정제하였다. 정제된 PCR 산물을 pENTR/D-TOPO 벡터 (인비트로젠)로 클로닝하고, TOP10 화학적 적격 *이. 콜라이* 세포 (인비트로젠, 캘리포니아주 칼스배드)로 형질전환시키고, 50 ppm 카나마이신을 함유하는 LA 플레이트 상에 플레이트하였다. 플라스미드 DNA를 퀴아스핀

(QIAspin) 플라스미드 제조 키트 (퀴아젠)를 사용하여 이. 콜라이 형질전환체로부터 수득하였다.

[0249] pENTR/D-TOPO 벡터 내에 삽입된 DNA에 대한 서열 테이터를 M13 정방향 및 역방향 프라이머를 사용하여 수득하였다. 오픈 리딩 프레임의 정확한 DNA 서열이 있는 pENTR/D-TOPO 벡터를 제조업체의 사용 설명서에 따라 LR 클로나제 반응 혼합물 (인비트로젠, 캘리포니아주 칼스배드)을 사용하여 pTrex3gM 데스티네이션(destination) 벡터 (도 2)와 재조합하였다.

[0250] LR 클로나제 반응의 산물을 이후에 TOP10 화학적 적격 이. 콜라이 세포로 형질전환시킨 다음에, 이를 50 ppm 카르베니실린을 함유하는 LA 상에 플레이트하였다. 생성된 pExpression 구축물은 Mg3A 오픈 리딩 프레임 및 아스페르길루스 투빙겐시스(*Aspergillus tubingensis*) 아세트아미다제 선택 마커 (*amdS*)를 함유하는 pTrex3gM이었다. pExpression 구축물의 DNA를 퀴아젠 미니프렙(miniprep) 키트를 사용하여 단리하고, 트리코테르마 레에세이의 형질전환을 위해 사용하였다.

[0251] pExpression 플라스미드 또는 발현 카세트의 PCR 산물 중 하나를 하기에 기재된 바와 같이 약간 변형을 가한 PEG-매개 원형질체 방법을 사용하여 *티. 레에세이* 6배-결실 균주로 형질전환시켰다 (예를 들어, 공개된 국제 특허 출원 공보 제WO 2010/141779호의 기재 참조). 원형질체 제조를 위해, 포자를 트리코테르마 최소 배지 MM - 20 g/L 글루코오스, 15 g/L KH₂PO₄, pH 4.5, 5 g/L (NH₄)₂SO₄, 0.6 g/L MgSO₄·7H₂O, 0.6 g/L CaCl₂ x 2H₂O, 1 mL의 1000 X *티. 레에세이* 미량 원소 용액 (이는 5 g/L FeSO₄ x 7H₂O, 1.4 g/L ZnSO₄ x 7H₂O, 1.6 g/L MnSO₄ x H₂O, 3.7 g/L CoCl₂ x 6H₂O를 함유함)을 함유함 - 에서 150 rpm으로 진탕하면서 24℃에서 16 내지 24시간 동안 성장시켰다. 발아 포자를 원심분리에 의해 수집하고, 50 mg/mL의 글루카넥스(Glucanex) G200 (노보자임즈 아게 (Novozymes AG)) 용액으로 처리하여, 진균 세포벽을 용해시켰다. 원형질체의 추가 제조는 문헌 [Penttilä *et al.* (1987) Gene 61: 155-164]에 기재된 방법에 따라 수행하였다. 200 µL의 총 부피에서 약 1 µg의 DNA 및 1 내지 5×10⁷개의 원형질체를 함유하는 형질전환 혼합물을 각각 2 mL의 25% PEG 용액으로 처리하고, 2개의 부피의 1.2 M 소르비톨/10 mM Tris, pH7.5, 10 mM CaCl₂로 희석하고, 5 mM 우리딘 및 20 mM 아세트아미드를 함유하는 3% 선택적 탑 아가로스(top agarose) MM과 혼합하였다. 생성된 혼합물을 우리딘 및 아세트아미드를 함유하는 2% 선택적 아가로스 플레이트에 부었다. 단일 형질전환체를 우리딘 및 아세트아미드를 함유하는 새로운 MM 플레이트 상에 취하기 전에, 플레이트를 28℃에서 7 내지 10일 동안 추가로 인큐베이션하였다. 독립 클론으로부터의 포자를 사용하여, 진탕 플라스크에서 발효 배지를 접종시켰다.

[0252] 발효 배지는 250 mL 톰슨 울트라 일드 플라스크 (Thomson Ultra Yield Flask) (톰슨 인스트루먼트 캄파니 (Thomson Instrument Co.), 캘리포니아주 오션사이드)에서, 글루코스/소포로오스 및 2 g/L 우리딘을 함유하는 36 mL의 규정된 브로쓰, 예를 들어, 글리신 최소 배지 (6.0 g/L 글리신; 4.7 g/L (NH₄)₂SO₄; 5.0 g/L KH₂PO₄; 1.0 g/L MgSO₄·7H₂O; 33.0 g/L PIPPS; pH 5.5)이었다 (탄소원으로서 ~2% 글루코오스/소포로오스 혼합물, 10 ml/L의 100g/L의 CaCl₂, 2.5 ml/L의 *티. 레에세이* 미량 원소 (400X) : 175g/L 무수 시트르산; 200 g/L FeSO₄·7H₂O; 16 g/L ZnSO₄·7H₂O; 3.2 g/L CuSO₄·5H₂O; 1.4 g/L MnSO₄·H₂O; 0.8 g/L H₃BO₃이 멸균 후 첨가됨).

[0253] 1-A-c. 효모 서틀 벡터 pSC11의 구축

[0254] 효모 서틀 벡터는 도 7의 벡터 맵에 따라 구축될 수 있다. 이러한 벡터를 사용하여 세포내에서 *사카로마이세스 세레비지에*에서 Mg3A 폴리펩티드를 발현시킬 수 있다. 셀로비오스 수송체는 공지된 방법, 예를 들어, 문헌 [Ha *et al.*, (2011), PNAS, 108(2): 504-509]에 기재된 것 등을 사용하여 동일한 서틀 벡터에서 또는 별개의 벡터에서 *사카로마이세스 세레비지에*에 도입될 수 있다.

[0255] 효모 EZ-형질전환 키트를 사용하여 발현 카세트의 형질 전환을 수행할 수 있다. 형질전환체는 20 g/L 셀로비오스를 함유하는 YSC 배지를 사용하여 선택될 수 있다. 발현 카세트의 효모로의 성공적 도입은 특이적 프라이머로 콜로니 PCR에 의해 확인될 수 있다.

[0256] 효모 균주는 공지된 방법 및 프로토콜에 따라 배양될 수 있다. 예를 들어, 균주는 20 g/L 글루코오스를 함유하는 YP 배지 (10 g/L 효모 추출물, 20 g/L 박토 펩톤)에서 30℃에서 배양될 수 있다. 아미노산 영양 요구성 마커를 사용하여 형질전환체를 선택하기 위해, 효모 합성 완전 (YSC) 배지가 사용될 수 있고, 이는 6.7 g/L 효모 질소 베이스에 더하여 20 g/L 글루코오스, 20 g/L 아가, 및 CSM-Leu-Trp-Ura를 함유하여 뉴클레오티드 및 아미노산을 제공한다.

[0257]

1-A-d. 지모모나스 모빌리스 통합 벡터 pZC11의 구축.

[0258]

지모모나스 모빌리스 통합 벡터 pZC11은 도 8의 벡터 맵에 따라 구축될 수 있다. 이러한 벡터를 사용하여 세포 내에서 지모모나스 모빌리스에서 Mg3A 폴리펩티드를 발현할 수 있다. 셀로비오스 수송체는 그러한 수송체를 세균 세포에 도입하는 공지된 방법, 예를 들어, 문헌 [Sekar *et al.*, (2012) Applied Environmental Microbiology, 78(5):1611-1614]에 기재된 것 등을 사용하여 동일한 통합 벡터에서 또는 별개의 벡터에서 지모모나스 모빌리스에 도입될 수 있다.

[0259]

통합 벡터뿐만 아니라 셀로비오스 수송체 유전자의 성공적 도입은 다양한 공지된 접근법을 사용하여, 예를 들어 이러한 목적으로 특이적으로 설계된 확인 프라이머를 사용하여 PCR에 의해 확인될 수 있다.

[0260]

지모모나스 모빌리스 균주는, 공지된 방법, 예를 들어, 미국 특허 제7,741,119호에 기재된 것 등에 따라 배양되고 발효될 수 있다.

[0261]

1-B. 티. 레에세이 Bgl1 및 Mg3A의 정제

[0262]

티. 레에세이 Bgl1을 6배 결실된 트리코테르마 레에세이 숙주 균주의 발효 브로쓰에서 과발현시켰고, 상기 브로쓰로부터 정제하였다 (예를 들어, 공개된 국제 특허 출원 공보 제WO 2010/141779호에서의 기재 참조). 농축 브로쓰를 G25 SEC 칼럼 (지이 헬스케어 바이오-사이언시즈(GE Healthcare Bio-Sciences)) 상에 로딩하고 50 mM 아세트산나트륨, pH 5.0에 대해 완충제-교환하였다. 그 다음에 완충제 교환된 Bgl1을 아미노 벤질-S-글루코피라노실 세파로오스 친화 매트릭스로 충전된 25 mL 칼럼 상에 로딩하였다. 50 mM 아세트산나트륨, pH 5.0 중 250 mM 염화나트륨으로 광범위하게 세척한 후, 결합된 분획을 50 mM 아세트산나트륨 및 250 mM 염화나트륨, pH 5.0 중 100 mM 글루코오스로 용리하였다. 클로로-니트로-페닐 글루코시드 (CNP)G 활성화에 양성으로 시험된 용리된 분획을 모으고 농축하였다. SDS-PAGE 상에 그리고 질량 분석법에 의해 확인된 티. 레에세이 Bgl1의 MW에 상응하는 단일 밴드는 용리된 Bgl1의 순도를 입증하였다. 최종 스톱 농도는 280 nm에서의 흡광도에 의해 2.2 mg/mL로 결정되었다.

[0263]

상기 기재된 바와 같은 트리코테르마 레에세이에 의해 발현된 Mg3A를 먼저 100 mg을 50 mM MES 완충제, pH 6.0로 희석함으로써 농축된 발효 브로쓰로부터 정제할 수 있다. Mg3A는 2 mg 단백질/mL 수지를 pH 6에서 충전된 SP 세파로오스 이온 교환 수지 (지이 헬스케어)에 로딩함으로써 풍부하게 될 수 있다. 그 다음에 Mg3A를 플로우-쓰루(flow-through)로 용리시킬 수 있다. 그 다음에 풍부하게 된 Mg3A를 10,000 MW 컷-오프(cut-off) 막 (비바스핀(Vivaspin), 지이 헬스케어)을 사용하여 원래 부피보다 5배 낮은 부피로 농축할 수 있다. 다른 백그라운드 성분들은 배치 모드(batch mode)로 40% 황산암모늄 (w/v)을 첨가함으로써 Mg3A로부터 제거될 수 있다. 그 다음에 순수 Mg3A를 원심분리 후 상청액 상에서 회수한다. 그 다음에 Mg3A를 10,000MW 컷 오프 막 (비바스핀, 지이 헬스케어)을 사용하여, 50 mM MES 완충제, pH 6.0에서 동시에 투석하고 농축한다. Mg3A의 활성 및 순도는 각각 클로로-니트로-페닐 글루코시드 검정 및 SDS-PAGE에 의해 평가될 수 있다. 그 다음에 상청액을 7,000 MW 컷-오프 막 투석 카세트 (피어스(PIERCE))를 사용하여 50 mM MES, 100 mM NaCl 완충제, pH 6.0에 대해 광범위하게 투석한다. 최종 Mg3A 배치의 활성은 클로로-니트로-페닐 글루코시드 검정에 의해 결정할 수 있다. 농도는 GPMW v 7.0에 의해 계산된 몰 흡광 계수를 사용하여 280 nm의 파장에서의 흡광도 검정 및 바이신 코닌산 검정 (피어스)에 의해 결정할 수 있다.

[0264]

실시예 2: 다양한 검정

[0265]

2-A. UPLC에 의한 단백질 농도 측정

[0266]

워터스(Waters) 액티비티(ACQUITY) UPLC BEH C4 칼럼 (1.7 μ m, 1 \times 50 mm)을 갖춘 애질런트(Agilent) HPLC 1290 인피니티(Infinity) 시스템을 단백질 정량에 사용하였다. 0.5분내 5% 내지 33% 아세토니트릴 (시그마-알드리치)의 초기 구매 후, 4.5분내 33% 내지 48%의 구매, 이어서 90% 아세토니트릴에 대한 단계 구매를 갖는 6분 프로그램을 사용하였다. 정제된 트리코테르마 레에세이 Bgl1을 기반으로 한 단백질 표준 곡선을 사용하여 Mg3A 폴리펩티드를 정량화하였다.

[0267]

2-B. 클로로-니트로-페닐-글루코시드 (CNP)G 가수분해 검정

[0268]

200 μ L의 50 mM 아세트산나트륨 완충제, pH 5를 마이크로타이터 플레이트의 각각의 웰에 첨가하였다. 50 mM 아세트산나트륨 완충제, pH 5에 희석시킨 5 μ L의 효소도 각각의 웰에 첨가하였다. 플레이트를 커버하여, 에펜도르프 썬믹서(Eppendorf Thermomixer)에서 15분 동안 37°C에서 평형화하였다. 밀리포어(Millipore) 물에서 제조한 20 μ L의 2 mM 2-클로로-4-니트로페닐-베타-D-글루코피라노시드 (CNP)G, 로즈 사이언티픽 리미티드(Rose

Scientific Ltd.), 캘리포니아주 에드먼턴)를 각각의 웰에 첨가하고, 플레이트를 신속하게 분광광도계 (스펙트라맥스(SpectraMax) 250, 몰레큘러 디바이시스(Molecular Devices))에 옮겼다. OD 405 nm에서 15분 동안 키네틱 리딩(kinetic read)을 행하고 데이터를 Vmax로서 기록하였다. CNP의 흡광 계수를 사용하여, Vmax를 OD/sec의 단위로부터 μM CNP/sec의 단위로 변환시켰다. μM CNP/sec를 검정에서 사용된 효소의 mg으로 나누어, 비활성도 (μM CNP/sec/mg 단백질)를 결정하였다. CNPG 검정에 대한 표준 오차는 3%로 결정되었다.

2-C. 셀로비오스 가수분해 검정

셀로비아제 활성을 문헌 [Ghose, T.K. *Pure & Applied Chemistry*, 1987, 59 (2), 257-268]의 방법을 사용하여 50°C에서 결정하였다. 셀로비오스 단위 (고세(Ghose)에 기재된 바와 같이 유래됨)는 검정 조건하에 0.1 mg 글루코오스를 방출하는데 필요한 효소의 양으로 나누어진 0.0926으로 정의된다. 셀로비아제 검정에 대한 표준 오차는 10%로 결정되었다.

2-D. 인산 팽윤된 셀룰로오스 (PASC)의 제조

인산 팽윤된 셀룰로오스 (PASC)를 문헌 [Walseth, TAPPI 1971, 35:228] 및 [Wood, *Biochem. J.* 1971, 121:353-362]의 개조된 프로토콜을 사용하여, 아비셀로부터 제조하였다. 요컨대, 아비셀 PH-101을 진한 인산에 가용화시킨 다음에, 냉각 탈이온수를 사용하여 침전시켰다. 셀룰로오스를 수집하고, 더 많은 물로 세정하여, pH를 중화시켰다. 이를 50 mM 아세트산나트륨 pH 5 중에서 1% 고형분으로 희석하였다.

실시예 3. CNPG 및 셀로비아제 검정에서 나타난 바와 같은, 기준 트리카테르마 레에세이 Bgl1에 비한 또는 기준 아스페르길루스 니케르 B-glu에 비한 Mg3A의 개선된 가수분해 성능

3-A. 진탕 플라스크에서 생성된 베타-글루코시다제의 CNPG 및 셀로비아제 활성

조 진탕 플라스크 브로쓰에서의 Mg3A의 농도를 UPLC (본원에 기재됨)에 의해 측정하고 0.041 g/L로 결정하였다. 두가지 셀로비오히드롤라제가 발현 균주 백그라운드 중 베타-글루코시다제 활성에 대한 대조군으로서 하기 실험에 포함되었고 검정의 검출 한계치 미만이었다. 2.2 mg/mL의 스톡으로부터 정제된 트리카테르마 레에세이 Bgl1을 사용하였다 (A280 측정). 정제된 *에이. 니케르* 베타-글루코시다제 B-glu를 BSA 없이, 메가자임 인터내셔널 (Megazyme International)로부터 구입하였다 (메가자임 인터내셔널 아일랜드 리미티드(Megazyme International Ireland Ltd.), 아일랜드 위클로, 로트(Lot) 번호 031809).

모델 기질 클로로-니트로-페닐-글루코시드 (CNPG) 및 셀로비오스에 대한 각각의 효소 활성을 측정하였다. 표준 프로토콜에서의 온도 (37°C에서 CNPG 및 50°C에서 셀로비오스)에서 검정을 각각 수행하였다.

[표 3-1]

		티. 레에세이 Bgl1에 대한 비	
효소	정제됨	CNPG	셀로비오스
티. 레에세이 Bgl1	Y	1	1
에이. 니케르 B-glu	Y	0.13	12
Mg3A	N	1.19	1.87

Mg3A는 트리카테르마 레에세이 Bgl1의 것보다 20% 더 높은 CNPG 활성을 갖고, 80% 초과로 더 높은 셀로비오스 가수분해 활성 (또는 셀로비아제 활성)을 가졌다. Mg3A는 아스페르길루스 니케르 B-glu의 것보다 CNPG에 대해 약 9배 더 높은 활성을 갖지만, 아스페르길루스 니케르 B-glu의 것보다는 셀로비오스에 대해 약 6배 더 낮은 활성을 가졌다. 셀로비오히드롤라제는 셀로비오스에 대한 어떤 활성도 갖지 않았다 (임의의 웰에 대해 어떤 글루코오스도 관찰되지 않았다).

[표 3-2]

효소	정제됨	CNPG/셀로비오스
티. 레에세이 Bgl1	Y	62
에이. 니케르 B-glu	Y	1.6
Mg3A	N	39

단백질 결정(protein determination)과 무관하게 각각의 분자의 활성을 비교하기 위해, 셀로비아제 활성에 대한 CNPG의 비를 계산하였다. Mg3A에 관한 CNPG/셀로비오스에 대한 가수분해 활성의 비는 트리카테르마 레에세이

Bgl1에 관한 CNPG/셀로비오스에 대한 가수분해 활성의 비의 절반보다 다소 적다. Mg3A에 관한 CNPG/셀로비오스에 대한 가수분해 활성의 비는 *아스페르길루스 니게르* B-glu에 관한 CNPG/셀로비오스에 대한 가수분해 활성의 비보다 20배 초과로 더 높다.

[0283] 실시예 4: PASC 기질에 대한 Mg3A 폴리펩티드의 개선된 가수분해 성능

[0284] 4-A. 공개된 국제 특허 출원 제WO 2011/038019호에 기재된 균주에 의해 생성된 백그라운드 전체 셀룰라제 조성물에서, PASC의 Mg3A 대 기준 *트리코테르마 레에세이* Bgl1 가수분해의 측정 및 비교를 도시하는 용량 곡선.

[0285] 베타-글루코시다제를 0-10 mg 단백질/g 셀룰로오스 내지 일정한 로딩 10 mg 단백질/g 글루칸 전체 셀룰라제 백그라운드 (공개된 국제 특허 출원 제WO 2011/038019호에 기재된 균주에 의해 생성됨) (이는 Fv43D, Fv3A, Fv51A, AfuXyn2, EG4 등을 발현한다)로 총 단백질 농도 88.8 g/L로 첨가하였다. 혼합물을 사용하여 인산 팽윤된 셀룰로오스 (PASC)를 가수분해하였다. 각각의 샘플 용량을 4벌로 검정하였다.

[0286] 모든 효소 희석은 50 mM 아세트산나트륨 완충제, pH 5.0로 이루어졌다. 150 μ L의 냉 0.6% PASC를 마이크로타이터 플레이트 (NUNC플랫 보텀(flat bottom) PS, 카탈로그 번호 269787)에서 30 μ L의 효소 용액에 첨가하였다. 따라서 효소 혼합물은 10 mg 단백질/g 글루칸의 전체 셀룰라제에 더하여 0-10 mg의 Mg3A 또는 Bgl1/g 글루칸을 함유하였다. 플레이트를 알루미늄 플레이트 시일(seal)로 커버하고, 이노바(Innova) 인큐베이터/진탕기에서 50 $^{\circ}$ C, 200 rpm에서 1.5시간 동안 인큐베이션하였다. 반응물을 100 μ L의 100 mM 글리신, pH 10으로 켄칭하고, 여과하고 (밀리포어 진공 필터 플레이트 카탈로그 번호 MAHVN45), 가용성 당을 아미넥스(Aminex) HPX-87P 칼럼을 구비한 애질런트 5042-1385 HPLC 상에서 측정하였다.

[0287] 글루칸 전환율을 (mg 글루코오스 + mg 셀로비오스 + mg 셀로트리오스) / mg 셀룰로오스 (반응물 중)로서 결정하였다.

[0288] 결과를 도 3a-3c에 나타냈다. 약 46% 및 62.5% 전환에서의 수평선은 바로 전체 셀룰라제 백그라운드 활성의 10 mg/g 및 20 mg/g 로딩을 나타낸다.

[0289] Mg3A는 동일 용량의 *티. 레에세이* Bgl1보다 글루코오스를 더 많이 생성하였다. 이는 *티. 레에세이* Bgl1보다 더 높은 셀로비아제 활성을 갖는 Mg3A와 일치하는 것이다. 글루코오스, 셀로비오스 및 셀로트리오스 농도를 함께 첨가하여 총 당을 측정하여 총 글루칸 전환율 (%)을 결정한 경우, Mg3A는 *티. 레에세이* Bgl1을 능가하였다.

[0290] 4-B. 4-A와 상이한 조건하에, 국제 공개 특허 출원 제WO 2011/038019호에 기재된 균주에 의해 생성된 백그라운드 전체 셀룰라제 조성물에서, PASC의 Mg3A 대 기준 *트리코테르마 레에세이* Bgl1 가수분해의 측정 및 비교를 도시하는 제2 세트의 용량 곡선.

[0291] 진탕 플라스크 생성된 Mg3A 배양 브로쓰를 10,000 분자량 컷 오프 PES 스핀 농축기(spin concentrator)를 사용하여 >20배로 농축하였다. *트리코테르마 레에세이* Bgl1 표준 곡선과 비교하여, 단백질 농도를 UPLC에 의해 결정하였다. 상기 4-A에서 보다 더 높은 단백질 로딩을 사용하여 PASC 가수분해 실험을 반복하였다. 농축된 샘플의 특이적 셀로비아제 활성은 원래 진탕 플라스크 상청액과 동등하였다. 상기 기재된 바와 같이 (1.5시간 동안 50 $^{\circ}$ C), PASC 검정을 수행하였다. 그 다음에 반응물을 pH 10에서 100 μ L의 100 mM 글리신 완충제로 켄칭하였다. 각각의 용량을 4벌로 검정하였다.

[0292] Mg3A의 농도는 UPLC를 사용하여 측정된 바, 1.55 g/L이었다. 기준 *트리코테르마 레에세이* Bgl1의 농도는 2.2 mg/L이었다.

[0293] 결과를 도 3d-3f에 나타냈다.

[0294] 결과는 Mg3A가 동일 용량의 *트리코테르마 레에세이* Bgl1이 생성하는 것보다 동일 PASC를 가수분해함으로써 글루코오스를 더 많이 생성하는 것을 확인해 주었다. Mg3A는 Bgl1보다 더 높은 셀로비아제 활성을 가지며, 이는 더 낮은 셀로비오스 농도, 동일 가수분해 조건하에 동일 PASC 기질로부터 산출된 것이다. 총 글루칸 전환율 (%)을 측정하기 위해, 총 농도의 글루코오스 및 셀로비오스를 첨가하여 총 당을 측정한 경우, Mg3A는 실질적으로 *티. 레에세이* Bgl1을 능가하였다.

[0295] 실시예 5: 희석 암모니아로 전처리된 옥수수 대 기질에 대한 Mg3A 폴리펩티드의 개선된 가수분해 성능.

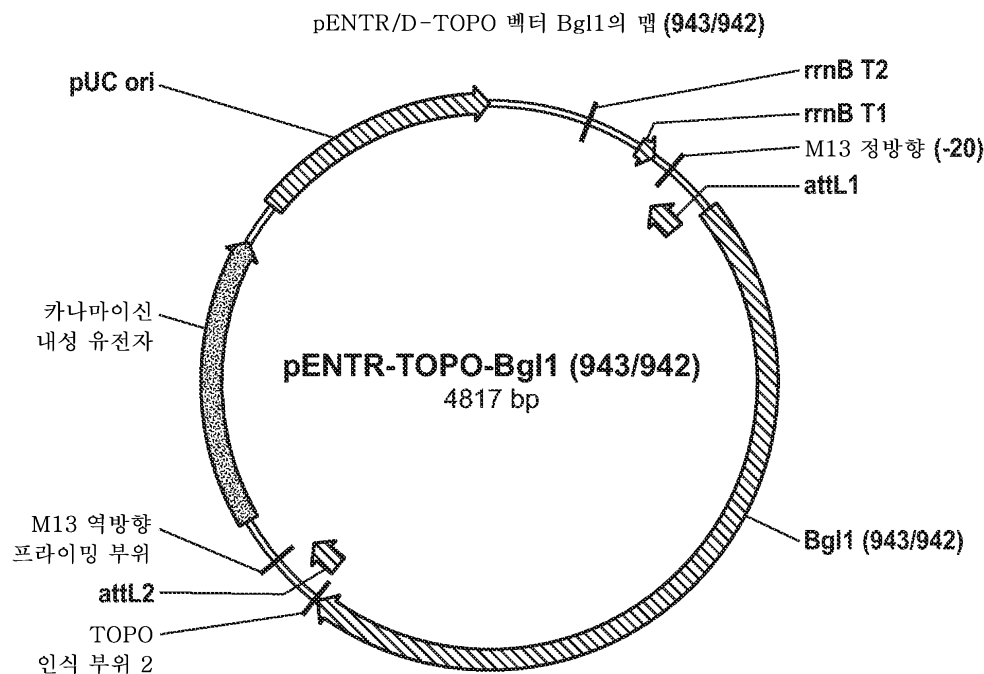
[0296] 5-A. 공개된 국제 특허 출원 제WO 2011/038019호에 기재된 균주에 의해 생성된 백그라운드 전체 셀룰라제 조성물에서, DACS의 Mg3A 대 기준 *트리코테르마 레에세이* Bgl1 가수분해의 측정 및 비교를 도시하는 용량 곡선.

- [0297] 진탕 플라스크 생성된 Mg3A 배양 브로쓰를 10,000 분자량 컷 오프 PES 스핀 농축기를 사용하여 >20배로 농축하였다. *트리코테르마 레에세이* Bgl1 표준 곡선과 비교하여, 단백질 농도를 UPLC에 의해 결정하였다. 농축된 진탕 플라스크 상청액을 기질로서 회석 암모니아로 전처리된 옥수수 대 (DACS)를 사용하여 당화 검정에서 사용하였다. 각각의 효소 혼합물 샘플을 10% Mg3A 또는 Bgl1, 국제 공개 특허 출원 제WO 2011/038019호에 기재된 바와 같은 조작된 *트리코테르마 레에세이* 균주에 의해 생성된 전체 셀룰라제와 블렌딩하였다. DACS 기질을 가수분해하기 위한 용량 반응 곡선은 3-53 mg 총 단백질/g 글루칸 효소 혼합물을 기질에 첨가함으로써 생성시켰다.
- [0298] 정제된 스탁 2.2 g/L 총 단백질로부터 *트리코테르마 레에세이* Bgl1을 가수분해 검정에 첨가하였다. 1.55 mg/mL 농축된 샘플로부터 Mg3A를 첨가하였다.
- [0299] 회석 암모니아로 전처리된 옥수수 대 (DACS)를 20 mM 아세트산나트륨, pH 5에서 슬러리화하여 최종 7% 글루칸 (21.5% 고형물) 함량을 수득하였다. 필요한 경우, 슬러리를 pH 5로 조정하고 슬러리를 96-웰 마이크로타이터 플레이트에 옮겼다.
- [0300] 모든 효소를 mg 단백질/g 글루칸 (기질 중)을 기준으로 로딩하였다. 모든 효소 회석은 50 mM 아세트산나트륨 완충제, pH 5.0로 이루어졌다. 30 μ L의 효소 용액 및 45 μ g DACS 기질을 96-웰 마이크로타이터 플레이트에 웰당 첨가하였다. 각각의 샘플 용량을 4별로 시험하였다. 플레이트를 알루미늄 시일로 커버하고, 이노바 인큐베이터/진탕기에서 50°C, 200 rpm에서 2일 동안 인큐베이션하였다. 반응물을 100 μ L의 100 mM 글리신, pH 10로 켄칭하고, 여과하고, 가용성 당을 HPLC에 의해 측정하였다.
- [0301] 글루칸 전환율 (%)은 (mg 글루코오스 + mg 셀로비오스 + mg 셀로트리오스) / mg 셀룰로오스 (DACS 기질 중)로서 정의된다.
- [0302] 용량 반응 곡선을 도 4a-4b에 나타냈다.
- [0303] 기준 *트리코테르마 레에세이* Bgl1을 동일한 전체 셀룰라제 백그라운드 조성물과 혼합한 경우보다, 또는 셀룰라제 백그라운드 조성물을 단독으로 사용하여 글루칸 전환 수준을 측정한 경우보다 Mg3A와 전체 셀룰라제 백그라운드 조성물을 혼합하는 것이 더 높은 수준의 글루칸 전환을 제공하였다. 베타-글루코시다제를 13 mg/g 미만으로 투여한 경우 개선이 특히 현저하였다.
- [0304] **5-B: DACS의 Mg3A 대 기준 *트리코테르마 레에세이* Bgl1 가수분해의 측정 및 비교를 도시하는 용량 곡선 (여기서 Mg3A 및 Bgl1은 국제 공개 특허 출원 제WO 2011/038019호에 기재된 균주에 의해 생성된 13.4 mg/g에서의 전체 셀룰라제 조성물에 증가하는 용량으로 첨가됨).**
- [0305] 본 실험에서, 베타-글루코시다제를 일정한 로딩의 13.4 mg 단백질/g 글루칸의 전체 셀룰라제 (국제 공개 특허 출원 제WO 2011/038019호에 따라서 조작된 *트리코테르마 레에세이* 균주에 의해 생성됨)에 증가하는 용량으로 첨가하였다. 혼합물을 사용하여 50°C에서 2일 동안 DACS (4% 글루칸)를 가수분해하였다. 혼합물을 제조하기 위해, *트리코테르마 레에세이* Bgl1을 2.2 g의 정제된 스탁/L 총 단백질로부터 혼합물에 첨가하고, 전체 셀룰라제를 88.8 g/L 총 단백질 스탁으로부터 혼합물에 첨가하였다. 1.55 mg/mL 농축물로부터 Mg3A를 첨가하였다.
- [0306] 마이크로타이터 플레이트에서 회석 암모니아로 전처리된 옥수수 대를 상기 기재된 바와 같이 제조하였다. 모든 효소를 mg 단백질/g 글루칸 (기질 중)을 기준으로 로딩하였다. 모든 효소 회석은 50 mM 아세트산나트륨 완충제, pH 5.0로 이루어졌다. 30 μ L의 효소 용액을 마이크로타이터 플레이트에서 웰당 45 μ g 기질에 첨가하였다. 플레이트를 호일 시일(foil seal)로 커버하고, 이노바 인큐베이터/진탕기에서 50°C, 200 rpm에서 2일 동안 인큐베이션하였다. 반응물을 100 μ L의 100 mM 글리신, pH 10으로 켄칭하고, 여과하고, HPLC (탈회(de-ashing) 칼럼 (바이오래드(Biorad) 125-0118) 및 탄수화물 칼럼 (아미넥스 HPX-87P)이 장착된 애질런트 100 시리즈)에 의해 가용성 당을 측정하였다. 이동상은 유량 0.6 mL/분 및 20분 시행 시간을 갖는 물이었다. 글루코오스 표준 곡선을 생성시키고 사용하여 정량화하였다.
- [0307] 글루칸 전환율 (%)은 (mg 글루코오스 + mg 셀로비오스 + mg 셀로트리오스) / mg 셀룰로오스 (기질 중)로서 정의된다.
- [0308] 결과를 도 5a-5b에 나타냈다.
- [0309] Mg3A는 모든 용량에서 *트리코테르마 레에세이* Bgl1을 능가하였다.
- [0310] **실시예 6: 회석 암모니아로 전처리된 옥수수 대 기질에 대한 Mg3A 폴리펩티드의 개선된 열안정성 및 고온 가수분해 성능**

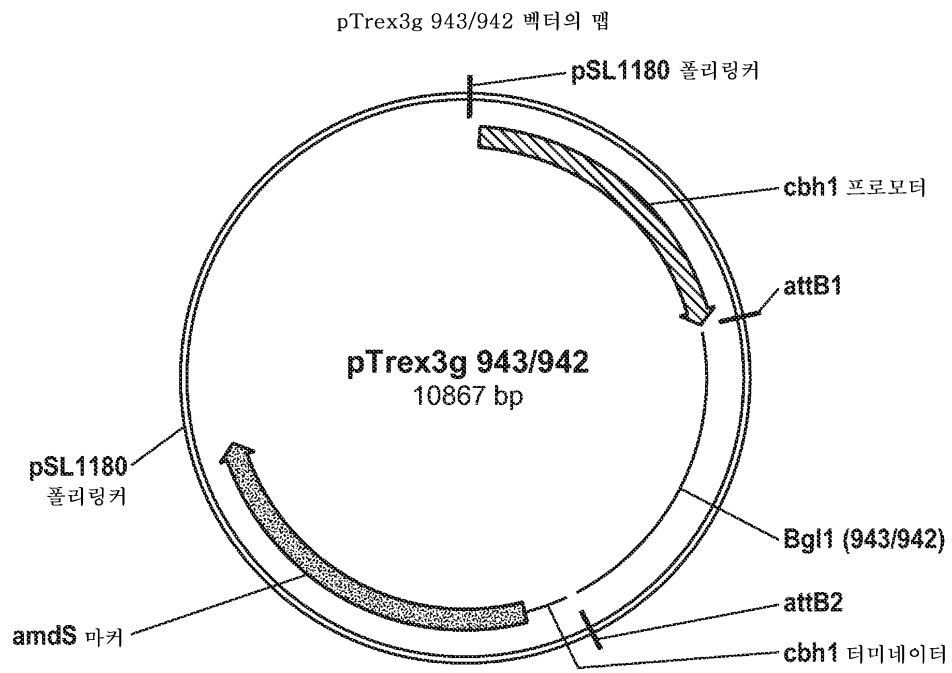
- [0311] 본 실험에서, 베타-글루코시다제를 일정한 로딩의 10 mg 단백질/g 글루칸의 전체 셀룰라제 (국제 공개 특허 출원 제WO 2011/038019호에 따라서 조작된 *트리코테르마 레에세이* 균주에 의해 생성됨)에 증가하는 용량으로 첨가하였다. 혼합물을 사용하여 55℃에서 2일 동안 DACS (4% 글루칸)를 가수분해하였다. 혼합물을 제조하기 위해, *트리코테르마 레에세이* Bgl1을 정제된 스톡 2.2 g/L 총 단백질로부터 혼합물에 첨가하고, 전체 셀룰라제를 88.8 g/L 총 단백질 스톡으로부터 혼합물에 첨가하였다. 상기 기재에 따라 제조된 1.55 mg/mL 농축물로부터 Mg3A를 첨가하였다.
- [0312] 마이크로타이터 플레이트에서 희석 암모니아로 전처리된 옥수수 대를 상기 기재된 바와 같이 제조하였다. 모든 효소를 mg 단백질/g 글루칸 (기질 중)을 기준으로 로딩하였다. 모든 효소 희석은 50 mM 아세트산나트륨 완충제, pH 5.0로 이루어졌다. 30 μ L의 효소 용액을 마이크로타이터 플레이트에서 웰당 75 μ g 기질에 첨가하였다. 플레이트를 호일 시일로 커버하고, 이노바 인큐베이터/진탕기에서 55℃, 200 rpm에서 2일 동안 인큐베이션하였다. 반응물을 100 μ L의 100 mM 글리신, pH 10으로 켄칭하고, 여과하고, HPLC (탈회 칼럼 (바이오래드 125-0118) 및 탄수화물 칼럼 (아미넥스 HPX-87P)이 장착된 애질런트 100 시리즈))에 의해 가용성 당을 측정하였다. 이동상은 유량 0.6 mL/분 및 20분 시행 시간을 갖는 물이었다. 글루코오스 표준 곡선을 생성시키고 사용하여 정량화하였다.
- [0313] 글루칸 전환율 (%)은 (mg 글루코오스 + mg 셀로비오스 + mg 셀로트리오스) / mg 셀룰로오스 (기질 중)로서 정의된다.
- [0314] 결과를 도 6a-5c에 나타냈다.
- [0315] Mg3A는 글루칸 전환, 글루코오스 생성에 관한 모든 용량에서 *트리코테르마 레에세이* Bgl1을 능가하였다. Mg3A 및 Bgl1 둘 다는 당화 혼합물에서 셀로비오스를 효과적으로 고갈시켰다.

도면

도면1



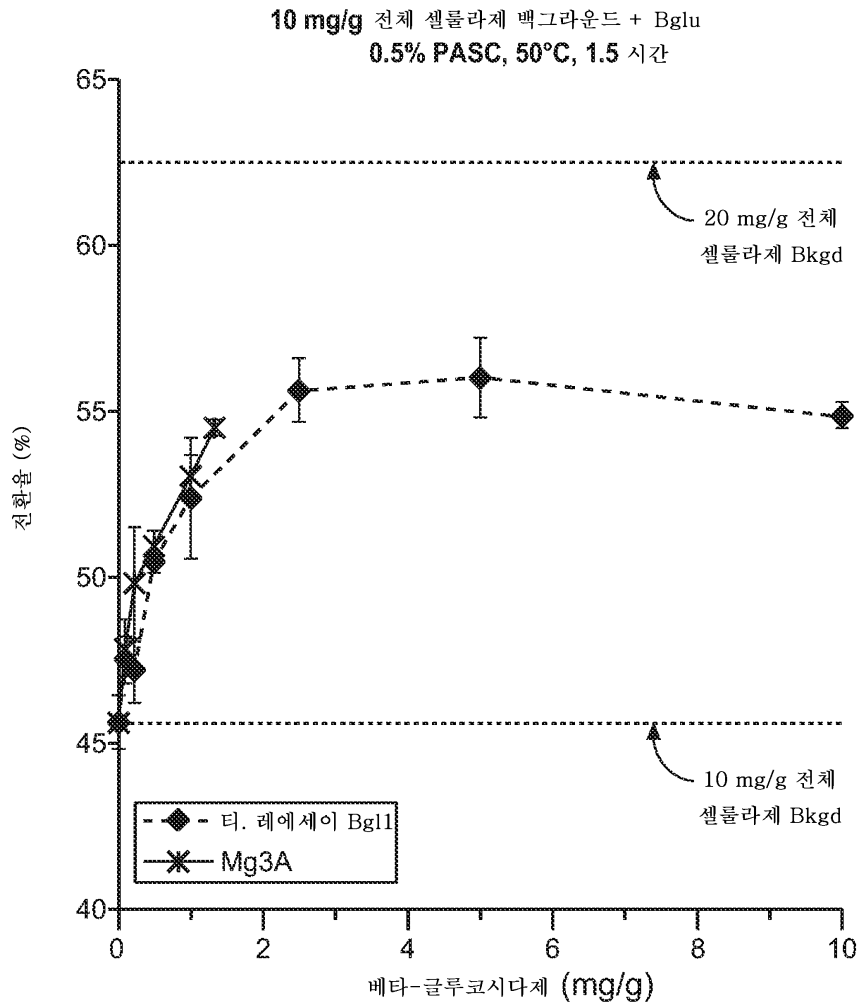
도면2



도면3a

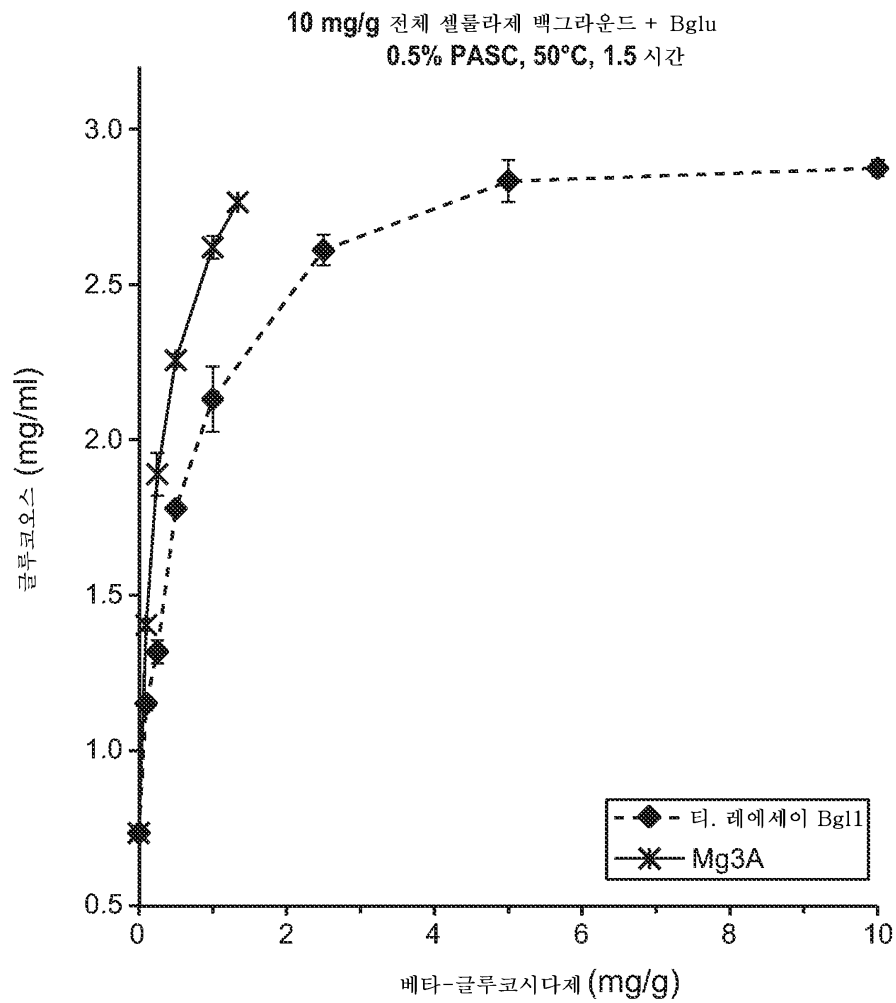
도 3: PASC (인산 팽윤된 셀룰로오스)의 가수분해를
측정하고 비교하는 용량 곡선

3a: 동일한 농도의 티. 레에세이 Bgl1 대 Mg3A를 포함하는 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된) 전체 셀룰라제 조성물에 의해, 소정의 PASC 기질로부터 총 글루칸 전환율 (%)의 측정 및 비교를 도시하는 용량 곡선. 실시예 4-A의 조건에 따라서 측정을 수행하였다.



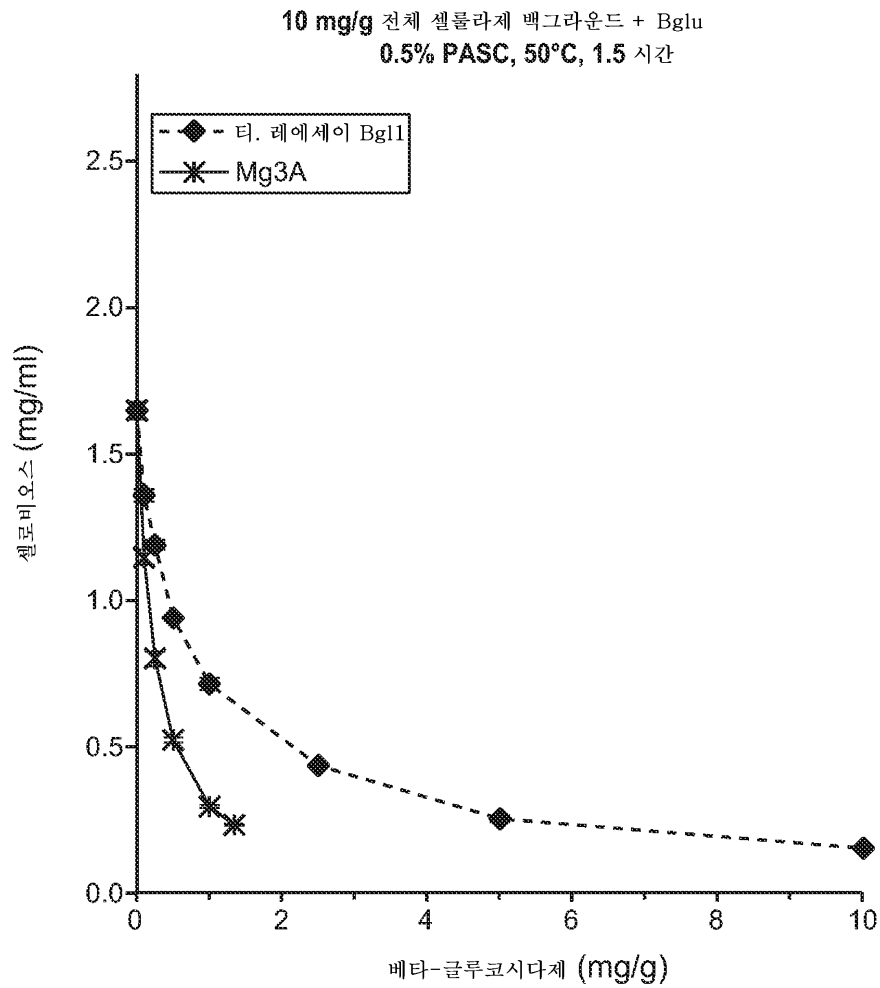
도면3b

3b: 동일한 농도의 티. 레에세이 Bgl1 대 Mg3A를 포함하는 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된) 전체 셀룰라제 조성물에 의해 소정의 PASC 기질의 가수분해로부터 글루코오스 수율의 측정 및 비교를 도시하는 용량 곡선. 실시예 4-A의 조건에 따라서 측정을 수행하였다.



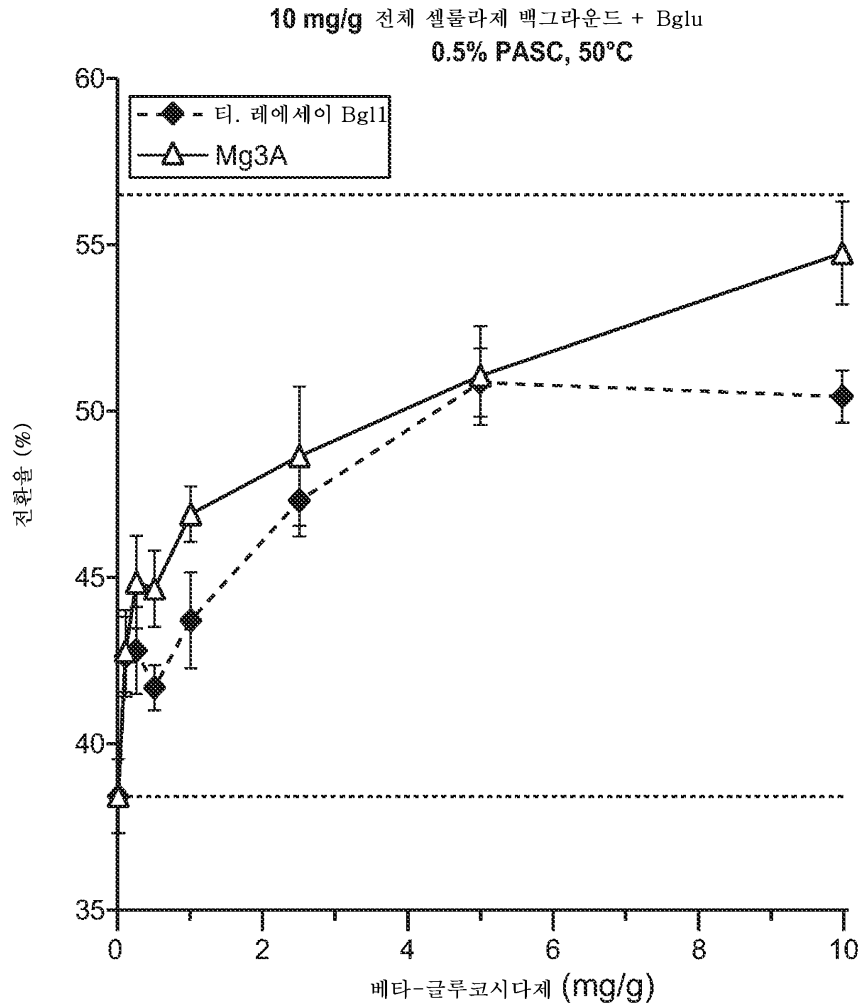
도면3c

3c: 동일한 농도의 티. 레에세이 Bgl1 대 Mg3A를 포함하는 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된) 전체 셀룰라제 조성물에 의해 소정의 PASC 기질의 가수분해로부터 셀로비오스 수율의 측정 및 비교를 도시하는 용량 곡선. 실시예 4-A의 조건에 따라서 측정을 수행하였다.



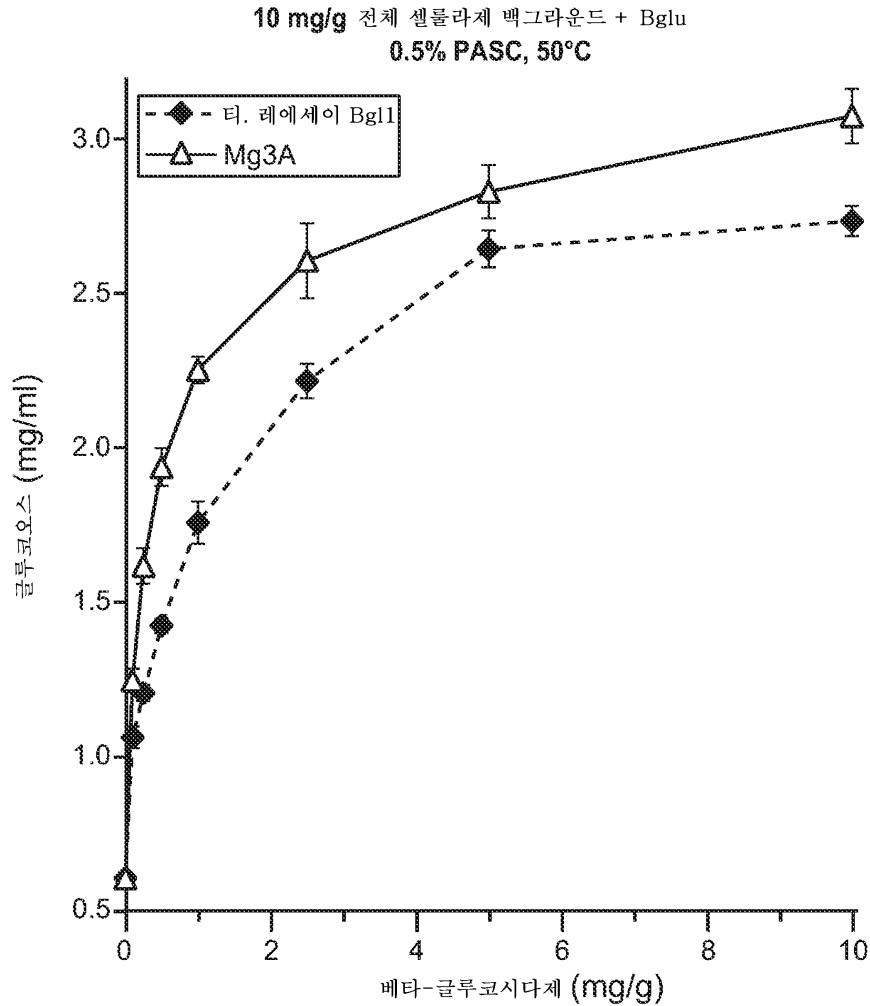
도면3d

3d: 동일한 농도의 티. 레에세이 Bgl1 대 Mg3A를 포함하는 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된) 전체 셀룰라제 조성물에 의해, 소정의 PASC 기질로부터 총 글루칸 전환율 (%)의 측정 및 비교를 도시하는 용량 곡선. 실시예 4-B의 조건에 따라서 측정을 수행하였다.



도면3e

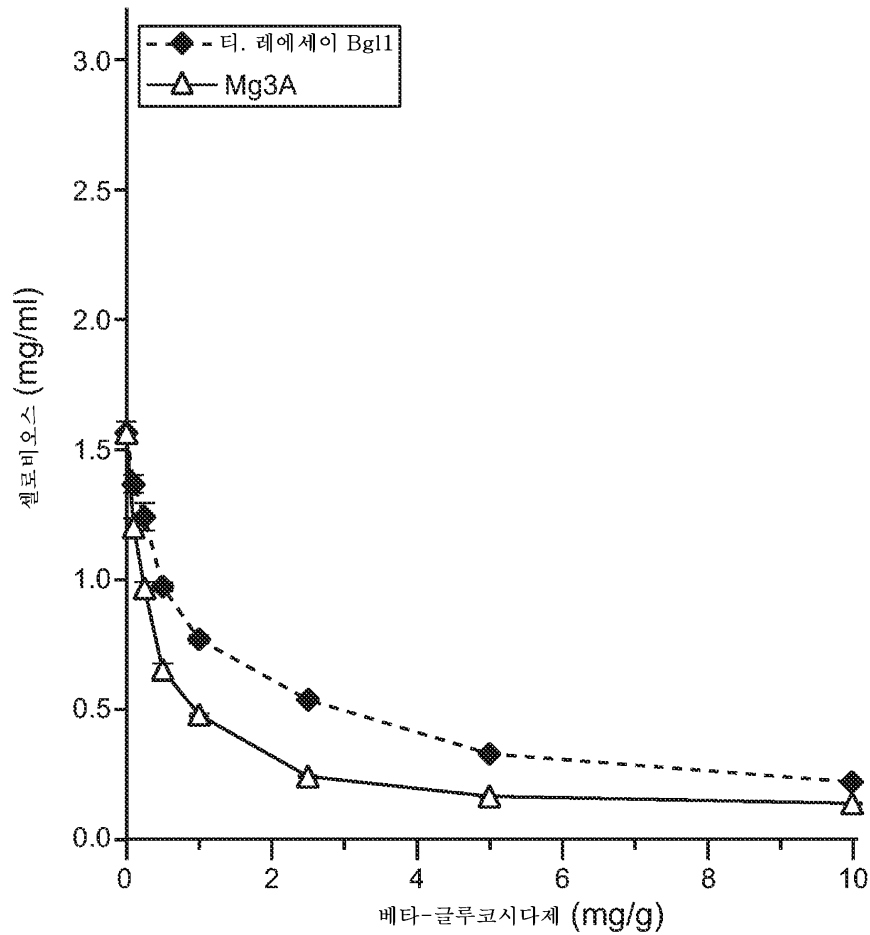
3e: 동일한 농도의 티. 레에세이 Bgl1 대 Mg3A를 포함하는 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된) 전체 셀룰라제 조성물에 의해 소정의 PASC 기질의 가수분해로부터 글루코오스 수율의 측정 및 비교를 도시하는 용량 곡선. 실시예 4-B의 조건에 따라서 측정을 수행하였다.



도면3f

3f: 동일한 농도의 티. 레에세이 Bgl1 대 Mg3A를 포함하는 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성된) 전체 셀룰라제 조성물에 의해 소정의 PASC 기질의 가수분해로부터 셀로비오스 수율의 측정 및 비교를 도출하는 용량 곡선. 실시예 4-B의 조건에 따라서 측정을 수행하였다.

10 mg/g 전체 셀룰라제 백그라운드 + Bglu
0.5% PASC, 50°C

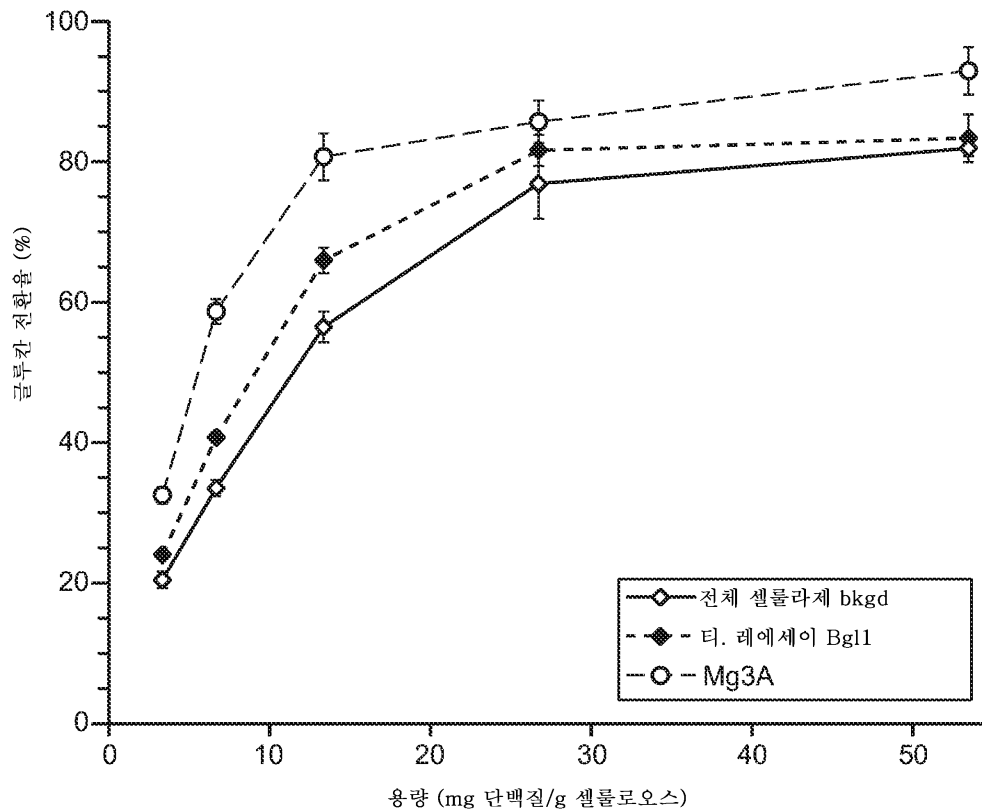


도면4a

도 4: 회석 암모니아로 전처리된 옥수수 대 (DACS)의 가수분해를
측정하고 비교하는 용량 곡선

4a: 동일한 농도의 티. 레에세이 Bgl1 대 Mg3A를 포함하는 (국제 공개 특허 출원
WO 2011/038019에 기재된 것에 따라서 조작된 트리코테르마 레에세이 균주로부터
생성된) 전체 셀룰라제 조성물에 의해, 소정의 DACS 기질로부터 총 글루칸 전환율의 측정
및 비교를 도시하는 용량 곡선.

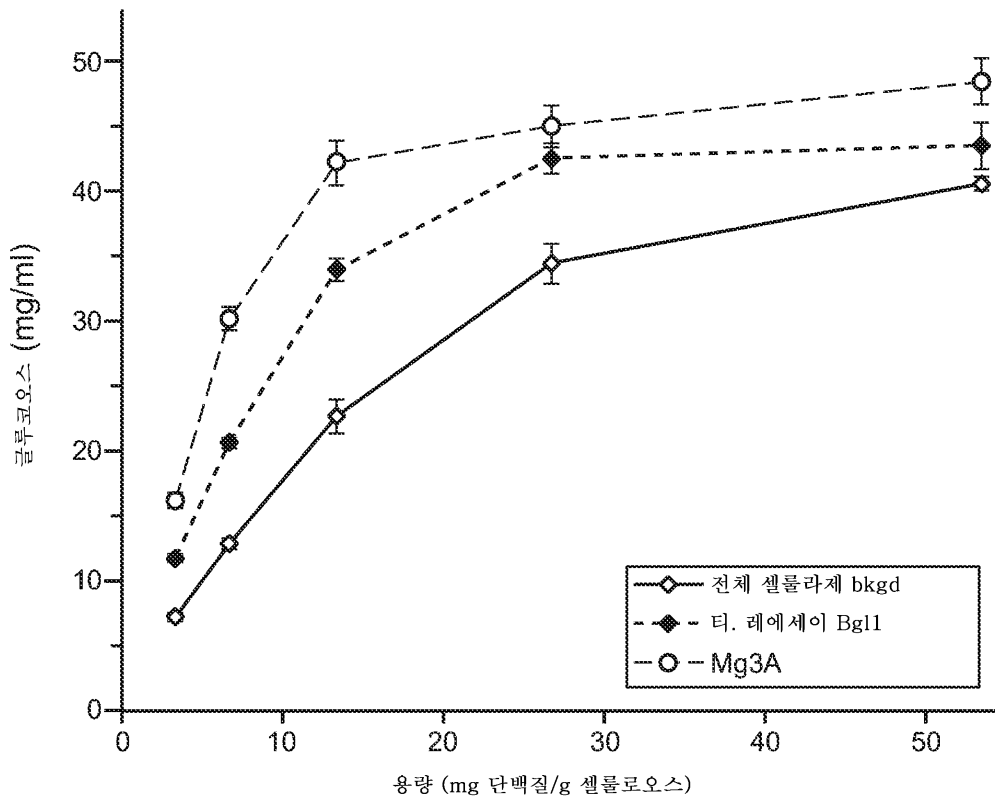
용량 반응: 전체 셀룰라제 백그라운드 +/- 10% Bglu
DACS, 4% 셀룰로오스, MTP, 50°C, 2일



도면4b

4b: 동일한 농도의 티. 레에세이 Bgl1 대 Mg3A를 포함하는 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 트리코테르마 레에세이 균주로부터 생성된) 전체 셀룰라제 조성물에 의해 소정의 DACS 기질의 가수분해로부터 글루코오스 수율의 측정 및 비교를 도시하는 용량 곡선.

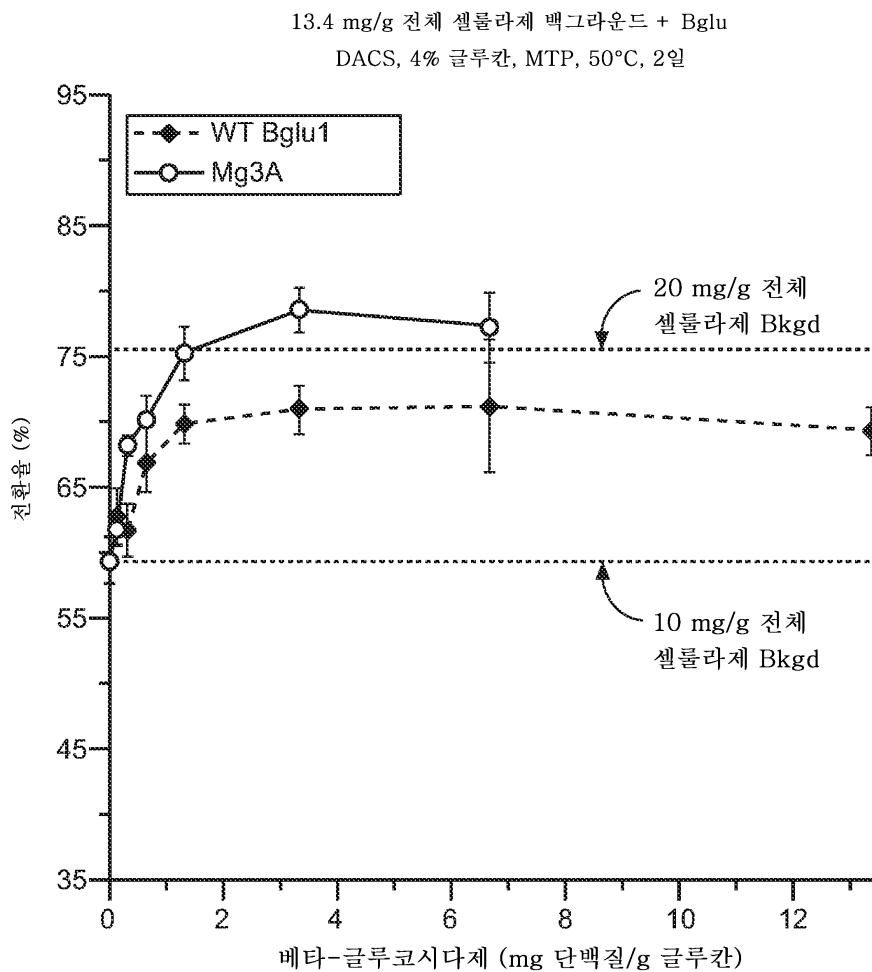
용량 반응: 전체 셀룰라제 백그라운드 +/- 10% Bglu
DACS, 4% 셀룰로오스, MTP, 50°C, 2일



도면5a

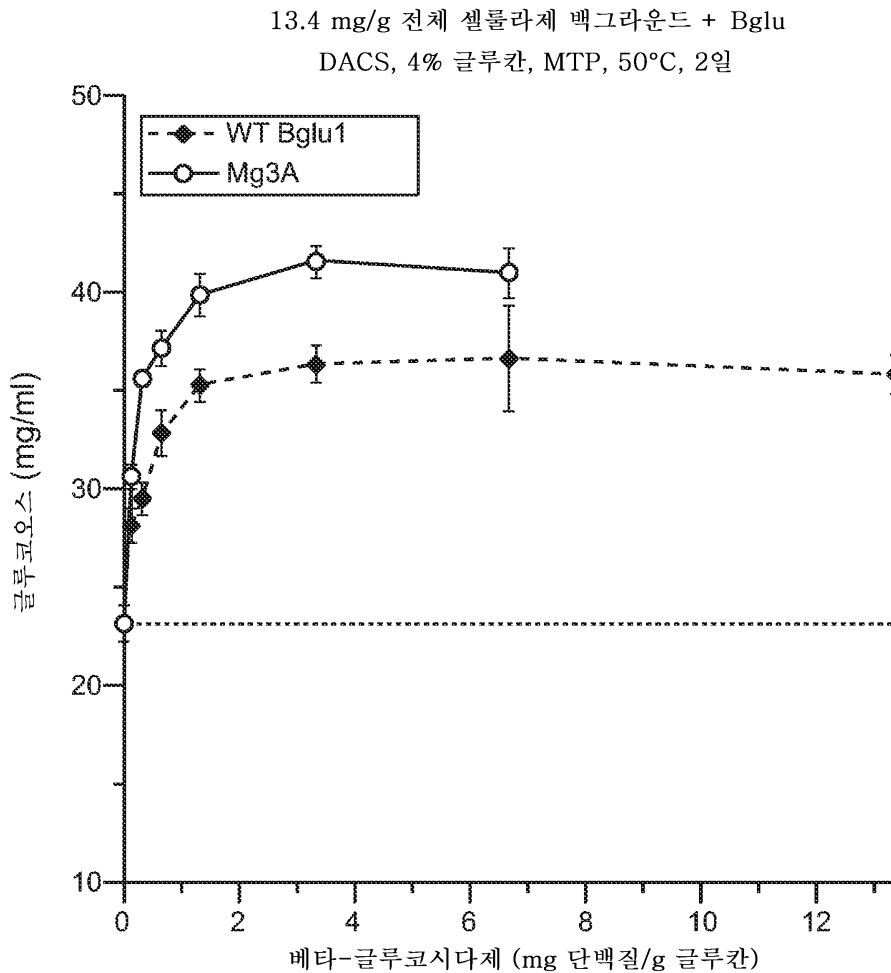
도 5: 상이한 용량의 Mg3A 및 *티. 레에세이* Bgl1을 13.4 mg/g 총 단백질로 전체 셀룰라제에 첨가하여, DACS 기질에 대한 가수분해 성능의 측정 및 비교를 도시하는 용량 곡선.

5a: 13.4 mg/g 전체 셀룰라제 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성됨)에 증가하는 용량으로 첨가된, Mg3A 대 *티. 레에세이* Bgl1의 혼합물에 의해, 소정의 DACS 기질로부터 총 글루칸 전환의 측정 및 비교를 도시하는 용량 곡선.



도면5b

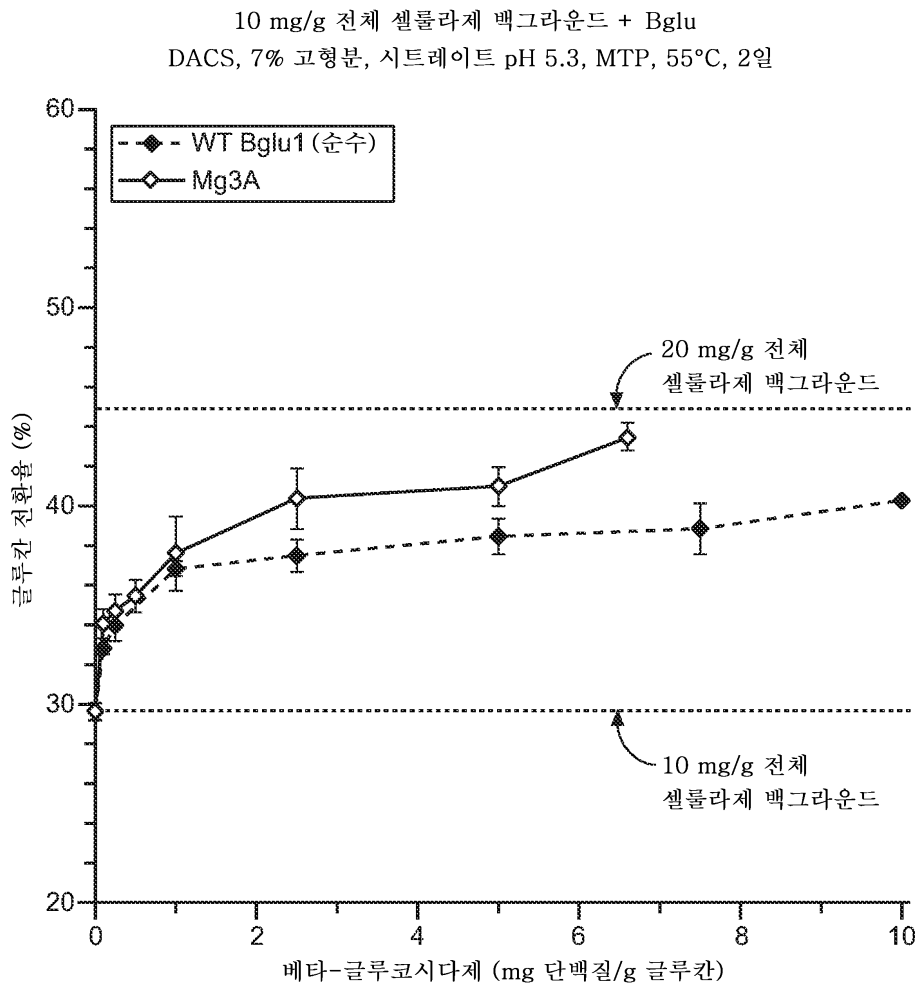
5b: 13.4 mg/g 전체 셀룰라제 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성됨)에 증가하는 용량으로 첨가된, Mg3A 대 티. 레에세이 Bgl1의 혼합물에 의해, 소정의 DACS 기질로부터 글루코오스 수율의 측정 및 비교를 도시하는 용량 곡선.



도면6a

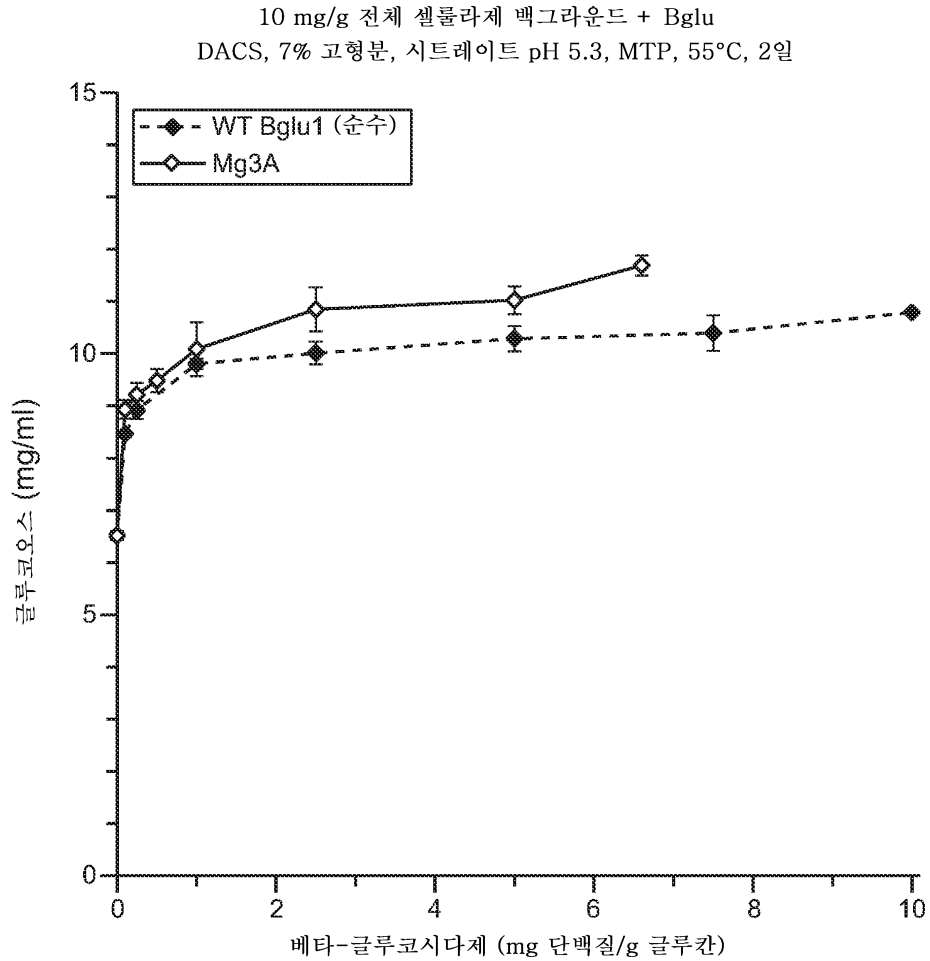
도 6: 상이한 용량의 Mg3A 및 *티. 레에세이* Bgl1을 10.0 mg/g 총 단백질로 전체 셀룰라제에 첨가하여, 55°C에서 DACS 기질에 대한 가수분해 성능의 측정 및 비교를 도시하는 용량 곡선.

6a: 10 mg/g 전체 셀룰라제 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성됨)에 증가하는 용량으로 첨가된, Mg3A 대 *티. 레에세이* Bgl1의 혼합물에 의해, 소정의 DACS 기질로부터 총 글루칸 전환의 측정 및 비교를 도시하는 용량 곡선.



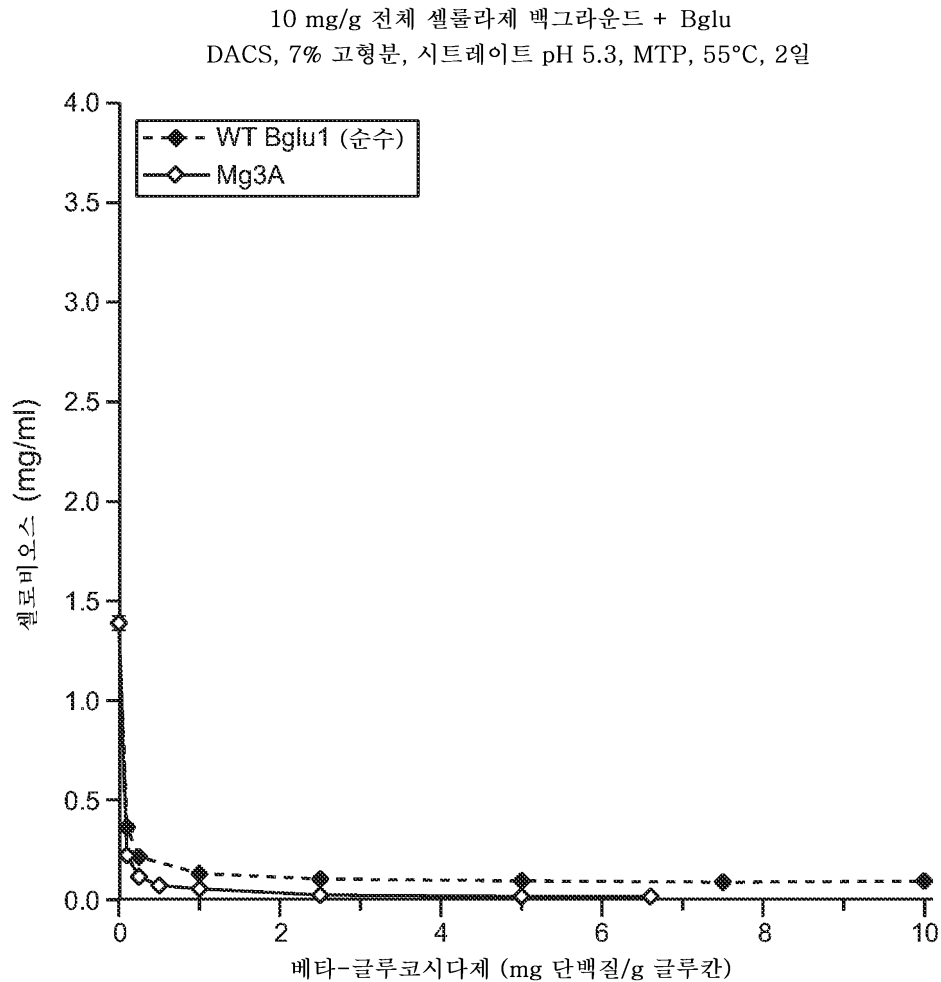
도면6b

6b: 10 mg/g 전체 셀룰라제 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성됨)에 증가하는 용량으로 첨가된, Mg3A 대 티. 레에세이 Bgl1의 혼합물에 의해, 소정의 DACS 기질로부터 글루코오스 수율의 측정 및 비교를 도시하는 용량 곡선.

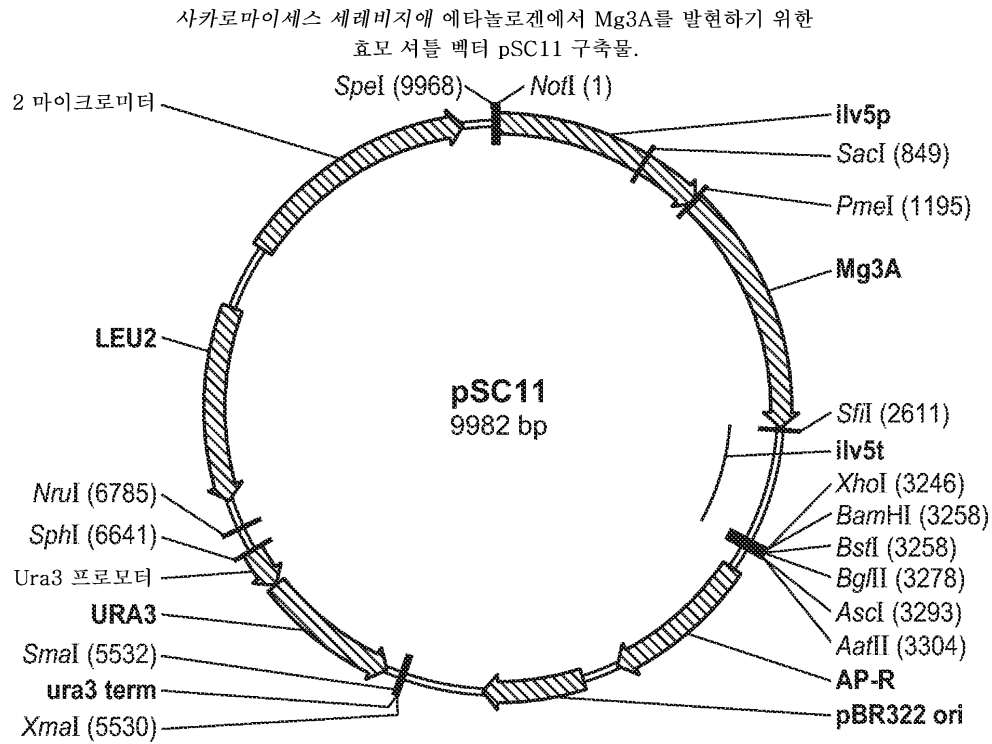


도면6c

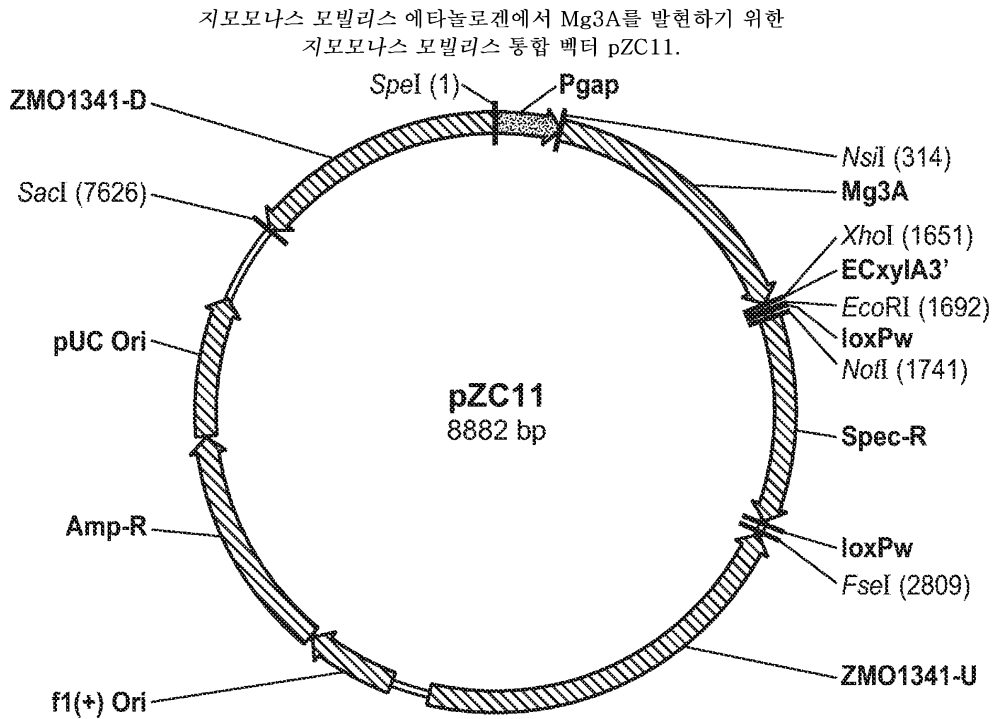
6c: 10 mg/g 전체 셀룰라제 (국제 공개 특허 출원 WO 2011/038019에 기재된 것에 따라서 조작된 *트리코테르마 레에세이* 균주로부터 생성됨)에 증가하는 용량으로 첨가된, Mg3A 대 *티. 레에세이* Bgl1의 혼합물에 의해, DACS 기질 상에서 소정의 당화 반응으로부터 셀로비오스 고갈의 측정 및 비교를 도시하는 용량 곡선.



도면7



도면8



도면9a

서열 번호 1: Mg3A를 암호화하는 해산 서열

atgcggtttctccgggatcgctgcgcacgcttgtggcgggtgccggcggtctgtcccaccctggtgacta
ctcaaagggtatgataaaaaataaaaaaaggctctcaagcctgacgactcgggaaccttttactt
tgtcccttgtttttgcgcgccgcatgctcagttttgttgacactcctttaggagaggctattgcctt
ttggtcaggttgtggtaggccttgttttgcctattctgtgacctctccacacaattctgaagcaat
tggcttaaatacgagggcaaagtccctctgtcgcgttgcgttgcagaataccctcgggatcgctca
tgtttctggtttcccatcatggcttccatcaatacaatagtgcatactatacatagcttgagcgtcg
ggcgtggccacatcagaacccactaccctcagccgtggatgaaccagatgcagacgggtggcagg
aggcctatgtcaaagcaaaggactttgtctcgcagatgacctgcttgagaaggtaacttgacaacc
gggtgtaggtaggtcaacaaacccctttttttctgcatgacccctgaatggcaaccggcaacccc
ctcccccaaaacacccgtagcgaatattgacactttttggtactggtgtttgtagctgggcaagcga
tctctgcgttggcaacgttgggtccgtccccgcctgggtctccgaagcctctgtctgcaggactctc
ccaccggtgtccgttttcgcgattgggtctcgggtctttcctgctggcatcaccaccggcgccacgttc
gataagggtctgtatgaccgcgttgacaagctatgggccaagaagccaaggacaaggcgcacacgt
cctgcttgggtccgttgcgtggaggacttggcgtgtgggtgctggtaagcctttgctcgatacactgg
ctgatgacaacatagagagatgtctttttctcacagtaaccgacatcataggtggacgagcttgg
gagcttttcggagctgacccagctcttaccggataggtatgattgagacaataaagggtatccagga
cactggtgtcattgtcactgccaaagcatttcatcggaacgagcaaggtaagtctatcggttgacctc
tttgccagctcgagatgcctgggtgcaagcttgcgtggtttccaggtccatttactaataccaatgg
catccaacagagcatttccgtcaggtaggcgaagagcgaggccgtggcgtcaacatcagcgagtcgct
ctcctccaacattgatgacaagactatgcacgagttgtacctctggttaagtccaacactgcattcagg
ggatggattggattagtgctgggtgacctgaccataataggccctttgcggatgcagttcgcgcagg
tgtcggctcgggtatgtgctcgtacacacaggtcaacaactcgtacggttgccagaacagcaagcttc
tcaacgggttgcgaaggacgagcttggcttccagggaattcgtcatgtctgactggcaggtcagcac
actggagcgccagcgctgcggccggtcttgatatgtccatggtcagtcgaagtcctcaagttctatt
ctagttcatgacaggcaacgatgctaacgtttctcgaacaaaagccgggtgacaccgagttcaacaca
ggctcttagtttctggggagccaacttgacctggctgtcgtcaacggaaactgttgccgaatggcgtat
cgacgacatggccatgcgaatcatggctgccttctcaagggttggcaacacgctggaccagccgaga
tcaacttttagctcttgaccaaggacaccttcgggtcccttccactcgtcatctgaaacagaatccag
cagatcaaccagcacgtcgatgtccgtcgcgaccacggcaacttgattcgcgaggttgccgccaaggg
aacctgtcttctcaagaacaccaacaacgcctcccgctcaacaagcccaagttcctggccgtcattg
gtgacgatgccggtcacaacccgcggtcccaacggctgcccggaaccgcggtgccttttggaact
ttgggcatggcctgggttctggcacggtgacttccctacctcatcaccgccgacgcggtctcca
agctcaggccatcgaggagggcacacgctacgagagcatcctgagcaactatgccacggcacagactc
aagctcttgcagccaaacctacgcaaccgcatgtctttgttgcgtctcttcgggagaaggctgt
gagctcgcctttctgtttcagctgaacgcacctgtttgaagcaacaccccttactgacaattctt
ctcttccaccagacatcgactttgacggaacaaaggagaccgcaacaacctgacctctggtatga
cggcgactcgtggtcaagaatgtcagctccgtgtgcaacaacacctcgtcgtcatccactccaccg
gcccgaccatcctgaccgagtggtacgacaaccccaacgtcaccgcatcgtctgggaggtgtccg
ggacagaggtccggcgtgcatcaccgacgtcctgtacggccggtcaaccccgccggccgctcgc
cttccactggggcaagaccgcgagagctacggcaccgacgtcatgtacaagcccaacaacggcaacg
aggccccgcagcaggactacaccgagggcgtcttcatcgaactaccgccactttgaccagcagaaggac
gagcccgctacgagttcggccacggcttgagctacaccacctttgagtaactccaacatccgcgtcga
caaggccccggcagcagtagtaaacgccaccacccggccagaccatccggcccccgcttttgccgcca
acgtctcaaggacgtgtccagtagactttcccgctccgacgagttcccgcacatctacctcttcatc
taccgtagctcaacaccagctctctcgggtgaggaggcctcgcgcgacccaagtacggcggaacggc

도면9b

cgaggagttcctgccgcccaggcgctcgatggctcgccccagcccttgccccgcgcgtcgggcaaga
actcgcccgccggcaaccgccaagctgtacgacacccctgtacaccgtcacggcaaccatcaccaacacc
ggcaagctcgtggcggaagaggtgccgcagttgtacgtgtcgacggcgccccgaggacccgcccgt
cgtcctgcgcggccttcgagcgcacccgtctcgaccccgccagagcgccaccttcaaggtcgacctga
ccaggagggacgtcagcaactgggacgtcaaggtccaggactgggtcatctctgagcaccccaagaag
gtgtttgtcggaagcagcagcaggaagctccatctgagtgctgacctgaactga

서열 번호 2: Mg3A의 폴리펩티드 서열 (밑줄 그어진 잔기는 예측된 신호 서열임)

mrfsqivatlvagaqvsahpgdysklerravatsephypqpwmnpdadgwqeyvkakdfvsqmtlle
kvnlttgvgwasdlcvngvavprlglrlslclqdsptgvrfadwvsvfpagittgatfdkglmyrrgq
amgqeakdkginvllgpgvaggglgrvaaggrawesfgadpvltygymietikgiqdtgviatakhfign
eqehfrqvgeergrgvniseslssniddktmhelylwpfadavragvgsvmcsytqvnsygcqnskl
lngllkdelgfggfvmwdwqaqhtgaasaaagldmsmpgdtefntglsfwganltlavngtvaewri
ddmamrimaaffkvngtldqpeinfsswtkdtfgplhsssgnriqqinqhvdvrrdhgnlirevaakg
tvllkntnnalplnkpkflavigddagsnprgngcpdrgcllgtlgmawgsqtadfpylitpdaalq
aqaidgtryesilsnyataqtgalvsqtyataivfvaassgegyidfdgnkgdrnnltlwydgdsly
knvssvcnntivvihstgptiltewydnpnvtaivwagvpggesgraitdvlygrvnpagrspftwgk
tresygtvmykpnngneapqdytegvfidyrhfdqgkdepvyefghglstyttfeysnirvdkapas
eykpttggtipapvfganvskdlsqytfpsdefphiylfiypylntsssggeasrdpkyggtaeeflp
pkaldgspqplprasgknspggnrqlydtlytvtatitntgklvgeevpqlyvshggpedppvvlrgf
erirldpgqsatfkvdltrrdvsnwdvkvqdwvisehpkkvfvgsssrklhlsadln

서열 번호 3: 성숙 Mg3A 폴리펩티드 서열

hpgdysklerravatsephypqpwmnpdadgwqeyvkakdfvsqmtllekvnlttgvgwasdlcvng
vgavprlglrlslclqdsptgvrfadwvsvfpagittgatfdkglmyrrgqamgqeakdkginvllgpgv
aggglgrvaaggrawesfgadpvltygymietikgiqdtgviatakhfigneqehfrqvgeergrgvni
seslssniddktmhelylwpfadavragvgsvmcsytqvnsygcqnsklngllkdelgfggfvmwd
wqaqhtgaasaaagldmsmpgdtefntglsfwganltlavngtvaewriddmamrimaaffkvngtld
qpeinfsswtkdtfgplhsssgnriqqinqhvdvrrdhgnlirevaakgtvllkntnnalplnkpkf
lavigddagsnprgngcpdrgcllgtlgmawgsqtadfpylitpdaalqaaidgtryesilsnya
taqtgalvsqtyataivfvaassgegyidfdgnkgdrnnltlwydgdslyknvssvcnntivvihstg
ptiltewydnpnvtaivwagvpggesgraitdvlygrvnpagrspftwgktresygtvmykpnngne
apqdytegvfidyrhfdqgkdepvyefghglstyttfeysnirvdkapasaykpttggtipapvfgan
vskdlsqytfpsdefphiylfiypylntsssggeasrdpkyggtaeeflppkaldgspqplprasgkn
spggnrqlydtlytvtatitntgklvgeevpqlyvshggpedppvvlrgferirldpgqsatfkvdltr
rdvsnwdvkvqdwvisehpkkvfvgsssrklhlsadln

도면9c

서열 번호 4: 티. 레에세이 BglI 폴리펩티드 서열 (예측된 신호 서열 잔기는 밑줄 그어짐)

mr yrtaaalalatqpfaradshstsgasaeavvppagtpwgtaydkakaalaklnlqdkvgivsgvwnggpcvg
ntspaskisypslclqdgplgvrystgstafpgvqaastwdvnlirergqfigeevkaasihvilgpgvaplglk
tpqggrrnwegfgvdpyltgiamgqtingiqsvgvqatakhylneqelnretissnpddrtlhelytwpfadavq
anvasvmcsynkvnttwacedqytlqtlvldqldgfgpyvmtwnaghttvqsansgldmsmpgtdfngnnrlwgp
altnavnsqyvptsrddmvttrilaawyltgqdgagypsfnisrnnvqgnhktnvraiardgivilkndanilplk
kpasiavvgsaaaiignharnspscndkgcddgalgmwgsavnyfpyfvapydaintrassqgtqvtlsntdnts
sgasaargkdvaivfitadsggyitvegnagdrnnldpwhngnalvqavagansnivvvhsvgaileqilal
pqvkavvwaglpsqesgnalvdvlwgdvpsgklvytiakspndytrivsggsdsfseglfidykhfddanitp
ryefgyglstytkfnysrlslstaksgpatgavvpggspdlfqnvatvtvdiansgqvtgaevaqlitytpssap
rtppkqlrgfaklnltppqsgtatfnirrrdlsywdtasqkwvpsgsfgisvgassrdirltstlsva

도면9d

서열번호: 13-39: 신호 펩티드 서열
 서열번호: 13: mryrtaaalalatgpfara
 서열번호: 14: mvsftsllaasppsrascrpaaevesvavekr
 서열번호: 15: mkanvilcllaplva
 서열번호: 16: mivgilttlatlatlaas
 서열번호: 17: myrklavisaflatara
 서열번호: 18: mllnlqvaasalslslgglaea
 서열번호: 19: mklnwvaalsigaagtds
 서열번호: 20: masirsvlvsgllaagvna
 서열번호: 21: mwltspilfastilglgtgvala
 서열번호: 22: mrfswllcpilamgsa
 서열번호: 23: mrlsfpsllvafitlkeass
 서열번호: 24: mqlkflssallslstgncaa
 서열번호: 25: mkvywlvawatsltpala
 서열번호: 26: mvrffsilaaaacfvaves
 서열번호: 27: mihlcpalaallalstqcva
 서열번호: 28: malqtffllaaamlana
 서열번호: 29: mklnkpflaiylafnlaea
 서열번호: 30: maplsrlalsllaltgaaaa
 서열번호: 31: mVRPTILLTSLLLAPFAAA
 서열번호: 32: mhmhslvaalaagtlpllasa
 서열번호: 33: mvhlsslaaalalplvyg
 서열번호: 34: mrfslaattllaglata
 서열번호: 35: mvvlsklvssilfaslvsa
 서열번호: 36: mvqikaaalamlfashvls
 서열번호: 37: mKASSVLLGLAPLAALA
 서열번호: 38: MRFPSIFTAVLFAASSALA
 서열번호: 39:
 MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYLDLEGDFDVAVL PFSNSTNNGLLFINTTIIASIAA
 KEEGVSLDKR

도면9e

서열번호: 40-42: 신호 펩티드 서열
 서열번호: 40: MLLQAFLLAGFAAKISAR
 서열번호: 41:
 ATGATAAAAGTCCCGCGGTTTCATCTGTATGATCGCGCTTACATCCAGCGTTCTGGCAAGCGGCCTTTCTCAAAGC
 GTTTCAGCTCAT
 서열번호: 42:
 ATGAAAAGAAAGCTTGGTCGTCGCCAGTTATTAACTGGCTTTGTTGCCCTTGGCGGTATGGCGATTACAGCTGGT
 AAGGCGCAGGCTTCT

서열 목록

SEQUENCE LISTING

<110> DANISCO US INC.

BOWER, Benjamin S.

FUJDALA, Meredith K.

<120> COMPOSITIONS AND METHODS OF USE

<130> 40298WO-2

<140> PCT/US13/67413
 <141> 2013-10-30
 <150> US 61/720,697
 <151> 2012-10-31
 <160> 42
 <170> PatentIn version 3.5
 <210> 1
 <211> 3454
 <212> DNA
 <213> Magnaporthe grisea
 <220><221> misc_feature
 <223> Nucleic acid sequence encoding Mg3A
 <400> 1
 atgcgtttct ccgggatcgt cgccacgctt gtggcgggtg ccggcgtgtc tgcccaccct 60
 ggtgactact caaaggtatg ataaaaata aaaataaaaa ggctctcaag cctgacgact 120
 cgggaacctt ttactttgtc ccttgttttt gcgcgccgca tgctcagttt tgttggcact 180
 cctttagtaga gagcctattg ccttttggtc aggttgtggt agggccttgt tttgcctatt 240
 ctgtgacct ctccacacaa ttctgaagca attggcttaa atacgagggc aaagttcctt 300
 ctgtcgcgtt gctgttgcaa gaaatccctc gggtacgtca tgtttctggt ttcctcatcat 360
 ggcttccatc aaatacaata gtgctaacta tacatagctt gacgctcggg ccgtggccac 420
 atcagaacct cactaccctc agccgtggat gaaccagat gcagacgggt ggcaggaggc 480
 ctatgtcaaa gcaaaggact ttgtctcgca gatgaccctg cttgagaagg tcaacttgac 540
 aaccggtgta gggtaggtea acaaaccccc ttttttttcg tcatgacccc ctgaatggca 600
 accggcaacc cctccccccc aaacaaaccg tagcgaatat tgacactttt tggtagtgg 660
 gttttagact gggcaagcga tctctgcgtt ggcaacgttg gtgccgtccc ccgcctgggt 720
 ctccgaagcc tctgtctgca ggactctccc accggtgtcc gtttcgccga ttgggtctcg 780
 gtctttctcg ctggcatcac caccggcgcc acgttcgata aggggctgat gtaccgccgt 840
 ggacaagcta tgggccaaga agccaaggac aagggcacat acgtcctgct tgggtccggt 900
 gctggaggac ttggccgtgt ggctgctggt aagcctttgc tcgatacact ggctgatgac 960
 aacatagaga gatgtctttt ttctcacagc taaccgcaca tcataggtgg acgagcttgg 1020
 gagtctttcg gagctgacct agtccttacc ggatatggta tgattgagac aatcaagggt 1080
 atccaggaca ctggtgtcat tgctactgcc aagcatttca tcggcaacga gcaaggtaag 1140

tctatcggtt gacctctttg gccagtcgag atgcctgggtg caagcttgct gtggtttcca	1200
gggtccattta ctaataccca atggcatcca acagagcatt tccgtcaggt aggccaagag	1260
cgaggccgtg gcgtcaacat cagcgagtcg ctctcctcca acattgatga caagactatg	1320
cacgagttgt acctctggta agtccaacac tgcattcagg ggatggattg gattagtgcc	1380
tggtcgacct gaccataata ggccctttgc ggatgcagtt cgcgaggtg tcggctcgggt	1440
tatgtgctcg tacacacagg tcaacaactc gtacggttgc cagaacagca agctttctca	1500
cggtttgctc aaggacgagc ttggcttcca gggattcgtc atgtctgact ggcaggctca	1560
gcacactgga gcggccagcg ctgcggccgg tcttgatatg tccatggta gtcgaagtct	1620
ccaagttcta ttctagtcca tgacaggcaa cgatgctaac gttttctgaa caaaagcccg	1680
gtgacaccga gttcaacaca ggtcttagtt tctggggagc caacttgacc ctggctgtcg	1740
tcaacggaac tgttgccgaa tggcgatcg acgacatggc catgcgaatc atggctgcct	1800
tcttcaaggt tggcaacacg ctggaccagc ccgagatcaa ctttagctct tggaccaagg	1860
acaccttcgg tccccctcac tcgtcatctg gaaacagaat ccagcagatc aaccagcacg	1920
tcgatgtccg tcgcgaccac ggcaacttga ttccgaggt tgcgccaag ggaaccgtgc	1980
ttctcaagaa caccaacaac gccctccgc tcaacaagcc caagttctg gccgtcattg	2040
gtgacgatgc cggtccaac ccgcgcggtc ccaacggctg cccggaccgc ggctgccttt	2100
tgggaacttt gggcatggcc tgggttctg gcacggctga cttcccctac ctcatcacc	2160
ccgacgcggc tctccaagct caggccatcg aggacggcac acgctacgag agcatcctga	2220
gcaactatgc cacggcacag actcaagctc ttgtcagcca aacctacga accgccattg	2280
tctttgttgc tgtctcttcg ggagaaggct gtgagtctgc cttttttgtt tcagctgaac	2340
gcaccatgtt tgaagcaaca ccccttact gacaattctt ctcttccac cagacatcga	2400
ctttgacgga aacaaggag accgcaaca cctgacctc tggatgacg gcgactcgct	2460
ggtcaagaat gtcagctccg tgtgcaaca caccatctc gtcattccact ccaccggccc	2520
gaccatcctg accgagtggc acgacaacc caacgtcacc gccatcgtct gggcaggtgt	2580
tccgggacag gagtccggcc gtgccatcac cgacgtctg tacggccgcg tcaacccgc	2640
cggccgctcg cctttacct ggggcaagac ccgcgagagc tacggcaccg acgtcatgta	2700
caagcccaac aacggcaacg agggcccgca gcaggactac accgaggcg tcttcacga	2760
ctaccgccac tttagaccgc agaaggacga gcccgctac gagttcgcc accgcttgag	2820
ctacaccacc tttagact ccaacatccg cgtcgacaag gccccggcca gcgagtacaa	2880
gcccaccacc ggccagacca tcccggcccc cgtctttggc gccacgtct ccaaggacct	2940

gtcccagtac actttcccggt ccgacgagtt cccgcacatc tacctcttca tctaccgta 3000

cctcaacacc agctcctcgg gtgaggaggc ctgcgcgcac cccaagtacg gcggcacggc 3060

cgaggagtgc ctgccgccca aggcgctcga tggctcgccc cagcccttgc cccgcgcgtc 3120

gggcaagaac tcgcccggcg gcaaccgccca gctgtacgac acctgtaca ccgtcacggc 3180

aaccatcacc aacaccggca agctcgtggg cgaagagggtg ccgcagttgt acgtgtcgca 3240

cggcggcccc gaggaccgcg ccgtcgtcct gcgcggcttc gagcgcatcc gtctcgaccc 3300

cggccagagc gccaccttca aggtcgacct gaccaggagg gacgtcagca actgggacgt 3360

caaggtccag gactgggtca tctctgagca cccaagaag gtgtttgtcg gaagcagcag 3420

caggaagctc catctgagtg ctgacctgaa ctga 3454

<210> 2

<211> 873

<212> PRT

<213> Magnaporthe grisea

<220><221> misc_feature

<223> Polypeptide sequence of Mg3A

<400> 2

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

1 5 10 15

Ser Ala His Pro Gly Asp Tyr Ser Lys Leu Glu Arg Arg Ala Val Ala

20 25 30

Thr Ser Glu Pro His Tyr Pro Gln Pro Trp Met Asn Pro Asp Ala Asp

35 40 45

Gly Trp Gln Glu Ala Tyr Val Lys Ala Lys Asp Phe Val Ser Gln Met

50 55 60

Thr Leu Leu Glu Lys Val Asn Leu Thr Thr Gly Val Gly Trp Ala Ser

65 70 75 80

Asp Leu Cys Val Gly Asn Val Gly Ala Val Pro Arg Leu Gly Leu Arg

85 90 95

Ser Leu Cys Leu Gln Asp Ser Pro Thr Gly Val Arg Phe Ala Asp Trp

100 105 110

Val Ser Val Phe Pro Ala Gly Ile Thr Thr Gly Ala Thr Phe Asp Lys

115 120 125
 Gly Leu Met Tyr Arg Arg Gly Gln Ala Met Gly Gln Glu Ala Lys Asp
 130 135 140
 Lys Gly Ile Asn Val Leu Leu Gly Pro Val Ala Gly Gly Leu Gly Arg
 145 150 155 160
 Val Ala Ala Gly Gly Arg Ala Trp Glu Ser Phe Gly Ala Asp Pro Val

 165 170 175
 Leu Thr Gly Tyr Gly Met Ile Glu Thr Ile Lys Gly Ile Gln Asp Thr
 180 185 190
 Gly Val Ile Ala Thr Ala Lys His Phe Ile Gly Asn Glu Gln Glu His
 195 200 205
 Phe Arg Gln Val Gly Glu Glu Arg Gly Arg Gly Val Asn Ile Ser Glu
 210 215 220
 Ser Leu Ser Ser Asn Ile Asp Asp Lys Thr Met His Glu Leu Tyr Leu

 225 230 235 240
 Trp Pro Phe Ala Asp Ala Val Arg Ala Gly Val Gly Ser Val Met Cys
 245 250 255
 Ser Tyr Thr Gln Val Asn Asn Ser Tyr Gly Cys Gln Asn Ser Lys Leu
 260 265 270
 Leu Asn Gly Leu Leu Lys Asp Glu Leu Gly Phe Gln Gly Phe Val Met
 275 280 285
 Ser Asp Trp Gln Ala Gln His Thr Gly Ala Ala Ser Ala Ala Ala Gly

 290 295 300
 Leu Asp Met Ser Met Pro Gly Asp Thr Glu Phe Asn Thr Gly Leu Ser
 305 310 315 320
 Phe Trp Gly Ala Asn Leu Thr Leu Ala Val Val Asn Gly Thr Val Ala
 325 330 335
 Glu Trp Arg Ile Asp Asp Met Ala Met Arg Ile Met Ala Ala Phe Phe
 340 345 350
 Lys Val Gly Asn Thr Leu Asp Gln Pro Glu Ile Asn Phe Ser Ser Trp

 355 360 365

Thr Lys Asp Thr Phe Gly Pro Leu His Ser Ser Ser Gly Asn Arg Ile
 370 375 380
 Gln Gln Ile Asn Gln His Val Asp Val Arg Arg Asp His Gly Asn Leu
 385 390 395 400
 Ile Arg Glu Val Ala Ala Lys Gly Thr Val Leu Leu Lys Asn Thr Asn
 405 410 415
 Asn Ala Leu Pro Leu Asn Lys Pro Lys Phe Leu Ala Val Ile Gly Asp
 420 425 430
 Asp Ala Gly Ser Asn Pro Arg Gly Pro Asn Gly Cys Pro Asp Arg Gly
 435 440 445
 Cys Leu Leu Gly Thr Leu Gly Met Ala Trp Gly Ser Gly Thr Ala Asp
 450 455 460
 Phe Pro Tyr Leu Ile Thr Pro Asp Ala Ala Leu Gln Ala Gln Ala Ile
 465 470 475 480
 Glu Asp Gly Thr Arg Tyr Glu Ser Ile Leu Ser Asn Tyr Ala Thr Ala
 485 490 495
 Gln Thr Gln Ala Leu Val Ser Gln Thr Tyr Ala Thr Ala Ile Val Phe
 500 505 510
 Val Ala Ala Ser Ser Gly Glu Gly Tyr Ile Asp Phe Asp Gly Asn Lys
 515 520 525
 Gly Asp Arg Asn Asn Leu Thr Leu Trp Tyr Asp Gly Asp Ser Leu Val
 530 535 540
 Lys Asn Val Ser Ser Val Cys Asn Asn Thr Ile Val Val Ile His Ser
 545 550 555 560
 Thr Gly Pro Thr Ile Leu Thr Glu Trp Tyr Asp Asn Pro Asn Val Thr
 565 570 575
 Ala Ile Val Trp Ala Gly Val Pro Gly Gln Glu Ser Gly Arg Ala Ile
 580 585 590
 Thr Asp Val Leu Tyr Gly Arg Val Asn Pro Ala Gly Arg Ser Pro Phe
 595 600 605
 Thr Trp Gly Lys Thr Arg Glu Ser Tyr Gly Thr Asp Val Met Tyr Lys

610	615	620	
Pro Asn Asn Gly Asn Glu Ala Pro Gln Gln Asp Tyr Thr Glu Gly Val			
625	630	635	640
Phe Ile Asp Tyr Arg His Phe Asp Gln Gln Lys Asp Glu Pro Val Tyr			
	645	650	655
Glu Phe Gly His Gly Leu Ser Tyr Thr Thr Phe Glu Tyr Ser Asn Ile			
	660	665	670
Arg Val Asp Lys Ala Pro Ala Ser Glu Tyr Lys Pro Thr Thr Gly Gln			
	675	680	685
Thr Ile Pro Ala Pro Val Phe Gly Ala Asn Val Ser Lys Asp Leu Ser			
690	695	700	
Gln Tyr Thr Phe Pro Ser Asp Glu Phe Pro His Ile Tyr Leu Phe Ile			
705	710	715	720
Tyr Pro Tyr Leu Asn Thr Ser Ser Ser Gly Glu Glu Ala Ser Arg Asp			
	725	730	735
Pro Lys Tyr Gly Gly Thr Ala Glu Glu Phe Leu Pro Pro Lys Ala Leu			
	740	745	750
Asp Gly Ser Pro Gln Pro Leu Pro Arg Ala Ser Gly Lys Asn Ser Pro			
755	760	765	
Gly Gly Asn Arg Gln Leu Tyr Asp Thr Leu Tyr Thr Val Thr Ala Thr			
770	775	780	
Ile Thr Asn Thr Gly Lys Leu Val Gly Glu Glu Val Pro Gln Leu Tyr			
785	790	795	800
Val Ser His Gly Gly Pro Glu Asp Pro Pro Val Val Leu Arg Gly Phe			
	805	810	815
Glu Arg Ile Arg Leu Asp Pro Gly Gln Ser Ala Thr Phe Lys Val Asp			
820	825	830	
Leu Thr Arg Arg Asp Val Ser Asn Trp Asp Val Lys Val Gln Asp Trp			
835	840	845	
Val Ile Ser Glu His Pro Lys Lys Val Phe Val Gly Ser Ser Ser Arg			
850	855	860	
Lys Leu His Leu Ser Ala Asp Leu Asn			

865

870

<210> 3

<211> 855

<212> PRT

<213> Magnaporthe grisea

<220><221> misc_feature

<223> Mature Mg3A polypeptide sequence

<400> 3

His Pro Gly Asp Tyr Ser Lys Leu Glu Arg Arg Ala Val Ala Thr Ser

1 5 10 15

Glu Pro His Tyr Pro Gln Pro Trp Met Asn Pro Asp Ala Asp Gly Trp

20 25 30

Gln Glu Ala Tyr Val Lys Ala Lys Asp Phe Val Ser Gln Met Thr Leu

35 40 45

Leu Glu Lys Val Asn Leu Thr Thr Gly Val Gly Trp Ala Ser Asp Leu

50 55 60

Cys Val Gly Asn Val Gly Ala Val Pro Arg Leu Gly Leu Arg Ser Leu

65 70 75 80

Cys Leu Gln Asp Ser Pro Thr Gly Val Arg Phe Ala Asp Trp Val Ser

85 90 95

Val Phe Pro Ala Gly Ile Thr Thr Gly Ala Thr Phe Asp Lys Gly Leu

100 105 110

Met Tyr Arg Arg Gly Gln Ala Met Gly Gln Glu Ala Lys Asp Lys Gly

115 120 125

Ile Asn Val Leu Leu Gly Pro Val Ala Gly Gly Leu Gly Arg Val Ala

130 135 140

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

145 150 155 160

Gly Tyr Gly Met Ile Glu Thr Ile Lys Gly Ile Gln Asp Thr Gly Val

165 170 175

Ile Ala Thr Ala Lys His Phe Ile Gly Asn Glu Gln Glu His Phe Arg

180	185	190	
Gln Val Gly Glu Glu Arg Gly Arg Gly Val Asn Ile Ser Glu Ser Leu			
195	200	205	
Ser Ser Asn Ile Asp Asp Lys Thr Met His Glu Leu Tyr Leu Trp Pro			
210	215	220	
Phe Ala Asp Ala Val Arg Ala Gly Val Gly Ser Val Met Cys Ser Tyr			
225	230	235	240
Thr Gln Val Asn Asn Ser Tyr Gly Cys Gln Asn Ser Lys Leu Leu Asn			
245	250	255	
Gly Leu Leu Lys Asp Glu Leu Gly Phe Gln Gly Phe Val Met Ser Asp			
260	265	270	
Trp Gln Ala Gln His Thr Gly Ala Ala Ser Ala Ala Ala Gly Leu Asp			
275	280	285	
Met Ser Met Pro Gly Asp Thr Glu Phe Asn Thr Gly Leu Ser Phe Trp			
290	295	300	
Gly Ala Asn Leu Thr Leu Ala Val Val Asn Gly Thr Val Ala Glu Trp			
305	310	315	320
Arg Ile Asp Asp Met Ala Met Arg Ile Met Ala Ala Phe Phe Lys Val			
325	330	335	
Gly Asn Thr Leu Asp Gln Pro Glu Ile Asn Phe Ser Ser Trp Thr Lys			
340	345	350	
Asp Thr Phe Gly Pro Leu His Ser Ser Ser Gly Asn Arg Ile Gln Gln			
355	360	365	
Ile Asn Gln His Val Asp Val Arg Arg Asp His Gly Asn Leu Ile Arg			
370	375	380	
Glu Val Ala Ala Lys Gly Thr Val Leu Leu Lys Asn Thr Asn Asn Ala			
385	390	395	400
Leu Pro Leu Asn Lys Pro Lys Phe Leu Ala Val Ile Gly Asp Asp Ala			
405	410	415	
Gly Ser Asn Pro Arg Gly Pro Asn Gly Cys Pro Asp Arg Gly Cys Leu			
420	425	430	

Leu Gly Thr Leu Gly Met Ala Trp Gly Ser Gly Thr Ala Asp Phe Pro
 435 440 445
 Tyr Leu Ile Thr Pro Asp Ala Ala Leu Gln Ala Gln Ala Ile Glu Asp
 450 455 460
 Gly Thr Arg Tyr Glu Ser Ile Leu Ser Asn Tyr Ala Thr Ala Gln Thr
 465 470 475 480
 Gln Ala Leu Val Ser Gln Thr Tyr Ala Thr Ala Ile Val Phe Val Ala
 485 490 495

 Ala Ser Ser Gly Glu Gly Tyr Ile Asp Phe Asp Gly Asn Lys Gly Asp
 500 505 510
 Arg Asn Asn Leu Thr Leu Trp Tyr Asp Gly Asp Ser Leu Val Lys Asn
 515 520 525
 Val Ser Ser Val Cys Asn Asn Thr Ile Val Val Ile His Ser Thr Gly
 530 535 540
 Pro Thr Ile Leu Thr Glu Trp Tyr Asp Asn Pro Asn Val Thr Ala Ile
 545 550 555 560

 Val Trp Ala Gly Val Pro Gly Gln Glu Ser Gly Arg Ala Ile Thr Asp
 565 570 575
 Val Leu Tyr Gly Arg Val Asn Pro Ala Gly Arg Ser Pro Phe Thr Trp
 580 585 590
 Gly Lys Thr Arg Glu Ser Tyr Gly Thr Asp Val Met Tyr Lys Pro Asn
 595 600 605
 Asn Gly Asn Glu Ala Pro Gln Gln Asp Tyr Thr Glu Gly Val Phe Ile
 610 615 620

 Asp Tyr Arg His Phe Asp Gln Gln Lys Asp Glu Pro Val Tyr Glu Phe
 625 630 635 640
 Gly His Gly Leu Ser Tyr Thr Thr Phe Glu Tyr Ser Asn Ile Arg Val
 645 650 655
 Asp Lys Ala Pro Ala Ser Glu Tyr Lys Pro Thr Thr Gly Gln Thr Ile
 660 665 670
 Pro Ala Pro Val Phe Gly Ala Asn Val Ser Lys Asp Leu Ser Gln Tyr

675 680 685
 Thr Phe Pro Ser Asp Glu Phe Pro His Ile Tyr Leu Phe Ile Tyr Pro
 690 695 700
 Tyr Leu Asn Thr Ser Ser Ser Gly Glu Glu Ala Ser Arg Asp Pro Lys
 705 710 715 720
 Tyr Gly Gly Thr Ala Glu Glu Phe Leu Pro Pro Lys Ala Leu Asp Gly
 725 730 735
 Ser Pro Gln Pro Leu Pro Arg Ala Ser Gly Lys Asn Ser Pro Gly Gly
 740 745 750

 Asn Arg Gln Leu Tyr Asp Thr Leu Tyr Thr Val Thr Ala Thr Ile Thr
 755 760 765
 Asn Thr Gly Lys Leu Val Gly Glu Glu Val Pro Gln Leu Tyr Val Ser
 770 775 780
 His Gly Gly Pro Glu Asp Pro Pro Val Val Leu Arg Gly Phe Glu Arg
 785 790 795 800
 Ile Arg Leu Asp Pro Gly Gln Ser Ala Thr Phe Lys Val Asp Leu Thr
 805 810 815

 Arg Arg Asp Val Ser Asn Trp Asp Val Lys Val Gln Asp Trp Val Ile
 820 825 830
 Ser Glu His Pro Lys Lys Val Phe Val Gly Ser Ser Ser Arg Lys Leu
 835 840 845
 His Leu Ser Ala Asp Leu Asn
 850 855

 <210> 4
 <211> 744
 <212> PRT
 <213> Trichoderma reesei
 <220><221> misc_feature
 <223> T.reesei Bgl1 polypeptide sequence
 <400> 4
 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
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
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
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> 5
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic Forward Primer SK943
 <400> 5
 caccatgaga tatagaacag ctgccgct 28
 <210> 6
 <211> 40
 <212> DNA
 <213> Artificial Sequence

 <220><223> Synthetic Reverse Primer SK941
 <400> 6
 cgaccgcctt gcggagtctt gccagtggt cccgcgacag 40
 <210> 7
 <211> 40
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic Forward Primer SK940
 <400> 7
 ctgtcgcggg accactgggc aagactccgc agggcgggtcg 40
 <210> 8
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic Reverse Primer SK942
 <400> 8
 cctacgctac cgacagagtg 20

 <210> 9
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic Forward Primer SK771

<400> 9
gtctagactg gaaacgcaac 20
<210> 10
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic Reverse Primer SK745
<400> 10
gagttgtgaa gtcggtaatc c 21
<210> 11
<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic primer Mg3A-F
<400> 11

caccatgcgt ttctccggga tcgt 24
<210> 12
<211> 28
<212> DNA
<213> Artificial Sequence
<220><223> Synthetic primer Mg3A-R
<400> 12
tcagttcagg tcagcactca gatggagc 28
<210> 13
<211> 19
<212> PRT
<213> Artificial Sequence
<220><223> Synthetic signal sequence
<400> 13
Met Arg Tyr Arg Thr Ala Ala Ala Leu Ala Leu Ala Thr Gly Pro Phe
1 5 10 15

Ala Arg Ala

<210> 14

<211> 32

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 14

Met Val Ser Phe Thr Ser Leu Leu Ala Ala Ser Pro Pro Ser Arg Ala

1 5 10 15

Ser Cys Arg Pro Ala Ala Glu Val Glu Ser Val Ala Val Glu Lys Arg

20 25 30

<210> 15

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 15

Met Lys Ala Asn Val Ile Leu Cys Leu Leu Ala Pro Leu Val Ala Ala

1 5 10 15

<210> 16

<211> 18

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 16

Met Ile Val Gly Ile Leu Thr Thr Leu Ala Thr Leu Ala Thr Leu Ala

1 5 10 15

Ala Ser

<210> 17

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 17

Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg

1 5 10 15

Ala

<210> 18

<211> 23

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 18

Met Leu Leu Asn Leu Gln Val Ala Ala Ser Ala Leu Ser Leu Ser Leu

1 5 10 15

Leu Gly Gly Leu Ala Glu Ala

20

<210> 19

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 19

Met Lys Leu Asn Trp Val Ala Ala Ala Leu Ser Ile Gly Ala Ala Gly

1 5 10 15

Thr Asp Ser

<210> 20

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 20

Met Ala Ser Ile Arg Ser Val Leu Val Ser Gly Leu Leu Ala Ala Gly

1 5 10 15

Val Asn Ala

<210> 21

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 21

Met Trp Leu Thr Ser Pro Leu Leu Phe Ala Ser Thr Leu Leu Gly Leu

1 5 10 15

Thr Gly Val Ala Leu Ala

20

<210> 22

<211> 16

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 22

Met Arg Phe Ser Trp Leu Leu Cys Pro Leu Leu Ala Met Gly Ser Ala

1 5 10 15

<210> 23

<211> 22

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 23

Met Arg Leu Leu Ser Phe Pro Ser His Leu Leu Val Ala Phe Leu Thr

1 5 10 15

Leu Lys Glu Ala Ser Ser

20

<210> 24

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 24

Met Gln Leu Lys Phe Leu Ser Ser Ala Leu Leu Leu Ser Leu Thr Gly

1 5 10 15

Asn Cys Ala Ala

20

<210> 25

<211> 18

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 25

Met Lys Val Tyr Trp Leu Val Ala Trp Ala Thr Ser Leu Thr Pro Ala

1 5 10 15

Leu Ala

<210> 26

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 26

Met Val Arg Phe Ser Ser Ile Leu Ala Ala Ala Ala Cys Phe Val Ala

1 5 10 15

Val Glu Ser

<210> 27

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 27

Met Ile His Leu Lys Pro Ala Leu Ala Ala Leu Leu Ala Leu Ser Thr

1 5 10 15

Gln Cys Val Ala

20

<210> 28

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 28

Met Ala Leu Gln Thr Phe Phe Leu Leu Ala Ala Ala Met Leu Ala Asn

1 5 10 15

Ala

<210> 29

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 29

Met Lys Leu Asn Lys Pro Phe Leu Ala Ile Tyr Leu Ala Phe Asn Leu

1 5 10 15

Ala Glu Ala

<210> 30

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 30

Met Ala Pro Leu Ser Leu Arg Ala Leu Ser Leu Leu Ala Leu Thr Gly

1 5 10 15

Ala Ala Ala Ala

20

<210> 31

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 31

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

1 5 10 15

Ala Ala Ala

<210> 32

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 32

Met His Met His Ser Leu Val Ala Ala Leu Ala Ala Gly Thr Leu Pro

1 5 10 15

Leu Leu Ala Ser Ala

20

<210> 33

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 33

Met Val His Leu Ser Ser Leu Ala Ala Ala Leu Ala Ala Leu Pro Leu

1 5 10 15

Val Tyr Gly

<210> 34

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 34

Met Arg Phe Ser Leu Ala Ala Thr Thr Leu Leu Ala Gly Leu Ala Thr

1 5 10 15

Ala

<210> 35

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 35

Met Val Val Leu Ser Lys Leu Val Ser Ser Ile Leu Phe Ala Ser Leu

1 5 10 15

Val Ser Ala

<210> 36

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 36

Met Val Gln Ile Lys Ala Ala Ala Leu Ala Met Leu Phe Ala Ser His

1 5 10 15

Val Leu Ser

<210> 37

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 37

Met Lys Ala Ser Ser Val Leu Leu Gly Leu Ala Pro Leu Ala Ala Leu

1 5 10 15

Ala

<210> 38

<211> 19

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 38

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

1 5 10 15

Ala Leu Ala

<210> 39

<211> 85

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 39

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

1 5 10 15

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln

20 25 30

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

35 40 45

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu

50 55 60

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val

65 70 75 80

Ser Leu Asp Lys Arg

85

<210> 40

<211> 20

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 40

Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys

1 5 10 15

Ile Ser Ala Arg

20

<210> 41

<211> 87

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 41

atgataaaag tcccgcggtt catctgtatg atcgcgctta catccagcgt tctggcaagc 60

ggcctttctc aaagcgtttc agtcat 87

<210> 42

<211> 90

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic signal sequence

<400> 42

atgaaaagaa agcttggtcg tcgccagtta ttaactggct ttgttgcct tggcggatg 60

gcgattacag ctggttaaggc gcaggcttct 90