



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0104586
(43) 공개일자 2013년09월25일

(51) 국제특허분류(Int. Cl.)

C12N 9/42 (2006.01) *C07K 19/00* (2006.01)

(21) 출원번호 10-2012-0026203

(22) 출원일자 2012년03월14일

심사청구일자 2012년03월14일

(71) 출원인

경상대학교산학협력단

경상남도 진주시 진주대로 501 (가좌동)

(72) 발명자

김선원

경상남도 진주시 내동면 독산리 휴먼빌아파트
101-1404

텔케 아마르

경남 진주시 가좌동 900번지

(뒷면에 계속)

(74) 대리인

리엔목특허법인

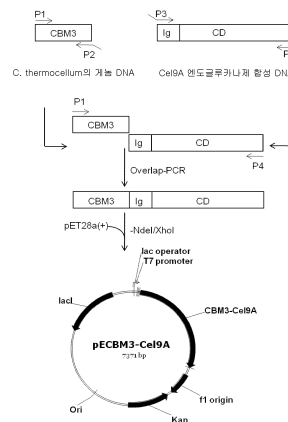
전체 청구항 수 : 총 9 항

(54) 발명의 명칭 **셀룰라제 융합 단백질 및 그를 이용한 셀룰로즈 물질의 분해 방법**

(57) 요약

개선된 특성을 갖는 셀룰라제 및 그를 이용한 셀룰로즈 물질 분해 방법을 제공한다.

대표도 - 도1



(72) 발명자

사바취 수닐

경남 진주시 가좌동 900번지

강서희

경상남도 진주시 옥봉동 176-1 대림아파트 504호

이 발명을 지원한 국가연구개발사업

과제고유번호 M0000035

부처명 지식경제부

연구사업명 국제공동기술개발사업

연구과제명 Development of platform strains for consolidated production of beta-1,3-glucan

주관기관 (주) DMJ Biotech.

연구기간 2010.03.01 ~ 2012.02.29

특허청구의 범위

청구항 1

클로스트리듐 터모셀룸 35319 유래의 패밀리 3 셀룰로즈결합모듈 (cellulose binding domain: CBM), 4 CBM, 및 30 CBM으로 이루어진 군으로부터 선택된 하나이상의 CBM의 C 말단과 글리코실 히드롤라제 패밀리 9 (glycosyl hydrolase family 9: GH9)의 N 말단이 융합된 융합 셀룰라제.

청구항 2

청구항 1에 있어서, 패밀리 3 CBM, 패밀리 4 CBM, 패밀리 30 CBM 및 GH9는 각각 서열번호 1, 2, 3, 및 4의 아미노산 서열을 갖는 것인 셀룰라제.

청구항 3

청구항 1에 있어서, 서열번호 5 내지 7 중 어느 하나의 아미노산 서열을 갖는 것인 셀룰라제.

청구항 4

청구항 1 내지 4 중 어느 하나에 있어서, 말단에 His₆ 태그 서열을 더 포함하는 것인 셀룰라제.

청구항 5

청구항 1 내지 5 중 어느 하나의 셀룰라제와 셀룰로즈 물질을 접촉시키는 단계를 포함하는, 셀룰로즈 물질을 분해시키는 방법.

청구항 6

청구항 5에 있어서, 셀룰로즈 물질은 섬유 물질, 동물 사료로 사용되는 식물, 목재 유래 펄프 또는 2차 섬유인 것인 방법.

청구항 7

청구항 5 또는 6에 있어서, 상기 접촉은 다른 당 분해 효소의 존재하에서 수행되는 것인 방법.

청구항 8

청구항 7에 있어서, 상기 효소는 셀로비오히드롤라제, 엔도글루카나제, 셀로텍스트리나제, β -글루코시다제 및 이들의 조합으로부터 선택되는 것인 방법.

청구항 9

청구항 5 또는 6에 있어서, 상기 접촉은 50℃ 내지 90℃의 온도에서 수행되는 것인 방법.

명세서

기술분야

[0001] 셀룰라제 및 그를 이용한 셀룰로즈 물질의 분해 방법에 관한 것이다.

배경기술

[0002] 작물 유래의 셀룰로즈 기질은 에탄올과 같은 생물연료의 생산에 사용될 수 있다. 셀룰로즈는 자연에서 가장 풍부한 탄소원이지만, 그의 불용해성 및 반-결정성 구조 (quasi-crystalline structure), 및 셀룰로즈 표면의 접근을 제한하는 다른 중합체와의 매트릭스 중에 식물 세포벽의 존재 때문에 분해하기 가장 어렵다. 셀룰로즈의 효율적인 분해가 생물연료 생산에 대한 주요 제한 단계이다. 따라서, 불용성 셀룰로즈에 작용하는 셀룰라제는 이 과정의 핵심 효소이다.

[0003] GH9는 다양한 미생물로부터 분리된 많은 셀룰라제를 포함한다. GH9는 넓은 범위 기질 특이성을 보이며, 단백질

분자 구조에 근거하여 서브그룹 A, B, C 및 D로 분류할 수 있다. 서브그룹 A는 셀룰로즈 결합 모듈 (CBM)이 없이 촉매 모듈로만 구성된 효소를 포함한다. 서브그룹 B는 촉매 모듈의 바로 하류에 패밀리 3 CBM을 갖는다. 서브그룹 C는 N-말단 면역글로불린 (Ig)-유사 도메인 및 뒤이어 하나의 촉매 모듈을 포함한다. 서브그룹 D는 Ig-유사 도메인뿐만 아니라 하나이상의 추가의 CBM을 포함한다. 불용성 셀룰로즈 기질을 이용하는 글리코시드 히드롤라제는 촉매 모듈과 하나이상의 비촉매 셀룰로즈 결합 모듈(CBM)을 포함하는 모듈성 단백질이다. CBM은 식물 구조 및 저장 다당의 효소적 가수분해에 중요한 역할을 한다. 현재 43 패밀리의 CBM이 알려져 있고, 이들 CBM은 리간드 특이성에 실질적 변이를 나타낸다. CBM은 20-200 아미노산으로 구성되고 하나, 둘 또는 3개 도메인으로 셀룰라제 촉매 모듈의 C-말단 또는 N-말단에 이웃하여 존재한다.

[0004] 패밀리 3 CBM, 패밀리 4 CBM, 및 패밀리 30 CBM은 GH9 촉매 모듈과 함께 흔하게 발견되는 것으로서, 촉매 모듈로부터 제거하는 경우, 효소 활성의 현저한 감소를 초래한다.

[0005] 상기한 바와 같이, 종래 알려진 셀룰라제에 비하여, 보다 개선된 효과를 갖는 셀룰라제가 여전히 요구되고 있다.

발명의 내용

해결하려는 과제

[0006] 일 양상은 개선된 특성을 갖는 융합 셀룰라제를 제공하는 것이다.

[0007] 다른 양상은 상기 셀룰로즈를 이용하여 셀룰로즈 물질을 분해하는 방법을 제공하는 것이다.

과제의 해결 수단

[0008] 본 발명의 일 양상은 클로스트리듐 터모셀룸 (*Clostridium thermocellum*) 35319 유래의 패밀리 3 셀룰로즈결합 모듈 (cellulose binding module: CBM), 4 CBM, 및 30 CBM로 이루어진 하나이상의 CBM의 C 말단과 글리코실 히드롤라제 패밀리 9 (glycosyl hydrolase family 9: GH9)의 N 말단이 융합된 융합 셀룰라제를 제공한다.

[0009] 패밀리 3 CBM, 패밀리 4 CBM, 및 패밀리 30 CBM은 각각 서열번호 1, 2, 및 3의 아미노산 서열을 갖는 것일 수 있다. GH9는 알리시클로바실러스 아시도칼드리우스 (*Alicyclobacillus acidocaldrius*) ATCC27009 유래의 GH9 엔도글루카나제 (Ce19A)일 수 있다. Ce19A는 서열번호 4의 아미노산 서열을 갖는 것일 수 있다. Ce19A는 2개 도메인: 1-85번 잔기의 Ig-유사 도메인과 그 다음의 86-537번 잔기의 촉매 도메인으로 구성된 537 아미노산 잔기의 모노머이다.

[0010] 상기 융합 셀룰라제는 서열번호 5 내지 7 중 어느 하나의 아미노산 서열을 갖는 것일 수 있다.

[0011] GH9는 다양한 미생물로부터 분리된 많은 셀룰라제를 포함한다. GH9는 넓은 범위 기질 특이성을 보이며, 단백질 분자 구조에 근거하여 서브그룹 A, B, C 및 D로 분류할 수 있다. 서브그룹 A는 셀룰로즈 결합 모듈 (CBM)이 없이 촉매 모듈로만 구성된 효소를 포함한다. 서브그룹 B는 촉매 모듈의 바로 하류에 패밀리 3 CBM을 갖는다. 서브그룹 C는 N-말단 면역글로불린 (Ig)-유사 도메인 및 뒤이어 하나의 촉매 모듈을 포함한다. 서브그룹 D는 Ig-유사 도메인뿐만 아니라 하나이상의 추가의 CBM을 포함한다.

[0012] 상기 융합 셀룰라제는 하나이상의 말단에 His6 태그 서열과 같은 융합 셀룰라제의 분리에 사용될 수 있는 서열 또는 모이어티를 포함할 수 있다. 융합 셀룰라제의 분리에 사용될 수 있는 서열 또는 모이어티는 융합 셀룰라제의 활성에는 실질적으로 영향을 미치지 않지만, 특정한 물질에 특이적 결합을 하는 것일 수 있다. 또한, 융합 셀룰라제는 CBM과 GH9 셀룰라제의 촉매 모듈 사이에 (Ig)-유사 도메인이 연결 서열 (junction sequence)의 역할을 하는 것으로 보여진다.

[0013] 패밀리 3 CBM, 패밀리 4 CBM, 및 패밀리 30 CBM은 GH9 촉매 모듈과 함께 흔하게 발견되는 것으로서, 촉매 모듈로부터 제거하는 경우, 효소 활성의 현저한 감소를 초래한다.

[0014] 본 발명의 다른 양상은 상기 융합 셀룰라제를 코딩하는 핵산 또는 상기 핵산을 포함하는 벡터를 제공한다. 상기 벡터는 플라스미드 또는 바이러스 계통을 포함하여, 숙주세포에서 복제될 수 있는 것이면 어느 것이나 포함된다. 또한, 상기 핵산은 조절 서열과 작동가능하게 연결된 것일 수 있다. 상기 조절은 리보솜 결합 서열,

Kozak 서열, 프로모터, 오퍼레이터, 폴리A 신호 및 이들의 조합으로부터 선택될 것일 수 있다.

- [0015] 또한, 본 발명의 다른 양상은 상기 핵산 또는 벡터를 포함하는 숙주세포를 제공한다. 상기 숙주세포는 상기 핵산을 발현할 수 있는 것이면 어느 것이나 될 수 있다. 예를 들면, 상기 숙주세포는 원핵 세포 또는 진핵세포일 수 있다. 상기 원핵 세포는 대장균과 같은 박테리아일 수 있다. 상기 진핵세포는 효모 세포와 같은 균류 세포를 포함하나, 여기에 한정되는 것은 아니다.
- [0016] 본 발명의 다른 양상은 상기 숙주세포를 배양하는 단계를 포함하는, 상기 셀룰라제를 생산하는 방법을 제공한다. 상기 방법은 얻어진 배양물로부터 상기 셀룰라제를 분리하는 단계를 더 포함할 수 있다. 또한, 얻어진 배양물을 그대로 사용하거나 세포를 파쇄하여 사용하는 것을 포함할 수 있다.
- [0017] 본 발명의 다른 양상은 상기한 셀룰라제와 셀룰로즈 물질을 접촉시키는 단계를 포함하는, 셀룰로즈 물질을 분해시키는 방법을 제공한다.
- [0018] 상기 접촉은 고상 또는 액상에서 이루어지는 것일 수 있다. 상기 접촉은 10℃ 내지 90℃의 온도, 예를 들면, 40℃ 내지 90℃의 온도, 50℃ 내지 90℃, 40℃ 내지 80℃의 온도, 50℃ 내지 80℃ 또는 55℃ 내지 80℃의 온도에서 수행되는 것일 수 있다. 또한, 상기 접촉은 상압 또는 가압된 조건에서 수행될 수 있다.
- [0019] 상기 접촉 단계에서 있어서, 상기 셀룰라제에 대하여는 상기한 바와 같다. 상기 셀룰라제는 균질하게 분리된 것이거나, 농축된 것 또는 배양물 또는 세포 파쇄물에 포함된 형태의 것일 수 있다.
- [0020] 상기 접촉 단계에서 있어서, 셀룰로즈 물질은 전체 또는 일부가 셀룰로즈를 포함하는 물질을 나타낸다. 셀룰로즈 물질은 수성 매질, 예를 들면 수용액에서 용해성 또는 수불용성의 것일 수 있다. 셀룰로즈 물질은 파쇄되거나 분말화된 것일 수 있다. 셀룰로즈 물질은 섬유 물질, 동물 사료로 사용되는 식물, 목재 유래 펄프 또는 2차 섬유인 것일 수 있다.
- [0021] 상기 접촉 단계에서 있어서, 상기 접촉은 다른 당 분해 효소의 존재하에서 수행되는 것일 수 있다. 상기 효소는 셀룰로즈 또는 그 분해산물을 단당류, 예를 들면 포도당으로 분해할 수 있는 효소 또는 효소 조합일 수 있다. 상기 효소는 셀로비오히드롤라제, 엔도글루카나제, 셀로텍스트리나제, β -글루코시다제 및 이들의 조합으로부터 선택되는 것일 수 있다. 다른 효소의 존재하에서 접촉시킴으로써, 셀룰로즈를 발효가능한 당으로 분해시킬 수 있으며, 얻어진 발효가능한 당을 탄소원으로 미생물의 존재하에서 배양함으로써, 에탄올과 같은 연료 물질을 생산할 수 있다.
- [0022] 따라서, 상기 방법은 얻어진 발효가능한 당을 미생물의 존재하에서 배양하여 상기 당을 연료물질로 전환시키는 단계를 포함할 수 있다. 발효가능한 당은 이당류, 또는 단당류일 수 있다. 발효가능한 당은 포도당일 수 있다. 상기 미생물은 메탄올 또는 에탄올과 같은 알콜 발효를 할 수 있는 미생물 또는 부탄디올과 같은 물질을 생성할 수 있는 미생물일 수 있다. 상기 미생물은 효모 또는 대장균과 같은 박테리아일 수 있다. 상기 연료물질은 메탄올, 에탄올, 부탄디올 및 이들의 조합으로부터 선택된 것일 수 있다.

발명의 효과

- [0023] 본 발명의 융합 셀룰라제에 따르면, 호열성이며, 넓은 기질특이성을 가지고 있고, 불용성 셀룰로즈 물질에 대하여도 높은 분해활성을 가질 수 있다.
- [0024] 본 발명의 셀룰로즈 물질을 분해시키는 방법에 의하면, 셀룰로즈 물질을 효율적으로 분해할 수 있다.

도면의 간단한 설명

- [0025] 도 1은 키메라 효소의 제작 과정을 도식적으로 나타낸 것이다. *C. thermocellum* 유래의 CBM3을 *A. acidocaldrius* 유래의 Cel9A 엔도글루카나제의 Ig-유사 도메인에 융합시켰다. *C. thermocellum* 유래의 CBM4와 CBM30과 Cel9A 엔도글루카나제의 융합도 동일한 전략을 따랐다.

도 2는 본 발명의 키메라 효소와 다른 패밀리 9 셀룰라제의 분자 구조를 나타낸 도면이다.

도 3은 키메라 효소의 발현 및 정제에 대한 SDS-PAGE 분석 결과를 나타낸 도면이다. (A) CBM3-Cel9A, (B) CBM4-Cel9A, 및 (C) CBM30-Cel9A. 레인은 다음과 같다: (W) 전체 세포 단백질, (S) 가용성 단백질, (P) 세포 펠렛 단백질, (HP) His-tag 정제된 단백질, 및 (M) 분자량 마커.

도 4는 키메라 효소의 최적 pH 및 온도를 나타낸 도면이다. 기호는 다음을 나타낸다: Cel9A (네모), CBM3-Cel9A (원), CBM4-Cel9A (삼각형), 및 CBM30-Cel9A (다이아몬드). 표준편차 오류는 10%내이다.

도 5는 PASC, CMC, 필터 페이퍼, Avicel, 셀로테트라오즈, 및 p-NPC에 대한 가수분해 산물의 TLC 분석 결과를 나타낸 도면이다. A: CMC 및 PASC의 광범위(5시간) 가수분해 산물, B: 0 내지 150분까지의 PASC의 가수분해 산물의 시간 코스, C: 필터 페이퍼 및 Avicel의 광범위 (16 시간) 가수분해 산물, D: 셀로테트라오즈의 광범위 (5 시간) 가수분해 산물, E: p-NPC 및 M의 광범위 가수분해 산물: 표준 마커로서, G1 내지 G4는 각각 글루코스, 셀로비오스, 셀로트리오스, 및 셀로테트라오스를 나타낸다.

PASC, CMC, 셀로테트라오즈 및 p-NPC를 60°C에서 0.1 nmol CBM3-Cel9A로 처리하였다. 동일한 반응을 0.1 nmol CBM3-Cel9A로 Avicel 및 필터 페이퍼에 수행하였다. 반응을 효소의 부재 (-) 및 존재하에서 수행하였다.

발명을 실시하기 위한 구체적인 내용

[0026] 이하 본 발명을 실시예를 통하여 보다 상세하게 설명한다. 그러나, 이들 실시예는 본 발명을 예시적으로 설명하기 위한 것으로 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다.

[0027] 재료 및 방법

[0028] 실시예에 사용된 재료 및 방법은 다음과 같다.

[0029] (1) 화학물질

[0030] Avicel PH101, 카르복시메틸 셀룰로즈 (CMC), 보리 β -글루칸, p-니트로페닐- β -D-글루코시드 (p-NPG) 및 p-니트로페닐- β -D-셀로비오시드 (p-NPC)는 Sigma Chemicals (St.Louis, MO, USA)로부터 구입하였다. TLC 플레이트는 Merck Laboratory (Germany)로부터 구입하였다.

[0031] (2) 박테리아 균주, 플라스미드 및 성장 조건

[0032] 플라스미드 pUCe19A 중의 Cel9A 엔도글루카나제를 코딩하는 코돈 최적화된 합성 유전자는 GenScript (NJ, USA)로부터 얻었다. 씨. 터모셀룸 (*C. thermocellum*) ATCC 35319의 게놈 DNA를 패밀리 3, 4 및 30 CBM 유전자 소스로서 사용하였다. 사카로파구스 데그라단스 (*Saccharophagus degradans*) 균주 2-40 ATCC43961의 게놈 DNA를 진행성 (processive) 엔도글루카나제 Cel15H로서 사용하였다. 발현 벡터 pET-28a(+)는 Novagen (Madison, WI, USA)으로부터 구입하였다. 대장균 (*E.coli*) DH5a 및 대장균 BL21 (DE3)를 각각 클로닝 및 발현을 위한 숙주로서 사용하였다. 카나미신으로 보충된 Luria-Bertani (LB) 배지 (50 μ g/ml)를 대장균 세포의 배양을 위하여 사용하였다.

[0033] (3) 키메라 유전자의 제조, 클로닝 및 서열분석

[0034] 키메라 유전자 3-단계 겹침 (overlapping) PCR를 사용하여 제조하였다 (Zhang et al., 2010, *Appl Environ Microbiol*, 76, 6870-6). 도 1은 키메라 유전자의 제조 과정을 나타낸다. pUCe19A 플라스미드를 전장 Cel9A 엔도글루카나제의 증폭을 위한 주형으로 사용하였다.

[0035] 셀룰로솜-통합 단백질 (cellulosome-integrating protein) CipA로부터 패밀리 CBM3, 셀로비오히드롤라제 A (CbhA)로부터 패밀리 CBM4 및 진행성 엔도글루카나제 Cel1J로부터 패밀리 CBM30을 *C. thermocellum*의 게놈 DNA로부터 PCR을 사용하여 증폭하였다. 진행성 엔도글루카나제 Cel15H의 전장 유전자는 *S. degradans* 균주 2-40의 게놈 DNA로부터 증폭하였다. CbhA (GH9-Fn3,2-CBDIII) 유전자를 *C. thermocellum*의 게놈 DNA로부터 PCR을 사용하여 클로닝하였다. 본 실시예에 사용된 프라이머의 서열은 표 1에 나타내었다. 제한 부위를 각 프라이머의 5'

및 3' 말단에 도입하였다. 프라이머는 Bioneer Co., Ltd. (South Korea)로부터 얻었다.

표 1

[0036]

프라이머	서열 (5'→3') ^a	제한효소
P1	TTGTCATATGA ATTTGAAGGTTGAATTCTACAACAGCAATC (서열번호 8)	NdeI
P2	GATTTAGGCACACGGCTCGGGGGTTCTTTACCCCATACAAG (서열번호 9)	
P3	GGGTAAGAACCCCGAGCCGTGTGCCTAAATC (서열번호 10)	
P1'	ATGTCATATG TTAGAAGATAATTCTTCGAC (서열번호 11)	NdeI
P2'	GATTTAGGCACACGGCTCGGAGGCTGCGGAAGTATATATT (서열번호 12)	
P3'	AATATATACTTCCGAGCCTCCGAGCCGTGTGCCTAAATC (서열번호 13)	
P1"	ATTGTCATATG GCTCCTGAAGGCTACAGGAAGC (서열번호 14)	NdeI
P2"	GATTTAGGCACACGGCTCGGGATTGCAGAGCGGACTTTTC (서열번호 15)	
P3"	CCGCTCCTGCAATCCCGAGCCGTGTGCCTAAATC (서열번호 16)	
P4	CCGATCTCGAG TAAACGGCCACGAGCCTCCAG (서열번호 17)	XhoI
Ce15H-F	CCCGGATCCA ATTCTTAGCGGTGGCCAGCAA (서열번호 18)	BamHI
Ce15H-R	CCCGCTCGAG CCAGCTACCAAATTGCAGGGTGT (서열번호 19)	XhoI
CbhA-F	CTAGCTAGC TATACTTCCGAGCCTGAT (서열번호 20)	NheI
CbhA-R	TATATAAGCT TAAACGCCCGGCGGTTCCCCA (서열번호 21)	HindIII

[0037]

a: 제한 효소 부위는 볼드체로 나타내었다. 오버랩 PCR 영역의 부분 (14-20 bp)은 밑줄로 표시하였다. P1-P3, P1'-P3' 및 P1"-P3" 프라이머는 CBM3-Ce19A, CBM4-Ce19A, 및 CBM30-Ce19A 키메라 유전자의 생성을 위하여 사용하였다. P4 프라이머는 공통으로 모든 키메라 유전자의 생성을 위하여 사용하였다.

[0038]

PCR은 다음과 같이 30 연속 사이클을 사용하여 수행되었다.: 변성 98℃에서 0.5분, 어닐링 56℃에서 0.5분, 및 연장 72℃에서 PCR 산물 길이에 따른 시간 동안.

[0039]

Phusion high-fidelity DNA 폴리머라제 (Finnzymes, USA)를 과정 동안 사용하였다. PCR 산물을 1% 아가로즈 겔 전기영동에 의하여 분리하고, 겔 추출 키트를 사용하여 겔로부터 추출하였다. 추출된 DNA 단편을 제한효소로 절단하고 벡터 pET-28a(+)에 연결하였다. 연결 혼합물을 화학적으로 컴페턴트한 대장균 DH5a를 형질전환하는데 사용하였다. 이들 형질전환체로부터 분리된 플라스미드를 제한 분석에 의하여 검증하고, 유전자 서열을 DNA 서열분석에 의하여 확인하였다. Ce15H, CbhA, Ce19A, CBM3-Ce19A, CBM4-Ce19A 및 CBM30-Ce19A에 대한 맞는 서열을 가진 플라스미드를 pECe15H, pECbhA, pECe19A, pECBM3-Ce19A, pECBM4-Ce19A 및 pECBM30-Ce19A로 각각 명명하였다.

[0040]

DNA 조작에 사용된 효소는 New England Biolabs (UK)으로부터 구입하였다. DNA 서열분석은 SolGent (South Korea)에 의하여 수행되었다. 게놈 DNA 추출은 이전에 기술된 방법에 따라 수행하였다 (Sambrook and Russell, 2001). 플라스미드 DNA는 Qiagen Spin Column 플라스미드 Mini-Preps kit (USA)를 사용하여 추출하였다. 제한된 플라스미드 및 PCR 산물을 Qiagen gel extraction kit (USA)를 사용하여 아가로즈 겔로부터 회수하였다.

[0041]

(4) 키메라 단백질의 발현 및 정제

[0042]

pECe15H, pECbhA, pECe19A, pECBM3-Ce19A, pECBM4-Ce19A 및 pECBM30-Ce19A를 가진 대장균 BL21(DE3) 세포를 카나미신 (50 µg/ml) 함유 LB 배지에서 37℃에서 성장시켰다. 배양물이 A600 0.7에 도달하였을 때, IPTG (isopropyl-β-D-thiogalactopyranoside)를 최종 농도 0.2 mM까지 첨가하였다. 다음으로, 배양물을 교반 조건 (180 rpm)에서 22℃에서 밤새 인큐베이션하였다. 밤새 유도 후 배양물을 원심분리하고 (10,000xg, 20 분, 4℃), 세포 펠렛을 회수하였다. 단백질을 제조자의 지침에 따라 CellLyticB system (Sigma)을 사용하여 추출하였다. 키메라 단백질을 천연 조건 (native conditions) 하에서 미리채워진 니켈 세파로즈 칼럼상에 정제하였다 (HisTrap, GE Healthcare/Amersham, Piscataway, NJ). 상기 키메라 단백질은 N 말단에 (His)₆ 서열을 포함한다.

[0043] (5) 효소 분석

[0044] 키메라 효소에 대한 대부분 분석은 50 mM 소듐 아세테이트 (pH 5.0), 200 mM NaCl, 및 10 mM CaCl₂ 함유 반응 버퍼 중 0.1-1.0 nmol 정제된 효소로 pH 5.0 및 70℃에서 수행하였다. 효소 활성 1 단위는 분당 1 μmol의 환원당을 방출하는 효소의 양으로 정의하였다. 촉매 활성에 대한 pH 및 온도의 영향은 각각 pH 범위 4.0-8.0 및 10-90℃에서 시험하였다. 소듐 아세테이트 (4.0 to 5.6) 및 피페라진-*N,N'*-비드(2-에탄술폰산) (6.0 내지 8.0) 버퍼 최종 농도 50 mM에서 pH를 유지하는데 사용하였다.

[0045] 카르복시메틸 셀룰로즈 (CMC) 및 보리 β-글루칸 분석은 총부피 0.5 ml 중의 1% (w/v) 기질로 10 분 반응 시간 동안 수행하였다. 인산-부풀려진 셀룰로즈 (Phosphoric acid-swollen cellulose: PASC)는 이전 보고에 기술된 방법에 따라 제조하였다 (Wood, 1971, *Biochem J*, 121, 353-62; Zhang et al., 2006, *Biomacromolecules*, 7, 644-8). PASC (0.1%, w/v) 분해 분석은 10 분 동안 0.5ml 총부피 중에서 수행하였다. 디니트로살리실산 (DNS) 분석은 이들 반응 중 산물 형성을 검출하기 위하여 수행하였다 (Ghose, 1987, *Pure Appl Chem*, 59, 257-268).

[0046] Avicel (1%, w/v) 또는 왓만 번호 1 필터 페이퍼 (3mg)의 분해는 총부피 0.25ml 중에서 60℃에서 2 시간 동안 수행하였다. 반응을 95℃에서 15 분 동안 인큐베이션시켜 중단시키고, 잔존 기질을 10,000xg에서 원심분리에 의하여 분리하였다. 다음으로, Avicel 또는 필터 페이퍼 분해의 산물을 총부피 0.35 ml 중 50 U의 β-글루코시다제 (Sigma, USA)로 37℃에서 1시간 동안 분해하였다. β-글루코시다제를 95℃에서 15분 동안 인큐베이션시켜 불활성화시켰다. 최종 반응 혼합물 중의 글루코스 농도를, 제조사의 지침에 따라, 글루코스 옥시다제 키트 (GAGO-20 from Sigma, USA)를 사용하여 측정하였다. 셀로비오스의 방출은 글루코스 축적 속도의 50% (Watson et al., 2009, *J Bacteriol*, 191, 5697-705)로서 계산하였다. *p*-니트로페닐-β-D-글리코시드 (10 mM) 및 *p*-니트로페닐-β-D-셀로비오시드 (10 mM)의 가수분해는 최종농도 0.1 M로 NaOH를 첨가한 후 410nm에서 방출된 *p*-니트로페놀 농도를 모니터링하여 분석하였다 (Eckert et al., 2002, *Appl Microbiol Biotechnol*, 60, 428-36). CMC 용액의 점도는 Brookfield DV-III viscometer (USA)를 사용하여 측정하였다. 측정은 100rpm 모터 속도로 25℃에서 제2 숫자 방추 (second number spindle)을 사용하여 수행하였다. 셀룰라제를 0.5% (wt/v) CMC 용액에 0.1nmol의 양으로 첨가하고, 60℃에서 30분 동안 인큐베이션하고, 점도 측정에 사용하였다.

[0047] (6) 단백질 정량 및 SDS-PAGE 분석

[0048] 단백질 농도는 표준으로서 정제된 소혈청알부민(BSA)을 사용하여 Pierce BCA 단백질 분석 키트 (Thermoscientific, USA)를 사용하여 결정하였다. 소듐 도데실 술페이트 폴리아크릴아미드 겔 전기영동 (SDS-PAGE)은 수직 폴리아크릴아미드 슬랩 겔 (vertical polyacrylamide slab gel)에서 수행하였다.

[0049] 전기영동은 변성 조건에서 5% 스택킹 겔 및 12% 폴리아크릴아미드 겔 상에서 수행되었다. 스택킹 겔 및 폴리아크릴아미드 겔에 대한 버퍼 용액은 각각 1 M Tris-HCl (pH 6.8) 및 1.5 M Tris-HCl (pH 8.8)이었다.

[0050] (7) 셀룰라제 결합 분석

[0051] 키메라 효소의 Avicel 결합 특성은 사소한 변형과 함께 이전에 기술된 바와 같이 연구하였다 (Li et al., 2007, *Appl Environ Microbiol*, 73, 3165-72). 결합 분석은 2.0 ml 튜브 내에서 수행하였다. Avicel 1 mg을 소듐 아세테이트 버퍼 (20 mM, pH 5) 중 0.5% BSA와 혼합하였다. 혼합물을 비특이적 결합을 방지하기 위하여 25℃에서 30 분 동안 인큐베이션하였다. 이 용액에 동일한 양 (2nmol)의 천연 또는 키메라 효소를 첨가하였다.

[0052] 다음으로, 회전속도 20rpm으로 1 시간 동안 intelli-mixer RM-2 (Rose Scientific Ltd., Canada) 내에 두었다. 반응 혼합물을 10,000xg에서 5분 동안 원심분리하고, 결합되지 않은 효소의 양을 상등액 중 잔존 셀룰라제 활성으로부터 추정하였다. Avicel 결합 효소의 양은 초기 효소 활성과 결합되지 않은 효소 활성 사이의 차이로부터 계산하였다. 대조군 반응은 Avicel이 없는 것을 제외하고, 동일한 조건에서 수행하였다. 또한, Avicel 결합 단백질은 SDS-PAGE를 사용하여 정성적으로 분석하였다.

[0053] (8) 셀룰라제 진행성 (processivity) 분석 및 상승적 상호작용

[0054] 셀룰라제 진행성은 Watson 등의 과정에 따른 불용성 환원당에 대한 가용성 환원당의 비율을 측정하여 평가하였다 (Watson et al., 2009, *J Bacteriol*, 191, 5697-705). 상기 비율은 가용성 환원당 μmol (셀로비오스 표준)을 불용성 당 μmol (글루코스 표준)으로 나누는 것과 같이 얻었다. 필터 페이퍼를 셀룰라제 진행성 시험을

위한 기질로서 사용하였다. 필터 페이퍼 가수분해 및 원심분리 후, 상등액 분획 중 가용성 환원당을 상기한 바와 같은 글루코스 옥시다제 방법을 사용하여 측정하였다. 침전물 분획 중의 불용성 환원당은 변형된 2,2-비친코니네이트 (bicinchoninate) 분석 (Doner and Irwin, 1992, *Anal Biochem*, 202, 50-3)을 사용하여 결정하였다. 침전된 필터 페이퍼를 6 M 구아니딘 히드로클리드로 세척하여 결합되지 않은 셀룰라제를 제거한 다음, 분석 버퍼 및 물로 다시 4회 세척하였다. 처리된 필터 페이퍼의 환원당은 표준으로서 글루코스를 사용하여 Pierce microBCA 시약 (USA)을 사용하여 측정하였다. 상기 진행성을 1 nmol의 CBM3-Ce19A 키메라 효소 및 *S. degradans* Ce15H 엔도글루카나제를 사용하여 조사하였다. 필터 페이퍼에 대한 CBM3-Ce19A 및 CbhA 사이의 상승적 상호작용을 조사하였다. 반응 혼합물은 7 mg 필터 페이퍼 디스크, 및 50 mM 소듐 아세테이트 (pH 5.0), 200 mM NaCl, 및 10 mM CaCl₂ 함유 반응 버퍼 중 CBM3-Ce19A 및 CbhA의 동일한 물 혼합물 (1:1)로 구성하였다. 반응을 60°C에서 10 시간 동안 수행하고 환원당 산물을 DNS 방법으로 측정하였다. 대조군으로서, 동일 반응을 CBM3-Ce19A 대신에 Ce19A를 사용하여 수행하였다. 상승적 효과의 정도 (degree of synergistic effect: DSE)는 관찰된 개별 활성의 합에 대한 조합된 효소의 관찰된 활성의 비로서 정의하였다. 모든 효소 활성 분석은 3배수로 진행하였고, 데이터는 평균±표준편차로서 나타내었다.

[0055] (9) 박막 크로마토그래피

[0056] 시료 1 마이크로리터를 silica 60 TLC plate (Merck, USA) 상에 스팟팅하고, 공기 건조시켰다. 크로마토그램을 니트로메탄, 1-프로판올, 및 물 (2:5:1.5) (v/v/v) (Watson et al., 2009, *J Bacteriol*, 191, 5697-705)을 사용하여 현상하였다. TLC 판을 메탄올 중 0.3% (w/v) α-나프톨 및 5% (v/v) 황산의 혼합물 중에 담그고 110°C까지 10 분 동안 가열하여 분할된 산물을 가시화하였다.

[0057] 실시예 1: 키메라 효소의 제조 및 정제

[0058] *C. thermocellum* 유래 패밀리 3, 4 및 30 CBM을 도 1에 기술된 바와 같이 오버랩 PCR 기법을 통하여 Ce19A 엔도글루카나제의 N-말단에 별개로 융합시켰다. 키메라 효소의 분자 구조 (molecular architecture) 및 이미 보고된 GH9 패밀리 셀룰라제와의 비교는 도 2에 나타내었다. 다음으로, 융합된 유전자를 N-말단 His₆ 태그를 코딩하는 발현 벡터 pET-28a(+)에 삽입하였다. 그 결과 얻어진 구조체를 대장균 BL21(DE3) 중에서 단백질 발현을 위하여 사용하였다. 발현된 단백질을 CellyticB 시약을 사용하여 추출하였다. 단백질의 가용성 발현은 SDS-PAGE에 의하여 확인하였다 (도 3). 이들 단백질을 천연 조건에서 미리 채워진 니켈 세파로즈 칼럼을 사용하여 정제하였다. 결합은 20 mM 이미다졸의 존재하에서 이루어졌고, 40 mM 이미다졸 함유 1X 소듐 포스페이트 버퍼 (pH 7.5-8.0)를 사용하여 세척하였다. 마지막으로, 250 mM 이미다졸을 사용하여 용출하였다. 원하는 단백질 함유 분획을 풀링하고, 50 mM 소듐 포스페이트 버퍼 (pH 7.0)에 대하여 4°C에서 밤새 투석하였다. 마지막으로 효소를 10% 글리세롤 함유 50 mM 피페라진-N,N'-비스(2-에탄술포산) 버퍼 (pH 7.0)에 저장하였다. 정제된 제조물 각각은 SDS-PAGE 상에서 단일 밴드를 나타내었고 분자량은 뉴클레오타이드 서열로부터 계산된 것과 잘 일치하였다 (도 3).

[0059] 실시예 2: 키메라 효소의 효소 특성

[0060] 키메라 효소의 최적 온도 및 pH를 온도 범위 10 내지 90°C 및 pH 범위 4.0 내지 8.0에서 결정하였다 (도 4). 도 4A의 온도 변화 실험 조건은 pH 5.0에서 기질로 1%(w/v) CMC를 사용하고 반응 용액 조성은 50mM 소듐 아세테이트 (pH 5.0), 200 mM NaCl 및 10 mM CaCl₂를 포함한다. 도 4B의 pH 변화 실험 조건은 70°C에서 기질 1%(w/v) CMC를 사용하고 pH 구간에 따라서 해당되는 버퍼를 사용하였다. 비교 목적을 위하여, Ce19A 엔도글루카나제를 또한 유사한 분석에 사용하였다. 그 결과, 모든 키메라 효소는 최적 pH 5.0을 보인 반면, 천연 Ce19A 엔도글루카나제는 최적 pH 5.5를 보였다. 놀랍지 않게도, 모든 천연 및 키메라 효소는 동일한 최적 온도 70°C를 보였다. 4°C에서 밤새 저장되었을 때, 키메라 효소는 pH 범위 5.0-8.0에서 안정하였다. 반면 60°C에서 1시간 후 약간의 효소 활성 손실이 있었다. 천연 Ce19A 엔도글루카나제의 반감기는 75°C에서 30분인 반면, 키메라 효소의 반감기는 동일 온도에서 20분이었다.

[0061] 키메라 효소는 넓은 기질 특이성을 가졌고 가용성뿐만 아니라 불용성 셀룰로즈 기질을 효율적으로 가수분해하였다 (표 2).

표 2

기질	비활성(U/μmol)			
	Ce19A (59kDa)	CBM3-Ce19A (76kDa)	CBM4-Ce19A (80kDa)	CBM30-Ce19A (80kDa)
CMC	4,325±70	6,600±90	7,320±80	7,720±90
보리 β-글루칸	11,800±200	12,800±190	15,000±170	15,000±180
PASC	649±50	1,000±60	1,260±70	1,280±70
Avicel	0.50±0.09	6.0±0.9	4.0±0.7	5.0±0.8
필터 페이퍼	0.70±0.08	7.0±1.0	5.4±1.0	6.0±0.9
pNPG	ND ^a	ND ^a	ND ^a	ND ^a
pNPC	118±30	120±30	130±20	135±20

[0063] ^a: 활성이 검출되지 않음

[0064] CMC에 대한 CBM4-Ce19A 및 CBM30-Ce19A 키메라 효소의 촉매 활성은 천연 Ce19A 엔도글루카나제의 그것에 비하여 1.7 및 1.8 배 증가하였다. 천연 효소에 비하여, PASC에 대한 키메라 효소의 촉매 활성은 2배 증가하였다. 불용성 다당류 (Avicel 및 필터 페이퍼)에 대한 키메라 효소의 촉매 활성은 천연 효소의 그것에 비하여 7.7-12.0배로 현저하게 증가하였다. 모든 키메라 효소 중, CBM3-Ce19A 키메라 효소가 불용성 다당류에 대하여 가장 높은 촉매 활성을 보였다. 그러나, p-NPC에 대한 키메라 효소의 효소 활성에는 현저한 증가가 없었다. 모든 효소에서 p-NPG에 대하여 검출가능한 활성이 없었다.

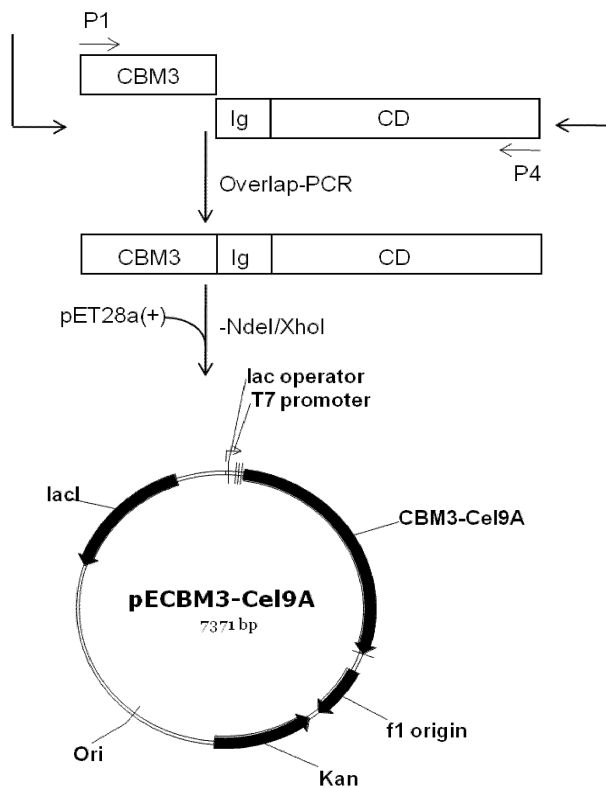
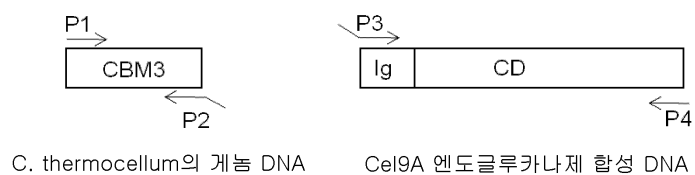
[0065] CMC 분해의 점도측정 분석 (viscometric assay)을 키메라 단백질의 엔도글루카나제 활성을 조사하기 위하여 수행하였다. CMC 용액 (0.5%, w/v)의 점도는 CBM3-Ce19A 키메라 효소와 인큐베이션한 후 감소하였다. CMC의 점도는 인큐베이션 30분 내에 80 cP에서 10 cP로 감소하였다. CMC 점도의 현저한 감소는 키메라 단백질의 엔도글루카나제 활성을 나타낸다. 키메라 단백질의 Avicel 결합 친화도를 결정하여 불용성 다당류에 대한 그들의 증가된 촉매 활성과의 상관관계를 발견하였다. 천연 Ce19A 엔도글루카나제에 비하여, 키메라 효소가 Avicel에 대한 가장 높은 친화도를 보였다. Ce19A, CBM3-Ce19A, CBM4-Ce19A 및 CBM30-Ce19A의 결합 비율은 각각 5%, 50%, 35% 및 37%이었다.

[0066] 실시예 3: 가수분해 산물 및 진행성의 분석

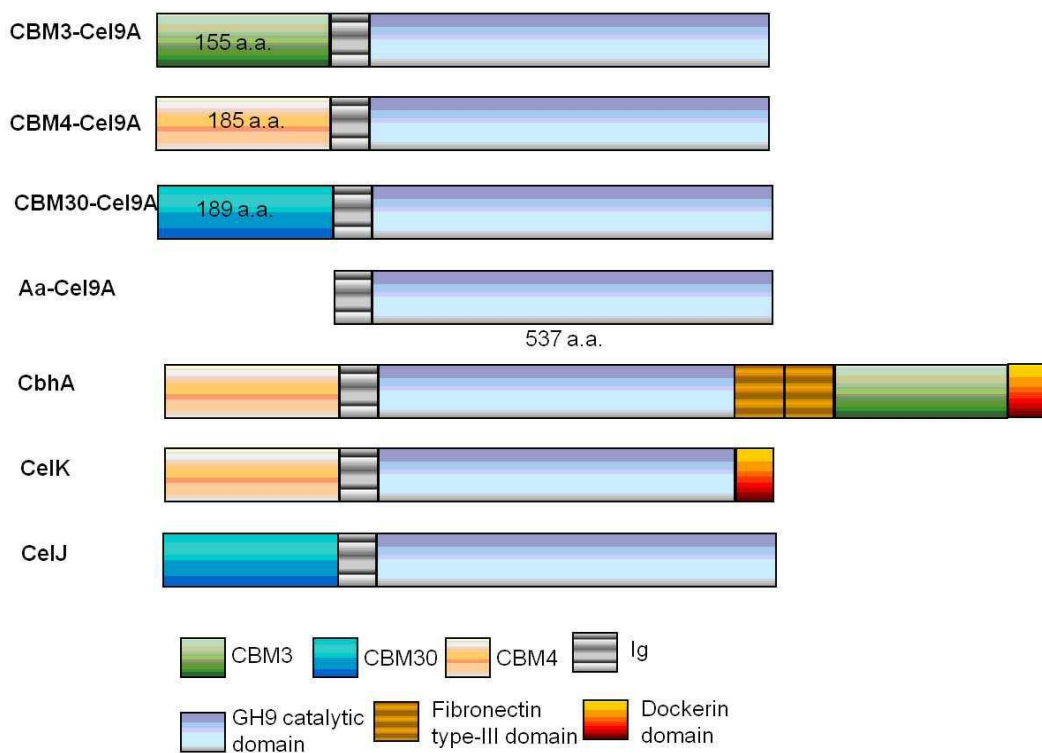
[0067] CBM3-Ce19A에 의한 CMC, PASC, 필터 페이퍼, Avicel, 셀로테트라오즈 및 p-NPC의 가수분해 산물을 TLC에 의하여 정성적으로 분석하였다 (도 5). 셀로비오즈 및 셀로테트라오즈를 CMC, PASC, 필터 페이퍼, 및 Avicel의 가수분해 후 검출하였다. PASC 가수분해의 시간에 따른 반응에서 (도 5B), 셀로비오즈 및 셀로테트라오즈 모두 전 시간에 관찰되었다. 키메라 효소는 절단 산물로서 셀로비오즈 및 셀로테트라오즈를 동시에 생성하는 것으로 여겨진다. 시간에 따른 반응에서, TLC 상의 셀로비오즈 스팟의 강도는 반응 시간 30분부터 150분까지의 증가에 따라 증가한 반면, 셀로테트라오즈 스팟 강도의 유의한 변화는 없었다. 이는 키메라 효소에 의하여 셀로테트라오즈가 셀로비오즈로 더 분해된다는 것을 나타낸다. 셀로테트라오즈 및 p-NPC가 기질로서 사용된 경우 (도 5D 및 E), 키메라 효소는 상기 기질을 효율적으로 가수분해시켜 동시에 셀로비오즈를 방출시켰다. 이들 결과는 주요 가수분해 산물로서 셀로텍스트린을 생성하는 엔도글루카나제의 통상적 특성과는 다르게, 키메라 효소는 셀룰로즈를 셀로비오즈로 가수분해하는 경향이 있다는 것을 암시한다. 셀로비오즈의 생산은 진행성 (processivity)의 표시일 수 있다. CBM3-Ce19A 키메라 효소의 진행성을 평가하기 위하여, 가용성 및 불용성 당의 농도를 필터 페이퍼의 가수분해 후에 측정하고, *S. degradans* 유래의 알려진 진행성 엔도글루카나제 Ce15H와 비교하였다. CBM3-Ce19A 키메라 효소는 필터 페이퍼의 가수분해 2시간 후에 81% 가용성 환원당 및 19% 불용성 환원당을 생성하였다. Ce15H 참조 효소는 80% 가용성 환원당 및 20% 불용성 환원당을 생성하였다. 키메라 효소 및 Ce15H의 진행성은 각각 4.26 및 4.0이었다.

도면

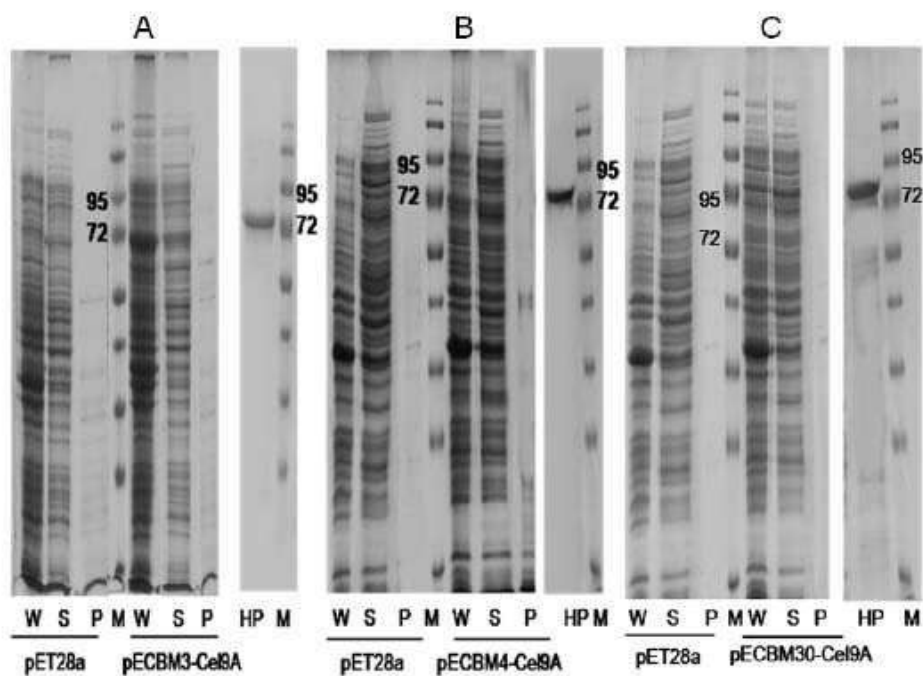
도면1



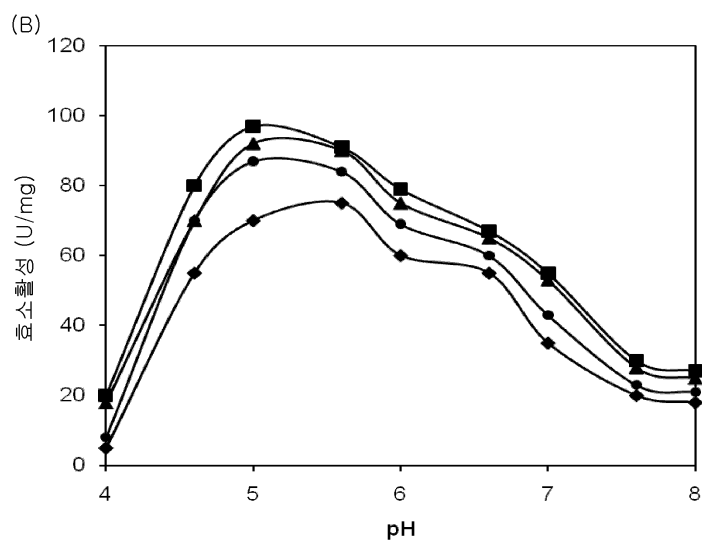
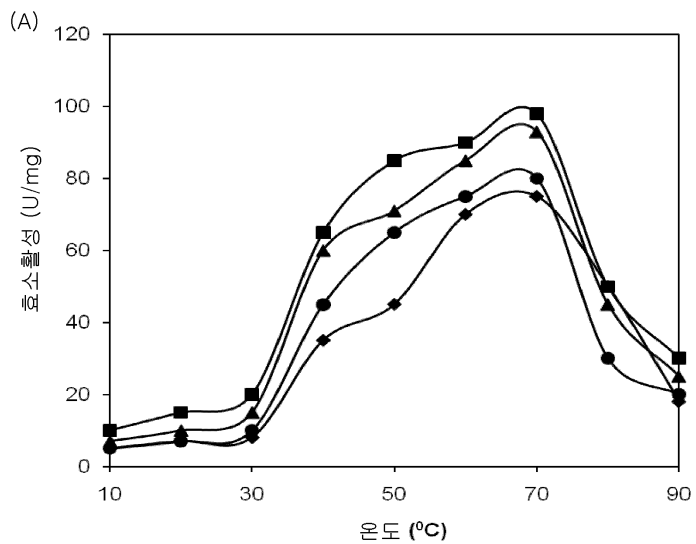
도면2



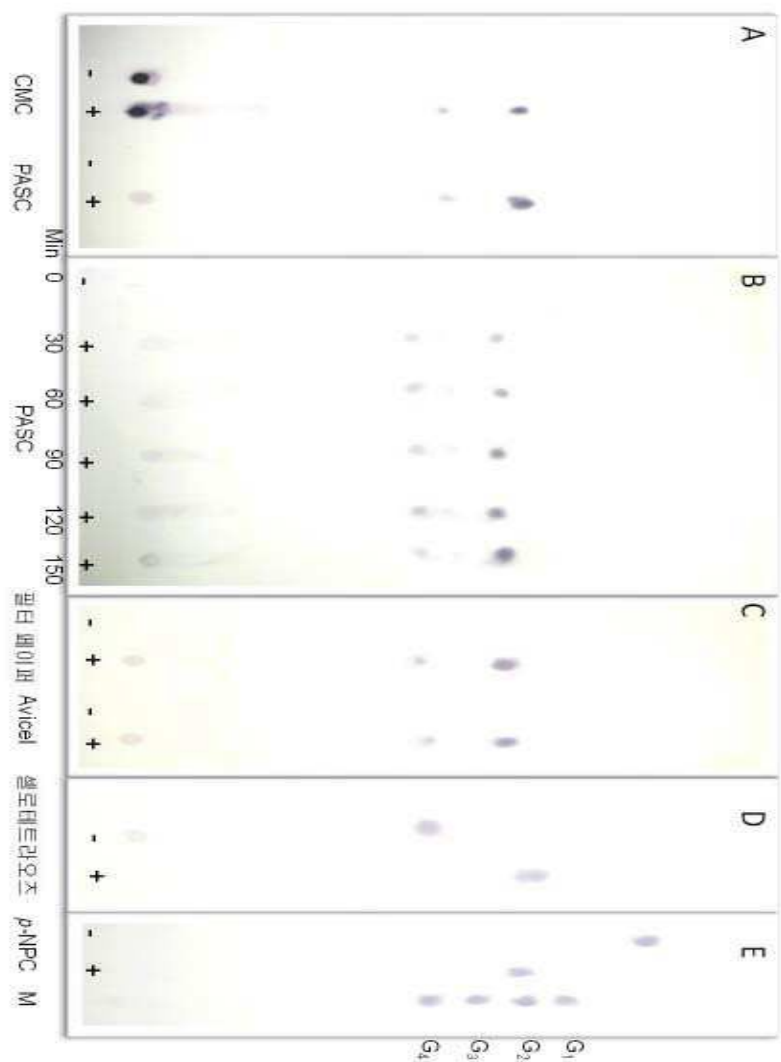
도면3



도면4



도면5



서열목록

- <110> INDUSTRY-ACADEMIC COOPERATION FOUNDATION GYENGsang UNIVERSITY
- <120> Cellulase fusion protein and method of degrading cellulosic material using the same
- <130> PN096065
- <160> 21
- <170> KopatentIn 2.0
- <210> 1
- <211> 155
- <212> PRT
- <213> Clostridium thermocellum
- <400> 1

Asn Leu Lys Val Glu Phe Tyr Asn Ser Asn Pro Ser Asp Thr Thr Asn
1 5 10 15
Ser Ile Asn Pro Gln Phe Lys Val Thr Asn Thr Gly Ser Ser Ala Ile
20 25 30
Asp Leu Ser Lys Leu Thr Leu Arg Tyr Tyr Tyr Thr Val Asp Gly Gln
35 40 45
Lys Asp Gln Thr Phe Trp Cys Asp His Ala Ala Ile Ile Gly Ser Asn
50 55 60
Gly Ser Tyr Asn Gly Ile Thr Ser Asn Val Lys Gly Thr Phe Val Lys
65 70 75 80
Met Ser Ser Ser Thr Asn Asn Ala Asp Thr Tyr Leu Glu Ile Ser Phe
85 90 95
Thr Gly Gly Thr Leu Glu Pro Gly Ala His Val Gln Ile Gln Gly Arg
100 105 110
Phe Ala Lys Asn Asp Trp Ser Asn Tyr Thr Gln Ser Asn Asp Tyr Ser
115 120 125
Phe Lys Ser Arg Ser Gln Phe Val Glu Trp Asp Gln Val Thr Ala Tyr
130 135 140
Leu Asn Gly Val Leu Val Trp Gly Lys Glu Pro
145 150 155
<210> 2
<211> 175
<212> PRT
<213> Clostridium thermocellum

<400> 2
Met Leu Glu Asp Asn Ser Ser Thr Leu Pro Pro Tyr Lys Asn Asp Leu
1 5 10 15
Leu Tyr Glu Arg Thr Phe Asp Glu Gly Leu Cys Tyr Pro Trp His Thr
20 25 30
Cys Glu Asp Ser Gly Gly Lys Cys Ser Phe Asp Val Val Asp Val Pro
35 40 45
Gly Gln Pro Gly Asn Lys Ala Phe Ala Val Thr Val Leu Asp Lys Gly

50 55 60
 Gln Asn Arg Trp Ser Val Gln Met Arg His Arg Gly Leu Thr Leu Glu

 65 70 75 80
 Gln Gly His Thr Tyr Arg Val Arg Leu Lys Ile Trp Ala Asp Ala Ser
 85 90 95
 Cys Lys Val Tyr Ile Lys Ile Gly Gln Met Gly Glu Pro Tyr Ala Glu
 100 105 110
 Tyr Trp Asn Asn Lys Trp Ser Pro Tyr Thr Leu Thr Ala Gly Lys Val
 115 120 125
 Leu Glu Ile Asp Glu Thr Phe Val Met Asp Lys Pro Thr Asp Asp Thr
 130 135 140

 Cys Glu Phe Thr Phe His Leu Gly Gly Glu Leu Ala Ala Thr Pro Pro
 145 150 155 160
 Tyr Thr Val Tyr Leu Asp Asp Val Ser Leu Tyr Asp Pro Glu Tyr
 165 170 175
 <210> 3
 <211> 185
 <212> PRT
 <213> Clostridium thermocellum
 <400> 3
 Ala Pro Glu Gly Tyr Arg Lys Leu Leu Asp Val Gln Ile Phe Lys Asp
 1 5 10 15
 Ser Pro Val Val Gly Trp Ser Gly Ser Gly Met Gly Glu Leu Glu Thr

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

85

90

95

Gly Glu Asp Phe Val Ile Gly Phe Arg Asp Lys Val Tyr Glu Arg Val

100

105

110

Tyr Gly Leu Glu Ile Asp Val Thr Thr Val Ile Ser Asn Tyr Val Thr

115

120

125

Val Thr Thr Asp Trp Gln His Val Lys Ile Pro Leu Arg Asp Leu Met

130

135

140

Lys Ile Asn Asn Gly Phe Asp Pro Ser Ser Val Thr Cys Leu Val Phe

145

150

155

160

Ser Lys Arg Tyr Ala Asp Pro Phe Thr Val Trp Phe Ser Asp Ile Lys

165

170

175

Ile Thr Ser Glu Asp Asn Glu Lys Ser

180

185

<210> 4

<211> 537

<212> PRT

<213> Clostridium thermocellum

<400> 4

Met Pro Ser Arg Val Pro Lys Ser Ile Phe Tyr Asn Gln Val Gly Tyr

1

5

10

15

Leu Ile Ser Gly Asp Lys Arg Phe Trp Ile Gln Ala His Glu Pro Gln

20

25

30

Pro Phe Ala Leu Arg Thr Pro Glu Gly Gln Ala Val Phe Ala Gly Met

35

40

45

Thr Lys Pro Val Gly Gly Asn Trp Tyr Val Gly Asp Phe Thr Ala Leu

50

55

60

Arg Val Pro Gly Thr Tyr Thr Leu Thr Val Gly Thr Leu Glu Ala Arg

65

70

75

80

Val Val Ile His Arg Arg Ala Tyr Arg Asp Val Leu Glu Ala Met Leu

85

90

95

Arg Phe Phe Asp Tyr Gln Leu Cys Gly Val Val Leu Pro Glu Asp Glu

100	105	110	
Ala Gly Pro Trp Ala His Gly Ala Cys His Thr Ser Asp Ala Lys Val			
115	120	125	
Phe Gly Thr Glu Arg Ala Leu Ala Cys Pro Gly Gly Trp His Asp Ala			
130	135	140	
Gly Asp Tyr Gly Lys Tyr Thr Val Pro Ala Ala Lys Ala Val Ala Asp			
145	150	155	160
Leu Leu Leu Ala His Glu Tyr Phe Pro Ala Ala Leu Ala His Val Arg			
165	170	175	
Pro Met Arg Ser Val His Arg Ala Pro His Leu Pro Pro Ala Leu Glu			
180	185	190	
Val Ala Arg Glu Glu Ile Ala Trp Leu Leu Thr Met Gln Asp Pro Ala			
195	200	205	
Thr Gly Gly Val Tyr His Lys Val Thr Thr Pro Ser Phe Pro Pro Leu			
210	215	220	
Asp Thr Arg Pro Glu Asp Asp Asp Ala Pro Leu Val Leu Ser Pro Ile			
225	230	235	240
Ser Tyr Ala Ala Thr Ala Thr Phe Cys Ala Ala Met Ala His Ala Ala			
245	250	255	
Leu Val Tyr Arg Pro Phe Asp Pro Ala Leu Ser Ser Cys Cys Ala Asp			
260	265	270	
Ala Ala Arg Arg Ala Tyr Ala Trp Leu Gly Ala His Glu Met Gln Pro			
275	280	285	
Phe His Asn Pro Asp Gly Ile Leu Thr Gly Glu Tyr Gly Asp Ala Glu			
290	295	300	
Leu Arg Asp Glu Leu Leu Trp Ala Ser Cys Ala Leu Leu Arg Met Thr			
305	310	315	320
Gly Asp Ser Ala Trp Ala Arg Val Cys Glu Pro Leu Leu Asp Leu Asp			
325	330	335	
Leu Pro Trp Glu Leu Gly Trp Ala Asp Val Ala Leu Tyr Gly Val Met			
340	345	350	

Asp Tyr Leu Arg Thr Pro Arg Ala Ala Val Ser Asp Asp Val Arg Asn
355 360 365

Lys Val Lys Ser Arg Leu Leu Arg Glu Leu Asp Ala Leu Ala Ala Met
370 375 380

Ala Glu Ser His Pro Phe Gly Ile Pro Met Arg Asp Asp Asp Phe Ile
385 390 395 400

Trp Gly Ser Asn Met Val Leu Leu Asn Arg Ala Met Ala Phe Leu Leu
405 410 415

Ala Glu Gly Val Gly Val Leu His Pro Ala Ala His Thr Val Ala Gln
420 425 430

Arg Ala Ala Asp Tyr Leu Phe Gly Ala Asn Pro Leu Gly Gln Cys Tyr
435 440 445

Val Thr Gly Phe Gly Gln Arg Pro Val Arg His Pro His His Arg Pro
450 455 460

Ser Val Ala Asp Asp Val Asp His Pro Val Pro Gly Met Val Val Gly
465 470 475 480

Gly Pro Asn Arg His Leu Gln Asp Glu Ile Ala Arg Ala Gln Leu Ala
485 490 495

Gly Arg Pro Ala Met Glu Ala Tyr Ile Asp His Gln Asp Ser Tyr Ser
500 505 510

Thr Asn Glu Val Ala Val Tyr Trp Asn Ser Pro Ala Val Phe Val Ile
515 520 525

Ala Ala Leu Leu Glu Ala Arg Gly Arg
530 535

<210> 5

<211> 692

<212> PRT

<213> Artificial Sequence

<220><223> CBM3-Cel9A cellulase fusion protein

<400> 5

Asn Leu Lys Val Glu Phe Tyr Asn Ser Asn Pro Ser Asp Thr Thr Asn
1 5 10 15

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

260	265	270	
His Gly Ala Cys His Thr Ser Asp Ala Lys Val Phe Gly Thr Glu Arg			
275	280	285	
Ala Leu Ala Cys Pro Gly Gly Trp His Asp Ala Gly Asp Tyr Gly Lys			
290	295	300	
Tyr Thr Val Pro Ala Ala Lys Ala Val Ala Asp Leu Leu Leu Ala His			
305	310	315	320
Glu Tyr Phe Pro Ala Ala Leu Ala His Val Arg Pro Met Arg Ser Val			
325	330	335	
His Arg Ala Pro His Leu Pro Pro Ala Leu Glu Val Ala Arg Glu Glu			
340	345	350	
Ile Ala Trp Leu Leu Thr Met Gln Asp Pro Ala Thr Gly Gly Val Tyr			
355	360	365	
His Lys Val Thr Thr Pro Ser Phe Pro Pro Leu Asp Thr Arg Pro Glu			
370	375	380	
Asp Asp Asp Ala Pro Leu Val Leu Ser Pro Ile Ser Tyr Ala Ala Thr			
385	390	395	400
Ala Thr Phe Cys Ala Ala Met Ala His Ala Ala Leu Val Tyr Arg Pro			
405	410	415	
Phe Asp Pro Ala Leu Ser Ser Cys Cys Ala Asp Ala Ala Arg Arg Ala			
420	425	430	
Tyr Ala Trp Leu Gly Ala His Glu Met Gln Pro Phe His Asn Pro Asp			
435	440	445	
Gly Ile Leu Thr Gly Glu Tyr Gly Asp Ala Glu Leu Arg Asp Glu Leu			
450	455	460	
Leu Trp Ala Ser Cys Ala Leu Leu Arg Met Thr Gly Asp Ser Ala Trp			
465	470	475	480
Ala Arg Val Cys Glu Pro Leu Leu Asp Leu Asp Leu Pro Trp Glu Leu			
485	490	495	
Gly Trp Ala Asp Val Ala Leu Tyr Gly Val Met Asp Tyr Leu Arg Thr			
500	505	510	

Pro Arg Ala Ala Val Ser Asp Asp Val Arg Asn Lys Val Lys Ser Arg
515 520 525
Leu Leu Arg Glu Leu Asp Ala Leu Ala Ala Met Ala Glu Ser His Pro
530 535 540
Phe Gly Ile Pro Met Arg Asp Asp Asp Phe Ile Trp Gly Ser Asn Met
545 550 555 560

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

625 630 635 640
Leu Gln Asp Glu Ile Ala Arg Ala Gln Leu Ala Gly Arg Pro Ala Met
645 650 655
Glu Ala Tyr Ile Asp His Gln Asp Ser Tyr Ser Thr Asn Glu Val Ala
660 665 670
Val Tyr Trp Asn Ser Pro Ala Val Phe Val Ile Ala Ala Leu Leu Glu
675 680 685

Ala Arg Gly Arg
690

<210> 6
<211> 712
<212> PRT

<213> Artificial Sequence
<220><223> CBM4-Cel9A cellulase fusion protein
<400> 6

Met Leu Glu Asp Asn Ser Ser Thr Leu Pro Pro Tyr Lys Asn Asp Leu
1 5 10 15

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

260	265	270	
Phe Phe Asp Tyr Gln Leu Cys Gly Val Val Leu Pro Glu Asp Glu Ala			
275	280	285	
Gly Pro Trp Ala His Gly Ala Cys His Thr Ser Asp Ala Lys Val Phe			
290	295	300	
Gly Thr Glu Arg Ala Leu Ala Cys Pro Gly Gly Trp His Asp Ala Gly			
305	310	315	320
Asp Tyr Gly Lys Tyr Thr Val Pro Ala Ala Lys Ala Val Ala Asp Leu			
325	330	335	
Leu Leu Ala His Glu Tyr Phe Pro Ala Ala Leu Ala His Val Arg Pro			
340	345	350	
Met Arg Ser Val His Arg Ala Pro His Leu Pro Pro Ala Leu Glu Val			
355	360	365	
Ala Arg Glu Glu Ile Ala Trp Leu Leu Thr Met Gln Asp Pro Ala Thr			
370	375	380	
Gly Gly Val Tyr His Lys Val Thr Thr Pro Ser Phe Pro Pro Leu Asp			
385	390	395	400
Thr Arg Pro Glu Asp Asp Asp Ala Pro Leu Val Leu Ser Pro Ile Ser			
405	410	415	
Tyr Ala Ala Thr Ala Thr Phe Cys Ala Ala Met Ala His Ala Ala Leu			
420	425	430	
Val Tyr Arg Pro Phe Asp Pro Ala Leu Ser Ser Cys Cys Ala Asp Ala			
435	440	445	
Ala Arg Arg Ala Tyr Ala Trp Leu Gly Ala His Glu Met Gln Pro Phe			
450	455	460	
His Asn Pro Asp Gly Ile Leu Thr Gly Glu Tyr Gly Asp Ala Glu Leu			
465	470	475	480
Arg Asp Glu Leu Leu Trp Ala Ser Cys Ala Leu Leu Arg Met Thr Gly			
485	490	495	
Asp Ser Ala Trp Ala Arg Val Cys Glu Pro Leu Leu Asp Leu Asp Leu			
500	505	510	

Pro Trp Glu Leu Gly Trp Ala Asp Val Ala Leu Tyr Gly Val Met Asp
515 520 525
Tyr Leu Arg Thr Pro Arg Ala Ala Val Ser Asp Asp Val Arg Asn Lys
530 535 540
Val Lys Ser Arg Leu Leu Arg Glu Leu Asp Ala Leu Ala Ala Met Ala
545 550 555 560

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

625 630 635 640
Val Ala Asp Asp Val Asp His Pro Val Pro Gly Met Val Val Gly Gly
645 650 655
Pro Asn Arg His Leu Gln Asp Glu Ile Ala Arg Ala Gln Leu Ala Gly
660 665 670
Arg Pro Ala Met Glu Ala Tyr Ile Asp His Gln Asp Ser Tyr Ser Thr
675 680 685
Asn Glu Val Ala Val Tyr Trp Asn Ser Pro Ala Val Phe Val Ile Ala
690 695 700

Ala Leu Leu Glu Ala Arg Gly Arg
705 710

<210> 7

<211> 722

<212> PRT

<213> Artificial Sequence

<220><223> CBM30-Cel9A cellulase fusion protein

<400> 7

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

245 250 255
 Leu Thr Val Gly Thr Leu Glu Ala Arg Val Val Ile His Arg Arg Ala
 260 265 270
 Tyr Arg Asp Val Leu Glu Ala Met Leu Arg Phe Phe Asp Tyr Gln Leu
 275 280 285
 Cys Gly Val Val Leu Pro Glu Asp Glu Ala Gly Pro Trp Ala His Gly
 290 295 300
 Ala Cys His Thr Ser Asp Ala Lys Val Phe Gly Thr Glu Arg Ala Leu
 305 310 315 320
 Ala Cys Pro Gly Gly Trp His Asp Ala Gly Asp Tyr Gly Lys Tyr Thr
 325 330 335
 Val Pro Ala Ala Lys Ala Val Ala Asp Leu Leu Leu Ala His Glu Tyr
 340 345 350
 Phe Pro Ala Ala Leu Ala His Val Arg Pro Met Arg Ser Val His Arg
 355 360 365
 Ala Pro His Leu Pro Pro Ala Leu Glu Val Ala Arg Glu Glu Ile Ala
 370 375 380
 Trp Leu Leu Thr Met Gln Asp Pro Ala Thr Gly Gly Val Tyr His Lys
 385 390 395 400
 Val Thr Thr Pro Ser Phe Pro Pro Leu Asp Thr Arg Pro Glu Asp Asp
 405 410 415
 Asp Ala Pro Leu Val Leu Ser Pro Ile Ser Tyr Ala Ala Thr Ala Thr
 420 425 430
 Phe Cys Ala Ala Met Ala His Ala Ala Leu Val Tyr Arg Pro Phe Asp
 435 440 445
 Pro Ala Leu Ser Ser Cys Cys Ala Asp Ala Ala Arg Arg Ala Tyr Ala
 450 455 460
 Trp Leu Gly Ala His Glu Met Gln Pro Phe His Asn Pro Asp Gly Ile
 465 470 475 480
 Leu Thr Gly Glu Tyr Gly Asp Ala Glu Leu Arg Asp Glu Leu Leu Trp
 485 490 495

Ala Ser Cys Ala Leu Leu Arg Met Thr Gly Asp Ser Ala Trp Ala Arg
500 505 510

Val Cys Glu Pro Leu Leu Asp Leu Asp Leu Pro Trp Glu Leu Gly Trp
515 520 525

Ala Asp Val Ala Leu Tyr Gly Val Met Asp Tyr Leu Arg Thr Pro Arg
530 535 540

Ala Ala Val Ser Asp Asp Val Arg Asn Lys Val Lys Ser Arg Leu Leu
545 550 555 560

Arg Glu Leu Asp Ala Leu Ala Ala Met Ala Glu Ser His Pro Phe Gly
565 570 575

Ile Pro Met Arg Asp Asp Asp Phe Ile Trp Gly Ser Asn Met Val Leu
580 585 590

Leu Asn Arg Ala Met Ala Phe Leu Leu Ala Glu Gly Val Gly Val Leu
595 600 605

His Pro Ala Ala His Thr Val Ala Gln Arg Ala Ala Asp Tyr Leu Phe
610 615 620

Gly Ala Asn Pro Leu Gly Gln Cys Tyr Val Thr Gly Phe Gly Gln Arg
625 630 635 640

Pro Val Arg His Pro His His Arg Pro Ser Val Ala Asp Asp Val Asp
645 650 655

His Pro Val Pro Gly Met Val Val Gly Gly Pro Asn Arg His Leu Gln
660 665 670

Asp Glu Ile Ala Arg Ala Gln Leu Ala Gly Arg Pro Ala Met Glu Ala
675 680 685

Tyr Ile Asp His Gln Asp Ser Tyr Ser Thr Asn Glu Val Ala Val Tyr
690 695 700

Trp Asn Ser Pro Ala Val Phe Val Ile Ala Ala Leu Leu Glu Ala Arg
705 710 715 720

Gly Arg

<210> 8

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> P1 primer

<400> 8

ttgtcatatg aatttgaagg ttgaattcta caacagcaat c 41

<210> 9

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> P2 primer

<400> 9

gatttaggca cacggctcgg gggttcttta cccatacaa g 41

<210> 10

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> P3 primer

<400> 10

gggtaaagaa ccccgagcc gtgtgcctaa atc 33

<210> 11

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> P1' primer

<400> 11

atgttcatat gttagaagat aattcttcga c 31

<210> 12

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> P2' primer

<400> 12

gatttaggca cacggctcgg aggctgcgga agtatatatt 40

<210> 13
 <211> 40
 <212> DNA
 <213> Artificial Sequence
 <220><223> P3' primer
 <400> 13
 aatatatact tccgcagcct ccgagccgtg tgcctaaatc 40
 <210> 14
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> P1'' primer
 <400> 14
 attgtcatat ggctcctgaa ggctacagga agc 33
 <210> 15
 <211> 41
 <212> DNA
 <213> Artificial Sequence
 <220><223> P2'' primer
 <400> 15
 gatttaggca cacggctcgg gattgcagga gcggactttt c 41
 <210> 16
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> P3'' primer
 <400> 16
 ccgctcctgc aatcccagac cgtgtgccta aatc 34
 <210> 17
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> P4 primer
 <400> 17

ccgatctcga gttaacggcc acgagcctcc ag

32

<210> 18

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Cel5H-F primer

<400> 18

cccggatcca attcttagcg gtggccagca a

31

<210> 19

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> Cel5H-R primer

<400> 19

cccgcctcgag ccagctacca aattgcaggg tgt

33

<210> 20

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> CbhA-F primer

<400> 20

ctagctagct atacttccgc agcctgat

28

<210> 21

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> CbhA-R primer

<400> 21

tatataagct taaccgcccgc gcggcggttcc cca

33