



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0139118
 (43) 공개일자 2014년12월04일

- (51) 국제특허분류(Int. Cl.)
 C12N 15/55 (2006.01) C12N 9/42 (2006.01)
 C12N 15/74 (2006.01) C12P 19/04 (2006.01)
 C02F 3/34 (2006.01) B09B 3/00 (2006.01)
- (21) 출원번호 10-2014-7030445
- (22) 출원일자(국제) 2013년03월12일
 심사청구일자 없음
- (85) 번역문제출일자 2014년10월29일
- (86) 국제출원번호 PCT/US2013/030571
- (87) 국제공개번호 WO 2013/148167
 국제공개일자 2013년10월03일
- (30) 우선권주장
 61/618,610 2012년03월30일 미국(US)
 61/660,556 2012년06월15일 미국(US)

- (71) 출원인
베레늄 코퍼레이션
 미국 캘리포니아주 92121 샌디에고 존 홉킨스 코
 트 3550
- (72) 발명자
탄 주키우
 미국 92121 캘리포니아주 샌디에고 존 홉킨스 코
 트 3550
베렛 케네쓰 이.
 미국 92121 캘리포니아주 샌디에고 존 홉킨스 코
 트 3550
 (뒷면에 계속)
- (74) 대리인
김성기, 김진희

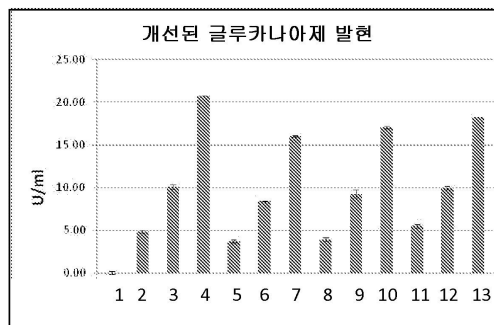
전체 청구항 수 : 총 44 항

(54) 발명의 명칭 **극한의 개정 조건 하에서 구아 파쇄액을 가수 분해하기 위한 셀룰라아제를 코딩하는 유전자**

(57) 요약

열 안정성 셀룰라아제를 코딩하고, 이 셀룰라아제의 발현을 증가시키는 폴리뉴클레오티드 서열과, 이러한 열 안
 정성 셀룰라아제를 포함하는 수압 파쇄 조성물이 제공된다.

대표도 - 도2



설명

- 레인 1 = 음성 대조군
- 레인 2, 5, 8, 11 = 서열 번호 3의 서열
- 레인 3, 6, 9, 12 = 서열 번호 4의 서열
- 레인 4, 7, 10, 13 = 서열 번호 1의 서열

(72) 발명자

데이븐포트 애드리엔 휴스턴

미국 92121 캘리포니아주 샌디에고 존 홉킨스 코트
3550

휘플 로렌스 이.

미국 92121 캘리포니아주 샌디에고 존 홉킨스 코트
3550

얼비나 휴고 디.

미국 92121 캘리포니아주 샌디에고 존 홉킨스 코트
3550

장 빈

미국 92121 캘리포니아주 샌디에고 존 홉킨스 코트
3550

월 마크 에이.

미국 92121 캘리포니아주 샌디에고 존 홉킨스 코트
3550

특허청구의 범위

청구항 1

서열 번호 1의 뉴클레오티드 서열.

청구항 2

제1항에 있어서, 상기 서열은 단백질을 코딩하는 것인 뉴클레오티드 서열.

청구항 3

T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C, A6C, G6C, A9C, G9C, A15G, C15G, T22C, G22C, A24T, C24T, T33C, G33C, T39C, G39C, T40C, G40C, T42C, G42C, T54C, G54C, T57C, G57C, A66C, G66C, C81A, T81A, T84C, G84C 또는 이의 임의의 조합으로부터 선택되는 하나 이상의 돌연 변이를 포함하는, 서모토가 마리티마(*Thermotoga maritima*) 유래 셀룰라아제를 코딩하는 뉴클레오티드 서열.

청구항 4

T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C, A6C, G6C, A9C, G9C, A15G, C15G, T22C, G22C, A24T, C24T, T33C, G33C, T39C, G39C, T40C, G40C, T42C, G42C, T54C, G54C, T57C, G57C, A66C, G66C, C81A, T81A, T84C, G84C 또는 이의 임의의 조합으로부터 선택되는 하나 이상의 돌연 변이를 포함하는 서열 번호 3의 뉴클레오티드 서열.

청구항 5

제3항 또는 제4항에 있어서, 상기 하나 이상의 돌연 변이는 침묵하는 것인 뉴클레오티드 서열.

청구항 6

제3항 또는 제4항에 있어서, 상기 하나 이상의 돌연 변이는, 하나 이상의 돌연 변이가 결실되고 그 외에는 제3항 또는 제4항의 뉴클레오티드 서열과 상이하지 않은 뉴클레오티드 서열보다 더욱 높은 수준으로 상기 셀룰라아제의 발현을 유도하는 하나 이상의 돌연 변이를 보유하는 뉴클레오티드 서열을 생성하는 것인 뉴클레오티드 서열.

청구항 7

하나 이상의 돌연변이를 가지는 서모토가 마리티마 유래 뉴클레오티드 서열로서, 서모토가 마리티마 게놈 서열에 비해 상기 뉴클레오티드 서열에 의해 코딩된 단백질의 발현 수준을 증가시키는 뉴클레오티드 서열.

청구항 8

제7항에 있어서, 상기 돌연 변이는 침묵하는 것인 뉴클레오티드 서열.

청구항 9

서열 번호 2의 폴리펩티드를 코딩하는 제1 뉴클레오티드 서열로서, 상기 뉴클레오티드 서열은, 상기 단백질의 발현 수준이 제2 뉴클레오티드 서열에 의해 코딩되는 상기 단백질의 발현 수준에 비해서 증가하도록 서열 번호 2의 서열을 코딩하는 상기 제2 서열에 대해서 돌연 변이된 것인 제1 뉴클레오티드 서열.

청구항 10

서열 번호 2의 서열과 적어도 50%, 60%, 70%, 80%, 90%, 95%, 97%, 98%, 99% 또는 100% 동일한 단백질 또는 이의 단편을 코딩하는 뉴클레오티드 서열로서, 상기 뉴클레오티드 서열은 T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C, A6C, G6C, A9C, G9C, A15G, C15G, T22C, G22C, A24T, C24T, T33C, G33C, T39C, G39C, T40C, G40C, T42C, G42C, T54C, G54C, T57C, G57C, A66C, G66C, C81A, T81A, T84C, G84C, 또는 이의 임의의 조합으로부터 선택되는 하나 이상의 돌연 변이를 포함하는 것인 뉴클레오티드 서열.

열.

청구항 11

제1항 내지 제10항 중 어느 하나의 항에 있어서, 상기 뉴클레오티드 서열은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩하고, 이 폴리펩티드는 재조합 세균 발현계 내에서 생산되는 것인 뉴클레오티드 서열.

청구항 12

제11항에 있어서, 상기 세균 발현계는 그람-음성 세균 발현계인 뉴클레오티드 서열.

청구항 13

제12항에 있어서, 그람-음성 세균은 슈도모나스(*Pseudomonas*), 이.콜라이(*E.coli*), 랄스토니아(*Ralstonia*) 또는 카울로박터(*Caulobacter*) 발현계인 그람-음성 세균 발현계.

청구항 14

제12항에 있어서, 슈도모나스 발현계는 슈도모나스 플루오레센스(*Pseudomonas fluorescens*) 발현계인 그람-음성 세균 발현계.

청구항 15

제12항에 있어서, 셀룰라아제는 적어도 1.0g/L, 2.0g/L, 3.0g/L, 4.0g/L, 5.0g/L, 6.0g/L, 7.0g/L, 8.0g/L, 9.0g/L, 10.0g/L, 11.0g/L, 12.0g/L, 13.0g/L, 14.0g/L, 15.0g/L, 16.0g/L, 17.0g/L, 18.0g/L, 19.0g/L, 20.0g/L, 21.0g/L, 22.0g/L, 23.0g/L, 24.0g/L, 25.0g/L, 26.0g/L, 27.0g/L, 28.0g/L, 29.0g/L, 30.0g/L, 31.0g/L, 32.0g/L, 33.0g/L, 34.0g/L, 또는 35.0g/L로 생산되는 것인 발현계.

청구항 16

서열 번호 2의 폴리펩티드로서, 아미노산 서열이 신호 서열, 프로프로테인 서열(proprotein sequence), 프로모터 서열 또는 이의 임의의 조합을 포함하지 않는 것인 폴리펩티드.

청구항 17

서열 번호 2의 폴리펩티드로서, 서열이 이중 서열을 추가로 포함하는 것인 폴리펩티드.

청구항 18

제17항에 있어서, 상기 이중 서열은 신호 서열, 태그, 에피토프, 프로모터 서열, N-말단 연장부, C-말단 연장부 및 이의 임의의 조합으로 이루어진 군으로부터 선택되는 것인 폴리펩티드.

청구항 19

서열 번호 2의 폴리펩티드를 포함하고, 제2 효소를 추가로 포함하는 조성물.

청구항 20

제19항에 있어서, 상기 제2 효소는 락타아제, 리파아제, 프로테아제, 카탈라아제, 크실라나아제, 셀룰라아제, 글루카나아제, 만난아제, 아밀라아제, 아미다아제, 에폭시드 하이드롤라아제, 에스테라아제, 포스포리파아제, 트랜스아민아제, 아민 옥시다아제, 셀로비오하이드롤라아제, 암모니아 분해 효소 또는 이의 임의의 조합으로 이루어진 군으로부터 선택되는 것인 조성물.

청구항 21

서열 번호 1의 서열을 포함하는 핵산 서열을 가지는 재조합 또는 합성의 단리된 뉴클레오티드로서, 상기 핵산 서열은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩하는 것인, 재조합 또는 합성의 단리된 뉴클레오티드.

청구항 22

서열 번호 1의 핵산 서열을 포함하는 재조합 또는 합성의 단리된 뉴클레오티드로서, 상기 핵산 서열은 셀룰라아

제 활성을 가지는 폴리펩티드를 코딩하고, 상기 폴리펩티드는 서열 번호 2의 아미노산 서열 또는 이의 효소 활성 단편을 포함하는 것인, 재조합 또는 합성의 단리된 뉴클레오티드.

청구항 23

셀룰라아제 활성을 가지는 폴리펩티드를 코딩하는 서열 번호 1의 서열을 포함하는 재조합 또는 합성의 단리된 핵산 서열로서, 상기 폴리펩티드는 서열 번호 2의 아미노산 서열을 포함하고, 상기 폴리펩티드는 재조합 슈도모나스 플루오레센스 발현계 내에서 생산되는 것인, 재조합 또는 합성의 단리된 핵산 서열.

청구항 24

고분자 증점제, 계면활성제, 열 안정화제, 그리고 초고온성 세균으로부터 유래하는 야생형 셀룰라아제 또는 이의 돌연 변이된 변이체를 포함하는 효소성 분해 물질(enzyme breaker)을 포함하는 조성물.

청구항 25

제24항에 있어서, 상기 증점제는 선형 구아, 가교형 구아 또는 이의 혼합물을 포함하는 구아 겔인 조성물.

청구항 26

제25항에 있어서, 상기 효소성 분해 물질은 구아 겔 내 β -1,4 글리코시드 결합들을 특이적으로 가수 분해하는 것인 조성물.

청구항 27

제25항에 있어서, 상기 효소성 분해 물질은 구아 겔 내 α -1,6 글리코시드 결합들을 특이적으로 가수 분해하지 않는 것인 조성물.

청구항 28

제24항 내지 제27항 중 어느 하나의 항에 있어서, 상기 효소성 분해 물질은 약 275°F 이하의 온도에서 구아 겔 내 β -1,4 글리코시드 결합들을 가수 분해하는 능력을 유지하는 것인 조성물.

청구항 29

제24항 내지 제28항 중 어느 하나의 항에 있어서, 상기 효소성 분해 물질은 pH 약 11 이하에서 구아 겔 내 β -1,4 글리코시드 결합들을 가수 분해하는 능력을 유지하는 것인 조성물.

청구항 30

제24항 내지 제29항 중 어느 하나의 항에 있어서, 상기 효소성 분해 물질은 야생형 셀룰라아제에 대해 12개의 돌연 변이를 포함하는 돌연 변이된 변이체인 조성물.

청구항 31

제30항에 있어서, 상기 12개의 돌연 변이는 F38Y, Y61Q, M69E, D70P, R71S, I94Q, I166V, S183R, S191A, E212P, L231V, M276A, E277S, R280G, T297P 및 T301Q로 이루어진 군으로부터 선택되는 것인 조성물.

청구항 32

제24항에 있어서, 상기 효소성 분해 물질은 서열 번호 2의 서열을 가지는 것인 조성물.

청구항 33

제24항에 있어서, 상기 효소성 분해 물질은 서열 번호 1의 서열을 가지는 폴리뉴클레오티드에 의해 코딩되는 것인 조성물.

청구항 34

제24항 내지 제29항 중 어느 하나의 항에 있어서, 상기 효소성 분해 물질은 야생형 셀룰라아제의 돌연 변이된 변이체로서, 용점이, pH 약 6.5에서 야생형 셀룰라아제의 용점보다 20°F 이상 높으며, pH 약 10.5에서 야생형

셀룰라아제의 용점보다 10°F 이상 높은 것인 조성물.

청구항 35

제24항에 있어서, 상기 효소성 분해 물질은 서열 번호 3의 서열에 의해 코딩되는 셀룰라아제의 돌연 변이체인 조성물.

청구항 36

β-1,4 글리코시드 결합들을 포함하는 다당체 겔의 점도를 감소시키는 방법으로서, 셀룰라아제가 다당체 겔 내 β-1,4 글리코시드 결합들을 가수 분해하기에 충분한 시간 동안, 허용된 조건 하에서 다당체 겔과 셀룰라아제 변이체를 접촉시킴으로써, 구아 겔의 점도를 감소시키는 단계를 포함하며, 이때 셀룰라아제 변이체는 초고온성 세균으로부터 유래하는 야생형 셀룰라아제에 비해 12개 이상의 돌연 변이를 포함하고, 셀룰라아제 변이체는 야생형 셀룰라아제에 비해 증가된 온도 관용성 및 pH 관용성을 나타내는 것인 방법.

청구항 37

제36항에 있어서, 상기 12개 이상의 돌연 변이는, 환원성 재편성(reductive reassortment), 재조합 및 선별의 반복 주기를 포함하는 유전자 위치 포화 돌연 변이 유발법(Gene Site Saturation Mutagenesis)에 의해 형성된 것인 방법.

청구항 38

제36항에 있어서, 상기 다당체 겔은 선형 구아, 가교형 구아 또는 이의 혼합물을 포함하는 것인 방법.

청구항 39

제36항에 있어서, 상기 셀룰라아제 변이체는 서열 번호 2의 서열을 가지는 것인 방법.

청구항 40

제36항에 있어서, 상기 셀룰라아제 변이체는 서열 번호 1의 서열을 가지는 폴리뉴클레오티드에 의해 코딩되는 것인 방법.

청구항 41

제36항에 있어서, 상기 허용된 조건은 약 180~약 275°F의 온도 범위를 포함하는 것인 방법.

청구항 42

제36항에 있어서, 상기 허용된 조건은 약 9~약 11의 pH 범위를 포함하는 것인 방법.

청구항 43

지하 형성물(subterranean formation)의 처리 방법으로서,
 다당체 겔 및 셀룰라아제 변이체를 포함하는 파쇄액을 지하 형성물에 도입하는 단계; 셀룰라아제 변이체가 상기 겔 내 α-1,6 글리코시드 결합들을 가수 분해하지 않고 β-1,4 글리코시드 결합들 중 적어도 일부를 가수 분해하기에 충분한 시간 동안, 허용된 조건 하에서 상기 겔과 셀룰라아제 변이체를 반응시킴으로써, 점도 감소 파쇄액을 형성하는 단계를 포함하고, 상기 셀룰라아제 변이체는 초고온성 세균으로부터 유래하는 모 야생형 셀룰라아제에 비해 12개 이상의 돌연 변이를 포함하고, 이 셀룰라아제 변이체는 야생형 셀룰라아제에 비해 증가된 온도 관용성 및 pH 관용성을 나타내는 것인 방법.

청구항 44

제36항에 있어서, 상기 셀룰라아제 변이체에 의해 형성된 점도 감소 파쇄액은, 화학적 분해 물질에 의해 형성된 유사 점도 감소 파쇄액보다 적은 잔류물을 포함하고, 상기 셀룰라아제 변이체와 화학적 분해 물질의 사용량이 점도에 있어 실질적으로 동일한 감소를 제공하는 것인 방법.

명세서

기술분야

관련 출원들

[0001]

[0002] 본 출원은 35 U.S.C. § 119(e) 하에, 그 자체로서 각각 본원에 참고용으로 인용되어 있는 미국 가 명세서 특허 출원 제61/618,610호(2012년 3월 30일 출원) 및 동 제61/660,556호(2012년 6월 15일 출원)의 우선권의 이익을 주장한다.

본 발명의 분야

[0003]

[0004] 셀룰라아제를 코딩하는 폴리뉴클레오티드 서열들이 제공된다. 특히 본 출원에 제공된 서열들은 특이적이고, 열 안정성이며, 열 내성임과 동시에 압력 안정성인 셀룰라아제, 예를 들어 극한의 갱정(we'll) 조건 하에서 구아 파쇄액을 가수 분해하기 위한 셀룰라아제의 발현을 증가시킬 수 있다.

서열 목록

[0005]

[0006] 본 출원은 USPTO EFS-WEB 서버를 통해 MPEP § 502.05에 따라서 승인 및 제시된 바와 같이 전자 출원되었으며, 본 전자 출원은 전자 제출된 서열 목록을 포함하는데; 본 서열 목록의 전체 내용은 본 출원의 명세서에 참고용으로 인용되어 있다. 본 서열 목록은 다음과 같이 전자 제출된 ASCII(.txt) 텍스트 파일인 것으로 확인된다.

파일명	생성일	크기
D2490_SEQLISTING	2013년 3월 9일	40.4 kb

[0007]

배경 기술

[0008]

0-글리코실 하이드롤라아제(EC 3.2.1.-)는, 2개 이상의 탄수화물 간 또는 탄수화물과 비 탄수화물 부 간 글리코시드 결합을 가수 분해하는 자연 발생 효소의 널리 보급된 군이다. 글리코실 하이드롤라아제(또는 글리코실라아제)의 국제 생화학 및 분자 생물학 연합회(IUBMB) 효소 명명법은, 원칙적으로 상기 글리코실 하이드롤라아제의 기질을 바탕으로 하며, 때로는 상기 글리코실 하이드롤라아제의 분자적 기작을 바탕으로 하기도 한다 [Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB), Accessed 10/24/2011].

[0009]

IUBMB 효소 명명법 EC 3.2.1.4은 "셀룰라아제"로 칭하여지는 글리코실라아제 류 효소의 하위 군을 명명한 바 있다. 이 군에 속하는 효소들에 사용되는 기타 명칭들로서는 다음과 같은 것들을 포함한다: 엔도글루카나아제, 엔도-1,4-베타-글루카나아제, 카복시메틸 셀룰라아제 및 베타-1,4-글루카나아제. 상기 군에 속하는 효소들에 의해 촉매화되는 반응으로서는 셀룰로스, 리케닌 및 곡물의 베타-D-글루칸(예를 들어 보리의 베타-글루칸) 내 1,4-베타-D-글리코시드 결합의 내부 가수 분해가 있다. 본원에 개시된 본 발명의 셀룰라아제의 주된 활성은 보리의 베타 글루칸 및 카복시메틸 셀룰로스의 내부 가수 분해이므로, 이의 명칭은 IUBMB 효소 명명법 EC 3.2.1.4.에 적절히 부합되는 것으로 생각된다.

[0010]

글리코실 하이드롤라아제의 대안적 분류법은 아미노산 서열의 유사성을 바탕으로 한다[Henrissat, B. UniProt 10/26/2011에 승인]. 이와 같은 분류 방식에 따르면, 글리코실 하이드롤라아제는 70개 이상의 군으로 구분될 수 있다. 본원에 개시된 본 발명의 셀룰라아제의 1차 아미노산 서열과, 공표된 데이터베이스에 포함되어 있는 기타 글리코실 하이드롤라아제의 서열을 비교한 결과를 바탕으로 하였을 때, 본원에 개시된 본 발명의 셀룰라아제는 글리코실 하이드롤라아제 제5군으로 분류될 수 있다. 상기 군은, 주된 촉매 활성이 셀룰로스 기질 내 베타-1,4-글리코시드 결합의 내부 가수 분해인 엔도글루카나아제 20개 이상을 포함한다(IUBMB 효소 명명법 EC 3.2.1.4). 이와 같은 제2의 효소 분류 방법이 이용되면, 본원에 개시된 본 발명의 셀룰라아제가 IUBMB 효소 명명법 EC 3.2.1.4에 부합되는 것으로 생각될 것이라는 결론이 추가로 뒷받침된다.

[0011]

셀룰라아제는 다양한 산업적 및 상업적 용도, 예를 들어 오일 및 가스 탐사, 식품 및 음료 생산, 음료 또는 연료로서 사용될 알코올의 생산, 예를 들어 양조, 에탄올, 와인, 조미료, 향수, 직물, 세제, 종이, 펄프의 생산에 사용되고 있으며, 환경 및 농업 분야에 사용될 수 있을 뿐만 아니라, 연구 목적으로도 사용되고 있으나, 이와 같은 용도에 한정되어 사용되는 것은 아니다[Rebecca S. Bryant, Erle C. Donaldson, Teh Fu Yen, George V. Chilingarian, Chapter 14 Microbial Enhanced Oil Recovery, In: Erle C. Donaldson, George V. Chilingarian and Teh Fu Yen, Editor(s), *Developments in Petroleum Science*, Elsevier, 1989, Volume 17, Part B, Pages 423-450; M. Karmakar and R.R. Ray, 2011. Current Trends in Research and Application of Microbial Cellulases. *Research Journal of Microbiology*, 6: 41-53].

[0012] 효소성 분해 물질(enzyme breaker)은 1990년대초 이래로 고분자 증점제를 가수 분해하기 위해 수계 파쇄액에 성공적으로 사용되어 왔다[Brannon HD, Tjon-Joe-Pin RM, Carman PS, Wood WD, "Enzyme breaker technology: a decade of improved well stimulation", paper SPE 84213 presented at the SPE Annual Technical Conference and Exhibition in Denver, Colorado, USA, 5-8 October 2003].

발명의 내용

[0013] **발명의 개요**

[0014] 효소는 다수의 파쇄 작업시 처할 수 있는 고온 및 고 알칼리 조건에 민감한 3차원 구조를 가지는 단백질이다. 험하다고 볼 수 있는 땅속 구멍의 환경(예를 들어 고온 및 높은 pH) 하에서도 촉매 활성을 유지하는 효소가 동정 및/또는 조작되면, 수계 파쇄시 효소성 분해 물질의 활용도가 더욱 넓어질 것이다.

[0015] 고분자 증점제, 계면활성제, 열 안정화제 및 효소성 분해 물질(초고온성 세균으로부터 유래하는 야생형 셀룰라아제 또는 이의 돌연 변이된 변이체 포함)을 포함하는 조성물이 본원에 개시되어 있다.

[0016] 몇몇 구체예에서, 증점제는 선형 구아, 가교형 구아 또는 이의 조합을 포함하는 구아 겔이다. 몇몇 구체예에서, 효소성 분해 물질은 구아 겔 내 β -1,4 글리코시드 결합들을 특이적으로 가수 분해한다. 몇몇 구체예에서, 효소성 분해 물질은 구아 겔 내 α -1,6 글리코시드 결합들을 특이적으로 가수 분해하지 않는다. 몇몇 구체예에서, 효소성 분해 물질은 약 275°F 이하의 온도에서 구아 겔 내 β -1,4 글리코시드 결합을 가수 분해하는 능력을 가진다. 몇몇 구체예에서, 효소성 분해 물질은 pH 약 11 이하에서 구아 겔 내 β -1,4 글리코시드 결합을 가수 분해하는 능력을 가진다.

[0017] 몇몇 구체예에서, 본 발명의 조성물 중 효소성 분해 물질은 야생형 셀룰라아제에 비하여 12개의 돌연 변이를 포함하는 돌연 변이된 변이체이다. 상기 돌연 변이 12개는 F38Y, Y61Q, M69E, D70P, R71S, I94Q, I166V, S183R, S191A, E212P, L231V, M276A, E277S, R280G, T297P 및 T301Q로 이루어진 군으로부터 선택될 수 있다. 몇몇 구체예에서, 효소성 분해 물질은 서열 번호 2의 서열을 가진다. 몇몇 구체예에서, 효소성 분해 물질은 서열 번호 1의 서열을 가지는 폴리뉴클레오티드에 의해 코딩된다. 몇몇 구체예에서, 효소성 분해 물질은 야생형 셀룰라아제의 돌연 변이된 변이체로서, 융점이, pH 약 6.5에서 야생형 셀룰라아제의 융점보다 20°F 이상 높으며, pH 약 10.5에서 야생형 셀룰라아제의 융점보다 10°F 이상 높다.

[0018] β -1,4-글루코시드 결합을 포함하는 다당체 겔의 점도를 감소시키기 위한 방법이 개시되어 있다. 상기 방법은, 셀룰라아제가 다당체 겔 내 β -1,4 글리코시드 결합들을 가수 분해하기에 충분한 시간 동안, 허용된 조건 하에서 다당체 겔과 셀룰라아제 변이체를 접촉시킴으로써 구아 겔의 점도를 감소시키는 단계를 포함한다. 셀룰라아제 변이체는 초고온성 세균으로부터 유래하는 야생형 셀룰라아제에 비해 12개 이상의 돌연 변이를 포함하며, 셀룰라아제 변이체는 야생형 셀룰라아제에 비해 증가된 온도 관용성 및 pH 관용성을 나타낸다.

[0019] 본 발명의 방법에 관한 몇몇 구체예에서, 셀룰라아제 변이체내 돌연 변이 12개 이상은, 환원성 재편성(reductive reassembly), 재조합 및 선별로 이루어진 주기가 반복 실시되는 유전자 위치 포화 돌연 변이 유발법에 의해 발생되었다.

[0020] 본 발명의 방법에 관한 몇몇 구체예에서, 다당체 겔은 선형 구아, 가교형 구아 또는 이의 조합을 포함한다.

[0021] 본 발명의 방법에 관한 몇몇 구체예에서, 셀룰라아제 변이체는 서열 번호 2의 서열을 가진다. 본 발명의 방법에 관한 몇몇 구체예에서, 셀룰라아제 변이체는 서열 번호 1의 서열을 가지는 폴리뉴클레오티드에 의해 코딩된다.

[0022] 본 발명의 방법에 관한 몇몇 구체예에서, 허용된 조건은 온도 범위 약 180~약 275°F를 포함한다. 본 발명의 방법에 관한 몇몇 구체예에서, 허용된 조건은 pH 범위 약 9~약 11을 포함한다.

[0023] 지하 형성물(subterranean formation)의 처리 방법도 개시되어 있다. 본 방법은 다음과 같은 단계들을 포함한다: 지하 형성물에 다당체 겔 및 셀룰라아제 변이체를 포함하는 파쇄액을 도입하는 단계; 그리고 셀룰라아제 변이체가 다당체 겔 내 α -1,6 글리코시드 결합들을 가수 분해하지 않고 β -1,4 글리코시드 결합들 중 적어도 일부를 가수 분해하기에 충분한 시간 동안, 허용된 조건 하에서 상기 겔과 셀룰라아제 변이체를 반응시킴으로써, 점도 감소 파쇄액을 생산하는 단계. 상기 셀룰라아제 변이체는 초고온성 세균으로부터 유래하는 모 야생형 셀룰라아제에 비해 12개 이상의 돌연 변이를 포함할 수 있으며, 이 셀룰라아제 변이체는 야생형 셀룰라아제에 비해 증가된 온도 관용성 및 pH 관용성을 나타낼 수 있다.

[0024] 본 발명의 방법에 관한 몇몇 구체예에서, 셀룰라아제 변이체에 의해 형성된 점도 감소 파쇄액은, 화학적 분해

물질에 의해 형성된 유사 점도 감소 파쇄액보다 적은 잔류물을 포함하고, 상기 셀룰라아제 변이체와 화학적 분해 물질의 사용량이 점도에 있어 실질적으로 동일한 감소를 제공한다.

- [0025] 본 발명의 방법에 관한 몇몇 구체예에서, 셀룰라아제 변이체에 의해 형성된 점도 감소 파쇄액은, 화학적 분해 물질에 의해 형성된 유사 점도 감소 파쇄액보다 큰 전도도를 유지하고, 상기 셀룰라아제 변이체와 화학적 분해 물질의 사용량이 점도에 있어 실질적으로 동일한 감소를 제공한다.
- [0026] 다른 구체예에서, 본 발명은 서열 번호 1의 서열을 포함하는데, 이 경우 상기 서열은 단백질을 코딩한다. 추가의 구체예에서, 본 발명은 T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C, A6C, G6C, A9C, G9C, A15G, C15G, T22C, G22C, A24T, C24T, T33C, G33C, T39C, G39C, T40C, G40C, T42C, G42C, T54C, G54C, T57C, G57C, A66C, G66C, C81A, T81A, T84C, G84C 또는 이의 임의의 조합으로부터 선택되는 돌연 변이 하나 이상을 포함하는 서모토가 마리티마(*Thermotoga maritima*) 유래 셀룰라아제 코딩 뉴클레오티드 서열이나, 또는 서열 번호 3의 서열을 포함하는데, 이 경우 상기 돌연 변이들 중 임의의 것은 임의로 침묵하기도 한다. 본 발명의 추가의 구체예에서, 이와 같은 침묵 돌연 변이 하나 이상이 포함된 뉴클레오티드 서열은 이와 같은 돌연 변이 하나 이상이 결실된 뉴클레오티드 서열보다 상기 셀룰라아제를 더욱 높은 수준으로 발현한다.
- [0027] 본 발명의 다른 구체예에서, 본 발명은 돌연 변이를 하나 이상 가지며, 코딩되는 단백질의 발현 수준이 서모토가 마리티마 야생형 게놈 서열에 의해 코딩된 단백질의 발현 수준과 비교되게 증가하게 되는, 서모토가 마리티마 유래 뉴클레오티드 서열을 포함하는데, 이 경우 상기 돌연 변이(들)는 임의로 침묵하기도 한다.
- [0028] 본 발명의 다른 구체예에서, 본 발명은 서열 번호 2의 폴리펩티드를 코딩하는 제1 뉴클레오티드 서열을 포함하는데, 이 경우 상기 제1 뉴클레오티드 서열은 서열 번호 2의 폴리펩티드를 코딩하는 제2 서열에 대하여 돌연 변이된 관계로, 상기 제1 뉴클레오티드 서열에 의해 코딩된 단백질의 발현 수준은, 상기 제2 뉴클레오티드 서열에 의해 코딩된 단백질의 발현 수준에 비해 증가하게 된다.
- [0029] 서열 번호 2의 폴리펩티드를 코딩하는 제1 뉴클레오티드 서열은 서열 번호 2의 폴리펩티드를 코딩하는 제2 서열에 대하여 돌연 변이된 관계로, 상기 제1 뉴클레오티드 서열에 의해 코딩된 단백질의 발현 수준은, 상기 제2 뉴클레오티드 서열에 의해 코딩된 단백질의 발현 수준에 비해 증가하게 된다.
- [0030] 본 발명의 또 다른 구체예에서, 본 발명은 서열 번호 2의 폴리펩티드와 50%, 60%, 70%, 80%, 90%, 95%, 97%, 98%, 99% 이상 또는 100% 동일한 단백질 또는 이의 단편을 코딩하는 뉴클레오티드 서열을 포함하는데, 이 경우 상기 뉴클레오티드 서열은 T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C, A6C, G6C, A9C, G9C, A15G, C15G, T22C, G22C, A24T, C24T, T33C, G33C, T39C, G39C, T40C, G40C, T42C, G42C, T54C, G54C, T57C, G57C, A66C, G66C, C81A, T81A, T84C, G84C, 또는 이의 임의의 조합으로부터 선택되는 하나 이상의 돌연 변이를 포함한다.
- [0031] 본 발명의 다른 구체예에서, 본 발명의 단백질들 중 임의의 것은 세균 발현계에서 발현되는데, 이 경우 세균 발현계는 그람-음성 세균 발현계, 예를 들어 슈도모나스(*Pseudomonas*), 이.콜라이(*E. coli*), 랄스토니아(*Ralstonia*) 또는 카울로박터(*Caulobacter*) 발현계이다.
- [0032] 본 발명의 다른 구체예에서, 본 발명의 셀룰라아제는 적어도 1.0g/L, 2.0g/L, 3.0g/L, 4.0g/L, 5.0g/L, 6.0g/L, 7.0g/L, 8.0g/L, 9.0g/L, 10.0g/L, 11.0g/L, 12.0g/L, 13.0g/L, 14.0g/L, 15.0g/L, 16.0g/L, 17.0g/L, 18.0g/L, 19.0g/L, 20.0g/L, 21.0g/L, 22.0g/L, 23.0g/L, 24.0g/L, 25.0g/L, 26.0g/L, 27.0g/L, 28.0g/L, 29.0g/L, 30.0g/L, 31.0g/L, 32.0g/L, 33.0g/L, 34.0g/L 또는 35.0g/L로 발현된다.
- [0033] 본 발명의 다른 구체예에서, 본 발명의 셀룰라아제는 제2 효소와 합하여지는데, 이 경우 상기 제2 효소는 다음과 같은 것들로 이루어진 군으로부터 선택된다: 락타아제, 리파아제, 프로테아제, 카탈라아제, 크실라나아제, 셀룰라아제, 글루카나아제, 만난아제, 아밀라아제, 아미다아제, 에폭시드 하이드롤라아제, 에스테라아제, 포스포리파아제, 트랜스아민아제, 아민 옥시다아제, 셀로비오하이드롤라아제, 암모니아 분해 효소 또는 이의 임의의 조합.
- [0034] 본 발명의 다른 구체예에서, 본 발명은 서열 번호 1의 서열을 포함하는 핵산 서열을 가지는 재조합 또는 합성의 단리된 뉴클레오티드를 포함하는데, 여기서 상기 핵산 서열은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩한다.
- [0035] 본 발명의 다른 구체예에서, 본 발명은 서열 번호 1의 핵산 서열을 포함하는 재조합 또는 합성의 단리된 뉴클레

오티드를 포함하는데, 여기서 상기 핵산 서열은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩하고, 상기 폴리펩티드는 서열 번호 2의 아미노산 서열 또는 이의 효소 활성 단편을 포함한다.

- [0036] 본 발명의 다른 구체예에서, 본 발명은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩하는 서열 번호 1의 서열을 포함하는 재조합 또는 합성의 단리된 핵산 서열을 포함하는데, 여기서 상기 폴리펩티드는 서열 번호 2의 아미노산 서열을 포함하고, 상기 폴리펩티드는 재조합 슈도모나스 플루오레센스(*Pseudomonas fluorescens*) 발현계 내에서 생산된다.
- [0037] 본 발명의 다른 구체예에서, 본 발명은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩하는 뉴클레오티드 서열을 포함하는데, 여기서 상기 폴리펩티드는 재조합 세균 발현계 내에서 생산된다.
- [0038] 본 발명의 다른 구체예에서, 본 발명은 서열 번호 1의 서열에 의해 코딩되는 폴리펩티드를 포함하는 조성물을 포함하는데, 여기서 상기 조성물은 임의로 제2 효소 하나 이상을 추가로 포함하기도 하며, 이때 제2 효소는 임의로 다음과 같은 것들로 이루어진 군으로부터 선택되기도 한다: 락타아제, 리파아제, 프로테아제, 카탈라아제, 크실라나아제, 셀룰라아제, 글루카나아제, 만난아제, 아밀라아제, 아미다아제, 에폭시드 하이드롤라아제, 에스테라아제, 포스포리파아제, 트랜스아민아제, 아민 옥시다아제, 셀로비오하이드롤라아제, 암모니아 분해 효소 또는 이의 임의의 조합.
- [0039] 본 발명의 다른 구체예에서, 본 발명은 서열 번호 1의 서열을 포함하는 핵산 서열을 가지는 재조합 또는 합성의 단리된 뉴클레오티드를 포함하는데, 여기서 상기 핵산 서열은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩한다.
- [0040] 본 발명의 다른 구체예에서, 본 발명은 서열 번호 1의 핵산 서열을 포함하는 재조합 또는 합성의 단리된 뉴클레오티드를 포함하는데, 여기서 상기 핵산 서열은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩하며, 상기 폴리펩티드는 서열 번호 2의 아미노산 서열 또는 이의 효소 활성 단편을 포함한다. 추가의 구체예에서, 서열 번호 1의 서열을 포함하는 재조합 또는 합성의 단리된 핵산 서열은 셀룰라아제 활성을 가지는 폴리펩티드를 코딩하는데, 여기서 상기 폴리펩티드는 서열 번호 2의 아미노산 서열을 포함하고, 상기 폴리펩티드는 재조합 슈도모나스 플루오레센스 발현계 내에서 생산된다.
- [0041] 본 발명의 다른 구체예에서, 본 발명은 고분자 증점제, 계면활성제, 열 안정화제 및 효소성 분해 물질(초고온성 세균으로부터 유래하는 야생형 셀룰라아제 또는 이의 돌연 변이된 변이체 포함)을 포함하는 조성물을 포함한다. 상기 조성물에 관한 추가의 구체예에서, 증점제는 선형 구아 겔, 가교형 구아 겔 또는 이의 조합을 포함하는 구아 겔이다. 본 발명의 조성물에 관한 추가의 구체예에서, 효소성 분해 물질은 구아 겔 내 β -1,4 글리코시드 결합을 특이적으로 가수 분해한다. 본 발명의 조성물에 관한 추가의 구체예에서, 효소성 분해 물질은 구아 겔 내 α -1,6 글리코시드 결합을 특이적으로 가수 분해하지 않는다. 본 발명의 조성물에 관한 추가의 구체예에서, 효소성 분해 물질은 약 275°F 이하의 온도에서 구아 겔 내 β -1,4 글리코시드 결합을 가수 분해하는 능력을 가진다. 본 발명의 조성물에 관한 추가의 구체예에서, 효소성 분해 물질은 pH 약 11 이하에서 구아 겔 내 β -1,4 글리코시드 결합을 가수 분해하는 능력을 가진다. 본 발명의 조성물에 관한 추가의 구체예에서, 효소성 분해 물질은 서열 번호 2의 폴리펩티드를 가진다. 본 발명의 조성물에 관한 추가의 구체예에서, 효소성 분해 물질은 서열 번호 1의 서열을 가지는 폴리뉴클레오티드에 의해 코딩된다. 상기 조성물들 중 임의의 것에 관한 추가의 구체예에서, 효소성 분해 물질은 야생형 셀룰라아제의 돌연 변이된 변이체로서, 융점이, pH 약 6.5에서 야생형 셀룰라아제의 융점보다 20°F 이상 높으며, pH 약 10.5에서 야생형 셀룰라아제의 융점보다 10°F 이상 높다. 추가의 구체예에서, 효소성 분해 물질은 야생형 셀룰라아제에 대하여 12개의 돌연 변이를 포함하는 돌연 변이된 변이체이다. 추가의 구체예에서, 변이체 셀룰라아제는 다음과 같은 돌연 변이들, 즉 F38Y, Y61Q, M69E, D70P, R71S, I94Q, I166V, S183R, S191A, E212P, L231V, M276A, E277S, R280G, T297P 및 T301Q 중 하나 이상을 포함한다.
- [0042] 추가의 구체예에서, 상기 효소성 분해 물질은 서열 번호 3의 서열에 의해 코딩되는 셀룰라아제의 돌연 변이체이다.
- [0043] 추가의 구체예에서, 본 발명은 β -1,4 글리코시드 결합들을 포함하는 다당체 겔의 점도를 감소시키는 방법을 포함하는데, 이 방법은, 셀룰라아제가 다당체 겔 내 β -1,4 글리코시드 결합들을 가수 분해하기에 충분한 시간 동안, 허용된 조건 하에서 다당체 겔과 셀룰라아제 변이체를 접촉시킴으로써 구아 겔의 점도를 감소시키는 단계를 포함하며, 이때 셀룰라아제 변이체는 초고온성 세균으로부터 유래하는 야생형 셀룰라아제에 비해 12개 이상의 돌연 변이를 포함하고, 셀룰라아제 변이체는 야생형 셀룰라아제에 비해 증가된 온도 관용성 및 pH 관용성을 나타낸다. 추가의 구체예에서, 상기 돌연 변이 12개는 환원성 재편성, 재조합 및 선별로 이루어진 주기가 반복 실

시되는 유전자 위치 포화 돌연 변이 유발법에 의해 발생되었다. 대안적 구체예에서, 상기 다당체 겔은 선형 구아 겔, 가교형 구아 겔 또는 이의 조합을 포함한다. 대안적 구체예에서, 상기 셀룰라아제 변이체는 서열 번호 1의 서열을 가지는 폴리뉴클레오티드에 의해 코딩된다. 대안적 구체예에서, 상기 허용된 조건은 약 180~약 275 °F의 온도 범위를 포함한다. 대안적 구체예에서, 상기 허용된 조건은 약 9~약 11의 pH 범위를 포함한다.

[0044] 본 발명의 하나의 구체예에서, 본 발명은 다당체 겔 및 셀룰라아제 변이체를 포함하는 과쇄액을 지하 형성물에 도입하는 단계; 셀룰라아제 변이체가 다당체 겔 내 α-1,6 글리코시드 결합들을 가수 분해하지 않고 β-1,4 글리코시드 결합들 중 적어도 일부를 가수 분해하기에 충분한 시간 동안, 허용된 조건 하에서 상기 겔과 셀룰라아제 변이체를 반응시킴으로써, 점도 감소 과쇄액을 생산하는 단계를 포함하는 지하 형성물 처리 방법을 포함하는데; 여기서 상기 셀룰라아제 변이체는 초고온성 세균으로부터 유래하는 모 야생형 셀룰라아제에 비해 12개 이상의 돌연 변이를 포함하고, 이 셀룰라아제 변이체는 야생형 셀룰라아제에 비해 증가된 온도 관용성 및 pH 관용성을 나타낸다.

[0045] 전술된 방법들에 관한 추가의 구체예에서, 상기 셀룰라아제 변이체를 이용한 가수 분해는, 화학적 분해 물질에 의해, 상기 셀룰라아제 변이체에 의해 점도가 감소한 과쇄액과 필적할만하게 점도가 감소한 과쇄액[상기 셀룰라아제 변이체와 화학적 분해 물질이 투여되면 점도가 실질적으로 동일한 정도로 감소함]을 이용한 가수 분해의 경우보다 잔류물이 소량으로 생산된다.

[0046] 전술된 방법들에 관한 추가의 구체예에서, 셀룰라아제 변이체에 의해 점도가 감소한 과쇄액은, 화학적 분해 물질에 의해, 상기 셀룰라아제 변이체에 의해 점도가 감소한 과쇄액과 필적할만하게 점도가 감소한 과쇄액[상기 셀룰라아제 변이체와 화학적 분해 물질이 투여되면 점도가 실질적으로 동일한 정도로 감소함]보다 큰 전도도를 나타낸다.

[0047] **본 발명의 상세한 설명**

[0048] 효소는 촉매로서 작용을 하는 단백질이다. 단백질은 탈수 반응을 통해 펩티드 결합들에 의해 결합한 아미노산들의 고분자이다. 아미노산들의 동일성과 이 아미노산들이 단백질을 이루는 아미노산 결합 순서는 소정의 단백질 활성을 결정한다. 아미노산들이 단백질로 조립되는 순서(단백질 "서열")는 궁극적으로 해당 단백질을 "코딩하는" DNA 사슬의 서열에 의해 결정된다.

[0049] 단백질로 조립될 소정의 아미노산을 특징하는 3 뉴클레오티드 서열은 "코돈"이라고 칭하여진다. 단백질을 구성하는 아미노산 20개는 총체적으로 64개의 3 뉴클레오티드 코돈 서열에 의해 코딩된다. 단백질을 특징하는 일련의 코돈들은 "개방 해독 틀(Open Reading Frame)"이라 칭하여진다. 아미노산은 하나 또는 6개의 변별적 코돈들에 의해 특정될 수 있다. 특정된 아미노산에 영향을 미치지 않는, 코돈의 3 뉴클레오티드 서열상 변화(또는 돌연 변이)는 "침묵" 돌연 변이라 칭하여진다.

[0050] DNA 서열들이 오로지 "침묵" 돌연 변이를 통해서만 서로 상이해질 경우, 결과적으로는 동일한 단백질을 코딩할 수 있는 DNA 서열들은 다수 존재하게 된다. 소정의 단백질을 코딩하는 코돈들 중 하나 이상에 변화가 가하여짐으로 인해서 유전자가 생산하는 단백질의 양은 상당히 증가하게 될 수 있으며, 이때 코딩된 단백질의 서열은 영향을 받지 않는다.

[0051] 몇몇 구체예에서, 본 발명은 서열 번호 1의 서열을 포함한다. 몇몇 구체예에서, 본 발명은 서열 번호 1의 폴리뉴클레오티드 서열을 포함한다. 몇몇 구체예에서, 이 서열은 단백질을 코딩한다. 몇몇 구체예에서, 이 단백질은 셀룰라아제 활성을 가지는 효소이다.

[0052] 본원에 개시된 개선 뉴클레오티드 서열은 서열 번호 1로서 제공되며, 이 뉴클레오티드 서열은 이전에 개시되었던 것으로서, 서모토가 마리티마 균주 MSB8로부터 기원하는 DNA 라이브러리로부터 단리된 모 셀룰라아제 효소로부터 유래한 셀룰라아제 효소(서열 번호 2)를 코딩한다. 본원에 개시된 셀룰라아제는 PCT 공개 공보 WO 제 2009/020459호에 서열 번호 9[상기 PCT 공개 공보 WO 제 2009/020459호에 개시된 서열 번호 8의 폴리뉴클레오티드에 의해 코딩되는 것으로서; 본원에는 서열 번호 3으로 개시됨]로서 개시되어 있다. 몇몇 구체예에서, 본 발명은 서열 번호 1의 폴리뉴클레오티드 서열 또는 이의 단편을 포함한다. 몇몇 구체예에서, 이와 같은 서열은 단백질을 코딩한다. 몇몇 구체예에서, 이 단백질은 셀룰라아제 활성을 가지는 효소이다.

[0053] 본 발명은 서열 번호 3의 서열에 대하여 다수 개의 뉴클레오티드 염기 변화를 포함한다. 이와 같은 변화들은 코딩된 단백질에 대해서 침묵한다. 14개의 염기 변화는 이하에 제시되어 있다. "위치"란, 개방 해독 틀 내 뉴클레오티드의 번호를 나타내는 것으로서, 제1 코돈의 첫 번째 뉴클레오티드는 1번으로 매겨진다. 서열 번호 1의 개방 해독 틀이 자체의 5' 말단에서 다른 핵산 서열과 결합함으로써 인하여, 이 개방 해독 틀이 서열 번호 1의 서열

의 5' 말단보다 훨씬 더 길게 연장되는 경우, 상기 "위치"는 서열 번호 1의 개방 해독 틀의 5' 말단으로부터 시작하여 계속해서 번호가 매겨지면서 염기들을 나타낼 것이다. 이와 유사하게, 만일 서열 번호 1의 개방 해독 틀이 절단되어 이 개방 해독 틀이 서열 번호 1과 관련된 서열의 5' 말단에서 시작되지 않는다면, 번호 매김 체계는 서열 번호 1의 5' 말단에 상응하는, 상기 서열의 5' 말단으로부터 시작되어 계속 진행될 것이다.

- [0054] 뉴클레오티드 염기 변화 또는 돌연 변이는 표기법 "(구 뉴클레오티드)(위치)(신규 뉴클레오티드)"을 사용하여 지정된다. 돌연 변이들로서는 다음과 같은 것들이 있다: T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C 또는 이의 임의의 조합.
- [0055] 본원에 개시된 셀룰라아제를 코딩하는 앞서 보고된 서열들과 서열 번호 1의 서열이 총체적이면서 개별적으로 구별되도록 만들어주는 염기 변화들을 통하여, 이전에 사용되었던 뉴클레오티드 서열로서 상기 단백질과 동일한 단백질을 코딩하는 뉴클레오티드 서열보다 단백질을 더욱 높은 수준으로 발현하는 개방 해독 틀이 만들어진다.
- [0056] 몇몇 구체예에는, 서모토가 마리티마로부터 유래하는 셀룰라아제를 코딩하는 뉴클레오티드 서열은 T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C, A6C, G6C, A9C, G9C, A15G, C15G, T22C, G22C, A24T, C24T, T33C, G33C, T39C, G39C, T40C, G40C, T42C, G42C, T54C, G54C, T57C, G57C, A66C, G66C, C81A, T81A, T84C, G84C 또는 이의 임의의 조합으로부터 선택되는 하나 이상의 돌연 변이를 포함하는 것으로 개시되어 있다. 이와 같은 구체예들에 관한 몇몇 양태에 있어서, 하나 이상의 돌연 변이는 코딩된 단백질의 서열에 대해서 침묵한다. 다른 양태에서, 하나 이상의 돌연 변이를 통해서는, 하나 이상의 돌연 변이가 결실되고 그 외에는 상기 뉴클레오티드 서열과 상이하지 않은 뉴클레오티드 서열보다 더욱 높은 수준으로 셀룰라아제의 발현을 유도하는 하나 이상의 돌연 변이를 보유하는 뉴클레오티드 서열이 생성된다.
- [0057] 몇몇 구체예에는, 셀룰라아제를 코딩하는 뉴클레오티드 서열로서, 서열 번호 3의 서열을 포함하고, T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C, A6C, G6C, A9C, G9C, A15G, C15G, T22C, G22C, A24T, C24T, T33C, G33C, T39C, G39C, T40C, G40C, T42C, G42C, T54C, G54C, T57C, G57C, A66C, G66C, C81A, T81A, T84C, G84C 또는 이의 임의의 조합으로부터 선택되는 하나 이상의 돌연 변이를 가지는 뉴클레오티드 서열이 개시되어 있다.
- [0058] 몇몇 구체예에는, 하나 이상의 돌연 변이를 가지는 서모토가 마리티마 유래 뉴클레오티드 서열로서, 서모토가 마리티마 게놈 서열에 비해 상기 뉴클레오티드 서열에 의해 코딩된 단백질 발현 수준을 증가시키는 뉴클레오티드 서열이 개시되어 있다. 몇몇 양태에서, 하나 이상의 돌연 변이는 침묵한다.
- [0059] 몇몇 구체예에서, 서열 번호 2의 폴리펩티드를 코딩하는 제1 뉴클레오티드 서열로서, 상기 뉴클레오티드 서열은, 상기 단백질의 발현 수준이 제2 뉴클레오티드 서열에 의해 코딩되는 상기 단백질의 발현 수준에 비해서 증가하도록 서열 번호 2의 서열을 코딩하는 상기 제2 서열에 대해서 돌연 변이된 것인 제1 뉴클레오티드 서열이 개시되어 있다.
- [0060] 서모토가 마리티마는 극한의 염 농도(즉 NaCl 0.25~6.00%)에서 생육하는 능력을 특징으로 하는 호열성 진정 세균이다. 서모토가 마리티마는 호열성이고, 막대 모양을 띠고 있으며, 혐기성이면서 그람-음성인 세균들로 이루어진 서모토갈레스(*Thermotogales*) 목에 속한다. 생육에 필요한 최저 온도는 약 55℃이고, 최적 온도는 80~85℃이며, 최고 온도는 약 90℃이다. 몇몇 구체예에서, 최저 온도는 55℃ 미만이고, 최고 온도는 90℃ 이상이다. 이 세균은 진정 세균 계에서 가장 아래에 있는 분과들 중 하나로부터 천천히 진화된 것이다. 서모토갈레스 목을 구성하는 세균들은 지열이 높은 지역에서 번성하는 것으로서, "널리 분포되어 있으며 보편적으로 존재하는 것"이라 기술된 바 있다(Huber, R et al., 2006). 서모토가 마리티마는 서모토가 네아폴리타나(*Thermotoga neapolitana*), 서모토가 페트로필라(*Thermotoga petrophila*) 및 서모토가 나프토포필라(*Thermotoga naphthophila*)와 밀접하게 관련되어 있다. 서모토가 마리티마 표본은 이탈리아 불카노섬; 아조레스 군도의 리베리아 퀸테섬 및 상 미구엘섬; 인도네시아 상제양섬; 그리고 피지섬의 해저로부터 얻어졌다(Huber, R et al., 2006).
- [0061] 균주 MSB8은 이탈리아 불카노섬의 지열로 가열된 해양 퇴적토로부터 단리되었다(Huber, 1986). 상기 균주 수집 위치의 온도는 70~100℃이었으며, pH는 6.5~7.0이었다. 상기 균주는 독일 생물 자원 센터(Deutsche Sammlung von Mikroorganismen)에 DSM 3109로서 기탁되었으며, 또한 ATCC에 ATCC 43589로서 기탁되었다(Huber, R et al., 2006).
- [0062] 서모토가 마리티마 균주 MSB8은, 이 균주가 생산하는 효소의 예상 밖 열 안정성으로 인해 이 균주 자체의 효소 코딩 유전자에 대해 연구되어 왔다. 리에블(Liebl)은 논문["Analysis of a *Thermotoga maritima* DNA fragment

encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes."]을 발표한 바 있다. 뿐만 아니라, 아밀로스 분해 효소 유전자(Bibel, M. *et al.*, 1998), 역 자이라아제 유전자(Bouthier de la Tour, C. *et al.*, 1998), 알파-아밀라아제 유전자(Liebl, W. *et al.*, 1997), 알파-글루쿠로니다아제 유전자(Ruile, P. *et al.*, 1997), 크실라나아제 유전자(Winterhalter, C. *et al.*, 1995), 베타-글루코시다아제 유전자(Liebl, W. *et al.*, 1994) 및 글루칸 전이 효소 유전자(Liebl, W. *et al.*, 1992)가 단리되어 분석된 바 있다. 브로넨마이어(Bronnenmeier)에 의한 연구(Liebl, W. *et al.*, 1992; "Purification of *Thermotoga maritima* enzymes for the degradation of cellulosic materials")는, 이와 같은 효소들이 셀룰로스 및 크실란을 분해하는데 유리하다는 것을 확인시켜 주었다.

[0063] 발현계

[0064] 몇몇 구체예에서, 본 발명의 셀룰라아제를 코딩하는 DNA는 플라스미드에 도입될 수 있거나, 예를 들어 임의의 수의 그람-음성 세균 계, 예를 들어 이.콜라이, 슈도모나스 종, 예를 들어 플루오레센스, 슈도모나스 푸티다(*Pseudomonas putida*), 슈도모나스 아에루기노사(*Pseudomonas aeruginosa*), 랄스토니아 종 또는 카울로박터 종의 계능에 안정적으로 도입되어 균주를 형질 전환할 수 있다. 이와 유사하게, 셀룰라아제는 임의의 수의 그람-양성 세균 발현계, 예를 들어 바실러스(*Bacillus*) 종, 예를 들어 바실러스 서브틸리스(*Bacillus subtilis*), 바실러스 메가테리움(*Bacillus megaterium*), 바실러스 브레비스(*Bacillus brevis*), 락토코커스(*Lactococcus*) 종, 예를 들어 락토코커스 락티스(*Lactococcus lactis*), 락토바실러스(*Lactobacillus*) 종, 스트렙토마이세스(*Streptomyces*) 종, 예를 들어 스트렙토마이세스 리비단스(*Streptomyces lividans*)에 도입될 수 있다. 기타 그람-음성, 그람-양성 또는 관련 없는 진정 세균 또는 고 세균 발현계도 셀룰라아제를 발현하는데 사용될 수 있다.

[0065] 몇몇 구체예에서, 서열 번호 1의 서열은 본원에 개시된 셀룰라아제 단백질이 발현될 수 있는 다수의 계들 내에서 발현 수준을 증가시키는데 사용된다. 서열 번호 1의 서열은 본원에 개시된 셀룰라아제를 개선된 축적 수준(accumulation level)으로 발현하기 위해 임의의 수의 발현계에 도입될 수 있다. 예를 들어 서열 번호 1의 서열은 플라스미드에 도입될 수 있거나, 예를 들어 임의의 수의 그람-음성 세균 계, 예를 들어 이.콜라이, 슈도모나스 종, 예를 들어 플루오레센스, 슈도모나스 푸티다, 슈도모나스 아에루기노사, 랄스토니아 종 또는 카울로박터 종의 계능에 안정적으로 도입되어 이 균주를 형질 전환할 수 있다. 이와 유사하게, 서열 번호 1의 서열은 임의의 수의 그람-양성 세균 발현계, 예를 들어 바실러스 종, 예를 들어 바실러스 서브틸리스, 바실러스 메가테리움, 바실러스 브레비스, 락토코커스 종, 예를 들어 락토코커스 락티스, 락토바실러스 종, 스트렙토마이세스 종, 예를 들어 스트렙토마이세스 리비단스에 도입될 수 있다. 기타 그람-음성, 그람-양성 또는 관련 없는 진정 세균 또는 고 세균 발현계도 서열 번호 1의 서열을 발현하는데 사용될 수 있다. 추가의 구체예에서, 서열 번호 1의 서열은 임의의 수의 진핵 생물 발현계, 예를 들어 사카로마이세스(*Saccharomyces*), 쉬조사카로마이세스 폼베(*Schizosaccharomyces pombe*), 피치아 파스토리스(*Pichia pastoris*) 및 한사누엘라 폴리몰파(*Hansanuela polymorpha*)에 도입될 수 있다.

[0066] 더욱 구체적으로, 서열 번호 1의 서열은 자체의 발현을 유도하도록 플라스미드에 도입될 수 있다. 서열 번호 1의 서열이 도입될 수 있는 플라스미드로서는, 예를 들어 pQE, pET 및 pASK 균의 이.콜라이 발현 벡터; pCN51 LT8, RSF1010, pWZ112T 및 pMYC 균의 슈도모나스 발현 벡터; pBAX, pHT01 및 pHIS1525의 바실러스 발현 벡터; pIJ6021 및 pIJ2460 균의 스트렙토마이세스 발현 벡터; 그리고 pNZ9530 및 pNZ8148 균의 락토코커스(*Lactococcus*) 발현 벡터를 포함한다. 이와 같은 예들은 예시적 목적으로 제시된 것이지만, 서열 번호 1의 폴리뉴클레오티드 서열이 발현될 수 있는 벡터의 세트 전체를 나타내는 것은 아니다.

[0067] 몇몇 구체예에서, 발현계는 당 업계에 알려진 임의의 슈도모나스 플루오레센스 발현계, 예를 들어 슈도모나스 플루오레센스 발현계[다우 글로벌 테크놀로지스 인코포레이티드(Dow Global Technologies Inc.)로부터 시판되는 것으로서, 균주 DC454(미국 특허 출원 공보 제20050130160호 및 미국 특허 출원 공보 제20050186666호)]일 수 있었다. 셀룰라아제 효소 또는 폴리펩티드를 코딩하는 핵산 서열은 pMYC 벡터(다우 글로벌 테크놀로지스 인코포레이티드, 미국 특허 출원 공보 제20050130160호) 또는 pDOW1169 벡터(다우 글로벌 테크놀로지스 인코포레이티드, 미국 특허 출원 공보 제20080058262호)에 삽입된 후, 전기 친공에 의해 슈도모나스 플루오레센스 숙주로 도입된다. 당 업자는 본 발명의 구체예로서 사용될 수 있는 대안적 벡터들을 알 것이다.

[0068] 몇몇 구체예에서, 셀룰라아제는 다음과 같은 발현 수준들 이상으로 발현될 것이다: 1.0g/L, 2.0g/L, 3.0g/L, 4.0g/L, 5.0g/L, 6.0g/L, 7.0g/L, 8.0g/L, 9.0g/L, 10.0g/L, 11.0g/L, 12.0g/L, 13.0g/L, 14.0g/L, 15.0g/L, 16.0g/L, 17.0g/L, 18.0g/L, 19.0g/L, 20.0g/L, 21.0g/L, 22.0g/L, 23.0g/L, 24.0g/L, 25.0g/L, 26.0g/L,

27.0g/L, 28.0g/L, 29.0g/L, 30.0g/L, 31.0g/L, 32.0g/L, 33.0g/L, 34.0g/L, 35.0g/L 또는 이 이상.

[0069] 핵산

[0070] 본 발명은, 본 발명의 핵산 서열에 완전히 상보성인 서열을 포함하는 합성 또는 재조합의 단리된 핵산을 제공한다(이하, 상보성 (비 코딩) 서열들 및 코딩 서열들도 총괄하여 본 발명의 핵산 서열이라 칭하여짐).

[0071] 본 발명은, 셀룰로스 분해 활성을 가지는 폴리펩티드 하나 이상을 코딩하는 핵산을 포함하는 합성 또는 재조합의 단리된 핵산을 제공하는데, 이 경우 상기 핵산은 본 발명의 예시적 핵산, 예를 들어 서열 번호 1의 서열에 대한 서열 동일성(상동성)이 약 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 이 이상이거나, 완전히 동일한(서열 동일성(상동성)이 100%인) 서열을 포함한다. 예를 들어 본 발명은 서열 번호 1의 핵산 서열 (본 발명의 예시적 폴리뉴클레오티드 서열)을 포함하는 합성 또는 재조합의 단리된 핵산을 제공한다. 본 발명은 서열 번호 2에 제시된 바와 같은 서열(본 발명의 예시적 폴리펩티드 서열)을 포함하는 폴리펩티드 및 이의 효소 활성 단편을 코딩하는 합성 또는 재조합의 단리된 핵산을 제공한다.

[0072] 폴리펩티드

[0073] 본 발명의 폴리펩티드 및 펩티드는 합성 또는 재조합의 단리된 폴리펩티드이다. 펩티드 및 단백질은 시험관 내 또는 생체 내에서 재조합을 통해 발현될 수 있다. 본 발명의 펩티드 및 폴리펩티드는 당 업계에 공지된 임의의 방법을 이용하여 제조 및 단리될 수 있다. 본 발명의 폴리펩티드 및 펩티드는 또한 당 업계에 널리 공지된 방법을 이용하여 전체적으로나 부분적으로 합성될 수도 있다. 예를 들어 셀룰라아제 폴리펩티드는 (본원에 기술된 바와 같은) 표준적 재조합 발현계에서 생산될 수 있거나 화학적으로 합성될 수 있거나, 아니면 이 폴리펩티드가 천연 발현되는 유기체로부터 정제될 수 있다.

[0074] 본 발명은, 본 발명의 예시적 아미노산 서열(예를 들어 서열 번호 2의 서열)에 대한 서열 동일성이 약 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 이 이상이거나 100%인(완전히 동일한) 아미노산 서열, 또는 이의 효소 활성 단편을 포함하고, 셀룰로스 분해 활성을 가지는 합성 또는 재조합의 단리된 폴리펩티드를 제공한다.

[0075] 본 발명은 서열 번호 2에 제시된 서열, 이의 효소 활성 단편 및 변이체를 포함하는 합성 또는 재조합의 단리된 폴리펩티드를 제공한다.

[0076] 대안적 구체예에서, 본 발명은 셀룰로스 분해 활성을 가지되 단일 서열, 프리프로 도메인(prepro domain), 도커린 도메인(dockerin domain) 및/또는 탄수화물 결합 모듈(CBM)은 가지지 않는 폴리펩티드(및 이를 코딩하는 핵산)를 제공하며; 하나의 양태에서, 탄수화물 결합 모듈(CBM)은 셀룰로스 결합 모듈, 리그닌 결합 모듈, 크실란 결합 모듈, 크실로스 결합 모듈, 만노스 결합 모듈, 크실로글루칸 특이적 모듈 및/또는 아라비노푸라노스 결합 모듈을 포함하거나 이것들로 이루어져 있다.

[0077] 대안적 구체예에서, 본 발명은 셀룰로스 분해 활성을 가지고, 이종 서열들을 추가로 포함하는 폴리펩티드(및 이를 코딩하는 핵산)를 제공하는데; 하나의 양태에서, 상기 이종 서열은 (i) 이종 신호 서열, 이종 탄수화물 결합 모듈, 이종 도커린 도메인, 이종 촉매 도메인(CD) 또는 이의 조합; (ii) 이종 신호 서열, 탄수화물 결합 모듈 또는 촉매 도메인(CD)이 이종 효소로부터 유래하는 것인 상기 (i)의 서열; 또는 (iii) 태그, 에피토프, 표적화 펩티드, 절단 가능 서열, 검출 가능 부 또는 효소를 코딩하는 서열을 포함하거나 이와 같은 서열로 이루어져 있고; 하나의 양태에서, 상기 이종 탄수화물 결합 모듈(CBM)은 셀룰로스 결합 모듈, 리그닌 결합 모듈, 크실란 결합 모듈, 크실로스 결합 모듈, 만노스 결합 모듈, 크실로글루칸 특이적 모듈 및/또는 아라비노푸라노스 결합 모듈을 포함하거나 이것들로 이루어져 있으며; 하나의 양태에서, 상기 이종 신호 서열은 코딩된 단백질을 액포, 소포체, 염록체 또는 전분 과립에 표적화한다.

[0078] 효소 활성

[0079] pNP-β-D-락토피라노시드의 효소에 의한 가수 분해(본원에 개시된 셀룰라아제에 의함)는 효소 활성의 척도로서 사용될 수 있다. p-니트로페놀의 유리는 분광 분석학적으로 405nm에서 추적될 수 있다. 405nm에서의 흡광도 증가량은 한정된 조건에서 표준적 흡광도를 이용함으로써 p-니트로페놀의 몰수(μmole)로 전환될 수 있다. 활성 1

유닛은, pH 7.00 및 80°C에서 1분 동안 pNP-β-D-락토피라노시드 2mM로부터 p-니트로페놀 0.42 μmole이 유리되는데 필요한 효소의 양으로서 정의된다[Advances in Carbohydrate Chemistry and Biochemistry, Academic Press, 1999].

[0080] 열 안정성

[0081] 몇몇 양태에서, 본 발명의 제조합 핵산은 열 안정성이면서 셀룰로스 분해 활성을 가지는 폴리펩티드를 코딩한다. 예를 들어 본 발명의 폴리펩티드, 즉 서열 번호 2의 서열 또는 본 발명의 변이체로부터 생산된 효소는 열 안정성일 수 있다. 본 발명에 의한 열 안정성 폴리펩티드는 37°C 이상~약 95°C, 또는 약 55~약 85°C, 또는 약 70~약 75°C, 또는 약 70~약 95°C, 약 90~약 95°C, 약 95~약 105°C, 또는 약 95~약 110°C의 범위에 속하는 온도를 포함하는 조건 하에서, 결합 활성 및/또는 효소 활성, 예를 들어 셀룰로스 분해 활성을 보유할 수 있다. 몇몇 양태에서, 본 발명의 폴리펩티드는 1~약 5°C, 약 5~약 15°C, 약 15~약 25°C, 약 25~약 37°C를 포함하는 조건 하에서, 결합 및/또는 효소 활성, 예를 들어 셀룰로스 분해 활성을 보유할 수 있다. 몇몇 양태에서, 본 발명의 폴리펩티드는 90°C, 91°C, 92°C, 93°C, 94°C, 95°C, 96°C, 97°C, 98°C, 99°C, 100°C, 101°C, 102°C, 103°C, 103.5°C, 104°C, 105°C, 107°C, 108°C, 109°C 또는 110°C, 또는 이 이상의 온도를 포함하는 조건 하에서, 결합 및/또는 효소 활성, 예를 들어 셀룰로스 분해 활성을 보유할 수 있다. 몇몇 양태에서, 본 발명에 의한 열 안정성 폴리펩티드는 pH 약 6.5, pH 약 6, pH 약 5.5, pH 약 5, pH 약 4.5 또는 pH 약 4 또는 이 pH 값들 이하인(즉, 산성이 더욱 강한) pH 값들을 포함하는 산성 조건 하에 전술된 범위의 온도에서 활성, 예를 들어 셀룰로스 분해 활성을 보유하거나; 또는 pH 약 6.5, pH 약 6, pH 약 5.5, pH 약 5, pH 약 4.5 또는 pH 약 4 또는 이 pH 값들 이하인(즉, 산성이 더욱 강한) pH 값들을 포함하는 산성 조건에 노출된 후에도 셀룰로스 분해 활성을 보유하거나; 또는 pH 약 7, pH 약 7.5, pH 약 8.0, pH 약 8.5, pH 약 9, pH 약 9.5, pH 약 10, pH 약 10.5, pH 약 11, pH 약 11.5, pH 약 12, pH 약 12.5 또는 이 pH 값들 이상인(즉 염기성이 더욱 강한) pH 값들을 포함하는 염기성 조건 하에 활성을 보유하거나; 또는 pH 약 7, pH 약 7.5, pH 약 8.0, pH 약 8.5, pH 약 9, pH 약 9.5, pH 약 10, pH 약 10.5, pH 약 11, pH 약 11.5, pH 약 12, pH 약 12.5 또는 이 pH 값들 이상인(즉 염기성이 더욱 강한) pH 값들을 포함하는 염기성 조건에 노출된 후에도 셀룰로스 분해 활성을 보유한다.

[0082] 열 내성

[0083] 몇몇 양태에서, 본 발명의 제조합 핵산은 열 내성이면서 셀룰로스 활성을 가지는 폴리펩티드를 코딩한다. 예를 들어 본 발명의 폴리펩티드, 즉 서열 번호 2의 서열 또는 본 발명의 변이체로부터 생산된 효소는 열 내성일 수 있다. 몇몇 양태에서, 셀룰로스 분해 활성은 열 내성인데, 예를 들어 본 발명의 폴리펩티드는 37°C 이상~약 95°C, 또는 약 55~약 85°C, 또는 약 70~약 75°C, 또는 약 70~약 95°C, 약 90~약 95°C, 약 95~약 105°C, 또는 약 95~약 110°C의 범위에 속하는 온도에 노출된 후에도 셀룰로스 분해 활성을 보유한다. 몇몇 양태에서, 본 발명의 폴리펩티드는 약 1~약 5°C, 약 5~약 15°C, 약 15~약 25°C, 약 25~약 37°C와 같은 온도 범위를 포함하는 조건에 노출된 후에도 셀룰로스 분해 활성을 보유한다. 몇몇 양태에서, 본 발명의 폴리펩티드는 90°C, 91°C, 92°C, 93°C, 94°C, 95°C, 96°C, 97°C, 98°C, 99°C, 100°C, 101°C, 102°C, 103°C, 103.5°C, 104°C, 105°C, 107°C, 108°C, 109°C 또는 110°C 이하, 또는 이 이상의 온도에 노출된 후에도 셀룰로스 분해 활성을 보유할 수 있다. 몇몇 양태에서, 본 발명의 핵산에 의해 코딩된 폴리펩티드는 pH 약 6.5, pH 약 6, pH 약 5.5, pH 약 5, pH 약 4.5 또는 pH 약 4 또는 이 pH 값들 이하인(즉, 산성이 더욱 강한) pH 값들을 포함하는 산성 조건 하에서도 셀룰로스 분해 활성을 보유하거나; 또는 pH 약 6.5, pH 약 6, pH 약 5.5, pH 약 5, pH 약 4.5 또는 pH 약 4 또는 이 pH 값들 이하인(즉, 산성이 더욱 강한) pH 값들을 포함하는 산성 조건에 노출된 후에도 셀룰로스 분해 활성을 보유하거나; 아니면 pH 약 7, pH 약 7.5, pH 약 8.0, pH 약 8.5, pH 약 9, pH 약 9.5, pH 약 10, pH 약 10.5, pH 약 11, pH 약 11.5, pH 약 12, pH 약 12.5 또는 이 pH 값들 이상인(즉 염기성이 더욱 강한) pH 값들을 포함하는 염기성 조건 하에서도 셀룰로스 분해 활성을 보유한다.

[0084] 셀룰로스 분해

[0085] 몇몇 양태에서, 본 발명의 조성물 및 방법은 바이오매스의 효소에 의한 분해에 사용되며, 다수의 상이한 효소들, 예를 들어 셀룰라아제 및 헤미셀룰라아제를 사용하는 것을 포함할 수 있다. 본 발명을 실행하는데 사용된 셀룰라아제는 셀룰로스를 글루코스로 분해할 수 있다. 몇몇 양태에서, 본 발명을 실행하는데 사용된 조성물은 효소들, 예를 들어 크실라나아제, 크실로시다아제(예를 들어 β-크실로시다아제), 셀로비오하이드롤라아제 및/또는 아라비노푸라노시다아제, 또는 헤미셀룰로스, 셀룰로스 및 리그노셀룰로스 물질을 발효 가능한 당 및/또는 단량체 당으로 분해할 수 있는 기타 효소들의 혼합물을 포함할 수 있다.

[0086] 본 발명의 효소, 예를 들어 엔도글루카나아제는 셀룰로스 또는 임의의 베타-1,4-결합 글루칸 포함 합성 또는 천연 물질, 예를 들어 임의의 식물성 물질에서 발견된 물질을 분해하는데 사용된다. 본 발명의 효소, 예를 들어 엔도글루카나아제는, 예를 들어 목재 가공, 펄프 및/또는 제지 산업, 식물 제조, 그리고 가정용 및 산업용 세정제, 및/또는 바이오매스 폐기물 처리에 있어서, 임의의 공급원, 예를 들어 모든 생물학적 공급원, 예를 들어 식물 바이오매스, 예를 들어 옥수수, 곡물, 목초(예를 들어 난지형 잔디류, 예를 들어 솔가스트럼 누탄스 (*Sorghastrum nutans*); 또는 지팡이풀(*switch grass*), 예를 들어 파니쿰(*Panicum*) 종, 예를 들어 파니쿰 비르가툼(*Panicum virgatum*)), 또는 목재나 목재 가공 부산물로부터 유래하는 셀룰로스를 분해하기 위한 상업용 효소로서 사용된다.

[0087] 음식물

[0088] 몇몇 구체예에서, 본 발명의 셀룰라아제는 동물 또는 사람이 먹는 식품, 식품 첨가제 또는 식이 보조제를 전처리, 개질 또는 분해를 촉진하거나 이의 소화를 촉진하는데 사용될 수 있다. 몇몇 구체예에서, 본 발명의 셀룰라아제는 동물 또는 사람이 먹는 식품, 식품 첨가제 또는 식이 보조제로서 사용될 수 있다. 몇몇 양태에서, 셀룰라아제는 소화기 질환을 치료하거나 예방하는 역할을 할 것이다. 본 발명의 다른 양태에서, 셀룰라아제는 분해능에 변화가 있거나 이 분해능이 강화될 것이다. 본 발명의 다른 양태에서, 셀룰라아제는 식료품의 분해를 촉진하거나 바꾸거나 도울 것이다. 본 발명의 추가의 양태에서, 셀룰라아제는 식료품의 영양 가치를 개선하거나, 보완하거나 바꿀 것이다. 추가의 양태에서, 셀룰라아제는 소화기관, 예를 들어 위 및/또는 장에서 활성을 갖는다.

[0089] 몇몇 구체예에서, 본 발명의 셀룰라아제는 동물 사료나 동물 사료의 첨가제로서 사용될 수 있다. 몇몇 구체예에서, 본 발명의 셀룰라아제의 열 안정성 및/또는 열 내성은 2차 제제, 예를 들어 염 또는 왁스가 사용되지 않더라도 펠릿을 생성시킬 수 있다. 셀룰라아제를 포함하는 동물 사료는 당 업자들에게 알려진 임의의 제형으로서 동물에 제공될 수 있다. 동물 사료 제형의 예로서는 다음과 같은 것들을 포함하나, 이에 한정되는 것은 아니다: 전달 매트릭스, 펠릿, 정제, 겔, 액체, 스프레이, 분쇄형 과립 또는 분말.

[0090] 본 발명은 열 안정성 재조합 셀룰라아제 효소, 예를 들어 본 발명의 폴리펩티드를 포함하는 식용 효소 전달 매트릭스를 제공한다. 본 발명은 셀룰라아제 보충제를 동물에 전달하기 위한 방법으로, 다음과 같은 단계들을 포함하는 방법을 제공한다: 과립형 식용 캐리어 및 열 안정성 재조합 셀룰라아제 효소를 포함하는 펠릿의 형태로 식용 효소 전달 매트릭스를 제조하는 단계로서, 상기 펠릿은 이 펠릿 자체 내에 포함된 셀룰라아제 효소를 수성 매질에 용이하게 분산하는 것인 단계, 및 식용 효소 전달 매트릭스를 동물에 투여하는 단계. 본 발명의 재조합 셀룰라아제 효소는 본 발명의 폴리펩티드를 포함할 수 있다. 과립형 식용 캐리어로서는 곡물의 씨눈, 지방 성분이 소모된 곡물의 씨눈, 건초, 자주개자리, 큰 조아재비, 대두 종피, 해바라기씨박 및 밀 미들링(*wheat middling*)으로 이루어진 군으로부터 선택되는 캐리어를 포함할 수 있다. 식용 캐리어는 지방 성분이 소모된 곡물의 씨눈을 포함할 수 있다. 셀룰라아제 효소는 글리코실화될 수 있는데, 이로써 펠릿 제조 공정 조건에서 열 안정성이 제공될 수 있다. 전달 매트릭스는 곡물의 씨눈과 셀룰라아제를 포함하는 혼합물을 펠릿으로 만듦으로써 제조될 수 있다. 펠릿 제조 공정 조건은 증기를 사용하는 과정을 포함할 수 있다. 펠릿 제조 공정 조건은 약 80°C가 넘는 온도를 약 5분 동안 적용하는 것을 포함할 수 있으며, 효소의 비활성은 효소 1 mg당 350 유닛 이상 ~ 약 900 유닛이다.

[0091] 에탄올 제조 방법

[0092] 본 발명은 전분 포함 조성물과 셀룰로스 분해 활성을 가지는 폴리펩티드를 접촉시키는 단계로서, 상기 폴리펩티드는 본 발명의 서열을 가지거나, 이 폴리펩티드 또는 이의 효소 활성 단편은 본 발명의 서열을 포함하는 핵산에 의해 코딩되는 것인 단계를 포함하는 에탄올 제조 방법을 제공한다. 본 발명은 전분 및 셀룰로스 분해 활성을 가지는 폴리펩티드를 포함하는 조성물을 제공하는데, 여기서 상기 폴리펩티드는 본 발명의 서열을 가지거나, 이 폴리펩티드 또는 이의 효소 활성 단편은 본 발명의 서열을 포함하는 핵산에 의해 코딩된다.

[0093] 양조 및 발효

[0094] 본 발명은, 본 발명의 셀룰라아제를 포함하는 맥주를 양조(예를 들어 발효)하는 방법을 제공한다. 하나의 예시적인 방법에 있어서, 전분 함유 원재료는 분해 및 가공되어 맥아가 된다. 본 발명의 효소는 발효 공정 중 임의의 시점에서 사용된다. 본 발명의 셀룰라아제는 양조 산업에서 베타-글루칸을 분해하기 위해 사용될 수 있다. 몇몇 양태에서, 본 발명의 셀룰라아제는 양조 산업에 있어서 음료를 투명하게 만들기 위해 사용된다. 본 발명의 효소는, 예를 들어 미국 특허 제4,746,517호에 개시된 바와 같이 음료 산업에서 맥아 즙 또는 맥주의 여과성을 개선하는데 사용될 수 있다.

- [0095] 몇몇 양태에서, 본 발명의 셀룰라아제는 담금 공정 및 전환 공정에 사용된다. 양조 및 발효 산업에 있어서, 담금 및 전환 공정은 수용성 글루칸, 만난, 아라비녹실란, 크실란 또는 기타 다당체의 적당한 분해가 촉진될 수 없을 정도로 낮은 온도에서 진행된다. 이와 같은 고분자들은 맥아즙 담금 공정에서 점도를 증가시킬 수 있는 점착성 기질을 생성하게 되고, 이로 인하여 담금 공정에서의 범람 시간이 길어지며, 비효율적인 여과와 낮은 추출 수율로 인하여 맥주 완제품에 침전물이 생기고, 잔류물로 인해 맥주가 탁해진다.
- [0096] 몇몇 양태에서, 본 발명의 셀룰라아제는 맥아 저장고 가동시 사용되는데, 예를 들어 글루카나아제가 공정 수에 첨가되면 발아 시간이 단축되고/단축되거나 품질이 떨어졌던 보리가 허용 가능한 수준의 품질을 가지는 맥아로 전환되는 것이 촉진된다. 몇몇 양태에서, 본 발명의 효소는 담금 공정에 사용되는데, 예를 들어 상기 효소는 맥아즙의 여과성을 증가시키고/증가시키거나 맥즙 여과(담금액으로부터 맥아즙을 분리하는 공정)를 촉진하기 위해 첨가된다. 몇몇 양태에서, 본 발명의 효소는, 예를 들어 부유물 제거를 돕고/돕거나 여과를 개선하기 위해 발효조 및/또는 침전조에서 사용된다. 몇몇 양태에서, 본 발명의 효소는 부가적 양조 공정에서 사용되는데, 예를 들어 본 발명의 글루카나아제는 보리, 밀 및/또는 기타 곡물로부터 유래하는 글루칸, 만난, 아라비녹실란, 크실란 또는 기타 다당체, 예를 들어 맥아 중 글리칸을 분해하기 위해 첨가된다. 몇몇 양태에서, 본 발명의 효소는 맥아가 사용되는 양조에 사용되는데, 예를 들어 본 발명의 글루카나아제는 글루칸 함량이 높은 저품질 맥아를 개선하기 위해 첨가된다.
- [0097] 본 발명의 셀룰라아제는, 예를 들어 미국 특허 제5,762,991호; 동 제5,536,650호; 동 제5,405,624호; 동 제5,021,246호; 및 동 제4,788,066호에 개시되어 있는 바와 같이, 임의의 맥주 또는 알코올 음료 생산 공정에 사용될 수 있다.
- [0098] 식품 처리 및 식품 가공
- [0099] 본 발명의 셀룰라아제는 식품 가공 산업에서 다양한 용도로 사용된다. 예를 들어 하나의 양태에서, 본 발명의 효소는 오일이 풍부한 식물성 물질, 예를 들어 오일이 풍부한 종자로부터의 오일 추출, 예를 들어 대두로부터의 오일 추출, 올리브로부터의 올리브 오일 추출, 평지 씨로부터의 평지 씨 오일 추출 및/또는 해바라기씨로부터의 해바라기 오일 추출을 개선하는데 사용된다.
- [0100] 본 발명의 셀룰라아제는 식물 세포 물질의 구성 성분들을 분리하는데 사용될 수 있다. 예를 들어 본 발명의 효소는 글루칸 풍부 물질(예를 들어 식물 세포)을 각각의 구성 성분별로 분단리하는데 사용될 수 있다. 몇몇 양태에서, 본 발명의 효소는 글루칸 풍부 농작물 또는 오일 풍부 농작물을, 유용한 단백질, 오일 및 외피 분획으로 분리하는데 사용될 수 있다. 분리 공정은 당 업계에 공지된 방법들을 사용함으로써 진행될 수 있다.
- [0101] 본 발명의 셀룰라아제는 생산량을 늘리기 위해 과일 또는 채소 주스, 시럽 및 추출물 등을 제조하는데 사용될 수 있다. 본 발명의 효소는, 예를 들어 곡물, 곡류, 와인 또는 주스 생산시 생성되는 다양한 식물 세포 벽 유래 물질 또는 폐기물, 또는 농업 부산물, 예를 들어 식물의 외피, 콩의 외피, 사탕무 펄프, 올리브 펄프 및 감자 펄프 등을 효소 처리(예를 들어 글루칸 포함 식물성 물질의 가수 분해)하는데 사용될 수 있다. 본 발명의 효소는 가공된 과일이나 채소의 외관 및 질기(consistency)를 개선하는데 사용될 수 있다. 본 발명의 효소는 식물성 물질, 예를 들어 식품의 가공과, 식물 구성 성분의 정제 또는 추출을 가속화하기 위해 식물성 물질을 처리하는데 사용될 수 있다. 본 발명의 셀룰라아제는 사료의 가치를 높이고, 수 결합 능(water binding capacity)을 감소시키며, 폐수 식물의 분해능을 개선하고/개선하거나 식물성 물질의 저장 목초로의 전환을 촉진하는 등의 목적으로 사용될 수 있다. 본 발명의 셀룰라아제는 또한 과일 가공 및 양조 산업에서 장비를 청결하게 유지하기 위해 사용될 수도 있다.
- [0102] 세제 조성물
- [0103] 본 발명은, 본 발명의 폴리펩티드 하나 이상을 포함하는 세제 조성물과, 이 조성물을 제조하는 방법 및 사용하는 방법을 제공한다. 본 발명은 세제 조성물을 제조 및 사용하는 모든 방법들을 채택하고 있는데, 예를 들어 미국 특허 제6,413,928호; 동 제6,399,561호; 동 제6,365,561호; 동 제6,380,147호를 참조. 세제 조성물은 1부 및 2부 수성 조성물, 비 수성 액체 조성물, 성형 고체형, 과립형, 입자형, 압착 정제형, 겔 형 및/또는 페이스트 형 및 슬러리 형일 수 있다. 본 발명은 또한 이와 같은 세제 조성물을 사용하여 다량의 음식물 얼룩, 음식물 찌꺼기 막 그리고 기타 소량의 식품 조성물을 신속하게 제거할 수 있는 방법도 제공한다. 본 발명의 효소는 전분 다당체의 촉매 가수 분해를 통해 점성이 있는 얼룩들이 제거되는 것을 촉진할 수 있다. 본 발명의 효소는 설거지용 세제와 직물 세탁용 세제에도 사용될 수 있다. 실제 활성 효소 함량은 세제 조성물의 제조 방법에 따라 달라지긴 하지만, 세제 용액이 원하는 효소 활성을 가지기만 한다면 그다지 중요한 문제는 아니다. 몇몇 양태에

서, 완제품 용액 중에 존재하는 글루코시다제의 양은, 세제 조성물 1g당 약 0.001~0.5mg이다. 본 발명의 방법과 제품에 사용되도록 선택된 구체적인 효소는 최종 효용의 조건, 예를 들어 제품의 물리적 형태, 적용 pH, 적용 온도, 그리고 분해 또는 변질될 오염물의 유형에 따라서 달라진다. 효소는 효용 조건에 관한 임의의 구체적 조합에 대하여 안정성과 최적 활성을 제공하는 것으로 선택될 수 있다. 본 발명의 세제는 양 이온성 계면활성제, 반 극성 비 이온성 계면활성제 또는 쌍 극성 계면활성제 또는 이의 혼합물을 포함할 수 있다.

[0104] 본 발명은 경질 표면 세정용 세제 조성물, 섬유 세정용 세제 조성물, 설거지용 조성물, 구강 청정 조성물, 틀니 세정 조성물 및 콘택트 렌즈 세정액을 포함하는 세정 조성물을 제공한다. 몇몇 양태에서, 본 발명은, 세척에 충분한 조건 하에서 물건과 본 발명의 폴리펩티드를 접촉시키는 단계를 포함하는 물건 세척 방법을 제공한다. 본 발명의 폴리펩티드는 세제 첨가제로서 포함될 수 있다. 본 발명의 세제 조성물은, 예를 들어 본 발명의 폴리펩티드를 포함하는 손세탁용 세제 조성물 또는 기계 세탁용 세제 조성물로서 제조될 수 있다. 오염된 섬유를 전처리하는데 적당한 세탁 첨가제는 본 발명의 폴리펩티드를 포함할 수 있다. 섬유 유연제 조성물은 본 발명의 폴리펩티드를 포함할 수 있다. 대안적으로, 본 발명의 폴리펩티드는 일반적으로 가정의 경질 표면 세정 작업에 사용하기 위한 세제 조성물로서 제조될 수 있다.

[0105] 오일 및 가스 탐사 및 정화

[0106] 오일 및 가스 갱정, 그리고 셰일 가스 저장소의 생산성을 높이기 위해서, 고도로 특화된 기술(소위 "수압 파쇄법(hydraulic fracturing)")이 점점 각광받고 있다. 통상의 수압 파쇄법 실시시에는 구아를 주성분으로 하는 다량의 유액(겔 형으로서, "파쇄액(fracturing fluid)"이라고도 칭하여짐)이 매우 높은 수압이 적용되는 가운데 갱정 구멍 내부로 펌핑되어 들어간다. 가압 유액은 갱정 주변 형성물에 열하들과 균열들을 새로이 발생시킨다. 파쇄액에 포함된 모래 입자들은, 새로이 발생한 균열 속으로 들어가서 이를 메우고, 채널이 뚫려있도록 지지해주는 역할을 함으로써, 오일 및 가스의 유량을 증가시킨다. 일단 모래가 균열 속에 퇴적되면, 겔의 형태를 띠었던 유액은 분해되어(즉 파괴되어) 균열 표면을 보강해주게 되고, 이로써 오일이나 가스의 흐름에 방해가 되는 어떠한 막힘 현상도 없어지게 된다. 산업 일각에서는 파쇄액을 분해하고, 열하들과 균열들로부터 임의의 고품질 형태의 잔류물을 제거하기 위해서 점도 감소 물질(viscosity breaker)(예를 들어 산화제, 산 또는 효소)을 사용하고 있다.

[0107] 몇몇 구체예에서, 본원에 개시된 셀룰라아제는 오일 및 가스의 가공을 촉진하기 위해 고온 점도 감소 물질로서 사용될 것이다. 더욱 구체적으로, 본원에 개시된 본 발명의 셀룰라아제는 오일 또는 가스 갱정 내에서 수압 파쇄가 행하여질 때 파쇄액에 첨가될 것이다.

[0108] 서열 번호 1의 서열에 의해 코딩된 효소, 그리고 기타 본원에 개시된 폴리뉴클레오티드에 의해 코딩되거나 본원에 개시된 방법에 의해 생산될 수 있는 셀룰라아제는 잠재적으로 광범위한 다당체를 가수 분해하는데 사용될 수 있다[상기 효소들 다수는 오일 및 가스 채취시 시추, 파쇄 및 갱정 정화 작업에 유용함]. 본원에 개시된 셀룰라아제는, 예를 들어 구아, 하이드록시프로필 구아, 카복시메틸 구아, 카복시메틸 하이드록시프로필 구아, 카복시메틸 셀룰로스, 보리의 β-글루칸 및 메뚜기콩 겉에 대해서 광범위한 β-글리코시다아제 활성을 나타낸다. 상기 효소의 작용 패턴은 바람직하게 엔도(endo) 작용 및 엑소(exo) 작용 둘 다이며, 이와 같은 패턴에 의해서(즉 길이가 긴 다당체 사슬 내부를 절단하기도 하고, 고분자 말단부로부터 이당체 단위씩 절단하기도 함으로써) 다당체(예를 들어 구아 및 유도체화된 구아 용액)의 점도가 효율적으로 감소할 수 있다. 전술된 다당체들 이외에, 본원에 개시된 효소의 기타 기질로서는 선형 또는 가교형 겔을 형성할 수 있는 것들을 포함한다. 적당한 다당체 기질의 예로서는 갈락토만난 겉, 구아, 유도체화된 구아, 셀룰로스 및 셀룰로스 유도체, 전분, 전분 유도체, 잔탄, 유도체화된 잔탄, 그리고 이의 조합을 포함한다. 구체예로서는 구아 겉, 구아 겉 유도체, 메뚜기콩 겉, 카라야 겉, 잔탄 겉, 셀룰로스와 셀룰로스 유도체 등도 포함하나, 이에 한정되는 것은 아니다. 본원에 개시된 효소들이 작용할 수 있는 통상의 고분자 또는 겔화제로서는 구아 겉, 하이드록시프로필 구아, 카복시메틸 하이드록시프로필 구아, 하이드록시에틸 셀룰로스, 카복시메틸 하이드록시에틸 셀룰로스, 카복시메틸 셀룰로스, 디알킬 카복시메틸 셀룰로스 등을 포함한다. 고분자의 기타 예로서는 포스포만난, 스크레로글루칸, 텍스트란 및 기타 고분자 류를 포함하나, 이에 한정되는 것은 아니다. 몇몇 구체예에서, 고분자 기질은 카복시메틸 하이드록시프로필 구아이다. 몇몇 구체예에서, 본원에 개시된 효소는 또한 바이오 겉(biogum)(예를 들어 데이트 시럽 또는 수크로스 제조된 숙시노글리칸 바이오 겉)을 가수 분해하는데에도 효과적일 수 있다. 몇몇 구체예에서, 본원에 개시된 효소는 셀룰로스 함유 고분자 또는 유도체화 셀룰로스 함유 고분자를 가수 분해하는데 사용될 수 있는데, 통상적으로 상기 효소는 셀룰로스 주쇄의 글루코시드 결합을 공격한다. 본원에 개시된 효소는, 몇몇 경우 임의의 셀로비오스 단편의 (1,4)-β-D-글루코시드 결합 내 셀룰로스 주쇄와 단당체 단위 사이의 엔도(1,4)-β-D-글루코시드 결합 및 엑소(1,4)-β-D-글루코시드 결합을 특이적으로 가수 분해함으로써, 고분자를 주로 단당체

단위로 분해하는데 적당할 수 있다.

- [0109] 본원에 개시된 셀룰라아제를 사용하는 각각의 파쇄 작업에 있어서, 현장 작업자들은 일반적으로 처음에 산업 현장 실험실에서 효소 투여량 최적화 연구를 진행할 것이다. 이와 같은 연구는, 셀룰라아제를 농도 10~400ppm으로 희석하고 나서, 선형 또는 가교형 구아 검(25~60lb/1,000gal)과 혼합하는 과정을 포함할 수 있다. 사용시 조건(특히 고온, 고압, 고 pH 조건이 조성되는 갱정의 조건)에 따라서, 구아 검은 아마도 가교제를 사용함으로써 가교될 것이다. 그 다음, 이와 같은 최적화 연구로부터 얻어진 효소 투여량 정보는 실제 파쇄 작업시 활용된다.
- [0110] 본원에 개시된 셀룰라아제 특유의 활성은 깊은 갱정, 즉 고온 및 고 pH 조건이 조성되는 곳에서 순조롭게 제어되는 방식으로, 구아를 주성분으로 하는 파쇄액을 가수 분해할 수 있다. 화학적 분해 물질과 비교되게, 본원에 개시된 본 발명의 셀룰라아제는 극단적이면서 비 선택적인 화학적 분해 물질의, 비부식성이면서 환경 친화적인 대안을 제공한다.
- [0111] 본 발명의 몇몇 구체예에서, 셀룰라아제는 오일 및 가스 탐사 활동에 사용되는 유액을 처리, 정화 또는 개질하는데 사용될 수 있다. 추가의 양태에서, 본 발명의 셀룰라아제는 유액을 부분적으로나 전체적으로 처리 또는 개질할 것인데, 그 결과 상기 유액은 추후의 오일 및 가스 탐사 활동에 재사용 또는 재활용될 수 있거나, 아니면 환경 친화적 방식으로 처분될 수 있게 될 것이다.
- [0112] 본원에는 오일 및 가스 파쇄 공정을 비롯한 여러 가지 용도로 사용될 조성물이 개시되어 있다. 몇몇 구체예에서, 본 발명의 조성물은 고분자 증점제와 함께, 초고온성 세균으로부터 유래하는 셀룰라아제 효소 및/또는 이의 비 자연 발생 변이체 하나 이상을 포함한다. 몇몇 구체예에서, 본 발명의 조성물은 임의로 부가 성분들, 예를 들어 유체 손실을 막고, 형성물의 손상을 줄이며, pH를 조정하고, 미생물 성장을 억제하며, 온도 안정성을 개선하는 제제들을 포함할 수도 있다. 몇몇 통상의 부가 성분들로서는 산, 향균제, 점토 안정화제, 부식 억제제, 가교제, 마찰 감소제, 철 생성 억제제, 관석 형성 억제제, 계면활성제 및 열 안정화제(예를 들어 티오황산나트륨)를 포함한다.
- [0113] 몇몇 구체예에서, 본원에 개시된 셀룰라아제 효소는 글루카나아제, 예를 들어 엔도글루카나아제, 만난아제, 크실라나아제 활성을 가지거나 또는 이 활성을 조합하여 가진다. 몇몇 양태에서, 글루카나아제 활성은 엔도글루카나아제 활성(예를 들어 엔도-1,4-베타-D-글루칸 4-글루카노 하이드롤라아제 활성)으로서, 셀룰로스, 셀룰로스 유도체(예를 들어 카복시메틸셀룰로스 및 하이드록시에틸셀룰로스) 및 리케닌의 1,4-베타-D-글리코시드 결합, 혼합형 베타-1,3-글루칸, 예를 들어 곡물의 베타-D-글루칸 또는 크실로글루칸 및 기타 셀룰로스부 함유 식물성 물질의 베타-1,4 결합의 가수 분해 활성을 포함한다. 대안적 양태에서, 이와 같은 글루카나아제, 예를 들어 엔도글루카나아제, 만난아제, 크실라나아제의 활성과 안정성, 예를 들어 열 내성 또는 열 안정성은 증가 또는 감소된 pH 및 온도에서 증가하였다.
- [0114] 몇몇 구체예에서, 본원에 개시된 효소성 분해 물질은 이 효소를 안정화하고, 열 안정성을 개선하며, 알칼리 pH 내성을 증가시키고, 제어된 방식으로 방출되도록 피포화될 수 있다. 가수 분해되는 코팅 또는 막을 가지는 피포화 분해 물질은 선행 기술의 효소성 분해 물질 계보다 효능이 월등할 수 있는데, 그 이유는 상기 피포화 분해 물질은 이전에는 볼 수 없었을 정도로 다루기가 용이하고 방출 시간을 훨씬 더 잘 제어할 수 있기 때문이다. 예를 들어 상기 분해 물질은 단순히 물속에서 용해되거나 소모되기보다는 물과 반응을 하는 물질 내에 피포화되기 때문에, 방출 과정은 코팅과 물의 반응 속도를 통해 제어될 수 있다. 이와 유사하게, 임의의 기간 동안 극한의 땅속 조건(고온 및 고 pH)으로부터 효소를 격리함으로써 점도 내림 현상(viscosity break)을 지연 및 종결시킬 수 있다. 당 업자들은, 코팅의 반응 속도(즉 분해 물질의 방출 프로필)는 적용된 고분자 화학적 피포화 원리에 따라서 광범위하게 달라질 수 있음을 이해할 것이다. 분해 물질 피포화 조성물 및 피포화 방법의 예는, 본원에 그 자체로서 각각 참고 문헌으로 인용되어 있는 미국 특허 제5,164,099호, 동 제6,163,766호, 동 제5,373,901호, 동 제5,437,331호 및 동 제6,357,527호에 제시되어 있다.
- [0115] DNA 서열 및 코딩된 셀룰라아제 서열
- [0116] 초고온성 세균으로부터 유래하는 몇몇 셀룰라아제 및/또는 이의 비 자연 발생 변이체는 전체 내용이 본원에 참고용으로 인용되어 있는 PCT 공개 공보 WO 제2009/02049호에 개시되어 있다. 이하 나열된 DNA와 아미노산의 서열 번호들은, 그 자체가 본원에 참고용으로 인용되어 있는 공보 WO 제2009/02049호의 전체 명세서 내에 포함되어 있다. 그 예로서는 다음과 같은 것들을 포함한다:
- [0117] * WO 제2009/02049호 서열 번호 1, 2(야생형 '모' 티.마리티마 셀룰라아제), 본원에는 서열 번호 5 및 6으로 개

시되어 있음.

- [0118] * WO 제2009/02049호 서열 번호 3(야생형 DNA, 대안적 개시 점들이 없어지도록 변형됨), 본원에는 서열 번호 7로 개시되어 있음.
- [0119] * WO 제2009/02049호 서열 번호 6, 7("7X" 조합형 유전자 위치 포화 돌연 변이 유발법("GSSM") 돌연 변이), 본원에는 서열 번호 8 및 9로 개시되어 있음.
- [0120] * WO 제2009/02049호 서열 번호 8, 9("12X-6" 조합형 GSSM 돌연 변이), 본원에는 서열 번호 3 및 2로서 개시되어 있음.
- [0121] * WO 제2009/02049호 서열 번호 10, 11("13X-1" 조합형 GSSM 돌연 변이), 본원에는 서열 번호 10 및 11로서 개시되어 있음.
- [0122] * WO 제2009/02049호 서열 번호 12, 13("12X-1" 조합형 GSSM 돌연 변이), 본원에는 서열 번호 12 및 13으로서 개시되어 있음.
- [0123] * WO 제2009/02049호 서열 번호 16, 17(서모토가 종으로부터 유래하는 대안적 셀룰라아제 분해 물질), 본원에는 서열 번호 14 및 15로서 개시되어 있음.
- [0124] * WO 제2009/02049호 서열 번호 18, 19(티.마리티마 셀룰라아제의 "7X" 코돈-최적화형, 옥수수 발현용), 본원에는 서열 번호 16 및 17로서 개시됨.
- [0125] * WO 제2009/02049호 서열 번호 20, 21(티.마리티마 셀룰라아제의 "12X-6" 코돈-최적화형, 옥수수 발현용), 본원에는 서열 번호 18 및 19로서 개시됨.

[0126] 서모토가 마리티마 균주 MSB8로부터 유래하는 셀룰라아제의 유도된 변이체와 야생형과 관련된, 상기 나열된 뉴클레오티드 및 아미노산 서열들 이외에도, 표 2 및 실시예 5에 나열된 부가의 돌연 변이체들(WO 제2009/02049호에 개시, 이하 발체 또는 반복 개시됨)도 본원에 개시된 조성물 및/또는 이 조성물을 제조하는 방법의 구성 성분 및/또는 이에 관여하는 구성 성분으로서 유용한 것으로 판단된다.

[0127] 효소 활성 분석법(WO 제2009/02049호의 실시예 5)

[0128] 이하 예에는, WO 제2009/02049호에서 서열 번호 2로서 확인된 "모" 또는 "야생형" 단백질의 변이체, 예시적 효소, 그리고 이것들의 활성을 입증하는 데이터가 기술되어 있다[여기서, 수치와 데이터는 참고용으로 인용된 것임].

[0129] 예를 들어 서열 번호 5의 서열(WO 제2009/02049호의 서열 번호 1의 서열과 상응하는 서열)에 의해 코딩된, 서열 번호 6의 "모"(즉 "야생형") 서열(WO 제2009/02049호의 서열 번호 2의 서열과 상응하는 서열)에 특이적인 잔기 변화가 일어난 서열들이 제공되어 있다[상기 표 1, 하기 표 2 및 3에 (부분적으로) 요약되어 있음].

[0130] [표 2]

위 치	38	61	69	70	71	94	166	183	191	212	231	276	277	280	297	301
돌연변이	Y	Q	E	P	S	Q	V	R	A	P	V	A	S	G	P	Q
7X	Y	Q	E			Q		R	A			A				
10X-1	Y	Q	E			Q	V	R	A	P		A		G		
10X-2	Y	Q	E			Q	V	R	A	P		A		G	P	
11X-1	Y	Q	E			Q	V	R	A	P		A		G	P	
11X-2		Q	E			Q	V	R	A	P		A		G	P	Q
12X-1		Q	E		S	Q	V	R	A	P		A		G	P	Q
12X-2	Y	Q	E		S	Q		R	A	P		A		G	P	Q
12X-3		Q		P	S	Q	V	R	A	P		A		G	P	Q
12X-4	Y	Q			S	Q	V	R	A	P		A		G	P	Q
12X-5		Q	E	P		Q	V	R	A	P		A		G	P	Q
12X-6		Q	E	P	S		V	R	A	P		A		G	P	Q
12X-7		Q	E			Q	V	R	A	P	V	A		G	P	Q
13X-1	Y	Q	E		S		V	R	A	P	V	A		G	P	Q
13X-2	Y	Q	E	P	S	Q	V	R	A	P		A		G	P	Q
13X-3	Y	Q		P	S	Q	V	R	A	P		A		G	P	Q
13X-4	Y	Q	E	P		Q	V	R	A	P	V	A		G	P	
13X-5	Y	Q	E	P	S	Q	V	R	A			A		G	P	Q
13X-6	Y	Q	E	P		Q	V	R	A	P		A	S	G	P	
13X-7	Y	Q	E	P	S	Q		R	A	P		A		G	P	Q
14X	Y	Q	E	P	S	Q	V	R	A	P	V	A		G	P	

[0131]

제 뉴클레오티드는 1번으로 매겨진다. 돌연 변이들로서는 다음과 같은 것들이 있다: T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C.

- [0139] 본원에 개시된 셀룰라아제를 코딩하는 앞서 보고된 서열들과 서열 번호 1의 서열이 총체적이면서 개별적으로 구별되도록 만들어주는 염기 변화들을 통하여, 단백질 발현 수준이, 이전에 사용되었던 뉴클레오티드 서열로서 상기 해당 단백질과 동일한 단백질을 코딩하는 뉴클레오티드 서열로부터의 단백질 발현 수준보다 더욱 높아진 개방 해독 틀이 만들어진다.
- [0140] 서열 번호 1의 서열은 본원에 개시된 셀룰라아제 단백질(서열 번호 2)이 발현될 수 있는 다수의 계 내에서 발현 수준을 증가시키는데 사용된다. 서열 번호 1의 서열은 임의의 수의 발현계 내에 도입되어, 본원에 개시된 셀룰라아제를 증가한 축적 수준으로 발현시킬 수 있다. 예를 들어 서열 번호 1의 서열은 플라스미드에 도입될 수 있거나, 예를 들어 임의의 수의 그람-음성 세균 계, 예를 들어 이.콜라이, 슈도모나스 종, 예를 들어 플루오레센스, 슈도모나스 푸티다, 슈도모나스 아에루기노사, 랄스토니아 종 또는 카울로박터 종의 계놈에 안정적으로 도입되어 이 균주를 형질 전환할 수 있다. 이와 유사하게, 서열 번호 1의 서열은 임의의 수의 그람-양성 세균 발현계, 예를 들어 바실러스 종, 예를 들어 바실러스 서브틸리스, 바실러스 메가테리움, 바실러스 브레비스, 락토코커스 종, 예를 들어 락토코커스 락티스, 락토바실러스 종, 스트렙토마이세스 종, 예를 들어 스트렙토마이세스 리비단스에 도입될 수 있다. 기타 그람-음성, 그람-양성 또는 관련 없는 진정 세균 또는 고 세균 발현계도 서열 번호 1의 서열을 발현하는데 사용될 수 있다.
- [0141] 더욱 구체적으로, 서열 번호 1의 서열은 자체의 발현을 유도하도록 플라스미드에 도입될 수 있다. 서열 번호 1의 서열이 도입될 수 있는 플라스미드로서는, 예를 들어 pQE, pET 및 pASK 균의 이.콜라이 발현 벡터; pCN51 LT8, RSF1010, pWZ112T 및 pMYC 균의 슈도모나스 발현 벡터; pBAX, pHT01 및 pHIS1525의 바실러스 발현 벡터; pIJ6021 및 pIJ2460 균의 스트렙토마이세스 발현 벡터; 그리고 pNZ9530 및 pNZ8148 균의 락토코커스 발현 벡터를 포함한다. 이와 같은 예들은 예시적 목적으로 제시된 것이지만, 서열 번호 1의 폴리뉴클레오티드 서열이 발현될 수 있는 벡터의 세트 전체를 나타내는 것은 아니다.
- [0142] 증점제 및 기타 첨가제
- [0143] 몇몇 구체예에서, 본 발명의 고분자 증점제는 구아, 가교형 구아, 하이드록시프로필 구아, 카복시메틸 구아, 카복시메틸 하이드록시프로필 구아, 카복시메틸 셀룰로스, 보리 β-글루칸 및 메뚜기콩 검 중 하나 이상으로부터 선택될 수 있다. 몇몇 구체예에서, 본원에 개시된 셀룰라아제의 가수 분해 패턴은 엔도 작용 방식 및 엑소 작용 방식 둘 다이며, 이와 같은 패턴에 의해서(즉 길이가 긴 다당체 사슬 내부를 절단하기도 하고, 고분자 말단부로부터 이당체 단위씩 절단하기도 함으로써) 다당체(예를 들어 구아 및 유도체화된 구아 용액)의 점도를 효율적으로 감소시킬 수 있다. 전술된 다당체 이외에, 본원에 개시된 효소의 기타 기질로서는 선형 또는 가교형 겔을 형성할 수 있는 것들(예를 들어 봉산염 가교형 구아)을 포함한다. 적당한 다당체 기질의 예로서는 갈락토만난 검, 구아, 유도체화된 구아, 셀룰로스 및 셀룰로스 유도체, 전분, 전분 유도체, 잔탄, 유도체화된 잔탄, 그리고 이의 조합을 포함한다. 구체예로서는 또한 구아 검, 구아 검 유도체, 메뚜기콩 검, 카라야 검, 잔탄 검, 셀룰로스 및 셀룰로스 유도체 등도 포함하나, 이에 한정되는 것은 아니다. 본원에 개시된 효소들이 작용할 수 있는 통상의 고분자 증점제 또는 겔화제로서는 구아 검, 하이드록시프로필 구아, 카복시메틸 하이드록시프로필 구아, 하이드록시메틸 셀룰로스, 카복시메틸 하이드록시메틸 셀룰로스, 카복시메틸 셀룰로스, 디알킬 카복시메틸 셀룰로스 등을 포함한다. 고분자의 기타 예로서는 포스포만난, 스크레로글루칸, 텍스트란 및 기타 고분자 류를 포함하나, 이에 한정되는 것은 아니다. 몇몇 구체예에서, 고분자 기질은 카복시메틸 하이드록시프로필 구아이다. 몇몇 구체예에서, 본원에 개시된 효소는 또한 바이오 검(biogum)(예를 들어 데이트 시럽 또는 수크로스 제조된 숙시노글리칸 바이오 검)을 가수 분해하는데에도 효과적일 수 있다. 몇몇 구체예에서, 본원에 개시된 효소는 셀룰로스 함유 고분자 또는 유도체화 셀룰로스 함유 고분자를 가수 분해하는데 사용될 수 있는데, 통상적으로 상기 효소는 셀룰로스 주쇄의 글루코시드 결합을 공격한다. 본원에 개시된 효소는, 몇몇 경우 임의의 셀로비오스 단편의 (1,4)-β-D-글루코시드 결합 중 셀룰로스 주쇄와 단당체 단위 사이의 엔도(1,4)-β-D-글루코시드 결합 및 엑소(1,4)-β-D-글루코시드 결합을 특이적으로 가수 분해함으로써, 고분자를 대부분 단당체 단위로 분해하는데 적당할 수 있다.
- [0144] 본원에 개시된 파쇄액 조성물은 또한 부가 성분들, 예를 들어 유체 손실을 막고, 형성물이 손상되는 것을 줄이며, pH를 조정 및/또는 완충하고, 미생물 성장을 억제하며, 온도 안정성을 개선하는 제제를 포함할 수도 있다. 가교제로서는 봉산염, 티탄산염, 지르콘산염 및 알루미늄 화합물을 포함할 수 있다. 통상의 수압 파쇄 혼합물은 물 약 98%와 모래, 그리고 산, 향균제, 분해 물질(효소 또는 화학적 분해 물질/산화 분해 물질), 점토

안정화제, 부식 억제제, 가교제, 마찰 감소제, 겔화제(증점제), 철 생성 억제제, pH 조정제, 관석 형성 억제제 및 계면활성제로부터 선택되는 기타 첨가제 약 2%를 포함한다.

- [0145] 본원에 개시된 셀룰라아제의 파쇄액 중 성능
- [0146] 본원에 개시된 셀룰라아제를 사용하는 각각의 파쇄 작업에 있어서, 현장 작업자들은 처음에 산업 현장 실험실에서 효소 투여량 최적화 연구를 진행하게 될 것이다. 본원에 개시된 셀룰라아제는 농도 10~400ppm으로 희석되고 나서, 다당체 생성 화합물, 예를 들어 선행 또는 가교형 구아 검(25~60lb/1,000gal)과 혼합될 것이다. 사용시 조건(특히 고온, 고압 및 고 pH 조건이 조성되는 갱정)에 따라서, 구아 검 또는 유도체화된 구아 검은 종종 봉산염 가교제를 사용함으로써 가교된다. 그 다음, 이와 같은 최적화 연구로부터 얻어진 효소 투여량 정보는 실제 파쇄 작업시 활용될 것이다.
- [0147] 본원에 개시된 셀룰라아제 특유의 활성 프로파일은 깊은 갱정, 즉 고온 및 고 pH 조건이 조성된 곳에서 순조롭게 제어되는 방식으로, 구아를 주성분으로 하는 파쇄액의 완전한 가수 분해를 허용한다. 화학적 분해 물질에 비해, 본원에 개시된 본 발명의 셀룰라아제는, 극단적이면서 비 선택적인 화학적 분해 물질에 대한 비부식성이면서 환경 친화적인 대안을 제공한다.
- [0148] 효소성 분해 물질은 다년간 150°F 이하의 온도에서, 예를 들어 구아 겔을 비롯한 다당체 증점제를 가수 분해하는데 사용되어 왔다. 그러나 이보다 더 고온(예를 들어 200~250°F) 및 극한의 pH(≥ 10.5) 조건에서 작용할 수 있는 효소성 분해 물질이 업계 전반에서 요망되고 있다. 이와 같은 요망을 충족시키기 위해, 본 발명의 발명자들은 예상외로 열 안정성인 셀룰라아제를 포함하는 파쇄액을 개발하였다. 특히, 단백질 및 이 단백질을 코딩하는 DNA에 대해 기술되어 있는 PCT 공개 공보 WO 제2009/02049호를 참조. 모 효소로부터 유도된 셀룰라아제로서, 본원에 개시된 것은 세일 가스 및 오일/가스 갱정 심부에서 처할 수 있는 극한의 땅속 조건에서 확실히 차별화된 성능을 나타내는 것이 확인되었다.
- [0149] 본원에 개시된 셀룰라아제는 넓은 범위의 온도(80~250°F) 및 pH(3.0~10.5)에서 선행 및 봉산염 가교형 구아를 효과적으로 분해할 수 있다. 레올로지 테스트 결과는, 완전한 분해를 달성하기 위해서는 단지 소량(100ppm 이하)만이 투여되어야 한다는 것을 보여주고 있다. 효소 반응은 파쇄 작업시 온도 및 pH에 일어나는 변화에 의해 촉발될 수 있다. 본원에 개시된 셀룰라아제는 또한 효소 투여량을 조정함으로써 작업자가 이상적인 점도-시간 프로파일을 얻을 수 있도록 하는 우수한 투여량 반응을 보인다. 투여량 의존성 행동은 이른 점도 손실과 원치 않는 프로판트 차단 현상(proppant screen-out)을 방지하므로 매우 유리하다는 것이 입증될 것이다. 본원에 개시된 셀룰라아제는, 심지어 유액인 첨가제, 예를 들어 완충액, 염, 안정화제 및 가교제 등이 존재할 때조차도 점도를 효과적으로 감소시킴에 있어 활성인 상태로 남게 된다.
- [0150] 본원에 개시된 셀룰라아제 분해 물질은 구아 내 만노스 단위들 사이의 β -1,4 글리코시드 결합들을 특이적으로 표적화함으로써 겔의 점도를 감소시킨다. 탄수화물 프로파일링 테스트는, 상기 효소가 길이가 긴 구아 고분자들을 작고 가용성인 단편들로 효과적이면서 효율적으로 분해함으로써 추후 겔이 다시 뭉쳐지는 현상(healing)을 방지할 것이라는 사실을 입증하였다. 전도도 테스트는, 구아가 완전히 가수 분해되었음과, 형성물을 손상시키고 갱정의 전도도를 감소시키는 고분자 잔류물들이 제거되었음을 입증해준다.
- [0151] 구아 분해 반응에 관여하는 생물 촉매로서, 만난아제 활성을 나타내는, 본원에 개시된 셀룰라아제는 통상의 산화 화학적 분해 물질과 비교되는 몇 가지 중요한 이점들을 가진다. 첫 번째, 본 발명의 효소는 갱정 구멍, 형성물 또는 파쇄 장비에 원치 않는 반응을 일으키지 않고, 엔도- 및 엑소-베타 만난아제 활성을 통해 장쇄 구아 검 고분자를 특이적으로 분해한다. 두 번째, 본 발명의 효소는 반응에서 소모되지 않으며, 구아 고분자가 존재하는 한 이 고분자에 계속 작용을 할 수 있어서, 연장 및 제어된 분해 프로파일을 제공한다. 세 번째, 본 발명의 효소는 구아 고분자들을 크기가 훨씬 작은 단편들로 분해할 수 있으며, 그 결과 구아는 거의 완전히 분해되고, 그로 인해 잔류물도 훨씬 줄어들게 된다. 이러한 이점들로 인하여 과학자들은 극한의 자연 환경으로부터 초 열 안정성 셀룰라아제를 발견 및 특성 규명하는데에 심혈을 기울여 왔다. 서모토가 마리티마 종, 즉 열수공(hydrothermal vent) 샘플로부터 동정한 초고온성 세균은 장쇄 구아 검의 다당체를 이루는 β -1,4 결합을 특이적으로 절단할 수 있는 엔도-만난아제 활성을 가지는 셀룰라아제를 가진다[Mathur EJ, Lam DE, "Carboxymethyl cellulase from *Thermotoga maritima*", 미국 특허 제5,962,258호(1999), 동 제6,008,032호(1999) 및 동 제6,245,547호(2001)]. 구체적으로, 상기 효소는 작용 최적 온도 180°F를 보였는데, 이 온도는 글리코시드 하이드롤라아제 군에 속하는 기타 공지의 엔도-만난아제의 작용 최적 온도보다 상당히 높은 온도이다[Pereira JH, Chen ZW, McAndrew RP, Sapra R, Chhabra SR, Sale KL, Simmons BA, Adams PD, "Biochemical characterization and crystal structure of endoglucanase Cel5A from the hyperthermophilic *Thermotoga*

maritima", Journal of Structure Biology, 2010, 172; 372-379; Wu TH, Huang CH, Ko TP, Lai HL, Ma YH, Chen CC, Cheng YS, Liu JR, Gup RT, "Diverse substrate recognition mechanism revealed by *Thermotoga maritima* Cel5A structures in complex with cellotetraose, cellobiose and mannotriose", Biochimica et Biophysica Acta, 2011, 1814; 1832-1840].

- [0152] 시차 주사 열량 측정법(DSC)에 의해 평가되는 셀룰라아제의 열 안정성
- [0153] 효소는 자체의 촉매 기능에 중요한 3-차원 구조를 가진다. 임의의 조건(열, 변성제의 존재 및 극한의 pH) 하에서 효소는 펴지므로(unfolding), 자체의 촉매 활성 3-차원 구조는 사라지게 된다. 용점(Tm)이란, 단백질의 50%가 펴지게 되는 온도이다. 이 매개 변수는 효소나 단백질의 열 안정성을 평가하는데 널리 사용되고 있다. Tm은 시차 주사 열량 측정법(DSC)을 통해 측정될 수 있다.
- [0154] 셀룰라아제 야생형 단백질과 선택된 셀룰라아제 변이체를 사용하여 DSC 테스트가 수행되었다. 모든 단백질 샘플들이 1mg/ml의 농도와 주사 속도 1°C/분에서 분석되었다. 각각의 주사가 진행될 때 온도 범위는 160~250°F였다. 모든 DSC 실험이 진행되는 동안에 가열시 일어날 수 있는 용액의 탈기 현상을 방지하기 위해서, 압력은 4.6atm으로 일정하게 유지되었다. Tm은 입수 가능한 소프트웨어 패키지로써 산정되었다. 도 7은, 야생형과 선택된 셀룰라아제 변이체의 Tm 측정 결과들을 보여주는 것이다. 마지막 후보 셀룰라아제에서 관찰되는 Tm 수치는 야생형 효소에서 관찰되는 Tm 수치보다 24°F 증가한다. 상기 Tm은 pH에 의존적이다. 변이체와 야생형 사이의 Tm 차는, pH6.5에서는 24°F이었으며, pH10.5에서는 13°F이었다.
- [0155] 화학적 대용물을 사용하는 셀룰라아제 변이체의 온도 프로파일링
- [0156] 셀룰라아제 변이체는 다양한 온도에서 자체의 1-4 결합 절단 활성에 대해 추가로 평가되었다. 분자 pNP-β-D-락토피라노시드가 대용 기질로서 사용되었다(도 8a). pNP-β-D-락토피라노시드의 셀룰라아제에 의한 효소 가수 분해를 통하여 유리 p-니트로페놀이 생성될 수 있는데, 이는 405nm에서 분광 분석법을 통해 확인될 수 있다. 도 8b는, 분광 분석법에 있어서 온도 한계치인 194°F 이하 온도에서의 셀룰라아제 변이체의 온도 프로필을 보여주는 것이다. 도 8b에 보인 바와 같이, 이 셀룰라아제 변이체의 최적 온도는 180°F이다.
- [0157] 비등점 이상의 온도(> 212°F)에서의 셀룰라아제 활성을 평가하기 위해, 셀룰라아제 변이체의 잔효성을 정량적으로 측정하는 분석법이 개발되었다. 효소는 1000gal당 25lb의 구아와 혼합되고(ppt)(1:1 비율), 이후에 225°F, 250°F 및 275°F에서 각각 10분, 20분 및 30분 동안씩 열 도전(heat challenge)이 진행되었다. 그 다음, pNP 흡광도 분석법(즉 열 처리된 효소 30uL을 2mM pNP-B-D-락토피라노시드 3000uL에 첨가한 후, 10분 동안 405nm에서의 흡광도를 측정함)을 이용하여 열 도전된 효소의 활성이 측정되었다. 도 8c는, 225°F, 250°F 및 275°F에서 셀룰라아제 변이체의 잔효성을 보여주는 것이다. 225°F에서는 셀룰라아제가 10~30분 동안 열 처리되었을 때 이 효소의 활성이 매우 컸던 것으로 관찰되었다. 열 처리된 셀룰라아제가 봉산염 가교형 구아에 적용되었을 때 (1000gal당 25lb), 상기 효소가 소량(20gpt)으로 존재하더라도 겔이 완전히 분해되는 것이 관찰되었다. 셀룰라아제 활성은 250°F 및 275°F에서 떨어졌는데(이 경우, 상대적 잔효성은 225°F에서의 잔효성의 11~14%에 불과하였음)(도 8c), 이는 고온(>250°F) 처리 이후 셀룰라아제가 낮은 백분율(%)로 존재할지라도 여전히 측정 가능한 촉매 반응을 진행시킬 수 있음을 암시하는 것이다.
- [0158] 형광 대용물을 사용하는 셀룰라아제 변이체의 pH 프로파일링
- [0159] 분자 4-메틸움벨리페릴 β-D-셀로비오시드는, 높은 pH 조건에서 셀룰라아제에 의해 절단될 수 있는 대용 기질이다(도 8a). 절단 생성물, 즉 4-메틸움벨리페론은 형광 분석적으로 정량될 수 있다[이때 적용된 여기 파장은 365nm이고 발광 파장은 455nm임]. 4-메틸움벨리페릴 β-D-셀로비오시드 2mM이 완충계에 용해되었는데, 이때 pH 수치는 9.5~10.5였다. 셀룰라아제 30uL를 2mM 4-메틸움벨리페릴 β-D-셀로비오시드 3000uL에 첨가하여, 상이한 pH 조건에서의 셀룰라아제 활성이 분광 분석계에서 동역학적으로 측정되었다. 선형 곡선 피팅법(linear curve fitting)을 통해 각각의 반응 곡선 기울기가 구하여졌다. 상대적 활성은 pH9.5에서 구하여진 결과들에 따라서 산정되었다. pH10.5, 즉 봉산염 가교형 구아 유액과 밀접하게 관련된 pH 조건에서, 셀룰라아제는 합리적인 수준의 활성을 보인다(도 9).
- [0160] HPLC를 통한, α-1,6-결합 다당체가 아닌 β-1,4-결합 다당체의 효소 분해 확인
- [0161] 만노스와 갈락토스 당으로 이루어진 장쇄 다당체인 구아는 물을 기반으로 하는 파쇄 작업에 주로 사용되는 증점제이다. 폴리만노스는 β-1,4-결합을 통해 장쇄인 주쇄를 형성하는데, 이때 갈락토스 단위는 α-1,6-결합을 통해 만노스 단위에 결합한다. 만노스 당 대 갈락토스 당의 비율은 1.6:1~1.8:1일 수 있다. 중요한 사실은, 구아의 폴리만노스 주쇄는 수 불용성이고, 갈락토스 분지들은 수용성을 상당 수준 증가시킨다는 것이다. 그러나, 6

개에 불과한 연속 비 분지형 만노스 단위는 완전히 불용성인 국소 나선형 구조를 형성할 수 있다고 보고된 바 있다(11). 그러므로 만일 셀룰라아제가 구아 분자의 α -1,6 결합을 분해하면, 불용성 잔류물이 유의적인 양만큼 생성될 수 있다. 결과적으로 프로판트 팩(proppant pack)의 전도도는 저하될 수 있다.

[0162] 셀룰라아제 변이체의 효소 활성이 다당체의 β -1,4 글리코시드 결합을 특이적으로 표적화한다는 것을 확인하기 위해서, 정상 HPLC 방법이 채택되었으며, 이로써 2개의 대용 기질들, 즉 1,4- β -D-만노펜타오스 및 6³,6⁴- α -D-갈락토실-만노펜타오스가 분석되었다(도 10a). 셀룰라아제 변이체(3유닛/ml)는 기질 1mg/ml와 함께 80°C 및 pH7.0에서 1시간 동안 항온 처리되었다. 만난 올리고당체에 대한 기준 분리(baseline separation)가 아미드 컬럼, 악토니트릴/H₂O/트리에틸아민(이동 상) 및 ELSD 검출 장치가 사용되는 HPLC를 통하여 이루어졌다. 도 10b는, 상이한 만난 올리고당체, 예를 들어 M1, M2, M3, M4, M6와, 갈락토스 표준들의 HPLC 용리 프로필을 도시한 것이다(M1~M6은 만난 올리고당체 사슬의 길이를 나타냄). 도 4c는, β -1,4-D-만노펜타오스, 즉 자체의 구조 내에 오로지 β -1,4 결합만을 포함하는 만난 올리고당체가 셀룰라아제 변이체에 의해 어떻게 분해되는지에 대한 일례를 보여주는 것이다. HPLC 프로필은, β -1,4-결합이 절단되면 M1, M2 및 M4가 생산됨과 아울러 바람직하게는 M3 올리고당체가 생산되는 것을 입증해주는데, 이는 만난 올리고당체가 분해될 때 셀룰라아제 변이체가 β -1,4-결합을 특이적으로 공격한다는 것을 확인시켜주는 것이다(도 10c의 (a)~(b)). 다른 한편, 6³,6⁴- α -D-갈락토실-만노펜타오스, 즉 β -1,4-결합과 α -1,6-결합을 둘 다 포함하는 화합물의 경우, HPLC 프로필은, 셀룰라아제 변이체가 오로지 β -1,4-결합만을 분해하여 OGGM4 피크와 하나의 만노스 단위의 피크를 나타낸다는 것을 보여준다(도 10c의 (a)). 기준으로서 갈락토스가 사용되면, 효소 처리된 6³,6⁴- α -D-갈락토실-만노펜타오스로부터 유래하는 하나의 갈락토스 단위에 대해서는 피크가 관찰되지 않았는데, 이는 만난아제 처리 후에도 α -1,6-결합은 원 상태 그대로 남아있다는 것을 확인시켜주는 것이다(도 10c의 (b)).

[0163] 분해 물질을 이용하는 가교형 구아 유액의 제조

[0164] 와류가 발생되기에 충분한 정도로 혼합기 내에서 교반된 물에 무수 고분자를 첨가함으로써 선형 구아 유액이 제조되었다. 약 30분 동안 계속해서 교반되어 수화가 완료되었다. pH는 수산화나트륨 용액으로 조정되었다. 가교된 유액, 점토 안정화제 및 세정용 계면활성제가 사용될 때, 이것들은 구아가 완전히 수화된 이후에 첨가되었다. 그 다음, 가성 물질이 첨가되었고, 이후에는 효소가 첨가되었다. 가교제는 마지막에 첨가되었다. 활성 봉소 종이 천천히 용해됨으로써 봉산염 가교제는 작용이 지연되었다. 지르코늄 가교제는 또한 세스퀴탄산염 완충액이 6ppt 만큼 사용됨으로 인해 작용이 지연되었다.

[0165] 병 테스트

[0166] 과황산암모늄(APS) 또는 셀룰라아제 분해 물질과 함께 구아 또는 가교된 구아가 담긴 자(jar)가 제조되었다. 자는 수조 내에서 숙성되었으며, 분해 수준은 병을 기울여 유액의 이동 양상을 관찰함으로써 눈으로 평가되었다. 병 테스트를 통하여, 소정의 과쇄액의 분해에 대한 농도, pH 및 온도의 영향력을 신속히 평가하는 방법이 실현될 수 있었다. 테스트는 원래 정성 분석을 위해 실시되었으므로, 결과들은 보이지 않았다.

[0167] 레올로지

[0168] 점도계는 컵(cup) 및 밥(bob) 기술(로터 1 및 밥 5)[모델 50 명세서 포함, 질소 압력 400psi 하에서 실시]을 이용하여 점도를 측정하는 기계이다. 유액의 온도는 오일 조를 통해 원하는 수치까지 증가하였다. 가열되어 원하는 온도에 도달될 때까지 20분이 소요되는데, 이 온도는 테스트가 진행되는 동안 유지되었다. 점도는 벽 전단 속도(wall shear rate) 100s⁻¹로 이동하는 회전 속도에서 계속 측정되었으며, 이때 전단 램프는 100, 75, 50, 25, 50, 75 및 100s⁻¹ 주기로 작동되었다. 전단 램프 테이터는 과쇄시 일반적으로 관찰되는 낮은 전단 속도들에서의 점도 예상을 위한 멱 법칙 매개 변수(n' 및 K')를 산정하는데 사용된다.

[0169] 도 11은, 180°F에서 1000 꺾론의 선형 구아 용액당 80lb를 분해하는데 필요한 활성의 범위를 보여주는 것이다. 셀룰라아제 농도는 200ppm으로 일정하게 유지되었다. pH가 9에서 11로 증가됨에 따라서 분해는 확연히 지연되었음이 관찰되었는데, 이와 같은 결과는 높은 pH 수준에서 효소 활성이 낮아지는 것과 일치한다. 심지어 pH 11일 때조차도 상기 용액은 효소를 포함하지 않는 대조군과 비교되게 느린 분해 효과를 보인다.

[0170] 도 12는, 200°F 및 pH10.5에서 봉산염 가교 구아 유액 분해에 효소 농도가 미치는 영향력을 보여주는 것이다. 셀룰라아제 투입량이 증가하면, 점도가 더욱 많이 내려간다. 경시적으로 점도가 계속 손실되는 것을 보기 위해 수 시간 동안 레올로지 결과가 관찰되었다. 구아가 고농도로 포함된 선형 겔과 비교되게, 가교 겔에는 더욱 많

은 효소가 필요함이 분명하다. 셀룰라아제 수준이 50ppm인 가교 유액의 점도는 대조군 가교 유액의 점도와 구별이 되지 않았던 반면, 효소의 수준이 이보다 높았을 때에는 분해 활성이 보였다.

[0171] 중요한 사실은, 열 안정화제가 10ppt의 수준으로 포함될 때 225°F 및 pH10에서도 효소가 여전히 활성을 띤다는 것이다(도 13). 대조군은 점진적 열 열화를 보이지만, 효소는 분명히 점도 저하 과정(viscosity degradation)을 가속화한다. 도 14는, 225°F 및 pH10에서 셀룰라아제 및 피포화된 과황산암모늄의, 카복시메틸하이드록시에틸셀룰로스(CMHPG)를 사용하는 지르코늄 가교 유액에 대한 비교 데이터를 제시하고 있다. 농도 100ppm은 과황산암모늄에 유사한 분해 프로필을 제공하는 반면에, 농도 200ppm은 점도의 즉각적인 손실을 초래한다. 피포화된 과황산암모늄은 지연 분해를 초래하는 피포화 물질의 열 분해 및/또는 이 피포화 물질을 통과하는 확산을 통해 활성 물질 일부를 방출한다는 것이 주목되어야 할 것이다.

[0172] 잔류물 분석

[0173] 과황산암모늄 또는 셀룰라아제 분해 물질 중 어느 하나와 함께 구아 또는 가교 구아가 담긴 자가 마련되었다. 자는 밤새도록 180°F의 수조 내에서 숙성되었다. 내용물은 3000rpm에서 5분 동안 원심 단리되었으며, 공칭 두께가 5 마이크로미터이고 무게가 나가는 페이퍼를 통해 진공 여과되었다. 상기 페이퍼는 110°F에서 밤새도록 건조되었으며, 이후 회수된 잔류물의 양을 산정하기 위해 페이퍼 무게가 다시 측정되었다. 이 수치는 용액 중 고분자의 원래 중량의 백분율로서 표현된다. 그 다음, 잔류물이 남아있는 필터 페이퍼는 16시간 더 건조되었으며, 수분이 증발되었는지 확인하기 위해 상기 필터 페이퍼의 무게가 측정되었다. 그러나, 데이터는 추가 건조 시간 16시간은 불필요한 것으로 보임을 암시하고 있다. 테스트된 유액은 pH10에서 선형 구아를 80ppt 포함하였다. 셀룰라아제는 200ppm 및 400ppm에서 테스트되었으며, 이때 과황산암모늄은 1ppt 및 5ppt 사용되었다.

[0174] 결과들은, 놀랍게도 셀룰라아제를 함유하는 용액이 이와 상응하는 용액으로서 과황산암모늄을 포함하는 용액보다 잔류물을 소량 포함하였음을 말해준다(도 15). 이와는 대조적으로, 산화 분해 물질은 더욱 심하게 무작위로 절단하는 것으로 보이며, 그 결과 잔류물로서 포획되는 불용성 단편들이 더욱 높은 수준으로 생성된다. 만일 자가 오랜 기간 동안 특정 온도(즉 180°F)에 방치되면, 효소 처리된 유액의 경우 잔류물은 훨씬 많이 감소할 것이지만, 산화 분해 물질을 포함하는 유액의 경우에는 그렇지 않을 수 있다. 잔류물은 과황산암모늄과 같은 분해 물질의 농도에 감수성이나, 이 잔류물의 효소 농도에 대한 감수성은 훨씬 작다.

[0175] 전도도 연구

[0176] 파쇄 전도도는 바닥에 피스톤이 고정되어 있는 변형 API 셀을 사용하여 측정되었다. 조건은 다음과 같았다: 180°F, 닫힘 응력(closure stress) 3000psi, 프로판트(20/40 메쉬 오타와 샌드(Ottwa sand)) 로딩률 21bm/ft², 그리고 누출 후 파쇄액 농축 배수 약 5배. 유액은 이 온도에 밤새도록 방치되었으며, 이후 4시간 동안 세정이 진행되었고, 그 다음 최소 1시간 동안 팩 흐름(pack flow)이 진행되었다. 세정은, 팩을 통과하는 흐름과 코어를 통과하는 흐름이 1시간마다 교대로 진행되는 방식으로 행하여졌으며, 이때 2wt% KCl 용액이 사용되었다. 2wt% KCl(유속 3ml/분)을 사용하여 최종 전도도 측정이 수행되었다(1시간 동안 4% 이내에서 안정적). 체류 전도도(retained conductivity)는 동일 조건에서 파쇄액 대신 2wt% KCl을 사용하여 진행된 전도도 측정 결과와 비교함으로써 산정되었다.

[0177] 선형 구아 및 붕산염 가교 구아가 과황산암모늄 또는 셀룰라아제와 함께 전개되었으며, 하나의 테스트에서는 상기 분해 물질 둘 다와 가교 구아가 사용되었다. 선형 구아의 경우 셀룰라아제가 사용되었을 때가 과황산암모늄이 사용되었을 때보다 높은 수준의 체류 전도도가 달성되었다. 이와 유사하게, 피포화된 과황산암모늄은 가교 붕산염 유액에 대해 200ppm의 셀룰라아제가 달성한 체류 전도도보다 더욱 낮은 체류 전도도를 달성하였다(도 16). 셀룰라아제와 피포화된 과황산염이 함께 사용되었을 때에는 이 분해 물질들 중 어느 하나만이 사용되었을 때보다 체류 전도도가 훨씬 더 높았다. 그러나 효소가 400ppm의 투여량으로 투여되었을 때 체류 전도도가 가장 높았다. 전도도 테스트는 밤새도록 투여되면서 진행되어야 하는 것이므로, 셀룰라아제는 구아의 분자량을 감소시키는 시간이 따로 필요하였다. 이는, 프로판트 팩의 전도도를 직접적으로 증가시키게 될 것이다.

[0178] 요약하면, 유도된 셀룰라아제 변이체는 붕산염 가교 구아 겔 유액 중 다양한 첨가제와 혼화 가능하다. 특히 이 셀룰라아제 변이체는 180~225°F의 온도 및 pH 10.5 이하일 때 탁월한 구아 겔 분해 프로필을 보인다. pH 제한은 더욱 높은 온도 범위에서 붕산염 가교 구아 유액에 대한 상기 효소의 테스트를 방해한다. 그러나 225°F의 온도에서 지르코늄 가교 CMHPG가 사용되면 양성 구아 분해 프로필이 관찰되었다. 주목할 점은, 잔류물 분석을 통해, 만난아제 처리된 유액에는 과황산암모늄 처리된 유액보다 훨씬 적은 잔류물이 생산된다는 것이 입증되었다는 점이다. 더욱 중요한 점은, 셀룰라아제 처리는 과황산암모늄 처리보다 훨씬 높은 체류 전도도를 달성하였다

는 점인데, 이는 효소가, 화학적 산화제에 비하여 구아를 더욱 완전히 분해할 수 있고, 갱정 조건을 개선할 수 있다는 것을 확신시켜주는 것이다. 더욱이, 셀룰라아제는 과황산암모늄과 함께 첨가되어 사용될 수 있으며, 이러한 점은 임의의 유전에서 사용됨에 있어 유용할 것이다.

도면의 간단한 설명

[0179]

도 1은, 실시예 2에 기술된 바와 같이 다양한 단백질 발현 수준을 보여주는 SDS PAGE 겔 전기 영동 이미지이다.

도 2는, 실시예 3에 기술된 바와 같이 단백질 제제의 활성 수준을 보여주는 막대 그래프이다.

도 3은, 서열 번호 1의 서열로서, 서열 번호 3의 서열과 비교되게 다음과 같은 14개의 침묵 돌연 변이들을 가지는 폴리뉴클레오티드이다: T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C.

도 4는, 서열 번호 2의 서열로서, 서열 번호 1, 3 및 4에 의해 코딩된 폴리펩티드이다.

도 5는, 서열 번호 3의 서열로서, 서열 번호 1 및 4의 비 변형 모 폴리뉴클레오티드 서열이다.

도 6은, 서열 번호 4의 서열로서, 서열 번호 3의 서열과 비교되게 14개의 침묵 돌연 변이[T6C, T9C, T15G, A22C, G24T, A33C, A39C, A40C, A42C, A54C, A57C, T66C, G81A, A84C]를 포함하고, 게다가 개시 코돈으로부터 상류 쪽으로 하나의 점 돌연 변이를 추가로 포함하는 폴리뉴클레오티드이다(추가 상류 서열이 표시됨).

도 7은, 야생형 셀룰라아제와 선택된 셀룰라아제 변이체의 DSC 테스트가 수행된 결과이다(산정된 Tm 수치가 표지화됨). 셀룰라아제 변이체가 사용되었을 경우, 야생형 셀룰라아제가 사용되었을 경우에 비해 Tm이 24°F 증가한 것이 관찰되었다.

도 8a는, pNP-β-D-락토피라노시드의 화학 구조이다.

도 8b는, pNP 흡광도 분석법(405nm)에 의해 측정된, 상이한 온도(70~200°F)에서의 셀룰라아제 온도 프로파일링 결과이다. 활성은, 180°F에서의 활성 백분율로서 산정되었다. 비등점 이상의 온도(>212°F)에서의 활성은 잔효성 분석으로 더 자세히 평가될 수 있다(도 8c).

도 8c는, 열 도전된 샘플이 사용된, 셀룰라아제 변이체의 잔효성 평가 결과이다. 상대적 활성은 225°F에서 얻어진 활성의 백분율로서 산정되었다.

도 9는, 형광 분석법을 통하여 셀룰라아제 변이체의 pH를 프로파일링한 결과이다. 상대적 활성은 pH9.5에서의 활성의 백분율로 산정되었다.

도 10a는, HPLC 연구에 사용된 화합물 2개의 구조를 보여주는 것이다. 1,4-β-D-만노펜타오스는 만노스 단위 5개가 1,4-β 결합에 의해 연결되어 있는 화합물이다. 6³,6⁴-α-갈락토실-만노펜타오스는 갈락토스 단위 2개가 α-1,6 결합을 통해 1,4-β-D-만노펜타오스에 연결되어 있는 구아 검 구조를 모의한 화합물로서, 셀룰라아제 처리로부터 생성된 가수 분해 생성물들의 분석을 가능하게 해준다.

도 10b는, 올리고당체 표준(M1~M6 및 갈락토스)의 HPLC 분석 결과를 보여주는 것이다. 만노스 단위 하나(M1, 적색 피크)는, HPLC의 용리 프로필 중 갈락토스 단위 하나(청색 피크)와 구분이 잘 된다는 점이 주목되어야 할 것이다.

도 10c는, 만노스 처리 후 1,4-β-D-만노펜타오스 및 6³,6⁴-α-갈락토실-만노펜타오스로부터 생성되는 가수 분해 생성물들을 HPLC 분석한 결과를 보여주는 것이다. a) 1,4-β-D-만노펜타오스의 효소 분해 결과. 효소 처리되지 않은 대조군 샘플의 HPLC 용리 프로필은 적색으로 표시되어 있으며, 셀룰라아제 변이체로 처리된 1,4-β-D-만노펜타오스는 청색으로 표시되어 있다. b) 6³,6⁴-γ-D-갈락토실-만노펜타오스의 효소 분해 결과. O-GGM5는 셀룰라아제 처리된 샘플(적색으로 표시)과 효소 처리되지 않은 샘플(청색으로 표시)의 6³,6⁴-α-D-갈락토실-만노펜타오스를 나타낸다. 하나의 만노스 단위(핑크색으로 표시)와 갈락토스 단위(녹색으로 표시)는 참고 기준으로 포함되어 있다. 셀룰라아제 처리된 O-GGM5의 HPLC 용리 프로필 중에는 갈락토스 피크는 보이지 않고 오로지 만노스 피크만이 보이는데, 이는 셀룰라아제 변이체가 만노스 단위들 사이의 β-1,4-글리코시드 결합들만을 특이적으로 절단한다는 것을 확신시켜주는 것임이 주목되어야 한다.

도 11은, 180°F에서 US 1000gal당 구아 801bm의 점도에 pH가 미치는 영향력을 보여주는 것이다. 명확성을 도모

하기 위해 전단 램프는 본 도면에서 제외되었다.

도 12는, 200°F 및 pH10.5에서 붕산염 가교 구아에 대한 셀룰라아제 변이체의 레올로지 연구 결과를 보여주는 것이다. 모든 유액들은 점토 안정화제와 계면활성제를 포함하였으며, 지연 가교제가 사용되었다.

도 13은, 225°F 및 pH10에서 붕산염 가교 구아에 대한 셀룰라아제 변이체의 레올로지 연구 결과를 보여주는 것이다. 모든 유액들은 US 1000gal당 열 안정화제 101bm를 포함하였으며, 지연 가교제가 사용되었다.

도 14는, 225°F 및 pH9.9에서 지르코늄 가교 CMHPG에 셀룰라아제 변이체가 사용되었을 때의 레올로지 연구 결과를 보여주는 것으로서, 이때 US 1000gal당 열 안정화제 101bm가 사용되었다.

도 15는, 과황산암모늄과 셀룰라아제가 사용되었을 때의 잔류물 분석 결과들을 비교한 것이다. 분해 물질의 수준이 높으면, 상기 과황산암모늄과 셀룰라아제가 사용된 경우 둘 다에서 잔류물이 감소하였지만, 산화 분해 물질이 사용된 경우에는 잔류물 양이 많았다.

도 16은, 선형 구아 제제 및 붕산염 가교 구아 제제 둘 다에 셀룰라아제와 과황산암모늄이 사용되었을 때 전도도를 비교한 결과를 보여주는 것이다. 선형: 구아 801bm/US 1000gal, 셀룰라아제 또는 생 과황산암모늄(live ammonium persulfate) 사용; 가교형: 구아 301bm/US 1000gal, 셀룰라아제 또는 피포화된 과황산암모늄 사용.

발명을 실시하기 위한 구체적인 내용

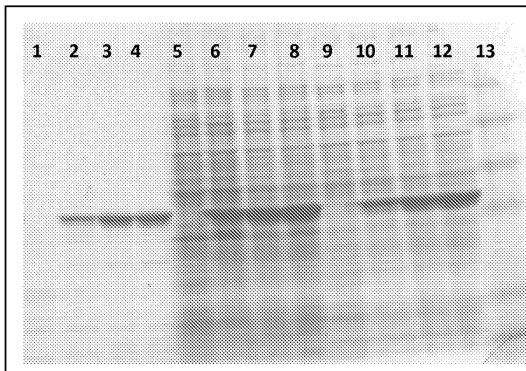
- [0180] 용어의 정의:
- [0181] "셀룰라아제"는 셀룰라아제, 엔도글루카나아제, 셀로비오하이드롤라아제, 베타-글루코시다아제, 크실라나아제, 만난아제, β-크실로시다아제, 아라비노푸라노시다아제 및/또는 올리고머라아제 활성을 가지는 효소이다.
- [0182] "셀룰로스 분해 활성"은 셀룰라아제, 엔도글루카나아제, 셀로비오하이드롤라아제, 베타-글루코시다아제, 크실라나아제, 만난아제, β-크실로시다아제, 아라비노푸라노시다아제 및/또는 올리고머라아제 활성을 가지는 효소의 활성이다.
- [0183] "코돈"은 단백질에 부가될 아미노산의 정체성을 특정하는 3 폴리뉴클레오티드 서열이다.
- [0184] "침묵 돌연 변이"는 아미노산의 본질을 상이하게 하지 않는, 코돈 내 돌연 변이이다.
- [0185] "개방 해독 틀"은 단백질 내 아미노산 서열을 특정하는 일련의 코돈들이다.
- [0186] 염기 "위치"는 개방 해독 틀이 시작되는 지점 또는 기타 몇몇 기준 마커들로부터 연속적으로 매겨지는, 폴리뉴클레오티드 내 염기의 수치상 자리이다.
- [0187] 단백질을 "코딩하다"란, 해당 단백질의 아미노산 서열을 특정하는 것을 의미한다.
- [0188] "돌연 변이"는 기준 서열과 비교되었을 때 뉴클레오티드 서열 또는 아미노산 서열에 일어난 변화이다.
- [0189] "뉴클레오티드"란, DNA 서열에 포함된 4개의 염기들(아데닌(A), 티미딘(T), 구아닌(G) 및 시토신(C)) 중 하나를 말한다.
- [0190] "서모토가 마리티마 게놈 서열"이란, 서모토가 마리티마 균주 MSB8 게놈 서열을 말한다(유전자 은행 기탁 번호 AE000512로 특정됨).
- [0191] 소정의 단백질에 있어서 "발현 수준"은, 발현계, 예를 들어 형질 전환된 세포 배양액에 의해 생산된, 세포 배양액 단위 부피당 측정되는 단백질의 양이다.
- [0192] 소정의 효소에 있어서 "발현 수준"은, 발현계, 예를 들어 형질 전환된 세포 배양액에 의해 생산된, 세포 배양액 단위 부피당 측정되는 효소 활성의 양이다.
- [0193] "야생형"이란, 자연에서 얻어질 수 있는 단백질 또는 핵산 서열을 말한다.
- [0194] 실시예 1 - 발현이 증가한 변이체를 제조하는 방법
- [0195] 유전자 발현을 개선하기 위해 DNA 수준에서 돌연 변이시킨 서열 번호 3을 바탕으로 2개의 변이체(서열 번호 1 및 서열 번호 4)를 디자인하였다. 디자인 시 유전자 발현에 영향을 미칠 수 있는 다수의 요소들을 고려하였다. 당 업자들에게 알려진 PCR 기술을 이용하여 PCR 프라이머에 돌연 변이를 도입하였다. 상기 유전자 둘 다를 PCR 증폭한 다음, 표준 분자 클로닝 기술을 이용하여 슈도모나스 벡터 pDOW1169(인디애나 소재, 다우 애그로사이언

시스(DOW AgroSciences))에 클로닝하였다. 생산된 발현 구조체를 슈도모나스 플루오레센스 DC454(인디애나 소재, 다우 애그로사이언시스)에 도입하여 형질 전환을 수행하였다. 서열 번호 1의 서열이 도입된 형질 전환체에서 발현이 가장 많이 증가하였으므로 이 형질 전환체를 선두로 정하였다.

- [0196] 실시예 2 - 서열 번호 1, 3 및 4의 서열들을 포함하는 구조체에 의해 발현된 서열 번호 2의 폴리펩티드의 발현 수준을 높이기 위한 비특이적 단백질 염색 및 SDS-PAGE 겔 전기 영동의 실시
- [0197] 크리테리온(Criterion)TM 예비 성형 트리스-HCl 폴리아크릴아미드 겔(바이오-래드 래보라토리즈 인코포레이티드 (Bio-rad Laboratories, Inc.))을 사용하여 단백질을 분리하였다. 겔에 150V의 전압을 걸어주고 트리스-글리신 완충액을 사용하여 전계를 진행하였다(도 1 참조). 각각의 레인에 대해 단백질 로딩량을 정규화하여 OD₆₀₀ 0.33이었던 세포로부터 유래하는 단백질을 로딩하였다. 씨블루(SeeBlue)[®]로 예비 염색한 단백질 표준을 사용하였다(라이프 테크놀로지스(Life Technologies)). 겔을 비특이적 염료로 염색하였으며, 각각의 레인에 서열 번호 2의 크기 약 37 kDa인 밴드가 나타났는지를 눈으로 관찰하였다.
- [0198] 결과들은, 축적 수준을 나타내는 밴드(샘플들마다 다르며, 음성 대조군에는 보이지 않음)가 하나 존재함을 보여 준다. 상기 밴드의 크기는 서열 번호 2의 서열에 대해 예상되는 크기였다.
- [0199] 상기 밴드의 축적 수준은, 서열 번호 1을 포함하는 구조체 보유 세포들로부터 유래하는 단백질 추출물들이 로딩된 레인들에서는 유의적으로 높았으나, 서열 번호 4를 포함하는 구조체 보유 세포들로부터 유래하는 단백질 추출물들이 로딩된 레인에서는 이보다 낮았으며, 서열 번호 3 또는 음성 대조군을 포함하는 구조체 보유 세포들로부터 유래하는 단백질 추출물들이 로딩된 레인에서도 마찬가지였다.
- [0200] 실시예 3 - 변이체들에 대한 상대적 발현 수준을 측정하는 방법
- [0201] 서열 번호 1, 서열 번호 3 및 서열 번호 4 유전자를 포함하는 핵산 서열을, 서열 번호 2의 단백질의 발현에 적당한 숙주 세포에 도입하여 이 세포를 형질 전환시켰다. 세포들을 플라스크 내에서 배양하여 코딩된 단백질이 발현되도록 만들었다. 배양액을 30℃ 및 220rpm에서 생육하여, 디자인된 복합 배지 중 OD₆₀₀이 약 0.9가 되도록 만든 다음, 0.3mM IPTG(이소프로필 β-D-1-티오갈락토피라노시드)로 24시간 동안 유도하였다. 세포들을 수집한 다음 음과 처리 또는 열 처리(80℃, 1시간)하여 용해하였다. pNP-β-D-락토피로아노시드를 기질로 사용하는 p-니트로페닐(pNP) 기반 분석법으로 셀룰라아제 활성을 측정하였다[Advances in Carbohydrate Chemistry and Biochemistry, Academic Press, 1999]. 활성 수준을 측정하여 도 2에 보였는데(U/ml), 이로써 각각의 배양액에서의 상대적 발현 수준을 측정할 수 있었다.
- [0202] 결과들은, 서열 번호 1의 서열을 포함하는 구조체를 보유하는 세포는, 서열 번호 4의 서열을 보유하는 세포보다 서열 번호 2의 활성을 유의적으로 높게 나타냈으며, 또한 서열 번호 1의 서열과 서열 번호 4의 서열 둘 다를 보유하는 세포는, 서열 번호 3의 서열을 보유하는 세포보다 발현된 단백질의 활성을 더욱 높게 보였음을 말해준다.

도면

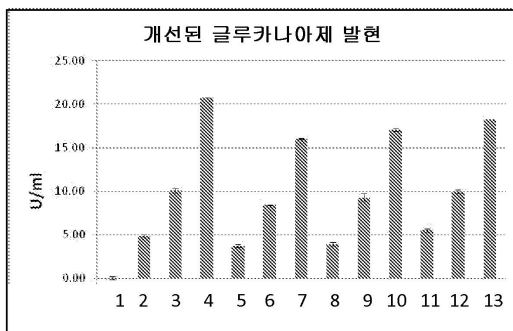
도면1



설명

레인 1, 5, 9 = 음성 대조군
 레인 2, 6, 10 = 서열 번호 3의 서열
 레인 3, 7, 11 = 서열 번호 4의 서열
 레인 4, 8, 12 = 서열 번호 1의 서열
 레인 13 = 분자량 사다리

도면2



설명

레인 1 = 음성 대조군
 레인 2, 5, 8, 11 = 서열 번호 3의 서열
 레인 3, 6, 9, 12 = 서열 번호 4의 서열
 레인 4, 7, 10, 13 = 서열 번호 1의 서열

도면3

서열 번호 1

ATGGGCGTCGATCCGTTTGAACGTAACAAAATCTTGGGCCGCGGCATTAATATCGGC
AATGCGCTCGAAGCACCAATGAAGGCGACTGGGGAGTGGTGATAAAAGATGAGTTC
TTCGACATTATAAAGAAGCCGGTTTCTCTCATGTTTCCGAATCCAATAAGATGGAGT
ACGCACGCTCAGGCGTTTCTCTCTTATAAAAATCGAGCCTTCTTTCTTCAAAAAGAGTG
GATGAAGTGATAAACGAGCCCTGAAAAGAGGACTGGCTGTTGTATAAATATTCAT
CACTACGAGGAGTAAATGAATGATCCAGAAGAACAAGGAAAGATTTCTTGCTCTT
TGGAACAAAATGCTGATCGTTATAAAGACTATCCCGAAACTCTATTTTTTGAAAT
CTGAATGAACCTCACGGAAATCTTACTCCGAAAAATGGAATGAACTGCTTGAGGAA
GCTCTAAAAGTTATAAGATCAATTGACAAAAAGCACACTGTGATTATAGGCACAGCT
GAATGGGGGGGTATATCTGCCCTTGA AAAACTGAGGGTCCCAAAATGGGAAAAAAT
GCGATAGTTACAATCACTACTACAATCCTTTTGAATTTACCCATCAAGGAGCTGAG
TGGGTGCCTGGATCTGAGAAATGGTTGGGAAGAAAGTGGGGATCTCCAGATGATCAG
AAACATTTGATAGAAGAATCAATTTTATAGAAGAATGGTCAAAAAAGAACAAGA
CCAATTTACATAGGTGAGTTTGGTGCCACAGAAAAGCTGACCTGAATCAAGAATA
AAATGGACCTCCTTTGCTGTTTCGGAAGCCGAGAAAAGGGGTGGAGCTGGGCATAC
TGGGAATTTTGTCCGGTTTTGGTGTATGATCCTCTGAGAAAACAGTGAATAAA
GATCTTTTAGAAGCTTAAATAGGAGGAGATAGCATTGAATGA

도면4

서열 번호 2

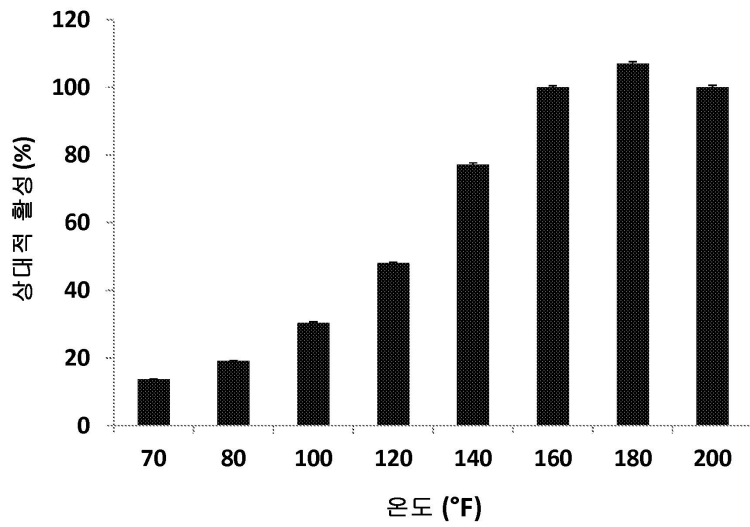
MGVDPFFERNKILGRGINIGNALEAPNEGDWGVVIKDEFDDIIKEAGFSHVRIPIRWS
THAQAFPPYKIEPSFKRVDEVINGALKRGLAVVINIHHYEELMNDPEEHKERFLAL
WKQIADRYKDYPETLFFEILNEPHGNLTPEKWNELLEALKVIRSIDKKHTVIIGTA
EWGGISALEKLRVPKWEKNAIVTIHYNPFETHQGAEWVPGSEKWLGRKWGSPDDQ
KHLIEEFNFIEEWSKKNKRIYIGEFGAYRKADLESRIKWTSFVVREAERKRGWSWAY
WEFCSGFVYDPLRKQWNKDLEALIGGDSIE

도면5

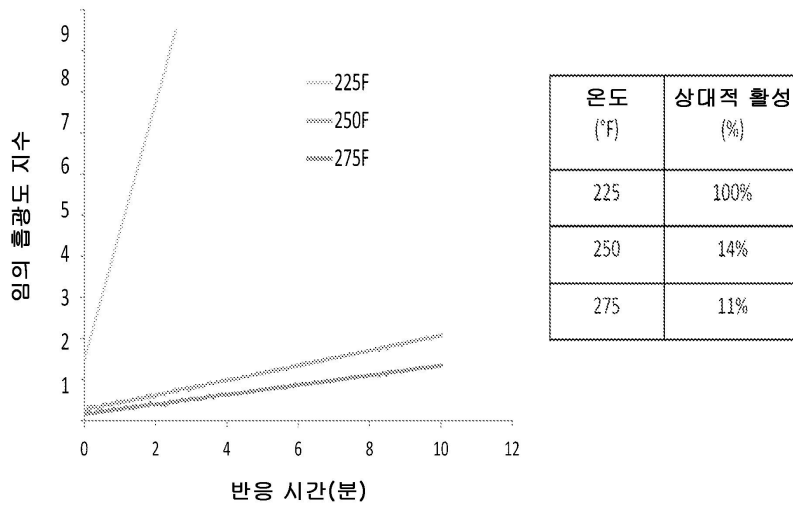
서열 번호 3

ATGGGTGTTGATCCTTTTGAAGGAACAAAATATTGGGAAGAGGCATTAATATAGGAAATGCGC
TTGAAGCACCAATGAGGGAGACTGGGGAGTGGTGATAAAAGATGAGTTCCTCGACATTATAAA
AGAAGCCGGTTTCTCTCATGTTTCCGAATCCAATAAGATGGAGTACGCACGCTCAGGCGTTTCTCT
CCTTATAAAAATCGAGCCTTCTTTCTTCAAAAAGAGTGGATGAAGTGATAAACGGAGCCCTGAAAA
GAGGACTGGCTGTTGTTATAAATATTCATCACTACGAGGAGTAAATGAATGATCCAGAAGAACA
CAAGGAAAGATTTCTTGCTCTTTGGAACAAAATGCTGATCGTTATAAAGACTATCCCGAACT
CTATTTTTTGAATTTGAAATGAACCTCACGGAAATCTTACTCCGAAAAATGGAATGAACTGC
TTGAGGAAGCTCTAAAAGTTATAAGATCAATTGACAAAAAGCACACTGTGATTATAGGCACAGC
TGAATGGGGGGGTATATCTGCCCTTGA AAAACTGAGGGTCCCAAAATGGGAAAAAATGCGATA
GTTACAATTCACTACTACAATCCTTTTGAATTTACCCATCAAGGAGCTGAGTGGGTGCCTGGAT
CTGAGAAATGGTTGGGAAGAAAGTGGGGATCTCCAGATGATCAGAAACATTTGATAGAAGAATT
CAATTTTATAGAAGAATGGTCAAAAAAGAACA AAAAGACCAATTTACATAGGTGAGTTTGGTGCC
TACAGAAAAGCTGACCTTGAATCAAGAATAAAAATGGACCTCCTTTGCTGTTTCGCGAAGCCGAGA
AAAGGGGTGGAGCTGGGCATACTGGGAATTTTGTCCGGTTTTGGTGTATGATCCTCTGAG
AAAACAGTGAATAAAGATCTTTTAGAAGCTTAAATAGGAGGAGATAGCATTGAATAA

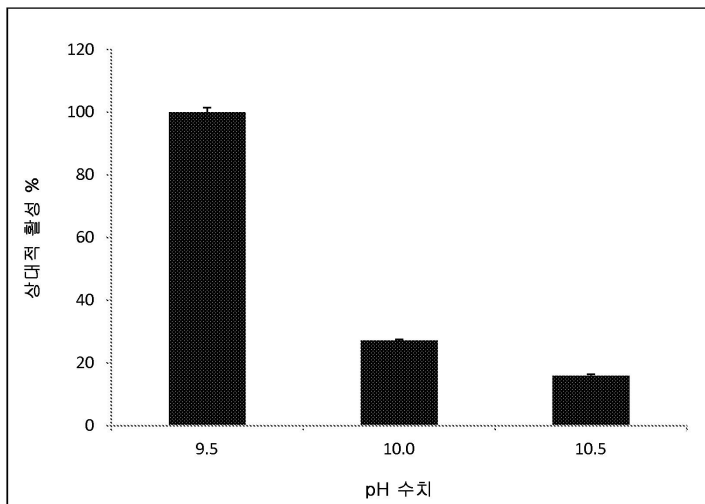
도면8b



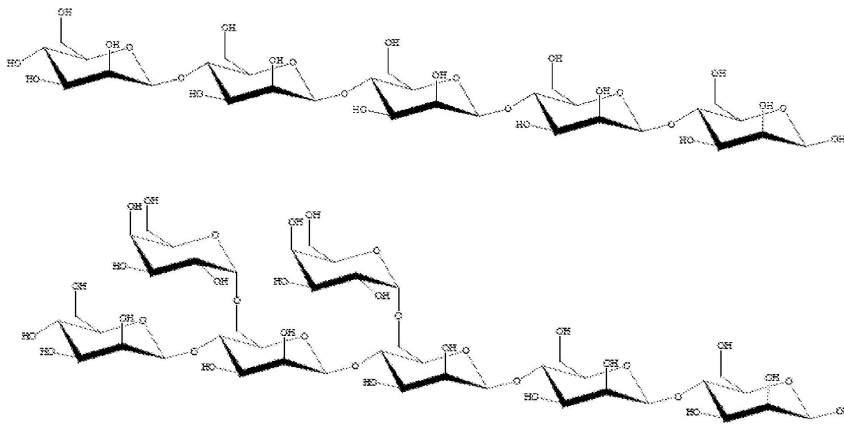
도면8c



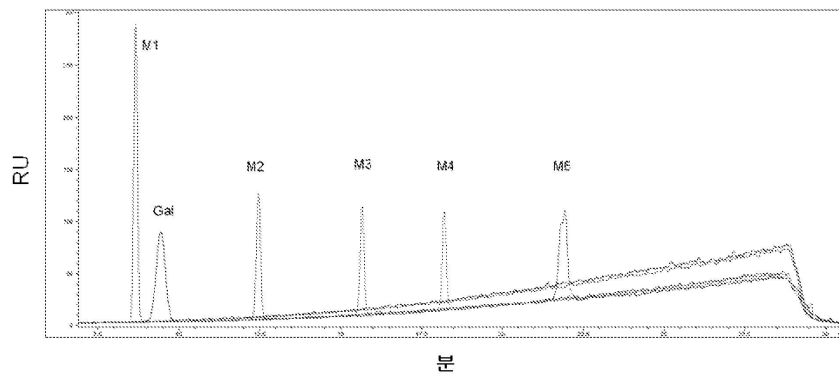
도면9



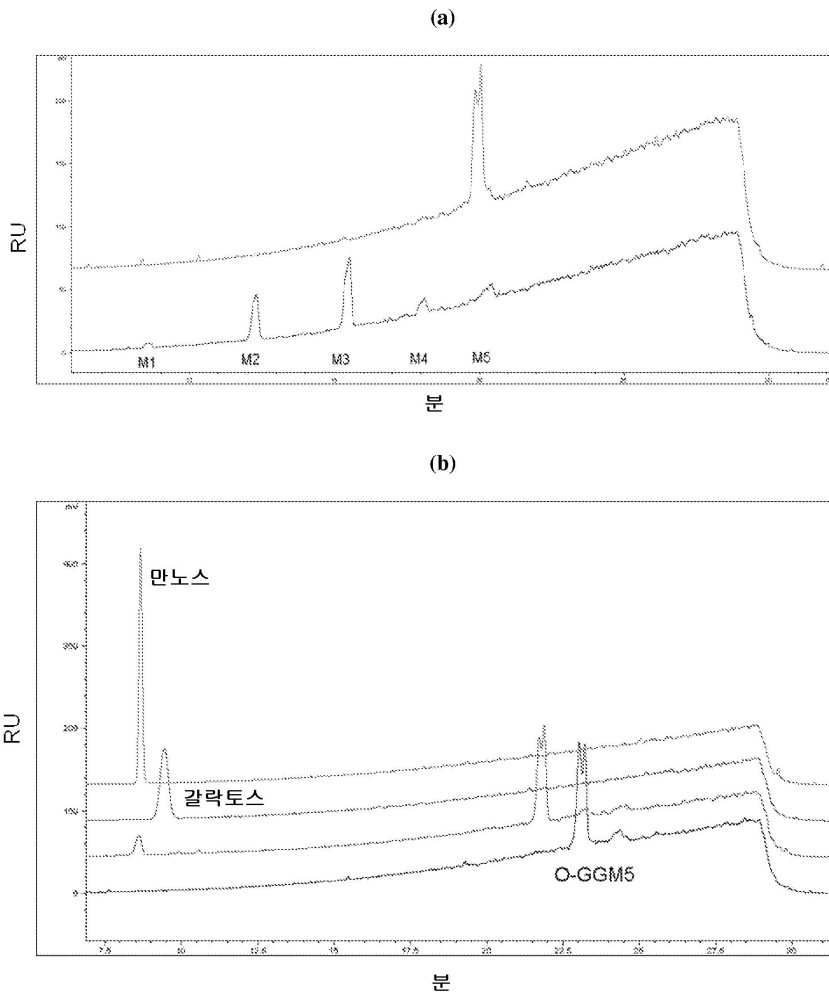
도면10a



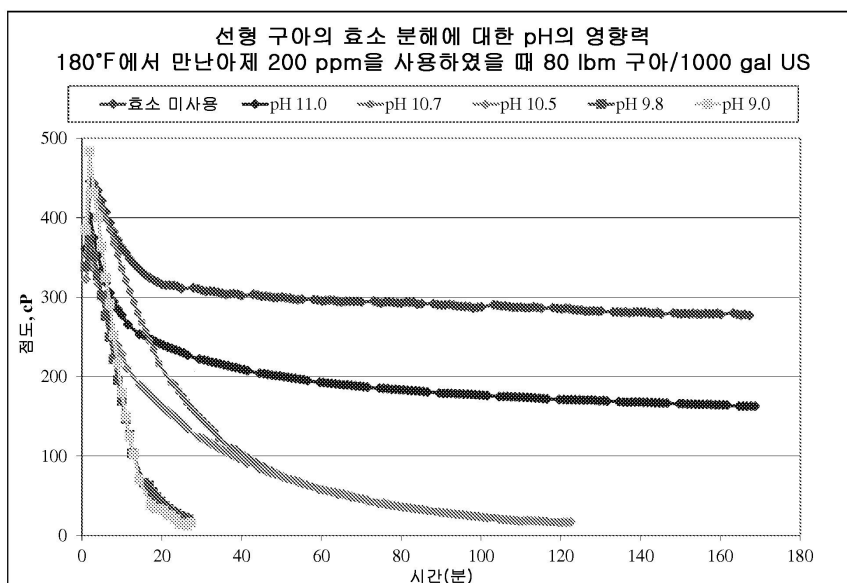
도면10b



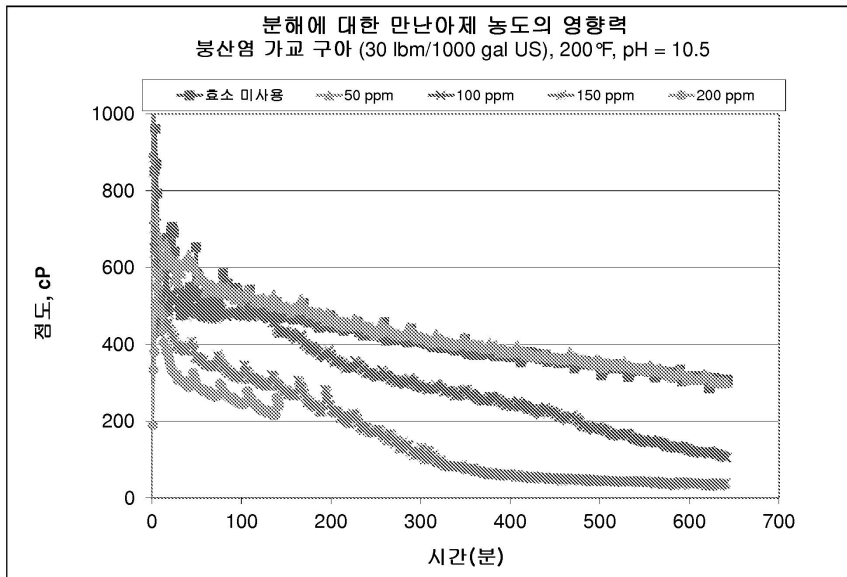
도면10c



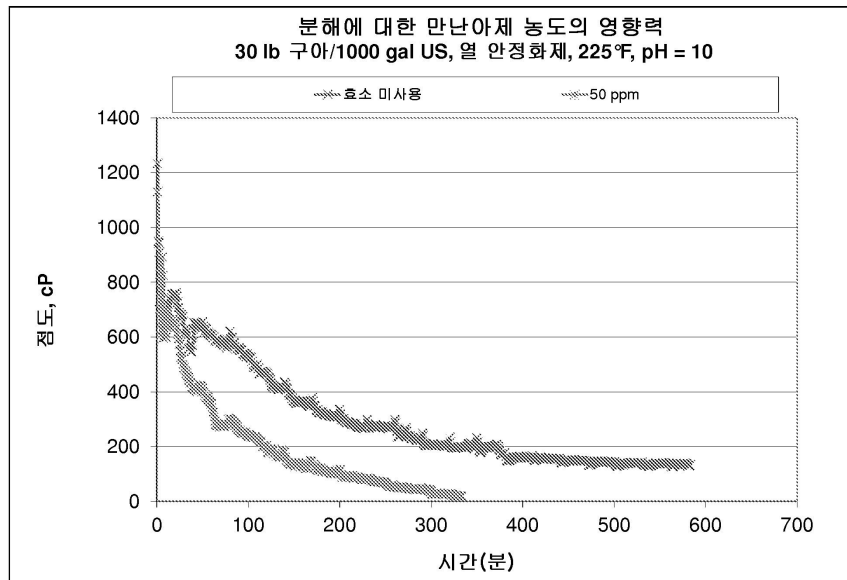
도면11



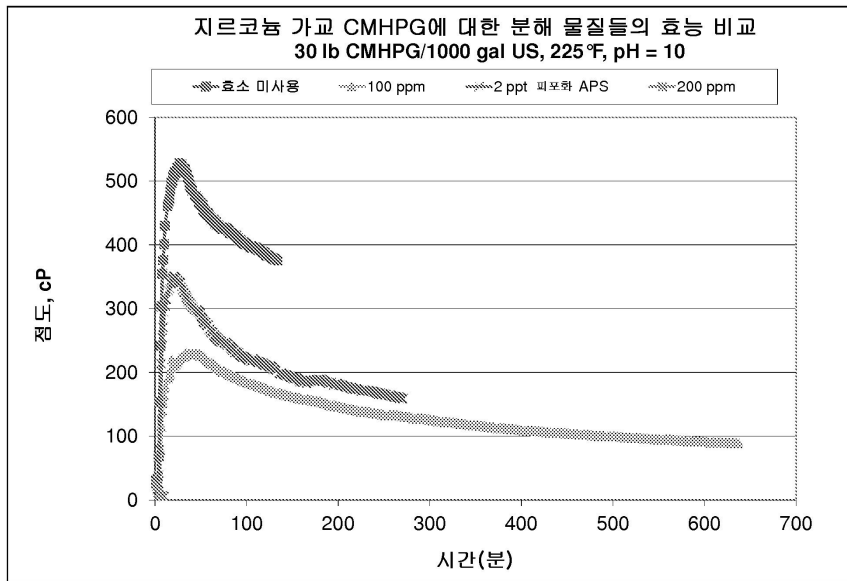
도면12



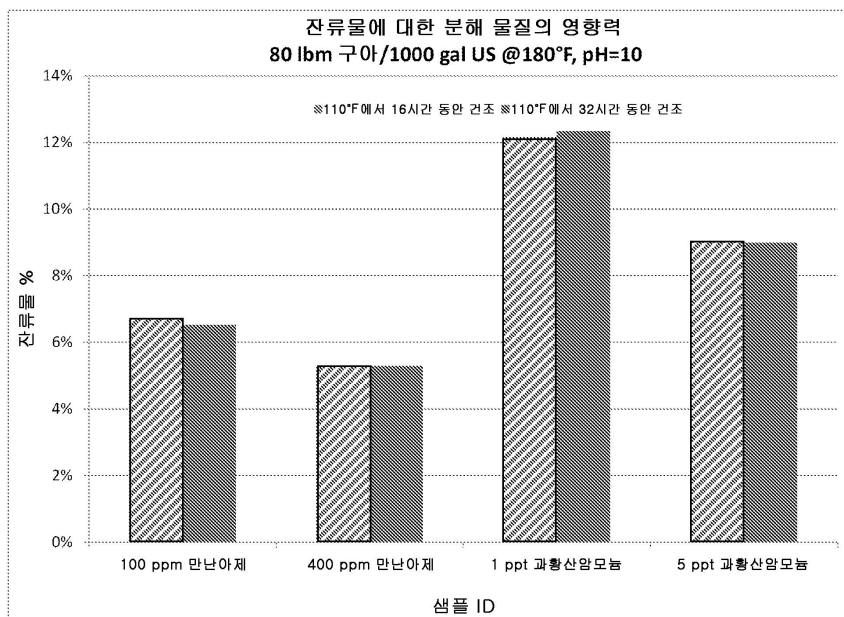
도면13



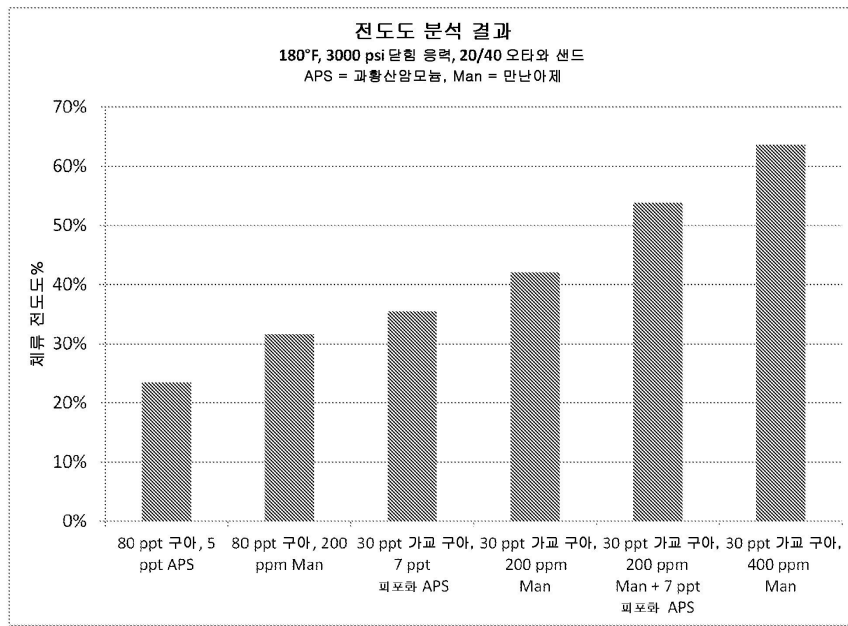
도면14



도면15



도면16



서열 목록

SEQUENCE LISTING

<110> Verenium Corporation

Zhang, Bin

Davenport, Adrienne Houston

Whipple, Lawrence A.

Urbina, Hugo D.

Barrett, Ken E.

Wall, Mark A.

Xuqiu, Tan

<120> GENES ENCODING CELLULASE FOR HYDROLYZING GUAR

FRACTURING FLUIDS UNDER EXTREME WELL CONDITIONS

<130> D2570-1

<150> 61/618610

<151> 2012-03-30

<150> 61/660556

<151> 2012-06-15

<160> 21

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 954

<212> DNA

<213> Artificial Sequence

<220>

<223> Open reading frame encoding SEQ ID NO:2 conveying
enhanced protein expression

<400> 1

atgggcgtcg atccgtttga acgtaacaaa atcttgggcc gcggcattaa tatcgcaat 60
gcgctcgaag caccaaatga aggcgactgg ggagtggatga taaaagatga gttcttcgac 120
attataaaag aagccggttt ctctcatggt cgaattccaa taagatggag tacgcacgct 180
caggcgtttc ctcttataa aatcgagcct tctttcttca aaagagtgga tgaagtgata 240
aacggagccc tgaaaagagg actggctgtt gttataaata ttcataccta cgaggagtta 300

atgaatgac cagaagaaca caaggaaaga tttcttgctc tttggaaaca aattgctgat 360
cgttataaag actatcccga aactctatct tttgaaattc tgaatgaacc tcacggaaat 420
cttactccgg aaaaatggaa tgaactgctt gaggaagctc taaaagtat aagatcaatt 480
gacaaaaagc aactgtgat tataggcaca gctgaatggg ggggtatatac tgccttgaa 540
aaactgaggg tcccaaatg ggaaaaaat gcgatagtta caattcacta ctacaatcct 600
ttcgaattta cccatcaagg agctgagtgg gtgcctggat ctgagaaatg gttgggaaga 660
aagtgggat ctccagatga tcagaaacat ttgatagaag aattcaattt tatagaagaa 720
tggtaaaaa agaacaaaag accaatctac ataggtgagt ttggtgccta cagaaaagct 780

gacctgaaat caagaataaa atggacctcc tttgtcgttc gcgaagccga gaaaaggggg 840
tggagctggg catactggga attttgttcc ggttttgggtg tttatgatcc tctgagaaaa 900
cagtggaata aagatctttt agaagcttta ataggaggag atagcattga atga 954

<210> 2

<211> 317

<212> PRT

<213> *Thermotoga maritima*

<400> 2

Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile

1 5 10 15

Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val

Val Arg Glu Ala Glu Lys Arg Gly Trp Ser Trp Ala Tyr Trp Glu Phe
 275 280 285

Cys Ser Gly Phe Gly Val Tyr Asp Pro Leu Arg Lys Gln Trp Asn Lys
 290 295 300

Asp Leu Leu Glu Ala Leu Ile Gly Gly Asp Ser Ile Glu
 305 310 315

<210> 3

<211> 954

<212> DNA

<213> Artificial Sequence

<220>

<223> Open reading frame encoding SEQ ID NO:2

<400> 3

atgggtgttg atccttttga aaggaacaaa atattgggaa gaggcattaa tataggaaat 60
 gcgcttgaag caccaaatga gggagactgg ggagtgtga taaaagatga gttcttcgac 120

attataaaag aagccggttt ctctcatggt cgaattccaa taagatggag tacgcacgct 180
 caggcgtttc ctcttataa aatcgagcct tctttcttca aaagagtgga tgaagtata 240
 aacggagccc tgaagagagg actggctggt gttataaata ttcactacta cgaggagta 300
 atgaatgatc cagaagaaca caaggaaaga tttcttgctc tttggaaaca aattgctgat 360
 cgttataaag actatcccga aactctatit tttgaaattc tgaatgaacc tcacggaaat 420
 cttactccgg aaaaatggaa tgaactgctt gaggaagctc taaaagtat aagatcaatt 480
 gacaaaaagc aactgtgat tataggcaca gctgaatggg ggggtatatc tgccttgaa 540
 aaactgaggg tcccaaatg ggaaaaaat gcgatagtta caattcacta ctacaatcct 600

ttcgaattta cccatcaagg agctgagtgg gtgcctggat ctgagaaatg gttgggaaga 660
 aagtgggat ctccagatga tcagaacat ttgatagaag aattcaattt tatagaagaa 720
 tggtaaaaa agaacaaaag accaatttac ataggtgagt ttggtgccta cagaaaagct 780
 gacctgaat caagaataaa atggacctcc tttgtcgttc gcgaagccga gaaaaggggg 840
 tggagctggg catactggga atttgttcc ggttttgggtg tttatgatcc tctgagaaaa 900
 cagtgaata aagatctttt agaagcttta ataggaggag atagcattga ataa 954

<210> 4

<211> 976

<212> DNA

<213> Artificial Sequence

<220>

<223> 5' region and open reading frame encoding SEQ ID

NO:2 having modest increased expression

<400> 4

tctactagtt aggaggtaac ttatgggcgt cgatccgttt gaacgtaaca aaatcttggg 60
 ccgcggcatt aatatcggca atgcgctcga agcaccaaat gaaggcgact ggggagtggg 120
 gataaaagat gaggttctcg acattataaa agaagccggt ttctctcatg ttcgaattcc 180
 aataagatgg agtacgcacg ctcaggcggt tcctccttat aaaatcgagc cttctttctt 240
 caaaagagtg gatgaagtga taaacggagc cctgaaaaga ggactggctg ttgttataaa 300
 tattcatcac tacgaggagt taatgaatga tccagaagaa cacaaggaaa gatttcttgc 360
 tctttggaaa caaattgctg atcgttataa agactatccc gaaactctat tttttgaaat 420

tctgaatgaa cctcacggaa atcttactcc ggaaaaatgg aatgaactgc ttgaggaagc 480
 tctaaaaagt ataagatcaa ttgacaaaaa gcacactgtg attataggca cagctgaatg 540
 ggggggtata tctgcccttg aaaaactgag ggtcccaaaa tgggaaaaaa atcgcatagt 600
 tacaattcac tactacaatc ctttcgaatt tacccatcaa ggagctgagt ggggtgcctg 660
 atctgagaaa tggttgggaa gaaagtgggg atctccagat gatcagaaac atttgataga 720
 agaattcaat tttatagaag aatggtcaaa aaagaacaaa agaccaatth acataggtga 780
 gtttgggtcc tacagaaaag ctgaccttga atcaagaata aatggacct cttttgtcgt 840
 tcgcgaagcc gagaaaaggg ggtggagctg ggcatactgg gaatthtgtt ccggttttgg 900

tgtttatgat cctctgagaa aacagtggaa taaagatctt ttagaagctt taataggagg 960
 agatagcatt gaatga 976

<210> 5

<211> 954

<212> DNA

<213> *Thermotoga maritima*

<400> 5

atgggtgttg atccttttga aaggaacaaa atattgggaa gaggcattaa tataggaaat 60
 gcgcttgaag caccaaatga gggagactgg ggagtgggtga taaaagatga gttcttcgac 120
 attataaaag aagccggttt ctctcatggt cgaattccaa taagatggag tacgcacgct 180
 tacgcgtttc ctcttataa aatcatggat cgcttcttca aaagagtgga tgaagtgata 240

aacggagccc tgaagaagg actggctggt gttataaata ttcacacta cgaggagtta 300

atgaatgac cagaagaaca caaggaaaga tttcttgctc tttgaaaca aattgctgat 360

cgttataaag actatcccga aactctatit tttgaaattc tgaatgaacc tcacggaaat 420

cttactccgg aaaaatggaa tgaactgctt gaggaagctc taaaagtat aagatcaatt 480

gacaaaaagc aactataat tataggcaca gctgaatggg ggggtatatac tgccttgaa 540

aaactgtctg tcccaaatg ggaaaaaat tctatagtta caattcacta ctacaatcct 600

ttcgaattta cccatcaagg agctgagtgg gtggaaggat ctgagaaatg gttggaaga 660

aagtgggat ctccagatga tcagaacat ttgatagaag aattcaattt tatagaagaa 720

tggtcaaaaa agaacaaaag accaatctac ataggtgagt ttggtgccta cagaaaagct 780

gaccttgaat caagaataaa atggacctcc tttgtcgttc gcgaaatgga gaaaaggaga 840

tggaactggg catactggga attttgttcc ggttttgggtg tttatgatac tctgagaaaa 900

acctggaata aagatctttt agaagcttta ataggaggag atagcattga ataa 954

<210> 6

<211> 317

<212> PRT

<213> Thermotoga maritima

<220>

<221> DOMAIN

<222> (19)...(296)

<223> Cellulase (glycosyl hydrolase family 5)

<400> 6

Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile

1 5 10 15

Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val

20 25 30

Val Ile Lys Asp Glu Phe Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser

35 40 45

His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Tyr Ala Phe Pro

50 55 60

Pro Tyr Lys Ile Met Asp Arg Phe Phe Lys Arg Val Asp Glu Val Ile

65 70 75 80

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetically generated polynucleotide

<400> 7

atgggtgttg atccttttga aaggaacaaa atattgggaa gaggcattaa tataggaaat 60

gcgcttgaag caccaaatga gggcgactgg ggagtcgtga taaaagatga gttcttcgac 120

attataaaag aagccggttt ctctcatggt cgaattccaa taagatggag tacgcacgct 180

tacgcgtttc ctcttataa aatcatggat cgcttcttca aaagagtgga tgaagtgata 240

aacggagccc tgaagagagg actggctggt gttataaata ttcactacta cgaggagtta 300

atgaatgac cagaagaaca caaggaaaga tttcttgctc tttggaaaca aattgctgat 360

cgttataaag actatcccga aactctatit tttgaaattc tgaatgaacc tcacggaaat 420

cttactccgg aaaaatggaa tgaactgctt gaggaagctc taaaagtat aagatcaatt 480

gacaaaaagc aactataat tataggcaca gctgaatggg ggggtatatac tgcccttgaa 540

aaactgtctg tcccaaatg ggaaaaaat tctatagtta caattcacta ctacaatcct 600

ttcgaattta cccatcaagg agctgagtgg gtggaaggat ctgagaaatg gttggaaga 660

aagtgggat ctccagatga tcagaaacat ttgatagaag aattcaattt tatagaagaa 720

tggtaaaaa agaacaaaag accaatttac ataggtgagt ttggtgccta cagaaaagct 780

gacctgaaat caagaataaa atggacctcc tttgtcgttc gcgaaatgga gaaaaggaga 840

tggagctggg catactggga attttgttcc ggttttgggtg tttatgatac tctgagaaaa 900

acctggaata aagatctttt agaagcttta ataggaggag atagcattga ataa 954

<210> 8

<211> 974

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetically generated polynucleotide

<400> 8

atgggtgttg atccttttga aaggaacaaa atattgggaa gaggcattaa tataggaaat 60

gcgcttgaag caccaaatga gggagactgg ggagtggtga taaaagatga gtatttcgac 120

attataaaag aagccggttt ctctcatggt cgaattccaa taagatggag tacgcacgct 180

caggcgtttc ctcttataa aatcgaggat cgcttcttca aaagagtgga tgaagtgata 240

aacggagccc tgaagagg actggctggt gttataaatc agcatcacta cgaggagtta 300
 atgaatgac cagaagaaca caaggaaaga tttcttgctc tttggaaaca aattgctgat 360

cgttataaag actatcccga aactetattt tttgaaattc tgaatgaacc tcacggaaat 420
 cttactccgg aaaaatggaa tgaactgctt gaggaagctc taaaagtat aagatcaatt 480
 gacaaaaagc aactataat tataggcaca gctgaatggg ggggtatatac tgccttgaa 540
 aaactgaggg tcccaaatg ggaaaaaat gcgatagtta caattcacta ctacaatcct 600
 ttcgaattta cccatcaagg agctgagtgg gtggaaggat ctgagaaatg gttgggaaga 660
 aagtggggat ctccagatga tcagaacat ttgatagaag aattcaattt tatagaagaa 720
 tgggtcaaaa agaacaaaag accaatctac ataggtgagt ttggtgccta cagaaaagct 780
 gaccttgaat caagaataaa atggacctcc tttgtcgttc gcgaagctga gaaaaggaga 840

tggagctggg catactggga attttgttcc ggttttgggtg tttatgatac tctgagaaaa 900
 acctggaata aagatctttt agaagcttta ataggaggag atagcattga ataacacat 960
 tccaagatgg cgtg 974

<210> 9

<211> 320

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetically generated polypeptide

<400> 9

Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile

1 5 10 15

Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val

20 25 30

Val Ile Lys Asp Glu Tyr Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser

35 40 45

His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Gln Ala Phe Pro

50 55 60

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

65 70 75 80

Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Val Ile Asn Gln His His

<213> Artificial Sequence

<220>

<223> Synthetically generated polynucleotide

<400> 10

atgggtgttg atccttttga aaggaacaaa atattgggaa gaggcattaa tataggaaat 60

gcgcttgaag caccaaatga gggagactgg ggagtgtga taaaagatga gtatttcgac 120

attataaaag aagccggttt ctctcatgtt cgaattccaa taagatggag tacgcacgct 180

caggcgtttc ctcttataa aatcgaggat tctttcttca aaagagtgga tgaagtgata 240

aacggagccc tgaagagg actggctggt gttataaata ttcataccta cgaggagta 300

atgaatgatc cagaagaaca caaggaaaga tttcttgctc tttgaaaca aattgctgat 360

cgttataaag actatcccga aactctatit tttgaaattc tgaatgaacc tcacggaaat 420

cttactccgg aaaaatggaa tgaactgctt gaggaagctc taaaagttat aagatcaatt 480

gacaaaaagc aactgtgat tataggcaca gctgaatggg ggggtatatac tgccttgaa 540

aaactgaggg tcccaaatg ggaaaaaat gcgatagtta caattcacta ctacaatcct 600

ttcgaattta cccatcaagg agctgagtgg gtgcctggat ctgagaaatg gttgggaaga 660

aagtgggat ctccagatga tcagaacat gtgatagaag aattcaattt tatagaagaa 720

tggtaaaaa agaacaaaag accaatctac ataggtgagt ttggtgccta cagaaaagct 780

gacctgaat caagaataaa atggacctcc tttgtcgttc gcgaagccga gaaaaggggg 840

tggagctggg catactggga atttgttcc ggttttgggtg tttatgatcc tctgagaaaa 900

cagtgaata aagatctttt agaagctcta ataggaggag atagcattga ataa 954

<210> 11

<211> 317

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetically generated polypeptide

<400> 11

Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile

1 5 10 15

Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val

20 25 30

Val Ile Lys Asp Glu Tyr Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser

Cys Ser Gly Phe Gly Val Tyr Asp Pro Leu Arg Lys Gln Trp Asn Lys

290

295

300

Asp Leu Leu Glu Ala Leu Ile Gly Gly Asp Ser Ile Glu

305

310

315

<210> 12

<211> 954

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetically generated polynucleotide

<400> 12

atgggtgttg atccttttga aaggaacaaa atattgggaa gaggcattaa tataggaaat 60
 gcgcttgaag caccaaatga gggagactgg ggagtgtga taaaagatga gttcttcgac 120
 attataaaag aagccggttt ctctcatggt cgaattccaa taagatggag tacgcacgct 180
 caggcgtttc ctcttataa aatcgaggat tctttcttca aaagagtgga tgaagtgata 240

aacggagccc tgaagagagg actggctgtt gttataaatc agcatcacta cgaggagtta 300
 atgaatgac cagaagaaca caaggaaaga tttcttgctc tttggaaaca aattgctgat 360
 cgttataaag actatcccga aactctatit tttgaaattc tgaatgaacc tcacggaaat 420
 ctactccgg aaaaatggaa tgaactgctt gaggaagctc taaaagtat aagatcaatt 480
 gacaaaaagc aactgtgat tataggcaca gctgaatggg ggggtatatac tgcccttgaa 540
 aaactgaggg tcccaaatg ggaaaaaat gcgatagtta caattcacta ctacaatcct 600
 ttcgaattta cccatcaagg agctgagtgg gtgcctggat ctgagaaatg gttgggaaga 660
 aagtgggat ctccagatga tcagaacat ttgatagaag aattcaattt tatagaagaa 720

tggtcaaaaa agaacaaaag accaatttac ataggtgagt ttggtgccta cagaaaagct 780
 gaccttgaat caagaataaa atggacctcc tttgtcgttc gcgaagccga gaaaaggggg 840
 tggagctggg cactactggga attttgttcc ggttttgggtg tttatgatcc tctgagaaaa 900
 cagtgggaata aagatctttt agaagcttta ataggaggag atagcattga ataa 954

<210> 13

<211> 317

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetically generated polypeptide

<400> 13

Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile
 1 5 10 15

Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val
 20 25 30

Val Ile Lys Asp Glu Phe Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser
 35 40 45

His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Gln Ala Phe Pro
 50 55 60

Pro Tyr Lys Ile Glu Asp Ser Phe Phe Lys Arg Val Asp Glu Val Ile
 65 70 75 80

Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Val Ile Asn Gln His His
 85 90 95

Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg Phe Leu
 100 105 110

Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr
 115 120 125

Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr Pro Glu
 130 135 140

Lys Trp Asn Glu Leu Leu Glu Glu Ala Leu Lys Val Ile Arg Ser Ile
 145 150 155 160

Asp Lys Lys His Thr Val Ile Ile Gly Thr Ala Glu Trp Gly Gly Ile
 165 170 175

Ser Ala Leu Glu Lys Leu Arg Val Pro Lys Trp Glu Lys Asn Ala Ile
 180 185 190

Val Thr Ile His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln Gly Ala
 195 200 205

Glu Trp Val Pro Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp Gly Ser
 210 215 220

Pro Asp Asp Gln Lys His Leu Ile Glu Glu Phe Asn Phe Ile Glu Glu

ggcaaagagt aa

972

<210> 15

<211> 323

<212> PRT

<213> Thermotoga sp.

<400> 15

Met Glu Gln Ser Val Ala Glu Ser Asp Ser Asn Ser Ala Phe Glu Tyr

1 5 10 15

Asn Lys Met Val Gly Lys Gly Val Asn Ile Gly Asn Ala Leu Glu Ala

20 25 30

Pro Phe Glu Gly Ala Trp Gly Val Arg Ile Glu Asp Glu Tyr Phe Glu

35 40 45

Ile Ile Lys Lys Arg Gly Phe Asp Ser Val Arg Ile Pro Ile Arg Trp

50 55 60

Ser Ala His Ile Ser Glu Lys Pro Pro Tyr Asp Ile Asp Arg Asn Phe

65 70 75 80

Leu Glu Arg Val Asn His Val Val Asp Arg Ala Leu Glu Asn Asn Leu

85 90 95

Thr Val Ile Ile Asn Thr His His Phe Glu Glu Leu Tyr Gln Glu Pro

100 105 110

Asp Lys Tyr Gly Asp Val Leu Val Glu Ile Trp Arg Gln Ile Ala Lys

115 120 125

Phe Phe Lys Asp Tyr Pro Glu Asn Leu Phe Phe Glu Ile Tyr Asn Glu

130 135 140

Pro Ala Gln Asn Leu Thr Ala Glu Lys Trp Asn Ala Leu Tyr Pro Lys

145 150 155 160

Val Leu Lys Val Ile Arg Glu Ser Asn Pro Thr Arg Ile Val Ile Ile

165 170 175

Asp Ala Pro Asn Trp Ala His Tyr Ser Ala Val Arg Ser Leu Lys Leu

180 185 190

Val Asn Asp Lys Arg Ile Ile Val Ser Phe His Tyr Tyr Glu Pro Phe

aacctgaccc cggagaagtg gaacgagctg ctggaggagg ccctgaaggt gatcaggagc 540
 atcgacaaga agcacaccat catcatcggc accgccgagt ggggcggcat cagcgcctg 600
 gagaagtga gggtgccgaa gtgggagaag aacgccatcg tgaccatcca ctactacaac 660
 ccgttcgagt tcaccacca gggcgccgag tgggtggagg gcagcgagaa gtggctgggc 720

 aggaagtggg gcagcccga cgaccagaag cacctgatcg aggagttaa cttcatcgag 780
 gagtggagca agaagaaca gaggccgatc tacatcggcg agttcggcgc ctacaggaag 840
 gccgacctgg agagcaggat caagtgacc agcttcgtgg tgaggaggc cgagaagagg 900
 aggtggagct gggcctactg ggagttctgc agcggcttcg gcgtgtacga caccctgagg 960
 aagacctgga acaaggacct gctggaggcc ctgatcggcg gcgacagcat cgagagcgag 1020
 aaggacgagc tgtgagagct ca 1042

<210> 17

<211> 341

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetically generated polypeptide

<400> 17

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

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

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetically generated polynucleotide

<400> 18

ggatccacca tgagggtgtt gctcgttgcc ctgctctcc tggetctcgc tgcgagcgcc 60
 accagcggcg tggacccgtt cgagaggaac aagatcctgg gcaggggcat caacatcggc 120
 aacgccttgg aggccccgaa cgagggcgac tggggcgtgg tgatcaagga cgagttcttc 180

gacatcatca aggaggccgg cttcagccac gtgagaatcc cgatcaggtg gagcacccac 240
 gccacggcct tcccgcgta caagatcgag ccgagcttct tcaagagggt ggacgaggtg 300
 atcaacggcg ccctgaagag gggcctggcc gtggtgatca acatccacca ctacaggag 360
 ctgatgaacg acccggagga gcacaaggag aggttcctgg ccctgtgaa gcagatcgcc 420
 gacaggtaca aggactacc ggagaccctg ttcttcgaga tcctgaacga gccgcacggc 480
 aacctgacc cggagaagtg gaacgagctg ctggaggagg ccctgaaggt gatcaggagc 540
 atcgacaaga agcacaccgt gatcatcggc accgccgagt ggggcggcat cagcgcctg 600
 gagaagtga gggtgccgaa gtgggagaag aacccatcg tgaccatcca ctactacaac 660

ccgttcgagt teaccacca gggcggcgg tgggtgccgg gcagcgagaa gtggtgggc 720
 aggaagtggg gcagcccgga cgaccagaag cacctgatcg aggagttaa cttcatcgag 780
 gagtggagca agaagaaca gaggccgatc tacatcggcg agttcggcgc ctacaggaag 840
 gccgacctgg agagcaggat caagtggacc agcttcgtgg tgaggaggc cgagaagagg 900
 ggctggagct gggcctactg ggagttctgc agcggcttcg gcgtgtacga cccgctgagg 960
 aagcagtgga acaaggacct gctggaggcc ctgatcggcg gcgacagcat cgagagcgag 1020
 aaggacgagc tgtgagagct ca 1042

<210> 19

<211> 341

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetically generated polypeptide

<400> 19

Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser

1 5 10 15

Ala Thr Ser Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg
 20 25 30

Gly Ile Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp
 35 40 45

Gly Val Val Ile Lys Asp Glu Phe Phe Asp Ile Ile Lys Glu Ala Gly
 50 55 60

Phe Ser His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Gln Ala
 65 70 75 80

Phe Pro Pro Tyr Lys Ile Glu Pro Ser Phe Phe Lys Arg Val Asp Glu
 85 90 95

Val Ile Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Val Ile Asn Ile
 100 105 110

His His Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg
 115 120 125

Phe Leu Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro
 130 135 140

Glu Thr Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr
 145 150 155 160

Pro Glu Lys Trp Asn Glu Leu Leu Glu Glu Ala Leu Lys Val Ile Arg
 165 170 175

Ser Ile Asp Lys Lys His Thr Val Ile Ile Gly Thr Ala Glu Trp Gly
 180 185 190

Gly Ile Ser Ala Leu Glu Lys Leu Arg Val Pro Lys Trp Glu Lys Asn
 195 200 205

Ala Ile Val Thr Ile His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln
 210 215 220

Gly Ala Glu Trp Val Pro Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp
 225 230 235 240

Gly Ser Pro Asp Asp Gln Lys His Leu Ile Glu Glu Phe Asn Phe Ile
 245 250 255

Glu Glu Trp Ser Lys Lys Asn Lys Arg Pro Ile Tyr Ile Gly Glu Phe

260 265 270
 Gly Ala Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Trp Thr Ser

275 280 285
 Phe Val Val Arg Glu Ala Glu Lys Arg Gly Trp Ser Trp Ala Tyr Trp
 290 295 300

Glu Phe Cys Ser Gly Phe Gly Val Tyr Asp Pro Leu Arg Lys Gln Trp
 305 310 315 320
 Asn Lys Asp Leu Leu Glu Ala Leu Ile Gly Gly Asp Ser Ile Glu Ser
 325 330 335
 Glu Lys Asp Glu Leu
 340

<210> 20

<211> 1042

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetically generated polynucleotide

<400> 20

ggatccacca tgagggtgtt gctcgttgcc ctgctctcc tggctctcgc tgcgagcgcc 60
 accagcggcg tggaccctgt cgagaggaac aagatcctgg gcaggggcat caacatcggc 120

aacgccttg aggccccgaa cgagggcgac tggggcgtgg tgatcaagga cgagtacttc 180
 gacatcatca aggaggccgg cttcagccac gtgagaatcc cgatcaggtg gagcaccac 240
 gccagccct tcccgccgta caagatcgag gacagcttct tcaagagggt ggacgaggtg 300
 atcaacggcg ccctgaagag gggcctggcc gtggtgatca acatccacca ctacaggag 360
 ctgatgaacg acccggagga gcacaaggag aggttctctg ccctgtggaa gcagatcgcc 420
 gacaggtaca aggactacc ggagaccctg ttcttcgaga tcctgaacga gccgcacggc 480
 aacctgacc cggagaagtg gaacgagctg ctggaggagg ccctgaaggt gatcaggagc 540
 atcgacaaga agcacaccgt gatcatcggc accgccgagt ggggcggcat cagcgcctg 600

gagaagctga ggggtgccgaa gtgggagaag aacccatcg tgaccatcca ctactacaac 660
 ccgttcgagt taccacca gggcgccgag tgggtgccgg gcagcgagaa gtggctgggc 720
 aggaagtggg gcagcccga cgaccagaag cacgtgatcg aggagttcaa cttcatcgag 780
 gagtggagca agaagaaca gaggccgatc tacatcggcg agttcggcgc ctacaggaag 840

gccgacctgg agagcaggat caagtgacc agcttcgtgg tgagggaggc cgagaagagg 900
 ggctggagct ggcctactg ggagtctgc agcggcttcg gcgtgtacga cccgctgagg 960
 aagcagtga acaaggacct gctggaggcc ctgatcggcg gcgacagcat cgagagcgag 1020
 aaggacgagc tgtgagagct ca 1042

<210> 21

<211> 341

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetically generated polypeptide

<400> 21

Met Arg Val Leu Leu Val Ala Leu Ala Leu Leu Ala Leu Ala Ala Ser

1 5 10 15

Ala Thr Ser Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg

20 25 30

Gly Ile Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp

35 40 45

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

50 55 60

Phe Ser His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Gln Ala

65 70 75 80

Phe Pro Pro Tyr Lys Ile Glu Asp Ser Phe Phe Lys Arg Val Asp Glu

85 90 95

Val Ile Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Val Ile Asn Ile

100 105 110

His His Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg

115 120 125

Phe Leu Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro

130 135 140

Glu Thr Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr

145 150 155 160

Pro Glu Lys Trp Asn Glu Leu Leu Glu Glu Ala Leu Lys Val Ile Arg

