(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0114003
(43) 공개일자 2019년10월08일

(51) 국제특허분류(Int. Cl.)
A61K 38/47 (2006.01) *A61K 47/64* (2017.01)
A61K 9/00 (2006.01) *A61P 21/00* (2006.01)
A61P 43/00 (2006.01) *C12N 1/14* (2018.01)
C12N 1/16 (2006.01) *C12N 9/26* (2006.01)
C12N 9/44 (2006.01) *C12N 9/62* (2006.01)
(52) CPC특허분류
A61K 38/47 (2013.01)
A61K 47/64 (2017.08)
(21) 출원번호 10-2019-7028133(분할)
(22) 출원일자(국제) 2013년03월14일
심사청구일자 2019년09월27일
(62) 원출원 특허 10-2014-7027709
원출원일자(국제) 2013년03월14일
심사청구일자 2018년03월13일
(85) 번역문제출일자 2019년09월25일
(86) 국제출원번호 PCT/IB2013/000912
(87) 국제공개번호 WO 2013/136189
국제공개일자 2013년09월19일
(30) 우선권주장
61/611,485 2012년03월15일 미국(US)

(71) 출원인
옥시레인 유케이 리미티드
영국 엠1 2에이피 맨체스터 피카딜리 49 피카딜리 하우스 수트 313
(72) 발명자
베르베켄, 와우터
벨기에 9860 랜드스카우터 스투이벤버그 35
파인스, 캐슬린 카밀라, 텔레스포레 앨리다 마리아
벨기에 비-9000 겐트 시타델란 54
(뒷면에 계속)
(74) 대리인
특허법인 무한

전체 청구항 수 : 총 71 항

(54) 발명의 명칭 **폼페병의 치료를 위한 방법 및 물질**

(57) 요 약

본 문서는, 포유동물 세포의 내부로 수송되도록, 분자 복합체의 증진된 능력을 결과적으로 나타내는, 산 알파 글루코시다아제 활성도 및 적어도 하나의 변형을 갖는 분자 복합체에 관한 것이다.

(52) CPC특허분류
A61K 9/0019 (2013.01)
A61P 21/00 (2018.01)
A61P 43/00 (2018.01)
C12N 1/14 (2013.01)
C12N 9/2408 (2013.01)
C12N 9/2451 (2013.01)
C12N 9/62 (2013.01)
C12Y 302/01003 (2013.01)
C12Y 302/0102 (2013.01)
(72) 발명자
스타우트, 잔 로버트, 루도
벨기에 비-9890 아스페르-가베레 클레인벨드 19
피나에르트, 그웬다, 노엘라
벨기에 비-9880 알테르 리에툴스트라트 7

명 세 서

청구범위

청구항 1

산 알파 글루코시다아제(acid alpha glucosidase, GAA) 활성도를 갖고, 적어도 두 개의 폴리펩티드를 포함하는 분리된 분자 복합체로서,

각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 아미노산 서열의 세그먼트(segment)와 적어도 85 % 서열 동일성을 갖고, 각각의 세그먼트는 아미노산 50 및 아미노산 74 사이의 하나 또는 그 이상의 부위(sites)에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해(proteolysis)에 의해 유도되고;

상기 분자 복합체는 포유동물의 세포의 내부로 수송되도록 상기 분자 복합체의 증진된 능력을 결과적으로 나타내는 적어도 하나의 변형을 포함하는 것인, 분리된 분자 복합체.

청구항 2

제1항에 있어서,

각각의 세그먼트는, 아미노산 60과 아미노산 65 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도된 것인, 분자 복합체.

청구항 3

제1항에 있어서,

상기 복합체는 적어도 세 개의 폴리펩티드를 포함하는 것인, 분자 복합체.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

SEQ ID NO: 1에 기재된 상기 아미노산 서열의 상기 단백질분해는, SEQ ID NO: 1에 기재된 상기 아미노산 서열의 아미노산 719과 아미노산 746 사이의 하나 또는 그 이상의 부위에서 분열(cleavage) 또는 아미노산 137과 아미노산 151 사이의 하나 또는 그 이상의 부위에서의 분열을 추가적으로 포함하는 것인, 분자 복합체.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서,

상기 복합체는 적어도 4 개의 폴리펩티드를 포함하는 것인, 분자 복합체.

청구항 6

제5항에 있어서,

상기 단백질분해는, SEQ ID NO: 1에 기재된 상기 아미노산 서열의 아미노산 719와 아미노산 746 사이의 하나 또는 그 이상의 부위에서의 분열 및 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 아미노산 137과 아미노산 151 사이의 하나 또는 그 이상의 부위에서의 분열을 추가적으로 포함하는 것인, 분자 복합체.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서,

상기 폴리펩티드의 적어도 하나는 하나 또는 그 이상의 인산화된 N-글리칸을 포함하고,

상기 변형은 적어도 하나의 인산화된 N-글리칸의 캡핑제거(uncapping) 및 디만노실레이션(demannosylation)을 포함하는 것인, 분자 복합체.

청구항 8

제5항에 있어서,

상기 폴리펩티드에서 상기 N-글리칸의 적어도 40 %는, 캡핑제거되고(uncapped), 디만노실레이트된(demannosylated) 것인, 분자 복합체.

청구항 9

제8항에 있어서,

상기 폴리펩티드에서 상기 N-글리칸의 적어도 60 %는, 캡핑제거되고, 디만노실레이트된 것인, 분자 복합체.

청구항 10

제8항에 있어서,

상기 폴리펩티드에서 상기 N-글리칸의 적어도 80 %는, 캡핑제거되고 디만노실레이트된 것인, 분자 복합체.

청구항 11

제8항에 있어서,

상기 폴리펩티드에서 상기 N-글리칸의 적어도 90 %는 캡핑제거되고 디만노실레이트된 것인, 분자 복합체.

청구항 12

제1항 내지 제11항 중 어느 한 항에 있어서,

각각의 폴리펩티드는, SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 90 % 서열 동일성을 갖는 것인, 분자 복합체.

청구항 13

제1항 내지 제11항 중 어느 한 항에 있어서,

각각의 폴리펩티드는, SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 95 %의 서열 동일성을 갖는 것인, 분자 복합체.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서,

상기 적어도 두 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 22 내지 57을 포함하고(said segment comprises amino acids 22 to 57 of SEQ ID NO:1),

상기 적어도 두 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 66 내지 896을 포함하는 것인, 분자 복합체.

청구항 15

제3항 내지 제14항 중 어느 한 항에 있어서,

상기 적어도 세 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 22 내지 57을 포함하고,

상기 적어도 세 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 66 내지 726을 포함하고,

상기 적어도 세 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 736 내지 896을 포함하는 것인, 분자 복합체.

청구항 16

제5항 내지 제14항 중 어느 한 항에 있어서,

상기 적어도 4 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 22 내지 57을 포함하고,

상기 적어도 4 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 66 내지 143을 포함하고,

상기 적어도 4 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 158 내지 726을 포함하고,

상기 적어도 4 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO: 1의 아미노산 736 내지 896을 포함하는 것인, 분자 복합체.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서,

상기 적어도 하나의 변형은, 적어도 하나의 상기 폴리펩티드에 융합된 세포외 수용체에 대한 리간드를 포함하는 것인, 분자 복합체.

청구항 18

제1항 내지 제16항 중 어느 한 항에 있어서,

상기 적어도 하나의 변형은, 적어도 하나의 상기 폴리펩티드에 융합된 타겟팅 도메인(targeting domain)을 포함하고,

상기 타겟팅 도메인은, 타겟 세포의 표면 상에 수용체의 세포외 도메인을 결합하는 것인, 분자 복합체.

청구항 19

제1항 내지 제16항 중 어느 한 항에 있어서,

상기 적어도 하나의 변형은, 상기 폴리펩티드의 적어도 하나에 융합된 우로키나아제-타입 플라스미노겐 수용체(urokinase-type plasminogen receptor)를 포함하는 것인, 분자 복합체.

청구항 20

제1항 내지 제16항 중 어느 한 항에 있어서,

상기 적어도 하나의 변형은, 적어도 하나의 상기 폴리펩티드에 융합된 인간 인슐린-유사 성장 인자 Ⅱ의 상기 인식 도메인(recognition domain)을 포함하는 것인, 분자 복합체.

청구항 21

제1항 내지 제20항 중 어느 한 항의 분자 복합체를 포함하는 조성물로서, 상기 분자 복합체는 동결 건조된 것인, 조성물.

청구항 22

제21항에 있어서,

상기 조성물은 단일 사용 바이알(single use vial)로서 포장되는 것인, 조성물.

청구항 23

제1항 내지 제20항 중 어느 한 항의 분자 복합체 및 약제학적으로 허용가능한 담체를 포함하는 약제학적 조성물.

청구항 24

제23항에 있어서,

상기 조성물은 정맥내 또는 피하의 투여를 위해 제형화된 것인, 조성물.

청구항 25

제23항에 있어서,

상기 조성물은 정맥내 투입을 위해 제형화된 것인, 조성물.

청구항 26

제23항 내지 제25항 중 어느 한 항의 조성물을 폼페병으로 진단된 환자에게 투여하는 것을 포함하는, 폼페병을 치료하는 방법.

청구항 27

제26항에 있어서,

상기 환자는 유아기 발병 폼페병(infantile onset Pompe's disease)으로 진단된 것인, 방법.

청구항 28

제26항에 있어서,

상기 환자는 후발성 폼페병으로 진단된 것인, 방법.

청구항 29

SEQ ID NO : 8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 프로테아제와 SEQ ID NO: 1에 기재된 상기 아미노산 서열을 갖는 폴리펩티드를 접촉하는 것을 포함하는, 분자 복합체를 제조하기 위한 방법으로서,

상기 프로테아제는 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위에서 상기 폴리펩티드를 분해하는 것(cleave)인, 방법.

청구항 30

제29항에 있어서,

상기 접촉시키는 것은 *생체 외에서* 실행되는 것인, 방법.

청구항 31

제29항 또는 제30항에 있어서,

상기 프로테아제는, 아미노산 60과 아미노산 65 사이의 하나 또는 그 이상의 부위에서의 상기 폴리펩티드를 분해하는 것인, 방법.

청구항 32

캡핑제거되고 디만노실레이트된 인산화된 N-글리칸(uncapped and demannosylated phosphorylated N-glycans)을 포함하는 분자 복합체를 제조하기 위한 방법으로서,

상기 방법은, (ⅰ) 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해하고, (ii) 말단 알파-1,2 만노스(terminal alpha-1,2 mannose), 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합(alpha-1,6 mannose linkages)을 가수분해할 수 있는, 만노시다아제와 분자 복합체를 접촉시키는 것을 포함하고, 상기 분자 복합체는 GAA 활성도를 가지고, 적어도 두 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는 SEQ ID NO: 1 에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 갖고, 각각의 세그먼트는 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도되고,

상기 폴리펩티드의 적어도 하나는 하나 또는 그 이상의 만노스-1-포스포-6-만노스 모이어티를 함유하는 인산화된 N-글리칸을 포함하는 것인, 방법.

청구항 33

제32항에 있어서,

상기 만노시다아제는 패밀리 38 글리코실 히드롤라제(family 38 glycosyl hydrolase)인 것인, 방법.

청구항 34

제32항 또는 제33항에 있어서,

상기 만노시다아제는 *카나발리아 엔스폴미스* 만노시다아제(*Canavalia ensiformis* mannosidase)인 것인, 방법.

청구항 35

제32항 또는 제33항에 있어서,

상기 만노시다아제는 *야로위야 리폴리카* 만노시다아제(*Yarrowia lipolytica* mannosidase)인 것인, 방법.

청구항 36

제32항 내지 제35항 중 어느 한 항에 있어서,

상기 접촉시키는 것은 재조합 균류 세포에서 발생하고, 상기 균류 세포는 상기 만노시다아제를 발현시키는 것인, 방법.

청구항 37

캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 포함하는 분자 복합체를 제조하는 방법으로서,

상기 방법은, 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있는 만노시다아제와 분자 복합체를 접촉하는 것을 포함하고, 상기 분자 복합체는 GAA 활성도를 가지고, 적어도 2 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 가지고, 각각의 세그먼트는 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도된 것이고,

상기 접촉시키기 전에, 상기 폴리펩티드의 적어도 하나는, 캡핑제거된 만노스-6-포스페이트 모이어티를 포함하는 인산화된 N-글리칸을 포함하는 것인, 방법.

청구항 38

제37항에 있어서,

상기 만노시다아제는 패밀리 47 글리코실 히드롤라아제(family 47 glycosyl hydrolase)인 것인, 방법.

청구항 39

제37항 또는 제38항에 있어서,

상기 만노시다아제는 *아스페르길루스 사토이* 만노시다아제(*Aspergillus satoi* mannosidase)인 것인, 방법.

청구항 40

제37항에 있어서,

상기 만노시다아제는 패밀리 92 글리코실 히드롤라제인 것인, 방법.

청구항 41

제37항 또는 제40항에 있어서,

만노시다아제는 *셀룰로시마이크로비움 셀룰란스* 만노시다아제(*Cellulosimicrobium cellulans* mannosidase)인 것인, 방법.

청구항 42

제37항에 있어서,

상기 만노시다아제는 패밀리 38 글리코실 히드롤라제인 것인, 방법.

청구항 43

제37항 또는 제42항에 있어서,

상기 만노시다아제는 *카나발리아 엔스폴미스* 만노시다아제인 것인, 방법.

청구항 44

제37항 내지 제43항 중 어느 한 항에 있어서,

상기 접촉시키는 것은 재조합 균류 세포에서 발생하고, 상기 균류 세포는 상기 만노시다아제를 발현시키는 것인, 방법.

청구항 45

캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 포함하는 분자 복합체를 제조하는 방법으로서,

상기 방법은, 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해할 수 있는 만노시다아제와 분자 복합체를 접촉시키는 것을 포함하고, 상기 분자 복합체는 GAA 활성도를 가지고, 적어도 2 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는 SEQ ID NO: 1 에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 갖고, 각각의 세그먼트는 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 아미노산 서열의 단백질분해에 의해 유도되고,

상기 접촉시키기 전에, 상기 폴리펩티드의 적어도 하나는 만노스-1-포스포-6-만노스 모이어티를 포함하는 것인, 방법.

청구항 46

제45항에 있어서,

상기 만노시다아제는 패밀리 38 글리코실 히드롤라제인 것인, 방법.

청구항 47

제45항 또는 제46항에 있어서,

상기 만노시다아제는 *카나발리아 엔스폴미스* 만노시다아제인 것인, 방법.

청구항 48

제45항 또는 제46항에 있어서,

상기 만노시다아제는 *야로위야 리폴리카* 만노시다아제인 것인, 방법.

청구항 49

캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 포함하는 분자 복합체를 제조하는 방법으로서,

상기 방법은,

a) 상기 분자 복합체에서 적어도 하나의 폴리펩티드에서 만노스-6-포스페이트 모이어티를 캡핑제거하기 위해, 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해할 수 있는 만노시다아제와 분자 복합체를 접촉시키는 단계로서, 상기 분자 복합체는 GAA 활성도를 가지고, 적어도 두 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 아미노산 서열의 세그먼트와 적어도 85 %의 서열 동일성을 가지고, 각각의 세그먼트는, 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도된 것인, 단계; 및

b) 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있는 만노시다아제와 상기 분자 복합체를 접촉시키는 단계;

를 포함하는 것인, 방법.

청구항 50

제49항에 있어서,

단계 (a) 및 단계 (b)는 두 가지 상이한 효소에 의해 촉매작용되는 것인, 방법.

청구항 51

제49항에 있어서,

단계 (a) 및 단계 (b)는 단일 효소에 의해 촉매작용되는 것인, 방법.

청구항 52

제49항 또는 제50항에 있어서,

상기 접촉시키는 것은 재조합 균류 숙주 세포(fungal host cell)에서 발생하고, 상기 균류 숙주 세포는, 단계 (a)를 촉매작용할 수 있는 만노시다아제 및 단계 (b)를 촉매작용할 수 있는 만노시다아제를 발현하는 것인(said fungal host cell expressing a mannosidase capable of catalyzing step (a) and a mannosidase capable of catalyzing step (b)), 방법.

청구항 53

제49항 또는 제51항에 있어서,

상기 접촉시키는 것은 재조합 균류 숙주 세포에서 발생하고, 상기 균류 숙주는 단계 (a) 및 (b)를 촉매작용할

수 있는 만노시다아제를 발현시키는 것인, 방법.

청구항 54

제32항 내지 제53항 중 어느 한 항에 있어서,

상기 방법은, 상기 캡핑제거되고 디만노실레이트된 N-글리칸을 포함하는 적어도 하나의 폴리펩티드를 포함하는 상기 분자 복합체와 포유동물 세포를 접촉시키는 것을 추가적으로 포함하고,

상기 접촉 후에, 상기 분자 복합체는 증진된 효율성을 갖는 상기 포유동물 세포의 내부로 수송되는 것인, 방법.

청구항 55

제54항에 있어서,

상기 포유동물 세포는 인간 세포인 것인, 방법.

청구항 56

세포의 내부로 GAA 활성도를 갖는 분자 복합체를 수송하는 방법으로서,

상기 방법은 상기 분자 복합체와 포유동물 세포를 접촉시키는 것을 포함하고, 상기 분자 복합체는 적어도 두 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 %의 서열 동일성을 갖고, 각각의 세그먼트는, 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도된 것이고;

상기 폴리펩티드의 적어도 하나에서 인산화된 N-글리칸은 제33항 내지 제55항 중 어느 한 항의 방법에 기재된 바와 같이 캡핑제거되고 디만노실레이트된 것인, 방법.

청구항 57

GAA 활성도를 갖는 분자 복합체를 세포의 내부로 수송하는 방법으로서,

상기 방법은 상기 분자 복합체와 포유동물 세포를 접촉하는 것을 포함하고, 상기 분자 복합체는 적어도 두 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 가지고, 각각의 세그먼트는 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도되고, 상기 분자 복합체는, 포유동물 세포의 상기 내부로 수송되도록, 상기 분자 복합체의 증진된 능력을 결과적으로 나타내는 적어도 하나의 변형을 포함하는 것인, 방법.

청구항 58

제56항 또는 제57항에 있어서,

상기 포유동물 세포는 *생체 외*에 있는 것인, 방법.

청구항 59

제56항 또는 제57항에 있어서,

상기 포유동물 세포는 포유동물 피검자에 있는 것인, 방법.

청구항 60

제58항 또는 제59항에 있어서,

상기 포유동물 세포는 인간 세포인 것인, 방법.

청구항 61

제57항 내지 제60항 중 어느 한 항에 있어서,

상기 변형은 상기 폴리펩티드의 적어도 하나에 융합된 타겟팅 도메인을 포함하고,

상기 타겟팅 도메인은 타겟 세포의 표면 상에 수용체의 세포외 도메인을 결합하는 것인, 방법.

청구항 62

제57항 내지 제60항 중 어느 한 항에 있어서,

상기 적어도 하나의 변형은 상기 폴리펩티드의 적어도 하나에 융합된 유로키나제-타입 플라스미노겐 수용체를 포함하는 것인, 방법.

청구항 63

제57항 내지 제60항 중 어느 한 항에 있어서,

상기 적어도 하나의 변형은, 상기 폴리펩티드의 적어도 하나에 융합된 인간 인슐린-유사 성장 인자 Ⅱ의 인식 도메인을 포함하는 것인, 방법.

청구항 64

SEQ ID NO : 8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 알칼리성 프로테아제(alkaline protease)를 코딩하는 핵산 및 SEQ ID NO: 1에 기재된 상기 GAA 아미노산 서열을 코딩하는 핵산을 포함하는 분리된 균류 세포로서,

상기 균류 세포는 GAA 활성도를 갖는 분자 복합체를 생산하고, 적어도 두 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는, SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 갖고, 각각의 세그먼트는 상기 알칼리성 프로테아제에 의해 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도되는 것인, 분리된 균류 세포.

청구항 65

제64항에 있어서,

각각의 세그먼트는, 상기 알칼리성 프로테아제에 의해 아미노산 60과 아미노산 65 사이의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도되는 것인, 분리된 균류 세포.

청구항 66

제64항 또는 제65항에 있어서,

상기 균류 세포는 만노시다아제를 코딩하는 핵산을 추가적으로 포함하고, 상기 만노시다아제는, (ⅰ) 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해하고, (ii) 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있는 것인, 균류 세포.

청구항 67

제64항 또는 제65항에 있어서,

상기 균류 세포는 만노시다아제를 코딩하는 핵산을 추가적으로 포함하고, 상기 만노시다아제는 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해할 수 있는 것인, 균류 세포.

청구항 68

제64항 또는 제65항에 있어서,

상기 균류 세포는 만노시다아제를 코딩하는 핵산을 추가적으로 포함하고, 상기 만노시다아제는, 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있는 것인, 균류 세포.

청구항 69

제65항에 있어서,

상기 균류 세포는, 만노실 인산화를 촉진할 수 있는 폴리펩티드를 코딩하는 핵산을 추가적으로 포함하는 것인, 균류 세포.

청구항 70

제64항 내지 제69항 중 어느 한 항에 있어서,

상기 균류 세포는 OCH1 활성도가 결핍되도록 유전적으로 조작된 것인, 균류 세포.

청구항 71

SEQ ID NO : 8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 알칼리성 프로테아제를 코딩하는 외인성 핵산(exogenous nucleic acid)을 포함하는 분리된 균류 세포.

발명의 설명

기 술 분 야

[0001] **관련된 출원에 대한 교차-참고**

[0002] 본 출원은, 2012년 3월 15일에 출원된, U.S. 가출원 시리즈 번호 No. 61/611,485의 이익을 청구한다. 상기 사전의 출원의 내용은, 본 출원의 내용의 일부(및 본 출원의 내용에서 참고문헌으로 포함됨)로 고려된다.

[0004] **기술 분야**

[0005] 본 발명은, 산 알파 글루코시다아제 활성도(acid alpha glucosidase activity)를 갖는 분리된 분자 복합체, 및 보다 특히, 전구체 분자로부터 단백질분해(proteolysis)에 의해 유도된 적어도 두 개의 폴리펩티드를 포함하는 분자 복합체에 관한 것으로, 상기 분자 복합체는 포유동물의 세포의 내부로 수송하도록 상기 분자 복합체의 증진된 활성도를 결과적으로 나타내는 적어도 하나의 변형을 함유한다.

배 경 기 술

[0007] **배경기술**

[0008] 폼페병[글리코겐-저장 질병 타입 Ⅱ(glycogen-storage disease type Ⅱ) 또는 산-말타아제 결핍증(acid-maltase deficiency)으로서 또한 언급됨]은, 산 알파 글루코시다아제(acid alpha glucosidase, GAA)의 결핍증으로 인하여 상기 리소좀에서 글리코겐의 축적을 결과적으로 나타내는 드문 상염색체 열성 질환(rare autosomal recessive disorder)이다. 글리코겐의 증가(build-up)는, 몸 전체를 통해 점진적인 근육 무기력증(progressive muscle weakness)(근육병)을 일으키고, 심장, 골격근(skeletal muscles), 간 및 신경계를 포함하는, 다양한 체조직에 영향을 미친다.

[0010] 폼페병은, 유아기 및 후발성의 형태(infantile and late onset forms)로 나눌 수 있다. 상기 유아기-발병 형태에서, 유아는, 약함(weakness) 및 느슨함(floppiness)을 갖는 영아초기(early infancy)(4 내지 8 세) 동안에 일반적으로 존재하고, 이들의 손을 들 수도 없고, 나가떨어지는 것(rolling over)과 같은, 이들의 나이에 대해 흔한 그 밖의 운동 과제를 할 수 없다. 치료 없이, 폼페병을 갖는 유아는 심부전 및 호흡근육의 약화로 인하여 12 개월 이전에 일반적으로 사망한다. United Pompe Foundation를 참고하라. 후발성 형태(청소년 및 성인의 형태를 포함함)는 후천적 발병을 가지고, 상기 유아기 형태보다 보다 더 느리게 진행한다. 재조합 인간 GAA(Myozyme® or Lumizyme®)는 폼페병을 치료하기 위해 사용된다. 그러나, Myozyme® 또는 Lumizyme®는, 1년당 $300,000 이상의 비용이 드는, 둘 다 매우 비싸다. 보통, 폼페병을 위한 개선된 치료에 대한 필요성이 있다.

발명의 내용

[0012] **요약**

[0013] 하나의 측면에서, 본 문서는 산 알파 글리코시다아제(GAA) 활성도를 갖는 분리된 분자 복합체를 특징으로 하고, 적어도 2 개의 폴리펩티드(예를 들어, 적어도 세 개 또는 적어도 네 개의 폴리펩티드)를 포함하고, 각각의 폴리펩티드는, SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트(segment)와 적어도 85 %(예를 들어, 적어도 90 %, 95 %, 99 %, 또는 100 %) 서열 동일성을 갖고, 각각의 세그먼트는, 아미노산 50과 아미노산 74 사이(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이)의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 제시된 상기 아미노산 서열의 단백질분해에 의해 유도된 것이다. 상기 분자 복합체는, 포유동물 세포의 내부로 수송되도록 상기 분자 복합체의 증진된 활성도를 결과적으로 나타내는 적어도 하나의 변형을 함유한다. SEQ ID NO : 1에 기재된 아미노산 서열의 단백질분해(Proteolysis)는, SEQ ID NO : 1에 기재된 아미노산 서열의 아미노산 719과 아미노산 746 사이의 하나 또는 그 이상의 부위에서의 분열(cleavage) 또는 아미노산 137과 아미노산 151 사이의 하나 또는 그 이상의 부위에서의 분열을 포함할 수 있다. 단백질분해는 추가적으로, SEQ ID NO : 1에 기재된 아미노산 서열의 아미노산 719과 아미노산 746 사이의 하나 또는 그 이상의 부위에서의 분열(cleavage) 및 SEQ ID NO : 1에 기재된 아미노산 서열의 아미노산 137과 아미노산 151 사이의 하나 또는 그 이상의 부위에서의 분열을 포함할 수 있다.

[0015] 본원에 기재된 상기 분자 복합체 중 어떠한 것도, 상기 폴리펩티드의 적어도 하나는 하나 또는 그 이상의 인산화된 N-글리칸을 함유할 수 있고, 상기 변형은 적어도 하나의 인산화된 N-글리칸의 캡핑제거(uncapping) 및 디만노실레이션(demannosylation)을 함유할 수 있다. 상기 폴리펩티드의 적어도 하나에서 적어도 40 %(예를 들어, 적어도 60 %, 80 %, 90 %, 95 %, 또는 99 %)의 N-글리칸은 캡핑제거되고 디만노실레이트될 수 있다(At least 40 % of the N-glycans on at least one of the polypeptides can be uncapped and demannosylated).

[0017] 본원에 기재된 상기 분자 복합체 중의 어떠한 것들에서, 적어도 두 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 22 내지 57을 포함하고(the segment includes amino acids 22 to 57 of SEQ

ID NO:1), 적어도 두 개의 폴리펩티드 중의 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 66 내지 896을 함유한다.

[0019] 적어도 세 개의 폴리펩티드를 함유하는 본원에 기재된 상기 분자 복합체 중의 어떠한 것에서, 적어도 세 개의 폴리펩티드 중의 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 22 내지 57을 함유하고, 적어도 세 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 66 내지 726을 함유하고, 적어도 세 개의 폴리펩티드 중의 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 736 내지 896을 함유한다.

[0021] 적어도 네 개의 폴리펩티드를 함유하는 본원에 기재된 상기 분자 복합체 중 어떠한 것에서, 상기 적어도 네 개의 폴리펩티드 중의 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 22 내지 57을 함유하고, 적어도 네 개의 폴리펩티드 중 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 66 내지 143을 함유하고, 상기 적어도 네 개의 폴리펩티드 중의 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 158 내지 726 을 함유하고, 적어도 네 개의 폴리펩티드 중의 하나에 대해, 상기 세그먼트는 SEQ ID NO : 1의 아미노산 736 내지 896을 함유한다.

[0023] 본원에 기재된 상기 분자 복합체 중의 어떠한 것에서, 상기 적어도 하나의 변형은, 상기 분자 복합체에서 적어도 하나의 폴리펩티드에 그 다음에 융합된 것 중의 어떠한 하나를 함유할 수 있다 : 세포외 수용체에 대한 리간드, 타겟 세포의 표면에서 수용체의 세포외 도메인을 결합하는 타겟팅 도메인, 유로키나제-타입 플라스미노겐 수용체, 또는 인간 인슐린-유사 성장 인자(Ⅱ)의 인식 도메인(In any of the molecular complexes described herein, the at least one modification can include any one of the following fused to at least one polypeptide in the molecular complex: a ligand for an extracellular receptor, a targeting domain that binds an extracellular domain of a receptor on the surface of a target cell, a urokinase-type plasminogen receptor, or the recognition domain of human insulin-like growth factor II).

[0025] 본 문서는, 본원에 기재된 상기 분자 복합체 중의 어떠한 것을 함유하는 조성물을 또한 특징으로 하고, 상기 분자 복합체는 동결건조된다. 상기 조성물은 단일 사용 바이알(single use vial)로서 포장될 수 있다.

[0027] 본 문서는 본원에 기재된 분자 복합체 중의 어떠한 것 및 약제학적으로 허용가능한 담체(carrier)를 함유하는 약제학적 조성물을 또한 특징으로 한다. 상기 조성물은 정맥내 또는 피하의 투여를 위해 제형화될 수 있다. 상기 조성물은 정맥내 주입을 위해 제형화될 수 있다.

[0029] 또 다른 측면에서, 본 문서는 폼페병을 치료하는 방법을 특징으로 한다. 상기 발명은, 폼페병으로 진단된 환자에게 본원에 기재된 상기 조성물 중의 어떠한 것을 투여하는 것을 포함한다. 상기 환자는 유아기 발병 폼페병 또는 후발성 폼페병으로 진단될 수 있다.

[0031] 본 문서는 분자 복합체를 제조하기 위한 방법을 또한 특징으로 한다. 상기 방법은, SEQ ID NO : 8 에 기재된 상기 아미노산 서열과 적어도 85 %(예를 들어, 적어도 90 %, 적어도 95 %, 적어도 99 %, 또는 100 %) 서열 동일성을 갖는 프로테아제(protease)와, SEQ ID NO : 1에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 폴리펩티드를 접촉시키는 것을 포함하고, 상기 프로테아제는, 아미노산 50과 아미노산 74 사이(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이)의 하나 또는 그 이상의 부위에서 상기 폴리펩티드를 분해한다(cleaves). 상기 접촉하는 단계는 *생체 외에서* 실행될 수 있다.

[0033] 본 문서는 또한, 캡핑제거되고 디만노실레이트된 인산화된 N-글리칸(uncapped and demannosylated phosphorylated N-glycans)을 함유하는 분자 복합체를 제조하기 위한 방법을 또한 특징으로 한다. 상기

방법은, (ⅰ) 만노스-1-포스포-6-만노스 모이어트를 만노스-6-포스페이트로 가수분해하고, (ii) 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있는 만노시다아제와 분자 복합체를 접촉시키는 것을 포함하고, 상기 분자 복합체는 GAA 활성도를 가지고, 적어도 2 개의 폴리펩티드를 함유하고, 각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 가지고, 각각의 세그먼트는, 아미노산 50과 아미노산 74 사이(예를 들어, 아미노산 56과 아미노산 58 사이 또는 아미노산 60 또는 아미노산 65 사이)의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도되고, 접촉하기 전에, 적어도 하나의 폴리펩티드는 하나 또는 그 이상의 만노스-1-포스포-6-만노스 모이어티를 함유하는 인산화된 N-글리칸을 함유한다[The method includes contacting a molecular complex with a mannosidase capable of (i) hydrolyzing a mannose-1-phospho-6-mannose moiety to mannose-6-phosphate and (ii) hydrolyzing terminal alpha-1,2 mannose, alpha-1,3 mannose and/or alpha-1,6 mannose linkages, the molecular complex having GAA activity and including at least two polypeptides, each polypeptide having at least 85% sequence identity to a segment of the amino acid sequence set forth in SEQ ID NO: 1, each segment being derived by proteolysis of the amino acid sequence set forth in SEQ ID NO: 1 at one or more sites between amino acid 50 and amino acid 74 (e.g., between amino acid 56 and amino acid 68 or between amino acid 60 and amino acid 65), wherein before the contacting, at least one of the polypeptides includes phosphorylated N-glycans containing one or more mannose-1-phospho-6-mannose moieties]. 상기 만노시다아제는, 패밀리 38 글리코실 히드롤라제[예를 들어, *카나발리아 엔스폴미스* 만노시다아제(*Canavalia ensiformis* mannosidase) 또는 *야로위야 리폴리카* 만노시다아제(*Yarrowia lipolytica* mannosidase)]일 수 있다. 상기 접촉시키는 것은, 상기 만노시다아제를 발현하는 재조합 균류 세포(recombinant fungal cell)에서 발생할 수 있다.

[0035] 본 문서는 또한, 캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 함유하는 분자 복합체를 제조하는 방법을 특징으로 한다. 상기 방법은, 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있는 만노시다아제와 분자 복합체를 접촉하는 것을 포함하고, 상기 분자 복합체는 GAA 활성도를 가지고, 적어도 두 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 가지고, 각각의 세그먼트는, 아미노산 50 과 아미노산 74 사이(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이)의 하나 또는 그 이상의 부위에서 SEQ ID NO : 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도되고, 상기 폴리펩티드 중의 적어도 하나는, 접촉시키기 전에, 캡핑제거된 만노스-6-포스페이트 모이어티를 포함하는 인산화된 N-글리칸을 함유한다(at least one of the polypeptides includes prior to contacting, phosphorylated N-glycans comprising uncapped mannose-6-phosphate moieties). 상기 만노시다아제는 패밀리 47 글리코실 히드롤라제[예를 들어, *아스페르길루스 사토이* 만노시다아제(*Aspergillus satoi* mannosidase)], 패밀리 92 글리코실 히드롤라제[예를 들어, *셀룰로시마이크로비움 셀룰란스* 만노시다아제(*Aspergillus satoi* mannosidase)], 또는 패밀리 38 글리코실 히드롤라제[예를 들어, *카나발리아 엔스폴미스* 만노시다아제(*Canavalia ensiformis* mannosidase)]일 수 있다. 상기 접촉시키는 것은, 상기 만노시다아제를 발현하는 재조합 균류 세포에서 발생할 수 있다.

[0037] 본 문서는 캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 함유하는 분자 복합체를 제조하는 방법을 또한 특징으로 한다. 상기 방법은, 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해할 수 있는 만노시다제와 분자 복합체를 접촉시키는 것을 포함하고, 상기 분자 복합체는 GAA 활성도를 가지고, 적어도 두 개의 폴리펩티드를 함유하고, 각각의 폴리펩티드는 SEQ ID NO : 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 가지고, 각각의 세그먼트는 아미노산 50과 아미노산 74 사이(예를 들어 아미노산 56과 아미노산 68 또는 아미노산 60과 아미노산 65 사이)의 하나 또는 그 이상의 부위에서 SEQ ID NO : 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도되고, 상기 폴리펩티드의 적어도 하나는, 접촉시키기 전에, 하나 또는 그 이상의 만노스-1-포스포-6-만노스 모이어티를 함유한다. 상기 만노시다아제는 패밀리 38 글리코실 히드롤라제일 수 있다(예를 들어, *카나발리아 엔스폴미스* 만노시다아제 또는 *야로위야 리폴리카* 만노시다아제).

[0039] 또 다른 측면에서, 본 문서는 캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 함유하는 분자 복합체를 제

조하는 방법을 특징으로 한다. 상기 방법은, a) 상기 분자 복합체에서 적어도 하나의 폴리펩티드에서 만노스-6-포스페이트 모이어티를 캡핑제거하기 위해, 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해할 수 있는 만노시다아제와 분자 복합체를 접촉시키는 단계로서, 상기 분자 복합체는 GAA 활성도를 가지고, 적어도 두 개의 폴리펩티드를 포함하고, 각각의 폴리펩티드는, SEQ ID NO : 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 가지고, 각각의 세그먼트는 아미노산 50과 아미노산 74 사이(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이)의 하나 또는 그 이상의 부위에서 SEQ ID NO : 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도되는 것인, 단계; 및 b) 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있는 만노시다아제와 상기 분자 복합체를 접촉시키는 단계를 포함한다. 단계 (a) 및 단계 (b)는, 단일 효소에 의해 촉매작용되거나 또는 두 개의 상이한 효소에 의해 촉매작용될 수 있다. 상기 접촉시키는 단계는 어떠한 순서로, 함께 또는 별도로 실행될 수 있다(The contacting steps can be performed together or separately, and in either order). 상기 접촉시키는 단계는 재조합 균류 숙주 세포에서 발생할 수 있고, 상기 균류 숙주 세포는 단계 (a)를 촉매작용할 수 있는 만노시다아제 및 단계 (b)를 촉매작용할 수 있는 만노시다아제를 발현한다(the fungal host cell expressing a mannosidase capable of catalyzing step (a) and a mannosidase capable of catalyzing step (b)) 상기 접촉시키는 단계는 재조합 균류 숙주 세포에서 발생할 수 있고, 상기 균류 숙주 세포는 단계 (a) 및 (b)를 촉매작용할 수 있는 만노시다아제를 발현한다.

[0041] 적어도 하나의 캡핑제거되고 디만노실레이트된 N-글리칸을 함유하는 본원에 기재된 상기 분자 복합체 중의 어떠한 것은, 포유동물 세포를 접촉하는데 사용될 수 있고, 접촉한 후에, 상기 분자 복합체는 증진된 효율을 갖는 상기 포유동물 세포의 내부로 수송된다. 상기 포유동물 세포는 인간 세포일 수 있다.

[0043] 본 문서는 세포의 내부로 GAA 활성도를 갖는 분자 복합체를 수송하는 방법을 또한 특징으로 한다. 상기 방법은 상기 분자 복합체와 포유동물 세포를 접촉하는 것을 포함하고, 각각의 폴리펩티드는 SEQ ID NO : 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 가지고, 각각의 세그먼트는, 아미노산 50과 아미노산 74 사이(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이)의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도된 것이고; 상기 적어도 하나의 폴리펩티드에서 인산화된 N-글리칸은 본원에 기재된 방법에 기재된 바와 같이 캡핑제거되고 디만노실레이트된다. 상기 포유동물 세포는 *생체 외* 또는 포유동물 피검자일 수 있다. 상기 포유동물의 세포는 인간 세포일 수 있다.

[0045] 또 다른 측면에서, 본 문서는 세포의 내부로 GAA 활성도를 갖는 분자 복합체를 수송하는 방법을 특징으로 한다. 상기 방법은, 적어도 두 개의 폴리펩티드를 포함하는 상기 분자 복합체와 포유동물 세포를 접촉시키는 것을 포함하고, 각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 가지고, 각각의 세그먼트는, 아미노산 50과 아미노산 74 사이(아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이)의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1에 기재된 아미노산 서열의 단백질분해에 의해 유도된 것이고, 상기 분자 복합체는, 포유동물 세포의 상기 내부에 수송되도록 상기 분자 복합체의 증진된 활성도를 결과적으로 나타내는 적어도 하나의 변형을 포함한다. 상기 포유동물 세포는 *생체 외* 또는 포유동물 피검자 일 수도 있다. 상기 포유동물 세포는 인간 세포일 수 있다. 상기 변형은 상기 분자 복합체에서 적어도 하나의 폴리펩티드로 그 다음에 융합된 것 중의 하나를 포함할 수 있다 : 세포외 수용체에 대한 리간드, 타겟 세포의 표면에서 수용체의 세포외 도메인에 결합하는 타겟팅 도메인, 유로키나아제-타입 플라스미노겐 수용체, 또는 인간 인슐린-유사 성장 인자 Ⅱ의 인식 도메인(The mammalian cell can be a human cell. The modification can include any one of the following fused to at least one polypeptide in the molecular complex: a ligand for an extracellular receptor, a targeting domain that binds an extracellular domain of a receptor on the surface of a target cell, a urokinase-type plasminogen receptor, or the recognition domain of human insulin-like growth factor II).

[0047] 또 다른 측면에서, 본 문서는, SEQ ID NO : 8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는

알칼리성 프로테아제를 코딩하는 외인성 핵산을 함유하는 분리된 균류 세포를 특징으로 한다.

[0049] 본 문서는, SEQ ID NO : 8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 알칼리성 프로테아제를 코딩하는 핵산 및 SEQ ID NO : 1에 기재된 상기 GAA 아미노산 서열을 코딩하는 핵산을 포함하는 분리된 균류 세포를 또한 특징으로 한다. 상기 균류 세포는 GAA 활성도를 갖고 적어도 두 개의 폴리펩티드를 포함하는 분자 복합체를 생산하고, 각각의 폴리펩티드는 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 갖고, 각각의 세그먼트는, 상기 알킬리성 프로테아제(alkaline protease)에 의해 아미노산 50과 아미노산 74 사이(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이)의 하나 또는 그 이상의 부위에서, SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도된다. 몇몇 실시형태에서, 상기 균류 세포는 만노시다아제를 코딩하는 핵산을 추가적으로 포함하고, 상기 만노시다아제는 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해할 수 있다. 몇몇 실시형태에서, 상기 균류 세포는 만노시다아제를 코딩하는 핵산을 추가적으로 함유하고, 상기 만노시다아제는 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있다. 몇몇 실시형태에서, 상기 균류 세포는 추가적으로 만노시다아제를 코딩하는 핵산을 함유할 수 있고, 상기 만노시다아제는, (ⅰ) 만노스-1-포스포-6-만노스 모이어티를 만노스-6-포스페이트로 가수분해할 수 있고, (ii) 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합을 가수분해할 수 있다. 이러한 균류 세포의 어떠한 것은 추가적으로 만노실 인산화를 촉진할 수 있는 폴리펩티드를 코딩하는 핵산을 함유할 수 있고, 및/또는 OCH1 활성도가 결핍되도록 유전적으로 조작된다(Any of such fungal cells further can include a nucleic acid encoding a polypeptide capable of promoting mannosyl phosphorylation and/or be genetically engineered to be deficient in OCH1 activity).

[0051] 다른 방식으로 정의되지 않는다면, 본원에 사용된 모든 기술 및 과학 용어는, 이러한 발명에 속하는 본 분야의 어떠한 숙련자에 의해 일상적으로 이해되는 바와 같은 동일한 의미를 갖는다. 본원에 기재된 이러한 것들과 유사하거나 또는 동일한 방법 및 물질은 본 분야의 실행 또는 테스팅에 사용될 수 있을지라도, 대표적인 방법 및 물질은 하기에 기재되어 있다. 본원에 언급된 모든 공개물, 특허 출원, 특허, Genbank® 접근 번호 및 그 밖의 참고문헌은 이의 전체가 참고문헌으로 포함된다. 충돌하는 경우에, 정의를 포함하는 본 출원은 조절될 것이다. 상기 물질, 방법 및 예는 단지 설명적인 것이고, 이로 제한하는 의도는 아니다.

[0053] 본 발명의 그 밖의 특징 및 장점은, 청구의 범위 및 하기의 상세한 설명으로부터 명백할 것이다.

도면의 간단한 설명

[0055] 도 1은, 상기 단일 서열의 분열 후에 인간 산 알파 글루코시다아제(GAA)의 상기 아미노산 서열(SEQ ID NO : 1)의 서술이다.

도 2a는, DsbA-*셀룰로시마이크로비움 셀룰란스* 만노시다아제 5(CcMan5)의 오픈 리딩 프레임(open reading frame, ORF)의 상기 뉴클레오티드 서열(SEQ ID NO : 2)의 서술이다.

도 2b는, 굵은 활자체(bold)와 함께 상기 신호 서열(signal sequence)과 함께, 상기 CcMan5 폴리펩티드의 상기 아미노산 서열의 서술(SEQ ID NO: 3)이다.

도 2c는, 신호 서열 없이 상기 CcMan5 폴리펩티드의 상기 아미노산 서열(SEQ ID NO : 4)의 서술이다. 상기 신호 서열 없이 상기 CcMan5 폴리펩티드의 상기 예측된 분자 중량은 173 kDa이다.

도 3a 및 3b는, CcMan5 및 JbMan로 처리된 rhGAA의 상기 N-글리칸 분석을 나타내는 일련의 일렉트로페로그램(electropherogram)이다. 분석은, DNA 시퀀서-보조, 형광단-보조 탄수화물 전기영동(DNA sequencer-assisted, fluorophore-assisted carbohydrate electrophoresis, DSA-FACE)을 사용하여 실행되었다. 상기 Y-축은, 각각의 N-글리칸 구조의 양의 표시로서 상대적인 형광 유닛(relative fluorescence units)을 나타낸다. 상기 X-축은, 모세혈관을 통해 각각의 N-글리칸 구조의 상기 상대적인 유동성(relative mobility)을 나타낸다. 도 3a 및 도 3b 둘 다에서, 패널 A 는, PNGaseF와 함께 RNaseB로부터 방출된 상기 N-글리칸을 함유하는 표준 시료

(reference sample)이다. 도 3a에서, 패널 B 및 C는, CcMan5 및 JbMan으로, 각각, 처리 전 및 후에, huGAA(76 kD 변이체)로부터 N-글리칸 분석을 함유한다. 도 3b에서, 패널 B, C, 및 D는, 각각 huGAA 76 kD 형태, 95 kD 형태, 및 110 kD 형태로부터 상기 N-글리칸 분석을 함유한다.

도 4는, 기질로서 토끼 간 글리코겐을 사용하여 마이오자임(Myozyme)(●), 76 kDa GAA(▲), 95 kDa GAA(▼), 및 110 kDa GAA(◆)을 갖는 분 당 형성된 글루코스의 양의 선 그래프이다.

도 5a는, 심장에서 개별적인 마우스의 상기 글리코겐 레벨(㎍/mg 단백질)의 두 개의 그림을 포함한다. 도 5b는, 골격근에서 개별적인 마우스의 상기 글리코겐 레벨(㎍/mg 단백질)의 두 개의 그림을 포함한다. 빨간 점은 암컷이고, 검은 점은 수컷이다. 선은 각각의 그룹의 중간값을 나타낸다.

도 6 은, *야로위야 리폴리카* AMS1 만노시다아제의 아미노산 서열의 서술을 함유한다(SEQ ID NO: 5).

도 7은, *아스페르길루스 사토이* 만노시다아제의 아미노산 서열의 서술을 포함한다(SEQ ID NO:6).

도 8은, 굵은 활자체로 신호 서열과 함께, *셀룰로시마이크로비움 셀룰란스* 만노시다아제 4(CcMan4, SEQ ID NO:7)의 아미노산 서열의 서술을 포함한다. 상기 신호 서열이 없는 상기 CcMan4 폴리펩티드의 상기 예측된 분자량은 184 kDa이다.

도 9는, 상기 신호 펩티드(21 개의 아미노산), 프로-펩티드(100 개의 아미노산) 및 성숙한 단백질(282 개의 아미노산)을 포함하는 *아스페르길루스 오리제* 알칼리성 프로테아제(*Aspergillus oryzae* alkaline protease)의 아미노산 서열의 서술을 포함한다.

도 10 은, 상기 *A. 오리제* 알칼리성 프로테아제를 코딩하는 상기 *Y. 리폴리티카* 코돈 최적화된 서열을 포함하는 융합 구조물의 상기 뉴클레오티드 서열의 서술을 포함한다(SEQ ID NO:10)(FIG. 10 contains a depiction of the nucleotide sequence of the fusion construct containing the *Y. lipolytica* codon optimized sequence encoding the *A.oryzae* alkaline protease). 클로닝을 위해 사용된 제한 부위(Restriction sites)는 밑줄쳐져 있다. 상기 링커를 코딩하는 상기 뉴클레오티드 서열은 굵은 글씨체이고, 상기 His 태그를 코딩하는 상기 뉴클레오티드 서열(10 His 잔기)은 이탤릭체로 되어 있다(italicized).

발명을 실시하기 위한 구체적인 내용

[0056] **상세한 서술(DETAILED DESCRIPTION)**

[0057] 일반적으로, 본 문서는, 포유동물 세포의 내부로 수송되도록, 증진된 능력을 결과적으로 나타내는 적어도 하나의 변형 및 산 알파-글리코시다아제(GAA) 활성도를 갖는 분리된 분자 복합체를 제공한다. GAA는 N-결합된 글리칸을 함유하는 110 kDa 전구체로서 합성되었다. 상기 전구체는 상기 신호 서열을 제거하기 위해 단백질 가수분해적으로 처리되고, 그 다음에 95 kDa, 76 kDa, 및 70 kDa의 주요한 종류로 추가적으로 단백질 가수분해적으로 진행된다(The precursor is proteolytically processed to remove the signal sequence and then further proteolytically processed to major species of 95 kDa, 76 kDa, and 70 kDa). 그러나, 상기 전구체로부터 방출된 상기 펩티드의 적어도 몇몇은 상기 주요한 종과 연관되어 남아있다. 예를 들어, Moreland *et al.*, *J. Biol. Chem.*, 280:6780-6791 (2005)를 참고하라. 따라서, 본원에 기재된 GAA 활성도를 갖는 상기 분자 복합체는, 하나 또는 그 이상의 부위에서 상기 전구체 분자의 단백질분해의 분열로부터 유도된 적어도 두 개의 폴리펩티드(적어도 두 개, 세 개, 또는 네 개의 폴리펩티드)를 포함한다. 상기 분자 복합체에서 적어도 두 개의 폴리펩티드는, 상기 전구체에서 하나 또는 그 이상의 부위에서 단백질분해의 분열을 결과적으로 나타낸다. 예를 들어, SEQ ID NO : 1 에 기재된 상기 아미노산 서열의 단백질분해는, 적어도 두 개의 폴리펩티드를 생산하도록, 아미노산 50과 아미노산 74 사이, 예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이일 수 있다. 두 개의 폴리펩티드를 포함하는 분자 복합체는 본원에서 상기 95 kDa 형태로서 언급되어 있다.

[0059] 몇몇의 실시형태에서, 상기 분자 복합체에서 적어도 세 개의 폴리펩티드는, 상기 전구체에서 둘 또는 그 이상의 부위에서 단백질분해의 분열을 결과적으로 나타낸다. 예를 들어, SEQ ID NO : 1에 기재된 상기 아미노산 서열의 아미노산 50과 아미노산 74 사이의 분열(예를 들어, 아미노산 50과 아미노산 74 사이 또는 아미노산 60과 아미노산 65 사이), 아미노산 719와 아미노산 746 사이의 하나 또는 그 이상의 부위에서의 분열 또는 아미노산 137과 아미노산 151 사이의 하나 또는 그 이상의 부위에서의 분열 뿐만 아니라 SEQ ID NO : 1 에 기재된 상기

아미노산 서열의 단백질분해가 포함될 수 있다. 상기 세 개의 폴리펩티드를 포함하는 분자 복합체는 본원에서 76 kDa 형태로서 언급된다.

[0061] 몇몇 실시형태에서, 상기 분자 복합체에서 적어도 네 개의 폴리펩티드는, 상기 전구체에서 3 또는 그 이상의 부위에서 단백질분해의 분열을 결과적으로 나타낸다. 예를 들어, SEQ ID NO : 1에 기재된 상기 아미노산 서열의 단백질분해는, 아미노산 50과 아미노산 74 사이의 분열(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이), SEQ ID NO : 1에 기재된 상기 아미노산 서열의 아미노산 719와 아미노산 746 사이의 하나 또는 그 이상의 부위에서의 분열, 및 SEQ ID NO : 1에 기재된 상기 아미노산 서열의 아미노산 137과 아미노산 151 사이의 하나 또는 그 이상의 부위에서의 분열에 더하여 포함할 수 있다. 4 개의 폴리펩티드를 함유하는 분자 복합체는 본원에서 70 kDa 형태로서 언급되어 있다.

[0063] 분열은 하나의 분자에서 하나 또는 그 이상의 부위에서 발생할 수 있고, 분열의 부위는 상이한 분자에서 상이할 수도 있음을 분명할 것이다.

[0065] *아스페르길루스 오리제*(예를 들어, Sigma 또는 NovozymesCorp로부터)로부터의 프로테아제를 함유하는 상업적으로 입수가능한 프로테아제는, 아미노산 50과 74 사이, 예를 들어 아미노산 56과 68 또는 아미노산 60과 65 사이의 SEQ ID NO : 1에 기재된 상기 아미노산 서열을 분해하기 위해 사용될 수 있다. 그 대신에, *아스페르길루스 오리제*(SEQ ID NO : 8)로부터 상기 알킬리성 프로테아제와 적어도 85 %(예를 들어, 적어도 90 %, 95 %, 97 %, 98 %, 99 %, 또는 100 %) 서열 동일성을 가지는 알칼리성 프로테아제가 사용될 수 있다. 예를 들어, 본원에 기재된 바와 같이, SEQ ID NO : 1에 기재된 상기 아미노산 서열을 갖는 GAA 폴리펩티드는, SEQ ID NO : 8 또는 SEQ ID NO : 9에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 프로테아제와 접촉될 수 있다. SEQ ID NO : 8 는 성숙한 *아스페르길루스 오리제* 알칼리성 프로테아제의 상기 아미노산 서열이다. SEQ ID NO : 9는, 상기 신호 펩티드, 프로-펩티드, 및 성숙한 단백질을 포함하는 상기 *아스페르길루스 오리제* 프로테아제의 상기 아미노산 서열이다. 상기 접촉하는 것은, *아스페르길루스 오리제*를 분리하거나 재조합적으로 생산되는 프로테아제를 사용하여 *생체 외에서* 발생할 수 있다. 선택적으로, 균류의 숙주(fungal host)는, 상기 GAA 폴리펩티드 및 알칼리성 프로테아제가 상기 배양 배지 내로 둘 다 분비되도록 조직될 수 있고, 상기 알칼리성 프로테아제는, 아미노산 50과 아미노산 74 사이(아미노산 56과 68 사이 또는 아미노산 60과 아미노산 65 사이)의 SEQ ID NO : 1에 기재된 상기 아미노산 서열을 분해할 수 있다.

[0067] 본원에 기재된 상기 분리된 분자 복합체는, 포유동물 세포의 내부로 수송되도록, 증진된 능력을 결과적으로 나타내는 적어도 하나의 변형을 가진다. 포유동물 세포의 내부로 수송되는 상기 복합체의 능력을 증진시키는 변형의 비-제한적인 예는, 수송을 가능하게 하는 인산화된 N-글리칸 또는 펩티드 태그(peptide tags)의 캡핑제거(uncapping) 및 디만노실레이션(demannosylation)을 포함한다. 방법 및 물질은, 태그 또는 캡핑제거되고 디만노실레이트된 N-글리칸을 포함하는 분자 복합체를 제조하기 위한 본원에 기재된 것이다(Methods and materials are described herein for preparing molecular complexes containing tags or uncapped and demannosylated N-glycans).

[0069] 본원에 기재된 상기 분리된 분자 복합체는, 폼페병으로 진단된 환자를 포함하는, 폼페병을 갖는 환자, 유아기 발병 폼페병(infantile onset Pompe's disease) 및 후발성 폼페병(late onset Pompe's disease)을 갖는 환자를 치료하기 위해 특히 유용하다. 폼페병은 GAA의 결핍으로 인하여 상기 리소좀에서 글리코겐의 축적으로 발생한다. 글리코겐의 증가(build-up)가 상기 몸을 통해 점진적인 근육 약화(근육병)을 일으키고, 심장, 골격근, 간, 및 신경계를 포함하는, 다양한 몸 조직에 영향을 미친다.

[0071] 상기 분자 복합체에서 상기 폴리펩티드 각각은, SEQ ID NO : 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성(예를 들어, 적어도 90 %, 95 %, 97 %, 98 %, 99 %, 또는 100 %)을 갖고, 각각의 세그먼트는, 아미노산 50과 아미노산 74 사이(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산

65 사이)의 하나 또는 그 이상의 부위에서 SEQ ID NO: 1 에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도된다. SEQ ID NO: 1 에 기재된 상기 아미노산 서열와 특정한 아미노산 서열 사이의 퍼센트 동일성은 하기로서 결정될 수 있다. 첫 번째로, 상기 아미노산 서열은, BLASTP version 2.0.14을 포함하는 BLASTZ의 단독 버전으로부터 상기 BLAST 2 서열(Bl2seq) 프로그램을 사용하여 배열되었다. BLASTZ의 이러한 단독 버전은, 미국 국가 생물공학 정보 웹 사이트(www.ncbi.nlm.nih.gov) 또는 피쉬 & 리처드슨 웹 사이트(e.g., www.fr.com/blast/)로부터 입수가능할 수 있다. 상기 Bl2seq 프로그램을 어떻게 사용하는지를 설명하는 사용설명서는, BLASTZ와 함께 동반하는 리드미 파일(readme file)에서 발견될 수 있다. Bl2seq는, 상기 BLASTP 알고리즘을 사용하여 두 개의 아미노산 서열 사이의 비교를 수행한다. 두 개의 아미노산 서열을 비교하기 위해, 상기 Bl2seq의 선택은 하기와 같다 : -i 는 비교되도록 상기 제1 아미노산 서열을 함유하는 파일로 설정되고(예를 들어, C:₩seq1.txt); -j는 비교하도록 제2 아미노산 서열을 함유하는 파일로 설정되고(예를 들어, C:₩seq2.txt); -p는 blastp로 설정되고; -o 는 어떠한 원하는 파일 이름으로 설정되고(예를 들어, C:₩output.txt); 및 모든 다른 옵션(options)은 이들의 디폴트 세팅(default setting)으로 놓아둔다. 예를 들어, 하기의 명령(command)은, 두 개의 아미노산 서열 사이의 비교를 포함하는 출력 파일을 생성하기 위해 사용될 수 있다: C:₩Bl2seq-i c:₩seq1.txt-j c:₩seq2.txt-p blastp -o c:₩output.txt. 만약 상기 두 개의 비교된 서열이 성동 관계(homology)를 공유한다면, 상기 지정된 출력 파일은 정렬된 서열로서 상동의 이러한 영역을 나타낼 것이다. 만약 상기 두 개의 비교된 서열은 상동 관계를 공유하지 않는다면, 상기 지정된 출력 파일은 정렬된 서열을 나타내지 않을 것이다. 유사한 절차는 blastn이 사용된 것을 제외하고 핵산 서열에 따를 것이다(Similar procedures can be following for nucleic acid sequences except that blastn is used).

[0073] 정렬된 다음에, 매치의 수는 동일한 아미노산 잔기가 서열 둘 다에 존재하는 위치의 수를 계산함으로써 결정된다. 퍼센트 동일성은, 매치의 수를 전체-길이 폴리펩티드 아미노산 서열의 길이로 나눈 다음에 결과적으로 생성된 수치를 100으로 곱함으로써 결정된다.

[0076] *퍼센트 동일성 수치가 소수점 첫 째자리에서 반올림하였음을 참고하라. 예를 들어, 78.11, 78.12, 78.13, 및 78.14는 78.1로 반올림되는 반면에 78.15, 78.16, 78.17, 78.18, 및 78.19는 78.2로 반올림된다. 상기 길이 수치는 항상 정수일 것임을 참고하라.

[0078] 핵산의 수가 특정한 아미노산 서열을 갖는 폴리펩티드를 코딩할 수 있음을 인정할 것이다. 상기 유전자 코드의 겹침(degeneracy)은 본 분야에서 널리 알려져 있다; 즉, 많은 아미노산에 대해, 상기 아미노산에 대한 코돈으로서 역할을 하는 하나 이상의 뉴클레오티드 트리플레트(nucleotide triplet)가 있다. 예를 들어, 제공된 GAA 폴리펩티드에 대한 상기 코딩 서열에서 코돈은, 특정한 종(예를 들어, 박테리아 또는 균류)에서 최적의 발현이 그러한 종에 대해 적절한 코돈 바이어스 테이블을 사용하여, 수득되도록, 변형될 수 있다(For example, codons in the coding sequence for a given GAA polypeptide can be modified such that optimal expression in a particular species is obtained, using appropriate codon bias tables for that species).

[0080] 하나의 실시형태에서, 분자 복합체는 적어도 두 개의 폴리펩티드를 포함할 수 있고, 상기 폴리펩티드 중 하나는 SEQ ID NO : 1의 아미노산 22 내지 57을 포함하고, 그 밖의 폴리펩티드는 SEQ ID NO : 1의 아미노산 66 내지 896을 포함한다.

[0082] 하나의 실시형태에서, 분자 복합체는 적어도 세 개의 폴리펩티드를 포함할 수 있고, 상기 폴리펩티드 중 하나는 SEQ ID NO : 1의 아미노산 22 내지 57을 포함하고, 하나의 폴리펩티드는 SEQ ID NO : 1의 아미노산 66 내지 726을 포함하고, 하나의 폴리펩티드는 SEQ ID NO : 1의 아미노산 736 내지 896을 포함한다.

[0084] 하나의 실시형태에서, 분자 복합체는 적어도 네 개의 폴리펩티드를 포함할 수 있고, 상기 폴리펩티드 중 하나는

SEQ ID NO : 1의 아미노산 22 내지 57을 포함하고, 하나의 폴리펩티드는 SEQ ID NO : 1의 아미노산 66 내지 143을 포함하고, 하나의 폴리펩티드는 SEQ ID NO : 1의 아미노산 158 내지 726을 포함하고, 및 하나의 폴리펩티드는 SEQ ID NO : 1의 아미노산 736 내지 896을 포함한다.

[0086] GAA의 생물학적으로 활성 변이체는 SEQ ID NO: 1에 기재된 상기 서열에 대해 부가, 삭제 또는 치환을 포함할 수 있다. 치환과 함께 GAA 단백질은 10 이하(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10)의 보존적 아미노산 치환을 일반적으로 가질 것이다. 보존적 치환은 유사한 특성을 갖는 다른 것에 대해 하나의 아미노산의 치환이다(A conservative substitution is the substitution of one amino acid for another with similar characteristics). 보존적 치환은 하기의 그룹 내의 치환을 포함한다 : 발린, 알라닌 및 글리신; 류신, 발린, 및 이소류신; 아스파르트산 및 글루탐산; 아스파라긴 및 글루타민; 세린, 시스테인 및 트레오닌; 리신 및 아르기닌; 및 페닐알라닌 및 티로신. 상기 비-극성 소수성 아미노산은, 알라닌, 류신, 이소류신, 발린, 프롤린, 페닐알라닌, 트립토판 및 메티오닌을 포함한다. 상기 극성 중성 아미노산은 글리신, 세린, 트레오닌, 시스테인, 티로신, 아스파라긴 및 글루타민을 포함한다. 상기 양전하 (염기성) 아미노산[positively charged (basic) amino acids]은, 아르기닌, 리신 및 히스티딘을 포함한다. 상기 음전하 (산성) 아미노산은 아스파르트산 및 글루탐산을 포한한다. 상기 언급된 극성, 염기성 또는 산성 기의 하나의 멤버(member)의 동일한 그룹의 또 다른 멤버로의 어떠한 치환은, 보존적 치환으로 여겨질 수 있다. 이와 대조적으로, 비-보존적 치환은, 하나의 아미노산이 다른 특성을 갖는 또 다른 것으로의 치환이다.

[0088] 몇몇 실시형태에서, GAA 폴리펩티드는, 재조합 단백질(예를 들어, FLAG, 폴리히스티딘(예를 들어, 헥사히스티딘), 헴타글루트타닌(hemagluttanin, HA), 글루타티온-S-전달효소(glutathione-S-transferase, GST), 또는 말토오스-결합 단백질(MBP))의 정제를 위해 사용된 서열과 같은 이종의 아미노산 서열과 융합 단백질일 수 있다.

[0090] 몇몇 실시형태에서, 상기 이종의 아미노산 서열은, 포유동물 세포 내로 상기 분자 복합체의 수송의 효율성을 증진시키는데 사용된다. 예를 들어, 복합체에서 상기 폴리펩티드의 적어도 하나는, 세포외 수용체에 대한 리간드, 타겟 세포의 표면 상에 수용체의 세포외 도메인을 결합하는 타겟팅 도메인(targeting domain), 유로키나아제-타입 플라스미노겐 슈용체, 또는 상기 만노스-6-포스페이트 수용체에 결합하는 인간 인슐린-유사 성장 인자 Ⅱ의 도메인에 융합될 수 있다(예를 들어, 아미노산 1-67 또는 1-87; 적어도 아미노산 48-55; 적어도 아미노산 8-28 및 41-61; 또는 인간 인슐린-유사 성장 인자의 적어도 아미노산 8-87; 인간 인슐린-유사 성장 인자 Ⅱ의 이의 서열 변이체(예를 들어, R68A) 또는 인간 인슐린-유사 성장 인자의 절단된 형태(truncated form)[예를 들어, 위치 62로부터 C-말단 절단됨(C-terminally truncated from position 62)]. 상기 이종의 아미노산 서열은, 상기 폴리펩티드의 N-말단 또는 C-말단에서 융합될 수 있다. 하나의 실시형태에서, 펩티드 태그(peptide tag)는, 스페이서(spacer)(예를 들어, 5-30 아미노산 또는 10-25 아미노산)에 의해 상기 폴리펩티드의 상기 N- 또는 C-말단에 융합된다. 예를 들어, U.S. 특허 No. 7,785,856를 참고하라.

[0092] 이종성 아미노 서열은 또한, 진단 또는 검출할 수 있는 마커, 예를 들어, 루시페라아제(luciferase), 녹색 형광 단백질(green fluorescent protein, GFP), 또는 클로람페니콜 아세틸 전달효소(chloramphenicol acetyl transferase, CAT)로서 유용한 단백질일 수 있다.

[0094] 특정한 숙주 세포(예를 들어, 효모 숙주 세포)에서, 상기 타겟 단백질의 발현 및/또는 분비는, 이종의 단일 서열의 사용을 통해 증가될 수 있다. 몇몇 실시형태에서, 상기 융합 단백질은, 담체(예를 들어, 항체 발생(antibody generation)에 대한 면역 반응을 끌어내는데 유용한 KLH) 또는 소포체 또는 골지체 체류 신호(endoplasmic reticulum or Golgi apparatus retention signals)를 포함할 수 있다. 이종의 서열은 길이가 다양한 것일 수 있고, 몇몇의 경우에서, 상기 이종의 서열(heterologous sequences)이 부착된 전체-길이의 타겟 단백질(full-length target proteins)보다 더 긴 서열일 수 있다.

[0096] *당단백질의 캡핑제거 및 디만노실레이팅, 또는 디만노실레이팅의 방법(Methods of Demannosylating, or Uncapping and Demannosylating Glycoproteins)*

[0097] 인산화된 N-글리칸을 함유하는 당단백질은 디만노실레이트될 수 있고, 만노실-1-포스포-6-만노스 결합 또는 모이어티를 함유하는 인산화된 N-글리칸을 포함하는 당단백질은, (ⅰ) 만노스-1-포스포-6-만노스 결합 또는 모이어티가 만노스-6-포스페이트로 가수분해하고, (ii) 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합 또는 모이어티를 가수분해할 수 있는 만노시다아제와 상기 당단백질을 접촉함으로써 캡핑제거되고 디만노실레이트될 수 있다. 이러한 만노시다이제의 비-제한적인 예는, *카나발리아 엔스폴미스* [잭 빈(Jack bean)] 만노시다아제 및 *야로위야 리폴리카* 만노시다아제(예를 들어, AMS1)를 포함한다. 잭 빈 및 AMS1 만노시다아제 둘 다는 패밀리 38 글리코시사이드 히드롤라제(family 38 glycoside hydrolases)이다.

[0099] 상기 잭 빈 만노시다아제는, 예를 들어, 황산암모늄 현탁액(ammonium sulfate suspension)(Catalog No. M7257) 및 프로테오믹스 등급 조제(proteomics grade preparation)(Catalog No. M5573)로서 Sigma-Aldrich (St. Louis, MO)로부터 상업적으로 입수가능하다. 이러한 상업적인 조제물은, 포스파타제(phosphatases)와 같은 오염 물질을 제거하기 위해, 예를 들어, 겔 여과 크로마토그래피(gel filtration chromatography)에 의해 추가적으로 정제될 수 있다. 상기 잭 빈 만노시다아제는, 그 다음의 아미노산 서열 NKIPRAGWQIDPFGHSAVQG (SEQ ID NO: 11)과 함께 세그먼트를 포함한다(The Jack bean mannosidase contains a segment with the following amino acid sequence NKIPRAGWQIDPFGHSAVQG). Howard *et al.*, *J. Biol. Chem.*, 273(4):2067-2072, 1998를 참고하라.

[0101] 상기 *야로위야 리폴리카* AMS1 만노시다아제는 재조합적으로 생산될 수 있다. 상기 AMS1 폴리펩티드의 상기 아미노산 서열은 SEQ ID NO:5에 기재되어 있다(도 6을 또한 참고하라).

[0103] 몇몇 실시형태에서, 상기 캡핑제거되고 디만노실레이팅 단계는, 두 개의 상이한 효소에 의해 촉매작용된다. 예를 들어, 만노스-1-포스포-6 만노스 결합 또는 모잉어티의 캡핑 제거는 *셀룰로시마이크로비움 셀룰란스*(예를 들어, CcMan5)으로부터 만노시다아제를 사용하여 실행될 수 있다. 상기 CcMan5 폴리펩티드를 코딩하는 상기 뉴클레오티드 서열은, SEQ ID NO : 2에 기재된 것이다(도 2a를 참고하라). 신호 서열을 포함하는 상기 CcMan5 폴리펩티드의 상기 아미노산 서열은 SEQ ID NO: 3에 기재된 것이다(도 2b를 참고하라). 신호서열 없는 상기 CcMan5 폴리펩티드의 상기 아미노산 서열은 SEQ ID NO : 4에 기재되어 있다(도 2c를 참고하라). 몇몇 실시형태에서, 상기 CcMan5 폴리펩티드의 생물학적으로 유효한 단편이 사용된다. 예를 들어, 생물학적으로 유효한 단편은, SEQ ID NO : 4에 기재된 상기 아미노산 서열의 잔기 1-774를 포함할 수 있다. WO 2011/039634를 참고하라. 상기 CcMan5 만노시다아제는 패밀리 92 글리코시사이드 히드롤라제(family 92 glycoside hydrolase)이다.

[0104]

[0105] 당단백질의 캡핑제거된 당단백질 또는 분자 복합체의 디만노실레이션은, *아스페르길루스 사토이*[*아스페르길루스 포에니시스(Aspergillus phoenicis)*로서 또한 알려짐]로부터의 만노시다아제 또는 *셀룰로시마이크로비움 셀룰란스*(예를 들어, CcMan4)로부터의 만노시다아제를 사용하여 촉매작용될 수 있다. 상기 *아스페르길루스 사토이* 만노시다아제는 패밀리 47 글리코시사이드 히드롤라제이고, 상기 CcMan4 만노시다아제는 패밀리 92 글리코시사이드 히드롤라제이다. 상기 *아스페르길루스 사토이* 만노시다아제의 상기 아미노산 서열은, 도 7(SEQ ID NO:6)에 기재되어 있고, GenBank 접근 번호 제BAA08634호이다. 상기 CcMan4 폴리펩티드의 상기 아미노산 서열은 도 8(SEQ ID NO : 7)에 기재되어 있다.

[0107] 당단백질의 분자 복합체 또는 캡핑제거된 당단백질의 디만노실레이션은 또한, *카나발리아 엔스폴미스* (잭 빈) 만노시다이제 또는 *야로위야 리폴리카* 만노시다아제(예를 들어, AMS1)와 같은 패밀리 38 글리코시사이드 히드롤라제로부터의 만노시다아제를 사용하여 촉매작용할 수 있다. 예를 들어, CcMan5는 당단백질(또는 당단백질의 분자 복합체)에서 만노스-1-포스포-6 만노스 모이어티를 캡핑제거하기 위해 사용될 수 있고, 상기 잭 빈 만노시다아제는 상기 캡핑제거된 당단백질(또는 당단백질의 분자 복합체)을 디만노실레이트하기 위해 사용될 수 있다.

[0109] 디만노실레이트된 당단백질(또는 당단백질의 분자 복합체) 또는 캡핑제거되고 디만노실레이트된 당단백질(또는 당단백질의 분자 복합체)을 생산하기 위해, 만노스-1-포스포-6 만노스 결합 또는 모이어티를 포함하는 타겟 분자(또는 분자 복합체)는, 적절하게 재조합적으로 생산된 만노시다제(들)을 포함하는 세포 용해물(cell lysate) 및/또는 적절한 만노시다아제(들)와 적절한 조건 하에서 접촉된다. 적절한 만노시다아제는 상기에 기재된 것이다. 상기 세포 용해물은 균류 세포, 식물 세포, 또는 동물 세포를 포함하는, 어떠한 유전적으로 조작된 세포로부터 일 수 있다. 동물 세포의 비-제한하는 예는, 선충(nematode), 곤충, 식물, 새, 파충류, 및 마우스, 랫, 토끼, 햄스터, 게르빌, 개, 고양이, 염소, 돼지, 소, 말, 고래, 원숭이와 같은 동물, 또는 인간을 포함한다.

[0111] 상기 정제된 만노시다아제 및/또는 세포 용해물과 상기 타겟 분자(예를 들어, 분자 복합체)를 접촉하면서, 상기 만노스-1-포스포-6-만노스 결합 또는 모이어티는 포스포-6-만노스로 가수분해될 수 있고, 글리칸을 포함하는 이러한 포스페이트의 상기 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합 또는 모이어티는, 캡핑제거되고 디만노실레이트된 타겟 분자를 생산하기 위해 가수분해될 수 있다. 몇몇 실시형태에서, 하나의 만노시다아제는, 상기 캡핑제거되거 디만노실레이팅 단계 둘 다를 가수분해하는데 사용된다. 몇몇 실시형태에서, 하나의 만노시다아제는 상기 캡핑제거하는 단계를 촉매작용하는데 사용되고, 상이한 만노시다아제는 상기 디만노실레이팅 단계를 촉매작용하는데 사용된다. 상기 만노시다아제에 의해 처리되는 그 다음에(Following processing by the mannosidase), 상기 타겟 분자 또는 분자 복합체가 분리될 수 있다.

[0113] 상기 용해물에서 상기 만노시다아제의 활성도 또는 온전함을 보존하는 세포 용해물을 수득하기 위한 적절한 방법은, 상기 세포 용해물에서 N-글리코실화 활성도에서 변화를 보존하거나 최소화하는, 뉴클레아제(nuclease), 프로테아제 및 포스파타아제 저해제를 포함하는, 적절한 완충용액 및/또는 저해제의 사용을 포함할 수 있다. 이러한 저해제는, 예를 들어, 에틸렌디아민 테트라아세트산(EDTA), 에틸렌 글리콜 비스(P-아미노에틸 에테르) N,N,N1,N1-테트라아세트산(EGTA)와 같은 킬레이터, 페닐메틸술포닐 플루오르화(phenylmethylsulfonyl fluoride, PMSF), 아프로티닌(aprotinin), 류펩틴(leupeptin), 안티페인(antipain) 등과 같은 프로테아제 저해제, 인산염, 플루오르화 나트륨(sodium fluoride), 바나듐산염 등과 같은, 포스파타아제 저해제를 포함한다. 효소적 활성도를 포함하는 세포 용해물을 수득하기 위한 적절한 완충용액 및 조건은, 예를 들어, Ausubel et al. Current Protocols in Molecular Biology (Supplement 47), John Wiley & Sons, New York (1999); Harlow and Lane, Antibodies: A Laboratory Manual Cold Spring Harbor Laboratory Press (1988); Harlow and Lane, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Press (1999); Tietz Textbook of Clinical Chemistry, 3rd ed. Burtis and Ashwood, eds. W.B. Saunders, Philadelphia, (1999)에 기재되어 있다.

[0115] 세포 용해물은, 적절한 바와 같이, 방해하는 물질(interfering substances)의 존재를 제거하거나 최소화하기 위해 추가적으로 처리될 수 있다. 만약 원한다면, 세포 용해물은, 초원심 세포분획, 및 이온 교환, 소수성 및 역상(reverse phase), 크기 배제(size exclusion), 친화도, 소수성 전하-유도 크로마토그래피 등과 같은 크로마토그래픽 기술을 포함하는, 본 분야에서 널리 알려진 다양한 방법에 의해 분별될 수 있다.

[0117] 몇몇 실시형태에서, 세포 용해물은, 전체 세포의 세포기관이 온전하고/온전하거나 기능적으로 남아있게 제조될 수 있다. 예를 들어, 용해물은, 온전한 조면 소포체, 온전한 활면 소포체, 또는 온전한 골지체의 하나 또는 그 이상을 포함할 수 있다. 온전한 세포의 세포기관을 포함하는 용해물을 제조하고, 상기 세포기관의 기능성을 위한 테스팅을 위한 적절한 방법은, 예를 들어, Moreau *et al.* (1991) *J. Biol. Chem.* 266(7):4329-4333; Moreau *et al.* (1991) *J. Biol. Chem.* 266(7):4322-4328; Rexach *et al.* (1991) *J. Cell Biol.* 114(2):219-229; and Paulik *et al.* (1999) *Arch. Biochem. Biophys.* 367(2):265-273에 기재되어 있다.

[0119] 캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 포함하는 분자 복합체와 포유동물 세포의 접촉시, 상기 분자 복합체는 상기 포유동물 세포(예를 들어, 인간 세포)의 내부로 수송될 수 있다. 캡핑제거되지만, 디만노실레이트되지 않고, 인산화된 N-글리칸을 갖는 분자 복합체는, 포유동물 세포에서 만노스-6-포스페이트 수용체에

실질적으로 결합하지 않고, 보통, 상기 세포의 내부에 효율적으로 수송되지 않는다. 본원에 사용된 바와 같이, "실질적으로 결합하지 않는다(does not substantially bind)"는, 상기 당단백질 분자의 15 % 미만(예를 들어, 14 %, 12 %, 10 %, 8 %, 6 %, 4 %, 2 %, 1 %, 0.5 % 미만, 또는 그 이하, 또는 0 %)가 포유동물 세포에서 만노스-6-포스페이트 수용체에 결합하는 것을 의미한다. 그러나, 만약 이러한 분자 복합체가, 상기 기본적인 만노스가 인산화된 경우에, 말단 알파-1,2 만노스 결합 또는 모이어티를 가수분해할 수 있는 만노시다아제와 접촉한다면, 디만노실레이트된 당단백질은 상기 포유동물 세포에서 상기 만노스-6-포스페이트 수용체에 실질적으로 결합되게 생산되고, 상기 세포의 내부로 효율적으로 수송된다. 본원에 사용된 바와 같은, "실질적으로 결합한다"는, 상기 분자 복합체의 15 % 또는 그 이상(예를 들어, 16 %, 18 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % 또는 95 % 초과)이 포유동물 세포에서 만노스-6-포스페이트 수용체에 결합하는 것을 의미한다. 인산화된 N-글리칸을 디만노실레이트하지 않지만 캡핑제거하는 효소를 포함하는 제조[예를 들어, 재조합 숙주 세포 또는 세포-프리 조제(cell-free preparation)]가 인산화된 N-글리칸을 디만노실레이트하는 효소와 혼성될 수 있는 것으로 이해된다. 이러한 조제와 접촉 후에 타겟 단백질 샘플은, 캡핑만 제거된 몇몇의 인산화된 N-글리칸 및 캡핑제거되고 디만노실레이트된 그 밖의 것들과 함께 단백질 분자를 포함할 수 있다. 캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 함유하는 자연적으로 이러한 단백질 분자는, 만노스-6-포스페이트 수용체에 실질적으로 결합할 수 있다. "실질적으로 결합하지 않는다"의 상기 정의는, 상기 단백질 분자에서 상기 인산화된 N-글리칸이 캡핑제거되지만 디만노실레이트되지 않는 것과 같이 특징될 수 없기 때문에, 이러한 타겟 단백질 샘플에 적용되지 않는다.

[0121] 따라서, 본 문서는, 포유동물 세포 상의 만노스-6-포스페이트 수용체에 결합하지 않는 제1 형태를 포유동물 세포 상의 만노스-6-포스페이트 수용체에 결합하는 제2 형태로의 분자 복합체로 전환하는 방법을 제공한다. 상기 제1 형태에서, 상기 복합체에서 상기 폴리펩티드의 적어도 하나를 포함하는 상기 분자 복합체는, 상기 6 위치에서 인산염 잔기를 포함하는 만노스 잔기로 상기 1 위치에 연결되는 하나 또는 그 이상의 만노스 잔기를 포함하는 하나 또는 그 이상의 N-글리칸을 포함한다. 이러한 방법에서, 상기 분자 복합체의 제1 형태는, 상기 말단 만노스가 되기 위해, 상기 6 위치에서 상기 인산염을 포함하는 상기 만노스를 결과적으로 나타내도록, 상기 말단 만노스 잔기를 디만노실레이트하는 만노시다아제와 접촉한다(the first form of the molecular complex is contacted with a mannosidase that demannosylates the terminal mannose residues to result in the mannose containing the phosphate at the 6 position to become the terminal mannose). 몇몇 실시형태에서, 상기 만노시다아제는 캡핑제거되고 디만노실레이팅 활성도(uncapping and demannosylating activity) 둘 다를 가진다(예를 들어, *카나발리아 엔스폴미스* (잭 빈) 또는 *야로위야 리폴리카* AMS1 만노시다아제). 몇몇 실시형태에서, 상기 만노시다아제는 캡핑제거하는 활성도(uncapping activity)를 가지지 않는다(예를 들어, *아스페르길루스 사토이*로부터의 만노시다아제 또는 *셀룰로시마이크로비움 셀룰란스*로부터의 만노시다아제(예를 들어, CcMan4)).

[0123] 상기 세포의 내부로 당단백질 또는 분자 복합체의 수송은, 세포 흡수 검정(cell uptake assay)을 사용하여 측정될 수 있다. 예를 들어, 캡핑제거되고 디만노실레이트된 인산화된 N-글리칸을 포함하는 분자 복합체 및 포유동물 세포가 배양될 수 있고, 그 다음에 상기 세포는 세척되고 용해된다. 세포 용해물은, 상기 세포 용해물에서 상기 GAA 복합체(예를 들어, 웨스턴블로팅(Western blotting)에 의해)의 존재 또는 GAA의 활성도에 의해 측정될 수 있다. 예를 들어, 흡수(uptake)는 산 알파 글리코시다아제 활성도가 결핍된 섬유아세포(fibroblast)일 수 있다. 알파 글리코시다아제의 세포외 활성도는, 4-메틸움벨릴페릴-알파-D-글루코피라노사이드(4-methylumbelliferyl-alpha-D-glucopyranoside, 4-MUG) 검정을 사용하여 평가될 수 있다. 글루코시다아제에 의해 상기 기질 4-MUG의 분열은, UV 빛으로 조사에 의해 검출되거나 또는 시각화될 수 있는, 상기 플루오리제닉 생산물(fluorigenic product) 4-MU의 발생을 유도한다.

[0125] *폴리펩티드의 재조합 생산*

[0126] 폴리펩티드(예를 들어, 만노시다아제, 알칼리성 프로테아제, 또는 GAA 또는 이의 단편)를 코딩하는 분리된 핵산 분자는 표준 기술에 의해 생산될 수 있다. 상기 용어 "핵산" 및 "폴리뉴클레오티드"는 본원에서 교환하여 사용되고, 핵산 유사체를 포함하는 cDNA, 게놈의 DNA(genomic DNA), 합성의 DNA, 및 DNA (또는 RNA)를 포함하는, RNA 및 DNA 둘 다를 언급한다. 폴리뉴클레오티드는 어떠한 3차 구조(any three-dimensional structure)를 가

질 수 있다. 핵산은 이중-가닥(double-stranded) 또는 단일-가닥(single-stranded)[즉, 센스 가닥(sense strand) 또는 안티센스 가닥(antisense strand)]일 수 있다. 폴리뉴클레오티드의 비-제한적인 예는, 유전자(genes), 유전자 단편(gene fragments), 엑손(exons), 인트론(introns), 메신저 RNA(messenger RNA, mRNA), 운반 RNA(transfer RNA), 리보솜 RNA(ribosomal RNA), siRNA, 마이크로-RNA(micro-RNA), 리보자임(ribozymes), cDNA, 재조합 폴리뉴클레오티드, 가지친 폴리뉴클레오티드(branched polynucleotides), 플라스미드(plasmids), 벡터, 어떠한 서열의 분리된 DNA, 어떠한 서열의 분리된 RNA, 핵산 프로브, 및 프라이머, 뿐만 아니라 핵산 유사체를 포함한다.

[0128] "분리된 핵산"은, 자연적으로-발생하는 게놈(예를 들어, 효모 게놈)에서 상기 핵산의 하나 또는 둘 다의 면의 측면에 있는 핵산을 포함하는, 자연적으로-발생하는 게놈에 존재하는 그 밖의 핵산 분자에서 분리되는 핵산을 나타낸다. 핵산에 대하여 본원에 사용된 바와 같은 상기 용어 "분리된"은, 이러한 비-자연적으로-발생하는 서열이 천연에서 발생하지 않고, 자연적으로-발생하는 게놈에서 바로 인접한 서열을 가지지 않도록, 어떠한 비-자연적으로-발생하는 핵산 서열을 포함한다(The term "isolated" as used herein with respect to nucleic acids also includes any non-naturally-occurring nucleic acid sequence, since such non-naturally-occurring sequences are not found in nature and do not have immediately contiguous sequences in a naturally-occurring genome).

[0130] 분리된 핵산은, 예를 들어, 자연적으로-발생하는 게놈에서 DNA 분자가 제거되거나 존재하지 않는 바로 옆에 배치되어 일반적으로 발견된, 핵산 서열 중 하나로 제공된, 예를 들어, DNA 분자일 수 있다(An isolated nucleic acid can be, for example, a DNA molecule, provided one of the nucleic acid sequences normally found immediately flanking that DNA molecule in a naturally-occurring genome is removed or absent). 따라서, 분리된 핵산은, 벡터, 자기복제 플라스미드(autonomously replicating plasmid), 바이러스(예를 들어, 파라믹소바이러스, 레트로바이러스, 렌티바이러스, 아데노바이러스, 또는 헤르페스(herpes) 바이러스), 또는 핵생물 또는 진핵생물의 게놈 DNA 내로 포함되는 DNA 뿐만 아니라, 그 밖의 서열과 독립한 분리 분자(예를 들어, 화학적으로 합성된 핵산, 또는 PCR에 의해 생성된 게놈 DNA 단편 또는 제한 효소 처리)로서 존재하는 DNA 분자를 포함하지만 이로 제한하지 않는다. 게다가, 분리된 핵산은, 하이브리드 또는 융합 핵산(hybrid or fusion nucleic acid)의 일부인 DNA 분자와 같은 제작된 핵산(engineered nucleic acid)을 포함할 수 있다. 예를 들어, cDNA 라이브러리 또는 게놈 라이브러리, 또는 게놈 DNA 제한 절단(genomic DNA restriction digest)을 포함하는 겔 슬라이스(gel slices) 내에 수백 내지 수백만의 그 밖의 핵산 중에 존재하는 핵산은, 분리된 핵산을 고려하지 않는다.

[0132] 핵산 및 특정한 숙주 세포와 관련하여 본원에 사용된 바와 같은 용어 "외인성(exogenous)"은, 자연에서 발견된 바와 같은 그러한 특정한 세포에서 발생하지 않는 (및 그러한 특정한 세포로부터 수득될 수 없는) 어떠한 핵산을 나타낸다. 따라서, 비-자연적으로-발생하는 핵산은, 상기 숙주 세포 내로 한 번 도입된 숙주 세포에 대한 외인성(exogenous)이 되도록 고려된다. 비-자연적으로-발생하는 핵산은, 전체로서 핵산이 자연에서 존재하지 않는다면, 자연에서 발견되는 핵산 서열의 단편 또는 핵산 서브시퀀스(subsequence)를 포함할 수 있음을 주목하는 것은 중요하다. 예를 들어, 발현 벡터 내의 게놈의 DNA 서열을 포함하는 핵산 분자는 비-자연적으로-발생하는 핵산이고, 따라서, 완전체로서 그러한 핵산 분자(게놈의 DNA 플러스 벡터 DNA)가 자연에 존재하지 않기 때문에, 상기 숙주 세포 내로 한번 도입되는 숙주 세포에 대해 외인성이다. 따라서, 전체로서 자연에 존재하지 않는 어떠한 벡터, 자기 복제 플라스미드, 또는 바이러스(예를 들어, 레트로바이러스, 아데노바이러스, 또는 헤르페스 바이러스)는, 비-자연적으로-발생하는 핵산일 수 있도록 고려된다. PCR에 의해 생산된 게놈 DNA 단편 또는 제한 효소 처리 뿐만 아니라 cDNAs가 이들이 자연에서 발견되지 않는 분리된 분자로서 존재하기 때문에, 비-자연적으로-발생하는 핵산이기 위해 고려되는 결과가 나온다(It follows that genomic DNA fragments produced by PCR or restriction endonuclease treatment as well as cDNAs are considered to be non-naturally-occurring nucleic acid since they exist as separate molecules not found in nature). 자연에서 발견되지 않는 배열에서 프로모터 서열 및 폴리펩티드-코딩하는 서열(예를 들어, cDNA 또는 게놈 DNA)을 포함하는 어떠한 핵산이 비-자연적으로-발생하는 핵산임을 또한 결과로 나온다(It also follows that any nucleic acid containing a promoter sequence and polypeptide-encoding sequence (e.g., cDNA or genomic DNA) in an

arrangement not found in nature is non-naturally-occurring nucleic acid). 자연적으로-발생하는 핵산은 특정한 세포에 대해 외인성일 수 있다. 예를 들어, 효모 x로부터 분리된 전체 염색체는, 염색체가 효모 y의 세포 내로 도입된다면, 효모 y의 세포에 대해 외인성 핵산이다.

[0134] 중합효소 연쇄 반응(PCR) 기술은 본원에 기재된 뉴클레오티드 서열을 포함하는 분리된 핵산을 수득하기 위해 사용될 수 있다. PCR은 전체 게놈의 DNA 또는 전체 세포의 RNA로부터 서열을 포함하는, DNA 뿐만 아니라 RNA로부터 특정한 서열을 증폭시키기 위해 사용될 수 있다. 일반적으로, 관심 또는 이후의 영역의 말단으로부터의 서열 정보는, 증폭시키기 위해 상기 템플레이트의 반대 가닥에 대한 서열과 동일하거나 또는 유사한 올리고뉴클레이티드 프라이머(oligonucleotide primers)를 설계하기 위해 사용된다. 다양한 PCR 계획은 또한, 부위-특이적인 뉴클레오티드 서열 변형이 템플레이트 핵산 내로 도입될 수 있는 것과 같이 또한 이용할 수 있다. 분리된 핵산은 또한, 단일 핵산 분자[예를 들어, 포스포라미디트 기술(phosphoramidite technology)을 사용하여 상기 3' 내지 5' 방향에서 자동화된 DNA 합성을 사용하여] 또는 일련의 올리고뉴클레오티드로서, 화학적으로 합성될 수 있다. 예를 들어, 긴 올리코뉴클레오티드(예를 들어, > 100 뉴클레오티드)의 하나 또는 그 이상의 쌍(pair)은, 상기 올리고뉴클레오티드 쌍이 어닐링된 경우에(when the oligonucleotide pair is annealed), 듀플렉스(duplex)가 형성되는 것과 같이, 상보성의 짧은 세그먼트(예를 들어, 약 15 개의 뉴클레오티드)를 포함하는 각각의 쌍과 함께, 원하는 서열을 포함하여 합성될 수 있다. DNA 중합효소는 올리고뉴클레오티드 쌍 당 단일, 이중-가닥 핵산 분자를 결과적으로 나타내는, 올리고뉴클레오티를 연장하는데 사용되고, 그리고 난 다음에 벡터 내로 결합될 수 있다. 분리된 핵산은 또한, 예를 들어, 자연적으로 발생하는 DNA의 돌연변이유발(mutagenesis)에 의해 수득될 수 있다.

[0136] 폴리펩티드(예를 들어, 만노시다아제, 알칼리성 프로테아제, 또는 GAA 또는 이의 단편)를 재조합적으로 생산하기 위해, 벡터는 상기 폴리펩티드를 코딩하는 핵산에 작동가능하게 연결된 프로모터를 포함하여 사용된다(a vector is used that contains a promoter operably linked to nucleic acid encoding the polypeptide). 본원에 기재된 바와 같이, "프로모터"는 유전자가 번역될 수 있도록 DNA 서열을 나타낸다. 상기 프로모터는 DNA 중합효소에 의해 인식되고, 그리고 난 다음에 전사(transcription)를 개시한다. 따라서, 프로모터는 RNA 중합효소의, 보충(recruitment)에서 포함되거나 또는 직접적으로 결합하는 DNA 서열을 포함한다(a promoter contains a DNA sequence that is either bound directly by, or is involved in the recruitment, of RNA polymerase). 프로모터 서열은, 유전자-클러스터(gene-cluster)에서 유전자의 전사 레벨을 증진시키기 위해(이런 이유로 명명됨), 단백질[즉, 전사 인자의 세트와 같이, 트랜스-작용 인자(trans-acting factor)]과 결합할 수 있는 DNA의 하나 또는 그 이상의 영역인 "인핸서 영역(enhancer regions)"을 또한 포함할 수 있다. 코딩 영역의 상기 5' 말단에 일반적인, 상기 인핸서는 프로모터 서열로부터 또한 분리될 수 있고, 예를 들어, 상기 유전자의 상기 코딩 영역에 대해 3' 또는 유전자의 인트론 영역(intronic region) 내에 일 수 있다.

[0138] 본원에 사용된 바와 같이, "작동가능하게 연결된(operably linked)"은, 발현 조절 서열(expression control sequences)이 관심의 코딩 서열의 발현을 효과적으로 조절하도록, 유전적 구성(예를 들어, 벡터) 내로 포함됨을 의미한다.

[0140] 발현 벡터는, 상기 코딩된 폴리펩티드의 발현을 위해 숙주 세포 내로 도입될 수 있고[예를 들어, 형질전환(transformation) 또는 형질주입(transfection)에 의해], 그 다음에 정제될 수 있다. 폴리펩티드의 작은 또는 큰 규모의 생산에 사용할 수 있는 발현 시스템(예를 들어, 만노시다아제, 알칼리성 프로테아제, 또는 GAA 또는 이의 단편)은, 재조합 박테리오파지 DNA, 플라스미드 DNA, 또는 상기 핵산 분자를 포함하는 코스미드 DNA 발현 벡터(cosmid DNA expression vectors)로 형질전환된 박테리아(예를 들어, *E. coli*)와 같은 미생물, 및 상기 핵산 분자를 포함하는 재조합 균류 발현 벡터로 형질전환된 균류[예를 들어, *야로위야 리폴리카, 아르슐라 아데니닌보란스(Arxula adeninivorans), 피치아 파스토리스(Pichia pastoris), 한세눌라 폴리모르파(Hansenula polymorpha), 오가테아 미누타(Ogataea minuta), 피치아 메탄올리카(Pichia methanolica), 아스페르길루스 니게르(Aspergillus niger), 트리코데르마 레세이(Trichoderma reesei), 및 사카로미세스 세레비시아(Saccharomyces cerevisiae)*]를 포함하지만, 이로 제한되지 않는다. 유용한 발현 시스템은 또한, 핵산 분자를

포함하는 재조합 바이러스 발현 벡터(예를 들어, 바큐로바이러스)로 감염된 곤충 세포 시스템, 및 상기 핵산 분자를 포함하는 재조합 플라스미드 발현 벡터(예를 들어, Ti 플라스미드)로 형질전환되거나 재조합 바이러스 발현 벡터(예를 들어, 담배 모자이크 바이러스)로 감염된 식물 세포 시스템을 또한 포함한다. 만노시다아제 또는 알칼리성 프로테아제 폴리펩티드는 또한, 포유동물 세포의 게놈으로부터 유도된 프로모터[예를 들어, 메틸로티오네인 프로모터(metallothionein promoter)] 또는 포유동물 바이러스로부터 유도된 프로모터[예를 들어, 아데노바이러스 후기 프로모터(adenovirus late promoter) 및 거대세포바이러스 프로모터(cytomegalovirus promoter)]를 포함하는 재조합 발현 구성물(constructs)을 포함하는, 세포[예를 들어, COS 세포, 중국 햄스터 난소 세포, 헬라 세포, 인간 배아 신장 293 세포, 및 3T3 L1 세포와 같은 불멸의 세포주(immortalized cell lines)]를 포함하는, 포유동물 발현 시스템을 사용하여 생산될 수 있다.

[0142] 만노시다아제와 같은 재조합 폴리펩티드는, 상기 단백질을 정제하는데 도움을 주기 위해 FLAG, 폴리히스티딘(예를 들어, 헥사히스티딘), 헤마글루트타닌(hemagluttanin, HA), 글루타티온-S-전이효소(glutathione-S-transferase, GST), 또는 말토스-결합 단백질(MBP)와 같은 이종의 아미노산 서열과 태그될 수 있다(be tagged with). 단백질을 정제하기 위한 그 밖의 방법은, 이온 교환, 소수성 및 역상, 크기 배제(size exclusion), 친화도(affinity),소수성 전하-유도 크로마토그래피(hydrophobic charge-induction chromatography) 등과 같은 크로마토그래픽 기술을 포함한다(예를 들어, Scopes, Protein Purification: Principles and Practice, third edition, Springer-Verlag, New York (1993); Burton and Harding, *J. Chromatogr. A* 814:71-81 (1998)를 참고하라).

[0144] *당단백질의 캡핑제거 및 디만노실레이팅의 생체 내 방법*

[0145] 본원에 기재된 유전적으로 조작된 세포는 GAA 활성도를 갖는 분자 복합체를 생산하기 위해 사용될 수 있다. 예를 들어, 유전적으로 조작된 세포는 적어도 2 개의 폴리펩티드를 포함하고 GAA 활성도를 갖는 분자 복합체를 생산하기 위해 사용될 수 있고, 각각의 폴리펩티드는, SEQ ID NO: 1에 기재된 상기 아미노산 서열의 세그먼트와 적어도 85 % 서열 동일성을 갖고, 각각의 세그먼트는, 아미노산 50과 아미노산 74 사이의 하나 또는 그 이상의 부위(예를 들어, 아미노산 56과 아미노산 68 사이 또는 아미노산 60과 아미노산 65 사이)에서 SEQ ID NO: 1에 기재된 상기 아미노산 서열의 단백질분해에 의해 유도된다. 예를 들어, 균류 세포는, SEQ ID NO : 1 에 기재된 상기 아미노산 서열을 코딩하는 핵산 및 SEQ ID NO:8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 알칼리성 프로테아제를 코딩하는 핵산을 포함하도록 조작될 수 있고, 코딩된 폴리펩티드의 각각은 상기 배양 배지 내로 분비되도록 하고(such that each of the encoded polypeptides are secreted into the culture medium), 상기 알칼리성 프로테아제는 SEQ ID NO : 1에 기재된 상기 아미노산 서열을 분열할 수 있다. 실시예 12에 기재된 바와 같이, 상기 재조합 GAA가 상기 알칼리성 프로테아제와 함께 상기 배양 배지 내로 분비된 경우에, 상기 110 kDa 전구체를 상기 95 kDa 형태로 처리하는 것이 완전하고, 즉 110 kDa 전구체가 검출되지 않았다.

[0146]

[0147] 본원에 기재된 유전적으로 조작된 세포는 또한 캡핑제거되고 디만노실레이트된 분자 복합체를 생산하기 위해 사용될 수 있다. 이러한 유전적으로 조작된 세포는, SEQ ID NO : 1에 기재된 상기 아미노산 서열을 갖는 폴리펩티드를 코딩하는 핵산, 본원에 기재된 바와 같은 만노시다아제를 코딩하는 핵산, 및 임의적으로 SEQ ID NO : 8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 알칼리성 프로테아제를 코딩하는 핵산을 포함할 수 있다.

[0149] 캡핑제거되고 디만노실레이트된 분자 복합체를 생산하는 세포 기저 방법은,만노스-1-포스포-6-만노스 결합 또는 모이어티를 포스포-6-만노스로 가수분해할 수 있는 만노시다아제를 코딩하는 핵산, SEQ ID NO: 1에 기재된 상기 아미노산 서열을 갖는 폴리펩티드를 코딩하는 핵산 및 임의적으로 SEQ ID NO:8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 알칼리성 프로테아제를 코딩하는 핵산을 포함하도록, 유전적으로 조작된 균류 세포 내로 도입시키는 것을 포함할 수 있고, 상기 세포는 캡핑제거된 인산화된 N-글리칸을 포함하는 본원 기재된 상기 분자 복합체를 생산한다. 이러한 인산화된 N-글리칸은 상기에 기재된 바와 같이 디만노실레이트될 수 있다. 몇몇 실시형태에서, 상기 만노시다아제 및 GAA를 코딩하는 상기 핵산은 상기 만노시다아제 및 GAA가 공

동-분비되도록 분비 서열을 포함한다. 알칼리성 프로테아제를 코딩하는 핵산을 포함하는 유전적으로 조작된 세포에서, 상기 분자 복합체는 95 kDa 형태로 처리될 수 있다.

[0151] 또 다른 세포 기저 방법(cell based method)은, (ⅰ) 만노스-1-포스포-6-만노스 결합 또는 모이어티가 포스포-6-만노스로 가수분해하고, (ii) 글리칸을 포함하는 포스페이트의 말단 알파-1,2 만노스, 알파-1,3 만노스 및/또는 알파-1,6 만노스 결합 또는 모이어티를 가수분해할 수 있는, 만노시다아제를 코딩하는 핵산, SEQ ID NO: 1에 기재된 상기 아미노산 서열을 갖는 폴리펩티드를 코딩하는 핵산, 및 임의적으로 SEQ ID NO:8에 기재된 상기 아미노산 서열과 적어도 85 % 서열 동일성을 갖는 알칼리성 프로테아제를 코딩하는 핵산을 포함하도록 유전적으로 조작된 균류 세포 내로 도입시키는 단계를 포함할 수 있고, 상기 세포는 캡핑제거되고 디만노실레이트된 분자 복합체를 생산한다. 몇몇 실시형태에서, 상기 만노시다아제 및 GAA를 코딩하는 상기 핵산은, 상기 만노시다아제 및 타겟 분자가 공동-분비되도록 분비 서열을 포함한다. 알칼리성 프로테아제를 코딩하는 핵산을 포함하는 유전적으로 조작된 세포에서, 상기 분자 복합체는 95 kDa 형태로 처리될 수 있다.

[0153] 타겟 분자 또는 분자 복합체의 생체내 생산에 적합한 세포는, *야로위아 리폴리카*, *아르슐라 아데니닌보란스*를 포함하는, 균류 기원(fungal origin), 메탄영양 효모[상기 속 *칸디다(Candida)*, *한세눌라(Hansenula)*, *우가타에(Oogataea)*, *피치아(Pichia)* 또는 *토룰롭시스(Torulopsis)*와 같은 메탄영양효모(methylotrophic yeast)], 속 *아스페르길루스(Aspergillus)*, *트리코데르마(Trichoderma)*, *뉴로스포라(Neurospora)*, *푸사리움(Fusarium)*, 또는 *크리소스포리움(Chrysosporium)*의 곰팡이(filamentous fungi)일 수 있다. 대표적인 균류 종은, *피치아 아노말라(Pichia anomala)*, *피치아 보비스(Pichia bovis)*, *피치아 카나데니시스(Pichia canadensis)*, *피차아 카르소니(Pichia carsonii)*, *피치아 파리노스(Pichia farinose)*, *피치아 페르멘탄스(Pichia fermentans)*, *피치아 플루시움(Pichia fluxuum)*, *피치아 멤브레인파시엔스(Pichia membranaefaciens)*, *피치아 멤브레인파시엔스(Pichia membranaefaciens)*, *칸디다 발리다(Candida valida)*, *칸디다 알비칸스(Candida albicans)*, *칸디다 아스칼라피다룸(Candida ascalaphidarum)*, *칸디다 암피시아에(Candida amphixiae)*, *칸디다 안타르크티카(Candida Antarctica)*, *칸디다 아틀란티카(Candida atlantica)*, *칸디다 아트모스페리카(Candida atmosphaerica)*, *칸디다 블라트테(Candida blattae)*, *칸디다 카르포필라(Candida carpophila)*, *칸디다 세람비시다룸(Candida cerambycidarum)*, *칸디다 차울리오데스(Candida chauliodes)*, *칸디다 코리달리스(Candida corydalis)*, *칸디다 도세일(Candida dosseyi)*, *칸디다 두블리니엔시스(Candida dubliniensis)*, *칸디다 에르가테니시스(Candida ergatensis)*, *칸디다 프룩투스(Candida fructus)*, *칸디다 칼라브라타(Candida glabrata)*, *칸디다 페르만타티(Candida fermentati)*, *칸디다 구일리에르몬디(Candida guilliermondii)*, *칸디다 헤물로니(Candida haemulonii)*, *칸디다 인섹타멘스(Candida insectamens)*, *칸디다 인섹토룸(Candida insectorum)*, *칸디다 인터메디아(Candida intermedia)*, *칸디다 제프레시(Candida jeffresii)*, *칸디다 케피르(Candida kefyr)*, *칸디다 크루세이(Candida krusei)*, *칸디다 루시타니아(Candida lusitaniae)*, *칸디다 직소오소필라(Candida lyxosophila)*, *칸디다 말토사(Candida maltosa)*, *칸디다 멤브라니파시엔스(Candida membranifaciens)*, *칸디다 밀레리(Candida milleri)*, *칸디다 올레오필라(Candida oleophila)*, *칸디다 오레고네시스(Candida oregonensis)*, *칸디다 파라피실로시스(Candida parapsilosis)*, *칸디다 퀘르시트루사(Candida quercitrusa)*, *칸디다 쉬하테아(Candida shehatea)*, *칸디다 트레모칠라에(Candida temnochilae)*, *칸디다 테누이스(Candida tenuis)*, *칸디다 트로피칼리스(Candida tropicalis)*, *칸디다 트스치야(Candida tsuchiyae)*, *칸디다 시놀라보란티움(Candida sinolaborantium)*, *칸디다 소자에(Candida sojae)*, *칸디다 비스와나티(Candida viswanathii)*, *칸디다 우틸리스(Candida utilis)*, *우가타에 미누타(Oogataea minuta)*, *피치아 멤브레인파시엔스(Pichia membranaefaciens)*, *피치아 실베스트리스(Pichia silvestris)*, *피치아 멤브레인파시엔스(Pichia membranaefaciens)*, *피치아 초다티(Pichia chodati)*, *피치아 멤브레인파시엔스(Pichia membranaefaciens)*, *피치아 멤브레인파시엔스(Pichia menbranaefaciens)*, *피치아 미누스큘레(Pichia minuscule)*, *피치아 파스토리스(Pichia pastoris)*, *피치아 프세우도폴리모르파(Pichia pseudopolymorpha)*, *피치아 퀘르쿰(Pichia quercuum)*, *피치아 로베르트시(Pichia robertsii)*, *피치아 사이토이(Pichia saitoi)*, *피치아 실베스트리시(Pichia silvestrisi)*, *피치아 스트라버겐시스(Pichia strasburgensis)*, *피치아 테르리콜라(Pichia terricola)*, *피치아 반리지(Pichia vanriji)*, *슈도지마 안타르시티카(Pseudozyma Antarctica)*, *로도스포리디움 토룰로이데스(Rhodosporidium toruloides)*, *로도토룰라 글루티니스(Rhodotorula glutinis)*, *사카로마이세스 바야뉴스(Saccharomyces bayanus)*, *사카로마이세스 바야뉴스(Saccharomyces bayanus)*, *사카로마이세스 몸드쉬리쿠스(Saccharomyces momdshuricus)*, *사카로마이세스 우바룸(Saccharomyces uvarum)*, *사카로마이세스 바야뉴스*

(Saccharomyces bayanus), 사카로마이세스 세레비시아(Saccharomyces cerevisiae), 사카로마이세스 비스포루스(Saccharomyces bisporus), 사카로마이세스 시바리에리(Saccharomyces chevalieri), 사카로마이세스 델브루에키(Saccharomyces delbrueckii), 사카로마이세스 엑시구오우스(Saccharomyces exiguous), 사카로마이세스 페르만타티(Saccharomyces fermentati), 사카로마이세스 프라길리스(Saccharomyces fragilis), 사카로마이세스 마르시아누스(Saccharomyces marxianus), 사카로마이세스 말리스(Saccharomyces mellis), 사카로마이세스 로세이(Saccharomyces rosei), 사카로마이세스 루시(Saccharomyces rouxii), 사카로마이세스 우바륨(Saccharomyces uvarum), 사카로마이세스 윌리아우스(Saccharomyces willianus), 사카로마이코데스 루드위기(Saccharomycodes ludwigii), 사카로마이코프시스 카프슐라리스(Saccharomycopsis capsularis), 사카로마이코프시스 피불리게라(Saccharomycopsis fibuligera), 사카로마이코프시스 피불리게라(Saccharomycopsis fibuligera), 엔도마이세스 호르데이(Endomyces hordei), 엔도마이코프시스 포블리게라(Endomycopsis fobuligera). 사튜르니스포라 사이토이(Saturnispora saitoi), 스키조사카로마이세스 옥토스포루스(Schizosaccharomyces octosporus), 스키조카로마이세스 폼베(Schizosaccharomyces pombe), 쉬반니오마세스 옥시덴탈리스(Schwanniomyces occidentalis), 토룰라스포라 델브루엑키(Torulaspora delbrueckii), 토룰라스포라 델브루엑키(Torulaspora delbrueckii), 사카로마이세스 다이레니시스(Saccharomyces dairensis), 토룰라스포라 델브루엑키(Torulaspora delbrueckii), 토룰라스포라 페르만타티(Torulaspora fermentati), 사카로마이세스 페르만타티(Saccharomyces fermentati), 토룰라스포라 델브루엑키(Torulaspora delbrueckii), 토룰라스포라 로세이(Torulaspora rosei), 사카로마이세스 로세이(Saccharomyces rosei), 토룰라스포라 델브루엑키(Torulaspora delbrueckii), 사카로마이세스 로세이(Saccharomyces rosei), 토룰라스포라 델브루엑키(Torulaspora delbrueckii), 사카로마이세스 델브룩엑키(Saccharomyces delbrueckii), 토룰라스포라 델브룩엑키(Torulaspora delbrueckii), 사카로마이세스 델브룩엑키(Saccharomyces delbrueckii), 자이고사카로마이세스 몬골리쿠스(Zygosaccharomyces mongolicus), 도룰라스포라 글로보사(Dorulaspora globosa), 데바르오마이세스 글로보수스(Debaryomyces globosus), 토룰로프시스 글로보사(Torulopsis globosa), 트리콜로스포론 쿠타니움(Trichosporon cutaneum), 트리고노프시스 바리아빌리스(Trigonopsis variabilis), 윌리오프시스 칼리포르니카(Williopsis californica), 윌리오프시스 사투르누스(Williopsis saturnus), 자이고사카로마이세스 비스포루스(Zygosaccharomyces bisporus), 자이고사카로이마이세스 비스포루스(Zygosaccharomyces bisporus), 데바르오마이세스 디스포루아(Debaryomyces disporua). 사카로마이세스 디스포라스(Saccharomyces bisporas), 자이고사카로마이세스 비스포루스(Zygosaccharomyces bisporus), 사카로마이세스 비스포루스(Saccharomyces bisporus), 자이고사카로마이세스 멜리스(Zygosaccharomyces mellis), 자이고사카로마아세스 피이오리아누스(Zygosaccharomyces priorianus), 자이고사카로마이세스 로우심(Zygosaccharomyces rouxiim), 자이고사카로마이세스 로우심(Zygosaccharomyces rouxii), 자이고사카로마이세스 바르케리(Zygosaccharomyces barkeri), 사카로마이세스 로우심(Saccharomyces rouxii), 자이고사카로마이세스 로우심(Zygosaccharomyces rouxii), 자이고사카로마이에스 마조르(Zygosaccharomyces major), 사카로마이세스 로우시(Saccharomyces rousii), 피치아 아노말라(Pichia anomala), 피치아 보비스(Pichia bovis), 피치아 칸나덴시스(Pichia Canadensis), 피치아 카르소니(Pichia carsonii), 피치아 파르노세(Pichia farinose), 피치아 페르만탄스(Pichia fermentans), 피치아 플루시움(Pichia fluxuum), 피치아 멤브레인파시엔스(Pichia membranaefaciens), 피치아 슈도폴리모르파(Pichia pseudopolymorpha), 피치아 퀘르큠(Pichia quercuum), 피치아 로베르트시(Pichia robertsii), 슈도자마 안타르티카(Pseudozyma Antarctica), 로도스포리디움 토룰로이데스(Rhodosporidium toruloides), 로도스포리디움 토룰로이데스(Rhodosporidium toruloides), 로도스토룰라 글루티니스(Rhodotorula glutinis), 사카로마이세스 바야누스(Saccharomyces bayanus), 사카로마이세스 바야누스(Saccharomyces bayanus), 사카로마이세스 비스포루스(Saccharomyces bisporus), 사카로마이세스 세레비시아(Saccharomyces cerevisiae), 사카로마이세스 시바리에리(Saccharomyces chevalieri), 사카로마이에스 델브룩엑키(Saccharomyces delbrueckii), 사카로마이세스 프로만타티(Saccharomyces fermentati), 사카로마이세스 프라질러스(Saccharomyces fragilis), 사카로마이세스 루드위기(Saccharomycodes ludwigii), 스키조사카로마이세스 폼베(Schizosaccharomyces pombe), 쉬반니오마이세스 옥시덴탈리스(Schwanniomyces occidentalis), 토룰라스포라 델브룩엑키(Torulaspora delbrueckii), 토룰라스포랄 글로보사(Torulaspora globosa), 트리고노프시스 바리아빌리스(Trigonopsis variabilis), 윌리오프시스 칼리포르니카(Williopsis californica), 윌리오프시스 사투르누스(Williopsis saturnus), 자이고사카로마이세스 비스포루스(Zygosaccharomyces bisporus), 자이고사카로마이세스 멜리스(Zygosaccharomyces mellis), 또는 자이고사카로마이세스 루드위기(Zygosaccharomyces rouxii)를 포함하지만 이로 제한되지 않는다. 대표적인 곰팡이(filamentous fungi)는, 다양한 종류의 아스페르길루스(Aspergillus), 아스페르길루스 카시엘루스(Aspergillus caesiellus), 아스페르길루스 칸디두스(Aspergillus candidus), 아스페르길루스 카르네우스(Aspergillus carneus), 아스페르길루스 클라바투스(Aspergillus clavatus), 아스페르길루스 데프렉투스(Aspergillus

deflectus), *아스페르길루스 플라부스(Aspergillus flavus)*, *아스페르길루스 푸미가투스(Aspergillus fumigatus)*, *아스페르길루스 글루우쿠스(Aspergillus glaucus)*, *아스페르길루스 니둘란스(Aspergillus nidulans)*, *아스페르길루스 니게르(Aspergillus niger)*, *아스페르길루스 오크라세우스(Aspergillus ochraceus)*, *아스페르길루스 오리제(Aspergillus oryzae)*, *아스페르길루스 파라시티커스(Aspergillus parasiticus)*, *아스페르길루스 페니실로이데스(Aspergillus penicilloides)*, *아스페르길루스 레스트릭투스(Aspergillus restrictus)*, *아스페르길루스 소재(Aspergillus sojae)*, *아스페르길루스 시도위(Aspergillus sydowi)*, *아스페르길루스 타마리(Aspergillus tamari)*, *아스페르길루스 테르레우스(Aspergillus terreus)*, *아스페르길루스 우스투스(Aspergillus ustus)*, 또는 아스페르길루스 베르시콜로르(*Aspergillus versicolor*)를 포함하지만 이로 제한되지 않는다. 본원에 특정된 바와 같이 유전적으로 조작되기 전에 이러한 세포는, 예를 들어, American Type Culture Collection(Rockville, MD)와 같은, 다양한 상업적인 공급자 및 연구 자원 시설로부터 수득될 수 있다.

[0155] 세포의 유전적인 조작은, 만노시다아제, GAA, 및/또는 알칼리성 프로테아제를 코딩하는 외인성 핵산을 더하여, 하기의 하나 또는 그 이상의 유전적인 변형을 포함할 수 있다: (i) 외부의 사슬 신장(Outer CHain elongation, OCH1) 단백질을 코딩하는 내인성 유전자의 삭제; (ii) 만노스 잔기의 인산화를 증가시키기 위해, 만노실 인산화를 촉진시킬 수 있는 폴리펩티드[예를 들어,*야로위야 리폴리카*, *S. 세레비시아(S. cerevisiae)*, *오가태아 미누타(Ogataea minuta)*, *피치아 파스토리스(Pichia pastoris)*, 또는 *C.알비칸스(C. albicans)*로부터의 MNN4 폴리펩티드, 또는 *P. 파스토리스*로부터 PNO1 폴리펩티드]를 코딩하는 재조합 핵산의 도입; (iii) CH1 단백질의 기능적인 발현을 방해하는 RNA 분자의 도입 또는 발현; (iv) N-글리코실화 활성도를 갖는(즉, N-글리코실화 활성도를 갖는 단백질을 발현하는) 야생형(예를 들어, 내인성 또는 외인성) 단백질을 코딩하는 재조합 핵산의 도입; 또는 (v) N-글리코실화 활성도를 갖는 단백질을 코딩하는 하나 또는 그 이상의 내인성 유전자의 프로모터 또는 인핸서를 변경하여, 이의 코딩된 단백질의 발현을 변경하는 것. RNA 분자는, 예를 들어, 짧은-간섭 RNA(small-interfering RNA, siRNA), 짧은 헤어핀 RNA(short hairpin RNA, shRNA), 안티-센스 RNA, 또는 마이크로 RNA(miRNA)를 포함한다. 유전적으로 조작하는 것은 또한, 첨가(예를 들어, 이종의 서열), 삭제, 또는 치환(예를 들어, 점 돌연변이. 보전 또는 비-보존 돌연변이와 같은 돌연변이)을 갖는 단백질을 생산하기 위해 N-글리코실화 활성을 갖는 단백질을 코딩하는 내인성 유전자를 변경하는 것을 포함한다. 돌연변이는, 특이적으로 도입될 수 있거나[예를 들어, 위치-지정된 돌연변이(site-directed mutagenesis) 또는 상동 재조합(homologous recombination)에 의해], 또는 임의적으로 도입될 수도 있다(예를 들어, 세포는 예를 들어, Newman and Ferro-Novick (1987) *J. Cell Biol.* 105(4):1587에 기재된 바와 같이, 화학적으로 돌연변이를 일으킬 수 있다).

[0157] 본원에 기재된 유전적인 변형은, (i) 상기 유전적으로 변형된 세포에서 하나 또는 그 이상의 활성도에서의 증가, (ii) 상기 유전적으로 변형된 세포에서 하나 또는 그 이상의 활성도에서의 감소, 또는 (iii) 상기 유전적으로 변형된 세포에서 하나 또는 그 이상의 활성도의 로컬라이제이션(localization) 또는 세포내 분포에서의 변화의 하나 또는 그 이상을 결과적으로 발생할 수 있다. 특정한 활성도의 양에서의 증가(예를 들어, 만노실 인산화의 촉진)은, 만노실 인산화를 촉진할 수 있는 하나 또는 그 이상의 단백질을 과잉발현, 내인성 유전자의 복제수(예를 들어, 유전자 중복)에서의 증가, 또는 상기 유전자에 의해 코딩된 상기 단백질의 발현에서의 증가를 촉진하는 내인성 유전자의 상기 프로모터 또는 인핸서에서의 변형 때문임을 이해할 수 있다. 하나 또는 그 이상의 특정한 활성도에서의 감소는, 돌연변이 형태[예를 들어, 우성 음성적인 형태(dominant negative form)]의 과잉발현, 특정한 활성도를 갖는 하나 또는 그 이상의 단백질의 발현을 감소시키는 하나 또는 그 이상의 방해하는 RNA 분자의 도입 또는 발현, 또는 상기 특정한 활성도를 갖는 단백질을 코딩하는 하나 또는 그 이상의 내인성 유전자의 삭제 때문일 수 있다.

[0159] 상동 재조합에 의해 유전자를 분열시키기 위해, "유전자 대체(gene replacement)" 벡터는 선택가능한 마커 유전자를 포함하도록, 이러한 방식으로 구성될 수 있다. 상기 선택가능한 마커 유전자는, 상동의 재조합을 조정하기 위해 충분한 길이의 유전자의 일부로, 5' 및 3' 말단에, 작동가능하게 연결될 수 있다. 상기 선택가능한 마커는, 숙주 세포 영양요구성을 보완하거나, URA3, LEU2 및 HIS3 유전자를 포함하는, 항생물질 내성을 제공하는, 많은 유전자 중의 하나일 수 있다(The selectable marker can be one of any number of genes which either complement host cell auxotrophy or provide antibiotic resistance, including URA3, LEU2 and HIS3

genes). 그 밖의 적절한 선택가능한 마커는, 효모 세포에 클로람페니콜 저항(chloramphenicol resistance)을 부여하는, 상기 CAT 유전자, 또는 β-갈락토시다아제의 발현으로 인하여 블루 콜리니를 결과적으로 나타내는, 상기 lacZ 유전자를 포함한다. 그리고 난 다음에, 상기 유전자 대체의 직선의 DNA 단편은, 본 분야에서 널리 알려진 방법을 사용하여 상기 세포 내에 도입된다(하기를 참고하라). 상기 게놈 내로 상기 선형 단편의 통합 및 상기 유전자의 분포는 상기 선택 마커를 기초로 결정될 수 있고, 예를 들어, 서던 블롯 분석(Southern blot analysis)에 의해, 확인될 수 있다. 선택가능한 마커는, 예를 들어, Cre-loxP 시스템에 의해 상기 숙주 세포의 게놈으로부터 제거될 수 있다(하기를 참고하라).

[0161] 선택적으로, 유전자 대체 벡터는, 이의 일부가 어떠한 내인성 유전자 프로모터 서열이 없고, 상기 유전자의 상기 코딩 서열의 불활성 단편 또는 어떠한 것도 없는 것을 코딩하는, 방해하도록 상기 유전자의 일부를 포함하는 것에 관해서, 이러한 방식으로 구성될 수 있다(a gene replacement vector can be constructed in such a way as to include a portion of the gene to be disrupted, which portion is devoid of any endogenous gene promoter sequence and encodes none or an inactive fragment of the coding sequence of the gene). "불활성 단편(inactive fragment)"은, 예를 들어, 상기 유전자의 전체-길이 코딩 서열로부터 생산된 상기 단백질의 활성도의 약 10 % 미만(예를 들어, 약 9 % 미만, 약 8 % 미만, 약 7 % 미만, 약 6 % 미만, 약 5 % 미만, 약 4 % 미만, 약 3 % 미만, 약 2 % 미만, 약 1 % 미만, 또는 0 %)을 갖는 단백질을 코드하는 유전자의 단편이다. 상기 유전자의 이러한 일부는, 어떠한 알려지지 않는 프로모터 서열이 상기 유전자 서열에 작동가능하게 연결되지 않지만, 정지 코돈 및 전사 종결 서열(transcription termination sequence)이 상기 유전자 서열의 일부에 작동가능하게 연결되는, 이러한 방식으로 벡터에 삽입된다. 이러한 벡터는 상기 유전자 서열의 일부에 그 뒤에 선형화될 수 있고, 세포 내로 형질전환될 수 있다. 그리고 난 다음에, 단일 상동 재조합에 의해, 이러한 선형화된 벡터는 상기 유전자의 내인성 대응물(endogenous counterpart)에 통합된다.

[0163] 발현 벡터는 자율적이거나 통합될 수 있다. 재조합 핵산(예를 들어, 만노시다아제, GAA, 또는 알칼리성 프로테아제)가, 플라스미드, 파지, 트랜스포존, 코스미드 또는 바이러스 입자와 같은 발현 벡터의 형태로 상기 세포 내로 도입될 수 있다. 상기 재조합 핵산은, 염색체 외에 유지될 수 있거나, 이는 상기 효모 세포 염색체 DNA 내로 통합될 수 있다. 발현 벡터는, 상기 원하는 핵산으로 형질전환된 이러한 세포의 선택 및/또는 검출을 허용하기 위해, 선택된 조건(예를 들어, 트립토판 생합성에 요구되는 효소를 코딩하는, 우라실 생합성 또는 TRP1에 대해 필요한 효소를 코딩하는, URA3) 하에서 세포의 활성도에 대해 필요로 하는 단백질을 코딩하는 선택 마커 유전자를 포함할 수 있다(예를 들어, U.S. Pat. No. 4,704,362를 참고하라). 발현 벡터는, 자동적인 복제 서열(autonomous replication sequence, ARS)을 또한 포함할 수 있다. 예를 들어, U.S. Pat. No. 4,837,148에는, *피치아 파스토리스*에서 플라스미드를 유지하기 위해, 적절한 수단을 제공하는, 자기 복제 서열이 기재되어 있다.

[0165] 통합 벡터는, 예를 들어 U.S. Pat. No. 4,882,279에 기재되어 있다. 통합 벡터는 일반적으로, 적어도 제1 삽입가능한 DNA 단편, 선택가능한 마커 유전자, 및 제2 삽입가능한 DNA 단편의 연속적으로 배열된 서열을 포함한다. 상기 제1 및 제2 삽입가능한 DNA 단편은, 길이에서 각각 약 200 (예를 들어, 약 250, 약 300, 약 350, 약 400, 약 450, 약 500, 또는 약 1000 또는 그 이상) 뉴클레오티드이고, 형질전환되도록, 상기 종의 상기 게놈의 DNA의 일부에 대해 상동인 뉴클레오티드 서열을 가진다. 발현을 위해 관심 유전자(예를 들어, GAA를 코딩하는 유전자)를 포함하는 뉴클레오티드는, 상기 마커 유전자 전 또는 후에, 제1 및 제2 삽입가능한 DNA 단편 사이에 이러한 벡터에서 삽입된다. 통합 벡터는, 상기 숙주 세포 게놈 내로 관심 상기 뉴클레오티드 서열의 통합을 용이하게 하도록 효모 형질전환 전에 선형화될 수 있다.

[0167] 발현 벡터는, 이들이 상기 균류 세포에서 발현되도록, 효모(예를 들어, *야로위아 리폴리카*, *아르술라 아데니닌보란스*, *P. 파스토리스*, 또는 그 밖의 적절한 균류 종) 프로모터의 조절 하에서 재조합 핵산을 특징으로 한다. 적절한 효모 프로모터는, 예를 들어, ADC1, TPI1, ADH2, hp4d, POX, 및 Gal10(예를 들어, Guarente *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79(23):7410을 참고한다) 프로모터를 포함한다. 추가적으로 적절한 프로모터는, 예를 들어, Zhu and Zhang (1999) *Bioinformatics* 15(7-8):608-611 and U.S. Patent No. 6,265,185에

기재되어 있다.

[0169] 프로모터는, 구성적이거나 또는 유도가능한(inducible)(조건부)일 수 있다. 구성적인 프로모터(constitutive promoter)는 이의 발현이 상기 표준 배양 조건 하에서 일정한 프로모터인 것으로 이해된다. 유도가능한 프로모터(Inducible promoters)는, 하나 또는 그 이상의 유도 신호(induction cues)에 반응하는 프로모터이다. 예를 들어, 상기 유도성 프로모터는, 화학적으로 조절(예를 들어, 알코올, 테트라사이클린, 스테로이드, 금속, 또는 그 밖의 작은 분자와 같은 화학적인 유도제의 존재 또는 부재에 의해 조절되는 프로모터)되거나, 물리적으로 조절(예를 들어, 전사 활성도가 밝음 또는 높은 또는 낮은 온도와 같은 물리적인 유도물질의 존재 또는 부재에 의해 조절되는 프로모터)될 수 있다. 유도성 프로모터는 또한, 화학적 또는 물리적은 유도에 의해 이들 자체를 직접적으로 조절하는, 하나 또는 그 이상의 전사 인자에 의해 간접적으로 조절될 수 있다.

[0171] 그 밖의 유전적으로 조절된 변형이 또한 조절적일 수 있음을 이해할 수 있다. 예를 들어, 유전자는, 예를 들어, Cre-loxP 시스템과 같은 부위-특이적인 DNA 재조합효소를 사용하여, 조건적으로 삭제될 수 있다(예를 들어, Gossen *et al.* (2002) *Ann. Rev. Genetics* 36:153-173 and U.S. Application Publication No. 20060014264를 참고하라).

[0173] 재조합 핵산은, 스페로플라스트 기술(spheroplast technique) 또는 전체-세포 염화리튬 효모 형질전환 방법(whole-cell lithium chloride yeast transformation method)과 같은 다양한 방법을 사용하여 본원에 기재된 세포 내로 도입될 수 있다. 세포 내로 플라스미드 또는 선형 핵산 벡터의 형질전환에 유용한 그 밖의 방법은, 예를 들어, 이의 각각의 내용이 이의 전체로서 참고문헌으로 본원에 포함되는, U.S. Patent No. 4,929,555; Hinnen *et al.* (1978) *Proc. Nat. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163; U.S. Patent No. 4,879,231; and Sreekrishna *et al.* (1987) *Gene* 59:115 에 기재되어 있다. 일렉트로포레이션 및 PEG1000 전체 세포 형질전환 방법은, Cregg and Russel, Methods in Molecular Biology: Pichia Protocols, Chapter 3, Humana Press, Totowa, N.J., pp. 27-39 (1998)에 의해 기재된 바와 같이, 또한 사용될 수도 있다.

[0175] 형질전환된 균류 세포는, 필요로 하는(상기 세포의 영양요구성으로 인하여) 상기 생화학적인 생산물의 존재에서 형질전환 후에 영양 요구성 세포(auxotrophic cell)를 배양, 새로운 표현형의 선택 및 검출, 또는 상기 형질전환에 포함된 저항 유전자의 존재에서 상기 효모에 대한 유독성을 갖는 항생물질의 존재에서의 배양을 포함하지만, 이로 제한되지 않는, 적절한 기술을 사용함으로써 선택될 수 있다. 형질전환체는 또한, 예를 들어, 서던 블롯 또는 PCR 분석에 의해 측정될 수 있는, 상기 게놈 내로 상기 발현 카세트(expression cassette)의 통합에 의해 선택되고, 및/또는 확인될 수 있다.

[0177] 관심 타겟 세포 내로 상기 벡터를 도입하기 전에, 상기 벡터는, 상기에 기재된 바와 같이, *대장균(Escherichia coli, E. coli)*와 같은 박테리아 세포에서 성장할 수 있다(예를 들어, 증폭될 수 있다). 상기 벡터 DNA는, 상기 박테리아의 환경(bacterial milieu)으로부터 벡터 DNA의 정제를 결과적으로 나타내는, 본 분야에서 공지된 어떠한 방법에 의해 박테리아 세포로부터 분리될 수 있다. 상기 정제된 벡터 DNA는, 이러한 단백질이 포유동물의 세포에 대한 독성일 수도 있기 때문에, 어떠한 *E. coli* 단백질도 상기 플라스미드 DNA 제조에 존재하는 것을 확인하기 위해, 페놀, 클로로포름과 함께 광범위하게 추출될 수 있다.

[0179] 몇몇 실시형태에서, 상기 유전적으로 조작된 균류 세포는, OCH1 유전자 또는 유전자 생산물(예를 들어, mRNA 또는 단백질)이 결여되어 있고, OCH1 활성도가 결여되어 있다. 몇몇 실시형태에서, 상기 유전적으로 조작된 세포는, 만노실 인산화를 촉진할 수 있는 폴리펩티드(예를 들어, *야로위아 리폴리카, S. 세레비시아, 오가태 미누타, 피치아 파스토리스, 또는 C. 알비칸스*로부터의 MNN4 폴리펩티드, 또는 *P. 파스토리스*로부터의 PNO1 폴리펩티드)를 발현한다. 예를 들어, 균류 세포는, *Y. 리폴리카*로부터의 MNN4 폴리펩티드(Genbank® 접근 번호 : XM_503217, Genolevures Ref: YALI0D24101g)를 발현할 수 있다. 몇몇 실시형태에서, 상기 유전적으로 조작된

세포는 OCH1 활성도가 결여되어 있고, 만노실 인산화를 촉진할 수 있는 폴리펩티드를 발현한다.

[0181] 캡핑제거되고 디만노실레이션 후에, 상기 분자 복합체는 분리될 수 있다. 몇몇 실시형태에서, 상기 분자 복합체는 상기 효모 세포 내로 유지되고 세포 용해시 방출된다. 몇몇 실시형태에서, 상기 분자 복합체는, 상기 세포로부터 상기 분자 분비에 직접적인, 코딩 서열[상기 외인성 핵산 그대로 또는 상기 발현 벡터 내로 조작됨(either native to the exogenous nucleic acid or engineered into the expression vector)]에 의해 제공된 매커니즘을 통해 상기 배양 배지 내로 분비된다. 상기 세포 용해물 또는 배양 배지에서 캡핑제거되고 디만노실레이트된 분자 복합체의 존재는, 상기 분자의 존재를 검출하기 위해 다양한 표준 프로토콜, 예를 들어, GAA에 특이적인 항체와 면역블로팅(immunoblotting) 또는 방사선 면역 침강(radioimmunoprecipitation), 또는 특정한 효소 활성도에 대한 테스트[예를 들어, 글리코겐 분해(glycogen degradation)]에 의해 확인될 수 있다.

[0183] 몇몇 실시형태에서, 분리한 후에, 상기 캡핑제거되고 디만노실레이티된 타겟 분자 또는 분자 복합체는, 이종의 모이어티, 예를 들어, 효소적이거나 화학적인 수단에 부착될 수 있다. "이종의 모이어티(heterologous moiety)"는, 구성성분이 상기 변경된 타겟 분자에 원래 존재하는 구성요소와 상이한, 상기 변경된 타겟 분자에 연결되는(예를 들어, 공유결합적으로 또는 비-공유결합적으로) 어떠한 구성요소에 관한 것이다. 이종의 모이어티는, 예를 들어, 중합체, 담체, 아쥬반트(adjuvants), 면역독소(immunotoxins), 또는 검출가능한(예를 들어, 형광성의, 발광성의, 또는 방사성의) 모이어티를 포함한다. 몇몇 실시형태에서, 추가적인 N-글리칸은 상기 변경된 타겟 분자에 첨가될 수 있다.

[0185] 분자의 글리코실화를 검출하기 위한 방법은, DNA 염기서열 분석기-보조된(DSA), 형광체-보조된 탄수화물 전기영동(FACE) 또는 표면-증가된 레이저 증착/이온화 자동화 질량 분광분석법(surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, SELDI-TOF MS)을 포함한다. 예를 들어, 분석은, 예를 들어, 당단백질이 변성된 다음에, 예를 들어 막 상에 고정화되는, DSA-FACE를 사용할 수 있다. 그리고 난 다음에, 상기 당단백질은, 디티오트레이톨(DTT) 또는 J-메르캅토에탄올과 같은 적절한 환원제로 환원될 수 있다. 상기 단백질의 술프히드릴기는, 요오드아세트산과 같은 산을 사용하여 카르복실화될 수 있다. 그 다음에, 상기 N-글리칸은, N-글리코시다제 F와 같은 효소를 사용하여 상기 단백질로부터 방출될 수 있다. N-글리칸은 임의적으로, 재구성되고, 환원성 아미노화(reductive amination)에 의해 유도체화된다. 예를 들어, 상기 방출된 N-글리칸은, 환원성 아미노화(reductive amination)에 의해 환원하는 말단(reducing terminus)에서 형광체, 이러한 APTS [8-아미노피레네-1,2,6-트리술폰산(8-aminopyrene-1,3,6-trisulfonic acid)]와 표시될 수 있다. 표시하는 화학량론이,오직 하나의 APTS 분자가 올리고당의 각각의 분자에 부착되게 한 것이다(The stoichiometry of labeling is such that only one APTS molecule is attached to each molecule of oligosaccharide). 상기 유도체화된 N-글리칸은 그리고 난 다음에 농축된다. N-글리칸 분석에 적절한 장치(Instrumentation)는, 예를 들어, ABI PRISM® 377 DNA 서열분석기(Applied Biosystems)를 포함한다. 데이터 분석은, 예를 들어, GENESCAN® 3.1 소프트웨어(Applied Biosystems)를 사용하여 실행될 수 있다. 분리된 만노단백질(Isolated mannoproteins)은, 이들의 N-글리칸 상태를 확인하기 위해 송아지의 창자 포스파타아제(calf intestine phosphatase)와 같은 하나 또는 그 이상의 효소로 추가적으로 처리될 수 있다. N-글리칸 분석의 추가적인 방법은, 예를 들어, 질량 분석법(예를 들어, MALDI-TOF-MS), 보통의 상(normal phase) 또는 역상(reversed phase)에서 고성능 액체 크로마토그래피(HPLC), 및 이온 교환 크로마토그래피(예를 들어, 글리칸이 적절하게 표시된다면, UV 흡수 또는 형광발광과 함께 및 글리칸이 표시되지 않는 경우에, 펄스된 전류 측정 검출(pulsed amperometric detection)과 함께] 를 포함한다. 또한, Callewaert *et al.* (2001) *Glycobiology* 11(4):275-281 and Freire *et al.* (2006) *Bioconjug. Chem.* 17(2):559-564를 참고하라.

[0187] *조작된 세포의 배양(Cultures of Engineered Cells)*

[0188] 본 문서는, 본원에 기재된 상기 유전적으로 조작된 세포 중 어떠한 것의 실질적인 순수 배양(substantially pure culture)을 또한 제공한다. 본원에 사용된 바와 같이, 유전적으로 조작된 세포의 "실질적인 순수 배양"은, 상기 유전적으로 조작된 세포, 예를 들어, 박테리아, 균류(효모를 포함함), 마이코플라즈마(mycoplasmal), 또는 원생 동물 세포 이외에 생균수(viable cell)가 상기 배양에서 전체 생균수의 약 40 % 미만

(즉, 약 미만 : 35 %; 30 %; 25 %; 20 %; 15 %; 10 %; 5 %; 2 %; 1 %; 0.5 %; 0.25 %; 0.1 %; 0.01 %; 0.001 %; 0.0001 %; 또는 그 이하)인, 상기 세포의 배양이다. 본 내용에서 용어 "약"은, 상기 명시된 퍼센트 이상 또는 이하의 특정한 퍼센트의 15 % 퍼센트일 수 있음을 의미한다. 따라서, 예를 들어, 약 20 %는, 17 내지 23 %일 수 있다. 유전적으로 조작된 세포의 이러한 배양물은, 세포 및 성장, 저장, 또는 수송 배지를 포함한다. 배지는, 액체, 반-고체[예를 들어, 젤리와 같은 배지(gelatinous media)], 또는 냉동될 수 있다. 상기 배양은, 상기 액체 또는 상기 반-고체 배지에서 성장하거나, 냉동 저장 또는 수송 배지를 포함하는 저장 또는 수송 배지에서 수송되거나 저장되는 세포를 포함한다. 상기 배양물은, 배양관(culture vessel) 또는 저장관(storage vessel) 또는 기질(예를 들어, 배양 접시, 플라스크, 또는 튜브 또는 저장 바이알 또는 튜브)에 있다.

[0190] 본원에 기재된 상기 유전적으로 조작된 세포는, 예를 들어, 냉동된 세포 현탁액, 예를 들어, 동결건조된 세포로서, 글리세롤 또는 수크로스와 같은 저온보호물질(cryoprotectant)을 포함하는 완충용액에 저장될 수 있다. 선택적으로, 이들은, 예를 들어, 유동화 베드 건조(fluidized bed drying) 또는 분무 건조(spray drying), 또는 어떠한 적절한 건조 방법에 의해 수득된, 예를 들어 건조된 세포 제조물을 저장할 수 있다.

[0192] *약제학적 조성물 및 처리 방법*

[0193] 본원에 기재된 GAA 분자 및 분자 복합체, 예를 들어, 포유동물 세포의 내부로 수송을 향상시키는 적어도 하나의 변형을 포함하는 분자 복합체는, 치료학적으로 유효한 양의 상기 분자 및 하나 또는 그 이상의 아쥬반트(adjuvants), 부형제(excipients), 담체, 및/또는 희석제의 치료학적으로 유효한 양을 포함하는 약제학적 조성물 내로 포함될 수 있다. 허용가능한 희석제, 담체 및 부형제는 일반적으로, 수용자의 항상성(예를 들어, 전해질 균형)에 부정적으로 영향을 미치지 않는다. 허용가능한 담체는, 생체에 적합한, 불활성(inert) 또는 생흡수가능한 염(bioabsorbable salts), 완충제, 올리고- 또는 다당류, 중합체, 점성-개선하는 제제(viscosity-improving agent), 보존제 등을 포함한다. 하나의 대표적인 담체는, 생리적 염화나트륨 수용액(0.15 M NaCl, pH 7.0 to 7.4)이다. 다른 대표적인 담체는, 50 mM 인산나트륨, 100 mM 염화나트륨이다. 약학적인 조성물의 제형 및 투여를 위한 기술에 대해 추가적인 세부사항은, 예를 들어, Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.)에서 발견될 수 있다. 보충의 활성 화합물은 상기 조성물 내로 또한 혼합될 수 있다.

[0195] 본원에 기재된 하나 또는 변형을 갖는 분자 복합체를 포함하는 약제학적 조성물의 투여는 전신적이거나 국소적일 수 있다. 약제학적 조성물은, 이들이 비경구 및/또는 비-비경구 투여(non-parenteral administration)에 적절하도록 제형화될 수 있다. 특정한 투여 양상은, 피하의, 정맥내, 근육내, 복강내, 경피투여(transdermal), 척추 강내(intrathecal), 경구의(oral), 직장의(rectal), 구강(buccal), 국부의(topical), 코의(nasal), 안과의(ophthalmic), 관절 내(intra-articular), 관절 내(intra-arterial), 지주막 아래의(sub-arachnoid), 기관지의(bronchial), 림프액의(lymphatic), 질의(vaginal), 및 자궁내(intra-uterine) 투여를 포함한다.

[0197] 투여는 상기 약제학적 조성물의 볼루스(bolus)의 주기적인 주사에 의할 수 있거나, 외부[예를 들어, IV 백(IV bag)] 또는 내부[예를 들어, 생체 부식성 임플란트(bioerodable implant), 생체인공 장기, 또는 이식된 변경된 N-글리코실화 분자 생산 세포의 콜로니] 인 것인, 저장소로부터 정맥내 또는 복강내 투여에 의해 지속적이거나 연속될 수 있다. 예를 들어, U.S. Pat. Nos. 4,407,957, 5,798,113, and 5,800,828를 참고하라. 약제학적 조성물의 투여가, 펌프(예를 들어, Annals of Pharmacotherapy, 27:912 (1993); Cancer, 41:1270 (1993); Cancer Research, 44:1698 (1984); 미세캡슐화(microencapsulation)(예를 들어, U.S. Pat. Nos. 4,352,883; 4,353,888; and 5,084,350를 참고하라); 지속된 방출 중합체 임플란트(continuous release polymer implants)(예를 들어, Sabel, U.S. Pat. No. 4,883,666); 마크로엔캡슐화(macroencapsulation)(예를 들어, U.S. Pat. Nos. 5,284,761, 5,158,881, 4,976,859 및 4,968,733 및 공개된 PCT 특허 출원 WO92/19195, WO 95/05452를 참고하라); 피하내, 정맥내, 동맥내, 근육내, 또는 그 밖의 적절한 부위에서의 주사; 또는 캡슐, 액체, 정제, 알약, 또는 지속된 방출 제형으로, 경구 투여와 같은 적절한 수송 수단을 사용하여 달성될 수 있다.

[0199] 비경구적 전달 시스템(parenteral delivery systems)의 예는, 에틸렌-비닐 아세테이트 공중합체 입자(copolymer particles), 삼투 펌프, 주입가능한 주입 시스템(implantable infusion systems), 펌프 전달, 캡슐화된 세포 전달, 리포솜의 전달, 바늘-전달된 주사(needle-delivered injection), 바늘을 제외한 주사(needle-less injection), 분무기(nebulizer), 에어로졸(aerosolizer), 일렉트로포레이션, 및 경피투여 패치(transdermal patch)를 포함한다.

[0201] 비경구 투여에 적절한 제형은 일반적으로, 바람직하게 수용자의 혈액과 등장인, 변경된 N-글루코실화 분자의 멸균된 수용액(예를 들어, 생리 식염수)을 포함한다. 제형은 단일-투여(unit-dose) 또는 다중-투여 형태로 나타낼 수 있다.

[0203] 경구 투여에 적절한 제형은, 각각 변경된 N-글루코실화 분자의 예정된 양을 포함하는, 캡슐, 카시에(cachets), 정제 또는 로젠지(lozenges) ; 또는 시럽(syrup), 엘릭시르(elixir), 또는 물약(draught)과 같은 수성의 액체 또는 비-수성의 액체에서의 현탁액과 같은 별개의 단위(discrete units)로서 나타낼 수 있다.

[0205] 국소 투여에 적절한 포유동물 세포의 내부로 상기 복합체의 수송을 향상시키는 적어도 하나의 변형을 포함하는 분자 복합체는, 예를 들어 크림, 스프레이, 폼(foam), 겔, 연고, 고약(salve), 또는 건조 루브(dry rub)로서 포유동물(예를 들어, 인간 환자)에게 투여될 수 있다. 건조 루브는 또한 투여 부위(site of administration)에서 다시 수화(rehydrated)될 수 있다. 이러한 분자는 붕대(bandage), 거즈(gauze), 또는 패치(patch) 내[예를 들어, 스며들게(soaked into)하고 건조시킴]에 직접적으로 스며들게 할 수 있고, 그리고 난 다음에 국소적으로 적용할 수 있다. 이러한 분자는 또한, 국소 투여를 위한 붕대, 거즈 또는 패치에 반액체(semi-liquid), 겔화(gelled) 또는 완전한 액체 상태로 유지시킬 수 있다(예를 들어, U.S. Patent No. 4,307,717을 참고하라).

[0207] 약제학적 조성물의 치료학적 유효량(Therapeutically effective amounts)은, 본 분야의 기술 중의 하나로 확인할 수 있는 투여량 처방(dosage regimen)에서 이를 필요로 하는 피검자에게 투여될 수 있다. 예를 들어, 조성물은, 투여 당(per dose), 피검자의 몸무게당 0.01 ㎍/kg 내지 10,000 ㎍/kg의 투여량으로 전신적으로, 피검자에게 투여될 수 있다. 또 다른 예에서, 투여량은, 투여 당 피검자의 몸무게 당 1 ㎍/kg 내지 100 ㎍/kg 이다. 또 다른 예에서, 투여량은, 투여당 피검자의 몸무게 당 1 ㎍/kg 내지 30 ㎍/kg, 예를 들어 투여당 피검자의 몸무게 당 1 ㎍/kg 내지 10 ㎍/kg 이다.

[0208]

[0209] 치료학적 효율(therapeutic efficacy)이 최적화되도록, 본원에 기재된 분자 복합체는, 상이한 투여 치료 계획(different dosing regimens)에서 첫 번째로 투여될 수 있다. 단위 투여 및 치료 계획은, 예를 들어 포유동물의 종(pecies of mammal), 이의 면역 상태(immune status), 포유동물의 몸무게를 포함하는 요인에 따라 달라진다. 일반적으로, 조직에서의 이러한 분자 복합체의 레벨은, 예를 들어 주어진 치료 계획의 효율성을 결정하기 위해, 임상적인 테스트 절차의 일부로서 적절한 스크리닝 검정(screening assays)을 사용하여 모니터될 수 있다.

[0211] 본원에 기재된 분자 복합체에 대한 투여 횟수는, 의료의 전문가(예를 들어, 의사나 간호사)의 임상적인 판단 및 기술 내에 있다. 일반적으로, 투여 치료 계획은 최적의 투여 파라미터를 설정할 수도 있는 임상적인 시도(clinical trials)에 의해 설정된다. 그러나, 전문가는, 피검자의 나이, 건강, 몸무게, 성별 및 의학적인 상태에 따라 이러한 투여 치료계획은 다양할 수도 있다. 투여량의 횟수는, 치료가 예방적(prophylactic) 또는 치료학적(therapeutic) 인지에 따라서 다양할 수 있다.

[0213] 이러한 분자 복합체 또는 이의 약제학적 조성물의 독성 및 치료학적 효과는, 예를 들어 세포 배양 또는 실험 동물에서의 공지된 약제학적 절차에 의해 결정될 수 있다. 이러한 절차는, 예를 들어, LD_{50}[50 % 의 개체군에게

치명적인 투여량 (dose lethal to 50% of the population)] 및 ED_{50}[50 %의 개체군에서의 치료학적 유효량(dose therapeutically effective in 50% of the population)]를 측정하기 위해 사용될 수 있다. 독성 및 치료학적 효과 사이의 투여량 비율은 치료 지수(therapeutic index)이고, 이는 비율 LD_{50}/ED_{50}로서 표현될 수 있다. 높은 치료학적 지수(high therapeutic indices)를 나타내는 약제학적 조성물이 바람직하다. 독성 부작용을 나타내는 약제학적 조성물이 사용될 수 있는 반면에, 주의(care)는, 정상 세포[예를 들어, 비-타겟 세포(non-target cells)]에 대한 잠재적인 손상을 최소화하기 위해 영향을 받은 조직(affected tissue)의 부위에 이러한 화합물을 타겟하는(targets) 전달 시스템(delivery system)을 설계하기 위한 조치를 취해야 하고, 이로 인해 부작용을 줄일 수 있다

[0215] 세포의 배양 검정 및 동물 연구로부터 획득된 데이터는 적절한 피검자(예를 들어, 인간 환자)에 사용하기 위한 투여량의 범위를 제형화하는데 사용될 수 있다. 이러한 약제학적 조성물의 투여량은, 독성이 적거나 없는 ED_{50}을 포함하는 순환 농도(circulating concentrations)의 범위 내에 일반적으로 놓인다. 상기 투여량은 사용된 제형(dosage form) 및 이용된 투여의 경로에 따라 이러한 범위 내에 다양할 수 있다. 본원에 사용된 바와 같은 약제학적 조성물에 대해서[예를 들어, 피검자에서의 대사 질환을 치료하기 위해], 치료학적 유효량은 세포 배양 검정으로부터 처음에 평가될 수 있다. 투여량은, 세포 배양에서 결정된 바와 같은 IC_{50} [즉, 징후의 최대한의 저해의 반(half-maximal inhibition)을 달성하는 약제학적 조성물의 농도]을 포함하는 순환 플라즈마 농도 범위(circulating plasma concentration range)를 달성하기 위해 동물 모델에서 제형화될 수 있다. 이러한 정보는 인간에서의 유효한 투여량을 보다 정확하게 측정하기 위해 사용될 수 있다. 플라즈마(plasma)에서의 레벨은, 예를 들어 고성능 액체 크로마토그래피에서 측정될 수 있다.

[0216]

[0217] 본원에서 나타낸 바와 같이, 분자 복합체의 "치료학적 유효량"은 처리된 피검자에게 의학상으로 원하는 결과(예를 들어, 폼베병의 하나 또는 그 이상의 징후의 개선)를 생산할 수 있는 상기 복합체의 양이다. 치료학적으로 유효량(즉, 유효한 투여량)은 피검자 또는 샘플 무게의 킬로그램 당 상기 복합체의 밀리그램(milligram) 또는 마이크로그램(microgram)의 양[예를 들어, 킬로그램당 약 1 마이크로그램 내지 킬로그램당 약 500 밀리그램, 킬로그램당 약 100 마이크로그램 내지 킬로그램당 약 5 밀리그램, 킬로그램당 약 1 마이크로그램 내지 킬로그램당 약 50 마이크로 그램]을 포함할 수 있다.

[0218]

[0219] 피검자는, 어떠한 동물, 예를 들어 인간(예를 들어, 인간 환자) 또는 비-인간 영장류(non-human primate)[예를 들어, 침팬치, 개코원숭이(baboon), 또는 원숭이], 마우스, 랫, 토끼, 기니피그(guinea pig), 게르빌루스쥐(gerbil), 햄스터, 말, 가축의 종류(예를 들어, 소, 돼지, 양, 또는 염소), 개, 고양이 또는 고래일 수 있다.

[0221] 본원에 기재된 분자 복합체 또는 이의 약제학적 조성물은, 폼페병에 사용된 또다른 치료와 병용요법(combination therapy)으로서 피검자에게 투여될 수 있다. 예를 들어, 병용요법은, 폼페병이 진행될, (또는 GAA를 코딩하는 유전자에서의 돌연변이로 인하여) 위험이 있거나 또는 이를 갖는 피검자에게 치료학적 이익을 제공하는 하나 또는 그 이상의 추가적인 제제를 피검자(예를 들어, 인간 환자)에게 투여하는 것을 포함할 수 있다. 따라서, 화합물 또는 약제학적 조성물 및 하나 또는 그 이상의 추가적인 제제는 동시에 투여될 수 있다. 선택적으로, 상기 분자 복합체는 첫 번째로 투여될 수 있고, 하나 또는 그 이상의 추가적인 제제는 두 번째로 투여될 수 있고, 또는 반대로 투여될 수 있다.

[0222]

[0223] 본원에 기재된 분자 복합체의 어떠한 것은 동결건조될 수 있다.

[0225] 본원에 기재된 약제학적 조성물의 어떠한 것은, 투여를 위한 설명서(instructions)와 용기(container), 팩(pack) 또는 디스펜서(dispenser)를 포함할 수 있다. 몇몇 실시형태에서, 상기 조성물은 단일 사용 바이알

(single use vial)로서 포장된다.

[0227] 하기의 내용은 본 발명의 실행의 예이다. 이들은 어떠한 방식으로든 본 발명의 범위를 한정하는 것으로 이해되지 않는다.

[0229] 실시예

[0230] 실시예 1

[0231] CcMan5 및 잭 빈 α-만노시다아제로 재조합 huGAA의 캡핑제거 및 디-만노실레이션(Uncapping and de-mannosylation of recombinant huGAA with CcMan5 and Jack bean a-mannosidase)

[0232] 재조합 인간 GAA(rhGAA)는, 상기 인간 알파 글루코시다아제 유전자(산 알파 글루코시다아제 또는 산 말타제 EC3.2.1.3로서 알려짐)의 세 개의 복사물 및 *Y. 리폴리카* MNN4 유전자의 두 개의 복사물을 포함하는, *Y. 리폴리카* 생산 스트레인 OXYY1589을 사용하여 WO2011/039634에 기재된 바와 같이 생산되었다. 인간 GAA의 상기 아미노산 서열은 도 1 에 기재되어 있다. 상기 스트레인 OXY1589(strain OXY1589)의 유전자형은 하기와 같다 :

[0233] *MatA, leu2-958, ura3-302, xpr2-322,*

[0234] *gut2-744, ade2-844*

[0236] *POX2-Lip2pre-huGAA:URA3Ex::zeta*

[0237] *POX2-Lip2pre-huGAA:LEU2Ex::zeta*

[0238] *POX2-Lip2pre-hGM-CSF:GUTEx::zeta*

[0239] Y1MNN4-POX2-hp4d-YLMNN4 :ADE2*::PT targeted*

[0241] RhGAA는 *셀룰로시마이크로비움 셀룰란스* 만노시다아제(CcMan5) 및 잭 빈 α 만노시다아제(JbMan)(Sigma Product M7257, 3.0 M 황산암모늄 현탁액)로 캡핑제거되고 디만노실레이트된 것이다. CcMan5는, DsbA 신호 서열을 포함하는, 벡터 pLSAH36 내로, 상기 CcMan5 폴리펩티드를 코딩하는 핵산(도 2a)를 첫 번째 클로닝함으로써 재조합적으로 생산되고, N-말단 HIS 태그를 갖는 단백질의 발현을 결과적으로 나타낸다. 도 2b 및 2c는, 각각 신호 서열이 있고, 없는 상기 CcMan5 폴리펩티드의 상기 아미노산 서열을 포함한다. 플라스미드 pLSAH36는 *E. coli* B21 세포 내로 클로닝되고, 상기 주변세포질에 거주하는 단백질은 Talon 컬럼을 사용하여 분리되고 정제된다. JbMan의 상기 황산암모늄 현탁액을 사용하기 전에, 이는 포스파타아제 활성도에 악영향을 주는 것을 제거하기 위해, Superdex 200 컬럼을 통해 젤 여과에 의해 추가적으로 정제되었다.

[0243] RhGAA[20 mM의 아세트산나트륨 (NaOAc) 완충용액, pH 5.0에서 약 5 mg/mL의 농도]는, huGAA:CcMan5:JbMan에 대해 100:5:10의 w:w 비율로, CcMan5(인산 완충 식염수(PBS)에서 약 0.15 - 0.30 mg/mL) 및 JbMan (약 0.5 - 1 mg/mL in PBS)과 배양함으로써 캡핑제거되고 디만노실레이트되었다. 상기 전체 반응 부피는 100 mM NaOAc and 2 mM $CaCl_2$의 최종 농도를 수득하기 위해, 500 mM NaOAc 완충용액, pH 5.0 및 100 mM $CaCl_2$으로 희석되었다. 상기 반응 혼합물은 16 시간 동안 30 ℃에서 배양되었다.

[0245] 상기 캡핑제거하는 과정을 평가하고, 상기 정제된 huGAA의 상기 N-글리칸 프로파일을 분석하기 위해, 5 ㎍의 당단백질의 상기 N-글리칸은 N-글리코시다아제 F(PNGaseF)와 방출되고, APTS[8-아미노-1,3,6-피레네트리술폰산(pyrenetrisulfonic acid); 트리소듐 염(trisodium salt)]으로 표시되고, 그 다음에 DSA-FACE[DNA 서열분석기-보조된 형광체(Aided Fluorophore)-보조된 탄수화물 전기영동(Assisted Carbohydrate Electrophoresis)]로 분석된다. 상기 방법은, Laroy *et al*, *Nature Protocols*, 1:397-405 (2006에 기재된 상기 프로토콜을 필수적으로 따른다.

[0247] CcMan5 및 JbMan으로 처리 전(패널 B) 및 처리 후(패널 C), GAA (76 kD form)로부터 상기 N-글리칸의 상기 DSA-FACE의 일렉트로페로그램(electropherogram)은, 도 3a에 기재되어 있다. 패널 A는, PNGaseF을 갖는 RNaseB로부터 방출된 상기 N-글리칸을 포함하는 표준 샘플이다. 캡핑된 huGAA로부터 방출된 상기 N-글리칸 혼합물은, ManP-$Man_8GlcNAc_2$ 및 $(ManP)_2$-$Man_8GlcNAc_2$으로 주로 구성된다(도 3a, 패널 B). 보다 약간 빠르게 움직이는 피크는 ManP-$Man_7GlcNAc_2$에 배정될 수 있다(A peak running slightly faster than ManP-$Man_8GlcNAc_2$ can be assigned to ManP-$Man_7GlcNAc_2$). 캡핑제거되고 디만노실레이션 후에 관찰된 상기 주요한 피크는, 상기 이중 인산화된 P_2-$Man_6GlcNAc_2$ 및 상기 모노인산화된(monophosphorylated) P-$Man_4GlcNAc_2$, P-$Man_5GlcNAc_2$ 및 P-$Man_6GlcNAc_2$(패널 C)에 배정된다.

[0249] huGAA 결과의 상이하게 처리된 형태의 캡핑제거는, 76 kD 형태(패널 B), 95 kD 형태(패널 C) 및 110 kD 형태(패널 D)에 대한 상기 동일한 N-글리칸 프로파일을 결과적으로 나타낸다(도 3b).

[0251] 실시예 2

[0252] 110 kDa rhGAA의 정제

[0253] rhGAA의 상기 110 kDa 형태는 하기와 같이 스트레인 OXYY1589로부터 분리되었다. 채취한 후에, 브로스(broth)는 원심분리하고, Durapore 멤브레인(Merck Millipore)을 사용하여 여과되었다. 황산암모늄(AMS)는 1 M의 농도로 첨가되었고, 상기 용질(solute)은, 20 mM의 인산나트륨 pH 6, 1 M 황산암모늄으로 평형된, 소수성 상호작용 크로마토그래피(HIC) 컬럼 상에 로딩(loading)하기 전에 여과되었다. 상기 생산물은 20 mM 인산나트륨 pH 6으로 용출되었다(elute).

[0255] 두 번째 크로마토그래피 컬럼 상에 로딩하기 전에, 상기 생산물은, 재생된 셀룰로오스 막 상에 접선 유동 여과(tangential flow filtration, TFF)를 통해 첫 번째로 농축된 다음에, 완충용액으로부터 20 mM의 아세트산나트륨 pH 4.5으로 교환되었다. 이러한 물질은, 20 mM 아세트산 나트륨 pH 4.5으로 평형된, 양이온 교환 크로마토그래피(CEX) 컬럼 상에 로딩되었다. 상기 컬럼을 로딩한 후에, 이는 상기 UV 흡광도 신호가 베이스라인(baseline)에 도달할 때까지, 평형 완충용액으로 세척되었고, 그리고 난 다음에, 20 mM 아세트산 나트륨 pH 4.5, 50 mM NaCl로 세척되었다. 상기 생산물은 20 mM 아세트산 나트륨 pH 4.5, 150 mM NaCl으로 용출되었고, 농축되고, 완충용액에서 20 mM 아세트산 나트륨 pH 5로 교환되었다(The product was eluted in 20 mM sodium acetate pH 4.5, 150 mM NaCl, and concentrated and exchanged from buffer to 20 mM sodium acetate pH 5). (하기를 참고하라)

[0257] 상기 샘플은 실시예 1에 기재된 바와 같이 캡핑제거되고 디만노실레이트된 다음에, D-만니톨은 100 mM의 농도에 첨가되었다. 그러한 물질의 4 분의 3(Three quarters of that material)은, 10 kDa 분자량 컷 오프(MWCO)를 갖는 재생된 셀룰로오스 막을 사용하여 TFF 통해 부피로 감소되고, 25 mM 인산 나트륨 pH 6, 150 mM NaCl, 100 mM D-만니톨로 4 ℃에서 평형유지된 Superdex 200 컬럼 상에서 크기 배제 크로마토그래피(SEC)를 통해 추가적으로 정제되었다. 분획은, 변성 조건(denaturating condition) 하에서 시바크론-블루 염색된 폴리아크릴아미드 겔에서 순도에 대해 나중에 선별되었다. 모여진 분획은 TFF를 통해 농축되었고, 10 kD MWCO(재생된 셀룰로오스 막, Merck Millipore)의 Amicon 원심여과기 필터를 사용하여 초원심분리되었다.

[0259] 실시예 3

[0260] 110 kDa rhGAA의 정제

[0261] rhGAA의 상기 110 kDa 형태는 하기와 같이 스트레인 OXYY1589로부터 분리되었다. 채취한 후에, 상기 물질은, 상기 AMS의 농도가 1 M으로 증가되기 전에, 원심분리하고, 여과되었다. 상기 용질(solute)은 또 다시 여과되었

고, 상기 생산물은 HIC 컬럼 상에 포획되고(captured), 20 mM 인산나트륨 pH 6, 1 M AMS으로 균형이 유지되었고, 20 mM의 인산나트륨 pH 6 완충용액에서 1 에서 0 M으로 AMS의 단계 기울기(step gradient)로 방출되었다.

[0263] 상기 용출액(eluate)은 농축되었고, 완충용액은, Vivaflow 200 모듈(PES membrane, 10 kD MWCO, Sartorius) 상에 TFF를 통해 10 mM BIS-TRIS, pH 6으로 교환되었다. 탈염된 물질은, 음이온 교환 크로마토그래피 (AEC) 컬럼 상으로 제공되었다. 상기 UV 신호가 거의 대부분 베이스라인에 도달할 때까지 상기 컬럼을 세척한 후에, 이-상 지속적인 염 기울기(two-phase continuous salt gradient)가 적용되었고; 0 에서 0.3 M NaCl까지 가는 첫 번째 기울기(the first gradient going from 0 to 0.3 M NaCl), 0.3 내지 1 M NaCl까지 가는 두 번째 기울기. 분획은 상기 기울기 동안에 수집되었고, 정성 4-메틸움벨리페릴(methylumbellifferyl)-a-D-글루코피라노사이드(qualitative 4-methylumbellifferyl-a-D-glucopyranoside, 4-MUG)를 통해 선별되었다. 상기 4-MUG 검정에서, 반응은, 10 μl의 용출 분획에 대한 10 : 1의 부피 비율로 0.35 mM 4-MUG, 0.1% BSA 및 100 mM 아세트산 나트륨 pH 4로 이루어진 반응 완충용액을 첨가함으로써 개시되었다. 37 ℃에서 30 분에서 1 시간 동안 배양한 후에, 100 mM 글리신 pH 11의 동일한 부피가 상기 반응을 중단하기 위해 첨가되었다. 상기 형광성 반응 생산물 4-메틸움벨리페론(4-methylumbelliferone)의 방출이 UV-광(UV-light) 하에서 관찰되었다.

[0265] GAA를 포함하는 분획이 모였고, 10 kD MWCO(재생된 셀룰로오스 막, Merck Millipore)의 Amicon 원심 여과기를 사용한 초원심분리 및 Vivaflow 200 모듈(PES membrane, 10 kD MWCO, Sartorius)에서 TFF를 통해 농축되었다.

[0267] 상기 농축된 물질은 두 개로 나뉘고, 50 mM 인산 나트륨 pH 6, 250 mM NaCl, 0.05 % Tween-20으로 4 ℃에서 평형된 Superdex 200 컬럼 상에서 추가적으로 정제되었다. 분획은, 변성 조건 하에서 시바크론-블루 염색된 폴리아크릴아미드 겔(cibacron-blue stained polyacrylamide gels)에서 순도를 위해 나중에 선별되었고, 포스파타아제 함량은, 405 nm의 파장에서 황색으로 착색된 p-니트로페놀레이트 반응 생산물의 효소의 방출(enzymatic release)을 측정하는, 파라-니트로페닐포스페이트(para-nitrophenylphosphate)를 사용한 비색 테스트(colormetric test)를 사용하여 측정되었다.

[0269] 파일럿 풀(Pilot pools)은 GAA를 함유하는 분획으로부터 제조되었다. 상기 파일럿 풀의 전체 단백질은 Bradford 정량(Bradford assay)을 통해 측정되었다. 선택된 분획은 Vivaflow 200 TFF 모듈(PES 막, 10 kD MWCO, Sartorius)로 농축을 위해 모아졌다. 상기 부피는 10 kD MWCO(재생된 셀룰로오스 막, Merck Millipore)의 15 ml Amicon 원심 분리기를 사용하여 추가적으로 감소되었다.

[0271] 상기 농축된 물질은, 상기 첫 번째 SEC 단계에 대한 바와 같이 상기 동일한 조건에 대한 상기 동일한 조건을 사용한, 크기 배제 크로마토그래피(SEC)를 두 번째로 하였다. 분획은, 변성 조건 하에서 시바크론-블루 염색된 폴리아크릴아미드 겔에서 순도를 위해 또 다시 선별되었다. 분획은 상기 선택된 파일럿 풀에 따라 모아졌고, 15 ml Amicon 원심여과기(10 kD MWCO, 재생된 셀룰로오스 막, Merck Millipore) 상에서 농축되었다.

[0273] 실시예 4

[0274] 95 kDa rhGAA의 정제

[0275] rhGAA의 상기 95 kDa 형태는 하기와 같이 스트레인 OXYY1589로부터 분리되었다. 채취한 후에, 상기 브로스(broth)는 원심분리되었고, Durapore 막을 사용하여 여과되었다(Merck Millipore). 상기 생산물은 그 후에, 10 kD의 분획분자량(molecular-weight-cut-off, MWCO)으로 변형된 폴리에테르술폰(PES) 막 상에서 TFF를 통해 농축된다. AMS는 1 M의 농도로 첨가되었고, 상기 용질은 HIC 컬럼 상에서 로딩(loading)하기 전에 여과되었고, 20 mM 인산나트륨 pH 6, 1 M AMS로 평형되었다. 상기 생산물은 물로 용리되었고, 상기 용출액의 pH는 100 mM BIS-TRIS pH 6의 스톡 완충용액(stock buffer)을 첨가하여 10 mM의 농도로 조절하였고, 이는 13일 동안 4 ℃에서 저장되었다.

[0277] AEX 컬럼에서 로딩하기 전에, 상기 생산물은 10 kD의 MWCO를 갖는 재생된 셀룰로오스 막 상에 TFF를 통해 농축되었고, 10 mM BIS-TRIS pH 6로 완충용액-교환되었다(the product was concentrated via TFF on a regenerated cellulose membrane with an MWCO of 10 kD and buffer-exchanged to 10 mM BIS-TRIS pH 6). 상기 탈염된 물질(desalted material)은, 실시예 3에 기재된 바와 같이 실행된, AEX 크로마토그래피를 통해 추가적으로 공정되었다. 분획은 상기 기울기 동안에 수집되었고, 정성의 4-MUG 검정을 통해 GAA에 대해 선별되었다. GAA를 포함하는 분획은 추가적인 정제를 위해 모아졌다.

[0279] 상기 세번째 크로마토그래피 단계에 대해, AMS의 농도는 1 M으로 증가되었고, 여과된 후에, 상기 물질은 HIC를 통해 추가적으로 정제되었다. 1 에서 0 M AMS까지의 지속적인 염 기울기(continuous salt gradient)가 적용되면서, 상기 기울기 동안에 분획이 수집되었다. 모든 분획은 정성 분석(qualitative assay)을 통해 GAA에 대해 선별되었고, GAA를 포함하는 이러한 것들은 추가적인 공정을 위해 모아졌다.

[0281] 상기 풀(pool)은 10 kD MWCO 재생된 셀룰로오스 막의 15 ml Amicon 원심여과기를 사용한 초원심분리를 통해 농축되었고, 추가적으로 실시예 3 에 기재된 바와 같이 동일한 공정을 사용하여 SEC를 통해 정제되었다. 분획은 그 다음에, 변성 조건 하에서 시바크론-블루 염색된 폴리아크릴아미드 겔에서 순도를 위해 또 다시 선별되었다. 상기 90 %의 순수한 GAA 분획이 모아졌고, TFF Vivaflow 200 모듈(PES 막, 10 kD MWCO, Sartorius)에서 첫 번째로 농축된 다음에, 15 ml Amicon 원심여과기(10 kD MWCO, 재생된 셀룰로오스 막, Merck Millipore)를 사용하여 초원심분리되었다. 상기 농축된 물질은, 상기 첫 번째 SEC 단계에 대한 바와 같이 상기 동일한 조건을 사용하여 SEC를 두 번째로 하였다. 분획은 정성 4-MUG GAA 검정을 사용하여 GAA에 대해 선별되었다(Fractions were screened for GAA using the qualitative 4-MUG GAA assay). GAA 활성도를 갖는 분획이 모아졌고, 농축되었다.

[0283] 실시예 1 에 기재된 바와 같이 캡핑제거되고 디만노실레이션 후에, D-만니톨은 100 mM의 농도로 첨가되었고, 상기 부피는 최종의 Superdex 200 겔 여과 컬럼으로 로딩되기 전에 또 다시 감소되었고, 25 mM 인산나트륨 pH 6, 150 mM NaCl, 100 mM D-만니톨로 4 ℃에서 평형되었다. 분획은 상기 4-MUG 정성 분석을 사용하여 GAA에 대해 선별되었고, 상기 생산물을 포함하는 이러한 것들이 모아졌고, 농축되었다.

[0285] 실시예 5

[0286] 95-110 kDa rhGAA mix의 정제

[0287] rhGAA의 110 kDa 전구체 및 95 kDa 형태 둘 다는 하기와 같이 스트레인 OXYY1589로부터 분리되었다. 채취한 후에, 상기 물질은, 실시예 2에 기재된 바와 같이 두 번째 크로마토그래피 단계로 처리되었다. 상기 HIC 단계 후에, 상기 생산물은 농축되었고, 상기 완충용액은 TFF를 통해 10 mM BIS-TRIS pH 6로 교환되었고, AEX 컬럼 상에서 로딩되었다. 상기 생산물은 pH 6(10 mM BIS-TRIS)에서 0 에서 300 mM까지의 NaCl의 단일 단계에서 용출된 다음에, Pellicon XL50 TFF 모듈(10 kD MWCO과 함께 재생된 셀룰로오스 막)을 사용하여 농축되었다. 상기 물질의 반은 크기 배제 크로마토그래피를 통해 추가적으로 정제되었다. 상기 크로마토그래피 단계는 실시예 3에 기재된 바와 같이 실행되었지만, 추가적인 공정(processing) 동안에 상기 분획의 선택은 변성 조건 하에서 시바크론-블루 염색된 폴리아크릴아미드 겔 상에 순도의 기초에서만 실행되었다.

[0289] 상기 풀(pool)의 반은 농축되었고, 상기 AEX-물질의 나머지와 결합되었다. 캡핑제거되고 디만노실레이션 후에, D-만니톨의 농도는 100 mM 및 그 다음 농도로 증가되었고, SEC 단계는 도 2에 기재된 바와 같은 상기 동일한 절차 후에 실행되었다. 분획은 4-MUG 정성 검정의 기초로 모아졌고, 실시예 6으로부터 캡핑제거된 생산물로 모아졌다.

[0291] 실시예 6

[0292] 95 kDa rhGAA의 정제

[0293] rhGAA의 95 kDa 형태가 하기와 같이 스트레인 OXYY1589로부터 분리되었다. 채취한 후에, 상기 물질은, 실시예 3에 기재된 바와 같이 상기 AEX 단계를 포함될 때까지 처리되었다(After harvest, the material was processed up to and including the AEX step as described in Example 3). 상기 AEX 단계에서, 상기 생산물의 중요한 양은, 상기 로딩 동안에서 전도성(conductivity)의 증가로 인하여 분획을 통해 상기 흐름에 존재하였다(a significant amount of the product resided in the flow through fraction due to an increase of conductivity during the loading). 물질을 통한 상기 흐름은 따라서, 상기 적절한 완충용액으로 또다시 초여과되었고, AEX 크로마토그래피를 두 번째 하였다(The flow through material was therefore again diafiltered to the appropriate buffer and subjected to a second round of AEX chromatography). 둘 다의 양(배치 A 및 배치 B)가 별도로 처리된 것에서 여기로부터 있었다(Both amounts (batch A and batch B) were from here on processed separately).

[0295] 배치 A 는 실시예 5로부터 상기 SEC 풀(pool)의 나머지와 결합되었고, 실시예 2로부터 상기 CEX 풀(pool)의 나머지 및 상기 풀(pool)은 그 다음에 농축되었고, 10 mM BIS-TRIS pH 6, 300 mM NaCl를 포함하는 완충용액으로 초여과되었다(diafiltered). 상기 물질은 65 h 동안 30 ℃에서 배양되었다. 그리고 난 다음에, 상기 pH는 125 mM의 농도로 아세트산 나트륨 pH 5의 1 M 스톡 완충용액을 첨가함으로써 pH 5로 낮아졌고(The pH then was lowered to pH 5 by adding a 1 M stock buffer of sodium acetate pH 5 to a concentration of 125 mM), 상기 샘플은 30 ℃에서 또 다시 배양되었다. 24 h 후에, 상기 생산물은, 상기 생산물 대 프로테아제 혼합물의 40 : 1 중량:중량 비율 전체 단백질 함량을 사용한, *아스페르길루스 오리제*로부터의 프로테아제 혼합물, Flavourzyme (Novozymes Corp)로 처리되었고, 마지막 시간 동안에 30 ℃에서 배양되었다(After 24h, the product was treated with Flavourzyme (Novozymes Corp), a protease mix from *Aspergillus oryzae*, using a 40:1 weight:weight ratio total protein content of the product versus protease mix, and was for the last time incubated at 30 ℃). 밤새 배양한 후에, 상기 물질은 첫 번째 SEC 단계를 통해 정제되었고, 실시예 3에 기재된 바와 같은 동일한 조건 하에서 실행되었다. 분획은 변성 조건 하에서 시바크론-블루 염색된 폴리아크릴아미드 겔의 기초 상에 순도 생산물을 포함하도록 추정된 것을 모아졌다(Fractions were pooled that were estimated to contain pure product on the basis of cibacron-blue stained polyacrylamide gels under denaturing conditions). 농축 및 20 mM 아세트산 나트륨 pH 5로 완충용액 교환한 후에, 상기 물질은 실시예 1에 기재된 바와 같이 캡핑제거되고 디만노실레이션되었다. D-만니톨은 100 mM의 농도로 첨가되었고, 상기 물질은 실시예 5로부터 캡핑제거되고 디만노실레이션된 물질과 함께 모아졌다. 상기 최종의 SEC 단계 및 그 다음의 샘플 처리는 실시예 2에 기재된 바와 같이 실행되었다.

[0297] 배치 B 는 실시예 3 으로부터 최종 물질과 함께 모아졌고, 상기 풀(pool)은 농축되었고, 10 mM BIS-TRIS pH 6, 300 mM NaCl를 포함하는 완충용액으로 초여과되었다. 상기 생산물은 그리고 난 다음에, 실시예 5에 기재된 바와 같은 동일한 중량 비율을 사용하여 30 ℃에서 밤새 배양 기간 동안에 *A. 오리제* 프로테아제 혼합물로 처리되었고, 그 후에, 첫 번째 SEC 단계를 통해 정제되었고, 실시예 3에 기재된 바와 같은 동일한 조건 하에서 실행되었다. 상기 생산물의 추가적인 공정은 실시예 5에 기재된 바와 같이 실행되었다.

[0299] 상기 최종의 배치에서, 배치 A 및 배치 B로부터의 생산물은 GAA 함량에서 14:1 비율로 혼합되었다.

[0301] 실시예 7

[0302] 76 kDa rhGAA의 정제

[0303] rhGAA의 76 kDa 형태는, 하기와 같이 스트레인 OXYY1589로부터 분리되었다. 채취한 후에, 상기 배양물은 2 회의 지속적인 원심분리를 하였다. 상기 상청액은 모아졌고, AMS는 대략 1 M의 농도로 도입되었다. 용해 후에, 1 용량의 HIC 레진은 20 mM 인산 나트륨 pH 6, 1 M AMS으로 예비-평형되었고, 배치 흡수 모드(batch uptake mode)에서 상기 생산물을 결합하기 위해 교반시키면서, 50 용량의 상청액에 첨가되었다(After dissolution, 1

volume of HIC resin, pre-equilibrated in 20 mM sodium phosphate pH 6, 1 M AMS, was added to 50 volumes of supernatant while stirring to bind the product in a batch uptake mode). 결과적으로 생성된 슬러리는 교반 없이 4 ℃에서 밤새 저장되었다. 이러한 기간 동안에서, 용질(solute)의 윗쪽에 놓여있는 갈색의 층은 조심스러운 흡인(gentle aspiration)을 통해 아침에 제거되었다. 상기 레진은, 이것이 컬럼 내로 포장되기 전에, 각각의 회에서 3 용량의 리드 완충용액(lead buffer)(20 mM 인산 나트륨 pH 6, 1 M AMS)으로 세 번 세척되었다. 상기 포장된 레진은 UV 신호가 거의 베이스라인에 도달할 때까지 세척되었고, 그 후에 상기 생산물은 물과 함께 용출되었다. 상기 용출액의 pH 는 100 mM BIS-TRIS pH 6의 스톡 완충용액을 첨가함으로써 10 mM의 농도로 조절되었다. 상기 물질은 백(bag)에서 서서히 여과된 다음에, 4 ℃에서 11 일의 기간 동안에 저장되었다.

[0305] 상기 두 번째 및 세 번째 크로마토그래피 단계 및 동반하는 상기 물질의 처리는 실시예 4에 기재된 바와 같이 실행되었다. 상기 세 번째 크로마토그래피 단계 후의 상기 풀(pool)은 동시에 결합된 두 개의 TFF Vivaflow 200 모듈(PES membrane, 10 kD MWCO, Sartorius)에서 대략 7 번으로 첫 번째로 농축되었고, 그 다음에 10 kD MWCO의 15 ml Amicon 원심 여과기(재생된 셀룰로오스 막, Merck Millipore)를 사용하여, 초-원심분리되었다(ultra-centrifuged). 상기 농축된 물질은 두 개로 분열되고, 실시예 4 에 기재된 상기 동일한 조건을 사용하여 SEC를 통해 추가적으로 정제되었다. 분획은 변성 조건 하에서 시바크론-블루 염색된 폴리아크릴아미드 겔 상에 순도에 대해 그 후에 선별되었다. 선택된 분획은 동시에 결합된 두 개의 Vivaflow 200 TFF 모듈(PES 막, 10 kD MWCO, Sartorius) 내로 농축을 위해 모아졌다. 상기 볼륨은 10 kD MWCO의 15 ml Amicon 원심 분리기(재생된 셀룰로오스 막, Merck Millipore)를 사용하여 추가적으로 감소되었다.

[0307] 캡핑제거되고 디만노실레이션 후에, D-만니톨은 100 mM의 농도로 첨가되었고, 상기 샘플은, 실시예 4 에 기재된 바와 같은 동일한 방식으로 실행하는, 최종의 SEC 컬럼 내로 로딩하기 전에, Vivaflow 50 TFF 모듈(PES 막, 10 kD MWCO, Sartorius)에서 또 다시 농축되었다. 분획을 포함하는 생산물이 모아졌고, 농축되었다.

[0309] 실시예 8

[0310] 인공 색원체성 기질(artificial chromogenic substrate) p-니트로페닐-α-D-글루코피라노사이드(p-nitrophenyl-α-D-glucopyranoside)를 사용하여 huGAA의 세 가지 상이한 변이체(76, 95 and 110 kD 변이체)의 효소적 특성

[0311] 인공 색원체성 기질 p-니트로페닐-α-D-글리코피라노사이드(PNPG)는 실시예 2에서 수득된 처리되지 않은 huGAA (110 kDa)의 동력학적 파라미터를 측정하기 위해 사용되었고, 상기 처리된 huGAA 변이체는 실시예 7(76 kDa), 실시예 6(95 kDa) 및 실시예 4(95 kDa)에서 수득되었다. 비교는 또한 통상적인 인간 α-글루코시다아제, Myozyme®(alglucosidase alpha, Genzyme)으로 제조되었다._

[0313] 상기 효소는, 0.1 % BSA 및 100 mM AMS(반응 완충용액)을 포함하는, 100 mM 아세트산 나트륨 완충용액 pH 4.0에서 20 ㎍/ml로 희석되었다. 10 μl의 상기 효소 용액은 세 배로 96-웰 플레이트에 첨가되었다(Ten μl of the enzyme solutions were added to a 96-well plate in triplicate). 상기 PNPG 기질(Sigma)은 반응 완충용액으로 다양한 기질 농도(10, 6, 4, 2, 1.6, 1.2, 0.8, 및 0.4 mM)로 희석되었고, 90 μl의 상기 희석된 기질은 각각의 웰에 첨가되었다. 상기 효소적 반응은, 상기 반응물을 퀀칭하기 위해, 37 ℃에서 60 min 동안 배양된 다음에 100 μl 10 % 탄산 나트륨, pH 12를 첨가하였다. 상기 흡광도는 405 nm에서 측정되었다. p-니트로페놀(PNP)를 갖는 표준 곡선은 분 당 형성된 생산물의 양을 측정하기 위해 측정되었다. μM/min로 나타낸 속도는 Michaelis-Menten 곡선을 생성하는 상이한 기질의 농도의 작용으로 표시되었다. GraphPad Prism는 미하엘리스-멘텐 방정식(Michaelis-Menten equation)에 대해 직접적으로 맞춰진 것에 따라 Vmax and Km을 계산하기 위해 사용되었다. 상기 촉매 상수 kcat 및 촉매 효율 kcat/Km가 계산되었다. 상기 다양한 효소의 특정한 활성도는 U/mg로서 보고되었고, 여기서 1 단위는, 100 mM 아세트산 나트륨 완충용액, pH 4.0 + 0.1 % BSA 및 100 mM AMS에서 2 mM 기질 농도에서 분 당 1 nmol 기질의 가수분해를 촉매작용하는 효소의 양으로서 정의된다. 상기 결과는 표 1에 나타내었다.

[0315] [표 1]

	Myozyme	95 kDa (Ex. 4)	76 kDa (Ex. 7)	110 kDa (Ex. 2)	95 kDa (Ex. 6)
Vmax (μM/min)	12	12	14	13	13
Km (mM)	4.4	4.4	4.3	4.4	4.7
kcat (min^{-1})	660	677	770	688	730
kcat/Km ($min^{-1}mM^{-1}$)	150	154	179	156	155
Sp. Activity (U/mg)	2000	1910	2415	1935	1980

[0316]

[0318] Myozyme 및 huGAA의 상기 처리되지 않고 처리된 형태는, 상기 기질 PNPG에 대한 비교가능한 동력학적 파라미터를 가진다. 이는 마우스 우유 및 중국 햄스터 난소(CHO) 배지로부터의 인간 산 α-글루코시다아제를 위한 문헌에 보고된 데이터에 따른 것이다(Bijvoet *et al* (1998), *Human Molecular Genetics*, 7, 1815-1824). 상기 처리되지 않는(110 kD) 및 상기 처리된 (76 kD) 형태는 인공 기질 4-메틸움벨리페릴-α-D-글루코시다아제에 대해 동일한 Km 및 kcat 수치를 갖는 것으로 보고되었다.

[0320] 실시예 9

[0321] 기질로서 토끼 간 글리코겐을 사용한 huGAA의 상이한 변이체(76, 95 및 110 kD 변이체)의 효소적 특성

[0322] 상기 처리되지 않는 huGAA(110 kD 변이체; 실시예 2) 및 상기 처리된 huGAA 변이체(76 kDa, 실시예 7; 및 95 kD, 실시예 6)의 효소적 파라미터는, 토끼 간 글리코겐(lot N° 099K37931V, Sigma)을 사용하여 테스트되었다. 비교는 통상적인 인간 α-글루코시다아제, Myozyme®(alglucosidase alpha, Genzyme)으로 만들어졌다. 상기 효소는 100 mM 아세트산 나트륨 완충용액 pH 4.0에서 500 ng/ml으로 희석되었다. 50 μl 의 상기 효소 용액은 2 통으로(in duplicate) 96-웰 플레이트에 첨가되었다. 상기 글루코겐 기질은 아세테이트 완충용액으로 다양한 기질 농도(250, 200, 150, 100, 75, 50, 25 mg/ml)로 희석되었고, 100 μl의 상기 희석된 기질은 각각의 웰에 첨가되었다. 상기 효소적 반응은 37 ℃에서 60 min 동안 배양되었다. 글루코스의 양은 암플렉스 레드 기질(amplex red substrate)과 함께 상기 글루코스-산화효소 방법(glucose-oxidase method)을 사용하여 측정되었다.

[0324] 글루코스 표준 곡선은 분 당 형성된 생산물의 양을 계산하기 위해 측정되었다. μM/min으로 나타낸 효소 속도가 미하엘리스-멘텔 곡선을 생성하는 상이한 기질 농도의 작용으로 표시되었다. GraphPad Prism는 미하엘리스-멘텐 방정식에 대해 직접적으로 맞춰진 것에 따라 Vmax and Km을 계산하기 위해 사용되었다. 상기 촉매 상수 kcat 및 촉매 효율 kcat/Km가 계산되었다. 상기 다양한 효소의 특정한 활성도는 U/mg로서 보고되었고, 여기서 1 단위는, 100 mM 아세트산 나트륨 완충용액, pH 4.0에서 2 mM 기질 농도에서 50 mg/ml 최종 기질 농도에서 분당 1 μmol 글루코스의 형성을 촉매작용하는 효소의 양으로서 정의된다. 상기 결과는 표 2에 나타내었다.

[0326] [표 2]

	Myozyme	76 kDa (Ex. 7)	95 kDa (Ex. 6)	110 kDa (Ex. 2)
Vmax (μM/min)	32 ±6	15± 2	13±1	11±1
Km (mM)	600 ±140	100±10	93±8	162±17
kcat (min^{-1})	21100	10000	8600	7260
kcat/Km ($min^{-1}mM^{-1}$)	35	100	92	45
Sp. Activity (U/mg)	14	32	27	16

[0327]

[0329] 본 실험에서, 기질 포화는 토끼 글리코겐의 상기 제한된 용해도로 인하여 도달할 수 없었다(도 4). Myozyme에 대해, 오직 명백한 Km 및 kcat 수치가 계산되었다. 상기 세 개의 huGAA 변이체에 대해, 보다 명백한 Km 수치가 측정되었다. 상기 처리된 형태의 상기 촉매 효율은 처리되지 않는 huGAA 및 Myozyme의 촉매 효율보다 두 배 더 높았다.

[0331] 실시예 10

[0332] 폼베병의 마우스 모델에서 글리코겐 제거(glycogen clearance)에서 산 알파 글루코시다아제의 효과

[0333] 실시예 7(76 kDa, 캡핑제거되고 디만노실레이트됨), 실시예 6(95 kDa, 캡핑제거되고 디만노실레이트됨), 및 실시예 2(110 kDa, 캡핑제거되고 디만노실레이트됨)으로부터의 상기 GAA 생산물이, 골격근 및 심장으로부터의 글리코겐의 제거를 측정하기 위해 폼페병의 마우스 모델에 투여되었다.

[0335] 상기 실험은 FVB GAA 녹아웃 마우스(FVB GAA knockout mice) 및 FVB 야생형 마우스로 실행되었다. 이러한 동물 모델은, 조기-발병 유아기 형태의 상기 질병에 대해 대표적인 좋은 예이기 때문에 테스트 시스템으로서 선택되었다. 출생 이후로부터, 상기 KO 마우스는 리소좀 글리코겐(lysosomal glycogen)의 일반화되고 진행되는 축적을 갖는다(Bijvoet *et al.*, 1998, *supra*). 수컷 및 암컷 FVB GAA KO 마우스는 Erasmus University, Rotterdam로부터 수득되었다. 처리의 출발에서 마우스는 나이 26 세 내지 49 세 사이에 있다.

[0337] 상기 테스트 기질 또는 비히클은, 4 주 동안 매주 한 번 10 ml/kg 몸무게(bw)의 투여량 볼륨으로 꼬리 막(tail veil)에서 서서히 볼루스(slow bolus)에 의해 정맥내로 투여되었다. 마우스는 검시 전에 16 시간 동안 단식시켰다. 동물은 마지막 주사 후 4 일에 희생시켰다. 상기 연구 그룹의 세부사항은 표 3에 기재되어 있다.

[0339] [표 3]

그룹/컬러 코드 (color code)	투여량 레벨 (mg/kg bw)	투여량 부피 (ml/kg bw)	마우스의 타입	마우스의 N°
1/화이트	0	10	WT	9
2/블루	0	10	KO	16
3/그린	20 mg/kg 76 kDa	10	KO	16
4/레드	20 mg/kg 95 kDa	10	KO	16
5/옐로우	20 mg/kg 110 kDa	10	KO	16
6/오렌지	20 mg/kg Myozyme	10	KO	16

[0340]

[0342] 장기의 관류 및 균질화(Perfusion and homogenization of organs)

[0343] 심장 및 골격근[사두근 페무랄리스(quadriceps femuralis), 양면(both sides)]은, PBS와 관류 후에 분리되었다. 조직은 울트라-투락스(ultra-turrax)를 사용함으로써 얼음처럼 차가운 PBS의 10 중량 볼륨(10 weight volumes)에서 균질화되었다. 그 후에, 상기 세포분쇄액(homogenate)은 15 min 동안 얼음에서 16 마이크론에서 두 번 초음파 처리되었다. 12000 g에서 30 min 동안 원심분리 후에, 상청액은 글루코겐의 측정을 위해 수집되었다.

[0345] 생분석(Bioanalysis)

[0346] 각각의 개별적인 마우스의 심장 및 근육근에서 상기 글리코겐 함량은, 입증된 정성 효소적 검정(validated quantitative enzymatic assay)을 사용하여 측정되었다. 상기 조직을 끓인 후에, 아밀로글루코시다아제 및 α-아밀라아제의 혼합물은, 글루코스로 글리코겐의 분해를 위해 *생체 외에* 첨가되었다. 글루코스의 양은 상기 암플렉스 레드 기질(amplex red substrate)과 함께 상기 글루코스-산화효소 방법을 사용하여 측정되었다. 글리코겐의 양은 ㎍ 글리코겐/mg 단백질로서 표시되었다.

[0348] 통계 분석

[0349] 그룹 2, 3, 4, 5로부터의 심장에서의 글리코겐 함량은 ANOVA에 의해 분석된 다음에, 던네트 테스트(Dunnet's ttest)에 의해 그룹 6(Myozyme) 및 그룹 2(플라세보)에 대한 사후 비교(post hoc comparison)를 하였다. 그룹 1은 통계 분석에서 누락되었고(Group 1 was left out of the statistical analysis), 상기 WT 마우스 모델에서 글리코겐 저장의 결핍에 대한 품질 조사(quality check)로서 사용되었다.

[0351] 사두근 데이터에서의 비본질적인 관찰의 존재 때문에, 크루스컬-왈리스 테스트(Kruskal-Wallis test)가 상이한 생산물의 상기 글리코겐 함량 데이터의 잠재적인 차이 분포를 평가하기 위해 사용되었다.

[0353] 사두근 데이터의 사후 분석은, 윌콕슨 순위 합계 테스트로 실행되었다. 각각의 생산물 그룹 및 Myozyme 그룹은 플라세보(그룹 2) 그룹과 비교되었고, 최종 생산물 그룹은 Myozyme와 비교되었다.

[0354]

[0355] 결과

[0356] 표 4 는, 그룹 당 16 마리의 마우스의 심장(A) 및 골격근(B)에서의 평균 글리코겐 레벨(㎍/mg 단백질)을 나타낸 것이다.

[0358] [표 4]

A. 심장에서 평균 글리코겐 레벨

요약	평균(Mean)	sd
WT	0.58	0.95
KO/플라세보	525.47	67.75
KO/76 kDa	377.75	80.20
KO/95 kDa	380.56	78.30
KO/110 kDa	416.56	106.77
KO/Myozyme	475.83	98.16

B. 골격근에서 평균 글리코겐 레벨

요약	평균	sd
WT	2.22	0.66
KO/플라세보	191.80	34.75
KO/76 kDa	152.27	35.35
KO/95 kDa	169.27	46.68
KO/110 kDa	160.39	36.46
KO/Myozyme	186.49	40.61

[0359]

[0361] 도 5는, 심장(5a) 및 골격근(5b)에서 개별적인 마우스의 상기 글루코겐 레벨(㎍/mg 단백질)을 나타낸 것이다. 상기 결과는, 본원에 생산된 상기 GAA 생산물(110 kDa, 95 kDa, 및 76 kDa)이 20 mg/kg에서 4 번의 정맥내 주사 후에 플라세보-처리된 마우스와 비교된 심장에서 글리코겐 레벨이 통계학적으로 감소됨을 나타내었다. 상기 동일한 Myozyme® 투여량은 상기 심장에서 글리코겐의 양을 감소시키지 않는다. 76 kDa 생산물 및 95 kDa 처리된 그룹 둘 다 에서 상기 글리코겐 레벨이 Myozyme®-처리된 그룹과 비교하여 통계적으로 상이하였다. 통계학적으로, 본원에 생성된 세 개의 상이한 GAA생산물 사이의 어떠한 차이도 없었다.

[0363] 본원에 생산된 상기 76 kDa 생산물은 또한, 플라세보-처리된 또는 Myozyme®-처리된 마우스와 비교된 상기 골격근에서 글리코겐의 양을 통계학적으로 감소시켰다. 95 kDa 및 110 kDa 생산물 둘 다에서 상기 글리코겐 레벨은, 상기 개별적인 마우스 사이의 더 높은 변화로 인한 가능성으로, 플라세보 및 Myozyme®-처리된 마우스와 비교되는 통계학적으로 다르지 않았다(The glycogen levels in both the 95 kDa and the 110 kDa product were not statistically different compared to placebo and Myozyme®-treated mice, likely due to a higher variation between the individual mice). 20 mg/kg에서 Myozyme®는 플라세보와 비교된 골격근에서 상기 글리코겐 레벨을 감소시킬 수 없었다.

[0365] 실시예 11

[0366] *아스페르길루스 오리제*로부터 프로테아제의 확인

[0367] GAA는, 산성의 pH에서, *아스페르길루스 오리제*로부터의 프로테아제 혼합물, 낮은 양의 Flavourzyme (Novozymes Corp)과 함께 배양에서 특정한 단백질분해의 분열을 받는다. 상기 결과적으로 생성된 GAA 생산물은, 환원 조건(reducing conditions) 하에서 SDS-PAGE에서 대략 95 kD의 분자량을 갖는다. 유사한 단백질분해 활성도는, 상기 생산 스트레인(*야로위야 리폴리카*)로부터 백그라운드 단백질을 포함하는 특정한 부분적으로 정제된 GAA 조제물에서 관찰되었다.

[0369] 상기 단백질분해 사건을 측정하기 위해, GAA의 상기 N-글리칸은, 단백질분해 처리(proteolytic treatment) 전에, EndoH를 사용하여 N-글리코실화 부위 당 단일 N-아세틸 글루코사민으로 제거되었다. 이는 SDS PAGE를 통해 보다 충분한 평가를 가능하게 한다. 상기 GAA 생산물은 그리고 난 다음에, 이의 정제된 샘플에서 Flavourzyme 프로테아제 칵테일과 함께 배양되었다. 상기 반응은 100 mM 아세트산 나트륨 완충용액 pH 5에서 30 ℃에서 실행되었다. 샘플은 상이한 시점에서 채취되었고, 환원 조건 하에서 SDS-PAGE를 통해 분석되었다. 0.5 ㎍의 GAA를 포함하는 부피는 레인 당 로딩되었다(Volumes containing 0.5 ㎍ of GAA were loaded per lane).

[0371] 상기 프로테아제 칵테일에서 GAA의 특정한 단백질분해(proteolysis)에 책임이 있음을 조사하기 위해, 프로테아제 저해제는, 상기 프로테아제의 확인을 가능하게 하도록 정의된 프로테아제 패밀리에 대해 특정한 것인 상기 검정에 포함되었다. 상기 반응은, 프로테아제 저해제가 상기 반응 혼합물에 현재 첨가되는 것을 제외하고, 상기에 기재된 바와 같이 실행되었다. 상기 비가역적 저해제 PMSF(Sigma-Aldrich prod. nr. E5134-500G) 및 E-64 (Calbiochem prod. nr. CALB324890-5)는,단백질분해 반응 전에, 각각 1 mM 및 10 μM의 농도로 상기 희석된 프로테아제 칵테일과 함께 배양되었다. 상기 비가역적 저해제 키모스테틴(Calbiochem prod. nr. CALB230790-5), EDTA, 및 류펩틴(leupeptin)(Calbiochem prod. nr. CALB108976-10MG)는, 각각 60 ㎍/ml, 50 mM 및 100 μM의 농도에서 상기 반응 혼합물에 직접적으로 첨가되었다.

[0373] GAA의 상기 특정한 단백질분해는, 세린 및 시스테인 프로테아제의 활성을 파괴하는, 프로테아제 저해제, PMSF 및 키모스타틴(chymostatin)에 의해 저해되었다(The specific proteolysis of GAA was inhibited by PMSF and chymostatin, protease inhibitors that abolish the activity of serine and cysteine proteases). 시스테인 프로테아제를 저해하는, 비가역적 저해제 E-64는, 단백질분해를 차단하지 않는다. 이러한 데이터는, 상기 특정한 프로테아제가 세린 프로테아제 패밀리 멤버(serine protease family member)임을 제시한다. 이러한 가설을 뒷받침하는 더 많은 증거는, 상기 프로테아제 칵테일이, 이황화 결합을 감소시키는 산화환원제 디티오테이톨(dithiotheitol)(DTT) 및 PMSF와 함께 배양되는 추가적인 검정에 의해 제공되었다. 이러한 환원제의 첨가는, 상기 시스테인 프로테아제 활성을 회복하는, 공유결합 비활성 시스테인 프로테아제: PMSF 어덕트(adduct)를 환원시킨다(Addition of this reducer reduces the covalent inactive cysteine protease:PMSF adduct, restoring the cysteine protease activity). PMSF에 의해 저해된 경우에, 세린 프로테아제의 상기 활성도는 DTT에 의해 회복되지 않을 수 있다. 행동에서 이러한 차이는, GAA에서 세린 및 시스테인 프로테아제 행동 사이의 추가적으로 구별하는데 사용되었다.

[0375] DTT와 상기 PMSF-저해된 프로테아제의 배양은, 상기 프로테아제 칵테일의 상기 GAA-특정한 단백질분해 활성도를 회복하지 않는다. 상기 GAA-특정한 단백질분해는 또한, 메탈로프로테아제 저해제(metallo-protease inhibitor) EDTA 및 광범위 스펙트럼 저해제 류펩틴(broad spectrum inhibitor leupeptin)에 의해 저해되지 않는다. 세린 프로테아제가 이러한 GAA 단백질분해 사건에 대해 책임이 있음을 나타낸다.

[0377] 상기 혼합물로부터 상기 프로테아제를 확인하기 위해, 상기 프로테아제는 일련의 크로마토그래피 단계를 사용하여 정제되었다. 상기 첫 번째 크로마토그래피 단계는 음이온 교환 크로마토그래피 레진(anion exchange chromatography resin)(Q-Sepharose FF, GE healthcare)이 사용되었다. 상기 프로테아제 칵테일 물질은, 로딩 전에 20 mM TRIS-HCl 완충용액 pH 7 에서 희석되었다. 20 mM TRIS-HCl 완충용액에서 100 mM, 300 mM 및 500 mM NaCl에서의 용리액 및 흐름(flow through)이 수집되었다. 모든 흐름 및 용리 분획은 상기에 기재된 바와 같이 상기 검정을 사용하여 분석되었다. GAA에서 작용하는 상기 프로테아제는 운행의 흐름 분획에 존재하였고, 출발 물질과 비교하여 현저하게 풍부하였다(The protease acting on GAA was present in the flow-through fraction of the run and was significantly enriched compared to the starting material).

[0379] 상기 흐름 물질(flow-through material)은 양이온 교환 크로마토그래피[SP sepharose XL (GE Heathcare) at pH

5 10mM Na Acetate; elution with 0-300 mM NaCl]를 통해 추가적으로 처리되었다. 용리 분획(Elution fractions)이 수집되었고, 즉각적인 블루 염색된 SDS PAGE를 통해 분석되었고, 상기에 기재된 바와 같은 상기 검정을 사용하여 관심의 프로테아제의 존재를 위해 분석되었다.

[0381] 상기 활성도의 대부분은, 상기 CEX 크로마토그래피 용리액의 마지막 분획에 존재하였다. 상기 마지막 두 번째 분획이 모아졌고, 하기와 같이 질량 분석법을 통해 분석되었다. 상기 단백질 혼합물은, 트립신 소화 전에 탈염되었고, 환원되었고, 알킬화되었고, 그 다음에 LC-MS/MS 방법을 하였다(The protein mixture was desalted, reduced and alkylated prior to trypsin digestion and subsequently subjected to an LC-MS/MS methodology). 획득된 스펙트럼은 Mascot 알고리즘을 사용하여 상기 NCBI 데이터베이스 상에 매치되었다. 하기의 세팅이 적용되었다 :

[0382] ● 트립신, 키모트립신[허용된 4 까지의 불완전한 분열(up to 4 miscleavages allowed)]

[0383] ● 산화(M,W), 탈아미드화(deamidation)(N,Q)(가변적인 변경)

[0384] ● 카르바미도메틸레이션(Carbamidomethylation)(고정된 변경)

[0385] ● 분류(Taxonomy): 진핵생물

[0386] ● MS 저항력(tolerance): 0.05 Da, MS/MS 저항력: 0.05 Da

[0387]

[0388] *아스페르길루스*로부터의 알칼리성 프로테아제(GenBank Accession No. BAA00258.1; gi 217809)는 상기 연구로부터 확인되었다. 상기 성숙한 프로테아제의 서열은 하기와 같다:

[0389] >gi|217809|dbj|BAA00258.1| 알칼리성 프로테아제[아스페르길루스 오리제]

[0390] GLTTQKSAPWGLGSISHKGQQSTDYIYDTSAGEGTYAYVVDSGVNVDHEEFEGRASKAYNAAGGQHVDSIGHGTHVSGTIAGKTYGIAKKASILSVKVFQGE
SSSTSVILDGFNWAANDIVSKKRTSKAAINMSLGGGYSKAFNDAVENAFEQGVLSVVAAGNENSDAGQTSPASAPDAITVAAIQKSNNRASFSNFGKVVDVF
APGQDILSAWIGSSSATNTISGTSMATPHIVGLSLYLAALENLDGPAAVTKRIKELATKDVVKDVKGSPNLLAYNGNA (SEQ ID NO:8).

[0392] *A. 오리제*로부터의 상기 정제된 프로테아제의 SDS-PAGE 겔 분석은, 약 30 kDa(성숙한 프로테아제)의 MW에서의 밴드 및 20과 10 kDa 사이의 MW를 갖는 몇몇의 밴드의 존재를 나타낸다. 상기 보다 낮은 MW 밴드는 겔로부터 잘라졌고(excised), 트립신 소화되고, 나노-LC-MS/MS에 의해 분석되었다. 이러한 밴드는, *A. 오리제*로부터의 상기 알칼리성 프로테아제를 나타내는, *A. 오리제* 알칼리성 프로테아제로부터의 생산물로서 확인되고, 자가단백질분해(autoproteolysis)된다(These bands were identified as products from the *A. oryzae* alkaline protease, indicating the alkaline protease from *A. oryzae* is susceptible to autoproteolysis).

[0394] 실시예 12

[0395] *야로위아 리폴리카*에서 *아스페르길루스 오리제* 프로테아제의 발현

[0396] 본 실시예는, 상기 성숙한 단백질 ALP를 발현하는, *Y. 리폴리카*의 구성을 기재하였다. *아스페르길루스* 오리제로부터의 상기 알칼리성 프로테아제(ALP)를 코딩하는 상기 유전자(EC. 3.4.21.63)는, *Y. 리폴리카* 발현에 최적화된 코돈이고, 융합 구성물로서 화학적으로 합성되었다. 상기 융합 구성물은, 단일 펩티드(21 개 아미노산), 프로-펩티드(pro-peptide)(100 개 아미노산) 및 성숙한 단백질(282 개 아미노산)을 포함하는 효소의 전체적인 오픈 리딩 프레임(ORF), 그 다음에 링커(SGGG) 및 His 태그(Tag)(10x His 잔기)를 코딩한다[The fusion construct encoded the entire open reading frame (ORF) of the enzyme including signal peptide (21 amino acids), pro-peptide (100 amino acids) and mature protein (282 amino acids) followed by a linker (SGGG) and a His Tag (10x His residue)]. 도 9를 참고하라. 상기 융합 구성물(fusion construct)의 상기 완전한 뉴클레오티드 서열은 도 10에 기재되어 있다.

[0398] ALP의 상기 합성의 ORF는, 상기 발현 카세트(expression cassette)의 안정한 통합을 위해 상이한 로커스

(different loci)를 이용하여, *Y. 리폴리카* 게놈 내로 타겟된 통합을 위해, BamHI/AvrII 단편으로서, 상기 pPT 벡터 시리즈 내로 클론되었다. 상기 pPT 벡터에서, 상기 박테리아 모이어티(bacterial moiety)는 상기 플라스미드 pHSS6로부터 유도되었고, 복제의 박테리아 오리진(bacterial origin of replication, ori) 및 카나마이신에 대한 저항을 부여하는 상기 카나마이신-저항 *유전자*(KanR)를 포함한다. 상기 통합 카세트는, **a)** *Y. 리폴리카*에 대한 형질전환을 위해 선택가능한 마커(*URA3; LEU2; ADE2*), **b)** 프로모터로 구성된 상기 발현 카세트(POX2; Hp4d), **c)** 상기 ALP 합성 구성물을 삽입하기 위한 다수의 클로닝 부위(MCS) 및 **d)** 상기 *YILIP2* 유전자의 종결(terminator)를 포함한다. 상기 통합 카세트는, 상동 재조합에 의해 Y. 리폴리카 내로 안정한 단일 카피 타겟된 통합에 대한 특정한 로커스의 업스트림(upstream)(P) 및 다운스트림(T)의 측면에 있다(The integration cassette is flanked by upstream (P) and downstream (T) sequences of a specific locus for stable single copy targeted integration into *Y. lipolytica* genome by homologous recombination). 두 개의 NotI 제한 부위(NotI restriction sites)가 상기 박테리아 모이어티의 통합을 방지하기 위해 형질전환 전에 상기 발현 카세트의 분리를 가능하게 한다.

[0400] *Y. 리폴리카*에 사용된 상기 배지 및 기술은, Barth and Gaillardin(*FEMS Microbiol Rev.*, 19(4):219-37, 1997)에 기재되어 있고, 효모 세포는, *E. coli*에 대해 사용된 표준 기술 및 1 μg의 정제된 통합 카세트를 사용하여, Le Dall *et al*. (*Curr Genet.*, 26(1):38-44, 1994)에 기재된 리튬 아세테이트에 의해 형질전환되었다.

[0402] 상기 발현 카세트 ALP의 상기 통합은 하나의 프리 로커스, 및 삽입이 상기 발효 공정 동안에 원하지 않는 높은 분비 단백질(리파아제 2 및 리파이제 8)의 발현의 제한을 제공하는 사실을 기초로 2 개의 특정한 로커스에서 실행되었다. 상기 최종의 스트레인 OXYY2184는 반-구성요소 Hp4D 프로모터에 의해 유도된 ALP의 3 개의 발현 카세트를 포함한다.

[0404] OXYY2184는, 평균으로 약 2 내지 2.5 g/L 발효 브로스(fermentation broth)를 수득하는, 상기 재조합 *아스페르르길루스 오리제* ALP 성숙한 형태(35 kDa)를 생산한다. 전체 단백질은 Bradford 기술을 사용하여 검정되었고, 상기 프로테아제 활성도는 기질로서 아조카세인(azocasein)과 함께 검정을 사용하여 측정되었다. 프로테아제는 프리 아조 염료 및 카세인에 대한 아조카세인을 분해한다(Proteases digest the azocasein towards casein and the free azo dye). 상기 분해된 단백질의 정제 및 원심분리는, 프리 아조 염료가 상기 단백질분해 활성의 지표(indication)인, 알칼리성 조건에서 측정되는 것을 가능하게 한다. 이러한 생성물의 흡광도는 OD 440 nm에서 측정되었다. 소화된 아조카세인의 양은, 흡광도가 OD 440 nm에서 측정되는, 공지된 농도와 함께 아조카세인 희석 시리즈(azocasein dilution series)와 연관성에 의해 계산될 수 있다.

[0406] 스트레인 OXYY2184의 상기 배양 상청액에서 ALP는 SDS-PAGE에 의해 분석되었고, 항-His 폴리클로날 항체와 함께 면역검출되었다(immunodetected). *Y. 리폴리카*에서 생산된 상기 재조합 ALP는 활성되고, 상기 정제된 천연의 효소(purified native enzyme)로서 유사한 특성을 갖는다. 상기 재조합 ALP의 이러한 효소 특성은, 95 kDa rhuGAA 형태를 수득하기 위해 상기 rhuGAA 전구체를 처리하도록 이의 사용을 가능하게 한다.

[0408] 스트레인 OXYY2122는, ALP 및 rhuGAA를 공동-발현하도록 구성되었다. 상기 ALP 발현 카세트의 하나의 복사본은 상기 rhuGAA를 발현하는 수용체 스트레인(recipient strain) 내로 통합되었다(rhuGAA의 4 개의 복사본). huGAA 및 ALP를 코딩하는 둘 다의 유전자는 상기 유도성 POX2 프로모터 하에서 유도되었다. 상기 결과적으로 생성된 스트레인 OXYY2122는 상기 rhuGAA 전구체(110Kda)와 함께 ALP의 상기 성숙한 형태를 생산한다. 스트레인 OXYY2122의 상기 배양 상청액에서 재조합 huGAA는 SDS-PAGE에 의해 검정된 다음에, 면역학적 블로팅(immunoblotting)을 하고, 상기 rhGAA는 상기 상청액에서 상기 95 kDa 형태로 처리됨을 확인하였다. 이러한 공정이 완료되었다: 어떠한 110 kDa 형태도 검출되지 않은 반면에, ALP 없이 상기 스트레인의 상기 동일한 배양에서 어떠한 처리도 발생하지 않았다(This processing was complete; no 110 kDa form was detected, whereas in the same cultivation of the strain without ALP no processing occurred).

[0410] 실시예 13

[0411] *야로위아 리폴리카*에서 발현된 *아스페르길루스 오리제* 알칼리성 프로테아제로 rhGAA 발효 브로스의 처리 후에 수득된 95 kDa rhGAA의 정제

[0412] rhGAA의 상기 95 kDa 형태는 하기와 같이 스트레인 OXYY1589로부터 분리되었다. 재취한 후에, 상기 브로스는 세라믹 막(ceramic membranes)(Pall Corporation)을 사용하여 투명해졌다. 상기 생산물은, 10 kD의 분획분자량(molecular-weight-cut-off, MWCO)으로 중공 섬유 막(hollow fiber membranes)을 통해 농축되었다. AMS 는 1 M의 농도로 첨가되었고, 상기 용질은 여과 전에 30 ℃로 가열되었다. 상기 여과액(filtrate)은 *야로위야 리폴리카*(스트레인 OXYY2184)에서 재조합적으로 발현된 *A. 오리제* 알칼리성 프로테아제로 처리되었고, 어떠한 추가적인 정제 없이 상기 발효 브로스의 정화(clarification) 후에 사용되었다. 30 ℃에서 16 h 동안 배양 및 전체 단백질 : 프로테아제에 대한 200 : 1의 중량:중량 비율이 95 kDa 생산물에 대해 완전한 단백질분해(full proteolysis)를 결과적으로 나타내었다.

[0413]

[0414] 추가적인 정제 및 상기 인산화된 N-글리칸의 캡핑제거 및 디만노실레이션 후의 분석은, 표 1 에 나타낸 바와 같이 PNPG 에서 유사한 특정한 활성도를 갖는 95 kDa GAA 생산물(SDS-PAGE에서 관찰된 바와 같음)을 나타내었다.

[0416] 실시예 14

[0417] 아스페르길루스 오리제 알칼리성 프로테아제(ALP)로 처리한 후 rhGAA에서 상기 단백질분해 분열 부위의 확인

[0418] rhGAA는 상기 아스페르길루스 오리제 ALP로 처리되었고, 상기 실시예에 기재된 바와 같이 추가적으로 정제되었다. 서열 분석을 용이하게 하기 위해, 상기 정제된 샘플은, PNGase F가 아스파르테이트(aspartate)에 대해 서열에서 N-글리코실화된 아스파라긴 잔기를 결정하는 바와 같이, 상기 rhGAA를 디글리코실화하기 위해 PNGaseF로 처리되었다(To facilitate sequence analysis, the purified sample was treated with PNGaseF to deglycosylate the rhGAA as PNGase F deaminates the N-glycosylated asparagine residues in the sequence to aspartate).

[0420] rhGAA의 서열을 확인하기 위해, 상기 디글리코실레이트된 단백질(deglycosylated protein)은 트립신을 사용하여 소화된 다음에, 이황화물 다리(disulfide bridge)를 환원하고 상기 시스테인 잔기를 알킬화반응하였다. 상기 결과적으로 생성된 펩티드 혼합물은 LC-MS 및 MS/MS 하였고, 상기 데이터는 상기 유전자-코딩된 단백질 서열 내로 일치시켜, 이로 인하여 동일성(identity)을 결정한다. 정확한 질량(<10 ppm) 및 분할 스펙트럼(fragmentation spectra)은 완전한 확인을 위해 사용된 표준이다.

[0422] 거의 완전한 서열 범위는 상기 펩티드 혼합물(잔기 23-60, 65-535, 및 538-898)로부터 수득되었고, 상기 단백질분해 분열 부위는, 아미노산 60에서 65 사이(SEQ ID NO: 1에 따른 서열 넘버링)인 것으로 밝혀졌다(determined). 잔기 60과 65 사이의 상기 rhGAA 서열에서의 갭(gap)은, Gly62 및/또는 Gly65 전에 단백질분해 사건으로부터 기인할 수 있었다(The gap in the rhGAA sequence between residues 60 and 65 could result from a proteolytic event before Gly62 and/or before Gly65). *아스페르길루스 오리제*로부터의 상기 알칼리성 프로테아제가 Leu과 Gly 사이의 상기 합성 펩티드 Ileu-Gln-Asn-Cys-Pro-Leu-Gly-NH2를 분해하도록 문헌에 보고되어 있다(Nakadai *et al.*, 1973, *Agr. Biol. Chem.*, 37, 2685-2694를 참고하라).

[0424] 이러한 실험에서 결정된 상기 단백질분해 분열 부위는 리소좀에서 관찰된 GAA의 단백질분해 공정에 따른 것에 있다. 95 kDa 폴리펩티드에 대해, 상기 분열 부위가 아미노산 59와 아미노산 58 사이(SEQ ID NO: 1에 따른 서열 넘버링)에서 확인된, Moreland et al., 2005, *J. Biol. Chem.*, 280, 6780-6791를 참고하라. 상기 분열된 N-말단 펩티드는 연관된 체인간 이황화물 결합을 통해 유지되었다(The cleaved N-terminal peptide remains associated via an interchain disulfide bond).

[0426] **그 밖의 실시형태**

[0427] 본 발명은 이의 상세한 설명과 함께 기재하긴 하였지만, 상기에 나타낸 상세한 설명은 본 발명을 설명하기 위한 의도이며, 이로 본 발명의 범위를 한정하지 않으며, 이는 청구의 범위의 범위에 의해 정의되어 있다. 그 밖의 측면, 장점 및 변형은 하기의 청구의 범위의 범위 내에 있다.

도면

도면1

AAQQGASRPGPRDAQAHPGRPRAVPTQCDVPPNSRFDCAPDKAITQEQCEARGCCYIPAKQGLQ
GAQMGQPWCFFPPSYPSYKLENLSSSEMGYTATLTRTTPTFFPKDILTLRLDVMMETENRLHFTIKD
PANRRYEVPLETPHVHSRAPSPLYSVEFSEEPFGVIVRRQLDGRVLLNTTVAPLFFADQFLQLSTSLPS
QYITGLAEHLSPLMLSTSWTRITLWNRDLAPTPGANLYGSHPFYLALEDGGSAHGVFLLNSNAMDV
VLQPSPALSWRSTGGILDVYIFLGPEPKSVVQQYLDVVGYPFMPPYWGLGFHLCRWGYSSTAITRQ
VVENMTRAHFPLDVQWNDLDYMDSRRDFTFNKDGFRDFPAMVQELHQGGRRYMMIVDPAISSS
GPAGSYRPYDEGLRRGVFITNETGQPLIGKVWPGSTAFPDFTNPTALAWWEDMVAEFHDQVPFD
GMWIDMNEPSNFIRGSEDGCPNNELENPPYVPGVVGGTLQAATICASSHQFLSTHYNLHNLYGLTE
AIASHRALVKARGTRPFVISRSTFAGHGRYAGHWTGDVWSSWEQLASSVPEILQFNLLGVPLVGAD
VCGFLGNTSEELCVRWTQLGAFYPFMRNHNSLLSLPQEPYSFSEPAQQAMRKALTLRYALLPHLYTL
FHQAHVAGETVARPLFLEFPKDSSTWTVDHQLLWGEALLITPVLQAGKAEVTGYFPLGTWYDLQTV
PVEALGSLPPPPAAPREPAIHSEGQWVTLPAPLDTINVHLRAGYIIPLQGPGLTTTESRQQPMALAV
ALTKGGEARGELFWDDGESLEVLERGAYTQVIFLARNNTIVNELVRVTSEGAGLQLQKVTVLGVATA
PQQVLSNGVPVSNFTYSPDTKVLDICVSLLMGEQFLVSWC* (SEQ ID NO:1)

도면2a

>DsbA-6xHis-CcMan5 (107bp - 5167bp, direct) 5061bp From pLSAHCcMan5

ATGAAAAAGATTTGGCTGGCGCTGGCTGGTTTAGTTTTAGCGTTTAGCGCATCGGCCGGCCATCACCATCATCACCA
CGTGGGGCCCGGCTCGGACGAAGTGGATGCACCGGAACCTCCGAGCGCAGATTATGCAAGCCTGGTTGATGTTTTT
GTTGGCACCGAAGGTGATTTTGGTAATGATATGCCTGCAGCACAGGCACCGAATGGTCTGGCAAAAGTTAATCCGC
GTACCACACCGGGTCGTAATAATACCGGTTATGATTATGCCCAGAGCAAAATTAGCGGTTTTACCCATACCAATCTG
GATGGTGTTGGTGGTAGCGGTGGTGGTGGTGATCTGCTGGTTGTTCCGACCAGCGGTAGCTATACCGCACGTCCGG
GTACAGGCACCTATGCACATCCGTTTAGCCATGATGATGAAGATGCAGGTCCGGGTTTTTATAGCGTTGGTCTGGGT
AATGTTGCAGGCACCGATGGTGCAATTACCGGTGCTCCGGGTACAATTGAAGCAGAAGTTGCAGCAGCAACCCGTA
GCGGTGTTCATCGTTATGCATTTCCGGCAGGTAGCACCCCGAGCCTGGTTGTTGATCTGGAAACCAATAATACCAGC
CGTCGTAGCAGCAGCGTTCAGGTTGAAACCCGTGCAGATGGCACCGTTGAACTGAGCGGTCAGGTTACCGGCTATT
TTTATAATGCAGCCTATACCCTGTATTATACCGCACGCACCCTGCAGCCTGCAACCGTTCAGACCTGGGGTGATGATG
ATCGTCTGGTTGATGCAACCGCACAGGATGGTGTTGATACCGGTGCAATTCTGACCTTTGATCCGGCAGATGCCGGT
GAAATTGGTCTGCAGGTTACCCTGTCTCCGGTTAGCGTTGAACAGGCACGTATTGATCAGCAGGTTGAACTGGGTG
ATCTGAGCTTTGATGCAATTCGTGATCGTACCCGTGCAGAATGGAATGCAACCCTGGGTCGTGTTGCAATTGATGCA
AGCACCGCAACCGATCCGACCGGTGAACTGCAGCGTCTGTTTTATACCCATCTGTATCGCATGTTTGCAATGCCGAT
GAATGCAACCAGCACCAGCGGCACCTATCGTGGTGTTGATGGTGCAGTTCATGCAGCACAGGGCTTTACCTATTATG
ATAGCTGGGCAACCTGGGATGATTTTCGCAAATTTAGCGTGATTGCCTATATTGATCCGGCACTGTATCGTGATATG
GTTCAGAGCCTGGTTTACCTGTTTGCAGATGCAGAAGCAACCGGTACAGGCGGTGGTCTGGGTGGTTTTGTTCATAG
CGTTCCGACCGTTCGTTGGGAACGTAGCAGCGTTGTTGTTGCAGATGCAATTGCCAAAGGCTTTGATGGTTTTGATC
GTCTGGATGAAGCATATCCGGCACTGCAGCGCCTGGTTGGTCAGTATAGCGCAGATGAACTGCGTCGTGGTTATGT
TGCAGGTAATCCGGGTGCAAGCGTTCAGCGTGGTTATGATCAGTATGGTCTGAGCGTTATTGCCGATGAACTGGGT
CTGACCGAAGAAGCAGAAACCCTGCGCGAACAGGCAAGCTGGCCGATTGAAAAACTGACCAAACCGGGTGCATGG
ACCGCAGCAGATGGTACACAGGTTGGTCTGCTGACACCGCGTGCAGCCGATGGTAGCTGGCAGAGCGCAGATCAT
GCCAAATTTGAAGCAGCAGGTCTGTATCAGGGCACCCTGTGGCAGTATCATTGGTATGATGCCTATGATATGGATGC
ACTGGTTGAAGCAATGGGTGGTCATGAAGCAGCCCGTCTGGGTATGCGTCATATGTTTGGTGAACATGCACCGGAT
GATGGTAAAGCAATGCTGCATAGCAATGCCAATGAAATTGATCTGCAGGCACCGTACCTGTTTAATTATACCGGTGA
ACCGAGCCTGACCCAGAAATGGGCACGTGCAATTTATACCAAAGAAACCTGGAATCGCTATATTGCAACCGGTAGC
AGCTCTGCAGTTCCGTCAGGTGGTGGTGAATTTACACCTCCGCTGAAAACCAAAGTTTATCGTCTGGACCCTCGTGG
TATGCTGCCGACCATGGATAATGATGCAGGTACAATGAGCACCATGTTTGTTGCAGCAGCCGTTGGTCTGTTTCCGG
TTACCGCAGGTAGCAGCCAGTTTCAGGTTGGTAGCCCGTTTTTTGATAGCACCACCATTACCTATGATGATGGTAGC
GCATTTACCGTTACCGCAGATGGTGTTAGCGAAGATGCCTTTTATGTTCAGAGCGCAACCCTGGATGGTGCAACCTT
TGGTAATACCTGGGTTGATTATGCAACCGTTGTTGGTGGTGCAGATCTGGCATTTCGTATGGGTGAACAGCCGAGC
GATTGGGGCACCGATACCGCACCGGCATTTAGCATGAGCACCGCCACCGATGAACCGGCAGAAGGTCCTCGCGTTA
GCGCAGAACCGACCACCGTGCAGACCGGTGATGGTGGTGCACTGGATGCAACCGTTACCCTGACACTGGATGGCGC
ACGTCTGGCAGCACCGGCAGGTACAGATCTGGTTACCAGCGGTGCAGCAAGCGTTGTTGGTCTGCCGGATGGTGTT
ACCGCAGCAGTTACCGTTGCAAGCCCGACCGCACTGACCGTTAGCCTGACCGGCACCGCATCAGCAGATGCACGTTT
TTTTGTGCATCTGCGTGATGCAGCACTGGCCGATGGTGTTGCAGCCGCAAGCCTGCAGGGTCAGGGTGTTAGCGTT
CGTTCTCCGCTGCGTCTGAGCGTTGCAAGCGCAGAACGTGATGCACTGGCAGCACTGGTTGATGATGCCGTTCTGGT
TCGTCATGGTAATTATAGCAGCGTTACCTTTGATCGTTTAGCACCGCTCTGACAAAAGCACAGGAAGCACTGGGCGA
CGAAGCAGCAACCAGCATTGCACTGCGTTTTGCAGCAGATCGTCTGGGTGCAGCAGCAGATGCACTGGATCTGACC
GGTGGTGGTTATCGTACCCTGGAAGCAGAACAGAGCGAAGCATGGTCTGGTGGTGAACTGAAAAATGAAGCCAAT
AGCAGCAGCGGTAATCTGGGTGGTGTTCGTAGCGGTAGCTGGGTTCAGTATCGCGATATGACCTTTGAAACCGCAG
CCGGTGATACACCTCCGCGTTTTCTGACCGTTCGTTATGATACCAGCTTTGCACCGACCGATACCCCGAGCACCGTTC
GTGTTCATGCCGGTGATGTTTCTGGTCCGGTTGTTGCAACCGTTGATCTGAAAGGCACCAGCGGTTGGGGTAAATAT
ACCGAAGTTACCGCAGAACTGGGTGATGTTCAGGCCCTGGTTGATGCCCAGGTTGTTACCTTTGAACTGCTGGCACC
GAGCGGTCGTAGCTGGGTTGGTAATTTTGATTGGTTTCGCTTTAGCGCAGAAGATCCGGCAGCACCGGGTCAGCCT
GGTGAAAGCCCGACCGTTACCATTGAAGCCGAAGATTGGACCGCAAGCAGCGGTCGTGGTCTGAAAAAAGAAAGC
AGCACCTGGACCAGCGGTCCGGTGACCAATGTTGGTGGTACAGCAGATGGTGATTGGATTGCCTATGGTGAAGTTG
ATCTGGGTGAACTGCCGCTGGGCGAACTGAGCGTTCATTATGTGCATAATAGCAATCGCAGCGGTAATAATAGCGC
ACTGAGCGTTTATCTGGATGCATTTGATCCGGCTAATCCGGGTGAACCGTTTGTTACCGTTCCGCTGCCGACCACCG
GTAGCAGTTGGACCGCAGATGGCACAGCCACCGTTGTTCTGCCGGAAACCGTGCAGGGCACCCATGAAGTTTTTGT
TCGTCTGAGCACCGAACCGTATGCAGATCATCCGTATGTTGCAAATCTGGATAGCCTGACCTTTGCACCGGGTGGTC
CGACCAGCGTTGTGGTTGAAAGCGAAGCCTGGACCAGCAATTCTGGTCGTGGCCTGAAAAATGAATCTTCTACCTG
GACCTCTGGTCCGGTTACAAATGTGGGTGGCACCGCTGATGGCGATTGGCTGGCATATGGCGAAATTGATCTGGGC
AGCGCAGCACTGGATCAGCTGTCTGTGCATTATGTTCATAATTCTAATCGCTCTGGTCGTAATTCTGCACTGTCTGTG
TATCTGGATGCCTTTGATCCGGCAAATCCGGGTGAACCGTTTGTGACAGTGCCGCTGGCAAATACCGGTAGCTCTTG
GACCACCGATGGTACTGCAGTTGTGGATCTGCCGTCTACCGTTCGTGGTAAACATCAGGTTTGGGTTCGTCTGTCTA
CCGAAGCATATGCCGATCATCCGTATGTGGCCAATCTGGATTCTATGCGCTTTTTTACCGATGCATATGATGTTGAAG
TTCCTCCGACCGATACAGCAGCACTGGCAGCCGTTGTTGATGCAGCAGGTACACCGGAAGCAGAAATTGCACGTTAT
GGTCGTATTGATGCCCGTGTTTTTACCCGTGAACTGGCAGCAGCACGTAGCGTTCTGGCCGATGCCGGTGCAACACA
GGCACAGGCAGATGAACGTGCTCGTCGTCTGGGTCTGGCAACCGATCAGCTGGTTCCGGCAGAACGTCGTCGTCTG
GAAAATCTGGTTGCCAGCGCAGAAGCACTGACCGACGAAGGTTATTCTCCGGAAAGCTGGCAGGCATTTCGTACCG
CACTGGCTGCTGCAACCGGCACCCTGGATGATGCAGCAGCATCTGATGAAGCACTGCATGATGCACGTCTGGCGCT
GCAGGGTGCAGTTGATGCACTGGAAGAACCGGCAGATGTTGTTCTGGTTGAAGTTGAAGTTTCTCCGCGTTGTCTG
GCAGGTAAACCGTATGTTGCCGTTCGTGCAGTTAATGTTTCTGATGCAGCCGTTGATGTTGAACTGGCAAGCTCTCT
GGGCACCCGTAGCTTTGTTGGTGTGGCACCGGGTGCGAGCGCATATCAGAGCTTTGCAGCCCGTAGCGCAACCGGT
GATCTGGATGTTACCGTGACCGCAACCGGTGCAGATGGTACTCAGACCGTTGAACAGGTTGTGACCGTTCCGAGCT
GTAGCTAATAA (SEQ ID NO:2)

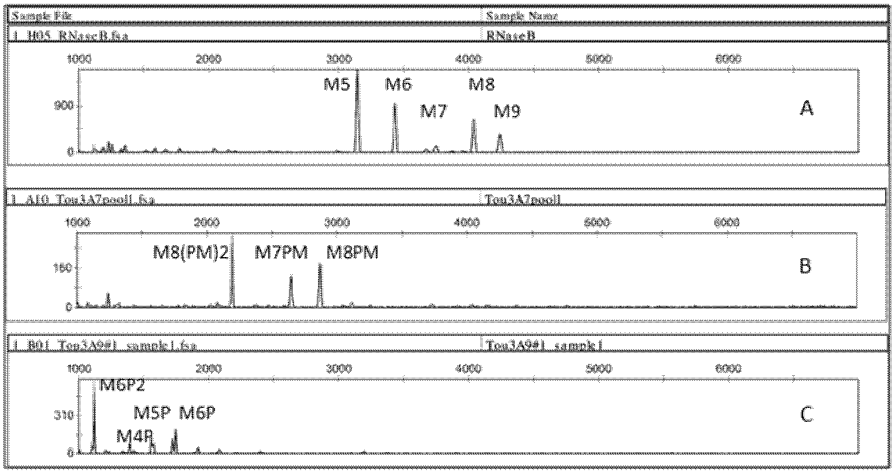
도면2b

ALAVVGLAPATAASAAPEPPSADYASLVDVFVGTEGDFGNDMPAAQAPNGLAKVNPRTT
PGRNNTGYDYAQSKISGFTHTNLDGVGGSGGGGDLLVVPTSGSYTARPGTGTYAHPFSH
DDEDAGPGFYSVGLGNVAGTDGAITGAPGTIEAEVAAATRSGVHRYAFPAGSTPSLVVD
LETNNTSRRSSSVQVETRADGTVELSGQVTGYFYNAAYTLYYTARTLQPATVQTWGDDD
RLVDATAQDGVDTGAILTFDPADAGEIGLQVTLSPVSVEQARIDQQVELGDLSFDAIRD
RTRAEWNATLGRVAIDASTATDPTGELQRLFYTHLYRMFAMPMNATSTSGTYRGVDGAV
HAAQGFTYYDSWATWDDFRKFSVIAYIDPALYRDMVQSLVYLFADAEATGTGGGLGGFV
HSVPTVRWERSSVVVADAIAKGFDGFDRLDEAYPALQRLVGQYSADELRRGYVAGNPGA
SVQRGYDQYGLSVIADELGLTEEAETLREQASWPIEKLTKPGAWTAADGTQVGLLTPRA
ADGSWQSADHAKFEAAGLYQGTLWQYHWYDAYDMDALVEAMGGHEAARLGMRHMFGEHA
PDDGKAMLHSNANEIDLQAPYLFNYTGEPSLTQKWARAIYTKETWNRYIATGSSSAVPS
GGGEFTPPLKTKVYRLDPRGMLPTMDNDAGTMSTMFVAAAVGLFPVTAGSSQFQVGSPF
FDSTTITYDDGSAFTVTADGVSEDAFYVQSATLDGATFGNTWVDYATVVGGADLAFRMG
EQPSDWGTDTAPAFSMSTATDEPAEGPRVSAEPTTVQTGDGGALDATVTLTLDGARLAA
PAGTDLVTSGAASVVGLPDGVTAAVTVASPTALTVSLTGTASADARFFVHLRDAALADG
VAAASLQGQGVSVRSPLRLSVASAERDALAALVDDAVLVRHGNYSSVTFDRFSTALTKA
QEALGDEAATSIALRFAADRLGAAADALDLTGGGYRTLEAEQSEAWSGGELKNEANSSS
GNLGGVRSGSWVQYRDMTFETAAGDTPPRFLTVRYDTSFAPTDTPSTVRVHAGDVSGPV
VATVDLKGTSGWGKYTEVTAELGDVQALVDAQVVTFELLAPSGRSWVGNFDWFRFSAED
PAAPGQPGESPTVTIEAEDWTASSGRGLKKESSTWTSGPVTNVGGTADGDWIAYGEVDL
GELPLGELSVHYVHNSNRSGNNSALSVYLDAFDPANPGEPFVTVPLPTTGSSWTADGTA
TVVLPETVQGTHEVFVRLSTEPYADHPYVANLDSLTFAPGGPTSVVVESEAWTSNSGRG
LKNESSTWTSGPVTNVGGTADGDWLAYGEIDLGSAALDQLSVHYVHNSNRSGRNSALSV
YLDAFDPANPGEPFVTVPLANTGSSWTTDGTAVVDLPSTVRGKHQVWVRLSTEAYADHP
YVANLDSMRFFTDAYDVEVPPTDTAALAAVVDAAGTPEAEIARYGRIDARVFTRELAAA
RSVLADAGATQAQADERARRLGLATDQLVPAERRRLENLVASAEALTDEGYSPESWQAF
RTALAAATGTLDDAAASDEALHDARLALQGAVDALEEPADVVLVEVEVSPRCLAGKPYV
AVRAVNVSDAAVDVELASSLGTRSFVGVAPGASAYQSFAARSATGDLDVTVTATGADGT
QTVEQVVTVPSCS (SEQ ID NO:3)

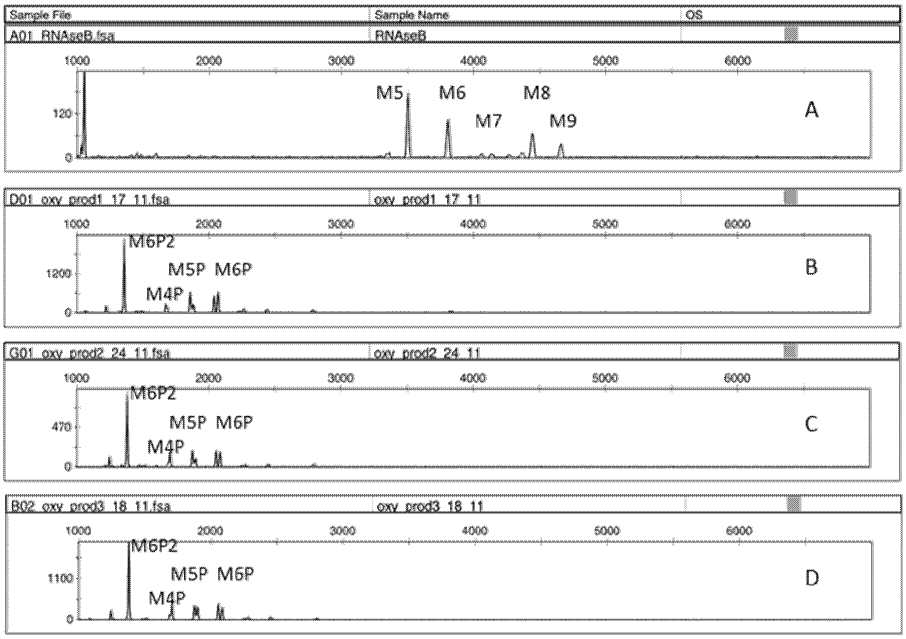
도면2c

APEPPSADYASLVDVFVGTEGDFGNDMPAAQAPNGLAKVNPRTTPGRNNTGYDYAQSKISGF
THTNLDGVGGSGGGGDLLVVPTSGSYTARPGTGTYAHPFSHDDEDAGPGFYSVGLGNVAGTD
GAITGAPGTIEAEVAAATRSGVHRYAFPAGSTPSLVVDLETNNTSRRSSSVQVETRADGTVE
LSGQVTGYFYNAAYTLYYTARTLQPATVQTWGDDDRLVDATAQDGVDTGAILTFDPADAGEI
GLQVTLSPVSVEQARIDQQVELGDLSFDAIRDRTRAEWNATLGRVAIDASTATDPTGELQRL
FYTHLYRMFAMPMNATSTSGTYRGVDGAVHAAQGFTYYDSWATWDDFRKFSVIAYIDPALYR
DMVQSLVYLFADAEATGTGGGLGGFVHSVPTVRWERSSVVVADAIAKGFDGFDRLDEAYPAL
QRLVGQYSADELRRGYVAGNPGASVQRGYDQYGLSVIADELGLTEEAETLREQASWPIEKLT
KPGAWTAADGTQVGLLTPRAADGSWQSADHAKFEAAGLYQGTLWQYHWYDAYDMDALVEAMG
GHEAARLGMRHMFGEHAPDDGKAMLHSNANEIDLQAPYLFNYTGEPSLTQKWARAIYTKETW
NRYIATGSSSAVPSGGGEFTPPLKTKVYRLDPRGMLPTMDNDAGTMSTMFVAAAVGLFPVTA
GSSQFQVGSPFFDSTTITYDDGSAFTVTADGVSEDAFYVQSATLDGATFGNTWVDYATVVGG
ADLAFRMGEQPSDWGTDTAPAFSMSTATDEPAEGPRVSAEPTTVQTGDGGALDATVTLTLDG
ARLAAPAGTDLVTSGAASVVGLPDGVTAAVTVASPTALTVSLTGTASADARFFVHLRDAALA
DGVAAASLQGQGVSVRSPLRLSVASAERDALAALVDDAVLVRHGNYSSVTFDRFSTALTKAQ
EALGDEAATSIALRFAADRLGAAADALDLTGGGYRTLEAEQSEAWSGGELKNEANSSSGNLG
GVRSGSWVQYRDMTFETAAGDTPPRFLTVRYDTSFAPTDTPSTVRVHAGDVSGPVVATVDLK
GTSGWGKYTEVTAELGDVQALVDAQVVTFELLAPSGRSWVGNFDWFRFSAEDPAAPGQPGES
PTVTIEAEDWTASSGRGLKKESSTWTSGPVTNVGGTADGDWIAYGEVDLGELPLGELSVHYV
HNSNRSGNNSALSVYLDAFDPANPGEPFVTVPLPTTGSSWTADGTATVVLPETVQGTHEVFV
RLSTEPYADHPYVANLDSLTFAPGGPTSVVVESEAWTSNSGRGLKNESSTWTSGPVTNVGGT
ADGDWLAYGEIDLGSAALDQLSVHYVHNSNRSGRNSALSVYLDAFDPANPGEPFVTVPLANT
GSSWTTDGTAVVDLPSTVRGKHQVWVRLSTEAYADHPYVANLDSMRFFTDAYDVEVPPTDTA
ALAAVVDAAGTPEAEIARYGRIDARVFTRELAAARSVLADAGATQAQADERARRLGLATDQL
VPAERRRLENLVASAEALTDEGYSPESWQAFRTALAAATGTLDDAAASDEALHDARLALQGA
VDALEEPADVVLVEVEVSPRCLAGKPYVAVRAVNVSDAAVDVELASSLGTRSFVGVAPGASA
YQSFAARSATGDLDVTVTATGADGTQTVEQVVTVPSCS (SEQ ID NO:4)

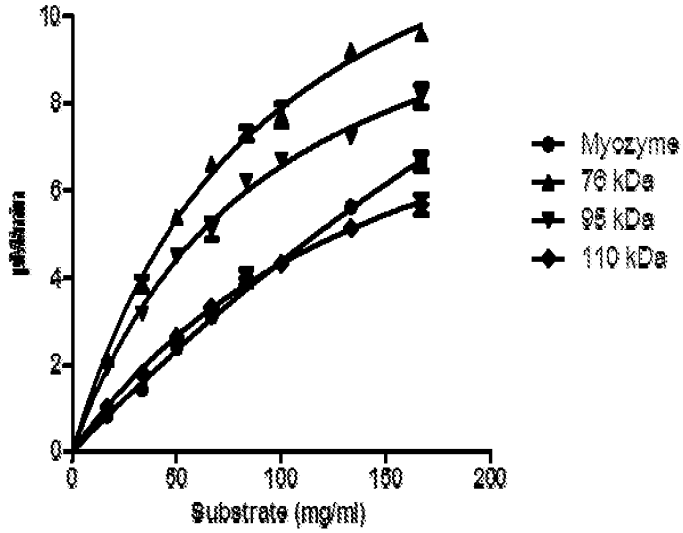
도면3a


Sample File
Sample Name
1_H05_RNaseB.fsa
RNaseB
M5
M6
M7
M8
M9
A
1_A10_Tou3A7pool1.fsa
Tou3A7pool1
M8(PM)2
M7PM
M8PM
B
1_B01_Tou3A9#1_sample1.fsa
Tou3A9#1_sample1
M6P2
M5P M6P
M4P
C

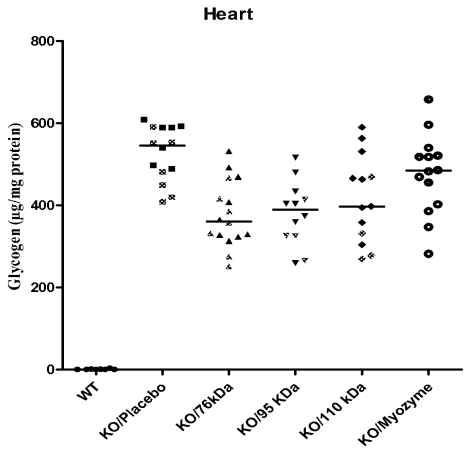
도면3b

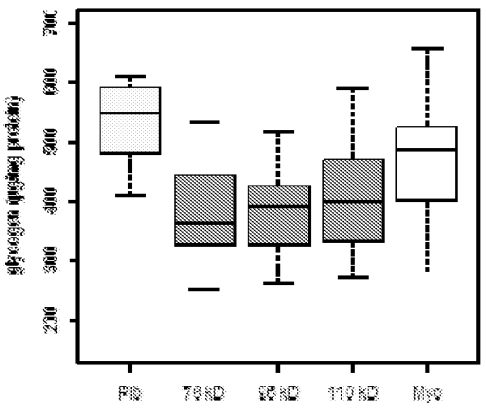

Sample File
Sample Name
OS
A01_RNAseB.fsa
RNAseB
M5
M6
M7
M8
M9
A
D01_oxy_prod1_17_11.fsa
oxy_prod1_17_11
M6P2
M5P M6P
M4P
B
G01_oxy_prod2_24_11.fsa
oxy_prod2_24_11
M6P2
M5P M6P
M4P
C
B02_oxy_prod3_18_11.fsa
oxy_prod3_18_11
M6P2
M5P M6P
M4P
D

도면4

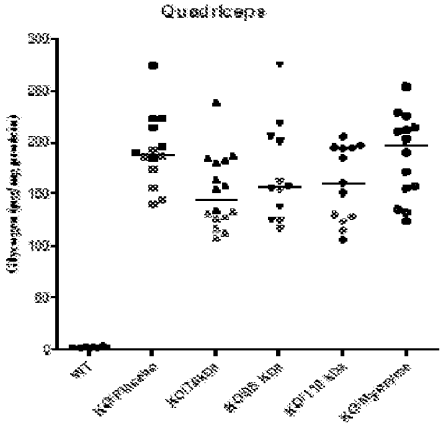

10
8
6
4
2
0
0
50
100
150
200
Substrate (mg/ml)
Myozyme
76 kDa
95 kDa
110 kDa

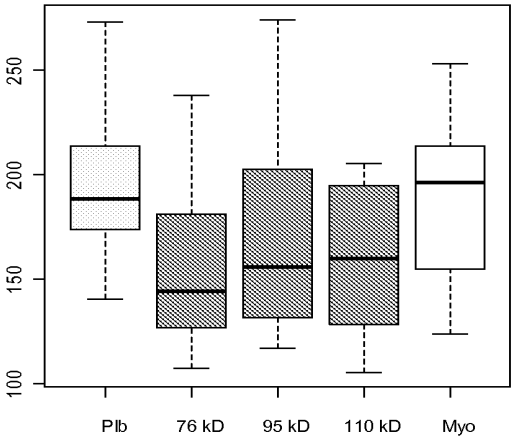
도면5a


Heart
800
600
400
200
0
Glycogen (μg/mg protein)
WT
KO/Placebo
KO/76kDa
KO/95 kDa
KO/110 kDa
KO/Myozyme



도면5b





도면6

MYSHFNNEPVAKRVNNLFTDRLRQFTSDGEYRSLNLPAFYERERLDGKNHVAIETYAVSD
LRRPLFKDALKEADGHWKPAKKGSEYGPSWATHWFKIQVCVPPEWKKNYYKKGDLVVFNW
NLNCEGLVFSESGEALIGLSGEERREWPIPDNWFDGKCHTFYIEASCNGMFGNATGSSIQ
PPSDNRYFRLDSADLVVINSEARHLFVDFWIIGDAAREFPGDSWQRGKALDVANKIMDAF
DPENPDESIAEGRKLAKEYLGDTTKAYKQQLPFADGLVYALGNCHIDTAWLWPFAETRRK
AGRSWASQLELIDKYPEYVFVASQAQQFKWLKEDYPDLFAKIQKQAKKGRFLPVGGAWTE
CDTNLPSGESLLRQFLLGQRFFLEHFGSLSDTFWLPDTFGYSAQVPQLCRLAGMDRFLTQ
KLSWNNINSFPNSTFNWVALDGSQVLCHMPPNNTYTSMANFGDVSRTQKQNKNLDTTRNS
LMLYGHGDGGGGPTAEMLEKLRRCRGVSNTVGELPPVIQGQSVTDFYNELLDQTNNGKDL
VTWVGELYFEFHRGTYTSQAQTKKGNRVSENLLHDVELLATLASIRDSSYKYPFAQLESL
WEDVCLCQFHDVLPGSCIEMVYKDVKKIHGRVIDTASHLIDKAASALGLSGHPSKDSFDC
TPVALNTMPWSRTEVVAVPQPHWDATVELAEGVEIQEDSGNALVMMSESGPVVTTQSVDL
FKSEDAYILENSQVKVTICKDDGTLTSIYDKENDRRVLSGTGNRLVLFDDQPLSWQAWDT
EVFSLGKKQYIGAENVTRHSIVSSGPLRSTVAFTYEFNKSVVTTEISLDANSPLVTFNTR
ADWHETCKFLKVEFPVDVHSESASYESQFGVVKRPTHYNTSWDVAKFEVCCHKFADLSEL
DYGVSILNDCKYGFATHGNLMRLSLLRAPKAPDAHADMGHHEFKYGVLAHKGPLGATTVR
AAYNFNNPLRVKYVGLSEVSTKQAFSLKGPANLVLSQVKRAEVDRSKKSTNVILRVYEAL
GGRTRGKLVIDLPNVVSVTKTCALEYSKEKQVVAKSEGVTSVDISLRAFEVATYKVELA
(SEQ ID NO:5)

도면7

1 mhlpslslsl talaiaspsa ayphfgssqp vlhsssdttq sradaikaaf shawdgylqy
61 afphdelhpv sngygdsrng wgasavdals tavimrnati vnqildhvgk idysktnttv
121 slfettiryl ggmlsgydll kgpvsdlvqn sskidvlltq sknladvlkf afdtpsgvpy
181 nnlnitsggn dgaktnglav tgtlalewtr lsdltgdtty adlsqkaesy llnpqpksae
241 pfpglvgsni nisngqftda qvswnggdds yyeylikmyv ydpkrfglyk drwvaaaqst
301 mqhlashpss rpdltflasy nngtlglssq hltcfdggsf llggtvlnrt dfinfgldlv
361 sgchdtynst ltgigpesfs wdtsdipssq qslyekagfy itsgayilrp eviesfyyaw
421 rvtgqetyrd wiwsafsavn dycrtssgfs gltdvnaang gsrydnqesf lfaevmkysy
481 mafaedaawq vqpgsgnqfv fnteahpvrv sst (SEQ ID NO:6)

도면8

MTRPLPPGRAVARSGSGRARPLGLVLAAALAVPLGVPLAAPAGALAAAPAAAAEPGDFS
SSFESGDPAALPTTVAERDGAPWQANVGSFTAGLPGSVLGQLKGVTASAQNLPNEGAAN
LADGSSGTKWLAFASTGWVRYEFAEPVSFVAYTMTSGDDAAGRDPKTWTVEGSNDGSTW
AALDRRTDEDFPNRQQTRTFELEAPTAAYTYLRLNVTANSGDSIVQLAGWDLSADLSAG
PSAAPMTTKVGTGPRVSFTNKAGVGFSGLHSLRYDGSHLADGETYATNVLYDDVDVVVG
EDTRLSYTIFPELLDDLQYPSTYAAVDVLFTDGTYLSDLGARDAHETVATAQAQGEGKI
LYADQWNSVRVDLGDVAEGKTVDQVLLGYDNPGGHAGTKFAGWLDDVEITAEPATIDGS
SLANYVDTRRGTLASGSFSRGNNIPATATPNGFNFWTPYTNASSQSWLYEYHKANNANN
KPVLQGFGISHEPSPWMGDRNQLTFLPSTASGTPDATLSTRGLEFDHADETARPDYYGV
TFTNGSAIEATPTDHGAVLRFSYPGAKGHVLVDKVDGSSKLTYDQATGTISGWVENGSG
LSVGRTRMFVAGTFDRSPTAVGTAAGNRADARFATFETSSDKTVELRVATSFISLDQAR
KNLDLEVTGKTFTEVKAAAAQAWNDRLGVIEVEGASEDQLVTLYSNLYRLNLYPNSQFE
NTGTAQEPVYRYASPVSATTGSATDTQTNAKIVDGKIYVNNGFWDTYRTAWPAYSLLYP
ELAAELVDGFVQQYRDGGWIARWSSPGYADLMTGTSSDVAFADAYLKGSLPTGTALEAY
DAALRNATVAPPSNAVGRKGLQTSPFLGFTPESTHESVSWGLEGLVNDFGIGNMAAALA
EDPATPEERRETLREESAYFLERATHYVELFDPEVDFFVPRHEDGTWAVDPETYDPEAW
GGGYTETNGWNFAFHAPQDGQGLANLYGGKQGLEDKLDEFFSTPEKGAGNGGIHEQREA
RDVRMGQWGMSNQVSHHIPWLYDAAGAPSKAQEKVREVTRRLFVGSEIGQGYPGDEDNG
EMSSWWIFASLGFYPLQVGSDQYAVGSPLFDKATVHLPDGDLVVNAENNSVDNVYVQSL
AVDGEARTSTSLSQADLSGGTTLDFVMGPEPSDWGTGEDDAPPSLTEGDEPPTPVQDAT
TAGLGTTTVADGDATTSAAALTDNTSGTRTTFATTTPSITWAGNGIRPTVGSYTLTSGA
SGTASPSAWTLEGSDDGETWTTLDERSGEQFRWALQTRPFTVAEPTAFARYRVTVTATS
GSGALSLAEVELLADPKESGAEELTLSAAPDRDGVTGREVSGSFATLTGVEGDVAALDV
QVAFGDGSEPVAGTLRAGAFGGYAVDAAHTWTAPGVYPVTVTVSGEGIETVSASSYVSV
SLLREGSLLAAYDNVCIGDAGTTVGSCDGQGVFFDRAQLAAKGFVQGERATVPGTDLAF
DVPAVPAGQPDNATGDGQTIELDVPADAEQLSVIGTGTEKNQQATGTLTFDDGSTQPID
LSFGDWSGAARNPVFGNIPVAVTDSRLRGGSPQTGTPAAFFATAPITLPEGKRPVSLTL
PDQPGELSRDGRIHVVAVAHDGTFAEHPALEVTAAEGVTLAVGQTSDVALAQVAGGREG
ADLRAAVTWGDGSDVAAGAVTDGSVSGSHAYTAAGTYTAYVVVDDGWTSQVVEVPVTVT
EAEPALAVDVTVSTRCLAGKAYVAVRAENGEDVPLAIRLVTPFGTKEVAAVAPGANAYS
FATRVTAVEAGTVTVEATRGTGDEEVTASIQADYAAVTCG (SEQ ID NO:7)

도면9

MQSIKRTLLLLGAILPAVLGAPVQETRRAAEKLPGKYIVTFKPGIDEAKIQEHTTWATNI
HQRSLERRGATGGDLPVGIERNYKINKFAAYAGSFDDATIEEIRKNEDVAYVEEDQIYYL
DGLTTQKSAPWGLGSISHKGQQSTDYIYDTSAGEGTYAYVVDSGVNVDHEEFEGRASKAY
NAAGGQHVDSIGHGTHVSGTIAGKTYGIAKKASILSVKVFQGESSSTSVILDGFNWAAND
IVSKKRTSKAAINMSLGGGYSKAFNDAVENAFEQGVLSVVAAGNENSDAGQTSPASAPDA
ITVAAIQKSNNRASFSNFGKVVDVFAPGQDILSAWIGSSSATNTISGTSMATPHIVGLSL
YLAALENLDGPAAVTKRIKELATKDVVKDVKGSPNLLAYNGNA (SEQ ID NO:9)

도면10

GGATCCATGCAGTCCATTAAGCGAACTCTGCTGCTGCTGGGAGCCATTCTGCCCGCCGTG
CTGGGAGCCCCCGTTCAGGAGACCCGACGAGCCGCCGAGAAGCTCCCCGGCAAGTACATT
GTCACCTTCAAGCCTGGTATCGACGAGGCTAAGATTCAGGAGCACACCACTTGGGCCACC
AACATCCATCAGCGATCCCTCGAGCGACGAGGAGCCACCGGCGGTGACCTGCCTGTGGGA
ATCGAGCGAAACTACAAGATTAACAAGTTCGCCGCTTACGCTGGATCTTTTGACGATGCC
ACCATCGAGGAGATTCGAAAGAACGAGGACGTCGCTTACGTGGAGGAAGACCAGATCTAC
TACCTCGATGGTCTGACCACTCAGAAGTCCGCTCCTTGGGGCCTGGGCTCCATCTCTCAC
AAGGGACAGCAGTCGACTGACTACATCTACGATACCTCCGCTGGCGAGGGTACTTACGCC
TACGTCGTGGACTCCGGTGTTAACGTCGATCACGAGGAGTTTGAGGGACGAGCCTCTAAG
GCTTACAACGCCGCTGGAGGCCAGCATGTGGACTCTATCGGACACGGCACCCATGTTTCG
GGTACTATTGCCGGAAAGACCTACGGCATCGCCAAGAAGGCTTCTATTCTCTCGGTGAAG
GTTTTCCAGGGAGAGTCCTCTTCGACCTCTGTCATCCTGGACGGCTTTAACTGGGCCGCT
AACGATATTGTGTCTAAGAAGCGAACCTCGAAGGCCGCTATCAACATGTCCCTCGGTGGA
GGCTACTCTAAGGCCTTCAACGACGCTGTTGAGAACGCCTTTGAGCAGGGTGTCCTGTCT
GTTGTGGCTGCTGGTAACGAGAACTCTGACGCTGGACAGACCTCCCCTGCTTCTGCTCCT
GATGCCATCACTGTGGCCGCTATTCAGAAGTCCAACAACCGAGCTTCGTTCTCCAACTTT
GGCAAGGTGGTTGACGTTTTCGCCCCCGGACAGGATATCCTCTCTGCTTGGATTGGCTCC
TCTTCGGCCACCAACACTATCTCGGGCACCTCCATGGCCACTCCCCACATTGTCGGTCTG
TCCCTCTACCTGGCTGCTCTGGAGAACCTGGACGGACCTGCCGCTGTTACCAAGCGAATC
AAGGAGCTGGCTACTAAGGACGTCGTGAAGGATGTCAAGGGTTCTCCTAACCTGCTCGCC
TACAACGGCAACGCTTCT**GGCGGCGGAGGA***CATCACCACCATCACCATCACCACCATCAT*
TGATAACCTAGG (SEQ ID NO:10)

서 열 목 록

SEQUENCE LISTING

<110> Oxyrane UK Limited

<120> METHODS AND MATERIALS FOR TREATMENT OF
POMPE'S DISEASE

<130> 18990-0034WO1

<150> 61/611,485

<151> 2012-03-15

<160> 12

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 898

<212> PRT

<213> Homo sapiens

<400> 1

Ala Ala Gln Gln Gly Ala Ser Arg Pro Gly Pro Arg Asp Ala Gln Ala
1 5 10 15

His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys Asp Val Pro Pro
20 25 30

Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile Thr Gln Glu Gln

35 40 45

Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys Gln Gly Leu Gln

50 55 60

Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr Pro

65 70 75 80

Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr Ala

85 90 95

Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr

100 105 110

Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr

115 120 125

Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr Pro

130 135 140

His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser

145 150 155 160

Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu Asp Gly Arg Val

165 170 175

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

180 185 190

Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu

195 200 205

His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu

210 215 220

Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser

225 230 235 240

His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val

245 250 255

Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro

260 265 270

Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe

275 280 285

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

340 345 350
Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg
355 360 365
Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr
370 375 380
Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser
385 390 395 400
Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr Asn

405 410 415
Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr Ala
420 425 430
Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp Met
435 440 445
Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly Met Trp Ile Asp
450 455 460
Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys Pro

465 470 475 480
Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly
485 490 495
Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser
500 505 510
Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala
515 520 525
Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile

530 535 540

Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr

545 550 555 560

Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu

565 570 575

Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val

580 585 590

Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr

595 600 605

Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu

610 615 620

Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala

625 630 635 640

Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr

645 650 655

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

660 665 670

Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His

675 680 685

Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala

690 695 700

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

705 710 715 720

Leu Gln Thr Val Pro Val Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro

725 730 735

Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr

740 745 750

Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr

755 760 765

Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln

770 775 780

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

785 790 795 800

Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg

805 810 815

Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val

820 825 830

Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln

835 840 845

Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser

850 855 860

Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val

865 870 875 880

Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser

885 890 895

Trp Cys

<210> 2

<211> 5060

<212> DNA

<213> Cellulosimicrobium cellulans

<400> 2

atgaaaaaga tttggctggc gctggctggt ttagttttag cgtttagcgc atcggccggc 60

catcaccatc atcaccacgt ggggcccggc tcggacgaag tggatgcacc ggaacctccg 120

agcgcagatt atgcaagcct ggttgatgtt tttgttggca ccgaaggtga ttttggtaat 180

gatatgcctg cagcacaggc accgaatggt ctggcaaaag ttaatccgcg taccacaccg 240

ggtcgtaata ataccggtta tgattatgcc cagagcaaaa ttagcggttt tacccatacc 300

aatctggatg gtgttggtgg tagcggtggt ggtggtgatc tgctggttgt tccgaccagc 360

ggtagctata ccgcacgtcc gggtacaggc acctatgcac atccgtttag ccatgatgat 420

gaagatgcag gtccgggttt ttatagcgtt ggtctgggta atgttgcagg caccgatggt 480

gcaattaccg gtgctccggg tacaattgaa gcagaagttg cagcagcaac ccgtagcggt 540

gttcatcgtt atgcatttcc ggcaggtagc accccgagcc tggttgttga tctggaaacc 600

aataatacca gccgtcgtag cagcagcgtt caggttgaaa cccgtgcaga tggcaccgtt 660

gaactgagcg gtcaggttac cggctatttt tataatgcag cctataccct gtattatacc 720

gcacgcaccc tgcagcctgc aaccgttcag acctggggtg atgatgatcg tctggttgat 780

gcaaccgcac aggatggtgt tgataccggt gcaattctga cctttgatcc ggcagatgcc 840

ggtgaaattg gtctgcaggt taccctgtct ccggttagcg ttgaacaggc acgtattgat 900

cagcaggttg aactgggtga tctgagcttt gatgcaattc gtgatcgtac ccgtgcagaa 960

tggaatgcaa ccctgggtcg tgttgcaatt gatgcaagca ccgcaaccga tccgaccggt 1020

gaactgcagc gtctgtttta tacccatctg tatcgcatgt ttgcaatgcc gatgaatgca 1080

accagcacca gcggcaccta tcgtggtgtt gatggtgcag ttcatgcagc acagggcttt 1140

acctattatg atagctgggc aacctgggat gattttcgca aatttagcgt gattgcctat 1200

attgatccgg cactgtatcg tgatatggtt cagagcctgg tttacctgtt tgcagatgca 1260

gaagcaaccg gtacaggcgg tggtctgggt ggttttgttc atagcgttcc gaccgttcgt 1320

tgggaacgta gcagcgttgt tgttgcagat gcaattgcca aaggctttga tggttttgat 1380

cgtctggatg aagcatatcc ggcactgcag cgcctggttg gtcagtatag cgcagatgaa 1440

ctgcgtcgtg gttatgttgc aggtaatccg ggtgcaagcg ttcagcgtgg ttatgatcag 1500

tatggtctga gcgttattgc cgatgaactg ggtctgaccg aagaagcaga aaccctgcgc 1560

gaacaggcaa gctggccgat tgaaaaactg accaaaccgg gtgcatggac cgcagcagat 1620

ggtacacagg ttggtctgct gacaccgcgt gcagccgatg gtagctggca gagcgcagat 1680

catgccaaat ttgaagcagc aggtctgtat cagggcaccc tgtggcagta tcattggtat 1740

gatgcctatg atatggatgc actggttgaa gcaatgggtg gtcatgaagc agcccgtctg 1800

ggtatgcgtc atatgtttgg tgaacatgca ccggatgatg gtaaagcaat gctgcatagc 1860

aatgccaatg aaattgatct gcaggcaccg tacctgttta attataccgg tgaaccgagc 1920

ctgacccaga aatgggcacg tgcaatttat accaaagaaa cctggaatcg ctatattgca 1980

accggtagca gctctgcagt tccgtcaggt ggtggtgaat ttacacctcc gctgaaaacc 2040

aaagtttatc gtctggaccc tcgtggtatg ctgccgacca tggataatga tgcaggtaca 2100

atgagcacca tgtttgttgc agcagccgtt ggtctgtttc cggttaccgc aggtagcagc 2160

cagtttcagg ttggtagccc gttttttgat agcaccacca ttacctatga tgatggtagc 2220

gcatttaccg ttaccgcaga tggtgttagc gaagatgcct tttatgttca gagcgcaacc 2280

ctggatggtg caacctttgg taatacctgg gttgattatg caaccgttgt tggtggtgca 2340

gatctggcat ttcgtatggg tgaacagccg agcgattggg gcaccgatac cgcaccggca 2400

tttagcatga gcaccgccac cgatgaaccg gcagaaggtc ctcgcgttag cgcagaaccg 2460

accaccgtgc agaccggtga tggtggtgca ctggatgcaa ccgttaccct gacactggat 2520

ggcgcacgtc tggcagcacc ggcaggtaca gatctggtta ccagcggtgc agcaagcgtt 2580
gttggtctgc cggatggtgt taccgcagca gttaccgttg caagcccgac cgcactgacc 2640
gttagcctga ccggcaccgc atcagcagat gcacgttttt ttgtgcatct gcgtgatgca 2700
gcactggccg atggtgttgc agccgcaagc ctgcagggtc agggtgttag cgttcgttct 2760
ccgctgcgtc tgagcgttgc aagcgcagaa cgtgatgcac tggcagcact ggttgatgat 2820
gccgttctgg ttcgtcatgg taattatagc agcgttacct ttgatcgttt agcaccgctc 2880
tgacaaaagc acaggaagca ctgggcgacg aagcagcaac cagcattgca ctgcgttttg 2940
cagcagatcg tctgggtgca gcagcagatg cactggatct gaccggtggt ggttatcgta 3000

ccctggaagc agaacagagc gaagcatggt ctggtggtga actgaaaaat gaagccaata 3060
gcagcagcgg taatctgggt ggtgttcgta gcggtagctg ggttcagtat cgcgatatga 3120
cctttgaaac cgcagccggt gatacacctc cgcgttttct gaccgttcgt tatgatacca 3180
gctttgcacc gaccgatacc ccgagcaccg ttcgtgttca tgccggtgat gtttctggtc 3240
cggttgttgc aaccgttgat ctgaaaggca ccagcggttg gggtaaatat accgaagtta 3300
ccgcagaact gggtgatgtt caggccctgg ttgatgccca ggttgttacc tttgaactgc 3360
tggcaccgag cggtcgtagc tgggttggta attttgattg gtttcgcttt agcgcagaag 3420
atccggcagc accgggtcag cctggtgaaa gcccgaccgt taccattgaa gccgaagatt 3480

ggaccgcaag cagcggtcgt ggtctgaaaa aagaaagcag cacctggacc agcggtccgg 3540
tgaccaatgt tggtggtaca gcagatggtg attggattgc ctatggtgaa gttgatctgg 3600
gtgaactgcc gctgggcgaa ctgagcgttc attatgtgca taatagcaat cgcagcggta 3660
ataatagcgc actgagcgtt tatctggatg catttgatcc ggctaatccg ggtgaaccgt 3720
ttgttaccgt tccgctgccg accaccggta gcagttggac cgcagatggc acagccaccg 3780
ttgttctgcc ggaaaccgtg cagggcaccc atgaagtttt tgttcgtctg agcaccgaac 3840
cgtatgcaga tcatccgtat gttgcaaatc tggatagcct gacctttgca ccgggtggtc 3900
cgaccagcgt tgtggttgaa agcgaagcct ggaccagcaa ttctggtcgt ggcctgaaaa 3960

atgaatcttc tacctggacc tctggtccgg ttacaaatgt gggtggcacc gctgatggcg 4020
attggctggc atatggcgaa attgatctgg gcagcgcagc actggatcag ctgtctgtgc 4080
attatgttca taattctaat cgctctggtc gtaattctgc actgtctgtg tatctggatg 4140
cctttgatcc ggcaaatccg ggtgaaccgt ttgtgacagt gccgctggca aataccggta 4200
gctcttggac caccgatggt actgcagttg tggatctgcc gtctaccgtt cgtggtaaac 4260
atcaggtttg ggttcgtctg tctaccgaag catatgccga tcatccgtat gtggccaatc 4320

tggattctat gcgctttttt accgatgcat atgatgttga agttcctccg accgatacag 4380

cagcactggc agccgttgtt gatgcagcag gtacaccgga agcagaaatt gcacgttatg 4440

gtcgtattga tgcccgtgtt tttacccgtg aactggcagc agcacgtagc gttctggccg 4500

atgccggtgc aacacaggca caggcagatg aacgtgctcg tcgtctgggt ctggcaaccg 4560

atcagctggt tccggcagaa cgtcgtcgtc tggaaaatct ggttgccagc gcagaagcac 4620

tgaccgacga aggttattct ccggaaagct ggcaggcatt tcgtaccgca ctggctgctg 4680

caaccggcac cctggatgat gcagcagcat ctgatgaagc actgcatgat gcacgtctgg 4740

cgctgcaggg tgcagttgat gcactggaag aaccggcaga tgttgttctg gttgaagttg 4800

aagtttctcc gcgttgtctg gcaggtaaac cgtatgttgc cgttcgtgca gttaatgttt 4860

ctgatgcagc cgttgatgtt gaactggcaa gctctctggg cacccgtagc tttgttggtg 4920

tggcaccggg tgcgagcgca tatcagagct ttgcagcccg tagcgcaacc ggtgatctgg 4980

atgttaccgt gaccgcaacc ggtgcagatg gtactcagac cgttgaacag gttgtgaccg 5040

ttccgagctg tagctaataa 5060

<210> 3

<211> 1665

<212> PRT

<213> Cellulosimicrobium cellulans

<400> 3

Ala Leu Ala Val Val Gly Leu Ala Pro Ala Thr Ala Ala Ser Ala Ala
1 5 10 15

Pro Glu Pro Pro Ser Ala Asp Tyr Ala Ser Leu Val Asp Val Phe Val
20 25 30

Gly Thr Glu Gly Asp Phe Gly Asn Asp Met Pro Ala Ala Gln Ala Pro
35 40 45

Asn Gly Leu Ala Lys Val Asn Pro Arg Thr Thr Pro Gly Arg Asn Asn
50 55 60

Thr Gly Tyr Asp Tyr Ala Gln Ser Lys Ile Ser Gly Phe Thr His Thr
65 70 75 80

Asn Leu Asp Gly Val Gly Gly Ser Gly Gly Gly Gly Asp Leu Leu Val
85 90 95

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

```
            100                 105                 110
Ala His Pro Phe Ser His Asp Asp Glu Asp Ala Gly Pro Gly Phe Tyr
        115                 120                 125
Ser Val Gly Leu Gly Asn Val Ala Gly Thr Asp Gly Ala Ile Thr Gly
    130                 135                 140
Ala Pro Gly Thr Ile Glu Ala Glu Val Ala Ala Ala Thr Arg Ser Gly
145                 150                 155                 160

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

Gln Pro Ala Thr Val Gln Thr Trp Gly Asp Asp Asp Arg Leu Val Asp
225                 230                 235                 240
Ala Thr Ala Gln Asp Gly Val Asp Thr Gly Ala Ile Leu Thr Phe Asp
                245                 250                 255
Pro Ala Asp Ala Gly Glu Ile Gly Leu Gln Val Thr Leu Ser Pro Val
            260                 265                 270
Ser Val Glu Gln Ala Arg Ile Asp Gln Gln Val Glu Leu Gly Asp Leu
        275                 280                 285

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

Ala Val His Ala Ala Gln Gly Phe Thr Tyr Tyr Asp Ser Trp Ala Thr
355 360 365
Trp Asp Asp Phe Arg Lys Phe Ser Val Ile Ala Tyr Ile Asp Pro Ala
370 375 380
Leu Tyr Arg Asp Met Val Gln Ser Leu Val Tyr Leu Phe Ala Asp Ala
385 390 395 400
Glu Ala Thr Gly Thr Gly Gly Gly Leu Gly Gly Phe Val His Ser Val
405 410 415

Pro Thr Val Arg Trp Glu Arg Ser Ser Val Val Val Ala Asp Ala Ile
420 425 430
Ala Lys Gly Phe Asp Gly Phe Asp Arg Leu Asp Glu Ala Tyr Pro Ala
435 440 445
Leu Gln Arg Leu Val Gly Gln Tyr Ser Ala Asp Glu Leu Arg Arg Gly
450 455 460
Tyr Val Ala Gly Asn Pro Gly Ala Ser Val Gln Arg Gly Tyr Asp Gln
465 470 475 480

Tyr Gly Leu Ser Val Ile Ala Asp Glu Leu Gly Leu Thr Glu Glu Ala
485 490 495
Glu Thr Leu Arg Glu Gln Ala Ser Trp Pro Ile Glu Lys Leu Thr Lys
500 505 510
Pro Gly Ala Trp Thr Ala Ala Asp Gly Thr Gln Val Gly Leu Leu Thr
515 520 525
Pro Arg Ala Ala Asp Gly Ser Trp Gln Ser Ala Asp His Ala Lys Phe
530 535 540

Glu Ala Ala Gly Leu Tyr Gln Gly Thr Leu Trp Gln Tyr His Trp Tyr
545 550 555 560
Asp Ala Tyr Asp Met Asp Ala Leu Val Glu Ala Met Gly Gly His Glu
565 570 575
Ala Ala Arg Leu Gly Met Arg His Met Phe Gly Glu His Ala Pro Asp
580 585 590
Asp Gly Lys Ala Met Leu His Ser Asn Ala Asn Glu Ile Asp Leu Gln

595 600 605

Ala Pro Tyr Leu Phe Asn Tyr Thr Gly Glu Pro Ser Leu Thr Gln Lys
610 615 620
Trp Ala Arg Ala Ile Tyr Thr Lys Glu Thr Trp Asn Arg Tyr Ile Ala
625 630 635 640
Thr Gly Ser Ser Ser Ala Val Pro Ser Gly Gly Gly Glu Phe Thr Pro
645 650 655
Pro Leu Lys Thr Lys Val Tyr Arg Leu Asp Pro Arg Gly Met Leu Pro
660 665 670

Thr Met Asp Asn Asp Ala Gly Thr Met Ser Thr Met Phe Val Ala Ala
675 680 685
Ala Val Gly Leu Phe Pro Val Thr Ala Gly Ser Ser Gln Phe Gln Val
690 695 700
Gly Ser Pro Phe Phe Asp Ser Thr Thr Ile Thr Tyr Asp Asp Gly Ser
705 710 715 720
Ala Phe Thr Val Thr Ala Asp Gly Val Ser Glu Asp Ala Phe Tyr Val
725 730 735

Gln Ser Ala Thr Leu Asp Gly Ala Thr Phe Gly Asn Thr Trp Val Asp
740 745 750
Tyr Ala Thr Val Val Gly Gly Ala Asp Leu Ala Phe Arg Met Gly Glu
755 760 765
Gln Pro Ser Asp Trp Gly Thr Asp Thr Ala Pro Ala Phe Ser Met Ser
770 775 780
Thr Ala Thr Asp Glu Pro Ala Glu Gly Pro Arg Val Ser Ala Glu Pro
785 790 795 800

Thr Thr Val Gln Thr Gly Asp Gly Gly Ala Leu Asp Ala Thr Val Thr
805 810 815
Leu Thr Leu Asp Gly Ala Arg Leu Ala Ala Pro Ala Gly Thr Asp Leu
820 825 830
Val Thr Ser Gly Ala Ala Ser Val Val Gly Leu Pro Asp Gly Val Thr
835 840 845

Ala Ala Val Thr Val Ala Ser Pro Thr Ala Leu Thr Val Ser Leu Thr
850 855 860

Gly Thr Ala Ser Ala Asp Ala Arg Phe Phe Val His Leu Arg Asp Ala
865 870 875 880
Ala Leu Ala Asp Gly Val Ala Ala Ala Ser Leu Gln Gly Gln Gly Val
885 890 895
Ser Val Arg Ser Pro Leu Arg Leu Ser Val Ala Ser Ala Glu Arg Asp
900 905 910
Ala Leu Ala Ala Leu Val Asp Asp Ala Val Leu Val Arg His Gly Asn
915 920 925

Tyr Ser Ser Val Thr Phe Asp Arg Phe Ser Thr Ala Leu Thr Lys Ala
930 935 940
Gln Glu Ala Leu Gly Asp Glu Ala Ala Thr Ser Ile Ala Leu Arg Phe
945 950 955 960
Ala Ala Asp Arg Leu Gly Ala Ala Ala Asp Ala Leu Asp Leu Thr Gly
965 970 975
Gly Gly Tyr Arg Thr Leu Glu Ala Glu Gln Ser Glu Ala Trp Ser Gly
980 985 990

Gly Glu Leu Lys Asn Glu Ala Asn Ser Ser Ser Gly Asn Leu Gly Gly
995 1000 1005
Val Arg Ser Gly Ser Trp Val Gln Tyr Arg Asp Met Thr Phe Glu Thr
1010 1015 1020
Ala Ala Gly Asp Thr Pro Pro Arg Phe Leu Thr Val Arg Tyr Asp Thr
1025 1030 1035 1040
Ser Phe Ala Pro Thr Asp Thr Pro Ser Thr Val Arg Val His Ala Gly
1045 1050 1055

Asp Val Ser Gly Pro Val Val Ala Thr Val Asp Leu Lys Gly Thr Ser
1060 1065 1070
Gly Trp Gly Lys Tyr Thr Glu Val Thr Ala Glu Leu Gly Asp Val Gln
1075 1080 1085
Ala Leu Val Asp Ala Gln Val Val Thr Phe Glu Leu Leu Ala Pro Ser

1090 1095 1100

Gly Arg Ser Trp Val Gly Asn Phe Asp Trp Phe Arg Phe Ser Ala Glu

1105 1110 1115 1120

Asp Pro Ala Ala Pro Gly Gln Pro Gly Glu Ser Pro Thr Val Thr Ile

1125 1130 1135

Glu Ala Glu Asp Trp Thr Ala Ser Ser Gly Arg Gly Leu Lys Lys Glu

1140 1145 1150

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

1155 1160 1165

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

1170 1175 1180

Leu Gly Glu Leu Ser Val His Tyr Val His Asn Ser Asn Arg Ser Gly

1185 1190 1195 1200

Asn Asn Ser Ala Leu Ser Val Tyr Leu Asp Ala Phe Asp Pro Ala Asn

1205 1210 1215

Pro Gly Glu Pro Phe Val Thr Val Pro Leu Pro Thr Thr Gly Ser Ser

1220 1225 1230

Trp Thr Ala Asp Gly Thr Ala Thr Val Val Leu Pro Glu Thr Val Gln

1235 1240 1245

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

1250 1255 1260

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

1265 1270 1275 1280

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

1285 1290 1295

Arg Gly Leu Lys Asn Glu Ser Ser Thr Trp Thr Ser Gly Pro Val Thr

1300 1305 1310

Asn Val Gly Gly Thr Ala Asp Gly Asp Trp Leu Ala Tyr Gly Glu Ile

1315 1320 1325

Asp Leu Gly Ser Ala Ala Leu Asp Gln Leu Ser Val His Tyr Val His

1330 1335 1340

Asn Ser Asn Arg Ser Gly Arg Asn Ser Ala Leu Ser Val Tyr Leu Asp
1345 1350 1355 1360
Ala Phe Asp Pro Ala Asn Pro Gly Glu Pro Phe Val Thr Val Pro Leu
1365 1370 1375

Ala Asn Thr Gly Ser Ser Trp Thr Thr Asp Gly Thr Ala Val Val Asp
1380 1385 1390
Leu Pro Ser Thr Val Arg Gly Lys His Gln Val Trp Val Arg Leu Ser
1395 1400 1405
Thr Glu Ala Tyr Ala Asp His Pro Tyr Val Ala Asn Leu Asp Ser Met
1410 1415 1420
Arg Phe Phe Thr Asp Ala Tyr Asp Val Glu Val Pro Pro Thr Asp Thr
1425 1430 1435 1440

Ala Ala Leu Ala Ala Val Val Asp Ala Ala Gly Thr Pro Glu Ala Glu
1445 1450 1455
Ile Ala Arg Tyr Gly Arg Ile Asp Ala Arg Val Phe Thr Arg Glu Leu
1460 1465 1470
Ala Ala Ala Arg Ser Val Leu Ala Asp Ala Gly Ala Thr Gln Ala Gln
1475 1480 1485
Ala Asp Glu Arg Ala Arg Arg Leu Gly Leu Ala Thr Asp Gln Leu Val
1490 1495 1500

Pro Ala Glu Arg Arg Arg Leu Glu Asn Leu Val Ala Ser Ala Glu Ala
1505 1510 1515 1520
Leu Thr Asp Glu Gly Tyr Ser Pro Glu Ser Trp Gln Ala Phe Arg Thr
1525 1530 1535
Ala Leu Ala Ala Ala Thr Gly Thr Leu Asp Asp Ala Ala Ala Ser Asp
1540 1545 1550
Glu Ala Leu His Asp Ala Arg Leu Ala Leu Gln Gly Ala Val Asp Ala
1555 1560 1565

Leu Glu Glu Pro Ala Asp Val Val Leu Val Glu Val Glu Val Ser Pro
1570 1575 1580
Arg Cys Leu Ala Gly Lys Pro Tyr Val Ala Val Arg Ala Val Asn Val

1585 1590 1595 1600

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

1605 1610 1615

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

1620 1625 1630

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

1635 1640 1645

Ala Asp Gly Thr Gln Thr Val Glu Gln Val Val Thr Val Pro Ser Cys

1650 1655 1660

Ser

1665

<210> 4

<211> 1650

<212> PRT

<213> Cellulosimicrobium cellulans

<400> 4

Ala Pro Glu Pro Pro Ser Ala Asp Tyr Ala Ser Leu Val Asp Val Phe

1 5 10 15

Val Gly Thr Glu Gly Asp Phe Gly Asn Asp Met Pro Ala Ala Gln Ala

20 25 30

Pro Asn Gly Leu Ala Lys Val Asn Pro Arg Thr Thr Pro Gly Arg Asn

35 40 45

Asn Thr Gly Tyr Asp Tyr Ala Gln Ser Lys Ile Ser Gly Phe Thr His

50 55 60

Thr Asn Leu Asp Gly Val Gly Gly Ser Gly Gly Gly Gly Asp Leu Leu

65 70 75 80

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

85 90 95

Tyr Ala His Pro Phe Ser His Asp Asp Glu Asp Ala Gly Pro Gly Phe

100 105 110

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

115 120 125

Gly Ala Pro Gly Thr Ile Glu Ala Glu Val Ala Ala Ala Thr Arg Ser

130 135 140

Gly Val His Arg Tyr Ala Phe Pro Ala Gly Ser Thr Pro Ser Leu Val

145 150 155 160

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

165 170 175

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

180 185 190

Gly Tyr Phe Tyr Asn Ala Ala Tyr Thr Leu Tyr Tyr Thr Ala Arg Thr

195 200 205

Leu Gln Pro Ala Thr Val Gln Thr Trp Gly Asp Asp Asp Arg Leu Val

210 215 220

Asp Ala Thr Ala Gln Asp Gly Val Asp Thr Gly Ala Ile Leu Thr Phe

225 230 235 240

Asp Pro Ala Asp Ala Gly Glu Ile Gly Leu Gln Val Thr Leu Ser Pro

245 250 255

Val Ser Val Glu Gln Ala Arg Ile Asp Gln Gln Val Glu Leu Gly Asp

260 265 270

Leu Ser Phe Asp Ala Ile Arg Asp Arg Thr Arg Ala Glu Trp Asn Ala

275 280 285

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

290 295 300

Gly Glu Leu Gln Arg Leu Phe Tyr Thr His Leu Tyr Arg Met Phe Ala

305 310 315 320

Met Pro Met Asn Ala Thr Ser Thr Ser Gly Thr Tyr Arg Gly Val Asp

325 330 335

Gly Ala Val His Ala Ala Gln Gly Phe Thr Tyr Tyr Asp Ser Trp Ala

340 345 350

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

355 360 365

Ala Leu Tyr Arg Asp Met Val Gln Ser Leu Val Tyr Leu Phe Ala Asp

370 375 380

Ala Glu Ala Thr Gly Thr Gly Gly Gly Leu Gly Gly Phe Val His Ser

385 390 395 400

Val Pro Thr Val Arg Trp Glu Arg Ser Ser Val Val Val Ala Asp Ala

405 410 415

Ile Ala Lys Gly Phe Asp Gly Phe Asp Arg Leu Asp Glu Ala Tyr Pro

420 425 430

Ala Leu Gln Arg Leu Val Gly Gln Tyr Ser Ala Asp Glu Leu Arg Arg

435 440 445

Gly Tyr Val Ala Gly Asn Pro Gly Ala Ser Val Gln Arg Gly Tyr Asp

450 455 460

Gln Tyr Gly Leu Ser Val Ile Ala Asp Glu Leu Gly Leu Thr Glu Glu

465 470 475 480

Ala Glu Thr Leu Arg Glu Gln Ala Ser Trp Pro Ile Glu Lys Leu Thr

485 490 495

Lys Pro Gly Ala Trp Thr Ala Ala Asp Gly Thr Gln Val Gly Leu Leu

500 505 510

Thr Pro Arg Ala Ala Asp Gly Ser Trp Gln Ser Ala Asp His Ala Lys

515 520 525

Phe Glu Ala Ala Gly Leu Tyr Gln Gly Thr Leu Trp Gln Tyr His Trp

530 535 540

Tyr Asp Ala Tyr Asp Met Asp Ala Leu Val Glu Ala Met Gly Gly His

545 550 555 560

Glu Ala Ala Arg Leu Gly Met Arg His Met Phe Gly Glu His Ala Pro

565 570 575

Asp Asp Gly Lys Ala Met Leu His Ser Asn Ala Asn Glu Ile Asp Leu

580 585 590

Gln Ala Pro Tyr Leu Phe Asn Tyr Thr Gly Glu Pro Ser Leu Thr Gln

595 600 605

Lys Trp Ala Arg Ala Ile Tyr Thr Lys Glu Thr Trp Asn Arg Tyr Ile

610 615 620

Ala Thr Gly Ser Ser Ser Ala Val Pro Ser Gly Gly Gly Glu Phe Thr
625 630 635 640
Pro Pro Leu Lys Thr Lys Val Tyr Arg Leu Asp Pro Arg Gly Met Leu
645 650 655
Pro Thr Met Asp Asn Asp Ala Gly Thr Met Ser Thr Met Phe Val Ala

660 665 670
Ala Ala Val Gly Leu Phe Pro Val Thr Ala Gly Ser Ser Gln Phe Gln
675 680 685
Val Gly Ser Pro Phe Phe Asp Ser Thr Thr Ile Thr Tyr Asp Asp Gly
690 695 700
Ser Ala Phe Thr Val Thr Ala Asp Gly Val Ser Glu Asp Ala Phe Tyr
705 710 715 720
Val Gln Ser Ala Thr Leu Asp Gly Ala Thr Phe Gly Asn Thr Trp Val

725 730 735
Asp Tyr Ala Thr Val Val Gly Gly Ala Asp Leu Ala Phe Arg Met Gly
740 745 750
Glu Gln Pro Ser Asp Trp Gly Thr Asp Thr Ala Pro Ala Phe Ser Met
755 760 765
Ser Thr Ala Thr Asp Glu Pro Ala Glu Gly Pro Arg Val Ser Ala Glu
770 775 780
Pro Thr Thr Val Gln Thr Gly Asp Gly Gly Ala Leu Asp Ala Thr Val

785 790 795 800
Thr Leu Thr Leu Asp Gly Ala Arg Leu Ala Ala Pro Ala Gly Thr Asp
805 810 815
Leu Val Thr Ser Gly Ala Ala Ser Val Val Gly Leu Pro Asp Gly Val
820 825 830
Thr Ala Ala Val Thr Val Ala Ser Pro Thr Ala Leu Thr Val Ser Leu
835 840 845
Thr Gly Thr Ala Ser Ala Asp Ala Arg Phe Phe Val His Leu Arg Asp

850 855 860
Ala Ala Leu Ala Asp Gly Val Ala Ala Ala Ser Leu Gln Gly Gln Gly

865 870 875 880

Val Ser Val Arg Ser Pro Leu Arg Leu Ser Val Ala Ser Ala Glu Arg

885 890 895

Asp Ala Leu Ala Ala Leu Val Asp Asp Ala Val Leu Val Arg His Gly

900 905 910

Asn Tyr Ser Ser Val Thr Phe Asp Arg Phe Ser Thr Ala Leu Thr Lys

915 920 925

Ala Gln Glu Ala Leu Gly Asp Glu Ala Ala Thr Ser Ile Ala Leu Arg

930 935 940

Phe Ala Ala Asp Arg Leu Gly Ala Ala Ala Asp Ala Leu Asp Leu Thr

945 950 955 960

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

965 970 975

Gly Gly Glu Leu Lys Asn Glu Ala Asn Ser Ser Ser Gly Asn Leu Gly

980 985 990

Gly Val Arg Ser Gly Ser Trp Val Gln Tyr Arg Asp Met Thr Phe Glu

995 1000 1005

Thr Ala Ala Gly Asp Thr Pro Pro Arg Phe Leu Thr Val Arg Tyr Asp

1010 1015 1020

Thr Ser Phe Ala Pro Thr Asp Thr Pro Ser Thr Val Arg Val His Ala

1025 1030 1035 1040

Gly Asp Val Ser Gly Pro Val Val Ala Thr Val Asp Leu Lys Gly Thr

1045 1050 1055

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

1060 1065 1070

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

1075 1080 1085

Ser Gly Arg Ser Trp Val Gly Asn Phe Asp Trp Phe Arg Phe Ser Ala

1090 1095 1100

Glu Asp Pro Ala Ala Pro Gly Gln Pro Gly Glu Ser Pro Thr Val Thr

1105 1110 1115 1120

Ile Glu Ala Glu Asp Trp Thr Ala Ser Ser Gly Arg Gly Leu Lys Lys
1125 1130 1135
Glu Ser Ser Thr Trp Thr Ser Gly Pro Val Thr Asn Val Gly Gly Thr
1140 1145 1150
Ala Asp Gly Asp Trp Ile Ala Tyr Gly Glu Val Asp Leu Gly Glu Leu
1155 1160 1165
Pro Leu Gly Glu Leu Ser Val His Tyr Val His Asn Ser Asn Arg Ser

1170 1175 1180
Gly Asn Asn Ser Ala Leu Ser Val Tyr Leu Asp Ala Phe Asp Pro Ala
1185 1190 1195 1200
Asn Pro Gly Glu Pro Phe Val Thr Val Pro Leu Pro Thr Thr Gly Ser
1205 1210 1215
Ser Trp Thr Ala Asp Gly Thr Ala Thr Val Val Leu Pro Glu Thr Val
1220 1225 1230
Gln Gly Thr His Glu Val Phe Val Arg Leu Ser Thr Glu Pro Tyr Ala

1235 1240 1245
Asp His Pro Tyr Val Ala Asn Leu Asp Ser Leu Thr Phe Ala Pro Gly
1250 1255 1260
Gly Pro Thr Ser Val Val Val Glu Ser Glu Ala Trp Thr Ser Asn Ser
1265 1270 1275 1280
Gly Arg Gly Leu Lys Asn Glu Ser Ser Thr Trp Thr Ser Gly Pro Val
1285 1290 1295
Thr Asn Val Gly Gly Thr Ala Asp Gly Asp Trp Leu Ala Tyr Gly Glu

1300 1305 1310
Ile Asp Leu Gly Ser Ala Ala Leu Asp Gln Leu Ser Val His Tyr Val
1315 1320 1325
His Asn Ser Asn Arg Ser Gly Arg Asn Ser Ala Leu Ser Val Tyr Leu
1330 1335 1340
Asp Ala Phe Asp Pro Ala Asn Pro Gly Glu Pro Phe Val Thr Val Pro
1345 1350 1355 1360
Leu Ala Asn Thr Gly Ser Ser Trp Thr Thr Asp Gly Thr Ala Val Val

1365 1370 1375

Asp Leu Pro Ser Thr Val Arg Gly Lys His Gln Val Trp Val Arg Leu

1380 1385 1390

Ser Thr Glu Ala Tyr Ala Asp His Pro Tyr Val Ala Asn Leu Asp Ser

1395 1400 1405

Met Arg Phe Phe Thr Asp Ala Tyr Asp Val Glu Val Pro Pro Thr Asp

1410 1415 1420

Thr Ala Ala Leu Ala Ala Val Val Asp Ala Ala Gly Thr Pro Glu Ala

1425 1430 1435 1440

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

1445 1450 1455

Leu Ala Ala Ala Arg Ser Val Leu Ala Asp Ala Gly Ala Thr Gln Ala

1460 1465 1470

Gln Ala Asp Glu Arg Ala Arg Arg Leu Gly Leu Ala Thr Asp Gln Leu

1475 1480 1485

Val Pro Ala Glu Arg Arg Arg Leu Glu Asn Leu Val Ala Ser Ala Glu

1490 1495 1500

Ala Leu Thr Asp Glu Gly Tyr Ser Pro Glu Ser Trp Gln Ala Phe Arg

1505 1510 1515 1520

Thr Ala Leu Ala Ala Ala Thr Gly Thr Leu Asp Asp Ala Ala Ala Ser

1525 1530 1535

Asp Glu Ala Leu His Asp Ala Arg Leu Ala Leu Gln Gly Ala Val Asp

1540 1545 1550

Ala Leu Glu Glu Pro Ala Asp Val Val Leu Val Glu Val Glu Val Ser

1555 1560 1565

Pro Arg Cys Leu Ala Gly Lys Pro Tyr Val Ala Val Arg Ala Val Asn

1570 1575 1580

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

1585 1590 1595 1600

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

1605 1610 1615

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

1620 1625 1630

Gly Ala Asp Gly Thr Gln Thr Val Glu Gln Val Val Thr Val Pro Ser

1635 1640 1645

Cys Ser

1650

<210> 5

<211> 1079

<212> PRT

<213> Yarrowia lipolytica

<400> 5

Met Tyr Ser His Phe Asn Asn Glu Pro Val Ala Lys Arg Val Asn Asn

1 5 10 15

Leu Phe Thr Asp Arg Leu Arg Gln Phe Thr Ser Asp Gly Glu Tyr Arg

20 25 30

Ser Leu Asn Leu Pro Ala Phe Tyr Glu Arg Glu Arg Leu Asp Gly Lys

35 40 45

Asn His Val Ala Ile Glu Thr Tyr Ala Val Ser Asp Leu Arg Arg Pro

50 55 60

Leu Phe Lys Asp Ala Leu Lys Glu Ala Asp Gly His Trp Lys Pro Ala

65 70 75 80

Lys Lys Gly Ser Glu Tyr Gly Pro Ser Trp Ala Thr His Trp Phe Lys

85 90 95

Ile Gln Val Cys Val Pro Pro Glu Trp Lys Lys Asn Tyr Tyr Lys Lys

100 105 110

Gly Asp Leu Val Val Phe Asn Trp Asn Leu Asn Cys Glu Gly Leu Val

115 120 125

Phe Ser Glu Ser Gly Glu Ala Leu Ile Gly Leu Ser Gly Glu Glu Arg

130 135 140

Arg Glu Trp Pro Ile Pro Asp Asn Trp Phe Asp Gly Lys Cys His Thr

145 150 155 160

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

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

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

Pro Val Gly Gly Ala Trp Thr Glu Cys Asp Thr Asn Leu Pro Ser Gly
355 360 365
Glu Ser Leu Leu Arg Gln Phe Leu Leu Gly Gln Arg Phe Phe Leu Glu
370 375 380
His Phe Gly Ser Leu Ser Asp Thr Phe Trp Leu Pro Asp Thr Phe Gly
385 390 395 400
Tyr Ser Ala Gln Val Pro Gln Leu Cys Arg Leu Ala Gly Met Asp Arg

405 410 415

Phe Leu Thr Gln Lys Leu Ser Trp Asn Asn Ile Asn Ser Phe Pro Asn
420 425 430
Ser Thr Phe Asn Trp Val Ala Leu Asp Gly Ser Gln Val Leu Cys His
435 440 445
Met Pro Pro Asn Asn Thr Tyr Thr Ser Met Ala Asn Phe Gly Asp Val
450 455 460
Ser Arg Thr Gln Lys Gln Asn Lys Asn Leu Asp Thr Thr Arg Asn Ser
465 470 475 480

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

Gly Glu Leu Tyr Phe Glu Phe His Arg Gly Thr Tyr Thr Ser Gln Ala
545 550 555 560
Gln Thr Lys Lys Gly Asn Arg Val Ser Glu Asn Leu Leu His Asp Val
565 570 575
Glu Leu Leu Ala Thr Leu Ala Ser Ile Arg Asp Ser Ser Tyr Lys Tyr
580 585 590
Pro Phe Ala Gln Leu Glu Ser Leu Trp Glu Asp Val Cys Leu Cys Gln
595 600 605

Phe His Asp Val Leu Pro Gly Ser Cys Ile Glu Met Val Tyr Lys Asp
610 615 620
Val Lys Lys Ile His Gly Arg Val Ile Asp Thr Ala Ser His Leu Ile
625 630 635 640
Asp Lys Ala Ala Ser Ala Leu Gly Leu Ser Gly His Pro Ser Lys Asp
645 650 655

Ser Phe Asp Cys Thr Pro Val Ala Leu Asn Thr Met Pro Trp Ser Arg
660 665 670

Thr Glu Val Val Ala Val Pro Gln Pro His Trp Asp Ala Thr Val Glu
675 680 685

Leu Ala Glu Gly Val Glu Ile Gln Glu Asp Ser Gly Asn Ala Leu Val
690 695 700

Met Met Ser Glu Ser Gly Pro Val Val Thr Thr Gln Ser Val Asp Leu
705 710 715 720

Phe Lys Ser Glu Asp Ala Tyr Ile Leu Glu Asn Ser Gln Val Lys Val
725 730 735

Thr Ile Cys Lys Asp Asp Gly Thr Leu Thr Ser Ile Tyr Asp Lys Glu
740 745 750

Asn Asp Arg Arg Val Leu Ser Gly Thr Gly Asn Arg Leu Val Leu Phe
755 760 765

Asp Asp Gln Pro Leu Ser Trp Gln Ala Trp Asp Thr Glu Val Phe Ser
770 775 780

Leu Gly Lys Lys Gln Tyr Ile Gly Ala Glu Asn Val Thr Arg His Ser
785 790 795 800

Ile Val Ser Ser Gly Pro Leu Arg Ser Thr Val Ala Phe Thr Tyr Glu
805 810 815

Phe Asn Lys Ser Val Val Thr Thr Glu Ile Ser Leu Asp Ala Asn Ser
820 825 830

Pro Leu Val Thr Phe Asn Thr Arg Ala Asp Trp His Glu Thr Cys Lys
835 840 845

Phe Leu Lys Val Glu Phe Pro Val Asp Val His Ser Glu Ser Ala Ser
850 855 860

Tyr Glu Ser Gln Phe Gly Val Val Lys Arg Pro Thr His Tyr Asn Thr
865 870 875 880

Ser Trp Asp Val Ala Lys Phe Glu Val Cys Cys His Lys Phe Ala Asp
885 890 895

Leu Ser Glu Leu Asp Tyr Gly Val Ser Ile Leu Asn Asp Cys Lys Tyr

900 905 910

Gly Phe Ala Thr His Gly Asn Leu Met Arg Leu Ser Leu Leu Arg Ala
915 920 925

Pro Lys Ala Pro Asp Ala His Ala Asp Met Gly His His Glu Phe Lys
930 935 940

Tyr Gly Val Leu Ala His Lys Gly Pro Leu Gly Ala Thr Thr Val Arg
945 950 955 960

Ala Ala Tyr Asn Phe Asn Asn Pro Leu Arg Val Lys Tyr Val Gly Leu
965 970 975

Ser Glu Val Ser Thr Lys Gln Ala Phe Ser Leu Lys Gly Pro Ala Asn
980 985 990

Leu Val Leu Ser Gln Val Lys Arg Ala Glu Val Asp Arg Ser Lys Lys
995 1000 1005

Ser Thr Asn Val Ile Leu Arg Val Tyr Glu Ala Leu Gly Gly Arg Thr
1010 1015 1020

Arg Gly Lys Leu Val Ile Asp Leu Pro Asn Val Val Ser Val Thr Lys
1025 1030 1035 1040

Thr Cys Ala Leu Glu Tyr Ser Lys Glu Lys Gln Val Val Ala Lys Ser
1045 1050 1055

Glu Gly Val Thr Ser Val Asp Ile Ser Leu Arg Ala Phe Glu Val Ala
1060 1065 1070

Thr Tyr Lys Val Glu Leu Ala
1075

<210> 6

<211> 513

<212> PRT

<213> Aspergillus satoi

<400> 6

Met His Leu Pro Ser Leu Ser Leu Ser Leu Thr Ala Leu Ala Ile Ala
1 5 10 15

Ser Pro Ser Ala Ala Tyr Pro His Phe Gly Ser Ser Gln Pro Val Leu

20 25 30

His Ser Ser Ser Asp Thr Thr Gln Ser Arg Ala Asp Ala Ile Lys Ala
35 40 45
Ala Phe Ser His Ala Trp Asp Gly Tyr Leu Gln Tyr Ala Phe Pro His
50 55 60
Asp Glu Leu His Pro Val Ser Asn Gly Tyr Gly Asp Ser Arg Asn Gly
65 70 75 80
Trp Gly Ala Ser Ala Val Asp Ala Leu Ser Thr Ala Val Ile Met Arg
85 90 95

Asn Ala Thr Ile Val Asn Gln Ile Leu Asp His Val Gly Lys Ile Asp
100 105 110
Tyr Ser Lys Thr Asn Thr Thr Val Ser Leu Phe Glu Thr Thr Ile Arg
115 120 125
Tyr Leu Gly Gly Met Leu Ser Gly Tyr Asp Leu Leu Lys Gly Pro Val
130 135 140
Ser Asp Leu Val Gln Asn Ser Ser Lys Ile Asp Val Leu Leu Thr Gln
145 150 155 160

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

Gln Lys Ala Glu Ser Tyr Leu Leu Asn Pro Gln Pro Lys Ser Ala Glu
225 230 235 240
Pro Phe Pro Gly Leu Val Gly Ser Asn Ile Asn Ile Ser Asn Gly Gln
245 250 255
Phe Thr Asp Ala Gln Val Ser Trp Asn Gly Gly Asp Asp Ser Tyr Tyr
260 265 270

Glu Tyr Leu Ile Lys Met Tyr Val Tyr Asp Pro Lys Arg Phe Gly Leu
275 280 285

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

Ile Asn Phe Gly Leu Asp Leu Val Ser Gly Cys His Asp Thr Tyr Asn
355 360 365
Ser Thr Leu Thr Gly Ile Gly Pro Glu Ser Phe Ser Trp Asp Thr Ser
370 375 380
Asp Ile Pro Ser Ser Gln Gln Ser Leu Tyr Glu Lys Ala Gly Phe Tyr
385 390 395 400
Ile Thr Ser Gly Ala Tyr Ile Leu Arg Pro Glu Val Ile Glu Ser Phe
405 410 415

Tyr Tyr Ala Trp Arg Val Thr Gly Gln Glu Thr Tyr Arg Asp Trp Ile
420 425 430
Trp Ser Ala Phe Ser Ala Val Asn Asp Tyr Cys Arg Thr Ser Ser Gly
435 440 445
Phe Ser Gly Leu Thr Asp Val Asn Ala Ala Asn Gly Gly Ser Arg Tyr
450 455 460
Asp Asn Gln Glu Ser Phe Leu Phe Ala Glu Val Met Lys Tyr Ser Tyr
465 470 475 480

Met Ala Phe Ala Glu Asp Ala Ala Trp Gln Val Gln Pro Gly Ser Gly
485 490 495
Asn Gln Phe Val Phe Asn Thr Glu Ala His Pro Val Arg Val Ser Ser
500 505 510
Thr

<210> 7

<211> 1810

<212> PRT

<213> Cellulosimicrobium cellulans

<400> 7

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

Gly Arg Ala Arg Pro Leu Gly Leu Val Leu Ala Ala Ala Leu Ala Val

20 25 30

Pro Leu Gly Val Pro Leu Ala Ala Pro Ala Gly Ala Leu Ala Ala Ala
35 40 45

Pro Ala Ala Ala Ala Glu Pro Gly Asp Phe Ser Ser Ser Phe Glu Ser
50 55 60

Gly Asp Pro Ala Ala Leu Pro Thr Thr Val Ala Glu Arg Asp Gly Ala
65 70 75 80

Pro Trp Gln Ala Asn Val Gly Ser Phe Thr Ala Gly Leu Pro Gly Ser

85 90 95

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

Asn Glu Gly Ala Ala Asn Leu Ala Asp Gly Ser Ser Gly Thr Lys Trp
115 120 125

Leu Ala Phe Ala Ser Thr Gly Trp Val Arg Tyr Glu Phe Ala Glu Pro
130 135 140

Val Ser Phe Val Ala Tyr Thr Met Thr Ser Gly Asp Asp Ala Ala Gly

145 150 155 160

Arg Asp Pro Lys Thr Trp Thr Val Glu Gly Ser Asn Asp Gly Ser Thr
165 170 175

Trp Ala Ala Leu Asp Arg Arg Thr Asp Glu Asp Phe Pro Asn Arg Gln
180 185 190

Gln Thr Arg Thr Phe Glu Leu Glu Ala Pro Thr Ala Ala Tyr Thr Tyr
195 200 205

Leu Arg Leu Asn Val Thr Ala Asn Ser Gly Asp Ser Ile Val Gln Leu

210 215 220

Ala Gly Trp Asp Leu Ser Ala Asp Leu Ser Ala Gly Pro Ser Ala Ala

225 230 235 240

Pro Met Thr Thr Lys Val Gly Thr Gly Pro Arg Val Ser Phe Thr Asn

245 250 255

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

260 265 270

Ser His Leu Ala Asp Gly Glu Thr Tyr Ala Thr Asn Val Leu Tyr Asp

275 280 285

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

290 295 300

Phe Pro Glu Leu Leu Asp Asp Leu Gln Tyr Pro Ser Thr Tyr Ala Ala

305 310 315 320

Val Asp Val Leu Phe Thr Asp Gly Thr Tyr Leu Ser Asp Leu Gly Ala

325 330 335

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

340 345 350

Lys Ile Leu Tyr Ala Asp Gln Trp Asn Ser Val Arg Val Asp Leu Gly

355 360 365

Asp Val Ala Glu Gly Lys Thr Val Asp Gln Val Leu Leu Gly Tyr Asp

370 375 380

Asn Pro Gly Gly His Ala Gly Thr Lys Phe Ala Gly Trp Leu Asp Asp

385 390 395 400

Val Glu Ile Thr Ala Glu Pro Ala Thr Ile Asp Gly Ser Ser Leu Ala

405 410 415

Asn Tyr Val Asp Thr Arg Arg Gly Thr Leu Ala Ser Gly Ser Phe Ser

420 425 430

Arg Gly Asn Asn Ile Pro Ala Thr Ala Thr Pro Asn Gly Phe Asn Phe

435 440 445

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

450 455 460

His Lys Ala Asn Asn Ala Asn Asn Lys Pro Val Leu Gln Gly Phe Gly

465 470 475 480

Ile Ser His Glu Pro Ser Pro Trp Met Gly Asp Arg Asn Gln Leu Thr

485 490 495

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

500 505 510

Arg Gly Leu Glu Phe Asp His Ala Asp Glu Thr Ala Arg Pro Asp Tyr

515 520 525

Tyr Gly Val Thr Phe Thr Asn Gly Ser Ala Ile Glu Ala Thr Pro Thr

530 535 540

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

545 550 555 560

Val Leu Val Asp Lys Val Asp Gly Ser Ser Lys Leu Thr Tyr Asp Gln

565 570 575

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

580 585 590

Val Gly Arg Thr Arg Met Phe Val Ala Gly Thr Phe Asp Arg Ser Pro

595 600 605

Thr Ala Val Gly Thr Ala Ala Gly Asn Arg Ala Asp Ala Arg Phe Ala

610 615 620

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

625 630 635 640

Ser Phe Ile Ser Leu Asp Gln Ala Arg Lys Asn Leu Asp Leu Glu Val

645 650 655

Thr Gly Lys Thr Phe Thr Glu Val Lys Ala Ala Ala Ala Gln Ala Trp

660 665 670

Asn Asp Arg Leu Gly Val Ile Glu Val Glu Gly Ala Ser Glu Asp Gln

675 680 685

Leu Val Thr Leu Tyr Ser Asn Leu Tyr Arg Leu Asn Leu Tyr Pro Asn

690 695 700

Ser Gln Phe Glu Asn Thr Gly Thr Ala Gln Glu Pro Val Tyr Arg Tyr
705 710 715 720
Ala Ser Pro Val Ser Ala Thr Thr Gly Ser Ala Thr Asp Thr Gln Thr

725 730 735
Asn Ala Lys Ile Val Asp Gly Lys Ile Tyr Val Asn Asn Gly Phe Trp
740 745 750
Asp Thr Tyr Arg Thr Ala Trp Pro Ala Tyr Ser Leu Leu Tyr Pro Glu
755 760 765
Leu Ala Ala Glu Leu Val Asp Gly Phe Val Gln Gln Tyr Arg Asp Gly
770 775 780
Gly Trp Ile Ala Arg Trp Ser Ser Pro Gly Tyr Ala Asp Leu Met Thr

785 790 795 800
Gly Thr Ser Ser Asp Val Ala Phe Ala Asp Ala Tyr Leu Lys Gly Ser
805 810 815
Leu Pro Thr Gly Thr Ala Leu Glu Ala Tyr Asp Ala Ala Leu Arg Asn
820 825 830
Ala Thr Val Ala Pro Pro Ser Asn Ala Val Gly Arg Lys Gly Leu Gln
835 840 845
Thr Ser Pro Phe Leu Gly Phe Thr Pro Glu Ser Thr His Glu Ser Val

850 855 860
Ser Trp Gly Leu Glu Gly Leu Val Asn Asp Phe Gly Ile Gly Asn Met
865 870 875 880
Ala Ala Ala Leu Ala Glu Asp Pro Ala Thr Pro Glu Glu Arg Arg Glu
885 890 895
Thr Leu Arg Glu Glu Ser Ala Tyr Phe Leu Glu Arg Ala Thr His Tyr
900 905 910
Val Glu Leu Phe Asp Pro Glu Val Asp Phe Phe Val Pro Arg His Glu

915 920 925
Asp Gly Thr Trp Ala Val Asp Pro Glu Thr Tyr Asp Pro Glu Ala Trp
930 935 940
Gly Gly Gly Tyr Thr Glu Thr Asn Gly Trp Asn Phe Ala Phe His Ala

945 950 955 960

Pro Gln Asp Gly Gln Gly Leu Ala Asn Leu Tyr Gly Gly Lys Gln Gly

965 970 975

Leu Glu Asp Lys Leu Asp Glu Phe Phe Ser Thr Pro Glu Lys Gly Ala

980 985 990

Gly Asn Gly Gly Ile His Glu Gln Arg Glu Ala Arg Asp Val Arg Met

995 1000 1005

Gly Gln Trp Gly Met Ser Asn Gln Val Ser His His Ile Pro Trp Leu

1010 1015 1020

Tyr Asp Ala Ala Gly Ala Pro Ser Lys Ala Gln Glu Lys Val Arg Glu

1025 1030 1035 1040

Val Thr Arg Arg Leu Phe Val Gly Ser Glu Ile Gly Gln Gly Tyr Pro

1045 1050 1055

Gly Asp Glu Asp Asn Gly Glu Met Ser Ser Trp Trp Ile Phe Ala Ser

1060 1065 1070

Leu Gly Phe Tyr Pro Leu Gln Val Gly Ser Asp Gln Tyr Ala Val Gly

1075 1080 1085

Ser Pro Leu Phe Asp Lys Ala Thr Val His Leu Pro Asp Gly Asp Leu

1090 1095 1100

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

1105 1110 1115 1120

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

1125 1130 1135

Asp Leu Ser Gly Gly Thr Thr Leu Asp Phe Val Met Gly Pro Glu Pro

1140 1145 1150

Ser Asp Trp Gly Thr Gly Glu Asp Asp Ala Pro Pro Ser Leu Thr Glu

1155 1160 1165

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

1170 1175 1180

Gly Thr Thr Thr Val Ala Asp Gly Asp Ala Thr Thr Ser Ala Ala Ala

1185 1190 1195 1200

Leu Thr Asp Asn Thr Ser Gly Thr Arg Thr Thr Phe Ala Thr Thr Thr
1205 1210 1215
Pro Ser Ile Thr Trp Ala Gly Asn Gly Ile Arg Pro Thr Val Gly Ser
1220 1225 1230
Tyr Thr Leu Thr Ser Gly Ala Ser Gly Thr Ala Ser Pro Ser Ala Trp

1235 1240 1245
Thr Leu Glu Gly Ser Asp Asp Gly Glu Thr Trp Thr Thr Leu Asp Glu
1250 1255 1260
Arg Ser Gly Glu Gln Phe Arg Trp Ala Leu Gln Thr Arg Pro Phe Thr
1265 1270 1275 1280
Val Ala Glu Pro Thr Ala Phe Ala Arg Tyr Arg Val Thr Val Thr Ala
1285 1290 1295
Thr Ser Gly Ser Gly Ala Leu Ser Leu Ala Glu Val Glu Leu Leu Ala

1300 1305 1310
Asp Pro Lys Glu Ser Gly Ala Glu Glu Leu Thr Leu Ser Ala Ala Pro
1315 1320 1325
Asp Arg Asp Gly Val Thr Gly Arg Glu Val Ser Gly Ser Phe Ala Thr
1330 1335 1340
Leu Thr Gly Val Glu Gly Asp Val Ala Ala Leu Asp Val Gln Val Ala
1345 1350 1355 1360
Phe Gly Asp Gly Ser Glu Pro Val Ala Gly Thr Leu Arg Ala Gly Ala

1365 1370 1375
Phe Gly Gly Tyr Ala Val Asp Ala Ala His Thr Trp Thr Ala Pro Gly
1380 1385 1390
Val Tyr Pro Val Thr Val Thr Val Ser Gly Glu Gly Ile Glu Thr Val
1395 1400 1405
Ser Ala Ser Ser Tyr Val Ser Val Ser Leu Leu Arg Glu Gly Ser Leu
1410 1415 1420
Leu Ala Ala Tyr Asp Asn Val Cys Ile Gly Asp Ala Gly Thr Thr Val

1425 1430 1435 1440
Gly Ser Cys Asp Gly Gln Gly Val Phe Phe Asp Arg Ala Gln Leu Ala

1445 1450 1455

Ala Lys Gly Phe Val Gln Gly Glu Arg Ala Thr Val Pro Gly Thr Asp

1460 1465 1470

Leu Ala Phe Asp Val Pro Ala Val Pro Ala Gly Gln Pro Asp Asn Ala

1475 1480 1485

Thr Gly Asp Gly Gln Thr Ile Glu Leu Asp Val Pro Ala Asp Ala Glu

1490 1495 1500

Gln Leu Ser Val Ile Gly Thr Gly Thr Glu Lys Asn Gln Gln Ala Thr

1505 1510 1515 1520

Gly Thr Leu Thr Phe Asp Asp Gly Ser Thr Gln Pro Ile Asp Leu Ser

1525 1530 1535

Phe Gly Asp Trp Ser Gly Ala Ala Arg Asn Pro Val Phe Gly Asn Ile

1540 1545 1550

Pro Val Ala Val Thr Asp Ser Arg Leu Arg Gly Gly Ser Pro Gln Thr

1555 1560 1565

Gly Thr Pro Ala Ala Phe Phe Ala Thr Ala Pro Ile Thr Leu Pro Glu

1570 1575 1580

Gly Lys Arg Pro Val Ser Leu Thr Leu Pro Asp Gln Pro Gly Glu Leu

1585 1590 1595 1600

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

1605 1610 1615

Phe Ala Glu His Pro Ala Leu Glu Val Thr Ala Ala Glu Gly Val Thr

1620 1625 1630

Leu Ala Val Gly Gln Thr Ser Asp Val Ala Leu Ala Gln Val Ala Gly

1635 1640 1645

Gly Arg Glu Gly Ala Asp Leu Arg Ala Ala Val Thr Trp Gly Asp Gly

1650 1655 1660

Ser Asp Val Ala Ala Gly Ala Val Thr Asp Gly Ser Val Ser Gly Ser

1665 1670 1675 1680

His Ala Tyr Thr Ala Ala Gly Thr Tyr Thr Ala Tyr Val Val Val Asp

1685 1690 1695

Asp Gly Trp Thr Ser Gln Val Val Glu Val Pro Val Thr Val Thr Glu
1700 1705 1710
Ala Glu Pro Ala Leu Ala Val Asp Val Thr Val Ser Thr Arg Cys Leu
1715 1720 1725
Ala Gly Lys Ala Tyr Val Ala Val Arg Ala Glu Asn Gly Glu Asp Val
1730 1735 1740
Pro Leu Ala Ile Arg Leu Val Thr Pro Phe Gly Thr Lys Glu Val Ala

1745 1750 1755 1760
Ala Val Ala Pro Gly Ala Asn Ala Tyr Ser Phe Ala Thr Arg Val Thr
1765 1770 1775
Ala Val Glu Ala Gly Thr Val Thr Val Glu Ala Thr Arg Gly Thr Gly
1780 1785 1790
Asp Glu Glu Val Thr Ala Ser Ile Gln Ala Asp Tyr Ala Ala Val Thr
1795 1800 1805
Cys Gly
1810

<210> 8

<211> 282

<212>
PRT

<213> Aspergillus oryzae

<400> 8

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

Gln His Val Asp Ser Ile Gly His Gly Thr His Val Ser Gly Thr Ile
65 70 75 80
Ala Gly Lys Thr Tyr Gly Ile Ala Lys Lys Ala Ser Ile Leu Ser Val

85 90 95

Lys Val Phe Gln Gly Glu Ser Ser Ser Thr Ser Val Ile Leu Asp Gly

100 105 110

Phe Asn Trp Ala Ala Asn Asp Ile Val Ser Lys Lys Arg Thr Ser Lys

115 120 125

Ala Ala Ile Asn Met Ser Leu Gly Gly Gly Tyr Ser Lys Ala Phe Asn

130 135 140

Asp Ala Val Glu Asn Ala Phe Glu Gln Gly Val Leu Ser Val Val Ala

145 150 155 160

Ala Gly Asn Glu Asn Ser Asp Ala Gly Gln Thr Ser Pro Ala Ser Ala

165 170 175

Pro Asp Ala Ile Thr Val Ala Ala Ile Gln Lys Ser Asn Asn Arg Ala

180 185 190

Ser Phe Ser Asn Phe Gly Lys Val Val Asp Val Phe Ala Pro Gly Gln

195 200 205

Asp Ile Leu Ser Ala Trp Ile Gly Ser Ser Ser Ala Thr Asn Thr Ile

210 215 220

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

225 230 235 240

Leu Ala Ala Leu Glu Asn Leu Asp Gly Pro Ala Ala Val Thr Lys Arg

245 250 255

Ile Lys Glu Leu Ala Thr Lys Asp Val Val Lys Asp Val Lys Gly Ser

260 265 270

Pro Asn Leu Leu Ala Tyr Asn Gly Asn Ala

275 280

<210> 9

<211> 403

<212> PRT

<213> Aspergillus oryzae

<400> 9

Met Gln Ser Ile Lys Arg Thr Leu Leu Leu Leu Gly Ala Ile Leu Pro

1 5 10 15

Ala Val Leu Gly Ala Pro Val Gln Glu Thr Arg Arg Ala Ala Glu Lys
20 25 30

Leu Pro Gly Lys Tyr Ile Val Thr Phe Lys Pro Gly Ile Asp Glu Ala
35 40 45
Lys Ile Gln Glu His Thr Thr Trp Ala Thr Asn Ile His Gln Arg Ser
50 55 60
Leu Glu Arg Arg Gly Ala Thr Gly Gly Asp Leu Pro Val Gly Ile Glu
65 70 75 80
Arg Asn Tyr Lys Ile Asn Lys Phe Ala Ala Tyr Ala Gly Ser Phe Asp
85 90 95

Asp Ala Thr Ile Glu Glu Ile Arg Lys Asn Glu Asp Val Ala Tyr Val
100 105 110
Glu Glu Asp Gln Ile Tyr Tyr Leu Asp Gly Leu Thr Thr Gln Lys Ser
115 120 125
Ala Pro Trp Gly Leu Gly Ser Ile Ser His Lys Gly Gln Gln Ser Thr
130 135 140
Asp Tyr Ile Tyr Asp Thr Ser Ala Gly Glu Gly Thr Tyr Ala Tyr Val
145 150 155 160

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

Ser Ser Thr Ser Val Ile Leu Asp Gly Phe Asn Trp Ala Ala Asn Asp
225 230 235 240
Ile Val Ser Lys Lys Arg Thr Ser Lys Ala Ala Ile Asn Met Ser Leu
245 250 255
Gly Gly Gly Tyr Ser Lys Ala Phe Asn Asp Ala Val Glu Asn Ala Phe

260 265 270

Glu Gln Gly Val Leu Ser Val Val Ala Ala Gly Asn Glu Asn Ser Asp

275 280 285

Ala Gly Gln Thr Ser Pro Ala Ser Ala Pro Asp Ala Ile Thr Val Ala

290 295 300

Ala Ile Gln Lys Ser Asn Asn Arg Ala Ser Phe Ser Asn Phe Gly Lys

305 310 315 320

Val Val Asp Val Phe Ala Pro Gly Gln Asp Ile Leu Ser Ala Trp Ile

325 330 335

Gly Ser Ser Ser Ala Thr Asn Thr Ile Ser Gly Thr Ser Met Ala Thr

340 345 350

Pro His Ile Val Gly Leu Ser Leu Tyr Leu Ala Ala Leu Glu Asn Leu

355 360 365

Asp Gly Pro Ala Ala Val Thr Lys Arg Ile Lys Glu Leu Ala Thr Lys

370 375 380

Asp Val Val Lys Asp Val Lys Gly Ser Pro Asn Leu Leu Ala Tyr Asn

385 390 395 400

Gly Asn Ala

<210> 10

<211> 1272

<212> DNA

<213> Artificial Sequence

<220>

<223> codon optimized sequence encoding alkaline

protease

<400> 10

ggatccatgc agtccattaa gcgaactctg ctgctgctgg gagccattct gcccgccgtg 60

ctgggagccc ccgttcagga gacccgacga gccgccgaga agctccccgg caagtacatt 120

gtcaccttca agcctggtat cgacgaggct aagattcagg agcacaccac ttgggccacc 180

aacatccatc agcgatccct cgagcgacga ggagccaccg gcggtgacct gcctgtggga 240

atcgagcgaa actacaagat taacaagttc gccgcttacg ctggatcttt tgacgatgcc 300

accatcgagg agattcgaaa gaacgaggac gtcgcttacg tggaggaaga ccagatctac 360

tacctcgatg gtctgaccac tcagaagtcc gctccttggg gcctgggctc catctctcac 420

aagggacagc agtcgactga ctacatctac gatacctccg ctggcgaggg tacttacgcc 480

tacgtcgtgg actccggtgt taacgtcgat cacgaggagt ttgagggacg agcctctaag 540

gcttacaacg ccgctggagg ccagcatgtg gactctatcg gacacggcac ccatgtttcg 600

ggtactattg ccggaaagac ctacggcatc gccaagaagg cttctattct ctcggtgaag 660

gttttccagg gagagtcctc ttcgacctct gtcatcctgg acggctttaa ctgggccgct 720

aacgatattg tgtctaagaa gcgaacctcg aaggccgcta tcaacatgtc cctcggtgga 780

ggctactcta aggccttcaa cgacgctgtt gagaacgcct ttgagcaggg tgtcctgtct 840

gttgtggctg ctggtaacga gaactctgac gctggacaga cctcccctgc ttctgctcct 900

gatgccatca ctgtggccgc tattcagaag tccaacaacc gagcttcgtt ctccaacttt 960

ggcaaggtgg ttgacgtttt cgcccccgga caggatatcc tctctgcttg gattggctcc 1020

tcttcggcca ccaacactat ctcgggcacc tccatggcca ctccccacat tgtcggtctg 1080

tccctctacc tggctgctct ggagaacctg gacggacctg ccgctgttac caagcgaatc 1140

aaggagctgg ctactaagga cgtcgtgaag gatgtcaagg gttctcctaa cctgctcgcc 1200

tacaacggca acgcttctgg cggcggagga catcaccacc atcaccatca ccaccatcat 1260

tgataaccta gg 1272

<210

> 11

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> peptide from Jack bean mannosidase

<400> 11

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

1 5 10 15

Ala Val Gln Gly

20

<210> 12

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> alkaline protease substrate

<400> 12

Ile Gln Asn Cys Pro Leu Gly

1 5