



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0021075
(43) 공개일자 2016년02월24일

- | | |
|--|--|
| <p>(51) 국제특허분류(Int. Cl.)
 <i>C12N 15/63</i> (2006.01) <i>A61K 39/108</i> (2006.01)
 <i>C07K 14/245</i> (2006.01) <i>C07K 14/605</i> (2006.01)
 <i>C12N 15/52</i> (2006.01) <i>C12N 9/10</i> (2006.01)
 <i>C12P 19/04</i> (2006.01) <i>C12P 19/18</i> (2006.01)
 <i>C12P 21/00</i> (2006.01)</p> <p>(52) CPC특허분류
 <i>C12N 15/63</i> (2013.01)
 <i>A61K 39/0258</i> (2013.01)</p> <p>(21) 출원번호 10-2015-7028859
 (22) 출원일자(국제) 2014년03월14일
 심사청구일자 없음
 (85) 번역문제출일자 2015년10월13일
 (86) 국제출원번호 PCT/US2014/026990
 (87) 국제공개번호 WO 2014/152137
 국제공개일자 2014년09월25일
 (30) 우선권주장
 61/785,586 2013년03월14일 미국(US)</p> | <p>(71) 출원인
 글리코비아, 인크.
 미국 14850 뉴욕주 이타카 스위트 104 쏘우드 드라이브 33</p> <p>(72) 발명자
 피셔, 아담, 씨.
 미국 14853 뉴욕주 이타카 웨일 홀 맥거번 센터 410 글리코비아 인크. 내
 메리트, 주디쓰, 에이치.
 미국 14853 뉴욕주 이타카 웨일 홀 맥거번 센터 410 글리코비아 인크. 내
 (뒷면에 계속)</p> <p>(74) 대리인
 양영준, 김영</p> |
|--|--|

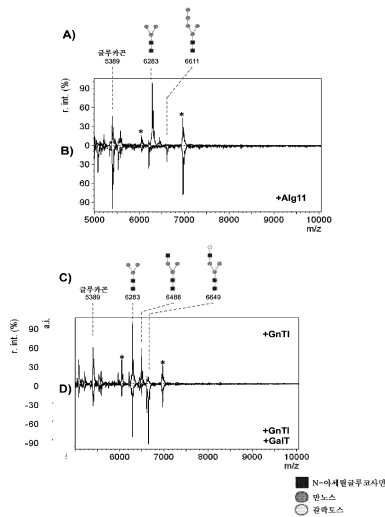
전체 청구항 수 : 총 70 항

(54) 발명의 명칭 **올리고사카라이드 조성물, 당단백질, 및 원핵생물에서의 이들의 제조 방법**

(57) 요약

다양한 올리고사카라이드 조성물 및 당단백질을 제조하기 위한 방법 및 조성물이 개시되어 있다. 원핵 속주 세포는 글리코실화 경로를 속주 세포로 도입함으로써 인간-유사, 예를 들어 고-만노스, 하이브리드 및 복합 글리코실화 패턴을 생성하기에 효과적인 조건 하에 배양된다.

대표도 - 도11



(52) CPC특허분류

C07K 14/245 (2013.01)
C07K 14/605 (2013.01)
C12N 15/52 (2013.01)
C12N 9/1051 (2013.01)
C12N 9/1081 (2013.01)
C12P 19/04 (2013.01)
C12P 19/18 (2013.01)
C12P 21/005 (2013.01)
Y02P 20/52 (2015.11)

(72) 발명자

해밀턴, 브라이언, 에스.

미국 14853 뉴욕주 이타카 웨일 홀 맥거번 센터
410 글리코비아 인크. 내

발데라마-린콘, 주안 디.

미국 14853 뉴욕주 이타카 웨일 홀 맥거번 센터
410 글리코비아 인크. 내

델리사, 매튜, 피.

미국 14853 뉴욕주 이타카 웨일 홀 맥거번 센터
410 글리코비아 인크. 내

명세서

청구범위

청구항 1

UDP-GlcNAc 잔기의 말단 만노스 잔기로의 전이를 촉매하는 하나 이상의 N-아세틸글루코사민일 트랜스퍼라제 효소 활성 (EC 2.4.1.101; EC 2.4.1.143; EC 2.4.1.145; EC 2.4.1.155; EC 2.4.1.201)을 발현하기 위해 말단 만노스 잔기를 갖는 올리고사카라이드 조성물을 생성하는 재조합 원핵 숙주 세포를 배양하는 단계를 포함하며, 상기 배양 단계는 말단 GlcNAc 잔기를 갖는 올리고사카라이드 조성물을 생성하기에 효과적인 조건 하에 수행되는 것인, 올리고사카라이드 조성물의 제조 방법.

청구항 2

제1항에 있어서, 배양 단계가 UDP-갈락토스 잔기의 상기 말단 GlcNAc 잔기로의 전이를 촉매하는 하나 이상의 갈락토실트랜스퍼라제 효소 활성 (EC 2.4.1.38)을 발현하기 위한 숙주 세포를 추가로 포함하며, 상기 배양 단계는 말단 갈락토스 잔기를 갖는 올리고사카라이드 조성물을 생성하기에 효과적인 조건 하에서 수행되는 것인 방법.

청구항 3

제1항에 있어서, 배양 단계가 CMP-NANA 잔기의 상기 말단 갈락토스 잔기로의 전이를 촉매하는 하나 이상의 시알릴트랜스퍼라제 효소 활성 (EC 2.4.99.4 및 EC 2.4.99.1)을 발현하기 위한 숙주 세포를 추가로 포함하며, 상기 배양 단계는 말단 시알산 잔기를 갖는 올리고사카라이드 조성물을 생성하기에 효과적인 조건 하에서 수행되는 것인 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 배양 단계가 하나 이상의 진행 UDP-GlcNAc 트랜스퍼라제 효소 활성 (EC 2.4.1.141 또는 EC 2.4.1.145) 및 하나 이상의 진행 만노실트랜스퍼라제 효소 활성 (EC 2.4.1.142, EC 2.4.1.132)을 발현하기 위한 숙주 세포를 포함하며, 여기서 효소는 숙주 세포에 진행 글리코실트랜스퍼라제 활성을 부여하는 것인 방법.

청구항 5

제1항 내지 제3항 중 어느 한 항에 있어서, 배양 단계가 하나 이상의 효소의 융합을 발현하기 위한 숙주 세포를 추가로 포함하며, 여기서 상기 융합은 다음의 용해도 인핸서: DsbA, GlpE, GST, MBP, MstX, NusA 및 TrxA로부터 선택되는 용해도 인핸서 중 적어도 하나를 포함하는 것인 방법.

청구항 6

제1항 내지 제3항 중 어느 한 항에 있어서, 배양 단계가 산화 조건 하에 있는 것인 방법.

청구항 7

제1항 내지 제3항 중 어느 한 항에 있어서, 배양 단계가 포스포마노뮤타제 효소 활성 (ManB) (EC 5.4.2.8), 만노스-1-포스페이트 구아닐릴트랜스퍼라제 효소 활성 (ManC) (EC 2.7.7.13) 및 글루타민-프록토스-6-포스페이트 트랜스아미나제 효소 활성 (GlmS) (EC 2.6.1.16) 중 하나 이상을 발현하기 위한 숙주 세포를 포함하며, 여기서 ManB 및 ManC는 GDP-만노스 합성을 촉매하고, GlmS는 UDP-GlcNAc 합성을 촉매하는 것인 방법.

청구항 8

제1항 내지 제3항 중 어느 한 항에 있어서, 배양 단계가 GDP-D-만노스 데히드라타제 효소 활성 (EC 4.2.1.47)에 있어서 감쇠를 갖는 숙주 세포를 포함하는 것인 방법.

청구항 9

제1항 내지 제3항 중 어느 한 항에 있어서, 배양 단계가 플립라제 효소 활성을 발현하기 위한 숙주 세포를 포함

하는 것인 방법.

청구항 10

제1항 내지 제3항 중 어느 한 항에 있어서, 배양 단계가 올리고사카릴 트랜스퍼라제 효소 활성 (EC 2.4.1.119) 을 발현하기 위한 숙주 세포를 포함하는 것인 방법.

청구항 11

제1항 내지 제3항 중 어느 한 항에 있어서, 관심 단백질을 코딩하는 유전자를 숙주 세포에 도입하여 숙주 세포 가 글리코실화 단백질을 생성하도록 하는 것을 추가로 포함하는 방법.

청구항 12

제9항에 있어서, 올리고사카라이드 조성물이 단백질에 N-연결된 것인 방법.

청구항 13

제11항에 있어서, 배양 단계가 고유 항원에 결합하는 Fv 부분 및 보존 아스파라긴 잔기에서 글리코실화되는 Fc 부분, 디아바디, scFv, scFv-Fc, scFv-CH, Fab 및 scFab 중 하나 이상을 포함하는 항체로서 상기 글리코실화 단백질을 발현하기 위한 재조합 원핵 숙주 세포를 포함하는 것인 방법.

청구항 14

제11항에 있어서, 글리코실화 단백질이 시토카인, 예컨대 인터페론, G-CSF, 응고 인자, 예컨대 인자 VIII, 인자 IX, 및 인간 단백질 C, 가용성 IgE 수용체 α -쇄, IgG, IgG 단편, IgM, 인터류킨, 우로키나제, 키마제, 및 우레 아 트립신 억제제, IGF-결합 단백질, 표피 성장 인자, 성장 호르몬-방출 인자, 아넥신 V 용합 단백질, 안지오스 타틴, 혈관 내피 성장 인자-2, 골수성 전구 억제 인자-1, 오스테오프로테게린, α -1 항트립신, DNase II, α -페 토(feto) 단백질, AAT, rhTBP-1 (일명 TNF 결합 단백질 1), TACI-Ig (막형단 활성화제 및 칼슘 조절제 및 시클 로필린 리간드 인터랙터(interactor)), FSH (여포 자극 호르몬), GM-CSF, 글루카곤, 글루카곤 펩티드, GLP-1 w/ 및 w/o FC (글루카곤 유사 단백질 1) IL-1 수용체 효능제, sTNFr (일명 가용성 TNF 수용체 Fc 용합), CTLA4-Ig (세포독성 T 림프구 조합 항원 4-Ig), 수용체, 호르몬, 예컨대 인간 성장 호르몬, 에리트로포이에틴, 펩티드, 스테이플드(stapled) 펩티드, 인간 백신, 동물 백신, 혈청 알부민 및 효소, 예컨대 ATIII, 재조합 인간 트롬빈(rhThrombin), 글루코세레브로시다제 및 아스파라기나제를 포함하는 것인 방법.

청구항 15

제1항 내지 제3항 중 어느 한 항에 있어서, 배양 단계가 MBP에 작동가능하게 융합된 하나 이상의 N-아세틸글루 코사민일 트랜스퍼라제, 갈락토실트랜스퍼라제 또는 시알릴트랜스퍼라제를 발현하기 위한 숙주 세포를 포함하는 것인 방법.

청구항 16

제2항에 있어서, 배양 단계가 상기 갈락토실트랜스퍼라제 효소 활성을 발현하기 위한 산화 숙주 세포를 포함하 는 것인 방법.

청구항 17

제1항에 있어서, 올리고사카라이드 조성물이 $GlcNAc_{1-5}Man_3GlcNAc_2$ 및 $Man_3GlcNAc_2$ 를 포함하는 것인 방법.

청구항 18

제1항에 있어서, 올리고사카라이드 조성물이 대부분 $GlcNAcMan_3GlcNAc_2$ 또는 $GlcNAc_2Man_3GlcNAc_2$ 인 방법.

청구항 19

제2항에 있어서, 올리고사카라이드 조성물이 $Gal_{1-5}GlcNAc_{1-5}Man_3GlcNAc_2$ 및 $Man_3GlcNAc_2$ 를 포함하는 것인 방법.

청구항 20

제2항에 있어서, 올리고사카라이드 조성물이 대부분 $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$, $\text{GalGlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 또는 $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 인 방법.

청구항 21

제3항에 있어서, 올리고사카라이드 조성물이 $\text{NANA}_{1-5}\text{Gal}_{1-5}\text{GlcNAc}_{1-5}\text{Man}_3\text{GlcNAc}_2$ 를 포함하는 것인 방법.

청구항 22

제3항에 있어서, 올리고사카라이드 조성물이 대부분 $\text{NANAGalGlcNAcMan}_3\text{GlcNAc}_2$ 또는 $\text{NANA}_2\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 인 방법.

청구항 23

제1항 내지 제3항 중 어느 한 항에 있어서, 숙주 세포가 글리세롤 또는 피루베이트의 존재 하에 배양되는 것인 방법.

청구항 24

UDP-GlcNAc 잔기의 말단 만노스 잔기로의 전이를 촉매하는 하나 이상의 N-아세틸글루코사민일 트랜스퍼라제 효소 활성 (EC 2.4.1.101; EC 2.4.1.143; EC 2.4.1.145; EC 2.4.1.155; EC 2.4.1.201)을 포함하는, 말단 만노스 잔기를 갖는 올리고사카라이드 조성물을 생성하는 재조합 원핵 숙주 세포이며, 여기서 상기 숙주 세포는 말단 GlcNAc 잔기를 갖는 올리고사카라이드 조성물을 생성하는 것인 재조합 원핵 숙주 세포.

청구항 25

제24항에 있어서, UDP-갈락토스 잔기의 상기 말단 GlcNAc 잔기로의 전이를 촉매하는 하나 이상의 갈락토실트랜스퍼라제 효소 활성 (EC 2.4.1.38)을 추가로 포함하며, 말단 갈락토스 잔기를 갖는 올리고사카라이드 조성물을 생성하는 숙주 세포.

청구항 26

제25항에 있어서, CMP-NANA 잔기의 상기 말단 갈락토스 잔기로의 전이를 촉매하는 하나 이상의 시알릴트랜스퍼라제 효소 활성 시알릴T (EC 2.4.99.4 및 시알릴T (EC 2.4.99.1)를 추가로 포함하며, 말단 시알산 잔기를 갖는 올리고사카라이드 조성물을 생성하는 숙주 세포.

청구항 27

제24항 내지 제26항 중 어느 한 항에 있어서, 하나 이상의 진행 UDP-GlcNAc 트랜스퍼라제 효소 활성 (EC 2.4.1.141 또는 EC 2.4.1.145) 및 하나 이상의 진행 만노실트랜스퍼라제 효소 활성 (EC 2.4.1.142, EC 2.4.1.132)을 포함하며, 상기 효소는 숙주 세포에 진행 글리코실트랜스퍼라제 활성을 부여하는 것인 숙주 세포.

청구항 28

제24항 내지 제26항 중 어느 한 항에 있어서, 하나 이상의 효소의 융합을 추가로 포함하며, 상기 융합은 DsbA, GlpE, GST, MBP, MstX, NusA 및 TrxA 중 적어도 하나를 포함하는 것인 숙주 세포.

청구항 29

제24항 내지 제26항 중 어느 한 항에 있어서, 산화 숙주인 숙주 세포.

청구항 30

제24항 내지 제26항 중 어느 한 항에 있어서, 다음의 효소: 포스포마노무타제 효소 활성 (ManB) (EC 5.4.2.8), 만노스-1-포스페이트 구아닐릴트랜스퍼라제 효소 활성 (ManC) (EC 2.7.7.13) 및 글루타민-프록토스-6-포스페이트 트랜스아미나제 효소 활성 (GlmS) (EC 2.6.1.16) 중 하나 이상을 포함하며, 여기서 ManB 및 ManC는 GDP-만노스 합성을 촉매하고, GlmS는 UDP-GlcNAc 합성을 촉매하는 것인 숙주 세포.

청구항 31

제24항 내지 제26항 중 어느 한 항에 있어서, GDP-D-만노스 테히드라타제 효소 활성 (EC 4.2.1.47)에 있어서 감쇠를 추가로 포함하는 숙주 세포.

청구항 32

제24항 내지 제26항 중 어느 한 항에 있어서, 플립파제 효소 활성을 추가로 포함하는 숙주 세포.

청구항 33

제24항 내지 제26항 중 어느 한 항에 있어서, 올리고사카틸 트랜스퍼라제 효소 활성 (EC 2.4.1.119)을 추가로 포함하는 숙주 세포.

청구항 34

제24항 내지 제26항 중 어느 한 항에 있어서, 관심 단백질을 코딩하는 유전자를 추가로 포함하여 글리코실화 단백질을 생성하는 숙주 세포.

청구항 35

제34항에 있어서, 단백질에 N-연결된 올리고사카라이드 조성물을 생성하는 숙주 세포.

청구항 36

제34항에 있어서, 고유 항원에 결합하는 Fv 부분 및 보존 아스파라긴 잔기에서 글리코실화된 Fc 부분, 디아바디, scFv, scFv-Fc, scFv-CH, Fab 및 scFab 중 적어도 하나를 포함하는 항체로서 상기 글리코실화 단백질을 발현하는 숙주 세포.

청구항 37

제34항에 있어서, 글리코실화 단백질이 시토카인, 예컨대 인터페론, G-CSF, 응고 인자, 예컨대 인자 VIII, 인자 IX, 및 인간 단백질 C, 가용성 IgE 수용체 α -쇄, IgG, IgG 단편, IgM, 인터류킨, 우로키나제, 키마제, 및 우레아 트립신 억제제, IGF-결합 단백질, 표피 성장 인자, 성장 호르몬-방출 인자, 아넥신 V 융합 단백질, 안지오탐타틴, 혈관 내피 성장 인자-2, 골수성 전구 억제 인자-1, 오스테오프로테게린, α -1 항트립신, DNase II, α -페토 단백질, AAT, rhTBP-1 (일명 TNF 결합 단백질 1), TACI-Ig (막형단 활성화제 및 칼슘 조절제 및 시클로필린 리간드 인터랙터), FSH (여포 자극 호르몬), GM-CSF, 글루카곤, 글루카곤 펩티드, GLP-1 w/ 및 w/o FC (글루카곤 유사 단백질 1) IL-I 수용체 효능제, sTNFr (일명 가용성 TNF 수용체 Fc 융합), CTLA4-Ig (세포독성 T 림프구 조합 항원 4-Ig), 수용체, 호르몬, 예컨대 인간 성장 호르몬, 에리트로포이에틴, 펩티드, 스테이플드 펩티드, 인간 백신, 동물 백신, 혈청 알부민 및 효소, 예컨대 ATIII, 제조한 인간 트롬빈, 글루코세레브로시다제 및 아스파라기나제를 포함하는 것인 숙주 세포.

청구항 38

제24항에 있어서, MBP에 작동가능하게 융합된 하나 이상의 만노실트랜스퍼라제, N-아세틸글루코사민일 트랜스퍼라제, 갈락토실트랜스퍼라제 또는 시알릴트랜스퍼라제를 발현하는 숙주 세포.

청구항 39

제24항에 있어서, 산화 숙주 세포가 갈락토실트랜스퍼라제 효소 활성을 발현하기 위하여 사용되는 것인 숙주 세포.

청구항 40

제24항에 있어서, $GlcNAc_{1-5}Man_3GlcNAc_2$ 및 $Man_3GlcNAc_2$ 를 포함하는 올리고사카라이드 조성물을 생성하는 숙주 세포.

청구항 41

제40항에 있어서, 올리고사카라이드 조성물이 대부분 $\text{GlcNAcMan}_3\text{GlcNAc}_2$ 또는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 인 숙주 세포.

청구항 42

제25항에 있어서, $\text{Gal}_{1-5}\text{GlcNAc}_{1-5}\text{Man}_3\text{GlcNAc}_2$ 및 $\text{Man}_3\text{GlcNAc}_2$ 를 포함하는 올리고사카라이드 조성물을 생성하는 숙주 세포.

청구항 43

제42항에 있어서, 올리고사카라이드 조성물이 대부분 $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$, $\text{GalGlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 또는 $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 인 숙주 세포.

청구항 44

제26항에 있어서, 올리고사카라이드 조성물이 $\text{NANA}_{1-5}\text{Gal}_{1-5}\text{GlcNAc}_{1-5}\text{Man}_3\text{GlcNAc}_2$ 를 포함하는 것인 숙주 세포.

청구항 45

제44항에 있어서, 올리고사카라이드 조성물이 대부분 $\text{NANAGalGlcNAcMan}_3\text{GlcNAc}_2$ 또는 $\text{NANA}_2\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 인 숙주 세포.

청구항 46

GDP-만노스 잔기의 말단 만노스 잔기로의 전이를 촉매하는 하나 이상의 알파-1,2-만노실트랜스퍼라제 효소 활성 (EC 2.4.1.-)을 발현하기 위해 말단 만노스 잔기를 갖는 올리고사카라이드 조성물을 생성하는 재조합 원핵 숙주 세포를 배양하는 단계를 포함하며, 상기 배양 단계는 말단 만노스 잔기를 갖는 올리고사카라이드 조성물을 생성하기에 효과적인 조건 하에 수행되는 것인 고-만노스 유형 올리고사카라이드 조성물의 제조 방법.

청구항 47

제46항에 있어서, 배양 단계가 하나 이상의 진행 UDP-GlcNAc 트랜스퍼라제 효소 활성 (EC 2.4.1.141 또는 EC 2.4.1.145) 및 하나 이상의 진행 만노실트랜스퍼라제 효소 활성 (EC 2.4.1.142, EC 2.4.1.132)을 발현하기 위한 숙주 세포를 포함하며, 여기서 효소는 숙주 세포에 진행 글리코실트랜스퍼라제 활성을 부여하는 것인 방법.

청구항 48

제46항에 있어서, 배양 단계가 하나 이상의 효소의 융합을 발현하기 위한 숙주 세포를 추가로 포함하며, 여기서 상기 융합은 다음의 용해도 인핸서: DsbA, GlpE, GST, MBP, MstX, NusA 및 TrxA로부터 선택된 용해도 인핸서 중 적어도 하나를 포함하는 것인 방법.

청구항 49

제46항에 있어서, 배양 단계가 포스포마뉴타제 효소 활성 (ManB) (EC 5.4.2.8), 만노스-1-포스페이트 구아닐릴트랜스퍼라제 효소 활성 (ManC) (EC 2.7.7.13) 및 글루타민-프록토스-6-포스페이트 트랜스아미나제 효소 활성 (GlmS) (EC 2.6.1.16) 중 하나 이상을 발현하기 위한 숙주 세포를 포함하며, 여기서 ManB 및 ManC는 GDP-만노스 합성을 촉매하고, GlmS는 UDP-GlcNAc 합성을 촉매하는 것인 방법.

청구항 50

제46항에 있어서, 배양 단계가 GDP-D-만노스 데히드라타제 효소 활성 (EC 4.2.1.47)에 있어서 감쇠를 갖는 숙주 세포를 포함하는 것인 방법.

청구항 51

제46항에 있어서, 배양 단계가 플립라제 효소 활성을 발현하기 위한 숙주 세포를 포함하는 것인 방법.

청구항 52

제46항에 있어서, 배양 단계가 올리고사카릴 트랜스퍼라제 효소 활성 (EC 2.4.1.119)을 발현하기 위한 숙주 세포를 포함하는 것인 방법.

청구항 53

제46항에 있어서, 관심 단백질을 코딩하는 유전자를 숙주 세포에 도입하여 숙주 세포가 글리코실화 단백질을 생성하도록 하는 것을 추가로 포함하는 방법.

청구항 54

제46항에 있어서, 올리고사카라이드 조성물이 단백질에 N-연결된 것인 방법.

청구항 55

제46항에 있어서, 숙주 세포가 MBP에 작동가능하게 융합된 하나 이상의 만노실트랜스퍼라제를 발현하는 것인 방법.

청구항 56

제46항에 있어서, 배양 단계가 백신으로서 상기 글리코실화 단백질을 발현하기 위한 재조합 원핵 숙주 세포를 포함하는 것인 방법.

청구항 57

제46항에 있어서, 올리고사카라이드 조성물이 $\text{Man}_2\text{GlcNAc}_2$, $\text{Man}_4\text{GlcNAc}$, $\text{Man}_3\text{GlcNAc}_2$, $\text{HexMan}_3\text{GlcNAc}_2$, $\text{HexMan}_5\text{GlcNAc}$, $\text{Man}_6\text{GlcNAc}$ 및 $\text{Man}_5\text{GlcNAc}_2$ 로 이루어진 군으로부터 선택되는 것인 방법.

청구항 58

제46항에 있어서, 올리고사카라이드 조성물이 대부분 $\text{Man}_5\text{GlcNAc}_2$ 인 방법.

청구항 59

GDP-만노스 잔기의 트리만노스 올리고사카라이드 조성물로의 전이를 촉매하는, a) 하나 이상의 진행 UDP-GlcNAc 트랜스퍼라제 효소 활성 (EC 2.4.1.141 또는 EC 2.4.1.145), b) 하나 이상의 진행 만노실트랜스퍼라제 효소 활성 (EC 2.4.1.142, EC 2.4.1.132), 및 c) 하나 이상의 알파-1,2-만노실트랜스퍼라제 효소 활성 (EC 2.4.1.-))을 포함하는 고-만노스 유형 올리고사카라이드 조성물을 생성하는 재조합 원핵 숙주 세포.

청구항 60

제59항에 있어서, 하나 이상의 진행 UDP-GlcNAc 트랜스퍼라제 효소 활성 (EC 2.4.1.141 또는 EC 2.4.1.145) 및 하나 이상의 진행 만노실트랜스퍼라제 효소 활성 (EC 2.4.1.142, EC 2.4.1.132)을 포함하며, 여기서 상기 효소는 숙주 세포에 진행 글리코실트랜스퍼라제 활성을 부여하는 것인 숙주 세포.

청구항 61

제59항에 있어서, 하나 이상의 효소의 융합을 추가로 포함하며, 여기서 상기 융합은 DsbA, G1pE, GST, MBP, MstX, NusA 및 TrxA 중 적어도 하나를 포함하는 것인 숙주 세포.

청구항 62

제59항에 있어서, 다음의 효소: 포스포마노무타제 효소 활성 (ManB) (EC 5.4.2.8), 만노스-1-포스페이트 구아닐릴트랜스퍼라제 효소 활성 (ManC) (EC 2.7.7.13) 및 글루타민-프룩토스-6-포스페이트 트랜스아미나제 효소 활성 (GlmS) (EC 2.6.1.16) 중 하나 이상을 포함하며, 여기서 ManB 및 ManC는 GDP-만노스 합성을 촉매하고, GlmS는 UDP-GlcNAc 합성을 촉매하는 것인 숙주 세포.

청구항 63

제59항에 있어서, GDP-D-만노스 데히드라타제 효소 활성 (EC 4.2.1.47)에 있어서 감쇠를 추가로 포함하는 숙주

세포.

청구항 64

제59항에 있어서, 플립라제 효소 활성을 추가로 포함하는 숙주 세포.

청구항 65

제59항에 있어서, 올리고사카릴 트랜스퍼라제 효소 활성 (EC 2.4.1.119)을 추가로 포함하는 숙주 세포.

청구항 66

제59항에 있어서, 관심 단백질을 코딩하는 유전자를 추가로 포함하여 글리코실화 단백질을 생성하는 숙주 세포.

청구항 67

제59항에 있어서, 단백질에 N-연결된 올리고사카라이드 조성물을 생성하는 숙주 세포.

청구항 68

제59항에 있어서, 백신으로서 상기 글리코실화 단백질을 발현하는 숙주 세포.

청구항 69

제59항에 있어서, 올리고사카라이드 조성물이 $Man_2GlcNAc_2$, $Man_4GlcNAc$, $Man_3GlcNAc_2$, $HexMan_3GlcNAc_2$, $HexMan_5GlcNAc$, $Man_6GlcNAc$ 및 $Man_5GlcNAc_2$ 로 이루어진 군으로부터 선택된 것인 방법.

청구항 70

제59항에 있어서, 올리고사카라이드 조성물이 대부분 $Man_5GlcNAc_2$ 인 방법.

발명의 설명

기술 분야

[0001] **연방 정부의 지원을 받은 연구 및 개발의 진술**

[0002] 본 발명은 미국 국립보건원에 의해 허가 번호 1R43GM088905-01, 2R44GM088905-02 및 5R44GM088905-03 하에 정부의 지원으로 완성되었다. 정부는 본 발명에 대해 특정 권리를 갖는다.

[0003] **관련 출원의 상호 참조**

[0004] 본 출원은 2013년 3월 14일에 출원된 미국 가출원 제61/785,586호에 관련되었으며, 상기 출원은 모든 목적을 위해 본원에 참조로 그의 전체가 포함된다.

[0005] **서열 목록**

[0006] 본 출원은 EFS-Web을 통해 제출된 서열 목록을 포함하며, 참조로 그의 전체가 포함된다. [DATE]에 생성된 상기 ASCII 사본은 [.txt]로 명명되었으며 크기는 [#####] 바이트이다.

[0007] **기술 분야**

[0008] 본 발명은 일반적으로 당생물학 및 단백질 조작 분야에 관한 것이다. 보다 구체적으로, 본원에 기술된 실시양태는 올리고사카라이드 조성물 및 원핵생물에서의 치료 단당백질 제조에 관한 것이다.

배경 기술

[0009] **당 치료제**

[0010] 단백질 기반의 치료제는 현재 FDA가 승인한 신약 4개 당 하나에 해당한다(문헌 [Walsh, G., "Biopharmaceutical Benchmarks," *Nat Biotechnol* 18:831-3 (2000)], [Walsh, G., "Biopharmaceutical Benchmarks," *Nat Biotechnol* 21:865-70 (2003)], 및 [Walsh, G., "Biopharmaceutical Benchmarks," *Nat*

Biotechnol 24:769-76 (2006)]).

[0011] 몇몇 단백질 치료제가 E. 콜라이와 같은 원핵 발현 시스템을 이용하여 제조될 수 있지만(예를 들어 인슐린), 압도적으로 대다수의 치료 단백질은 생물학적 기능을 완전하게 얻기 위해서는 원핵생물에는 존재하지 않는 것으로 생각되는 부가적인 번역 후 변형을 필요로 한다. 특히, N-연결된 단백질 글리코실화는 모든 진핵 단백질 중의 절반 초과에 영향을 미칠 것으로 예상되며(문헌 [Apweiler et al., "On the Frequency of Protein Glycosylation, as Deduced From Analysis of the SWISS-PROT Database," *Biochim Biophys Acta* 1473:4-8 (1999)]), 흔히 적절한 폴딩, 약동학 안정성, 조직 타겟팅 및 다수의 단백질의 효능에 필수적이다(문헌 [Helenius et al., "Intracellular Functions of N-linked Glycans," *Science* 291:2364-9 (2001)]). 대부분의 박테리아는 그들 자신의 단백질을 글리코실화하지 않기 때문에, 항체를 비롯한 대부분의 치료상 적절한 당단백질의 발현은 포유동물 세포로 격하된다. 그러나, 포유동물 세포 배양은 이하를 비롯하여 많은 결점이 있다: (i) 매우 높은 제조 비용 및 박테리아에 관하여 CHO 세포와 같은 진핵 숙주의 낮은 용적 생산성; (ii) 레트로바이러스 오염; (iii) 안정한 세포주를 생성하는데 비교적 긴 시간이 요구됨; (iv) 유전자 변형을 통해 안정하고 "높은 생산성"의 진핵 세포주를 신속하게 생성하는 것의 상대적 불능; 및 (v) 내인성 비-인간 글리코실화 경로를 갖는 CHO와 같은 숙주 세포를 사용할 때 발생하는 당형태 이질성에 의해 생성되는 높은 생성물 가변성(문헌 [Choi et al., "Use of Combinatorial Genetic Libraries to Humanize N-linked Glycosylation in the Yeast *Pichia pastoris*," *Proc Natl Acad Sci U S A* 100:5022-7 (2003)]). 다른 한편으로, E. 콜라이에서의 발현은 이러한 제한이 없다.

[0012] **E. 콜라이 중 치료 단백질의 발현**

[0013] 많은 치료 재조합 단백질이 현재 숙주 유기체로서 E. 콜라이를 사용하여 발현되고 있다. 가장 좋은 예 중 하나는 인간 인슐린이며, 이는 일라이 릴리에 의해 1982년 처음으로 제조되었다. 그 후로, E. 콜라이 발현에 의존하는 많은 인간 치료 단백질이 미국 및 유럽에서 승인되었으며, 인간 성장 호르몬(hGH), 과립구 대식세포 콜로니 자극 인자(GM-CSF), 인슐린-유사 성장 인자(IGF-1, IGFBP-3), 각질세포 성장 인자, 인터페론(IFN- α , IFN- β 1b, IFN- γ 1b), 인터류킨(IL-1, IL-2, IL-11), 조직 괴사 인자(TNF- α), 및 조직 플라스미노겐 활성화제(tPA)를 포함한다. 그러나, 거의 모든 당단백질이 포유동물 세포에서 생성된다. 정상적으로 글리코실화된 단백질이 E. 콜라이에서 발현되는 경우, 해당 숙주에서 글리코실화의 결여는 기능이 손상된 단백질을 생성한다. 예를 들어, 탈글리코실화 인간 모노클로날 항체(mAbs)(예를 들어 항-조직 인자 IgG1)는 E. 콜라이에서 가용한 형태로 높은 수준으로 발현될 수 있다(문헌 [Simmons et al., "Expression of Full-length Immunoglobulins in *Escherichia coli*: Rapid and Efficient Production of Aglycosylated Antibodies," *J Immunol Methods* 263:133-47 (2002)]). 그러나, E. 콜라이-유도된 mAbs가 이들의 동족 항원 및 신생 수용자에 견고한 결합을 유지하고 포유동물 세포-유도된 항체에 필적하는 순환 반감기를 나타내는 반면, 이들은 N-글리칸이 없기 때문에 C1q 및 Fc γ RI 수용자에 결합할 수 없다.

[0014] **진핵 및 원핵 N-연결된 단백질 글리코실화**

[0015] N-연결된 단백질 글리코실화는 진핵 유기체의 세포질 세망(ER)에서 발생하는 필수적이고 보존적인 과정이다 (문헌 [Burda et al., "The Dolichol Pathway of N-linked Glycosylation," *Biochim Biophys Acta* 1426:239-57 (1999)]). 이는 분비 및 막 단백질의 단백질 폴딩, 올리고머화, 품질 관리, 소팅 및 이송에 중요하다(문헌 [Helenius et al., "Intracellular Functions of N-linked Glycans," *Science* 291:2364-9 (2001)]). 진핵 N-연결된 단백질 글리코실화 경로는 이하의 2가지 상이한 공정으로 나뉠 수 있다: (i) 지질-연결된 올리고사카라이드를 세포질 세망의 막에서 조립, 및 (ii) 지질 앵커 돌리콜 피로포스페이트로부터 신생 폴리펩티드의 선택된 아스파라긴 잔기에 올리고사카라이드를 이송. N-연결된 단백질 글리코실화의 특징, 즉 (i) 올리고사카라이드 어셈블리를 위한 담체로서 돌리콜 피로포스페이트 (DoI-PP)의 사용, (ii) 완전히 조립된 Glc₃Man₉GlcNAc₂ 올리고사카라이드만의 이송, 및 (iii) 서열 N-X-S/T(N은 아스파라긴, X는 프롤린을 제외한 임의의 아미노산, 그리고 S/T는 세린/트레오닌임)를 특징으로 하는 아스파라긴 잔기의 인식(문헌 [Gavel et al., "Sequence Differences Between Glycosylated and Non-glycosylated Asn-X-Thr/Ser Acceptor Sites: Implications for Protein Engineering," *Protein Eng* 3:433-42 (1990)])은 진핵생물에 고도로 보존되어 있다. 올리고사카릴트랜스퍼라제(OST)는 지질 공여자 돌리킬피로포스페이트로부터 수용자 단백질로의 올리고사카라이드의 이송을 촉매한다. 효모에서, 8개의 상이한 막 단백질이 생체 내 촉매를 구성하는 것으로 확인되었다(문헌 [Kelleher et al., "An Evolving View of the Eukaryotic Oligosaccharyltransferase," *Glycobiology* 16:47R-62R (2006)]). STT3은 OST의 촉매적 서브유닛을 나타내는 것으로 생각된다(문헌 [Nilsson et al., "Photocross-linking of Nascent

Chains to the STT3 Subunit of the Oligosaccharyltransferase Complex," *J Cell Biol* 161:715-25 (2003)] 및 [Yan et al., "Studies on the Function of Oligosaccharyl Transferase Subunits. Stt3p is Directly Involved in the Glycosylation Process," *J Biol Chem* 277:47692-700 (2002)]. 이는 OST 작용에서 가장 보존된 서브유닛이다(문헌 [Burda et al., "The Dolichol Pathway of N-linked Glycosylation," *Biochim Biophys Acta* 1426:239-57 (1999)]).

[0016]

반대로, 박테리아 내의 글리코실화 경로의 결여는, 특히 "자연의 모든 단백질의 절반 초과가 궁극적으로는 당단백질로 밝혀질 것이다"(문헌 [Apweiler et al., "On the Frequency of Protein Glycosylation, as Deduced From Analysis of the SWISS-PROT Database," *Biochim Biophys Acta* 1473:4-8 (1999)])라는 특정 추정에 의해, 치료 단백질의 제조를 위한 원핵 발현 숙주의 유용성을 크게 제한하였다. 그러나, 최근 병원성 박테리아, 캄필로박터 제주니의 게놈이 N-연결된 단백질 글리코실화의 경로를 인코딩하는 것이 발견되었다(문헌 [Szymanski et al., "Protein Glycosylation in Bacterial Mucosal Pathogens," *Nat Rev Microbiol* 3:225-37 (2005)]). 이 경로를 위한 유전자는 지만스키(Szymanski)와 동료들에 의해 1999년에 처음 확인되었으며(문헌 [Szymanski et al., "Evidence for a System of General Protein Glycosylation in *Campylobacter jejuni*," *Mol Microbiol* 32:1022-30 (1999)]), 단백질 글리코실화를 위한 *pgl*로 명명된 17-kb 좌위를 포함한다. *pgl* 좌위의 발견에 이어, 2002년에 린톤(Linton) 등은 PEB3 및 CgpA의 두 캄필로박터 제주니 당단백질을 확인하였으며, 이들과 같은 캄필로박터 제주니-유도된 당단백질이 N-아세틸 갈락토사민(GalNAc)-특이적 렉틴 대두 응집소(SBA)에 결합하는 것을 밝혀냈다(문헌 [Linton et al., "Identification of N-acetylgalactosamine-containing Glycoproteins PEB3 and CgpA in *Campylobacter jejuni*," *Mol Microbiol* 43:497-508 (2002)]). 그로부터 얼마 후에 영(Young) 등은 PEB3 및 CgbA를 포함하는 30 초과의 잠재적 캄필로박터 제주니 당단백질을 확인했으며, 질량 분광측정법 및 NMR을 이용하여 N-연결된 글리칸이 GalNAc- α 1,4-GalNAc- α 1,4-[Glc β 1,3]GalNAc- α 1,4-GalNAc- α 1,4-GalNAc- α 1,3-Bac- β 1,N-Asn(GalNAc₅GlcBac(여기서 Bac은 바실로사민 또는 2,4-디아세트아미도-2,4,6-트리데옥시글루코스))(문헌 [Young et al., "Structure of the N-linked Glycan Present on Multiple Glycoproteins in the Gram-negative Bacterium, *Campylobacter jejuni*," *J Biol Chem* 277:42530-9 (2002)])의 구조를 갖는 헵타사카라이드인 것을 밝혀냈다. 분지형 헵타사카라이드는 지질 담체 운데카프레닐피로포스페이트(Und-PP)에 내막의 세포질 측에 뉴클레오티드-활성화된 당을 순차적으로 첨가하여 합성되고(문헌 [Feldman et al., "Engineering N-linked Protein Glycosylation with Diverse O Antigen Lipopolysaccharide Structures in *Escherichia coli*," *Proc Natl Acad Sci U S A* 102:3016-21 (2005)]), 조립된 후에는 추정 ATP-결합 카세트(ABC) 수송체 WlaB에 의해 막을 가로질러 플립핑된다(문헌 [Alaimo et al., "Two Distinct But Interchangeable Mechanisms for Flipping of Lipid-linked Oligosaccharides," *Embo J* 25:967-76 (2006)] 및 [Kelly et al., "Biosynthesis of the N-linked Glycan in *Campylobacter jejuni* and Addition Onto Protein Through Block Transfer," *J Bacteriol* 188:2427-34 (2006)]). 이어서, 주변세포질 내 기질 단백질로의 헵타사카라이드의 이송은 진핵 OST STT3의 촉매적 서브유닛에 상당한 서열 유사성을 갖는 단일의 내재성 막 단백질인 PglB로 명명된 OST에 의해 촉매된다(문헌 [Young et al., "Structure of the N-linked Glycan Present on Multiple Glycoproteins in the Gram-negative Bacterium, *Campylobacter jejuni*," *J Biol Chem* 277:42530-9 (2002)]). PglB는 모티프 D/E-X₁-N-X₂-S/T(여기서 D/E는 아스파르트산/글루탐산, X₁ 및 X₂는 프롤린을 제외한 임의의 아미노산, N은 아스파라긴, S/T는 세린/트레오닌임) 내에서 헵타사카라이드를 아스파라긴에 부착시키며, 진핵 글리코실화 과정에서 사용된 것과 유사한 시퀀스이다(N-X-S/T)(문헌 [Kowarik et al., "Definition of the Bacterial N-glycosylation Site Consensus Sequence," *Embo J* 25:1957-66 (2006)]).

[0017]

미생물의 당조작

[0018]

포유동물, 효모, 또는 심지어 박테리아 숙주 세포에서 치료 당단백질의 발현시 접하게 되는 주요 문제는 비-인간 글리칸의 첨가이다. 예를 들어, 치료 당단백질의 제조에 가장 흔하게 사용되는 두 시스템 중 하나인 효모는, 고도 면역원성 만난-유형 N-글리칸(만노스 잔기를 100개까지 함유)을 재조합 당단백질로 이송한다. 포유동물 발현 시스템은 또한 비-인간 당 잔기, 예컨대 N-글리코실뉴라민산(Neu5Gc) 형태의 시알산(CHO 세포 또는 우유에서 제조) 또는 말단 α (1,3)-갈락토스(Gal)(뮤린 세포에서 생성)로 치료 단백질을 변형시킬 수 있다. 비-인간 당을 지닌 치료 단백질을 반복적으로 투여하면 인간에서의 면역 반응을 포함하여 유해 반응을 유발할 수 있다.

[0019]

치료 당단백질의 제조에 천연 글리코실화 시스템을 사용하는 것에 대한 대안으로서, 당-조작 발현 시스템의 가

용성은 치료 단백질의 글리코실화의 커스터마이징의 기회를 제공하고, 개선된 치료 당단백질의 개발로 유도할 수 있다. 그러한 시스템은 바람직하지 않은 글리칸을 제거하고 고도의 균질성으로 인간 글리코실화를 수행할 수 있는 잠재력을 가질 것이다. 효모인 피키아 파스토리스는 당-조작되어 특정 치료 기능을 위한 글리코실화를 위한 능력을 발현 시스템에 제공한다(문헌 [Gerngross, T. U., "Advances in the Production of Human Therapeutic Proteins in Yeasts and Filamentous fungi," *Nat Biotechnol* 22:1409-14 (2004)], [Hamilton et al., "Glycosylation Engineering in Yeast: The Advent of Fully Humanized Yeast," *Curr Opin Biotechnol* 18:387-92 (2007)] 및 [Wildt et al., "The Humanization of N-glycosylation Pathways in Yeast," *Nat Rev Microbiol* 3:119-28 (2005)]).

[0020]

예를 들어, 다양한 모노클로날 항체 리톡산(항-CD20 IgG1 항체)의 당형태를 제조하는데 당-조작된 피키아 파스토리스 균주의 패널이 사용되었다(문헌 [Li et al., "Optimization of Humanized IgGs in Glycoengineered *Pichia pastoris*," *Nat Biotechnol* 24:210-5 (2006)]). 이러한 항체는 시판의 리톡산과 동일한 아미노산 서열을 갖지만, 특정 글라이코폼은 관련 Fc γ RIII 수용자에 대해 대략 100배 높은 결합 친화도를 나타내었으며, 개선된 시험관내 인간-B-세포 고갈을 나타내었다(문헌 [Li et al., "Optimization of Humanized IgGs in Glycoengineered *Pichia pastoris*," *Nat Biotechnol* 24:210-5 (2006)]). 당-조작된 피키아 파스토리스의 큰 성공 및 잠재성에 약간의 결점이 없는 것은 아니다. 예를 들어, 효모 및 다른 모든 진핵생물에서 N-연결된 글리코실화는 실행 가능성을 위해 필수적이다(문헌 [Herscovics et al., "Glycoprotein Biosynthesis in Yeast," *FASEB J* 7:540-50 (1993)] 및 [Zufferey et al., "STT3, a Highly Conserved Protein Required for Yeast Oligosaccharyl Transferase Activity In vivo," *EMBO J* 14:4949-60 (1995)]). 게른그로스(Gerngross) 및 동료들은 원하지 않은 효모 N-글리코실화 반응을 체계적으로 제거하고 재조작하였다(문헌 [Choi et al., "Use of Combinatorial Genetic Libraries to Humanize N-linked Glycosylation in the Yeast *Pichia pastoris*," *Proc Natl Acad Sci U S A* 100:5022-7 (2003)]). 그러나, 만난-유형 N-글리칸의 제거는 효모의 글리코실화 스토리에 있어서 단지 절반이다. 이것은 효모가 또한 O-연결된 글리코실화를 수행하여 O-글리칸이 당단백질 내 Ser 또는 Thr 잔기에 연결되기 때문이다(문헌 [Gentzsch et al., "The PMT Gene Family: Protein O-glycosylation in *Saccharomyces cerevisiae* is Vital," *EMBO J* 15:5752-9 (1996)]). N-연결된 글리코실화에서와 같이, O-글리코실화는 실행가능성을 위해 필수적이고(문헌 [Gentzsch et al., "The PMT Gene Family: Protein O-glycosylation in *Saccharomyces cerevisiae* is Vital," *EMBO J* 15:5752-9 (1996)]), 따라서 당-조작된 효모에서 유전자적으로 제거될 수 없다. 효모와 인간의 O-글리코실화 기구에는 차이가 있으므로, 당-조작된 효모 균주에 의한 O-글리칸의 가능한 첨가는 면역 반응을 비롯한 유해 반응을 불러올 잠재성이 있다.

[0021]

에비(Aebi) 및 그의 동료들은 캄필로박터 제주니 글리코실화 좌위를 E. 콜라이로 이송하고, 이들 세포에 N-글리칸으로 단백질을 번역 후 변형할 수 있는 탁월한 능력을 부여하였다(문헌 [Wacker et al., "N-linked Glycosylation in *Campylobacter jejuni* and its Functional Transfer into *E. coli*," *Science* 298:1790-3 (2002)]). 그러나, 원핵 및 진핵 글리코실화 메커니즘이 공유하는 기능적 유사성에도 불구하고, 진핵 글리코실화 기구에 의해 부착된 올리고사카라이드 쇠(GalNAc₅GlcBac)는 진핵 글리코실화 경로에 의해 부착된 것과는 구조적으로 구분된다(문헌 [Szymanski et al., "Protein Glycosylation in Bacterial Mucosal Pathogens," *Nat Rev Microbiol* 3:225-37 (2005)], [Young et al., "Structure of the N-linked Glycan Present on Multiple Glycoproteins in the Gram-negative Bacterium, *Campylobacter jejuni*," *J Biol Chem* 277:42530-9 (2002)] 및 [Weerapana et al., "Asparagine-linked Protein Glycosylation: From Eukaryotic to Prokaryotic Systems," *Glycobiology* 16:91R-101R (2006)]). E. 콜라이를 진핵 N-글리코실화 경로로 리프로그래밍하여 구조적으로 균질한 인간-유사 글리칸을 갖는 N-연결된 당단백질을 발현하기 위한 많은 시도가 있었다(성공하지는 못했음).

[0022]

보다 최근, 바데라마-린콘 등(Vaderrama-Rincon et al.; 문헌 ["An engineered eukaryotic protein glycosylation pathway in *E. coli*," *Nat Chem Bio* 8, 434-436 (2012)])은 원핵 숙주 세포를 진핵 글리코실트랜스퍼라제로 당조작할 수 있음을 밝혀냈다. 구체적으로, 원핵 숙주 세포에서의 UDP-GlcNAc 트랜스퍼라제 및 GDP-만노스 트랜스퍼라제의 발현은 거의 모든 진핵 N-연결된 올리고사카라이드 구조의 기본인 트리만노실 코어 구조의 제조를 나타내었다. 그러나 완전히 정교한 인간-유사 글리칸은 여전히 추가의 당조작을 필요로 한다.

[0023]

따라서 본 발명은 예컨대 고-만노스, 하이브리드 및 복합 유형의 인간-유사 글리칸의 제조에 관한 것이다.

발명의 내용

[0024]

발명의 요약

[0025]

본 발명은 올리고사카라이드 조성물의 제조 및 원핵 숙주 세포에서의 제조합 당단백질의 제조를 위한 방법 및 물질을 제공하는 것이다. 특정 N-글리칸을 포함하는 다양한 당단백질 조성물이 본 발명의 방법을 이용하여 제조된다. 특정 실시양태에서, 목적하는 당형태가 우세 중으로서 생성된다.

[0026]

본 발명은 또한 예컨대 대상체 내의 면역 또는 면역학적 내성(예를 들어 무반응)을 유도하기 위한, 특정 올리고사카라이드 조성물을 포함하는 백신 항원의 제조를 위한 방법 및 물질을 제공한다. 본 발명의 다양한 측면은 수지상 세포 및/또는 다른 항원-제공 세포의 표면 상에서 발현 가능한 렉틴을 결합할 수 있는 항원-탄수화물 접합체에 관한 것이다.

[0027]

본 발명의 제1 측면은 올리고사카라이드 조성물의 제조 방법에 관한 것으로, 이 방법은, 말단 만노스 잔기에 UDP-GlcNAc 잔기의 이송을 촉매하는 하나 이상의 N-아세틸글루코사미드 트랜스퍼라제 효소 활성화(EC 2.4.1.101; EC 2.4.1.143; EC 2.4.1.145; EC 2.4.1.155; EC 2.4.1.201)을 나타내기 위해 상기 말단 만노스 잔기를 갖는 올리고사카라이드 조성물을 생산하는 재조합 원핵 숙주 세포를 배양하는 것을 포함하며, 상기 배양 단계는 말단 GlcNAc 잔기를 갖는 올리고사카라이드 조성물을 제조하는데 유효한 조건 하에 수행된다.

[0028]

본 발명의 제2 측면은 올리고사카라이드 조성물의 제조 방법에 관한 것으로, 이 방법은, 상기 말단 GlcNAc 잔기에 UDP-갈락토스 잔기의 이송을 촉매하는 하나 이상의 갈락토실트랜스퍼라제 효소 활성화(EC 2.4.1.38)을 나타내기 위해 숙주 세포를 배양하는 것을 포함하며, 상기 배양 단계는 말단 갈락토스 잔기를 갖는 올리고사카라이드 조성물을 제조하는데 유효한 조건 하에 수행된다.

[0029]

본 발명의 제3 측면은 올리고사카라이드 조성물의 제조 방법에 관한 것으로, 이 방법은, 상기 말단 갈락토스 잔기에 CMP-NANA 잔기의 이송을 촉매하는 하나 이상의 시알릴트랜스퍼라제 효소 활성화(EC 2.4.99.4 및 EC 2.4.99.1)을 나타내기 위해 숙주 세포를 배양하는 것을 포함하며, 상기 배양 단계는 말단 시알산 잔기를 갖는 올리고사카라이드 조성물을 제조하는데 유효한 조건 하에 수행된다.

[0030]

본 발명의 다른 측면들은 용해도가 향상된 용합 단백질로서의 1종 이상의 효소의 발현에 관한 것이다. 본 발명의 또 다른 측면들은 관심 단백질을 인코딩하는 유전자에 글리칸을 이송하고, 이에 따라 숙주 세포가 글리코실화 단백질을 제조하는 것을 포함한다.

[0031]

추가 측면들은 주요 당형태의 제조를 위한 추가적인 효소의 배양 조건 및 과다발현을 포함한다. 본 발명에 개시된 측면들은 원핵 숙주 세포를 제공하여 다양한 글리코실트랜스퍼라제 활성을 발현하여 고-만노스, 하이브리드 및/또는 복합 올리고사카라이드 조성물뿐만 아니라 고-만노스, 하이브리드 및/또는 복합 글리코실화 단백질을 제조한다.

[0032]

바람직한 측면에서, 본 발명은 당단백질 치료제 및 진단의 고안, 발견 및 개발을 위한 기술을 상업화한다. 구체적으로, 본 발명은 미생물 세포에서 실제 인간 당단백질의 효율적인 제조를 위한 효율적이고 저비용의 전략의 개발을 제공한다. 다양한 측면에서, 본 발명의 당-조작 박테리아는 N-연결된 당단백질의 입체특이적 제조를 가능하게 한다. 일 실시양태에서, 박테리아는, 특정 아스파라긴 수용자 부위에서 표적 단백질을 효율적으로 글리코실화할 수 있는(예를 들어 N-연결된 글리코실화) 신규 글리코실화 경로를 인코딩하는 유전자로 형질전환된다. 이러한 특수하게 조작된 세포주를 이용하여, 목적하는 다양한 제조합 단백질을 발현시키고 글리코실화시킬 수 있다.

[0033]

또한, 본 발명은 원핵생물에서 올리고사카라이드 구조의 순열을 조작하기 위한 방법을 제공하며, 이는 예를 들어 단백질의 약동학 특성을 변경하고 생물학적 현상에서 글리코실화의 역할을 규명할 것으로 기대된다. 다양한 측면들에서, 본 발명은 연구, 공업 및 치료 용도의 치료 단백질, 신규 당접합체, 면역자극제(예를 들어 백신)의 생물공학적 합성을 제공한다.

도면의 간단한 설명

[0034]

도 1. 고-만노스 유형 Man₃GlcNAc₂ 당형태의 제조. (A) GLY02로부터 추출한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼(예상의 Man₃GlcNAc₂와 일치(m/z 1257.6)) 및 (B) α1,2-만노시다제로 추가 처리한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼(예상의 Man₃GlcNAc₂ 당형태와 일치(m/z 933.4))

도 2. 하이브리드 GlcNAcMan₃GlcNAc₂ 당형태의 제조. (A) GLY03으로부터 추출한 지질-방출된 글리칸의 MALDI-

TOF 질량 스펙트럼(예상의 $\text{GlcNAcMan}_3\text{GlcNAc}_2$ 당형태와 일치(m/z 1136.5)) 및 (B) β -N-아세틸글루코사미니다제로 추가 처리한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼(예상의 $\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치(m/z 933.5))

도 3. 복합 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태의 제조. GLY06.1로부터 추출한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼(예상의 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치(m/z 1339.8)).

도 4. 하이브리드, 분지된 당형태의 제조. (A) GLY05로부터 추출한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼(예상의 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치(m/z 1339.7)) 및 (B) β -N-아세틸글루코사미니다제로 추가 처리한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼(예상의 $\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치(m/z 933.5)).

도 5. 다중-안테나 당형태의 제조. (A) 생체의 합성한 글리칸의 MALDI-TOF 질량 스펙트럼 및 (B) GLY06.1로부터 추출한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼(예상의 $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ 와 일치(m/z 1543.1)).

도 6. $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$ 당형태의 제조. GLY04.1로부터 추출한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼(예상의 $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$ 당형태와 일치(m/z 1298.7)).

도 7. $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태의 제조. (A) 생체의 합성한 글리칸의 MALDI-TOF 질량 스펙트럼 및 (B) GLY04.2로부터 추출한 지질-방출된 글리칸의 질량 스펙트럼(예상의 $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 와 일치(m/z 1662.2)).

도 8. $\text{NANAGalGlcNAcMan}_3\text{GlcNAc}_2$ 당형태의 제조. 생체의 합성한 글리칸의 양성 이온 모드에서의 MALDI-TOF 질량 스펙트럼(예상의 $\text{NANAGalGlcNAcMan}_3\text{GlcNAc}_2$ 와 일치(m/z 1565.7)).

도 9. 향상된 글리칸 수율. (A) $\text{Man}_3\text{GlcNAc}_2$ 글리칸 표준(M3GN2 Std)을 이용하여, ManC/B의 과다발현과 함께 (GLY01.2) 또는 과다발현 없이(GLY01) 수행한 E. 콜라이로부터 추출한 지질-방출된 글리칸의 형광단-보조 탄수화물 전기영동(FACE)($\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치)(좌측) 및 GlmS의 과다발현과 함께(GLY01.1) 또는 과다발현 없이(GLY01.3) 수행한 E. 콜라이로부터 추출한 지질-방출된 글리칸의 형광단-보조 탄수화물 전기영동(FACE)($\text{GlcNAcMan}_3\text{GlcNAc}_2$ 당형태(GNM3GN2)와 일치함(우측)). (B) 나타낸 ManC/B의 과다발현 및 글리세롤 추가와 함께 GLY01.2로부터 추출한 지질-방출된 글리칸의 정량. (C) 0.2% 글리세롤 또는 피루베이트 추가와 함께 GLY01.2로부터 추출한 지질-방출된 글리칸의 FACE.

도 10. 향상된 생물형성. 균주 (A) GLY01, (B) GLY02, 및 (C) ManC/B의 과다발현 없이 GLY01.1 및 (D) GLY01.4, (E) GLY02.1 및 ManC/B의 과다발현과 함께 (F) GLY01.5로부터 추출한 지질-방출된 글리칸의 MALDI-TOF 질량 스펙트럼. 중간체 당형태에 해당하는 피크의 손실이 ManC/B의 첨가와 함께 관찰되었다.

도 11. 글리코실화 글루카곤 제조. 다양한 당조작된 균주로부터 C-말단 글리코실화 부위가 달린(M3, M5, $\text{GlcNAcMan}_3\text{GlcNAc}_2$ 및 $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$ 글리코펩티드를 제조함) 부분적으로 정제된 글루카곤의 MALDI-TOF MS.

도 12. 글리코실화 항원. 4개의 연속된 C-말단 글리코실화 부위가 달린 장의 병원성 E. 콜라이(ExPEC)로부터 유래되고 GLY01에서 발현된 부분적으로 정제된 (A) 3473 및 (B) 1275 단백질의 웨스턴 블롯(항-헥사히스티딘 항체로 검출(좌측) 및 말단 알파-D-만노스에 대해 특이적인 콘카나발린 A 렉틴으로 검출(우측)).

발명을 실시하기 위한 구체적인 내용

정의

이하의 용어의 정의 및 방법은 본 명세서를 보다 잘 설명하고 본 명세서의 실행시 통상의 기술자를 안내하기 위해 제공된다.

본원에 언급된 모든 공보, 특허 및 다른 참고문헌은 그의 전체가 참조로 인용된다.

EC 번호는 생화학 및 분자 생물학의 국제 연합의 명명 위원회(NC-IUBMB)에 의해 확립되었다(<http://www.chem.qmul.ac.uk/iubmb/enzyme/>에서 이용 가능). 본원에 기술된 EC 번호는 부분적으로 교토 대학교의 후원을 받은 교토 유전자 및 게놈 백과사전의 지원을 받는 KEGG 리간드 데이터베이스를 출처로 한다. 달

리 언급되지 않는 한, EC 번호는 2013년 3월의 데이터베이스에 제공된 것이다.

- [0039] 본원에 기술된 수탁 번호는 미국 국립 보건원의 지원을 받는 NCBI 데이터베이스(미국 국립 생물 정보 센터)를 출처로 한다. 달리 언급되지 않는 한, 수탁 번호는 2013년 3월의 데이터베이스에 제공된 것이다.
- [0040] 본 발명의 방법 및 기술은 달리 언급되지 않는 한 일반적으로 당업계에 공지된 통상의 방법에 따라, 그리고 본 명세서에 걸쳐 인용되고 논의된 다양한 일반적 및 구체적 참고문헌에 기술된 바에 따라 수행되었다(예를 들어 문헌 [Sambrook et al., Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)], [Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2002)], [Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1990)], [Taylor and Drickamer, Introduction to Glycobiology, Oxford Univ. Press (2003)], [Worthington Enzyme Manual, Worthington Biochemical Corp., Freehold, N.J.], [Handbook of Biochemistry: Section A Proteins, Vol I, CRC Press (1976)], [Handbook of Biochemistry: Section A Proteins, Vol II, CRC Press (1976)], [Essentials of Glycobiology, Cold Spring Harbor Laboratory Press (1999) 참조).
- [0041] 달리 설명하지 않는 한, 본원에 사용된 모든 기술적 및 과학적 용어는 본 발명이 속하는 업계의 통상의 기술자에게 통상적으로 이해되는 것과 동일한 의미를 갖는다. 본원에 기재된 것과 유사하거나 동등한 방법 및 물질이 본 발명의 실행 및 시험에 사용될 수 있으나, 이하에 적절한 방법 및 물질을 기재한다. 물질, 방법 및 실시예는 단지 예시를 위한 것이며, 한정하려는 의도는 없다. 본 발명의 다른 특징은 이하의 상세한 설명 및 특허청구범위로부터 명백하다.
- [0042] 가출원에 있어서 "특허청구범위(claim)"라는 용어는 실시양태 또는 바람직한 실시양태와 동의이다.
- [0043] 본원에 사용된 "~로 구성(comprising)"은 "~을 포함(including)"을 의미하며, 단수 형태 "a" 또는 "an" 또는 "the"는 문맥에서 달리 분명히 명시하지 않는 한 복수의 참조를 포함한다. 예를 들어 "세포로 구성"의 언급은 하나 또는 복수의 해당 세포를 포함한다. 용어 "또는"은 문맥에서 달리 분명히 명시하지 않는 한 언급된 대체 요소 또는 2 이상의 요소의 조합의 단일 요소를 지칭한다.
- [0044] 당단백질에 관한 "인간-유사"의 용어는 인간에서 생산된 것과 유사하거나 심지어는 동일한, 단백질 내 아스파라긴 잔기의 아미드 질소에 연결된(N-연결된) N-아세틸글루코사민(GlcNAc) 잔기가 부착된 단백질을 지칭한다.
- [0045] "N-글리칸" 또는 "N-연결된 글리칸"은 N-연결된 올리고사카라이드 구조를 지칭한다. N-글리칸은 단백질 또는 합성 당단백질 중간체에 부착될 수 있으며, 이는 시험관내 또는 생체내에서 추가로 조작될 수 있다. 당단백질에서 발견되는 주요 당은 글루코스(Glu), 갈락토스(Gal), 만노스(Man), 푸코스(Fuc), N-아세틸갈락토사민(GalNAc), N-아세틸글루코사민(GlcNAc), 및 시알산(예를 들어 N-아세틸-뉴라민산(NeuAc 또는 NANA)이다. 헥소스(Hex)가 또한 발견될 수 있다. N-글리칸은 "트리아마노실 코어"에 첨가된 주변 당(예를 들어 GlcNAc, 갈락토스, 푸코스 및 시알산)을 포함하는 가지("안테나" 또는 "아암")의 수에 있어서 상이하다. "M3", "M3GN2"로도 지칭되는 용어 "트리아마노실 코어", "트리아마노스 코어", "펜타사카라이드 코어" 또는 "소수-만노스 코어"는 디-GlcNAc 구조(GlcNAc₂)로부터 Man α1,3 아암 및 Man α1,6 아암이 연장되는 Man₃GlcNAc₂ 올리고사카라이드 구조: β1,4GlcNAc-β1,4GlcNAc를 나타낸다. N-글리칸은 그의 분지 구성성분에 따라 분류된다(예를 들어 고-만노스, 복합 또는 하이브리드).
- [0046] "고-만노스" 유형 N-글리칸은 디-GlcNAc 올리고사카라이드 구조에 4 이상의 만노스 잔기를 포함한다. "M4"는 Man₄GlcNAc₂를 나타낸다. "M5"는 Man₅GlcNAc₂를 나타낸다.
- [0047] "하이브리드" 유형 N-글리칸은 트리만노스 코어의 α1,3 만노스 (Man α1,3)아암의 말단에 하나 이상의 GlcNAc 잔기를 갖고 트리만노스 코어의 α1,6 만노스(Man α1,3) 아암에 0 이상의 만노스를 갖는다. 다양한 N-글리칸이 "당형태"로도 지칭된다. 하이브리드 글리칸의 예는 "GNM3GN2"이며, 이는 GlcNAcMan₃GlcNAc₂이다.
- [0048] "복합" 유형 N-글리칸은 전형적으로 Man α1,3 아암에 부착된 하나 이상의 GlcNAc 잔기를 갖고 트리만노스 코어의 Man α1,6 아암에 부착된 하나 이상의 GlcNAc를 갖는다. 복합 N-글리칸은 또한 시알산 또는 유도체로 개질될 수 있는 N-아세틸갈락토사민 잔기 또는 갈락토스를 가질 수 있다(예를 들어 "Neu"는 뉴라민산을 지칭하고 "Ac"는 아세틸을 지칭한다). 복합 N-글리칸은 또한 "2분(bisecting)" GlcNAc 및 코어 푸코스를 포함하는 쇠내 치환체를 가질 수 있다. 복합 N-글리칸은 또한 트리만노스 코어 상에 다중 안테나를 가질 수 있으며, 흔히 "다중 안테나 글리칸"으로 지칭되거나 또한 "다중-분지 글리칸"으로도 지칭되며, 트리-안테나, 테트라-안테나 또는 펜

타-안테나 글리칸일 수 있다.

- [0049] 용어 "G0"는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 를 지칭한다. 용어 "G0(1)"는 $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ 를 지칭하고, 용어 "G0(2)"는 $\text{GlcNAc}_4\text{Man}_3\text{GlcNAc}_2$ 를 지칭하며 용어 "G0(3)"는 $\text{GlcNAc}_5\text{Man}_3\text{GlcNAc}_2$ 를 지칭한다. 용어 "G1"은 $\text{GalGlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 를 지칭하고 "G2"는 $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 를 지칭하며 "G3"는 $\text{Gal}_3\text{GlcNAc}_{3-5}\text{Man}_3\text{GlcNAc}_2$ 를 지칭하고 "G4"는 $\text{Gal}_4\text{GlcNAc}_{4-5}\text{Man}_3\text{GlcNAc}_2$ 를 지칭하고 "G5"는 $\text{Gal}_5\text{GlcNAc}_5\text{Man}_3\text{GlcNAc}_2$ 를 지칭한다. 용어 "S1"는 $\text{NANAGal}_{1-5}\text{GlcNAc}_{1-5}\text{Man}_3\text{GlcNAc}_2$ 를 지칭하고 "S2"는 $\text{NANA}_2\text{Gal}_{2-5}\text{GlcNAc}_{2-5}\text{Man}_3\text{GlcNAc}_2$ 를 지칭한다. "S3"는 $\text{NANA}_3\text{Gal}_{3-5}\text{GlcNAc}_{3-5}\text{Man}_3\text{GlcNAc}_2$ 를 지칭하고 "S4"는 $\text{NANA}_4\text{Gal}_{4-5}\text{GlcNAc}_{4-5}\text{Man}_3\text{GlcNAc}_2$ 를 지칭하며 "S5"는 $\text{NANA}_5\text{Gal}_5\text{GlcNAc}_5\text{Man}_3\text{GlcNAc}_2$ 를 지칭한다.
- [0050] 본원에 사용된 용어 "대부분", 또는 "대부분의" 또는 "대부분인"과 같은 변형어는 당단백질이 제거된 후(예를 들어 PNGase로 처리되고 글리칸이 방출된 후), 그리고 질량 분광분석법, 예를 들어 MALDI-TOF MS로 분석하여 가장 높은 몰 퍼센트(%)의 총 N-글리칸을 갖는 것으로 측정된 글리칸 중을 의미하는 것으로 이해될 것이다. 환언하면, "대부분"이라는 어구는 다른 모든 개별 독립체보다 높은 몰 퍼센트로 존재하는 특정 당형태와 같은 개별 독립체로서 정의된다. 예를 들어, 조성물이 A종 40 몰 퍼센트, B종 35 몰 퍼센트 및 C종 25 몰 퍼센트로 이루어지는 경우, 조성물은 대부분 중 A를 포함한다. 용어 "풍부", "균일", "균질" 및 "본질적으로 이루어진"도 또한 글리칸에 대해 "대부분"과 동의어이다.
- [0051] 양성 모드로 MALDI-TOF-MS에 의해 측정된 전체 N-글리칸의 몰%는 전체 N-글리칸 몰%에 대해 이송된 사카라이드 몰%를 지칭한다. K⁺ 및 Na⁺와 같은 특정 양이온 부가물은 일반적으로 개별 부가물의 분자량에 의해 N-글리칸의 질량을 증가시키는 용리된 피크를 수반한다.
- [0052] 달리 나타내지 않는 한, 그리고 본원에 일반적 포맷 "서열"로 기술한 모든 서열에 대한 예로서, "서열 1을 포함하는 핵산"은 적어도 (i) 서열 1의 서열, 또는 (ii) 서열 1에 대한 상보적 서열의 부분을 갖는 핵산을 지칭한다. 둘 중에서의 선택은 문맥에서 지시된다. 예를 들어, 핵산이 프로브로서 사용되는 경우, 둘 중에서의 선택은 목적 표적에 대해 프로브가 상보적이어야 하는 요건에 의해 지시된다.
- [0053] "단리된" 또는 "실질적으로 순수한" 핵산 또는 폴리뉴클레오티드(예를 들어 RNA, DNA 또는 혼합된 중합체) 또는 당단백질은 천연 숙주 세포에 고유한 폴리뉴클레오티드를 자연적으로 동반하는 다른 세포 성분, 예를 들어 자연적으로 관련된 리보솜, 폴리머라제 및 게놈 서열로부터 실질적으로 분리된 것이다. 이 용어는 (1) 자연 발생 환경으로부터 제거된, (2) "단리된 폴리뉴클레오티드"가 자연에서 발견되는 폴리뉴클레오티드의 전부 또는 일부와 관련되지 않은, (3) 자연적으로 연결되지 않은 폴리뉴클레오티드에 인위적으로 연결된, 또는 (4) 자연 발생하지 않은 핵산, 뉴클레오티드를 포함한다. 용어 "단리된" 또는 "실질적으로 순수한"은 또한 재조합 또는 클론화 DNA 단리물, 화학적으로 합성된 폴리뉴클레오티드 유사체, 또는 이중 시스템에 의해 생물학적으로 합성된 폴리뉴클레오티드 유사체에 관해 사용될 수 있다.
- [0054] 그러나, "단리된"은 기술된 핵산, 폴리뉴클레오티드 또는 당단백질이 자체로 그의 고유 환경으로부터 물리적으로 제거되었을 것을 요구하지는 않는다. 예를 들어, 유기체의 게놈에서 내인성 핵산 서열은, 내인성 핵산 서열에 인접하여 이중 서열이 놓임으로써 상기 내인성 핵산 서열의 발현이 변경된 경우 "단리된" 것으로 간주된다. 본 문맥에서, 이중 서열은 해당 이중 서열이 자체로 내인성(동일 숙주 세포 또는 그의 후대에서 유래)인지 또는 외인성(다른 숙주 세포 또는 그의 후대에서 유래)인지와 관계 없이 내인성 핵산 서열에 고유하게 인접하지 않은 서열이다. 예를 들어, 프로모터 서열은 숙주 세포의 게놈에서 유전자의 고유 프로모터에 대해 치환(예를 들어 동종 재조합에 의해)되어 상기 유전자가 변경된 발현 패턴을 가질 수 있다. 상기 유전자는 고유하게 부속된 서열의 적어도 일부로부터 분리되었기 때문에 이제 "단리"되었을 것이다.
- [0055] 핵산은 또한 게놈에서 대응하는 핵산에 자연적으로 발생하지 않는 임의의 변형을 함유하는 경우 "단리된" 것으로 간주된다. 예를 들어, 내인성 코딩 서열은 삽입, 결실, 또는 인공적으로, 예를 들어 인간의 개입에 의해 도입된 점 돌연변이를 함유하는 경우 "단리된" 것으로 고려된다. "단리된 핵산"은 또한 예외적으로 존재하는 이중 부위 및 핵산 구조에서 숙주 세포 염색체에 통합된 핵산을 포함한다. 더욱이, "단리된 핵산"은 다른 세포 물질을 실질적으로 함유하지 않거나 재조합 기술로 제조된 경우 배양 매질을 실질적으로 함유하지 않거나, 화학적으로 합성된 경우 화학적 전구체 또는 기타 화학 물질을 실질적으로 함유하지 않는다.
- [0056] 본원에 사용된 용어인 치료 단백질의 "치료 유효량"은 소정의 질환 및/또는 그의 합병증의 임상적 징후를 치료하고 경감하거나 부분적으로 저지하기에 충분한 양을 지칭한다. 이를 달성하기에 적합한 양이 "치료 유효량"으

로서 정의된다. 각각의 목적을 위한 유효량은 질환 또는 부상의 중증도에 따라, 그리고 대상체의 체중 및 일반적인 상태에 따라 달라질 것이다. 적절한 투여량의 결정은 통상적인 실험을 이용하여, 관심 매트릭스를 구성하고 매트릭스 내의 상이한 지점을 시험하여 달성될 수 있고, 이는 모두 훈련된 의사 또는 수의사의 일반적인 기술 수준 내이다.

[0057] 용어 "치료", "치료하는" 및 그의 다른 변형어는 본원에서 질환 또는 장애와 같은 상태에 대처할 목적의 환자 또는 대상체의 처방 또는 관리를 지칭하는 데 사용된다. 상기 용어들은 환자가 앓는 주어진 상태에 대한 치료의 모든 범위를 포함하며, 예컨대 증상 또는 그의 합병증을 완화시키기 위해, 질환, 장애 또는 상태의 진행을 늦추기 위해, 질환, 장애 또는 상태를 치료 또는 제거하기 위해, 그리고/또는 상태를 방지하기 위해(방지는 질환, 상태 또는 장애에 대처하기 위한 목적으로 환자를 처방 또는 관리하는 것으로 이해된다) 당해 활성 화합물(들)을 투여하는 것을 포함하며, 증상 또는 합병증의 발생을 방지하기 위해 해당 활성 화합물(들)을 투여하는 것을 포함한다. 치료할 환자는 바람직하게는 포유동물, 특히 인간이지만, 다른 동물, 예컨대 개, 고양이, 소, 말, 양, 염소 또는 돼지의 치료가 본 발명의 범위에 포함된다.

[0058] 예를 들어, 심한 저혈당증(저혈당)으로 인한 인슐린 혼수상태 또는 인슐린 반응을 겪는 환자를 위한 본 발명의 글루카곤 펩티드의 치료 유효량은 성인의 경우 1mg(1유닛)이다. 44lb(20kg) 미만의 어린이에게 상기 유효량은 0.5mg이다. 글루카곤은 (1) 환자가 무의식이거나, (2) 환자가 당 또는 당-스위트닝된 제품을 섭취하지 못하거나, (3) 환자가 발작을 나타내거나, 또는 (4) 당 또는 당-스위트닝된 제품, 예컨대 통상의 소프트드링크 또는 과일 주스의 반복된 투여가 환자의 상태를 개선하지 못할 때 제공된다. 다른 경우, 용량은 0.25유닛 내지 2유닛의 범위일 수 있고, 근육내 또는 정맥내 투여될 수 있다. 1mg의 순수한 글루카곤은 대략 1유닛에 상응한다. 투여 스케줄은 변경될 수 있으나, 대략 1일 1회 내지 이벤트마다 필요한 만큼으로 달라질 수 있다. 실제 스케줄은 환자에게 투여되는 글루카곤의 종류(글루카곤 또는 글리코실화-글루카곤) 및 개별 환자의 반응을 비롯한 많은 인자에 의존할 것이다. 높은 용량 범위는 일반적으로는 저혈당증 용도에는 사용되지 않지만 다른 치료 용도에서는 유용할 수 있다. 환자에 대한 적절한 용량을 달성 및 수립하는 수단은 관련 기술분야에서 공지되었으며 통상 시행되고 있다.

[0059] 본원에서 사용된 용어 "제약상 허용되는"은 그의 일반적인 의미로 사용된다. 제약상 허용되는 조성물은 일반적으로 제제의 다른 물질과 상용성이며, 일반적으로 대상체에게 해롭지 않다.

[0060] 본 발명의 모든 조성물은 대상체에 치료 유효 용량으로 투여될 수 있다. 백신의 경우, 본원에 사용된 "치료 유효" 또는 "유효" 양 또는 용량은 대상체 내에 면역성 또는 내성을 유도하는 데 필요한 양이고/이거나 대상체가 보다 효과적으로 질환에(예를 들어 외래 병원체, 암, 자가면역 질환 등에 대해) 저항할 수 있게 하는 데 필요한 양을 의미한다. 대상체에 투여되었을 때, 유효량은 치료되는 특정 상태 및 목적하는 결과에 의존할 것이다. 치료 유효 용량은 예를 들어 이하에 추가로 기술하는 것과 같은 요인을 채용하고 상용 실험의 범위 내에서 통상의 기술자가 결정할 수 있다.

[0061] 일부 실시양태에서, 치료 유효량은 초기에 세포 배양 검정으로부터 결정될 수 있다. 예를 들어 수지상 세포 반응을 유도하기에 유용한 본 발명의 조성물의 유효량은 자극 지수에 대해 시험관내 검정을 이용하여 가늠할 수 있다. 자극 지수는 특정 대상체에 대한 본 발명의 특정 조성물의 유효량을 결정하는 데 사용될 수 있으며, 대상체 내 목적하는 수준을 달성하기 위해 용량을 상향 또는 하향 조정할 수 있다. 치료 유효량은 동물 모델로부터 결정할 수도 있다. 적용된 용량은 투여된 조성물의 상대적 생체이용률 및 효력을 기초로 조정될 수 있다. 상기 방법 및 다른 방법에 기초하여 최대 효능을 얻기 위해 용량을 조정하는 것은 통상의 기술자의 능력 범위 내이다. 상기 용량은 상용 실험의 범위 내에서 조정될 수 있다.

[0062] 본 발명의 조성물을 대상체에게 투여시, 복용량, 복용 스케줄, 투여 경로 등은 상기 조성물의 공지된 활성을 나타내도록 선택될 수 있다. 용량은 선택적으로 본 발명의 조성물의 검정 결과와 조합하여, 실험 모델 결과를 기초로 추정할 수 있다. 용량은 투여 모드에 따라 목적하는 조성물 농도, 국부 또는 전신을 달성하도록 적절하게 조정될 수 있다. 용량은 1일 1회 또는 수회 투여로 제공될 수 있다. 특정 대상체의 반응이 그러한 용량으로 불충분한 경우에는, 대상체 내성이 허용하는 한도까지 보다 높은 용량(또는 상이한, 보다 국부적인 전달 경로에 의해 실질적으로 더 높은 용량)이 채용될 수 있다. 1일 당 다중 복용 또한 어떤 경우 대상체 내 또는 대상체의 활성 부위 내 조성물의 적절한 전신 농도를 달성하도록 고려될 수 있다.

[0063] 대상체에 대한 조성물의 용량은 치료 유효량의 조성물이 대상체 내 조성물 활성 부위, 즉 체내의 수지상 세포 및/또는 다른 항원-제공 세포에 도달하도록 하는 양일 수 있다. 용량은 어떤 경우 대상체 내 임의의 잠재적으로 해로운 부작용을 피하거나 최소화하면서 최대량으로 제공될 수 있다. 실제로 투여되는 조성물 용량은 활성

부위에서의 최종 목표 농도, 대상체로의 투여 방법, 조성물의 효능, 대상체 내 조성물의 수명, 투여 타이밍, 공동 치료의 효과(예를 들어 혼합제에서) 등과 같은 인자들에 의존한다. 전달되는 용량은 또한 대상체 내 관련 상태에도 의존할 수 있고, 어떤 경우 대상체마다 달라질 수 있다. 예를 들어, 대상체의 연령, 성별, 체중, 크기, 환경, 신체 상태 또는 현재 건강 상태 또한 활성 부위에서의 조성물의 요구되는 용량 및/또는 농도에 영향을 줄 수 있다. 상이한 개인별로 또는 심지어 상이한 날에 동일 개인에서 용량을 변화시킬 수 있다. 타당한 의학적 판단에 따라 최대 용량, 즉 최고 안전 용량을 사용하는 것이 바람직할 수 있다. 바람직하게는, 복용 형태는 대상체에게 크게 해를 주지 않도록 하는 형태이다. 특정 실시양태에서, 조성물을 예방 수단으로 대상체에게 투여할 수 있다. 일부 실시양태에서, 본 발명의 조성물은 인구통계 또는 역학 연구에 기초하여 대상체에게 투여하거나 특정 분야 또는 직업의 대상체에게 투여할 수 있다.

[0064] 본 발명의 조성물의 투여는 그의 표적, 즉 체내의 수지상 세포 및/또는 다른 항원-제공 세포에 조성물이 도달하게 하는 의학적으로 허용되는 임의의 방법으로 달성될 수 있다. 선택된 특정 모드는 물론 상술한 것과 같은 인자들, 예를 들어 특정 조성물, 치료 대상체의 상태 중증도, 치료 효능에 요구되는 용량 등에 의존할 것이다. 본원에서 사용된 "의학적으로 허용되는" 치료 모드는 임상적으로 허용 가능하지 않은 부작용을 유발하지 않으면서 대상체 내 조성물의 유효 수준을 제공할 수 있는 모드이다.

[0065] 대상체에게 조성물을 투여하는 데 임의의 의학적으로 허용되는 방법이 사용될 수 있다. 투여는 치료할 상태에 따라 국부적(즉, 특정 영역, 생리학적 시스템, 조직, 기관 또는 세포 유형에) 또는 전신적일 수 있다. 예를 들어, 조성물은 폐로, 비강으로, 경피로, 비경구 주사를 통해 또는 피하 주입으로, 수술적 투여로, 또는 본 발명의 조성물에 의한 표적으로의 접근이 달성되는 임의의 다른 방법으로 투여될 수 있다. 본 발명에서 사용될 수 있는 비경구 모드의 예는 정맥내, 피내, 피하, 공동내, 근육내, 복강내, 경막외, 또는 척수강내를 포함한다. 피하 주입 방식의 예는 임의의 피하주입 가능한 또는 주사가 가능한 약물 전달 시스템을 포함한다.

[0066] 본 발명의 특정 실시양태에서, 본 발명의 조성물의 투여는 특정 시간 주기, 예를 들어 시간, 일, 주, 달 또는 년에 걸쳐 조성물에 순차적으로 노출되도록 고안될 수 있다. 이는 예를 들어 상술한 방법 중 하나에 의해 본 발명의 조성물의 반복된 투여에 의해, 또는 반복된 투여 없이 장기간에 걸쳐 조성물이 전달되는 지속 또는 제어 방출 전달 시스템에 의해 달성될 수 있다. 이러한 전달 시스템을 이용한 조성물의 투여는 예를 들어 경구 투여 형태, 볼루스 주사, 경피 패치 또는 피하 임플란트에 의해 수행될 수 있다. 조성물의 실질적으로 일정한 농도를 유지하는 것이 일부 경우 바람직할 수 있다.

[0067] 조성물은 또한 통상적인 스케줄로 투여될 수 있으나, 대안적으로는 증상이 나타날 때마다 투여될 수 있다. 본원에서 사용된 "통상적인 스케줄"은 소정의 지정 기간을 지칭한다. 통상적인 스케줄은 동일하거나, 스케줄이 예정된 것인 한 길이가 다른 기간을 포함할 수 있다. 예를 들어, 통상적인 스케줄은 1일 1회, 2일 1회, 3일 1회, 4일 1회, 5일 1회, 6일 1회, 주당 1회, 2주당 1회, 1개월당 1회, 2개월당 1회 또는 그 사이의 임의의 설정된 수의 일 또는 주, 격월, 3개월, 4개월, 5개월, 6개월, 7개월, 8개월, 9개월, 10개월, 11개월, 12개월 당 1회 조성물의 투여등을 포함할 수 있다. 대안적으로, 소정의 통상적인 스케줄은 첫 주에는 매일 조성물을 투여하고, 이어서 몇 달 동안 월 단위로 투여하며, 이어서 3개월마다 투여하는 것을 포함할 수 있다. 미리 결정되어 적절한 스케줄이 특정 일의 투여를 포함하는 한 임의의 특정 조합이 통상적인 스케줄에 포함될 수 있다.

[0068] 일부 경우, 조성물은 알레르기 이벤트를 방지하기 위해 알레르기 이벤트가 예상되는 환자에게 투여될 수 있다. 알레르기 이벤트는 천식 발작, 계절성 알레르기성 비염 (예를 들어 고초열, 화분, 돼지풀 과민증) 또는 통년성 알레르기성 비염 (예를 들어 예컨대 본원에 기술된 알레르겐에 대한 과민증)일 수 있으나, 이에 한정되는 것은 아니다. 일부 경우, 조성물은 알레르기 이벤트의 훨씬 전에 투여될 수 있다. 본원에서 사용된 "훨씬 전"은 알레르기 이벤트의 적어도 6개월, 적어도 5개월, 적어도 4개월, 적어도 3개월, 적어도 2개월, 적어도 1개월, 적어도 3주, 적어도 2주, 적어도 1주, 적어도 5일, 또는 적어도 2일 전을 의미한다.

[0069] 유사하게, 조성물은 알레르기 이벤트 직전(예를 들어 알레르기 이벤트의 48시간 내, 24시간 내, 12시간 내, 6시간 내, 4시간 내, 3시간 내, 2시간 내, 1시간 내, 30분 내 또는 10분 내), 알레르기 이벤트와 실질적으로 동시에(예를 들어 대상체가 알레르겐에 접촉하는 시간 도중 또는 대상체가 알레르기 증상을 겪는 시간 도중), 또는 알레르기 이벤트 후 투여될 수 있다. 특정 알레르겐에 대해 대상체를 탈감작시키기 위해서는, 종래의 탈감작 요법에 따라 항원 또는 알레르겐을 함유하는 접합체를 매우 작은 용량으로 일정 기간 동안 투여할 수 있다.

[0070] 본 발명에서 사용하기에 적합한 다른 전달 시스템은 시간-방출, 지연 방출, 지속 방출 또는 제어 방출 전달 시스템을 포함한다. 그러한 시스템은 많은 경우 조성물의 반복된 투여를 피할 수 있고 대상체에게 편리함을 증가시킬 수 있다. 많은 종류의 지연 전달 시스템이 가용하며 통상의 기술자에게 공지되어 있다. 이들은 예를 들

어 중합체-기재 시스템, 예컨대 폴리락트산 및/또는 폴리글리콜산, 폴리무수물, 폴리카프로락톤 및/또는 이들의 조합; 스테롤, 예컨대 콜레스테롤, 콜레스테롤 에스테르, 및 지방산 또는 중성 지방, 예컨대 모노-, 디- 및 트리글리세라이드를 포함하는 지질-기재인 비중합체 시스템; 히드로겔 방출 시스템; 리포솜-기재 시스템; 인지질-기재 시스템; 시알 시스템; 펩티드 기재 시스템; 왁스 코팅; 종래의 결합제 및 부형제를 사용한 압축 정제; 또는 부분적으로 융합된 임플란트를 포함한다. 제제는 예를 들어 마이크로구체, 히드로겔, 중합체 저장소, 콜레스테롤 매트릭스, 또는 중합체 시스템일 수 있다. 일부 실시양태에서, 시스템은 예를 들어 조성물을 함유하는 제제의 확산 또는 침식/분해 속도의 제어를 통해 조성물의 지속 또는 제어된 방출을 허용할 수 있다. 또한, 펌프 기반의 하드웨어 전달 시스템이 사용되어 본 발명의 하나 이상의 실시양태를 전달할 수 있다.

[0071] 장기 방출 장치의 사용이 본 발명의 일부 실시양태에서 특히 적합할 수 있다. 본원에서 사용된 "장기 방출"은 조성물을 함유하는 장치가 치료상 유효한 수준의 조성물을 적어도 30일 또는 45일, 바람직하게는 적어도 60일 또는 90일, 또는 심지어 이보다 긴 기간 동안 전달하도록 구성 및 설치된 것을 의미한다. 장기 방출 임플란트는 통상의 기술자에게 공지되었으며, 상기 방출 시스템의 일부를 포함한다.

[0072] **글리코실화 조작**

[0073] 본원에 제공된 신규 발현 시스템 및 방법을 사용하여, 원핵 숙주 세포의 당조작을 통해 고-만노스, 하이브리드 및 복합 글리칸의 제조를 위해 본 발명의 다양한 측면들이 제공된다. 본 발명의 한 측면은 구조적으로 균질 인간-유사 글리칸으로 N-연결된 당단백질을 발현시키기 위한 생합성 경로를 포함하는 재조합 원핵 숙주에 관한 것이다. 본 발명의 적용은 치료 단백질에 대한 개선된 생화학적 및 약동학적 안정성을 포함한다. 추가의 실시양태는 대상체 내에서 보호 면역을 유발할 수 있는 탄수화물-접합된 백신을 제공하는 방법 및 조성물을 제공한다. 안전하고 보다 효과적인 당단백질 및 백신을 제조하기 위한 신속한 미생물-기반의 제조 방법이 본 발명의 목적이다.

[0074] **원핵생물에서의 고-만노스 유형 글리칸 제조**

[0075] 트리만노실 코어로부터 출발하여, 본 발명은 도 1에 나타낸 바와 같은 고-만노스 유형 글리칸을 제조하기 위한 만노실트랜스퍼라제 효소의 재조합 발현을 위한 방법을 제공한다. 일 실시양태에서, 상기 방법은 2개의 GDP-만노스 잔기의 원핵 숙주 세포 내의 트리만노스 올리고사카라이드 조성물로의 이송을 촉매하는 하나 이상의 알파-1,2-만노실트랜스퍼라제 효소 활성(EC 2.4.1.131)을 발현하기 위해 재조합 원핵 숙주 세포를 배양하는 것을 제공한다. **실시예 3**은 α -1,2 만노실트랜스퍼라제 효소 활성(EC 2.4.1.131)의 발현을 기재한다. 바람직한 α -1,2-만노실트랜스퍼라제 효소 활성은 용해도 인핸서인 GST에 융합된 S. 세레비지아에 *alg11*에 의해 인코딩된다. **표 1**은 다양한 용해도 인핸서를 열거한다.

[0076] 따라서, 본 발명은 고-만노스 유형 올리고사카라이드 조성물의 제조 방법을 제공하며, 이 방법은, 말단 만노스 잔기에 GDP-만노스 잔기의 이송을 촉매하는 하나 이상의 알파-1,2-만노실트랜스퍼라제 효소 활성(EC 2.4.1.131)을 발현하기 위해 말단 만노스 잔기를 갖는 올리고사카라이드 조성물을 제조하는 재조합 원핵 숙주 세포를 배양하는 것을 포함하며, 상기 배양 단계는 적어도 4개의 말단 만노스 잔기를 갖는 올리고사카라이드 조성물을 제조하는데 유효한 조건 하에 수행된다. 특정 실시양태에서, 올리고사카라이드 조성물은 트리만노스 코어에 적어도 2개의 추가의 만노스 잔기를 포함한다. 바람직한 실시양태에서, 백신 후보는 원핵 숙주 세포 내에서 이들이 M5 당형태에 N-연결된 곳에서 재조합적으로 발현된다. **도 1**에 나타낸 예상되는 주요 당형태의 구조는 $\text{Man } \alpha 1-2 \text{ Man } \alpha 1-3(\text{Man } \alpha 1-6)-\text{Man } \beta 1-4-\text{GlcNAc } \beta 1-4-\text{GlcNAc}$ 이다.

[0077] 본 발명의 원핵 숙주 세포에서, 글리코실화 효소는 당단백질의 글리코실화 전에 지질-연결된 글리칸에 작용한다. 진핵생물에서, 알파-1,2-만노실트랜스퍼라제는 세포질 세망 막의 시토솔 측 상의 돌리콜 피로포스페이트에 연결된 트리만노스 코어 글리칸에 작용한다. 이어서 $\text{Man}_3\text{GlcNAc}_2$ -돌리콜 피로포스페이트는 $\text{Man}_5\text{GlcNAc}_2$ -돌리콜 피로포스페이트에 매우 특이적인 내인성 플리파제 효소에 의해 내인성 세망 내로 플립핑되어 플립핑 전에 올리고사카라이드의 완전한 조립을 보장한다(Sanyal & Menon, PNAS 2009). 원핵생물에서, $\text{Man}_3\text{GlcNAc}_2$ 지질은 플립핑될 수 있는 것이 밝혀졌으며(문헌 [Valderrama-Rincon, et. al. "An engineered eukaryotic protein glycosylation pathway in Escherichia coli," Nat. Chem. Biol. AOP (2012)]), 트리만노스 코어 너머 올리고사카라이드의 조립을 플립핑하고 저해하는 공지된 특이성은 없다. 따라서, 본 발명의 목적은 M3 올리고사카라이드 기질 상에 만노스 잔기를 이송하고 또한 올리고사카라이드의 주변 세포질로의 플립핑 활성을 촉매하는 원핵 시스템 내 $\text{Man}_{7-9}\text{GlcNAc}_2$, $\text{Man}_6\text{GlcNAc}$, $\text{Man}_5\text{GlcNAc}_2$ 및 $\text{Man}_4\text{GlcNAc}_2$ 를 포함하는 고-만노스 유형 올리고사카라이드 조성물을 제조하는 것이다. 바람직한 실시양태에서, 숙주 세포는 50몰% 이상의 고-만노스 유형 글리칸을 제

조한다.

[0078] 원핵생물에서의 GnT 발현

[0079] 특정 측면에서, 올리고사카라이드 조성물의 제조 방법이 제공되며, 이 방법은, 말단 만노스 잔기에 UDP-GlcNAc 잔기의 이송을 촉매하는 하나 이상의 N-아세틸글루코사미드 트랜스퍼라제 효소 활성(EC 2.4.1.101; EC 2.4.1.143; EC 2.4.1.145)을 발현하기 위해 상기 말단 만노스 잔기를 갖는 올리고사카라이드 조성물을 제조하는 제조함 원핵 숙주 세포를 배양하는 것을 포함하며, 상기 배양 단계는 말단 GlcNAc 잔기를 갖는 올리고사카라이드 조성물을 제조하는데 유효한 조건 하에 수행된다. 진핵생물에서, N-아세틸글루코사미드 트랜스퍼라제는 글리코실화 단백질의 아스파라긴 잔기에 공유 결합으로 연결된 올리고사카라이드에 작용한다. 원핵생물에서, 올리고사카라이드는 하이브리드 및 복합 올리고사카라이드의 제조를 저해하는 단백질 글리코실화 과정과는 독립적으로 제조된다.

[0080] 하이브리드 당형태를 제조하기 위해, UDP-GlcNAc 잔기는 수용자 기재인 트리만노실 코어 올리고사카라이드 구조의 Man α 1,3 아암에 이송된다. 모식적 실시양태에서, 본 발명은 Alg13, Alg14, Alg1 및 Alg2를 발현하는 숙주 세포 내 MBP 용해도 인핸서에 융합된 N. 타바쿰 GnTI를 인코딩하는 유전자로 형질 전환된 원핵 숙주 세포를 제공한다. 하이브리드 당형태 GlcNAcMan₃GlcNAc₂는 **도 2의 A**에 나타낸 바와 같이 제조된다. 도시된 당형태의 예상 구조는 β 1-2-GlcNAcMan α 1-3(Man α 1-6)-Man β 1-4-GlcNAc β 1-4-GlcNAc이다.

[0081] 복합 당형태를 제조하기 위해, UDP-GlcNAc 잔기는 수용자 기재인 트리만노실 코어 올리고사카라이드 구조의 Man α 1,3 및 Man α 1,6 아암 둘 다에 이송된다. 본 실시양태에서, 원핵 숙주 세포는 Alg13, Alg14, Alg1, Alg2 및 GnTI를 발현하는 숙주 세포 내 MBP에 융합된 인간 GnTII를 인코딩하는 유전자로 형질 전환된다. 복합 GlcNAc₂Man₃GlcNAc₂ (GO) 당형태가 **도 3**에 나타낸 바와 같이 제조되며, 예상되는 구조는 β 1-2-GlcNAcMan α 1-3(β 1-2-GlcNAc Man α 1-6)-Man β 1-4-GlcNAc β 1-4-GlcNAc이다.

[0082] 본 발명의 다른 측면에서, 다중-안테나 글리칸이 제조된다. 예를 들어, 원핵 숙주 세포는 Alg13, Alg14, Alg1, Alg2 및 GnTI를 발현하는 숙주 세포 내 MBP에 융합된 소 GnTIV를 인코딩하는 유전자로 형질 전환된다. **도 4의 A**는 본 발명의 방법을 이용하여 제조된 GlcNAc₂Man₃GlcNAc₂ 하이브리드 당형태를 나타내며, 여기서 2개의 UDP-GlcNAc 잔기가 트리만노실 코어의 Man α 1,3 아암에 이송되었다. 도시된 당형태의 예상 구조는 β 1-2-GlcNAc(β 1-2-GlcNAc) Man α 1-3(Man α 1-6)-Man β 1-4-GlcNAc β 1-4-GlcNAc이다.

[0083] 대안적 실시양태에서, 글리칸은 또한 생체의, 예를 들어 **실시예 7**에 기재된 바와 같이 올리고사카라이드의 효소적 합성을 통해 형성될 수 있다. 예를 들어, **도 5**는 GlcNAc₃Man₃GlcNAc₂ 당형태를 포함하는 복합, 다중-안테나 글리칸의 MS를 도시하며, 이는 GnTI, GnTII, GnTIV(생체외), Alg13, Alg14, Alg1 및 Alg2를 발현하여 트리만노실 코어 올리고사카라이드 구조의 Man α 1,3 상으로의 2개의 UDP-GlcNAc 잔기의 이송 및 Man α 1,6 아암 상으로 하나의 UDP-GlcNAc 잔기의 이송을 유발함으로써 제조된다. 도시된 당형태의 예상 구조는 (β 1-2-GlcNAcMan α 1-3)β 1-2-GlcNAc(β 1-2-GlcNAc Man α 1-6)-Man β 1-4-GlcNAc β 1-4-GlcNAc이다.

[0084] GnTV(EC 2.4.1.155) 및 GnTVI(2.4.1.201)과 같은 추가의 GnT 활성이 원핵 시스템에서 발현될 수 있다. 그 결과, 본 발명의 방법을 이용하여 트리만노스 코어에 5개까지의 분지가 있는 다중 안테나 글리칸이 가능하다. 다중 분지 글리칸은 예를 들어 에리트로포이에틴 상의 향상된 시알화를 가능하게 하여 혈청 반감기 및 효력을 향상시킨다(문헌 [Elliot, Nature Biotech 2003; Misaizu, Blood 1995]).

[0085] 글리코실트랜스퍼라제 용해도 인핸서

[0086] 다양한 GnT가 숙주 세포에서 발현될 수 있으나, 바람직한 실시양태에서는, GnT는 예를 들어 MBP에 융합되고 융합 단백질로서 발현되어 말단 UDP-GlcNAc 잔기를 트리만노실 코어로 전이시키고, 사실상 글리코실트랜스퍼라제의 용해도를 향상시킬 수 있다. **표 1**은 막 표적화 도메인 및 용해도 인핸서의 부류를 제공하는 목록을 제공한다.

표 1

용해도 인핸서

	글리칸 합성	
융합 파트너	Alg11	GnTI
없음	-	-

DsbA	-	+
GlpF	+/-	+
GST	+	+
MBP (EC# POAEX9)	+/-	+
MstX	+	+
NusA	-	N/A
TrxA	-	N/A

[0088] 융합 라이브러리를 사용하여, $GlcNAc_{(1-5)}Man_3GlcNAc_2$ 와 같은 글리칸이 본 발명의 원핵 시스템에서 생성된다. 본 발명의 특정 측면에서, MBP-융합된 글리코실트랜스퍼라제는 원핵 숙주에서 발현된다. 다른 막 표적화 도메인 및 용해도 인핸서, 예컨대 MstX가 또한 발현될 수 있다. 그러한 *N*-아세틸글루코사민일 트랜스퍼라제-MBP 또는 *N* 아세틸글루코사민일 트랜스퍼라제-MstX 융합은 UDP-GlcNAc 잔기의 수용자 올리고사카라이드 기질에의 부가에 대해 스크리닝된다. 바람직한 실시양태에서, 하기 융합, 즉, *N. 타바쿰* GnTI-MBP, *H. 사피엔스* GnTII-MBP, *B. 타우루스* GnT IV-MBP가 UDP-GlcNAc의 트리만노실 코어로의 전이를 제공한다. 따라서, GnT 융합의 라이브러리를 만들어 원핵 숙주 세포에서 하이브리드, 복합 및 멀티-안테나(multi-antennary) 글리칸을 생성할 수 있다. 다양한 GnT 융합 구축물이 본 발명의 방법을 사용하여 제조될 수 있다. 그러한 융합 구축물은 본 발명의 범위에 속하고 보다 우수한 활성 또는 향상된 용해도에 대해 스크리닝될 수 있다.

[0089] **원핵생물에서의 갈락토실트랜스퍼라제 발현**

[0090] 본 발명의 추가의 측면에서, 올리고사카라이드 조성물을 제조하기 위한 방법이 제공되며, 상기 방법은 숙주 세포를 배양하여 UDT-갈락토스 잔기의 상기 말단 GlcNAc 잔기로의 전이를 촉매하는 하나 이상의 갈락토실트랜스퍼라제 효소 활성화 (EC 2.4.1.38, EC 2.7.8.18)을 발현하는 단계를 포함하며, 상기 배양 단계는 말단 갈락토스 잔기를 갖는 올리고사카라이드 조성물을 생성하기에 효과적인 조건하에서 실시된다. **도 6**은 *E. 콜라이*에서 생성된 하이브리드 당형태 $GalGlcNAcMan_3GlcNAc_2$ 의 MS를 도시한다. **실시예 5**는 *E. 콜라이*에서의 *헬리코박터 필로리* β -1,4GalT의 발현을 기재하며, 이는 UDT-갈락토스 잔기를 $GlcNAcMan_3GlcNAc_2$ 수용자 올리고사카라이드로 전이시킨다.

[0091] 원핵세포에서 하이브리드 갈락토실화 당형태를 생성하기 위해, UDT-갈락토스 잔기를 트리만노실 코어의 β -1,2GlcNAcMan α 1,3으로 전이시키며, 복합 당형태를 위해서는 트리만노실 코어의 β -1,2GlcNAcMan α 1,3 및 β -1,2GlcNAcMan α 1,6 아암 둘 다로 전이시킨다. 그러한 실시양태에서, 원핵 숙주 세포는 숙주 세포에서 *H. 필로리* GalT를 코딩하는 유전자로 형질전환되어 Alg13, Alg14, Alg1, Alg2, GnTI 및 GnTII을 발현한다. **실시예 8**은 복합 $Gal_2GlcNAc_2Man_3GlcNAc_2$ 당형태의 제조 방법을 제공한다. **도 7**은 m/z 1662.2에서의 피크를 보여주는데, 이것은 복합 갈락토실화 글리칸 $Gal_2GlcNAc_2Man_3GlcNAc_2$ 의 질량과 상관된다. $Gal_{(1-4)}GlcNAc_2Man_3GlcNAc_2$ 를 포함하는 추가의 갈락토실화 당형태가 생성될 수 있다. 하이브리드 말단 갈락토스 글리칸의 예상 구조는 β 1-4Gal β 1-2-GlcNAcMan α 1-3(Man α 1-6)-Man β 1-4-GlcNAc β 1-4-GlcNAc이고, 복합 말단 갈락토스 글리칸은 β 1-4Gal β 1-2-GlcNAcMan α 1-3(β 1-4Gal β 1-2-GlcNAc Man α 1-6)-Man β 1-4-GlcNAc β 1-4-GlcNAc이다.

[0092] 다양한 다른 유기체로부터의 갈락토실트랜스퍼라제가 발현될 수 있으며, 이는 *헬리코박터 필로리*, *나이세리아 매닝기티데스*, *나이세리아 고노레아에*, *리슈마니아 도노바니*, *호모 사피엔스* (GALT), *보스 타우루스*, *드로소피야*, *펠라노가스터*, *래터스 노르베기쿠스* (GalT I), *무스 무스쿨루스*, *크리세톨루스 그리세우스*, *에쿠우스 카발루스*, *마크로푸스 유레니이* (4 β -GalT), *다니오 레리오* (GalT I) 및 *수스 스크로파*, *오비스 아리에스*를 포함하지만 이에 한정되지 않는다.

[0093] 일부 실시양태에서, 다양한 갈락토실트랜스퍼라제 효소 활성이 MBP 또는 mstX와 같은 용해도 인핸서와 융합되고, UDT-갈락토스의 수용자 올리고사카라이드 기질에의 부가에 대해 스크리닝된다. GnT와 달리, 인간 및 소 GalT-mstX 융합은 UDP-갈락토스를 말단 GlcNAc 올리고사카라이드 기질로 전이시키지 않는 것으로 나타났다.

[0094] 더 바람직한 실시양태에서, 산화성 박테리아 균주를 *H. 필로리* β -1,4-GalT의 발현에 사용한다.

[0095] 예시적인 실시양태에서, 하기 효소가 원핵 숙주에서 발현된다: Alg13, Alg14, Alg1, Alg2, *니코티아나 타바쿰* GnTI, 인간 GnTII, 소 GnTIV, *헬리코박터 필로리* β -1,4GalT. GnT 및 GalT가 산화성 박테리아 숙주에서 발현

된다.

[0096] **원핵생물에서의 시알릴트랜스퍼라제 발현**

[0097] 완전 복합 올리고사카라이드 구조는 말단 시알산, 예를 들어, NANA 잔기로 종결된다. 원핵생물에서의 시알릴트랜스퍼라제의 발현은 상당한 관심사였다. 몇몇 그룹은 수년 동안 당단백질 생성을 위한 시알산 전이에 관한 과제를 수행하였으나, 현재까지 원핵생물에서 인간-유사 글리칸을 생성하는 시알산 전이의 생성에 관한 보고는 존재하지 않는다.

[0098] 따라서, 본 발명은 재조합 원핵 숙주를 배양하여, CMP-NANA 잔기의 상기 말단 갈락토스 잔기로의 전이를 촉매하는 하나 이상의 시알릴트랜스퍼라제 효소 활성 (EC 2.4.99.4 및 EC 2.4.99.1)을 발현하며, 상기 배양 단계는 말단 시알산 잔기를 갖는 올리고사카라이드 조성물을 제조하기에 효과적인 조건하에서 실시되는, 올리고사카라이드 조성물의 제조 방법을 제공한다. 다양한 시알릴트랜스퍼라제가 본 발명의 방법을 사용하여 *생체내* 또는 *생체의* 발현된다. 한 실시양태에서, α-2,3 시알릴트랜스퍼라제 (EC 2.4.99.4)가 숙주 세포에서 또는 배양 배지에서 발현된다. 추가의 실시양태에서, α-2,6 시알릴트랜스퍼라제 (EC 2.4.99.1)가 숙주 세포에서 또는 배양 배지에서 발현된다.

[0099] 바람직한 실시양태에서, 하기 효소가 원핵 숙주에서 발현된다: Alg13, Alg14, Alg1, Alg2, *니코티아나 타바쿰* GnTI, 소 GnTIV, *헬리코박터 필로리* β-1,4-GalT 및 *P. 담셀라에* ST6. 상기 방법은 CMP-NANA의 올바른 올리고사카라이드 기질로의 적당한 전이를 나타내는 *생체내* 및 *생체의* 반응의 조합을 허용한다. 도 8에 나타난 바와 같이, 하이브리드 시알릴화 당형태가 생성되며, 나타난 당형태의 예상 구조는 2,6NANA β 1-4Gal β 1-2-GlcNAcMan α 1-3(Man α 1-6)-Man β 1-4-GlcNAc β 1-4-GlcNAc이다.

[0100] **당 뉴클레오티드 전구체**

[0101] 또 다른 실시양태에서, 상기 방법은 숙주 세포를 배양하여 당 뉴클레오티드 전구체를 증가시킨다. 예를 들어, GDP-만노스 합성을 촉매하는 효소가 시스템에서 발현된다. 포스포만노뉴타제 효소 활성 (ManB) (EC 5.4.2.8) 및 만노스-1-포스페이트 구아닐릴트랜스퍼라제 효소 활성 (ManC) (EC 2.7.7.13)이 본 발명의 숙주 세포에 도입된다. 도 9A (좌측)는 ManC/B가 과다발현될 경우 트리만노실 코어의 증가된 생성을 보여준다.

[0102] 추가의 실시양태에서, 세포질 내의 글리코실 공여자의 충분한 풀(pool)이 생성된다. GnTI 및 GnTII을 위한 기질인 UDP-GlcNAc는 E. 콜라이 세포질 내에 고유하게 존재하지만, 숙주 세포를 조작하여 UDP-GlcNAc 합성을 증가시킬 수 있다. 그러한 실시양태에서, 상기 방법은 UDP-GlcNAc 합성을 촉매하는 글루타민-프룩토스-6-포스페이트 트랜스아미나제 효소 활성 GImS (EC 2.6.1.16), GImU (EC 2.7.7.23 & EC 2.3.1.157), GImM (EC 5.4.2.10)을 발현함으로써 UDP-GlcNAc를 증가시키는 숙주 세포의 배양을 제공한다. FIG. 9A (우측)는 GImS가 과다발현될 경우 GlcNAcMan₃GlcNAc₂의 증가를 보여준다. ManC/B를 갖는 글리세롤의 부가는 도 9B에 나타난 바와 같이 증가된 글리칸 수율을 초래한다. 피루베이트 또한 도 9C에 나타난 바와 같이 글리칸 수율을 증가시키는 것으로 나타난다.

[0103] ManC/B의 과다발현은 도 10에서 입증되는 바와 같이 생성되는 글리칸의 균질성에 극적인 효과를 미친다. M3 당형태 (D), M5 당형태 (E) 및 GNM3GN2 (F)는 반응의 불완전한 뉴클레오티드 당 전이로 인한 것일 수 있는 피크들이 제거된 것으로 나타나는 우세한 글리칸을 생성하였다. 따라서, 숙주 세포는 생성되는 정확한 당형태를 생성 및 조절할 수 있다.

[0104] 효모에서, 보브로비츠(Bobrowicz) 등은 *P. 파스토리스*에서의 UDP-갈락토스 수송체, UDP-갈락토스 4-에피머라제 및 β1,4GalT의 발현을 통해 말단 갈락토실화 글리칸 *피치아*의 생성이 증가됨을 보였다 ([Bobrowicz et al., Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. Glycobiology 2004 Sep;14(9):757-66.]). UDP-갈락토스는 또한 E. 콜라이의 세포질에 고유하게 존재하지만, UDP-갈락토스의 입수가능성은 우리딜레이트 키나제 (*pyrH*), Glc-1-P 우리딜트랜스퍼라제 (*galU*), Gal-1-P 우리딜트랜스퍼라제 (*galT*), 갈락토키나제 (*galK*), 및 UDP-갈락토스 에피머라제 (*galE*)를 포함하는 UDP-Gal 합성 유전자의 과다발현에 의해 증가될 수 있는 것으로 연구되었다 ([Chung, S., et al., Galactosylation and sialylation of terminal glycan residues of human immunoglobulin G using bacterial glycosyltransferases with in situ regeneration of sugar-nucleotides. Enzyme and Microbial Technology, 2005. 39(1): p. 60-66]). 따라서, 바람직한 실시양태에서, E. 콜라이 K12로부터의 *galETK*, *galU*, 및 *pyrH*로부터 선택된 하나 이상의 유전자를 효모-기반 재조합을 이용하여 클로닝하고 이어서 숙주 균주에서 발현시켜, 갈락토스의 수용자 올리고

고사카라이드 조성물로의 전이를 위한 글리코실 공여자 기질의 충분한 UDP-Gal 풀을 보장한다.

[0105] CMP-NANA 수준의 조절은 효모 및 곤충 세포 둘 다에서 나타났다. 해밀톤(Hamilton) 등은 CMP-시알산 수송자, UDP-GlcNAc 2-에피머라제/N-아세틸만노사민 키나제, CMP-시알산 신타제, N-아세틸뉴라미네이트-9-포스페이트 신타제, 및 시알릴트랜스퍼라제를 사용한 *P. 파스토리스*에서의 성공적인 시알산 전이를 위한 증가된 세포 CMP-NANA 풀을 제시하였다 ([Hamilton, S.R., et al., Production of complex human glycoproteins in yeast. *Science*, 301, 1244 (2003)]). 로런스(Lawrence) 등은 곤충 세포에서 증가된 CMP-SA 기질 생성을 위한 N-아세틸만노사민 피딩(feeding)을 이용한 시티딘 모노포스페이트 시알산 신타제 (CMP-SA) 및 시알산 포스페이트 신타제 (SAS) 유전자의 공발현을 제시하였다 ([Lawrence et al., Cloning and expression of human sialic acid pathway genes to generate CMP-sialic acids in insect cells. *Glycoconj J*. 2001 Mar;18(3):205-13]). 단지 선택된 몇몇 숙주 세포, 예컨대 E. 콜라이만이 내인성 CMP-NANA 메카니즘을 갖지만, 많은 원핵생물은 CMP-NANA를 생성하기 위한 기구(machinery)를 결여하며, 적어도 증가된 CMP-NANA 수준이 원핵생물에서의 적당한 시알릴화에 필요하다는 것이 예상된다.

[0106] 진핵 단백질, 특히 막 단백질의 E. 콜라이 및 다른 박테리아에서의 성공적인 발현은 중대한 과제이다 ([Baneyx et al., "Recombinant Protein Folding and Misfolding in Escherichia coli," *Nat Biotechnol* 22:1399-1408 (2004)]). 따라서, 올바르게 접히고 올바르게 국재화된 단백질의 높은 발현 수율을 얻기 위해 다수의 문제를 감안해야 한다(예를 들어, 내막으로의 삽입). 이러한 모든 요인이 집합적으로, 진핵 단백질이 E. 콜라이 세포 내부에서 발현될 경우 가능성을 가질지를 결정한다.

[0107] **추가외 당조작**

[0108] 특정 효소 활성을 결여하는 숙주 세포, 예컨대 당 생합성과 경쟁하는 특정 효소 (예를 들어, 만노실트랜스퍼라제)가 발현하지 않거나 감소되는 숙주 세포가 바람직하다. 바람직한 실시양태에서, 상기 방법은 발데라마-린콘(Valderrama-Rincon) 등이 제시한 바와 같이, GDP-D-만노스 데히드라타제 효소 활성 (EC 4.2.1.47)이 약화되는 숙주 세포의 배양을 제공한다. GDP-만노스 데히드라타제 (GMD)를 코딩하는 *gmd* 유전자를 결여하는 E. 콜라이 균주가 구축되며, 이는 GDP-L-푸코스의 합성에서 첫 단계로서 GMD에 의해 GDP-4-케토-6-데옥시만노스로 변환되는 Alg1 및 Alg2를 위한 기질, GDP-만노스의 입수가능성을 사실상 증가시킬 것이다 ([Ruffing, A. & Chen, R.R. Metabolic engineering of microbes for oligosaccharide and polysaccharide synthesis. *Microb Cell Fact* 5, 25 (2006)]). 숙주 세포의 추가의 조작이 특정 경쟁 경로를 녹아웃(knock-out)하는 데에 필요할 수 있다.

[0109] **코돈 최적화**

[0110] 본 발명의 추가의 실시양태에서, 진핵 글리코실트랜스퍼라제를 코돈 최적화하여 E. 콜라이 (및 다른 박테리아) 및 보다 고등한 유기체, 예컨대 효모 및 포유동물 세포 사이의 코돈 용법 바이어스와 연관된 한계를 극복한다. 코돈 용법 바이어스란, 단백질을 코딩하는 DNA 서열 (유전자)에서의 코돈 발생 빈도의 유기체 간 차이를 지칭한다. 코돈은 폴리펩티드 쇠에서 특정 아미노산 잔기를 코딩하는 연속된 3개의 뉴클레오티드 (트리플릿)이다. 코돈 최적화는 특정한 전환(transversion) 뉴클레오티드 변화, 즉, 퓨린에서 피리미딘, 또는 피리미딘에서 퓨린으로의 뉴클레오티드 변화, 또는 이행(transition) 뉴클레오티드 변화, 즉, 퓨린에서 퓨린 또는 피리미딘에서 피리미딘으로의 뉴클레오티드 변화를 만듦으로써 이루어질 수 있다. 일부 경우, 코돈 최적화된 폴리펩티드 변이체는 코돈 최적화되지 않은 폴리펩티드와 동일한 생물학적 기능을 유지한다. E. 콜라이에서의 발현을 위해, 하나 이상의 코돈을 예를 들어 문헌 [Grosjean et al., *Gene* 18:199-209 (1982)]에 기재된 바와 같이 최적화할 수 있다. 본 명세서에 사용된 "*"은 정지 코돈을 나타낸다.

[0111] 본 명세서에 기재된 *alg*, N-아세틸글루코사민일 트랜스퍼라제, 갈락토실트랜스퍼라제, 시알릴트랜스퍼라제, ManB/C, *gImS*, 올리고사카릴 트랜스퍼라제의 핵산 분자, 폴리펩티드 분자 및 동족체, 변이체 및 유도체 또한 폴리뉴클레오티드 및 폴리펩티드 변이체를 포함하며, 이것은 자연 발생하거나 공지된 유전자 조작 기술을 이용한다. 화학적 합성을 포함하여 *시험관내* 생성될 수 있다. 일부 실시양태에서, 폴리뉴클레오티드 서열은 서열 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 또는 29과 적어도 75%, 77%, 80%, 85%, 90% 또는 95% 동일성을 갖는다. 다른 실시양태에서, 폴리펩티드 변이체는 서열 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 31, 32 또는 33에 적어도 약 50%, 55%, 60%, 65%, 70%, 75%, 77%, 80%, 85%, 90% 또는 95% 상동성을 갖는다.

[0112] 본 발명은 또한 엄격한 조건하에서 상술한 핵산 분자에 혼성화하는 핵산 분자를 포함한다. 상기 정의된 바와

같이, 그리고 관련 기술 분야에 널리 공지된 바와 같이, 엄격한 혼성화는 특정한 설정 조건하에서 특정 DNA 하이브리드를 위한 열 용융 온도 (T_m)보다 약 25°C 낮은 온도에서 수행되며, T_m 은 표적 서열의 50%가 완벽하게 매칭된 프로브에 혼성화하는 온도이다. 엄격한 세척은 특정 설정 조건하에서 특정 DNA 하이브리드를 위한 T_m 보다 약 5°C 낮은 온도에서 수행할 수 있다.

[0113] 본 발명의 폴리뉴클레오티드 또는 핵산 분자는 적어도 염기 10개 길이의 뉴클레오티드의 중합체 형태를 지칭한다. 이것은 DNA 분자(예를 들어, 선형, 환형, cDNA, 염색체, 게놈, 또는 합성, 이중 가닥, 단일 가닥, 삼중-가닥, 사중(quadruplexed), 부분 이중-가닥, 분지형, 머리핀형, 환형, 또는 자물쇠(padlock) 형태) 및 RNA 분자(예를 들어, tRNA, rRNA, mRNA, 게놈, 또는 합성) 및 상기한 것들의 DNA 또는 RNA 분자의 유사체 뿐 아니라 비-천연 뉴클레오티드 유사체, 비-천연 인터-뉴클레오티드 결합, 또는 둘 다를 함유하는 DNA 또는 RNA의 유사체를 포함한다. 본 발명의 단리된 핵산 분자는 핵산이 유도되는 유기체의 염색체 DNA에 천연 플랭킹 서열(즉, 핵산 분자의 5' 및 3' 말단에 위치한 서열)이 없는 핵산 분자를 포함한다. 다양한 실시양태에서, 단리된 핵산 분자는 핵산 분자가 유도되는 미생물의 천연 플랭킹 뉴클레오티드 염색체 DNA 서열을 약 10 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, 0.1 kb, 50 bp, 25 bp 또는 10 bp 미만으로 함유할 수 있다.

[0114] 이중 핵산 분자는 올바른 리딩 프레임 및 프로모터 및 임의의 다른 5' 조절 분자에 대해 적당한 센스(5'→3') 배향으로 발현 시스템 또는 벡터에 삽입된다. 핵산 구축물의 제조는, 문헌 [Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Springs Laboratory Press, Cold Springs Harbor, New York (1989)]에 기재된 바와 같은, 관련 기술분야에 널리 공지된 표준 클로닝 방법을 사용하여 실시할 수 있다. 코헨(Cohen) 및 보이어(Boyer)의 미국 특허 제4,237,224호 또한 DNA 리가제를 이용한 라이게이션 및 제한 효소 절단을 이용한 재조합 플라스미드 형태의 발현 시스템의 제조를 기재한다.

[0115] 적합한 발현 벡터는, 숙주 세포와 양립가능한 종으로부터 유도된 레플리콘 및 제어 서열을 함유하는 것을 포함한다. 예를 들어, E. 콜라이를 숙주 세포로 사용하는 경우, pUC19, pUC18 또는 pBR322와 같은 플라스미드를 사용할 수 있다. 다른 적합한 발현 벡터는 문헌 [*Molecular Cloning: a Laboratory Manual: 3rd edition*, Sambrook and Russell, 2001, Cold Spring Harbor Laboratory Press]에 기재되어 있다. 핵산의 조작, 예를 들어, 핵산 구축물의 제조, 돌연변이 유발, 시퀀싱, 세포 내로의 DNA 도입 및 유전자 발현, 및 단백질 분석을 위한 많은 공지 기술 및 프로토콜이 문헌 [*Current Protocols in Molecular Biology*, Ausubel et al. eds., (1992)]에 상세히 기재되어 있다.

[0116] 상이한 유전자 신호 및 처리 이벤트가 많은 수준의 유전자 발현(예를 들어, DNA 전사 및 메신저 RNA ("mRNA") 번역) 및 후속적으로 리보솜 표면 상에 디스플레이되는 융합 단백질의 양을 제어한다. DNA의 전사는 프로모터의 존재에 의존하며, 이것은 RNA 폴리머라제의 결합을 지시함으로써 mRNA 합성을 촉진하는 DNA 서열이다. 프로모터는 그의 "강도"(즉, 전사를 촉진하는 그의 능력)가 다양하다. 클로닝된 유전자를 발현하기 위해서는 고수준의 전사 및 이에 따른 발현 및 표면 디스플레이를 얻기 위해 강한 프로모터를 사용하는 것이 바람직하다. 따라서, 이용되는 숙주 시스템에 따라, 수많은 적합한 프로모터 중 어느 하나를, 스톨(stall) 서열에 커플링된 관심 단백질을 코딩하는 데옥시리보핵산 분자를 지닌 발현 벡터 내에 혼입할 수 있다. 예를 들어, E. 콜라이를 사용할 경우, 그의 박테리오파지, 또는 플라스미드, 프로모터, 예컨대 T7 파지 프로모터, lac 프로모터, trp 프로모터, recA 프로모터, 리보솜 RNA 프로모터, 콜리파지 람다 및 lacUV5, ompF, bla, lpp 등을 포함하지만 이에 한정되지 않는 다른 것들의 P_R 및 P_L 프로모터를 사용하여 인접 DNA 절편의 고수준의 전사를 지시할 수 있다. 추가로, 하이브리드 trp-lacUV5 (tac) 프로모터 또는 재조합 DNA 또는 합성 DNA 기술에 의해 생성된 기타 E. 콜라이 프로모터를 사용하여 삽입된 유전자의 전사를 제공할 수 있다.

[0117] 원핵생물 내의 mRNA의 번역은 적당한 원핵 신호의 존재에 의존하며, 이것은 진핵생물의 것과 상이하다. 원핵생물 내의 mRNA의 효율적인 번역은 샤인-달가노(Shine-Dalgarno)("SD") 서열이라 불리우는, mRNA 상의 리보솜 결합 부위를 필요로 한다. 이 서열은 개시 코돈, 통상 AUG 앞에 위치하는 mRNA의 짧은 뉴클레오티드 서열이며, 이것은 단백질의 아미노-말단 메티오닌을 코딩한다. SD 서열은 16S rRNA(리보솜 RNA)의 3'-말단에 상보적이며, rRNA와 이중화(duplexing)하여 리보솜의 올바른 포지셔닝을 허용함으로써 mRNA의 리보솜에의 결합을 촉진하는 것으로 추정된다. 유전자 발현의 최대화에 대한 개관은 문헌 [Roberts and Lauer, *Methods in Enzymology*, 68:473 (1979)]을 참조하라.

[0118] **숙주 세포**

[0119] 본 발명에 따라서, 숙주 세포는 원핵생물이다. 그러한 세포는 관심 재조합 치료 단백질의 제조를 위한 재조합

단백질의 발현을 위한 숙주로서 기능한다. 예시적인 숙주 세포는 E. 콜라이 및 기타 엔테로박테리아세아에, 에스케리키아 종, 캄필로박터 종, 윌리벨라 종, 데솔포비브리오 종, 비브리오 종, 슈도모나스 종, 바실루스 종, 리스테리아 종, 스타필로코쿠스 종, 스트렙토코쿠스 종, 펩토스트렙토코쿠스 종, 메가스파에라 종, 펙티나투스 종, 셀레노모나스 종, 짐필루스 종, 악티노미세스 종, 아르트로박터 종, 프란키아 종, 마이크로모노스포라 종, 노카르디아 종, 프로피오니박테리움 종, 스트렙토미세스 종, 락토바실루스 종, 락토코쿠스 종, 류코노스톡 종, 페디오코쿠스 종, 아세트박테리움 종, 유박테리움 종, 헬리오박테리움 종, 헬리오스피릴룸 종, 스포로무사 종, 스피로플라스마 종, 우레아플라스마 종, 에리시켈로트릭스 종, 코리네박테리움 종, 엔테로코쿠스 종, 클로스트리디움 종, 미코플라스마 종, 미코박테리움 종, 악티노박테리아 종, 살모넬라 종, 시겔라 종, 모락셀라 종, 헬리코박터 종, 스테노트로포모나스 종, 미크로코쿠스 종, 네이세리아 종, 브델로비브리오 종, 헤모필루스 종, 클레브시엘라 종, 프로테우스 미라빌리스, 엔테로박터 클로아카에, 세라티아 종, 시트로박터 종, 프로테우스 종, 세라티아 종, 예르시니아 종, 아시네토박터 종, 악티노바실루스 종, 보르데텔라 종, 브루셀라 종, 카프노시토파가 종, 카르디오박테리움 종, 에이케넬라 종, 프란시셀라 종, 헤모필루스 종, 킨겔라 종, 파스테우렐라 종, 플라보박테리움 종, 크산토모나스 종, 부르크홀데리아 종, 아에로모나스 종, 플레시오모나스 종, 레지오넬라 종 및 알파-프로테오박테리아, 예컨대 울바키아 종, 시아노박테리아, 스피로카에테스, 녹색 황 및 녹색 비-황 박테리아, 그람-음성 구균, 까다로운(fastidious) 그람 음성 바실루스, 엔테로박테리아세아에-글루코스-발효 그람-음성 바실루스, 그람 음성 바실루스 - 비-글루코스 발효기, 그람 음성 바실루스 - 글루코스 발효, 옥시다제 양성을 포함한다.

[0120] 본 발명의 한 실시양태에서, E. 콜라이 숙주 균주 C41(DE3)를 사용하며, 이는 이 균주가 종전에 일반적 막 단백질 과다발현에 최적화되었기 때문이다 ([Miroux et al., "Over-production of Proteins in Escherichia coli: Mutant Hosts That Allow Synthesis of Some Membrane Proteins and Globular Proteins at High Levels," *J Mol Biol* 260:289-298 (1996)]). 숙주 균주의 추가의 최적화는 DnaJ 단백질 (예를 들어, $\Delta dnaJ$ 세포)을 코딩하는 유전자의 결실을 포함한다. 이 결실의 이유는 dnaJ의 불활성화가 과다발현된 막 단백질의 축적을 증가시키고 막 단백질 과다발현과 흔히 연관되는 심한 세포독성을 억제하는 것으로 공지되어 있기 때문이다 ([Skretas et al., "Genetic Analysis of G Protein-coupled Receptor Expression in Escherichia coli: Inhibitory Role of DnaJ on the Membrane Integration of the Human Central Cannabinoid Receptor," *Biotechnol Bioeng* (2008)]). 본 출원인은 Alg1 및 Alg2의 발현 후 이를 관찰하였다. 또한, 경쟁하는 당 생합성 반응의 결실이 최적 수준의 N-글리칸 생합성을 보장하기 위해 필요하다. 예를 들어, E. 콜라이 O16 항원 생합성 경로에서의 유전자 결실 ([Feldman et al., "The Activity of a Putative Polyisoprenol-linked Sugar Translocase (Wzx) Involved in Escherichia coli O Antigen Assembly is Independent of the Chemical Structure of the O Repeat," *J Biol Chem* 274:35129-35138 (1999)])은 박테로프레놀-GlcNAc-PP 기질이 원하는 포유동물 N-글리칸 반응에 이용가능한 것을 보장할 것이다. 원치 않는 부반응을 제거하기 위해서, E. 콜라이에서 결실되는 대표적인 유전자는 다음과 같다: *wbbL*, *gltC*, *gltF*, *gltT*, *wzx*, *wzy*, *waaL*. 또 다른 균주는 MC4100, BL21, 오리가미 (ORIGAMI)TM, 셔플(Shuffle)[®]을 포함한다.

[0121] 숙주 세포를 발현 벡터로 형질전환/형질감염하는 방법은 관련 기술분야에 널리 공지되어 있으며, 문헌 [Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Springs Laboratory Press, Cold Springs Harbor, New York (1989)]에 기재된 바와 같이 선택된 숙주 시스템에 의존한다. 진행 세포의 경우, 적합한 기술은 인산칼슘 형질감염, DEAE-텍스트란, 전기천공, 리포솜-매개 형질감염, 및 레트로바이러스 또는 다른 바이러스, 예를 들어 백시니아, 또는 곤충 세포의 경우 바클로바이러스를 이용한 형질도입을 포함할 수 있다. 박테리아 세포의 경우, 적합한 기술은 염화칼슘 형질전환, 전기천공, 및 박테리오파지를 이용한 형질감염을 포함할 수 있다.

[0122] 본 발명의 원핵 숙주 세포의 주요 이점은 (i) 박테리아의 유전자 조작에 관한 방대한 양의 데이터; (ii) 단백질 생성을 위한 박테리아 이용에 관한 확립된 실적, 2003년 이래 FDA에 의해 승인된 단백질 치료제의 약 30%는 E. 콜라이 박테리아에서 생성; 및 (iii) 단백질 약물의 박테리아 제조를 위한 많은 회사 내의 기존 기반시설을 포함한다.

[0123] 다양한 진행 단백질 발현 시스템과 비교해, 본 발명의 방법 및 조성물을 사용하는 공정은 확장가능하고 (scalable), 비용효과적이며, 인간 병원체가 없고, 면역원성 N- 및 O-연결된 글리코실화 반응이 없으며, 신속한 클로닝 및 빠른 성장 속도, 빠른 배가 시간 (약 20분), 높은 성장 (높은 OD), 높은 역치 및 단백질 수율 (총 가용성 단백질 (TSP)의 50% 범위), 주변세포질 또는 상청액으로부터의 용이한 생성물 정제가 가능하며, 유전학적으로 추적가능하고, 철저히 연구되고, 광범위한 발현 최적화 방법 (예를 들어, 프로모터 조작, mRNA 안정화 방

법, 사폐론 공발현, 프로테아제 고갈 등)과 양립가능한 최적의 재조합 당단백질 발현을 제공한다.

[0124] 당단백질 발현을 위한 숙주로서 원핵생물, 예를 들어 E. 콜라이의 또 다른 주요 장점은 효모 및 모든 다른 진핵 생물과 달리 고유 글리코실화 시스템이 없다는 것이다. 따라서, 글리코실화-관련 유전자의 첨가 (또는 후속 제거)는 당조작된 E. 콜라이 세포의 생존성에 거의 내지 전혀 영향을 미치지 않을 것으로 예상된다. 또한, 내인성 글리코실화 반응에 의한 표적 단백질에의 비-인간 글리칸 부착의 가능성은 이들 세포에서 본질적으로 제거된다.

[0125] 따라서, 다양한 실시양태에서, 당단백질 발현 및 다양한 올리고사카라이드 조성물(예를 들어, 고-만노스, 하이브리드 복합체)의 제조를 위한 대안으로서, 원핵 숙주 세포를 사용하여 이를 제조하고 N-연결된 당단백질을 생성하는 방법이 개시되며, 이는 진핵 세포 배양과 연관된 상당한 장애물을 피하기 위한 매력적인 해결책을 제공한다. 구조적으로 균질한 인간-유사 N-글리칸을 생성하는 동시에 단백질 약물 개발 및 제조와 연관된 비용 및 시간을 극적으로 낮추는 제조 비히클로서 박테리아의 사용이 본 발명의 목적이다.

[0126] **원핵생물에서 올리고사카라이드의 표적 단백질로의 부위-특이적 전이**

[0127] 발테라마-린콘 등의 문헌에 기재된 바와 같이, 박테리아 글리코실화 기구의 "인간화"를 시작하기 위해, E. 콜라이에서 지질-연결된 생합성을 포함하는 재조합 경로를 통해 Man₃GlcNAc₂ 올리고사카라이드 구조를 생성한다. 특히, 몇몇 진핵 글리코실트랜스퍼라제 중 하나를 E. 콜라이에서 기능적으로 발현시키고, 얻어진 지질-연결된 올리고사카라이드를 올리고사카릴 트랜스퍼라제를 통해 단백질로 전이시킨다.

[0128] 원핵 숙주 세포 내의 글리칸 어셈블리는, 그것이 돌리콜 포스페이트 (DoI-P) 상에 조립되는 진핵생물과 달리 운데카프레닐 포스페이트 (Und-P) 상에 지질-연결된다. C. 제주니에서, N-연결된 글리코실화는 뉴클레오티드-활성화된 당의 지질 담체로의 순차 부가를 통해 진행되어 분지형 헵타사카라이드를 형성한다. 이 글리칸은 이후 PglK (이전의 WlaB)에 의해 내막을 가로질러 플립되고, 이후 OTase PglB가 글리칸의 아스파라긴 측쇄로의 전이를 촉매한다. Bac는 2,4-디아세트아미도-2,4,6-트리테옥시글루코스이고; GalNAc는 N-아세틸갈락토사민이고; HexNAc는 N-아세틸헥소사민이고; Glc는 글루코스이다. 문헌 [Szymanski et al., "Protein Glycosylation in Bacterial Mucosal Pathogens," *Nat Rev Microbiol* 3:225-37 (2005)]을 참조하라. PglK 플립퍼제는 지질-연결된 C. 제주니 헵타사카라이드를 내막을 가로질러 이동시키는 역할을 한다. 우연하게도, PglK는 지질-연결된 올리고사카라이드 중간체의 글리칸 구조에 대해 완화된 특이성을 나타낸다 ([Alaimo et al., "Two Distinct But Interchangeable Mechanisms for Flipping of Lipid-linked Oligosaccharides," *Embo J* 25:967-76 (2006)] 및 [Wacker et al., "Substrate Specificity of Bacterial Oligosaccharyltransferase Suggests a Common Transfer Mechanism for the Bacterial and Eukaryotic Systems," *Proc Natl Acad Sci U S A* 103:7088-93 (2006)]).

[0129] 바람직한 실시양태에서, 본 발명의 숙주 세포는 플립퍼제 효소 활성 (젠뱅크 AN AP009048.1)을 발현하며, 이는 운데카프레놀-연결된 올리고사카라이드를 내막을 가로질러 이동시킨다. 그러한 효소 활성은 내인성 또는 이종이거나 또는 발현 중에 변형되도록 조작될 수 있다. 추가의 실시양태에서, 원핵 숙주 세포는 *pglK* 및 *rft1*을 포함한 플립퍼제 활성을 포함한다.

[0130] 원핵생물에서 인간-유사 올리고사카라이드 구조의 생성은 다양한 올리고사카라이드의 폴리펩티드 쇠 상의 N-X-S/T 부위의 전이를 수반한다. 이것은 올리고사카라이드의 표적 단백질로의 전이를 책임지는 올리고사카릴트랜스퍼라제 (OST)로서 공지된 단백질 복합체 또는 내재성 막 단백질의 기능적 발현을 필요로 한다. 다양한 원핵 및 진핵 OST가 지질-연결된 올리고사카라이드를 표적 단백질로 전이시키는 능력을 갖는다. 본 발명은 고-만노스, 하이브리드 및 복합 글리칸의 단백질로의 전이를 나타내는 원핵 시스템을 개시한다. 따라서, 원핵 단백질 발현 시스템은 적어도 하나의 OST 활성을 포함하여 글리코실화 표적 단백질을 생성한다. 그러한 실시양태에서, 숙주 세포는 글리코실트랜스퍼라제 효소 이외에도 올리고사카릴 트랜스퍼라제 효소 활성 (EC 2.4.1.119)을 발현한다. 다양한 OST (표 2)가 발현될 수 있으며, 내인성 또는 이종이거나 또는 발현 중에 변형되도록 조작될 수 있다. 추가의 실시양태에서, 원핵 숙주 세포는 적어도 하나의 올리고사카릴 트랜스퍼라제 활성, 예컨대 C. 제주니로부터의 PglB (아에비 등) 또는 C. 라리 (발테라마-린콘 등)를 포함한다. 단백질로 전이되는 올리고사카라이드는 단백질에 N-연결된다.

표 2

[0131]

올리고사카릴트랜스퍼라제의 목록

단백질 EC #	유기체	젠 बैं크
CCC13826_0460	캄필로박터 콘시수스 13826	EAT99324.2
CFE8240_1383	캄필로박터 페투스 아종 페투스 82-40	ABK82109.1
CHAB381_0954	캄필로박터 호미니스 ATCC BAA-381	ABS52339.1
OrfA (단편)	캄필로박터 제주니 NCTC 11351	AAD09300.1
W1aF	캄필로박터 제주니 81116	CAA72355.1 ABV52665.1
W1aF	캄필로박터 제주니 D450	AAK97437.1
CJE1268	캄필로박터 제주니 RM1221	AAW35590.1
JJD26997_0595	캄필로박터 제주니 아종 도일레이 269.97	ABS43894.1
OST (PglB;W1aF) EC 2.4.1.119	캄필로박터 제주니 아종 제주니 81-176	AAK97438.1 AAD51383.1
OST (PglB;W1aF;Cj1126c) EC 2.4.1.119	캄필로박터 제주니 아종 제주니 NCTC 11168	CAB73381.1 NP_282274.1 CAL35243.1 AAD09293.1
C1a_1253 (PglB)	캄필로박터 라리 RM2100 RM2100; ATCC BAA-1060D	ACM64573.1
Ddes_0746	데슬포비브리오 데슬푸리칸스 아종 데슬푸리칸스 균주 ATCC 27774	ACL48654.1
Dde_3699	데슬포비브리오 데슬푸리칸스 아종 데슬푸리칸스 균주 G20	ABB40492.1
DvMF_0846	데슬포비브리오 불가리스 균주 '미야자키 F'	ACL07802.1
Dvu1_1810	데슬포비브리오 불가리스 DP4	ABM28827.1
DVU1252	데슬포비브리오 불가리스 균주 힐텐보로우	AAS95730.1
Geob_1424 Geob_2990	게오박터 종 FRC-32	ACM19784.1
NAMH_1652	나우틸리아 프로폰디콜라 AmH	ACM92784.1
NIS_1250	니트라티류프토르 종 SB155-2	BAF70358.1
Tmden_1474	슈푸리모나스 데니트리피칸스 DSM 1251	ABB44751.1
SUN_0103	슈푸로볼 종 NBC37-1	BAF71063.1
WSO043 (W1aF)	윌리넬라 속시노게네스 DSM 1740	CAE09214.1 NP_906314.1
OST, STT3 서브유닛	캄필로박테탈레스 박테리움 GD 1	EDZ62411.1
BACPLE_02950	박테로이데스 플레베이우스 DSM 17135	EDY94544.1
BACPLE_02943	박테로이데스 플레베이우스 DSM 17135	EDY94539.1
RHECIAT_CH0002772	리조비움 에틀리 CIAT 652	ACE91723.1
BACINT_01142	박테로이데스 인테스티날리스 DSM 17393	EDV06057.1
IMP (가능한 OST)	히드로게니비르가 종 128-5-R1-1	EDP74595.1
OST (PglB)	캄필로박터 콜리 RM2228	EAL57053.1
OST (PglB)	캄필로박터 읍살리엔시스 RM3195	EAL53100.1

[0132]

올리고사카라이드 조성물

[0133]

최근, 몇몇 진핵 발현 숙주가 N-당단백질 제조를 위한 포유동물 세포 배양의 대안으로서 도입되었다. 이는 유전자 조작된 효모 *피키아 파스토리스* ([Hamilton, S.R., et al., *Humanization of yeast to produce complex terminally sialylated glycoproteins*. Science, 2006. **313**(5792): p. 1441-3]), 재조합 바칼로바이러스를 위한 숙주로서의 배양된 곤충 세포 ([Aumiller, J.J., J.R. Hollister, and D.L. Jarvis, *A transgenic insect cell line engineered to produce CMP-sialic acid and sialylated glycoproteins*. Glycobiology, 2003. **13**(6): p. 497-507]), 및 식물 세포 ([Aviezer, D., et al., *A plant-derived recombinant human glucocerebrosidase enzyme--a preclinical and phase I investigation*. PLoS One, 2009. **4**(3): p. e4792])를

포함한다. 불운하게도, 포유동물, 식물, 곤충 및 효모 세포를 포함한 임의의 진핵 숙주 세포를 사용할 경우 비인간 당형태가 고유 글리코실화 경로로부터 발생한다. 포유동물 숙주 세포는 제어불가능한 수준의 만노스-6-포스페이트 및 푸코스를 글리칸에 부가시키며 종종 말단 시알산을 결여하는 것으로 나타났다 ([Van Patten, S.M., et al., *Effect of mannos chain length on targeting of glucocerebrosidase for enzyme replacement therapy of Gaucher disease*. *Glycobiology*, 2007. **17**(5): p. 467-78]). 식물 세포는 면역원성 베타-1,2 크실로스 및 코어 알파-1,3 푸코스를 부가하며 ([Bardor, M., et al., *Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha(1,3)-fucose and core xylose*. *Glycobiology*, 2003. **13**(6): p. 427-34]), 후자는 또한 곤충 세포에서도 발견된다 ([Bencurova, M., et al., *Specificity of IgG and IgE antibodies against plant and insect glycoprotein glycans determined with artificial glycoforms of human transferrin*. *Glycobiology*, 2004. **14**(5): p. 457-66]). O-연결된 글리코실화는 또한 효모에서 필수 과정이며 ([Gentsch, M. and W. Tanner, *The PMT gene family: protein O-glycosylation in Saccharomyces cerevisiae is vital*. *Embo J*, 1996. **15**(21): p. 5752-9]), 원치 않은 O-글리칸은 표적 당단백질에 공유 부착될 수 있다.

[0134] 원핵 글리코실화 기구에 의해 부착된 올리고사카라이드 체는 더 고등한 진핵 및 인간 글리코실화 경로에 의해 부착된 것과는 구조적으로 구분된다 ([Weerapana et al., "Asparagine-linked Protein Glycosylation: From Eukaryotic to Prokaryotic Systems," *Glycobiology* 16:91R-101R (2006)]. 원핵생물에서 및 본 발명의 방법으로부터 생성된 올리고사카라이드 조성물은 또한 효모, 곤충, 포유동물 및 심지어 인간 세포와 같은 진핵 시스템과 구별된다.

[0135] 몇몇 특징은 본 발명의 방법에 의해 생성된 올리고사카라이드 조성물을 진핵 숙주 세포 발현 시스템, 예를 들어, CHO, NS0, 램나(1emna), Sf9와 비교하여 구분한다. 예를 들어, 본 발명의 올리고사카라이드 조성물은 푸코스를 결여한다. 항체 내 푸코스의 부재는 증가된 ADCC 및 CDC 활성과 연관된다 ([Shinkawa T et al., The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J Bio Chem*, 278, 3466-73, 2003])). 또한, 원핵생물은 본래 O-연결된 글리칸을 결여하며, 이는 면역원성과 연관된다. 본 발명의 올리고사카라이드 조성물은 고-만노스 또는 만노스 포스페이트와 같은 많은 진핵 발현 시스템에 존재하는 혐오(abhorrent) 글리칸을 발현하지 않는다. 또한, 당조작된 E. 콜라이는 (i) N- 또는 C-말단에 포함된 글리코실화의 특이 부위 및 화학량론의 제어, (ii) 당형태의 선택, (iii) 글리코실화는 E. 콜라이에서 필수 과정이 아니기 때문에 신규 당형태를 조작할 수 있는 능력, 및 (iv) 생성물 균일성을 향상시키고 만노스 6-포스페이트 수용체와 같은 인간 숙주 내 다른 수용체에의 국제화오류의 방지를 도울 수 있는, O-글리코실화 및 만노스 6-포스페이트를 포함하는 경쟁 글리코실화 경로의 결여를 제공한다 ([Hayette, M.P. et al. Presence of human antibodies reacting with Candida albicans O-linked oligomannosides revealed by using an enzyme-linked immunosorbent assay and neoglycolipids. *J Clin Microbiol* **30**, 411-417 (1992)] 및 [Podzorski, R.P., Gray, G.R. & Nelson, R.D. Different effects of native Candida albicans mannan and mannan-derived oligosaccharides on antigen-stimulated lymphoproliferation in vitro. *J Immunol* **144**, 707-716 (1990)]).

[0136] 본 발명의 올리고사카라이드 조성물은 균일하고 또한 항염증 특성을 증가시키도록 강화(enrich)될 수 있으며, 예를 들어, 정맥내 Ig (IVIG)의 Fc 상의 α2,6 시알산이 강화될 수 있다 ([Anthony et al., Identification of a receptor required for the anti-inflammatory activity of IVIG. *Natl Acad Sci USA* 2008 Dec 16;105(50):19571-8]). 추가의 연구는 모든 인간 내 Neu5Gc-특이적 항체의 존재 (때로는, 고수준으로 존재)를 나타냈다 ([Ghaderi et al., Implications of the presence of N-glycolylneuraminic acid in recombinant therapeutic glycoproteins. *Nat Biotechnol*, 2010 Aug;28(8): 863-7]). 따라서, 치료 단백질, 예를 들어, 특정 시알산 잔기를 갖는 항체의 강화 (예를 들어, Neu5Ac, Neu5Gc와는 대조적으로 NeuNAc)는 단백질 치료제의 면역원성 또는 무효과와 같은 유해 반응을 감소시킬 수 있다.

[0137] 본 명세서에 반영된 바와 같이, 원핵 시스템은 균질 글리칸을 비교적 높은 수율로 생성할 수 있다. 바람직한 실시양태에서, 올리고사카라이드 조성물은 본질적으로 적어도 50, 60, 70, 80, 90, 95, 99 몰%의 단일 당형태로 이루어진다. 추가의 실시양태에서, 올리고사카라이드 조성물은 본질적으로 적어도 50, 60, 70, 80, 90, 95, 99 몰%의 2개의 바람직한 당형태로 이루어진다. 추가의 실시양태에서, 올리고사카라이드 조성물은 본질적으로 적어도 50, 60, 70, 80, 90, 95, 99 몰%의 3개의 바람직한 당형태로 이루어진다.

- [0138] 특정 실시양태에서, 생성된 올리고사카라이드 조성물은 $\text{GlcNAc}_{1-5}\text{Man}_3\text{GlcNAc}_2$ 및 $\text{Man}_3\text{GlcNAc}_2$ 이다. 특정 글리콜-조작된 숙주 세포는 주로 $\text{GlcNAcMan}_3\text{GlcNAc}_2$ 또는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 인 올리고사카라이드 조성물을 생성한다.
- [0139] 다른 실시양태에서, 생성된 올리고사카라이드 조성물은 $\text{Gal}_{1-5}\text{GlcNAc}_{1-5}\text{Man}_3\text{GlcNAc}_2$ 및 $\text{Man}_3\text{GlcNAc}_2$ 이다. 특정 글리콜-조작된 숙주 세포는 주로 $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$, $\text{GalGlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 또는 $\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 인 올리고사카라이드 조성물을 생성한다.
- [0140] 또 다른 실시양태에서, 생성된 올리고사카라이드 조성물은 $\text{NANA}_{1-5}\text{Gal}_{1-5}\text{GlcNAc}_{1-5}\text{Man}_3\text{GlcNAc}_2$ 이다. 특정 글리콜-조작된 숙주 세포는 주로 $\text{NANAGalGlcNAcMan}_3\text{GlcNAc}_2$ 또는 $\text{NANA}_2\text{Gal}_2\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 인 올리고사카라이드 조성물을 생성한다.
- [0141] 또 다른 실시양태에서, 생성된 올리고사카라이드 조성물은 $\text{Man}_2\text{GlcNAc}_2$, $\text{Man}_4\text{GlcNAc}$, $\text{Man}_3\text{GlcNAc}_2$, $\text{HexMan}_3\text{GlcNAc}_2$, $\text{HexMan}_5\text{GlcNAc}$, $\text{Man}_6\text{GlcNAc}$ 및 $\text{Man}_5\text{GlcNAc}_2$ 이다. 특정 글리콜-조작된 숙주 세포는 주로 $\text{Man}_5\text{GlcNAc}_2$ 인 올리고사카라이드 조성물을 생성한다.
- [0142] 따라서, 본 발명은 신규 올리고사카라이드 조성물 및 *N*-연결된 당단백질의 방대한 어레이의 입체특이적 생합성을 제공한다. 특정 실시양태에서, 대사 경로 및 단백질 조작 기술을 이용한 E. 콜라이에서의 진행 *N*-글리코실화 경로의 재구성은 구조적으로 균질한 인간-유사 글리칸을 갖는 *N*-당단백질을 생성한다. 이는 각각의 당조작된 세포주가 독특한 탄수화물 시그니처(signature)에 상응하는 것을 보장한다.
- [0143] 글리칸은 세포를 ^3H -GlcNAc 및 ^3H -만노스로 또는 형광 렉틴 (예를 들어, AlexaFluor-ConA)으로 대사 표지화하여 분석할 수 있다. 글리칸은 또한 PNGase와 함께 방출되고 MALDI/TOF-MS에서 검출될 수 있다.
- [0144] 글리칸의 정량화는 MS로 추정하거나, 또는 더욱 정확하게는 HPLC를 통해 이루어질 수 있다. NMR은 글리칸 구조의 글리코시드 연결을 측정할 수 있다.
- [0145] **표적 당단백질**
- [0146] 다양한 관심 당단백질을 제조하기 위해, 표적 단백질을 코딩하는 유전자를 숙주 세포에 도입한다.
- [0147] "표적 단백질", "관심 단백질" 또는 "치료 단백질"은 시토카인, 예를 들어, 인터페론, G-CSF, 응고 인자, 예컨대 인자 VIII, 인자 IX, 및 인간 단백질 C, 가용성 IgE 수용체 α -쇄, IgG, IgG 단편, IgM, 인터류킨, 우로키나제, 키마제, 및 우레아 트립신 억제제, IGF-결합 단백질, 표피 성장 인자, 성장 호르몬-방출 인자, 아벡신 V 융합 단백질, 안지오테닌, 혈관 내피 성장 인자-2, 미엘로이드 프로게니터 억제 인자-1, 오스테오프로테게린, α -1 항트립신, DNase II, α -페토(feto) 단백질, AAT, rhTBP-1 (일명 TNF 결합 단백질 1), TACI-Ig (막형단 활성화제 및 칼슘 조절제 및 시클로필린 리간드 인터랙터(interactor)), FSH (여포 자극 호르몬), GM-CSF, 글루카곤, 글루카곤 펩티드, GLP-1 w/ 및 w/o FC (글루카곤 유사 단백질 1) IL-1 수용체 효능제, sTNFr (일명 가용성 TNF 수용체 Fc 융합), CTLA4-Ig (세포독성 T 림프구 결합 항원 4-Ig), 수용체, 호르몬, 예컨대 인간 성장 호르몬, 에리트르포이에틴, 펩티드, 스테이플드(stapled) 펩티드, 인간 백신, 동물 백신, 혈청 알부민 및 효소 예컨대 ATIII, 재조합 인간 트롬빈(rhThrombin), 글루코세레브로시다제 및 아스파라기나제를 제한 없이 포함한다.
- [0148] 이미 승인된 E. 콜라이로부터의 치료제 또한 표적 단백질이다. 이는 호르몬 (인간 인슐린 및 인슐린 유사체, 칼시토닌, 부갑상선 호르몬, 인간 성장 호르몬, 글루카곤, 소마트로핀 및 인슐린 성장 인자 1), 인터페론 (α 1, α 2a, α 2b 및 γ 1b), 인터류킨 2 및 11, 중앙 피사 인자 α , 혈관 내피 성장 인자- α 에 대해 상승된 경쇄 및 중쇄, 콜레라 B 서브유닛 단백질, B-유형 나트륨이노 펩티드, 과립구 콜로니 자극 인자 및 조직 플라스미노겐 활성화제를 포함한다.
- [0149] 표적 단백질은 또한 단백질, 및 상기 단백질에 융합된 D-X₁-N-X₂-T 모티프 (여기에서 D는 아스파르트산이고, X₁ 및 X₂는 프롤린 이외의 임의의 아미노산이고, N은 아스파라긴이고, T는 트레오닌임)를 포함하는 적어도 하나의 펩티드를 포함하는 당단백질 접합체를 포함한다.
- [0150] 바람직한 실시양태에서, 적어도 30, 50, 70, 90, 95 및 바람직하게는 100 몰%의 글리칸이 OST에 의해 표적 단백질로 전이된다.
- [0151] **배양 조건**

- [0152] 다른 실시양태에서, 상기 방법은 산화 조건하에서의 숙주 세포의 배양을 제공한다. 바람직하게는, 산화성 박테리아 균주를 사용한다. 배양 조건은 당단백질 및 글리칸의 증가된 수율 및 역치를 생성할 수 있다. 그러한 방법 조건 및 파라미터는 조절 pH, 온도, 오스몰랄농도, 배양 기간, 배지, 영양분, 용존 산소 농도, 질소, 뉴클레오티드 당의 수준 또는 입수가능성을 포함하며, 심지어 탄소원, 예를 들어, 글리세롤 (**도 9B**)도 생성 시스템에 영향을 줄 수 있다. 배양 조건은 생성물 및 사용되는 특정 숙주 세포에 따라 변할 수 있다. 시스템의 생산성 또한 배양 조건에 의해 영향을 받을 것이다. 추가의 대사 조작이 최대 생산성 및 성장-억제 대사물의 제한을 위해 필요할 수 있다.
- [0153] **올리고사카라이드의 효소 합성**
- [0154] 본 발명의 대안적 측면에서, 글리칸은 **실시예 7**에 기재된 바와 같이 수용자 글리칸, 정제된 효소/용해물을 사용하고 뉴클레오티드 당을 첨가하여 세포가 없는 추출물에서 합성된다.
- [0155] 특정 실시양태에서, 본 발명은 말단 GlcNAc 잔기를 갖는 올리고사카라이드 조성물의 제조에 효과적인 조건하에서 배양된, 재조합 원핵생물, UDP-GlcNAc 및 GnT (EC 2.4.1.101; EC 2.4.1.143; EC 2.4.1.145)를 포함하며 상기 GnT는 UDP-GlcNAc 잔기의 상기 말단 만노스 잔기로의 전이를 촉매하는 세포 배양물을 제공한다.
- [0156] 추가의 실시양태에서, 본 발명은 말단 갈락토스 잔기를 갖는 올리고사카라이드 조성물의 제조에 효과적인 조건하에서 배양된, 재조합 원핵생물, UDP-갈락토스 및 GalT (EC 2.4.1.38)를 포함하며 상기 GalT는 UDP-갈락토스 잔기의 상기 말단 GlcNAc 잔기로의 전이를 촉매하는 세포 배양물을 제공한다.
- [0157] 바람직한 실시양태에서, 본 발명은 말단 시알산 잔기를 갖는 올리고사카라이드 조성물의 제조에 효과적인 조건하에 배양된, 재조합 원핵생물, CMP-NANA 및 시알릴트랜스퍼라제 (EC 2.4.99.4 및 EC 2.4.99.1)를 포함하며 상기 시알릴트랜스퍼라제는 CMP-NANA 잔기의 상기 말단 갈락토스 잔기로의 전이를 촉매하는 세포 배양물을 제공한다.
- [0158] **아글리코실화 대 글리코실화 IgG**
- [0159] 본 발명의 또 다른 측면은 고유 항원을 인식하고 그에 결합하는 Fv 부분 및 보존 아스파라긴 잔기에서 글리코실화되는 Fc 부분을 포함하는 글리코실화 항체에 관한 것이다. 대안적 실시양태는 디아바디, scFv, scFv-Fc, scFv-CH, Fab 및 scFab를 포함한다.
- [0160] 본 발명의 글리코실화 항체는 모노클로날 또는 폴리클로날 항체의 형태일 수 있다.
- [0161] 단일 이뮤노글로불린 분자는 2개의 동일한 경쇄 (L) 및 2개의 동일한 중쇄 (H)를 포함한다. 경쇄는 1개의 불변 도메인 (C_L) 및 1개의 가변 도메인 (V_L)으로 구성되고, 중쇄는 3개의 불변 도메인 (C_{H1}, C_{H2} and C_{H3}) 및 1개의 가변 도메인 (V_H)으로 이루어진다. V_H 및 V_L 도메인은 함께 Fv로서 공지된 분자의 항원-결합 부분을 구성한다. Fc 부분은 보존 Asn297 잔기에서 글리코실화된다. 이 위치에서의 N-글리칸의 부착은 이펙터 상호작용에 필수적인 "개방" 입체형태를 생성한다.
- [0162] 모노클로날 항체는 카빌리(Cabilly) 등의 미국 특허 제4,816,567호 및 문헌 [Anderson et al., "Production Technologies for Monoclonal Antibodies and their Fragments," *Curr Opin Biotechnol.* 15:456-62 (2004)]에 기재된 바와 같이 재조합 DNA 방법을 사용하여 생성될 수 있다. 모노클로날 항체를 코딩하는 폴리뉴클레오티드는 예컨대 성숙 B-세포 또는 하이브리도마 세포로부터, 예컨대 항체의 중쇄 및 경쇄를 코딩하는 유전자를 특이적으로 증폭시키는 올리고뉴클레오티드 프라이머를 사용하는 RT-PCR에 의해 단리되고, 그의 서열은 통상의 절차를 사용하여 결정된다. 중쇄 및 경쇄를 코딩하는 단리된 폴리뉴클레오티드는 이어서 적합한 발현 벡터로 클로닝되고, 이어서 본 발명의 숙주 세포를 형질감염시키고, 모노클로날 항체가 생성된다. 한 실시양태에서, 중쇄 및 경쇄를 N-말단 엑스포트 신호 펩티드 (예를 들어, PelB 신호 펩티드)로 개질하여 중쇄 및 경쇄 폴리펩티드를 박테리아 주변세포질로 향하게 하는데 재조합 DNA 기술이 사용된다. 또한, 중쇄 및 경쇄는 비시스트론 구축물 (예를 들어, 번역되어 2개의 폴리펩티드가 수득되는 단일 mRNA)로부터, 또는 대안적으로 2 시스트론 시스템 (예를 들어, 중쇄 및 경쇄 각각에 대해 2개의 별도의 mRNA가 생성됨)으로부터 발현될 수 있다. 박테리아 주변세포질에서 전장 IgGs의 고-수준의 발현 및 효율적인 여셀블리를 달성하기 위해서, 비시스트론 구축물 및 2 시스트론 구축물 둘 모두를 조작하여 양호한 발현 비를 달성할 수 있다. 예를 들어, 발현 벡터에서 비시스트론 구축물 및 2 시스트론 구축물의 바로 상류에 삽입되는 일련의 번역 개시 영역 (TIRs)을 사용하여 번역 수준이 높아지거나 또는 낮아질 수 있다 (문헌 [Simmons et al., "Translational Level is a Critical Factor for the Secretion of Heterologous Proteins in Escherichia coli," *Nat Biotechnol* 14:629-34 (1996)]). 이 항체

생산 플라스미드가 글리코실화 효소를 발현시키기 위한 플라스미드- 또는 게놈-코딩된 유전자를 또한 번식시키는 박테리아 숙주 내에 도입될 때, 생성된 항체는 주변세포질에서 글리코실화된다. 목적하는 종의 재조합 모노클로날 항체 또는 그의 단편은 또한 (문헌 [McCafferty et al., "Phage Antibodies: Filamentous Phage Displaying Antibody Variable Domains," *Nature* 348:552-554 (1990)]; [Clackson et al., "Making Antibody Fragments using Phage Display Libraries," *Nature* 352:624-628 (1991)]; 및 [Marks et al., "By-Passing Immunization. Human Antibodies from V-Gene Libraries Displayed on Phage," *J. Mol. Biol.* 222:581-597 (1991)])에 기재된 바와 같은 파지 디스플레이 라이브러리로부터 단리될 수 있다.

[0163] 모노클로날 항체를 코딩하는 폴리뉴클레오티드(들)는 재조합 DNA 기술을 사용하는 다수의 상이한 방식으로 추가로 개질되어 대안적 항체를 생성할 수 있다. 한 실시양태에서, 예를 들어, 마우스 모노클로날 항체의 경쇄 및 중쇄의 불변 도메인이 인간 항체의 이들 영역에 치환되어 키메라 항체를 생성할 수 있다. 대안적으로, 마우스 모노클로날 항체의 경쇄 및 중쇄의 불변 도메인은 비-이뮤노글로불린 폴리펩티드에 치환되어 융합 항체를 생성할 수 있다. 다른 실시양태에서, 불변 영역은 말단절단되거나 또는 제거되어 모노클로날 항체의 목적하는 항체 단편이 생성된다. 게다가, 모노클로날 항체의 특이성 및 친화도를 최적화하는 가변 영역의 부위-지정 또는 고-밀도 돌연변이유발이 사용될 수 있다.

[0164] 일부 실시양태에서, 본 발명의 항체는 인간화 항체이다. 인간화 항체는 가변 영역 내에 비-인간 (예를 들어, 뮤린) 항체로부터 최소량의 서열을 함유하는 항체이다. 그러한 항체는 인간 대상체에 투여될 때 항원성 및 인간 항-마우스 항체 반응을 감소시키기 위해 치료상 사용된다. 실제로, 인간화 항체는 전형적으로 비-인간 서열이 최소량 내지 부재하는 인간 항체이다. 인간 항체는 인간에 의해 생성된 항체 또는 인간에 의해 생성된 항체에 상응하는 아미노산 서열을 갖는 항체이다.

[0165] 인간화 항체는 관련 기술분야에 공지된 다양한 기술을 사용하여 생성될 수 있다. 항체는 인간 항체의 상보성 결정 영역 (CDR)을 목적하는 특이성, 친화도, 및 능력을 갖는 비-인간 항체 (예를 들어, 마우스, 래트, 토끼, 햄스터 등)의 것으로 치환함으로써 인간화될 수 있다 (문헌 [Jones et al., "Replacing the Complementarity-Determining Regions in a Human Antibody With Those From a Mouse," *Nature* 321:522-525 (1986)]; [Riechmann et al., "Reshaping Human Antibodies for Therapy," *Nature* 332:323-327 (1988)]; [Verhoeyen et al., "Reshaping Human Antibodies: Grafting an Antilysozyme Activity," *Science* 239:1534-1536 (1988)]). 인간화 항체는 Fv 프레임워크 영역에서 및/또는 교체된 비-인간 잔기 내에서 추가의 잔기의 치환에 의해 추가로 개질되어 항체의 특이성, 친화도 및/또는 능력을 개선 및 최적화할 수 있다.

[0166] 또한 이중특이적 항체는 본 발명의 방법에 사용하기에 적합하다. 이중특이적 항체는 적어도 2개의 상이한 에피토프를 특이적으로 인식 및 결합할 수 있는 항체이다. 이중특이적 항체는 무손상 항체 또는 항체 단편일 수 있다. 이중특이적 항체의 생성 기술은 관련 기술분야에 공지되어 있다 (문헌 [Traunecker et al., "Bispecific Single Chain Molecules (Janusins) Target Cytotoxic Lymphocytes on HIV Infected Cells," *EMBO J.* 10:3655-3659 (1991)] 및 [Gruber et al., "Efficient Tumor Cell Lysis Mediated by a Bispecific Single Chain Antibody Expressed in Escherichia coli," *J. Immunol.* 152:5368-74 (1994)]).

[0167] **원핵생물에서 글리코실화 글루카곤 펩티드의 생성**

[0168] 글루카곤 펩티드를 개선하기 위한 간단한 시험관내 당접합 기술이 입증되어 있지만, 그러한 펩티드의 치료적 결점이 여전히 존재하는데, 즉 이들은 작고 일반적으로 단량체성이고 일반적으로 몇 시간 미만의 짧은 반감기를 가지며 작은 펩티드로는 PEG화가 거의 제대로 동작하지 않는다. 현재 접근법은 현저히 저해된 활성을 여전히 겪고 있다.

[0169] 본 발명은 목적하는 글리칸을 갖는 신규 글리코실화 펩티드에 관한 것이다. 글리코실화 글루카곤 펩티드의 이점은 개선된 용해도, 겔 및 원섬유 형성에 대한 개선된 물리적 안정성과 함께, 증가된 반감기 및 개선된 활성 및 약동학 특성을 포함한다. 다른 이점은 단백질 및 글리칸 둘 모두를 생성하여 다중 단계를 피하는 단일 또는 동시 생체내 과정의 능력을 포함한다. 일부 실시양태에서, 신규 글리코실화 글루카곤 펩티드는 연장된 혈장 제거 반감기로 인한 생체내 연장된 노출 및 연장된 흡수 상 및 중성 pH 또는 약염기성 pH에서 개선된 수용해도를 갖는다. 다른 실시양태에서, 본 발명은 수용액 중에 겔 및 원섬유의 형성에 대한 개선된 안정성을 갖는다. 바람직한 실시양태에서, 우세형 N-글리칸은 포유동물에게 면역원성을 금지하지 않는다.

[0170] N-글리코실화 부위 점유는 진핵 시스템, 예를 들어 생성된 임의의 특정한 당단백질의 경우 CHO 및 효모에서 다양할 수 있다. 부위에서의 점유를 제어하는 성장 조건이 만들어질 수 있다.

- [0171] 전형적으로, 글루카곤 펩티드는 글리코실화를 갖지 않는다. 특정 실시양태에서, 글리코실화 부위는 펩티드 상에서 조작된다. 예시적 실시양태에서, 본 발명의 글루카곤 펩티드는 하나의 글리코실화 부위를 갖는다. 특정 실시양태에서, 방법은 펩티드마다 다중 글리칸을 첨가하여 더 나은 활성을 부여하는 것을 제공한다. 추가 실시양태에서, 숙주 세포는 우세형 종으로서 특정한 *N*-글리칸을 갖는 글루카곤 펩티드를 생성하도록 조작된다. 예시적인 글리코실화 패턴이 도 11에 나타난다.
- [0172] 따라서, 본 발명의 방법은 하나 이상의 당형태를 포함하는 당단백질 및 당펩티드를 제공한다. 바람직하게는, 당형태는 개선된 용해도 또는 안정성 특성뿐만 아니라 증가된 수용체 결합 활성을 부여하는 M4, M5, G0, G0(1), G0(2), G0(3), G1, G2, G3, G4, G5, S1, S2, S3, S4, S5를 포함한다. 글루카곤과 같은 아글리코실화 펩티드와 비교하여, 본 발명은 펩티드에 대한 반감기를 증가시키는 것으로 기대된다. 추가의 펩티드는 본 발명의 방법, 예컨대 hGH, ASNase 및 IL1-Ra에 의해 생성되고 있다. 다른 펩티드의 생성은 본 발명의 범위 내에 있다. 바람직한 실시양태에서, 적어도 50 몰%의 글루카곤 펩티드가 글리코실화된다.
- [0173] **백신 제조**
- [0174] 일반화된 후보 항원의 면역원성을 증가시키는 방법은 백신 개발의 초기 단계에서 투자되는 시간 및 비용을 감소시킬 수 있고 임의의 관심 질환에 적용될 수 있다. 면역원성을 증가시키는 하나의 방증된 전략은 만노실화, 즉 단백질로의 만노스-말단 글리칸의 접합이다. 만노스는 항원을 수용체-매개 세포내이입에 의한 내재화를 위해 항원 제시 세포 (APC) 상에 CD206 및 CD209를 포함하는 특정한 수용체에 표적화하여 음세포작용을 통해 차지되는 항원에 비해 항원 제시에서 200배 이하로 증가한다 (문헌 [Engering, A., et al., *The mannose receptor functions as a high capacity and broad specificity antigen receptor in human dendritic cells*. *Eur J Immunol*, 1997. **27**(9): p. 2417-25.], [Lam, J.S., et al., *A Model Vaccine Exploiting Fungal Mannosylation to Increase Antigen Immunogenicity*. *The Journal of Immunology*, 2005. **175**(11): p. 7496-7503.]). 항원의 만노실화는 (i) APC에 의한 증가된 항원 흡수, (ii) 10,000배 이하까지 증가된 MHC 클래스 II-매개 항원 제시, (iii) T 세포 증식 및 성숙의 촉진, 및 (iv) 혈청의 살박테리아 활성을 포함하는 개선된 체액성 면역 반응을 포함한 몇 가지 이점을 부여한다 (문헌 [(Arigita, C., et al., *Liposomal Meningococcal B Vaccination: Role of Dendritic Cell Targeting in the Development of a Protective Immune Response*. *Infection and Immunity*, 2003. **71**(9): p. 5210-5218.]). *E. 콜라이*는 주로 백신 생산용 플랫폼으로서 사용되지 않았는데 그 이유는 *N*-글리코실화를 위한 경로를 자연적으로 코딩할 수 없고 따라서 당단백질의 제조에 적합하지 않았기 때문이다.
- [0175] 특정 실시양태에서, 본 발명은 *E. 콜라이*의 당조작된 균주를 통한 만노실화 백신 항원을 위한 방법 및 조성물을 제공한다. 면역원성에 대한 만노실화의 영향은 마우스 모델에서 평가된다. 박테리아에서 백신 후보를 생산하는 능력은 다중 이점을 제공한다. *E. 콜라이*는 ExPEC (장외 병원성) 및 다른 박테리아 단백질의 발현을 위한 탁월한 플랫폼이고, 손쉬운 재조합 DNA 조작을 제공하고, 큰 조합적 라이브러리를 생성하는데 사용될 수 있고, 신속하고 낮은 비용 균주 개발 및 생성에 이르는 빠른 랩프-업을 가능하게 하며, 진핵 발현 시스템과 조우되는 바이러스 오염에 대한 위험성을 제거한다 (문헌 [Aguilar-Yanez, J., et al., *An influenza A/H1N1/2009 hemagglutinin vaccine produced in Escherichia coli*. *PLoS One*, 2010. **5**(7): p. e11694. Choi, B.-K., et al., *Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast Pichia pastoris*. *Proceedings of the National Academy of Sciences*, 2003. **100**(9): p. 5022-5027.]). *E. 콜라이*에서 만노실화 후보 항원의 생성은 추가의 화학적 또는 효소적 개질을 필요로 하지 않고 생체내 목적하는 당단백질의 합성을 가능하게 할 수 있다. 따라서, 백신 개발을 위한 신규한 패러다임이 *E. 콜라이*-생성된 백신 후보의 효능을 늘리는 방법에 의해 제공된다.
- [0176] 본 발명의 당조작된 *E. 콜라이*는 증가된 면역원성을 갖는 만노실화 단백질을 생성하기 위해 고려된다. *E. 콜라이*에서 만노실화 항원의 합성은, 증가된 면역원성 특성을 갖는 후보 단백질의 저렴하고 신속한 제조를 가능하게 하는 백신 개발의 상당한 진보를 제시한다. 과거에, 만단 또는 만노스-말단 글리칸의 시험관내 화학적 접합, 효모 고 만노스 올리고사카라이드에 의한 글리코실화를 위한 피키아 파스토리스에서 단백질의 생체내 발현, 또는 만노실화 리포솜에서 항원의 시험관내 캡슐화를 포함하는 항원을 만노실화하는 몇몇 전략이 적용되었다 (문헌 [(Lam, J.S., et al., Arigita, C., et al., Apostolopoulos, V., et al., *Oxidative/reductive conjugation of mannan to antigen selects for T1 or T2 immune responses*. *Proceedings of the National Academy of Sciences*, 1995. **92**(22): p. 10128-10132.], [Sheng, K., et al., *Delivery of antigen using a novel mannosylated dendrimer potentiates immunogenicity in vitro and in vivo*. *Eur J Immunol*, 2008. **38**(2): p. 424-36.]). 그러나, 지금까지, 백신 개발을 위한 박테리아에서 단백질에 만노스-말단 글리칸의 직접

적인 생체내 접합은 전혀 달성되지 않았다. E. 콜라이 발현 플랫폼은 일반적 단백질 생산 및 글리칸의 발현을 위한 이소성 숙주 둘 모두의 면에서 현존하는 기술에 비해 다중 이점을 제공할 것이다.

[0177] 예비 발현 및 글리코실화에 요로병원성 E. 콜라이 (UPEC) 항원, c1275가 선택되었다. c1275 단백질은 당조작된 E. 콜라이의 주변세포질에 표적화되고, GlycTag로 개질되며 (도 3의 A)), C. 제주니 헵타사카라이드 글리칸을 어셈블리하는데 필요한 OST PgIB 및 글리코실트랜스퍼라제로 공동-발현되었다. 정제 시에, 글리코실화 c1275는 웨스턴 블롯 및 GalNAc 특이적 렉틴 대두 응집소 (SBA)를 갖는 이들 생성물의 반응성 상의 더 느린-이동 밴드의 외관에 기초하여 뚜렷하였고, 이는 이 올리고사카라이드를 인식하는 것으로 공지되어 있다 (도 12) (문헌 [Young, N.M., et al., *Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, Campylobacter jejuni*. J Biol Chem, 2002. **277**(45): p. 42530-9.]). 흥미롭게도, c1275의 글리코실화는 GlycTag의 존재에 반드시 의존하는 것은 아니다. 서열 분석에 의해 공개된 PgIB의 수용자 부위 요건을 충족시키는 이 단백질에서 천연 DSNAT 모티프의 존재가 드러난다. 성공적인 c1275의 글리코실화는 E. 콜라이 주변세포질에서 후보 백신 항원이 성공적으로 발현되고 글리코실화될 수 있다는 가설을 뒷받침한다.

[0178] 통상적인 백신은 단지 체액성 면역 반응을 유도한다. DNA 백신은 살아있는 바이러스 백신과 관련된 동일한 위험 없이 체액 및 세포-매개 면역 둘 모두를 불러 일으키기 때문에 큰 가능성을 유지한다. 살아있는 약독화 바이러스 백신과는 반대로 DNA 백신은 특정한 수용체에 결합해야 하는 살아있는 바이러스와는 달리 동일한 또는 상이한 조직 또는 세포에 전달될 수 있다. 그의 천연 형태로 항원의 생성은 숙주 면역계에 대한 항원의 제시를 개선한다. 살아있는 약독화 백신과 달리, DNA 백신은 감염성이 없고 병독성으로 변환할 수 없다.

[0179] 후보 항원은 다양한 올리고사카라이드 예컨대 Man₃GlcNAc₂로 개질된다. 이는 백신 제제에 사용하기 위한 진핵 만노스-말단 글리칸으로 개질된 항원의 생성을 초래할 수 있다. 다수의 표적 항원이 마우스 모델에서 보호를 부여하는 것으로 공지된 ExPEC 백신 후보의 공개된 평가로부터 선택된다. 그러나, 본 발명은 고도로 모듈성이어서 다양한 단백질 및 펩티드 후보의 백신 개발을 강화하는데 널리 적용될 수 있음을 주목해야 한다.

[0180] **제약 제제**

[0181] 당단백질의 치료 제제는 동결건조 제제 또는 수용액의 형태로, 목적하는 순도를 갖는 당단백질을 임의의 생리학 상 허용되는 담체, 부형제 또는 안정화제와 함께 혼합하여 제조될 수 있다 (문헌 [Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)]). 허용되는 담체, 부형제 또는 안정화제는 적용된 투여량 및 농도에서 수용자에게 비독성이며, 완충제, 예컨대 포스페이트, 시트레이트 및 다른 유기 산; 아스코르브산 및 메티오닌을 포함한 항산화제; 보존제 (예컨대, 옥타데실디메틸벤질 암모늄 클로라이드; 헥사메토늄 클로라이드; 벤즈알코늄 클로라이드, 벤제토늄 클로라이드; 페놀, 부틸 또는 벤질 알콜; 알킬 파라벤, 예컨대 메틸 또는 프로필 파라벤; 카테콜; 레조르시놀; 시클로헥산올; 3-펜탄올; 및 m-크레졸); 저분자량 (약 10개 미만의 잔기) 폴리펩티드; 단백질, 예컨대 혈청 알부민, 젤라틴 또는 이뮤노글로불린; 친수성 중합체, 예컨대 폴리비닐피롤리돈; 아미노산, 예컨대 글리신, 글루타민, 아스파라긴, 히스티딘, 아르기닌 또는 리신; 모노사카라이드, 디사카라이드, 및 글루코스, 만노스, 또는 텍스트린을 포함한 다른 탄수화물; 킬레이트화제, 예컨대 EDTA; 당, 예컨대 수크로스, 만니톨, 트레할로스 또는 소르비톨; 염-형성 반대-이온, 예컨대 소듐; 금속 착물 (예를 들어, Zn-단백질 복합체); 및/또는 비-이온성 계면활성제, 예컨대 트윈(TWEEN)TM, 플루로닉스트 (PLURONICST)TM 또는 폴리에틸렌 글리콜 (PEG)을 포함한다.

[0182] 글리칸은 PEG 중합체에 편리한 앵커인데 그 이유는 특정 당, 예컨대 만노스 또는 갈락토스가 순한 산화제, 예컨대 과아이오딘산나트륨의 존재하에 반응성 알데히드로 쉽게 변환될 수 있기 때문이다 (문헌 [Soares, A.L., et al., *Effects of polyethylene glycol attachment on physicochemical and biological stability of E. coli L-asparaginase*. Int J Pharm, 2002. **237**(1-2): p. 163-70]). 히드라진 기로 관능화된 PEG 중합체는 이어서 글리코PEG화 생체접합체 생성에 사용될 수 있다. 이는 부위-특이적이고 고도로 제어되고 균질하며 활성인 단백질 접합체의 합성을 허용한다. PEG화는 비-특이적 과정의 결과로서 이질성 및 활성 손실의 문제를 종종 초래한다. 부위-특이적 PEG화 방법은 (i) 리신에 돌연변이를 일으켜 PEG를 특정 리신에 표적화시키는 것 (문헌 [Narimatsu, S., et al., *Lysine-deficient lymphotoxin-alpha mutant for site-specific PEGylation*. Cytokine, 2011. **56**(2): p. 489-93.], [Youn, Y.S. and K.C. Lee, *Site-specific PEGylation for high-yield preparation of Lys(21)-amine PEGylated growth hormone-releasing factor (GRF) (1-29) using a GRF(1-29) derivative Fmoc-protected at Tyr(1) and Lys(12)*. Bioconjug Chem, 2007. **18**(2): p. 500-6]) 또는 N-말단의 아민 기에 표적화시키는 것 (문헌 [Lee, H., et al., *N-terminal site-specific mono-PEGylation of epidermal*

growth factor. Pharm Res, 2003. **20**(5): p. 818-25.], [Yamamoto, Y., et al., *Site-specific PEGylation of a lysine-deficient TNF-alpha with full bioactivity*. Nat Biotechnol, 2003. **21**(5): p. 546-52]) 또는 (ii) 짝이 없는 시스테인 잔기를 첨가하여 유리 티올 기를 표적화시키는 것 (문헌 [Shaunak, S., et al., *Site-specific PEGylation of native disulfide bonds in therapeutic proteins*. Nat Chem Biol, 2006. **2**(6): p. 312-3.], [Doherty, D.H., et al., *Site-specific PEGylation of engineered cysteine analogues of recombinant human granulocyte-macrophage colony-stimulating factor*. Bioconjug Chem, 2005. **16**(5): p. 1291-8.], [Manjula, B.N., et al., *Site-specific PEGylation of hemoglobin at Cys-93(beta): correlation between the colligative properties of the PEGylated protein and the length of the conjugated PEG chain*. Bioconjug Chem, 2003. **14**(2): p. 464-72])과 관련된다. 이들 접근법은 몇몇 주요 결점을 갖는다. 첫째, 양으로 하전된 리신은 종종 단백질 구조/기능에 중요하다 (문헌 [Yoshioka, Y., et al., *Optimal site-specific PEGylation of mutant TNF-alpha improves its antitumor potency*. Biochem Biophys Res Commun, 2004. **315**(4): p. 808-14]). 둘째, 시스테인 잔기의 첨가는 가용성 발현 및 디설피드 결합 형성과 함께 심각한 문제를 야기하고, 심지어 포유동물 발현 숙주로의 이동을 필요로 할 수 있다 (문헌 [Constantinou, A., et al., *Site-specific polysialylation of an antitumor single-chain Fv fragment*. Bioconjug Chem, 2009. **20**(5): p. 924-31]). 셋째, 부위-특이적 PEG화는 연결된 PEG 분자의 수를 엄격히 제한한다. 글리코PEG화는, 단백질 내에 특정한 잔기에 이미 부착되어 있는 글리칸에 대한 PEG의 접합과 관련된다. 이점은 다음과 같다: (i) 과정이 부위-특이적이고, (ii) 글리코실화 부위가 활성 부위(들)로부터 떨어져 조작될 수 있으며, (iii) 생성물은 고도로 활성이고 비교적 균질일 수 있다.

[0183] 본원에서의 제제는 또한 처리될 특정한 지표에 필요한 것으로 하나 초과와 활성 화합물, 바람직하게는 서로 약 영향을 주지 않는 상보적 활성을 갖는 것을 함유할 수 있다. 예를 들어, 제제는 또 다른 항체 또는 화학요법제를 추가로 포함할 수 있다. 그러한 분자는 의도된 목적에 유효한 양으로 조합되어 적합하게 존재한다.

[0184] 또한 활성 성분은 예를 들어, 코아세르베이션 기술에 의해 또는 계면 중합에 의해 제조된 마이크로캡슐, 예를 들어 각각, 콜로이드성 약물 전달 시스템 (예를 들어, 리포솜, 알부민 마이크로구체, 마이크로에멀전, 나노입자 및 나노캡슐)에서 또는 마크로에멀전 중에 히드록시메틸셀룰로스 또는 젤라틴-마이크로캡슐 및 폴리-(메틸메타크릴레이트) 마이크로캡슐 내에 포착될 수 있다. 그러한 기술은 문헌 [Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)]에 개시된다.

[0185] 생체내 투여에 사용될 제제는 멸균되어야 한다. 이는 멸균 여과 막에 의한 여과에 의해 쉽게 달성된다. 지효성 제제가 제조될 수 있다. 지효성 제제의 적합한 예는 당단백질을 함유하는 고체 소수성 중합체의 반투과성 매트릭스를 포함하며, 여기서 매트릭스는 성형품의 형태, 예를 들어 필름 또는 마이크로캡슐이다. 지효성 매트릭스의 예는 폴리에스테르, 히드로겔 (예를 들어, 폴리(2-히드록시에틸-메타크릴레이트), 또는 폴리(비닐알콜)), 폴리락티드 (미국 특허 제3,773,919호), L-글루탐산 및 γ 에틸-L-글루타메이트의 공중합체, 비-분해성 에틸렌-비닐 아세테이트, 분해성 락트산-글리콜산 공중합체, 예컨대 루프론 데포(LUPRON DEPOT)TM (락트산-글리콜산 공중합체 및 류프롤리드 아세테이트로 구성되는 주사가능한 마이크로구체), 및 폴리-D-(-)-3-히드록시부티르산을 포함한다. 에틸렌-비닐 아세테이트 및 락트산-글리콜산과 같은 중합체는 100일이 넘는 동안 분자의 방출이 가능하지만, 특정 히드로겔은 더 짧은 시간 주기 동안 단백질을 방출한다. 캡슐화된 항체가 긴 시간 동안 몸체 내에 잔류할 때, 이들은 37°C에서 수분에 노출된 결과로서 변성되거나 또는 응집하여, 생물학적 활성의 손실 및 면역원성의 가능한 변화를 초래할 수 있다. 관련된 메카니즘에 따라서 합리적 전략이 안정화를 위해 도모될 수 있다. 예를 들어, 응집 메카니즘이 티오-디설피드 교환을 통한 분자간 S-S 결합 형성인 것으로 밝혀진다면, 안정화는 술폰히드릴 잔기를 개질하고, 산성 용액으로부터 동결건조하고, 수분 함량을 조절하고, 적절한 첨가제를 사용하며, 특정한 중합체 매트릭스 조성물을 전개함으로써 달성될 수 있다.

[0186] 제약 조성물은 동결건조될 수 있다. 동결건조된 항체 제제는 미국 특허 제6,267,958호에 기재된다. 안정한 수성 항체 제제는 미국 특허 제6,171,586B1호에 기재된다.

[0187] 특정 측면에서, 본 발명의 방법 및 조성물은 비-치료 목적, 예컨대 검정, 진단, 시약 및 키트에 사용될 수 있다.

[0188] **키트**

[0189] 본 발명은 올리고사카라이드 물질을 함유하는 제조 물품 및 키트를 추가로 제공한다. 제조 물품은 표지를 구비한 용기를 포함한다. 적합한 용기는 예를 들어, 병, 바이알, 및 시험 튜브를 포함한다. 용기는 다양한 물질, 예컨대 유리 또는 플라스틱으로 형성될 수 있다. 용기는 본원에 기재된 올리고사카라이드 제제를 포함하는 조

성물을 보유한다. 다른 실시양태에서, 키트는 당단백질을 포함한다. 용기 상의 표지는, 조성물이 특정한 질환 또는 장애의 치료 또는 예방에 사용되는 것을 나타내며, 또한 생체내, 예컨대 상기 기재된 것에 대한 방향을 나타낼 수 있다. 본 발명의 키트는 상기 기재된 용기 및 완충제를 포함하는 제2 용기를 포함한다. 그것은 상업 및 사용자 관점에서 바람직한 다른 물질 (다른 완충제, 희석제, 필터, 바늘, 시린지, 및 사용 지침서가 구비된 포장 삽입물을 포함함)을 추가로 포함할 수 있다.

[0190] 최종적으로, 원핵생물 (예를 들어, *E. 콜라이*)에서 다양한 당형태의 합성은 단백질에 대한 부착, 글리칸 어레이로의 혼입, 및 기재로서의 이용을 용이하게 하여, 다른 인간-유사, *N*-연결된 글리칸, 진단, 키트 또는 시약을 제조할 수 있다.

[0191] 상기 개시내용은 일반적으로 본 발명을 기재한다. 이하 하기 실시예에서 보다 구체적인 설명이 제공된다. 실시예는 단지 설명을 목적으로 하는 것이고 본 발명의 범주를 제한하려는 의도는 아니다. 등가물의 형태 및 치환의 변화는 상황이 제시되거나 또는 방편을 제공함에 따라서 고려된다. 본원에는 특정한 용어가 사용되었지만, 이러한 용어는 기술적인 의미로 의도되는 것이고 제한하려는 것이 아니다.

[0192] **실시예**

[0193] **실시예 1 - 플라스미드 구축**

[0194] 발데라마-린콘 등은 최근 *E. 콜라이*의 세포질 막에서 Und-PP 상의 Man₃GlcNAc₂의 생합성 및 어셈블리를 위한 생합성 경로를 개시하였다. 야생형 뉴클레오티드 서열 또는 코돈 최적화 서열을 갖는 Alg13 Alg14 Alg1 및 Alg2 활성을 포함하는 경로는 원핵 숙주 세포에 진핵 글리코실트랜스퍼라제 활성을 부여한다. 이 경로는 GlcNAc 및 만노스 단위를 운데카프레놀-연결된 담체 기재에 첨가하여 트리만노실 코어 올리고사카라이드 구조를 획득하는데 기여한다. *E. 콜라이*는 UDP-GlcNAc로부터의 GlcNAc-1-포스페이트의 운데카프레닐 포스페이트 (Und-P)로의 전달을 매개하여 Und-PP-GlcNAc를 형성하는 내재성 막 단백질 WecA를 포함한다 (문헌 [Rick, P.D. & Silver, R.P. in *Escherichia coli and Salmonella: Cellular and Molecular Biology*. (ed. F.C.a.o. Neidhardt) 104-122 (American Society for Microbiology, Washington, D.C.; 1996)]). 따라서, 천연 생성된 Und-PP-GlcNAc는 목적하는 Man₃GlcNAc₂ 글리칸을 위한 후보 전구체로서 존재한다. 제2 GlcNAc 잔기의 첨가 동안, 2개의 서브유닛, Alg13 및 Alg14로 구성되는 사카로미세스 세레비지아에 β1,4-GlcNAc 트랜스퍼라제가 발견되었다. 효모에서, Alg14는 ER의 시토졸 면 (여기서, Do1-PP-GlcNAc₂에 대한 촉매작용이 발생함)에 가용성 Alg13을 구성하는 막 앵커로서 기능하는 내재성 막 단백질이다 (문헌 [Bickel, T. et al., *Biosynthesis of lipid-linked oligosaccharides in Saccharomyces cerevisiae: Alg13p and Alg14p form a complex required for the formation of GlcNAc(2)-PP-dolichol*. *J Biol Chem* **280**, 34500-34506 (2005)]). *E. 콜라이*에서 공동-발현될 때, Alg14는 막 분획에서 국부화하는 것으로 관찰되었던 반면 Alg13은 효모에서의 상황과 일치하는 세포질 및 막 분획 둘 모두에서 발견되었다. 후속 단계 동안, 제1 만노스를 글리칸에 부착하는 *S. 세레비지아*에 β1,4-만노실트랜스퍼라제 Alg1, 및 α1,3- 및 α1,6-만노스 잔기 둘 모두를 글리칸에 첨가하는 것을 촉매하는 이관능성 만노실트랜스퍼라제 Alg2가 발견되었다 (문헌 [O'Reilly, M.K., et al., *In vitro evidence for the dual function of Alg2 and Alg11: essential mannosyltransferases in N-linked glycoprotein biosynthesis*. *Biochemistry* **45**, 9593-9603 (2006)]). *E. 콜라이*에서의 발현에 이어서, Alg1 및 Alg2 둘 모두가 세포 막에서 국부화되었다. 정확하게 국부화된 Alg 효소가 Und-PP 상에 Man₃GlcNAc₂를 생성할 수 있었는지를 결정하기 위해서, Alg13, Alg14, Alg1 및 Alg2의 동시 발현을 허용하는 플라스미드 pYCG (발데라마-린콘 등)가 구축되었다.

[0195] 플라스미드 pMW07이 구축되었다 (발데라마-린콘 등). 플라스미드 pMQ70 (문헌 [Shanks et. al, 2006 AEM. 72(7)5027-5036.])이 Eam11051의 이소스키조머인 Ahd1로 선형화되었다. p15a ori 및 cat 유전자는 pBAD33으로부터 증폭되었고 선형화 벡터 pMQ70으로 효모를 공동-형질전환하는데 사용되었다. 효모에서 상동 재조합은 colE1 ori 및 *bla* 유전자 조작 벡터 pMW07 (발데라마-린콘 등)의 대체를 초래하였다. 표 3에 다양한 균주의 구조 및 유전자형이 열거된다.

[0196] **실시예 2 - 분석 프로토콜**

[0197] *N*-연결된 올리고사카라이드의 추출 및 정제 방법이 가오(Gao) 등 (문헌 [Gao et al., "Non-radioactive analysis of lipid linked oligosaccharide composition by fluorophore-assisted carbohydrate electrophoresis," *Method Enzymol* 415: 3-20])에 기재된 바와 같이 후속되었다. 정제된 올리고사카라이드는 매트릭스 (에이비 사이엑스(AB Sciex) TOF/TOF 5800)로서 디히드록시벤조산 (DHB)을 사용한 MALDI-TOF 질량 분

광측정법에 의해 분석되었다.

[0198] 글리칸 도먼은 표준 CFG (기능 유전체학용 컨소시엄: Consortium for Functional Genomics) 흑색 및 백색 표기법이며, 이는 글리코 워크벤치(Glyco Workbench) 2.0으로 생성되었다.

[0199] **실시예 3 - E. 콜라이에서 인간-유사 N-연결된 Man₃GlcNAc₂ 고 만노스 올리고사카라이드의 생성**

[0200] 인간 및 다른 진핵생물에서, Man₃GlcNAc₂ 당형태는 글리칸 합성의 주요 중간체이다. 진핵생물에서, 이 주요 당형태는 세포질 세망 막의 시토졸 측 상에서 합성된다. 효소 Alg11은 Man₃GlcNAc₂ 글리칸 코어의 α1,3 만노스로의 2개의 α1,2-만노스 잔기의 첨가를 촉매한다. 사카로미세스 세레비지아에로부터의 Alg11을 코딩하는 유전자는 Man₃GlcNAc₂ 트리만노실 코어의 생성에 사용되는 플라스미드 pMW07-YCG-Pg1B.CO로의 글루타티온 S-트랜스퍼라제를 코딩하는 유전자 (*gst*)에 대한 융합체로서 클로닝되었다 (발데라마-린콘 등). 생성된 플라스미드는 전기천공에 의해 E. 콜라이 MC4100 Δ*waaL gmd::kan*으로 형질전환되었다 (Gly02). Gly02는 100 mL의 루리아-베르타니(Luria-Bertani) (LB) 브로쓰 중에서 성장되었고 0.2% (v/v) 아라비노스를 첨가함으로써 유도되었고 단번에 배양액은 3.0의 광학 밀도에 이르렀다.

[0201] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석에 의해 목적하는 Man₃GlcNAc₂ 당형태와 일치하는 우세형 피크 (m/z 1257.6 Na+)가 드러났다 (도 1의 A). 몇몇 샘플에서, Man₃GlcNAc₂ 당형태와 일치하는 지엽적인 피크가 나타났다 (m/z 933.5 Na+). 다른 예에서, Man₂GlcNAc₂, Man₄GlcNAc, Man₃GlcNAc₂, HexMan₃GlcNAc₂, HexMan₅GlcNAc, Man₆GlcNAc와 일치하는 글리칸을 포함하는 지엽적인 피크가 나타났다. Man₃GlcNAc₂ 글리칸 코어의 기대되는 α1,2 만노스 잔기의 첨가를 확인하기 위해서, 정제된 글리칸은 제조업체의 프로토콜에 따라서 α1,2 만노시다제 (프로자임(Prozyme))로 처리되었다. 효소와의 인큐베이션에 이어서, 글리칸은 표지되었고 질량 분광측정 및 가오 등의 방법에서의 페이스(FACE) 겔에 의해 분석되었다. 미처리된 샘플에서, Man₃GlcNAc₂ 당형태와 일치하는 우세형 피크가 관찰되었다 (도시하지 않음). 처리된 샘플에서, Man₃GlcNAc₂ 당형태와 일치하는 우세형 피크 (m/z 933.4 Na+)가 관찰되었다 (도 1의 B). 이에 의해 Man₃GlcNAc₂ 글리칸 코어의 기대되는 2개의 α1,2-만노스 잔기의 첨가가 확인된다. 결과로서, 인간-유사 Man₃GlcNAc₂ 당형태는 E. 콜라이에서 Alg11의 발현에 의해 생성될 수 있다. Man₃GlcNAc₂ 당형태의 단리는, 진핵생물에서 그것은 일과성 올리고사카라이드이기 때문에 다른 수단에 의해 시도되고 있다. 이 시스템에서 Man₃GlcNAc₂의 합성은 또한 충분한량의 활성 효소의 발현에 있어서의 어려움으로 인해 시도되고 있었다. 다양한 융합 파트너가 오직 Alg11과 함께, 분석되었고 대부분의 시험된 Alg11 모이어티의 경우 효율적인 생물물 형성의 기여가 초래되었다. Alg11에 융합된 *gst* 및 *mstX*는 이 시스템에서 Man₃GlcNAc₂ 당형태를 생성하였다.

[0202] **실시예 4 - E. 콜라이에서 하이브리드 N-연결된 GlcNAcMan₃GlcNAc₂ 올리고사카라이드의 생성**

[0203] 인간 및 다른 진핵생물에서, GlcNAcMan₃GlcNAc₂ 당형태는 글리칸 합성에 주요 중간체이다. 이 당형태는 전형적으로 오직 진핵생물의 골지에서의 단백질에 부착된 N-연결된 글리칸 상에서 발견된다. 여기서 글리칸은 E. 콜라이에서 지질 담체 상에 어셈블리되었다. 이를 달성하기 위해서, 니코티아나 타바쿰 N-아세틸글루코사미닐트랜스퍼라제 I (GnTI)의 말단절단된 형태 (잔기 30-446)를 코딩하는 유전자가 합성되었다. GnTI 유전자는 PCR에 의해 증폭되었고 그의 천연 신호 서열이 부족한 E. 콜라이 말토스 결합 단백질 (MBP)을 코딩하는 유전자 (*malE*)에 대한 융합물로서 플라스미드 pMQ70으로 서브클로닝되었다. 생성된 pMQ70-MBP-NtGnTI는 전기천공에 의해 Man₃GlcNAc₂ 트리만노실 코어의 생성을 위한 제2 플라스미드 pMW07-YCG-Pg1B.CO와 함께 E. 콜라이 MC4100 Δ*waaL gmd::kan* (Gly03) 및 오리가미(Origami)2 *gmd::kan* (Gly03.1)으로 형질전환되었고 (발데라마-린콘 등) 100 mL의 루리아-베르타니 (LB) 브로쓰 중에서 성장되었다. 글리코실트랜스퍼라제 발현은 0.2% (v/v) 아라비노스를 첨가함으로써 유도되었고 단번에 배양액은 3.0의 광학 밀도에 이르렀다.

[0204] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석에 의해 목적하는 GlcNAcMan₃GlcNAc₂ 당형태와 일치하는 우세형 피크가 드러났다 (m/z 1136.5 Na+) (도 2의 A). 지엽적인 피크는 Man₃GlcNAc₂ 당형태와 일치하였다 (m/z 933.4 Na+). Man₃GlcNAc₂ 글리칸 코어의 기대되는 GlcNAc의 첨가를 확인하기 위해서, 정제된 글리칸은 제조업체의 프로토콜에 따라서 β-N-아세틸글루코사미니다제 (뉴잉글랜드 바이오랩스(New England Biolabs))로

처리되었다. 효소와의 인큐베이션에 이어서, 글리칸은 표지되었고 질량 분광측정법, 및 가오 등의 방법에서 페이스 질에 의해 분석되었다. 미처리 샘플에서, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치하는 우세형 피크가 관찰되었다 (도시하지 않음). 처리된 샘플에서, 우세형 피크는 $\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치한다 (도 2의 B)). 이에 의해 $\text{Man}_3\text{GlcNAc}_2$ 글리칸 코어로서의 기대되는 β -GlcNAc의 첨가가 확인된다. 결과로서, 인간-유사 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태는 *E. 콜라이*에서 GnTI의 발현에 의해 생성될 수 있다. $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태의 단리는 진핵생물에서 그것은 일과성 올리고사카라이드이기 때문에 다른 수단에 의해 시도되고 있다. 장애물이 또한 이 시스템을 사용하여 접해지고, *E. 콜라이*에서 단독으로, 또는 *mstX*에 융합된 인간 GnTI의 발현이 우선 시도되었고 목적하는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태를 효율적으로 생성하지 않았다 (도면에 도시하지 않음). 더욱이, MBP에 융합되지 않는 경우, *N. 타바쿰* GnTI는 목적하는 생성물 생성에 실패하였다 (도면에 도시하지 않음).

[0205] **실시예 5 - *E. 콜라이*에서 *N*-연결된 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 복합 올리고사카라이드의 생성**

[0206] 인간 및 다른 진핵생물에서, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 복합 당형태 ("G0")는 글리칸 합성에 주요 중간체인데, 그것은 글리칸이 완전 테코레이트되는 "코어"이기 때문이다. 이 당형태는 전형적으로 오직 진핵생물의 골지에서 단백질에 부착된 *N*-연결된 글리칸 상에서 발견된다. 여기서 글리칸은 *E. 콜라이*에서 지질 담체 상에 어셈블리되었다. 이를 달성하기 위해서, 인간 *N*-아세틸글루코사미닐트랜스퍼라제 II (GnTII)의 말단절단된 형태 (잔기 30-447)를 코딩하는 유전자가 합성되었다. GnTII 유전자는 PCR에 의해 증폭되었고 그의 천연 신호 서열이 부족한 MBP에 대한 융합체로서 플라스미드 pMQ70으로 서브클로닝되었다. 생성된 pMQ70-MBP-hGnTII는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 기제 올리고사카라이드의 생성을 위한 제2 플라스미드 pMW07-YCG-MBP-NtGnTI와 함께 전기천공에 의해 *E. 콜라이* MC4100 $\Delta waal\ gmd::kan$, (*gly06*) 오리가미2 gmd::kan (*Gly06.1*), DR473 gmd::kan (*Gly06.2*) 및 셔플(Shuffle) $\Delta waal\ gmd::kan$ (*Gly06.3*)으로 형질전환되었다. 글리코실트랜스퍼라제 발현은 1 L의 루리아-베르타니 (LB) 브로쓰 내로의 접종 즉시 첨가된 0.2% (v/v) 아라비노스로 유도되었다.

[0207] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석에 의해 목적하는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치하는 우세형 피크 (*m/z* 1339.8 Na+)가 드러났다 (도 3). 지엽적인 피크는 $\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치하였다 (*m/z* 933.5 Na+). $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 와 일치하는 제2 지엽적인 피크 (*m/z* 1136.6 Na+)가 또한 스펙트럼에서 관찰되었다. 당조작된 *E. 콜라이*에서 GnTII의 발현은 시도되고 있는 것으로 증명되었는데, 여기서 3개의 유기체로부터의 GnTII는 발현 단독에 의해, 또는 *mstX* 또는 MBP에 융합된 경우에 시험되었다. 추가로, GnTII 발현은 산화적 및 비-산화적 박테리아 균주 둘 모두에서 시험되었다. 시험된 6개의 GnTII 모이머티 및 4개의 박테리아 균주 중에서, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 글리칸의 효율적인 생성은 4개의 박테리아 균주 중 하나에서 MBP-융합된, 인간 GnTII에서 나타났다 (도면에 도시하지 않음).

[0208] **실시예 6 - *E. 콜라이*에서 분지형 *N*-연결된 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 하이브리드 올리고사카라이드의 생성**

[0209] 멀티안테너리 *N*-연결된 글리칸의 합성은 인간 및 다른 진핵생물에서 공통 특징이다. 트리안테너리 올리고사카라이드의 생성은 *N*-아세틸글루코사미닐트랜스퍼라제 IV (GnTIV)에 의한 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 로의 GlcNAc 잔기의 첨가에 의해 달성된다. GnTIV는 또한 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 복합 글리칸의 구조 이성질체인 비안테너리 하이브리드 올리고사카라이드를 생성하는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 상에서 작용할 수 있다. 소(bovine) GnTIV의 말단절단된 형태 (잔기 93-535)를 코딩하는 박테리아 코돈 최적화된 유전자가 합성되었다. GnTIV 유전자는 PCR에 의해 증폭되었고 그의 천연 신호 서열이 부족한 MBP에 대한 융합체로서 플라스미드 pMQ70으로 서브클로닝되었다. 생성된 pMQ70-MBP-hGnTIV는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 기제 올리고사카라이드의 생성을 위한 제2 플라스미드 pMW07-YCG-MBP-NtGnTI와 함께 전기천공에 의해 *E. 콜라이* MC4100 $\Delta waal\ gmd::kan$ (*Gly05*) 및 오리가미2 gmd::kan (*Gly05.1*)으로 형질전환되었다. 글리코실트랜스퍼라제 발현은 1 L의 루리아-베르타니 (LB) 브로쓰로의 접종 즉시 첨가된 0.2% (v/v) 아라비노스로 유도되었다.

[0210] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석에 의해 목적하는 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치하는 우세형 피크 (*m/z* 1339.7 Na+)가 드러났다 (도 4의 A)). 일부 샘플에서, 지엽적인 피크는 $\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치하였다 (*m/z* 933.5 Na+). $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 글리칸으로의 기대되는 GlcNAc 의 첨가를 확인하기 위해서, 정제된 글리칸은 제조업체의 프로토콜에 따라서 β -*N*-아세틸글루코사미니다제 (뉴잉글랜드 바이오랩스)로

처리되었다. 효소와의 인큐베이션에 이어서, 글리칸은 표지되었고 질량 분광측정법에 의해 분석되었다. 미처리 샘플에서, $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치하는 우세형 피크가 관찰되었다 (도시하지 않음). 처리된 샘플에서, 우세형 피크는 $\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치한다 (도 4의 B)). 이에 의해 $\text{GlcNAcMan}_3\text{GlcNAc}_2$ 글리칸 코어의 기대되는 β - GlcNAc 의 첨가가 확인된다. 당조작된 *E. 콜라이*에서 GnTIV의 발현은 시도되고 있는 것으로 증명되었는데, 여기서 GnTIV 발현은 산화적 및 비-산화적 박테리아 균주 둘 모두에서 시험되었다. $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 글리칸의 효율적인 생성은 오직 산화적 박테리아 균주에서만 나타났다 (도면에 도시하지 않음).

[0211] 실시예 7 - *E. 콜라이*에서 다중 안테너리 *N*-연결된 $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ 하이브리드 올리고사카라이드의 생성

[0212] 트리안테너리, *N*-연결된 글리칸의 합성은 인간 및 다른 진핵생물에서 발견되는 특징이다. 하나의 그러한 트리안테너리 올리고사카라이드의 생성은 *N*-아세틸글루코사미닐트랜스퍼라제 IV (GnTIV)에 의한 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 로의 UDP-GlcNAc 잔기의 첨가에 의해 달성된다. 소 GnTIV를 코딩하는 코돈 최적화된 유전자가 합성되었다. GnTIV 유전자는 PCR에 의해 증폭되었고 플라스미드 pMQ70-MBP-hGnTII에서 인간 GnTII 유전자의 3'-말단을 지나서 브클로닝되었다. 생성된 구축물은 균주 GLY06.1을 생성하는 $\text{GlcNAcMan}_3\text{GlcNAc}_2$ 기제 올리고사카라이드의 생성을 위한 제2 플라스미드 pMW07-YCG-MBP-NtGnTI와 함께 전기천공에 의해 *E. 콜라이* 세포 (오리가미2 *gmd::kan*) 로 형질전환되었다. 글리코실트랜스퍼라제 발현은 1 L의 루리아-베르타니 (LB) 브로쓰로의 접종 즉시 첨가된 0.2% (v/v) 아라비노스로 유도되었다. *N*-연결된 올리고사카라이드의 추출 및 정제 방법은 가오 등에 기재된 바와 같이 후속되었다. 정제된 올리고사카라이드는 매트릭스 (에이비 사이엑스 TOF/TOF 5800)로서 DHB를 사용한 MALDI-TOF 질량 분광측정법에 의해 분석되었다.

[0213] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석에 의해 각각 m/z 값이 1339.9 (Na⁺) 및 1543.1 (Na⁺) 이고 기제 $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 및 $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ 생성물 당형태와 일치하는 2개의 우세형 피크가 드러났다 (도 5의 B)).

[0214] 기제 올리고사카라이드 G0(1)을 생성하기 위해서, 1 L의 GLY06 농축 배양액이 30°C에서 20시간 동안 0.2% v/v 아라비노스로 유도되었다. G0 올리고사카라이드는 가오 등에 기재된 방법에 따라서 단리되었다. 글리코실트랜스퍼라제는 별도의 100 mL 배양액에서 25°C에서 16시간 동안 0.2% v/v 아라비노스로 유도함으로써 발현되었다. 이 배양액은 원심분리에 의해 펠릿화되었고 2 mL의 GnTIV 활성 완충제 (50 mM 트리스, 10 mM MnCl_2 , pH 7.5) 중 에 재현탁되었고 초음파처리되었다. 용해물은 원심분리에 의해 분류되었고 20 μL 가 건조된 트리만노실 코어 기제 (약 5 μg)에 첨가되었다. 과량의 뉴클레오티드-당 (20 μg)이 반응에 첨가되고 후속적으로 30°C에서 배양되었다. 반응은 24시간 기간에 걸쳐 다양한 시점에서 MALDI-TOF 질량 분광측정법에 의해 모니터링되었다.

[0215] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석에 의해 목적하는 $\text{GlcNAc}_3\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치하는 피크 (m/z 1542.9 Na⁺)가 드러났다 (도 5의 A)).

[0216] 실시예 7 - *E. 콜라이*에서 하이브리드 *N*-연결된 $\text{GalGlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 하이브리드 올리고사카라이드의 생성

[0217] 인간 및 다른 진핵생물에서, $\text{GalGlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ 당형태는 글리칸 합성의 중간체이다. 이 당형태는 건강한 성인에게는 다소 이례적이지만, 전립선암이 걸린 개체에게서 나타난다 (문헌 [Kyselova et al., "Alterations in the serum glycome due to metastatic prostate cancer," J. Proteome Res. (2007)]). 여기서 글리칸은 *E. 콜라이*에서 지질 담체 상에 어셈블리되었다. 헬리코박터 필로리 β -1,4-갈락토실트랜스퍼라제 (GalT)를 코딩하는 유전자가 합성되었고, PCR에 의해 증폭되었고 플라스미드 pMQ70으로 서브클로닝되었다. 생성된 pMQ70-HpGalT는 $\text{GlcNAcMan}_3\text{GlcNAc}_2$ 기제 올리고사카라이드의 생성을 위한 제2 플라스미드 pMW07-YCG-MBP-NtGnTI와 함께 전기천공에 의해 MC4100 $\Delta waal gmd::kan$ (Gly04) 및 오리가미2 *gmd::kan* (Gly04.1)으로 형질전환되었다. 글리코실트랜스퍼라제 발현은 1 L의 루리아-베르타니 (LB) 브로쓰로의 접종 즉시 첨가된 0.2% (v/v) 아라비노스로 유도되었다.

[0218] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석에 의해 목적하는 $\text{GalGlcNAcMan}_3\text{GlcNAc}_2$ 당형태와 일치하는 우세형 피크 (m/z 1298.7 Na⁺)가 드러났다 (도 6). 몇몇 샘플에서, 지엽적인 피크가 $\text{Man}_3\text{GlcNAc}_2$ 당형태와 일치하였다 (m/z 933.5 Na⁺) (도면에 도시하지 않음). 당조작된 *E. 콜라이*에서 GalT의 발현은 시도되고 있는 것으로 증명되었는데, 여기서 소 및 인간으로부터의 GalT는 MBP 및 mstX에 미융합 및 융합된 둘 모두이고,

네이세리아 메닌기티데스는 *E. 콜라이*에서 목적하는 올리고사카라이드를 생성하지 않았다 (도시하지 않음). 더욱이, *H. 파일로리* GalT의 발현은 산화적 및 비-산화적 박테리아 균주 둘 모두에서 시험되었고 효율적인 갈락토실화는 오직 산화적 박테리아 균주에서만 나타났다.

[0219] **실시예 8 - *E. 콜라이* 중 *N*-연결된 Gal₂GlcNAc₂Man₃GlcNAc₂ 복합 올리고사카라이드의 제조**

[0220] 인간, 및 다른 진핵생물에서, Gal₂GlcNAc₂Man₃GlcNAc₂ 당형태는 글리칸 합성에서 주요 중간체이다. 이 당형태는 전형적으로 진핵생물에서 단백질에 부착되는 *N*-연결된 글리칸 상에서만 발견된다.

[0221] 하나의 샘플에서, 글리칸을 Gly06으로부터 생성된 GO 올리고사카라이드를 이용하여 생체의 생성하였으며, 방법은 실시예 7에 기재된 바와 같다. 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석은 목적하는 GalGlcNAc₂Man₃GlcNAc₂ 당형태와 일치하는 주요 피크 (m/z 1664.1 Na+)를 나타내었다 (도 7A).

[0222] 말단 갈락토실화 글리칸의 생체내 합성의 경우, 헬리코박터 필로리 β-1,4-갈락토실트랜스퍼라제 (GalT)를 코딩하는 유전자를 합성하고, PCR에 의해 증폭하고, 플라스미드 pMQ132 상에 서브클론화하였다. 생성된 pMQ132-HpGalT를 GlcNAc₂Man₃GlcNAc₂ 기재 올리고사카라이드의 제조를 위해 제2 플라스미드 pMW07-YCG-MBP-NtGnTI 및 제3 플라스미드 pMQ70-MBP-hGnTII와 함께 전기천공에 의해 Origami2 *gmd::kan* (Gly04.2)으로 형질전환시켰다. 글리코실트랜스퍼라제 발현을 0.2% (v/v) 아라비노스로 유도하였고, 1 L의 루리아-베르타니(Luria-Bertani) (LB) 브로쓰에 접종시 즉시 첨가하였다.

[0223] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석은 목적하는 Gal₂GlcNAc₂Man₃GlcNAc₂ 당형태와 일치하는 주요 피크(m/z 1664.1 Na+)를 나타내었다 (도 7A). Gly04.2에서 합성된 글리칸의 분석은 G2 당형태와 일치하는 피크 (m/z 1662.2 Na+), G1 당형태와 일치하는 피크 (m/z 1500.0 Na+), 및 GO 당형태와 일치하는 피크 (m/z 1337.9 Na+)를 나타내었다. *E. 콜라이*에서 Gal₂GlcNAc₂Man₃GlcNAc₂ 당형태를 제조할 때 실시예 7에 기재된 동일한 난점이 나타났는데, 왜냐하면 동일한 효소를 사용하여 두 생성물을 제조했기 때문이다.

[0224] **실시예 8 - *E. 콜라이* 중 *N*-연결된 NANAGalGlcNAcMan₃GlcNAc₂ 하이브리드 올리고사카라이드의 제조**

[0225] 기재 올리고사카라이드 GalGlcNAcMan₃GlcNAc₂의 생성을 위해, GLY04.1의 1 L의 조밀한 배양을 30°C에서 20시간 동안 0.2% v/v 아라비노스로 유도하였다. 기재 올리고사카라이드를 [Gao et al.]에 기재된 방법을 따라 단리시켰다. ST6을 별도의 100 mL의 배양물에서 16시간 동안 25°C에서 0.2% v/v 아라비노스에 의한 유도에 의해 발현시켰다. 이 배양물을 원심분리로 펠렛화하고, 2 ml의 ST6 활성 완충제 (50 mM 트리스, 10 mM MnCl₂, pH 7.5)에 재현탁시키고, 초음파처리하였다. 용해물을 원심분리에 의해 청정하고, 20 μL를 건조된 트리만노실 코어 기재 (~5 μg)에 첨가하였다. 과량의 뉴클레오티드-당 (20 μg)을 반응물에 첨가한 후, 30°C에서 인큐베이션하였다. 반응물을 24시간이라는 기간에 걸쳐 다양한 시점에서 MALDI-TOF 질량 분광측정법에 의해 모니터링하였다.

[0226] 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석은 목적하는 NANAGalGlcNAcMan₃GlcNAc₂ 당형태와 일치하는 피크 (m/z 1565.7 Na+)를 나타내었다 (도 8).

[0227] **실시예 9 - *E. 콜라이* 중 *N*-연결된 글리칸 수율의 최적화**

[0228] (i) 증가된 당단백질 제조 (ii) 및 글리칸 어레이와 같은 당분석 도구의 제조의 용이성을 포함한 글리코엔지니어드(glycoengineered) *E. 콜라이*에서 생성되는 *N*-연결된 글리칸의 양을 증가시키기 위한 다수의 방법이 존재한다. 따라서, Man₃GlcNAc₂ 글리칸, 트리만노실 코어 글리칸의 수율의 개선, 및 트리만노실 코어에 대한 GlcNAc 잔기의 첨가를 착수하였다. *E. 콜라이* 중 뉴클레오티드-당 풀이 제한적일 수 있다는 이해를 바탕으로, 뉴클레오티드-당 생합성 경로에서의 효소는 글리코엔지니어드 *E. 콜라이*에서 과다발현을 목적으로 하였다. 구체적으로는, 포스포마노뮤타제 (ManB), 만노스-1-포스페이트 구아닐릴트랜스퍼라제 (ManC), 및 글루타민-프룩토스-6-포스페이트 트랜스아미나제 (GlmS)를 조사하였으며, 여기서 ManB 및 ManC는 GDP-만노스의 형성에 포함되고, GlmS는 UDP-GlcNAc의 형성에 포함된다.

[0229] *E. 콜라이*로부터 ManB 및 ManC를 코딩하는 유전자를 비시스트론적으로(bicistronically) (ManC/ManB) 플라스미드 pMQ70내로 클로닝하고, 전기천공에 의해 Man₃GlcNAc₂ 트리만노실 코어 ([Valderrama-Rincon et al.])의 제조를 위해 pMW07-YCG-Pg1B.CO와 함께 *E. 콜라이* MC4100 Δ*waaL gmd::kan*으로 형질전환시켰다 (Gly01.2). *E. 콜라이*로부터 GlmS를 코딩하는 유전자를 플라스미드 pTrc99Y 내로 클로닝하고 ([Valderrama-Rincon et al.]), 전

기천공에 의해 pMW07-YCG-MBP-NtGnTI와 함께 *E. 콜라이* MC4100 $\Delta waal\ gmd::kan$ 으로 형질전환시켰다 (Gly01.3). pMW07-YCG-PglB.CO를 함유하는 *E. 콜라이* MC4100 $\Delta waal\ gmd::kan$ (Gly01) 및 pMW07-YCG-MBP-NtGnTI를 함유하는 *E. 콜라이* MC4100 $\Delta waal\ gmd::kan$ (Gly01.1)을 대조군으로 사용하였다. Gly01 및 Gly01.2를 100 mL의 루리아-베르타니 (LB) 브로쓰에서 성장시키고, 광학 밀도 (O.D.) 3에서 0.2% (v/v) 아라비노스로 발현을 유도하였다. Gly01.1 및 Gly01.3을 100 mL의 LB 브로쓰에서 성장시키고, O.D. 3에서 0.2% (v/v) 아라비노스 및 1 mM IPTG (오직 Gly01.3)로 발현을 유도하였다. *N*-연결된 올리고사카라이드의 추출 및 정제 방법은 [Gao et al.]에서 기재된 바에 따라 수행하였다. 정제된 올리고사카라이드를 [Gao et al.]에서 기재된 방법을 이용하여 형광단-보조 탄수화물 전기영동 (FACE)에 의해 분석하였다.

[0230] Gly01.2의 경우, Gly01과 비교하였을 때 트리만노실 코어의 제조에서 큰 증가가 관찰되었다 (도 9A, 좌측 패널). 그러나, Gly01 트리만노실 코어 밴드는 사실상 검출 불가능하기 때문에 수율의 차이를 정량화하는 것에는 어려움이 있었다. 유사하게, Gly01.3의 경우, Gly01.1과 비교하였을 때 글리칸 수율의 큰 증가가 관찰되었다 (도 9A, 우측 패널). 추가로, Gly01.1과 비교하여 GlcNAcMan₃GlcNAc₂에서 큰 증가가 관찰되었고, 이는 과다발현되는 이러한 효소를 표적 목표로 한 것이었다. 뉴클레오티드-당 생합성에 포함되는 효소가 다수 존재하므로, 글리코엔지니어드 *E. 콜라이* 중 과다발현을 표적으로 하는 효소가 어느 것인지에 대한 결정이 주의깊게 고려되어야 하며, 여기서 다수의 효소는 글리칸 수율에 거의 영향을 주지 않을 수도 있다.

[0231] cAMP 수준은 과량의 글리세롤을 갖는 *E. 콜라이*에서 높게 유지되므로, 글리세롤은 프로모터 억제를 통해 플라스미드로부터의 유전자 발현에 영향을 주지 않도록 글루코스에 대한 대안으로 탄소 공급원을 제공한다. 글리세롤의 사용은 도 9B에 나타난 바와 같이 글리칸 수율을 증가시키는 것으로 보였다. 피루베이트는 크렙스 주기에서 GDP에서 GTP로의 재순환에서 중요한 역할을 한다. GTP는 GDP-만노스 형성에 필요한 GDP-만노스 피로포스포릴라아제의 기체이다. 글리칸 수율의 증가는 또한 피루베이트의 첨가와 함께 고려된다 (도 9C).

[0232] ManC/B를 과다발현하는 숙주 세포의 질량 분광측정법에 의한 정제된 올리고사카라이드의 분석은 ManC/B를 과다발현하지 않는 숙주 세포와 비교하여 사실상 소소 피크의 제거를 나타내었다. GLY01.4는 목적하는 M3 당형태와 일치하는 단일의 주요한 피크 (m/z 933.5 Na⁺)를 생성하였다 (도 10D). GLY02.1은 목적하는 M5 당형태와 일치하는 단일의 주요한 피크 (m/z 1257.7 Na⁺)를 생성하였다 (도 10E). GLY01.5는 목적하는 하이브리드 GlcNAcMan₃GlcNAc₂ 당형태와 일치하는 단일의 주요한 피크 (m/z 1136.9 Na⁺)를 생성하였다 (도 10F).

[0233] **실시예 10 - *E. 콜라이* 중 글리코실화 글루카곤 제조**

[0234] 글루카곤 구축물은 *N*-연결된 글리코실화 부위 (DQNAT)에 이어서, C-말단에서 6-히스티딘 태그를 갖는 글루카곤으로 구성된다. 글루카곤은 벡터 pTrc99Y에서 3개의 연속적인 C-말단 TEV 프로테아제 부위 이후 MBP의 C-말단에 대한 융합체로서 발현된다. ManC 및 ManB를 코딩하는 유전자는 또한 글루카곤 코딩 영역의 3' 말단을 지나 이 벡터 내로 클로닝된다. 생성된 플라스미드를 상용하는 글리코실트랜스퍼라제 플라스미드와 함께 전기천공에 의해 *E. 콜라이* 세포 (Origami2 $\Delta waal$, $gmd::kan$)로 형질전환시켰다. 각 균주의 배양물 100 mL를 600 nm에서 ~2.0의 광학 밀도로 성장시키고, 16시간 동안 0.2% v/v 아라비노스로 유도한 후, 8시간 동안 30°C에서 0.1 mM IPTG로 유도하였다. 세포를 원심분리에 의해 수거하고, 용해 완충제 (50 mM P04 완충제, 300 mM NaCl, pH 8.0)에서 재현탁하고, 초음파처리하고, 스피너서 (spin) 잔해를 제거하였다. 청정한 세포 용해물을 예비 평형화된 Ni-NTA 스피너 칼럼 (키아젠(Qiagen)) 상에 로딩하고, 30 mM 이미다졸을 함유하는 완충제로 세척하였다. 융합 단백질을 200 μ L의 300 mM 이미다졸로 용리시켰다. 후속적으로, 용리된 단백질을 30°C에서 1 μ g의 TEV 프로테아제 (시그마 알드리치)로 인큐베이션하였다. 샘플을 24시간이라는 기간에 걸쳐 다양한 시점에서 질량 분광측정법에 의해 분석하였다.

[0235] C-말단 글리코실화 부위가 부가된 부분적으로 정제된 글루카곤의 MALDI-TOF MS 분석은, 예상 Man₃GlcNAc₂ 당펩티드 (m/z 6283)와 일치하는 균주 (도 11A) GLY01.6, 예상 GlcNAcMan₃GlcNAc₂ 당펩티드 (m/z 6611)와 일치하는 균주 (도 11B) GLY02.2, 예상 GlcNAcMan₃GlcNAc₂ 당펩티드 (m/z 6488)와 일치하는 균주 (도 11C) GLY01.7, 및 예상 GalGlcNAcMan₃GlcNAc₂ 당펩티드 (m/z 6649)와 일치하는 균주 (도 11D) GLY04.3로부터의 것이다. 별표는 글리코실트랜스퍼라제와 관계없이 모든 샘플에서 존재하는 배경 신호를 나타낸다.

[0236] **실시예 11 - *E. 콜라이* 중 만노실화 백신 제조**

[0237] **후보 항원로서의 유전자 클로닝 및 발현.** 글리코실화 연구를 위한 제조에서 후보 항원의 성공적인 발현은 단백질이 수용자 아스파라긴을 코딩하고 주변세포질에서 발현되는 것이 요구된다. 박테리아 OST PglB에 대해 최적

화된 N-글리코실화 시퀀스의 4개의 반복을 함유하는 GlycTag을 사용하였다. Sec 경로를 통해 방출되며 글리코실화를 위한 이소성 단백질의 방출에서 잘 작용하는 E. 콜라이 말토스 결합 단백질 (MBP)부터의 신호 펩티드를 이용하였다 ([Fisher, A.C., et al., *Production of Secretory and Extracellular N-Linked Glycoproteins in Escherichia coli*. Applied and Environmental Microbiology, 2011. 77(3): p. 871-881]). 글리코실화 연구 ([Fisher et al.])에 이용하기 위해 단백질을 이소프로필-β-D-티오갈락토피라노시드 (IPTG)-유도성 TRC 프로모터로부터 발현시켜 적절한 발현 수준을 제공하였다.

[0238]

E. 콜라이 백신 항원의 만노실화. 병원성 ExPEC E. 콜라이로부터 후보 백신 항원 c1275 및 3473 (Moriel, et al., PNAS 2010)을 코딩하는 유전자를 융합체로서 성숙 MBP의 C-말단에 대해 클로닝하고, 이들의 C-말단에서 4개의 연속적인 글리코실화 시퀀스 (4xGlycTag) 및 헥사히스티딘 태그로 개질시켰다. DsbA로부터의 신호 펩티드를 이용하여 단백질을 주면세포질에 대해 표적화시켰다. 생성된 플라스미드를 전기천공에 의해 E. 콜라이 세포 (MC4100 ΔwaaLΔgmd::kan) 중 pMW07-YCG-PgIB와 쌍형성하였다. 접종 후, 각 균주의 배양물 10 L를 600 nm에서 3.0의 적절한 광학 밀도로 성장시키고, 0.2% (v/v) 아라비노스 및 1 mM IPTG의 첨가로 유도하였다. 이전과 같이 ([Valerrama-Rincon, et al.]), 당단백질을 ConA 친화성 크로마토그래피에 이어서 니켈 친화성 크로마토그래피로 분리하였다. 부분적으로 정제된 샘플을 항-헥사히스티딘 항체 및 ConA 렉틴을 사용하여 웨스턴 블롯 분석하였다 (도 12).

[0239]

아스파라긴-연결된 만노스-말단 글리칸으로의 항원 개질. C. 제주니 헵타사카라이드의 성공적인 부착 후, 만노스-말단 글리칸을 후보 항원에 부착시켰다. 소수-만노스 올리고사카라이드 구조는 정상 인간 N-글리칸으로서 존재하며, 현재 글리칸 자체가 인간에게 내성이 있음을 나타내는 인간 치료([Van Patten, S.M., et al., *Effect of mannose chain length on targeting of glucocerebrosidase for enzyme replacement therapy of Gaucher disease*. Glycobiology, 2007. 17(5): p. 467-478])에서 사용된다. 플라스미드 (pYCG-PgIB)는 OST PgIB 및 S. 세레비리아에로부터의 4개의 글리코실트랜스퍼라제: Alg13, Alg14, Alg1, 및 Alg2를 발현한다 ([Valderrama et al.]). 이들 단백질은 Man₃GlcNAc₂ 글리칸 및 그의 유도체의 합성 및 접합을 조정하며, 이들은 인간 복합 N-글리칸의 염기를 형성한다.

[0240]

C. 제주니 헵타사카라이드에 의한 글리코실화에 의해 확인된 후보 항원은 개별적으로 글리코실화 숙주 균주 MC4100 ΔwaaL gmd::kan에서 pYCG-PgIB와 공동 발현된다. 글리코실화 경로 효소 및 항원의 유도 이후, 세포를 용균시키고, 표적 단백질을 α-연결된 만노스 잔기에 결합하는 ConA 렉틴으로 분리시켰다. 엔지니어드 글리칸이 α-만노스 잔기로 종결되기 때문에, ConA로의 정제가 완전히 목적하는 글리칸으로 개질된 단백질의 회수에 유리하다. 니켈-친화성 크로마토그래피를 이용하여 만노실화 항원을 추가로 정제하고, 단백질의 일부를 PNGase F로 처리하여 글리칸을 절단 제거하였다. SDS-PAGE에 이어서 ConA 및 αHis 항체로의 면역블롯팅에 의한 분석은 예상 만노실화 단백질의 회수를 증명하였다. 부착된 글리칸이 Man₃GlcNAc₂라는 것을 확인하기 위하여, PNGase F-방출 글리칸에 대해 ([Valderrama-Rincon et al.])에 기재된 바와 같이 질량 분광측정법을 수행하였다. 이 과정은 균질한 박테리아-유도 만노실화 표적 항원을 산출할 것으로 예상되었다. 만노실화 항원을 렉틴 ConA으로의 웨스턴 블롯에 의해 검출하였고, 질량 분광측정법에 의해 글리칸 동일성을 확인하였다.

[0241]

만노실화 항원의 면역원성 평가. 만노실화 및 아글리코실화(aglycosylated) 백신 항원을 정제하고, 표적 항원의 면역원성 특성을 평가하였다. 마우스 모델을 이용하여, 만노실화 ExPEC 항원에 대한 체액 및 세포-매개 면역 반응 둘 다의 마커를 검사하였다.

[0242]

만노실화 항원의 정제. 만노실화 항원을 ConA 칼럼 상에서의 렉틴-친화성 크로마토그래피에 이어서 니켈 정제를 이용하여 정제하였다. 아글리코실화 항원을 니켈-친화성 크로마토그래피에 의해 유사하게 정제하였다. 은 염색에 의해 유사한 순도를 보장하도록 제제를 비교하고, 각 샘플에 대해 내독소 수준을 측정하였다. 유사한 순도의 아글리코실화 단백질 및 만노실화 단백질을 적합한 양으로 회수하여 면역원성 연구를 수행하였다.

[0243]

인간 골수양 (mDC)에 의한 만노실화 항원의 결합 시험. 만노실화 또는 아글리코실화 항원을 mDC로 인큐베이션하여 만노스 수용체에 대한 결합을 평가하였다 ([Wieser, A., et al., *A Multi-epitope Subunit Vaccine Conveys Protection against Extraintestinal Pathogenic Escherichia coli in Mice*. Infection and Immunity, 2010. 78(8): p. 3432-3442]). 세척 후, 세포를 고정시키고, 표면-결합된 항원을 유동 세포측정법을 이용하여 αHis-FITC 항체로 검출하였다. 유리 만노스 또는 만난과의 경쟁은 만노실화 항원의 특이적 결합에서 만노스 수용체의 역할을 입증하였다 ([Wieser et al.]). 이러한 단계는 마우스 모델에서 면역원성의 평가 전에 만노실화 항원의 예비 검증으로서 작용한다.

- [0244] **만노실화 항원의 면역원성 측정.** 만노실화 항원의 피하 투여 후 마우스의 면역 반응을 아글리코실화 대조군과 비교하여 평가하였다. 이 단계는 백신 후보에 대한 항원성의 인헨서로서 Man₃GlcNAc₂ 글리칸의 사용에 있어 중요한 검증이다. 6마리의 CD1 마우스 (찰스 리버 레보리토리즈(Charles River Laboratories))의 균을, 예를 들어 20 µg의 항원으로 제1일, 제21일, 및 제35일에 피하 면역화시켰다. CD1 마우스는 동일한 면역화 시간대를 이용한 ExPEC 백신 연구에 대해 폐혈증 모델로서 사전에 이용되었고 ([Moriel, D.G., et al.]), 따라서 이들 실험은 미래 시험감염 연구에 대한 토대를 마련하였다. 최종 면역화 2 주 후 수집된 혈청 및 ELISA를 이용하여 IgG 및 IgM의 역가를 포함한 체액 반응을 정량화 하였다 ([Park, S.-U., et al., *Immunization with a DNA vaccine cocktail induces a Th1 response and protects mice against Mycobacterium avium subsp. paratuberculosis challenge*. Vaccine, 2008. 26(34): p. 4329-4337]). GlycTag 및 6x-His 태그 둘 다를 보유하는 관련없는 대조군 단백질과 비교하여 항원-특이적 반응을 평가하였다. 체액 및 세포-매개 면역 둘 다 ExPEC 감염에 대해 저항하는 기능을 할 수 있으므로 ([Thumbikat, P., et al., *Antigen-Specific Responses Accelerate Bacterial Clearance in the Bladder*. The Journal of Immunology, 2006. 176(5): p. 3080-3086]. [Nallamsetty, S. and D. Waugh, *Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners*. Protein Expr Purif, 2006. 45(1): p. 175-82]), CD8+ T 세포 검정을 이용하여 세포성 반응을 정량화하였다 ([Sivick, K.E. and H.L.T. Mobley, *Waging War against Uropathogenic Escherichia coli: Winning Back the Urinary Tract*. Infection and Immunity, 2010. 78(2): p. 568-585]).
- [0245] 모델 항원을 정상 *E. 콜라이* 주변세포질 상주 MBP를 갖는 융합 단백질로서 구성하였다. N-말단 MBP 융합체는 적절한 폴딩을 촉진시킬 수 있으며, 이어서 글리코실화를 향상시킬 수 있는 세포질로부터 방출될 수 있다 ([Nallamsetty et al.]). 신호 펩티드 서열이 교대로 존재하는 시험은 부적절한 국제화를 설명할 수 있다. 말단 GlycTag에 의해 개질된 단백질에 대한 *C. 제주니* 헵타사카라이드의 부착은, 충분한 표적 단백질이 주변세포질에 대해 적절히 국제화되며 Man₃GlcNAc₂ 글리칸으로의 글리코실화 성공에 대해 예상 지시약으로서 작용하는 모든 경우에서 확실히 달성된다. 그러나, 글리코실화는 GlycTag의 위치를 조정함으로써, 또는 상이한 만노스-말단 글리칸, 예컨대 박테리아 N-글리코실화 반응에서 단백질에 사전에 접합된 *E. 콜라이* 09로부터의 폴리-만노스 LPS를 사용함으로써 개선될 수도 있다 ([Wacker, M., et al., *Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems*. Proceedings of the National Academy of Sciences, 2006. 103(18): p. 7088-7093]).
- [0246] mDC에 대한 항원 결합에서의 만노실화의 효과 및 면역원성을 마우스 모델을 이용하여 평가하였다. 항원 내재화의 동역학은 mDC 결합 및 표면-결합된 항원에 있어 본 출원인의 가시화 능력에 영향을 줄 것이며, 필요하다면 엔도솜 트래피킹(trafficking)의 억제제를 적용하거나 투과성 세포 내 내재화 항원을 평가한다. 세포-매개 및 항원-특이적 체액성 면역 반응의 정량화를 이용하여 백신 항원 후보의 만노실화가 이들 면역원성 지시약에 영향을 주는지 측정하였다. *E. 콜라이* 09로부터의 폴리만노스 LPS와 같은 교대로 나타나는 만노스글리칸으로 글리코실화된 다양한 항원을 평가하였다. 대안적으로, 항원은 추가의 글리코실화 부위에 의해 개질되어 다중 글리칸의 부착을 촉진시킬 수 있다.

[0247]

[표 3] 균주 및 플라스미드 목록.

균주 명칭	플라스미드 1	플라스미드 2	플라스미드 3	이. 콜라이 균주	생성물
GLY01	pMW07-YCG-PglB.CO	—	—	MC4100Δ Δ <i>ΔwaaL</i> <i>gmd::kan</i>	Man ₃ GlcNAc ₂
GLY01.1	YCG-MBP-NtGnTI-PglB.CO	—	—	MC4100Δ Δ <i>ΔwaaL</i> <i>gmd::kan</i>	GlcNAcMan ₃ GlcNAc ₂
GLY01.2	pMW07-YCG-PglB.CO	pMQ70-ManC/B	—	MC4100Δ Δ <i>ΔwaaL</i> <i>gmd::kan</i>	Man ₃ GlcNAc ₂
GLY01.3	YCG-MBP-NtGnTI-PglB.CO	pTrc99Y-GlmS	—	MC4100Δ Δ <i>ΔwaaL</i> <i>gmd::kan</i>	GlcNAcMan ₃ GlcNAc ₂
GLY01.4	pMW07-YCG-PglB.CO	pMQ70-ManC/B	—	Origami2 <i>gmd::kan</i>	GlcNAcMan ₃ GlcNAc ₂

[0248]

GLY01.5	YCG-MBP-NtGnTI-PglB.CO	pMQ70-ManC/B	—	Origami2 <i>gmd::kan</i>	GlcNAcMan ₃ GlcNAc ₂
GLY01.6	pMW07-YCG-PglB.CO	pTrc99Y-MBP-글루카곤-ManC/B	—	Origami2 <i>gmd::kan</i> Δ <i>waaL</i>	Man ₃ GlcNAc ₂ -글루카곤
GLY01.7	YCG-MBP-NtGnTI-PglB.CO	pTrc99Y-MBP-글루카곤-ManC/B	—	Origami2 <i>gmd::kan</i> Δ <i>waaL</i>	GlcNAcMan ₃ GlcNAc ₂ -글루카곤
GLY02	pMW07-YCG-GST-Alg11-PglB.CO	—	—	MC4100Δ Δ <i>ΔwaaL</i> <i>gmd::kan</i>	Man ₅ GlcNAc ₂
GLY02.1	pMW07-YCG-GST-Alg11-PglB.CO	pMQ70-ManC/B	—	Origami2 <i>gmd::kan</i>	Man ₅ GlcNAc ₂
GLY02.2	pMW07-YCG-GST-Alg11-PglB.CO	pTrc99Y-MBP-글루카곤-ManC/B	—	Origami2 <i>gmd::kan</i> Δ <i>waaL</i>	Man ₅ GlcNAc ₂ -글루카곤
GLY03	pMW07-YCG-PglB.CO	pMQ70-MBP-NtGnTI	—	MC4100Δ Δ <i>ΔwaaL</i> <i>gmd::kan</i>	GlcNAcMan ₃ GlcNAc ₂
GLY03.1	pMW07-YCG-PglB.CO	pMQ70-MBP-NtGnTI	—	Origami2 <i>gmd::kan</i>	GlcNAcMan ₃ GlcNAc ₂
GLY04	YCG-MBP-NtGnTI-PglB.CO	pMQ70-HpGalT	—	MC4100Δ Δ <i>ΔwaaL</i> <i>gmd::kan</i>	GalGlcNAc-Man ₃ GlcNAc ₂

[0249]

GLY04.1	YCG-MBP-NtGnTI-PglB.CO	pMQ70-HpGalT	—	Origami2 <i>gmd::kan</i>	GalGlcNAc-Man ₃ GlcNAc ₂
GLY04.2	YCG-MBP-NtGnTI-PglB.CO	pMQ70-MBP-hGnTII	pMQ132-HpGalT	Origami2 <i>gmd::kan</i>	Gal ₂ GlcNAc ₂ -Man ₃ GlcNAc ₂
GLY04.3	YCG-MBP-NtGnTI-HpGalT-PglB.CO	pTrc99Y-MBP-글루카곤-ManC/B	—	Origami2 <i>gmd::kan</i> <i>ΔwaaL</i>	GalGlcNAc-Man ₃ GlcNAc ₂ -글루카곤
GLY05	YCG-MBP-NtGnTI-PglB.CO	pMQ70-MBP-bGnTIV	—	Origami2 <i>gmd::kan</i>	GlcNAc ₂ -Man ₃ GlcNAc ₂
GLY06	YCG-MBP-NtGnTI-PglB.CO	pMQ70-MBP-hGnTII	—	Origami2 <i>gmd::kan</i>	GlcNAc ₂ Man ₃ GlcNAc ₂
GLY06.1	YCG-MBP-NtGnTI-PglB.CO	pMQ70-MBP-hGnTII-bGnTIV	—	Origami2 <i>gmd::kan</i>	GlcNAc ₃ Man ₃ GlcNAc ₂

[0250]

[0251]

바람직한 실시양태가 본원에서 상세하게 설명 및 기재되어 있으나, 당업자에게는 다양한 변형, 첨가, 치환 등이 본 발명의 취지를 벗어나지 않으면서 이루어 질 수 있으며, 따라서 이들은 하기 청구범위에 정의된 바와 같은 본 발명의 범주 내에서 고려될 수 있다는 것이 명백할 것이다.

[0252]

비공식적 서열 목록

[0253]

서열 1 *alg13* 코돈 최적화

```

ATGGGTATCATCGAAGAAAAAGCTCTGTTTCGTTACCTGCGGT
GCTACCGTTCGGTTCGCGAAACTGGTTTCTTGCCTTCTGTCTGACGAATC
TGCCAGGAAGTATCCAGTACGGTTTCGTTTCGTTCTGATCATCCAGTTCGGT
CGTAACTACTCTTCTGAATTCGAACACCTGGTTCAGGAACGTGGTGGTCA
GCGTGAATCTCAGAAAATCCCGATCGACCAGTTCGGTTGCGGTGACACCG
CTCGTCAGTACGTTCTGATGAACGGTAAACTGAAAGTTATCGGTTTCGACT
TCTCTACCAAAATGCAGTCTATCATCCGTGACTACTCTGACCTGGTTATCT
CTCACGCTGGTACCGTTCTATCTGGACTCTCTGCGTCTGAACAAACCGC
TGATCGTTTTCGTTAACGACTCTCTGATGGACAACCACCAGCAGCAGATC
GCTGACAAATTCGTTGAACTGGGTTACGTTTGGTCTTTCGCTCCGACCGAA
ACCGGTCTGATCGCTGGTCTGCGTGCTTCTCAGACCGAAAAACTGAAACC
GTTCCCGGTTTCTCACAACCCGTTTTCGAACGCTGCTGGTTGAAACCAT
CTACTCTAA
    
```

[0254]

[0255]

서열 2 *alg13*

```

MGIIEEKALFVTCGATVFPFKLVSCVLSDFCQELIQYGFVRLIIQFG
RNYSSFEHLVQERGGQRESQKIPIDQFGCGDTARQYVLMNGKLVIGDFST
KMQSIIRDYSDLVISHAGTGSILDSLRLNKPLIVCVNDSLMDNHQQIADKQV
ELGYVWSCAPTETGLIAGLRASQTEKLPFPVSHNPSFERLLVETIYS*
    
```

[0256]

[0257]

서열 3 alg14 코돈 최적화

ATGAAAACCGCTTACCTGGCTTCTCTGGTTCTGATCGTTTCT
ACCGCTTACGTTATCCGTCTGATCGCTATCCTGCCGTCTTCCACACCCAG
GCTGGTACCGAAAAAGACACCAAAGACGGTGTTAACCTGCTGAAAATCCG
TAAATCTTCTAAAAAACCGCTGAAAATCTTCGTTTTCTGGGTCTGGTGG
TCACACCGGTGAAATGATCCGTCTGCTGGAAAACCTACCAGGACCTGCTGC
TGGGTAATCTATCGTTTACCTGGGTTACTCTGACGAAGCTTCTCGTCAGC
GTTTCGCTCACTTCATCAAAAAATTCGGTCACTGCAAAGTTAAATACTACG
AATTCATGAAAGCTCGTGAAGTTAAAGCTACCCTGCTGCAGTCTGTAAA
ACCATCATCGGTACCCTGGTTCAGTCTTTCGTTTACGTTGTTTCGTATCCGTT
TCGCTATGTGCGGTTCTCCGCACCTGTTCCCTGCTGAACGGTCCGGGTACCT
GCTGCATCATCTTTCTGGCTGAAAATCATGGAACCTGCTGCTGCCGCTGC
TGGGTTCTTCTCACATCGTTTACGTTGAATCTCTGGCTCGTATCAACACCC
CGTCTCTGACCGGTAAAATCCTGTAAGTGGTGGTGGTACGAATTCATCGTTC
AGTGGCAGGAACCTGCGTGACAACCTGCCGCGTTCTAAATGGTTCGGT
ATCCTGGTTTAA.

[0258]

[0259]

서열 4 alg14

MKTAYLASLVLIVSTAYVIRLIALPFHTQAGTEKDTKDGVNLLKI
RKSSKKPLKIFVFLGSGGHTGEMIRLLENYQDLLGKSIVYLGYSDEASRQRF
AHFIKKFGHCKVKYYEFMKAREVKATLLQSVKTIIGTLVQSFVHVVRIRFAM
CGSPHLFLLNGPGTCCIIISFWLKIMELLPLLGSSHIVYVESLARINTPSLTGKIL
YWVVDEFIVQWQELRDNYLPRSKWFGILV*

[0260]

[0261]

서열 5 alg1 코돈 최적화

ATGTTCTCTGGAAATCCCGCGTTGGCTGCTGGCTCTGATCATC
 CTGTACCTGTCTATCCCGCTGGTTGTTACTACGTTATCCCGTACCTGTTCT
 ACGGTAACAAATCTACCAAAAAACGTATCATCATCTTCGTTCTGGGTGAC
 GTTGGTCACTCTCCGCGTATCTGCTACCACGCTATCTTTCTCTAAACTG
 GGTTGGCAGGTTGAACTGTGCGGTTACGTTGAAGACACCCTGCCGAAAAT
 CATCTCTTCTGACCCGAACATCACCGTTCACCACATGTCTAACCTGAAACG
 TAAAGGTGGTGGTACCTCTGTTATCTTCATGGTTAAAAAAGTTCTGTTCCA
 GGTTCTGTCTATCTTCAAACCTGCTGTGGGAACTGCGTGGTTCTGACTACAT
 CCTGGTTCAGAACCCGCCGTCTATCCCGATCCTGCCGATCGCTGTTCTGTA
 CAAACTGACCGGTTGCAAACCTGATCATCGACTGGCACAACCTGGCTTACT
 CTATCCTGCAGCTGAAATTCAAAGGTAACCTTCTACCACCCGCTGGTTCTGA
 TCTCTTACATGGTTGAAATGATCTTCTCTAAATTCGCTGACTACAACCTGA
 CCGTTACCGAAGCTATGCGTAAATACCTGATCCAGTCTTTCACCTGAACC
 CGAAACGTTGCGCTGTTCTGTACGACCGTCCGGCTTCTCAGTTCAGCCGC
 TGGCTGGTGACATCTCTCGTCAGAAAGCTCTGACCACCAAAGCTTTCATCA
 AAAACTACATCCGTGACGACTTCGACACCGAAAAAGGTGACAAAATCATC
 GTTACCTCTACCTCTTTCACCCCGACGAAGACATCGGTATCCTGCTGGGT
 GCTCTGAAAATCTACGAAAACCTTACGTTAAATTCGACTCTTCTCTGCCG
 AAAATCCTGTGCTTCATCACCGTAAAGTCCGCTGAAAGAAAAATACAT
 GAAACAGGTTGAAGAATACGACTGGAAACGTTGCCAGATCGAATTCGTTT
 GGCTGTCTGCTGAAGACTACCCGAAACTGCTGCAGCTGTGCGACTACGGT
 GTTTCTCTGCACACCTCTTCTTCTGGTCTGGACCTGCCGATGAAAATCCTG
 GACATGTTCCGGTCTGGTCTGCCGTTATCGCTATGAACTACCCGGTTCTG
 GACGAACTGGTTCAGCACAACGTTAACGGTCTGAAATTCGTTGACCGTCG
 TGAAGTGCACGAATCTCTGATCTTCGCTATGAAAGACGCTGACCTGTACC
 AGAAACTGAAAAAAAACGTTACCCAGGAAGCTGAAAACCGTTGGCAGTC
 TAACTGGGAACGTACCATGCGTGACCTGAAACTGATCCACTAA.

[0262]

[0263]

서열 6 alg1

MFLEIPRWLLALILYLSIPLVYYYYVIPYLFYGNKSTKKRIIFVLGDV
 GHSPRICYHAISFSKLGWQVELCGYVEDTLPKIISDPNITVHHMSNLKRKGG
 GTSVIFMVKVLFQVLSIFKLLWELRGSYILVQNPPSIPILPIAVLYKLTGCKL
 IIDWHNLAYSILQLKFKGNFYHPLVLISYVEMIFSKFADYNLTVTEAMRKYL
 IQSFHLNPKRCAVLYDRPASQFQPLAGDISRQKALTTKAFIKNYIRDDFDTEK
 GDKIIVTSTSTPDEDIGILLGALKIYENSYVKFDSSLPKILCFITGKGPLKEKYM
 KQVEEYDWRKRCQIEFVWLSAEDYPKLLQLCDYGVSLHTSSSGLDLPMKILD
 MFGSGLPVIAMNYPVLDLQHNVNGLKFVDRRELHESLIFAMKDADLYQK
 LKKNVTQEAENRWQSNWERTMRDLKLIH*

[0264]

[0265]

서열 7 alg2 코돈 최적화

ATGATCGAAAAAGACAAACGTACCATCGCTTTCATCCACCC
 GGACCTGGGTATCGGTGGTGTGAACGTCTGGTTGTTGACGCTGCTCTGG
 GTCTGCAGCAGCAGGGTCACTCTGTTATCATCTACACCTCTCACTGCGACA
 AATCTCACTGCTTCGAAGAAGTTAAAAACGGTCAGCTGAAAGTTGAAGTT
 TACGGTGACTTCCTGCCGACCAACTTCTGGGTCTGTTTCTTCATCGTTTTCG
 CTACCATCCGTCAGCTGTACCTGGTTATCCAGCTGATCCTGCAGAAAAAA
 GTTAACGCTTACCAGCTGATCATCATCGACCAGCTGTCTACCTGCATCCCG
 CTGCTGCACATCTTCTTCTGCTACCCTGATGTTCTACTGCCACTTCCCGG
 ACCAGCTGCTGGCTCAGCGTGTGGTCTGCTGAAAAAATCTACCGTCTG
 CCGTTCGACCTGATCGAACAGTTCTCTGTTTCTGCTGCTGACACCGTTGTT
 GTTAACTCTAACTTACCAAAAAACACCTTCCACCAGACCTTCAAATACCTG
 TCTAACGACCCGGACGTTATCTACCCGTGCGTTGACCTGTCTACCATCGAA
 ATCGAAGACATCGACAAAAAATTCTTCAAACCGTTTTCAACGAAGGTGA
 CCGTTTCTACCTGTCTATCAACCGTTTCGAAAAAAAAAAGACGTTGCTCT
 GGCTATCAAAGCTTTCGCTCTGTCTGAAGACCAGATCAACGACAACGTAA

[0266]

AACTGGTTATCTGCGGTGGTTACGACGAACGTGTTGCTGAAAACGTTGAA
 TACCTGAAAGAAGTGCAGTCTCTGGCTGACGAATACGAACTGTCTCACAC
 CACCATCTACTACCAGGAAATCAAACGTGTTTCTGACCTGGAATCTTTCAA
 AACCAACAACCTCTAAAATCATCTTCTGACCTCTATCTTCTTCTCTGAA
 AGAACTGCTGCTGGAACGTACCGAAATGCTGCTGTACACCCCGGCTTACG
 AACACTTCGGTATCGTTCCGCTGGAAGCTATGAAACTGGGTAAACCGGTT
 CTGGCTGTTAACAACGGTGGTCCGCTGGAAACCATCAAATCTTACGTTGCT
 GGTGAAAACGAATCTTCTGCTACCGTTGGCTGAAACCGGCTGTTCCGAT
 CCAGTGGGCTACCGCTATCGACGAATCTCGTAAAAATCCTGCAGAACGGTT
 CTGTTAACTTGAACGTAACGGTCCGCTGCGTGTTAAAAAATACTTCTCTC
 GTGAAGCTATGACCCAGTCTTTCGAAGAAAACGTTGAAAAAGTTATCTGG
 AAAGAAAAAATACTACCCGTGGGAAATCTTCGGTATCTTCTTCTCTAA
 CTTTCATCTGCACATGGCTTTCATCAAAATCCTGCCGAACAACCCGTGGCC
 GTTCTGTTTCAATGGCTACCTTTCATGGTTCTGTACTTCAAAAACTACCTGTG
 GGGTATCTACTGGGCTTTCGTTTTTCGCTCTGTCTTACCCGTACGAAGAAAT
 CTAA

[0267]

[0268]

서열 8 alg2

MIEKDKRTIAFIHPDLGIGGAERLVVDAALGLQQGHSVIIYTSHCD
 KSHCFEEVKNQQLKVEVYGDFLPTNFLGRFFIVFATIRQLYLVIQLILQKKN
 AYQLIIIDQLSTCIPLLHIFSSATLMFYCHFPDQLLAQRAGLLKKIYRPLPDLIEQ
 FSVSAADTVVNSNFTKNTFHQTFKYLSNDPDIYPCVDLSTIEIEDIDKFFK
 TVFNEGDRFYLSINRFEKKDVALAIKAFALSEDQINDNVKLVICGGYDERVA
 ENVEYLKELQSLADEYELSHTTIYYQEIKRVSDLESFKTNNSKIIFLTSISSSLKE
 LLLERTEMLLYTPAYEHFIVPLEAMKLGKPVLAVNNGGPLETIKSYVAGEN
 ESSATGWLKPAPIQWATAIDESRKILQNGSVNFERNGPLRVKKYFSREAMT
 QSFEENVEKVIWKEKKYYPWEIFGISFSNFILHMAFIKILPNNPWPFLEMATFM
 VLYFKNYLWGIYWAFVVALSYPIEIEI*

[0269]

[0270]

서열 9 alg11

ATGGGCAGTGCTTGGACAAACTACAATTTTGAAGAGGTAA
 GTCTCATTGTTGGTTCAAAAAATATGTTGTATCATCTTTAGTACTAGTGTA
 TGGACTAATTAAGGTTCTCACGTGGATCTCCGTCAATGGGTGTATTCCAG
 CTTGAATCCGTTCTCCAAAAATCTTCATTAAGTGAACAGAGCAGTTGCCTC

[0271]

CTGTGGTGAGAAGAATGTGAAAGTTTTTGGTTTTTTTCATCCGTATTGTAA
 TGCTGGTGGTGGTGGGGAAAAAGTGCTCTGGAAAGCTGTAGATATCACTT
 TGAGAAAAGATGCTAAGAACGTTATTGTCATTTATTCAGGGGATTTTGTG
 AATGGAGAGAATGTTACTCCGGAGAATATTCTAAATAATGTGAAAGCGAA
 GTTCGATTACGACTTGGATTCCGATAGAATATTTTTCATTCATTGAAGCT
 AAGATACTTGGTGGATTCTCAACATGGAAGCATTTCACGTTGATTGGAC
 AAGCAATTGGATCAATGATTCTCGCATTGAAATCCATTATTCAGTGTCCAC
 CTGATATATGGATTGATACAATGGGGTACCCTTTCAGCTATCCTATTATTG
 CTAGGTTTTTGAGGAGAATTCCTATCGTCACATATACGCATTATCCGATAA
 TGTCAAAGACATGTTAAATAAGCTGTTCAAATGCCCAAGAAGGGTATC
 AAAGTTTACGGTAAAATATTATACTGGAAGTTTTTATGTTAATTTATCAA
 TCCATTGGTTCTAAAATTGATATTGTAATCACAACCTCAACATGGACAAAT
 AACCACATAAAGCAAATTTGGCAATCCAATACGTGTAATAATTATATATCC
 TCCATGCTCTACTGAGAAATTAGTAGATTGGAAGCAAAGTTTGGTACTG
 CAAAGGGTGAGAGATTAATCAAGCAATTGTGTTGGCACAATTTTCGTCCT
 GAGAAACGTCATAAGTTAATCATTGAGTCCTTTGCAACTTTCTTGAAAAAT
 TTACCGGATTCTGTATCGCCAATTAATGATAATGGCGGGTCCACTAG
 ATCCAAGCAAGATGAAAATTATGTTAAAAGTTTACAAGACTGGTCAGAAA
 ATGTATTAATAAATTCCTAAACATTGATATCATTGAAAAAATCTGCCCT
 TCGATAAGATTGAAATATTACTAAACAAATCTACTTTCGGTGTAAATGCCA
 TGTGGAATGAGCACTTTGGAATTGCAGTTGTAGAGTATATGGCTTCCGGTT
 TGATCCCCATAGTTCATGCCTCGGCGGGCCATTGTTAGATATAGTTACTC
 CATGGGATGCCAACGGGAATATCGGAAAAGCTCCACCACAATGGGAGTT
 ACAAAGAAATATTTGCAAACTCGAAGATGATGGTGAACCTACTGGAT
 TTTCTTTAAAGAGCCGAGTGATCCTGATTATAACACAACCAAAGATCCTC
 TGAGATACCCTAATTTGTCCGACCTTTTCTTACAAATTACGAACTGGACT
 ATGACTGCCTAAGGGTGATGGGCGCAAGAAACCAGCAGTATTCATTGTAT
 AAATCTCTGATTTGAAGTTGATAAAGATTGGGAAAACCTTGTACTGAAT
 CCTATTTGTAAATTATTAGAAGAGGAGGAAAGGGGCTGA

[0272]

[0273]

서열 10 Alg11 단백질

MGSAWTNYNFEEVKSHFGFKKYVVSSLVLVYGLIKVLTWIFR
 QWVYSSLNPFSSKSSLLNRAVASCGEKNVVFVGFHPYCNAGGGGKVLWK
 AVDITLRKDAKNVIVISGDFVNGENVTPENILNNVAKAFDYDLSDRIFLISL
 KLRYLVDSSTWKHFTLIGQAIGSMILAFESIQCPPDIWIDTMGYPFSPHARFL
 RRIPIVYTYTHYPIMSKDMLNKLKMPKKGIVYVKILYWKVFMLIYQSIGSKI
 DIVITNSTWTNNHIKQIWQSNCTKIIYPPCSTEKLVWQKQKFGTAKGERLNQA
 IVLAQFRPEKRHKLIIESFATFLKNLPDSVSPIKLIMAGSTRSKQDENYVKSQD
 WSENVLKIPKHLISFEKNLFPDKIEILLNKSTFGVNMWNEHFGIAVVEYMAS
 GLIPIVHASAGPLLDIVTPWDANGNIGKAPPQWELQKKYFAKLEDDGETTGFF
 FKEPSDPDYNTTKDPLRYPNLSDLFLQITKLDYDCLRVMGARNQQYSLYKFS
 DLKFDKDWENFVLNPICKLLEEEERG*

[0274]

[0275]

서열 11 malE (MBP)

AAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGA
 TAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATA
 CCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTC
 CCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACA
 CGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCC
 GGACAAAGCGTTCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTAC
 GTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGC
 TGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAG
 ATCCCGGCGCTGGATAAAGAAGTGAAGCGAAAGGTAAGAGCGCGCTGA
 TGTTCAACCTGCAAGAACCCTACTTACCTGGCCGCTGATTGCTGCTGACG
 GGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGTG
 GCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCT
 GATTAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAG
 CTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCCGTGGGCA
 TGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAAACGGTACTGCC
 GACCTTCAAGGGTCAACCATCAAACCGTTTCGTTGGCGTGCTGAGCGCAG
 GTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCGAAAGAGTTCCTCGAA
 AACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACC
 GCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATC
 CACGTATTGCCGCCACCATGGAAAACGCCAGAAAGGTGAAATCATGCCG
 AACATCCCGCAGATGTCCGCTTCTGGTATGCCGTGCGTACTGCGGTGATC
 AACGCCGCCAGCGGTCTCAGACTGTTCGATGAAGCCCTGAAAGACGCGCA
 GACTCGTATCACCAAGTAA

[0276]

[0277]

서열 12 Ma1E 단백질 (MBP)

KIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKL
EEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWD
AVRYNGKLIAYPIAVEALSIIYNDLLPNPPKTWEEIPALDKELKAKGKSALM
FNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAGLTFLVDLI
KNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTF
KGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV
ALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGR
QTVDEALKDAQTRITK*

[0278]

[0279]

서열 13 mstX

ATGTTTTGTACATTTTTTGAAAAACATCACCGGAAGTGGGAC
ATACTGTTAGAAAAAGCACGGGTGTGATGGAAGCTATGAAAGTGACGA
GTGAGGAAAAGGAACAGCTGAGCACAGCAATCGACCGAATGAATGAAGG
ACTGGACGCGTTTATCCAGCTGTATAATGAATCGGAAATTGATGAACCGC
TTATTCAGCTTGATGATGATACAGCCGAGTTAATGAAGCAGGCCCGAGAT
ATGTACGGCCAGGAAAAGCTAAATGAGAAATTAATAACAATTATTAACA
GATTTTATCCATCTCAGTATCTGAAGAAGGAGAAAAAGAA

[0280]

[0281]

서열 14 MstX 단백질

MFCTFFEKHHRKWDILLEKSTGVMEAMKVTSEEKEQLSTAIRMN
EGLDAFIQLYNESEIDEPLIQLDDDTAELMKQARDMYGQEKLNEKLNTHIKQIL
SISVSEEGEKE*

[0282]

[0283]

서열 15 GnTI (EC 2.4.1.101)

GCGACACAGTCAGAATATGCAGATCGCCTTGCTGCTGCAAT
 TGAAGCAGAAAATCATTGTACAAGCCAGACCAGATTGCTTATTGACCAGA
 TTAGCCTGCAGCAAGGAAGAATAGTTGCTCTTGAAGAACAAATGAAGCGT
 CAGGACCAGGAGTGCCGACAATTAAGGGCTCTTGTTTCAGGATCTTGAAAG
 TAAGGGCATAAAAAAGTTGATCGGAAATGTACAGATGCCAGTGCGCTGCTG
 TAGTTGTTATGGCTTGAATCGGGCTGATTACCTGGAAAAGACTATTAAT
 CCATCTTAAAATACCAAATATCTGTTGCGTCAAAATATCCTCTTTTCATAT
 CCCAGGATGGATCACATCCTGATGTCAGGAAGCTTGCTTTGAGCTATGAT
 CAGCTGACGTATATGCAGCACTTGGATTTTGAACCTGTGCATACTGAAAG
 ACCAGGGGAGCTGATTGCATACTACAAAATTGCACGTCATTACAAGTGGG
 CATTGGATCAGCTGTTTTACAAGCATAATTTTAGCCGTGTTATCATACTAG
 AAGATGATATGGAAATTGCCCTGATTTTTTTGACTTTTTTTGAGGCTGGAG
 CTACTCTTCTTGACAGAGACAAGTCGATTATGGCTATTTCTTCTTGAATG
 ACAATGGACAAATGCAGTTTGCCAAGATCCTTATGCTCTTACCGCTCAG
 ATTTTTTTCCCGGTCTTGGATGGATGCTTTCAAATCTACTGGGACGAAT
 TATCTCCAAAGTGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAA
 GAGAATCACAGAGGTCGACAATTTATTCGCCCAGAAGTTTGCAGAACATA
 TAATTTTGGTGAGCATGGTTCTAGTTTGGGGCAGTTTTTCAAGCAGTATCT
 TGAGCCAATTAACATAAATGATGTCCAGGTTGATTGGAAGTCAATGGACC
 TTAGTTACCTTTTGGAGGACAATTACGTGAAACACTTTGGTGACTTGGTTA
 AAAAGGCTAAGCCCATCCATGGAGCTGATGCTGTCTTGAAAGCATTTAAC
 ATAGATGGTGATGTGCGTATTCAGTACAGAGATCAACTAGACTTTGAAAA
 TATCGCACGGCAATTTGGCATTTTTGAAGAATGGAAGGATGGTGTACCAC
 GTGCAGCATATAAAGGAATAGTAGTTTTCCGGTACCAAACGTCCAGACGT
 GTATTCTTGTGGCCATGATTCGCTTCAACAACCTCGGAATTGAAGATACT
 TAA

[0284]

[0285]

서열 16 GnTI 단백질

ATQSEYADRLAAAIEAENHCTSQRLLIDQISLQQGRIVALEEQ
 MKRQDQECRQLRALVQDLESKGIKKLIGNVQMPVAAVVVMACNRADYLEK
 TIKSILKYQISVASKYPLFISQDGSHPDVRKLALSVDQLTYMQHLDPEPVHTER
 PGEIAYYKIARHYKWALDQLFYKHNFSRVIILEDDMEIAPDFDFFEAGATL
 LDRDKSIMAISSWNDNGMQMFVQDPYALYRSDFPGLGWMLSKSTWDELSP
 KWPKAYWDDWLRLENHRGRQFIRPEVCRTYNFGEHGSSLGQFFKQYLEPI
 KLNDVQVDWKSMDLSYLLEDNYVKHFGDLVKKAKPIHGADAVLKAFNIDG
 DVRIQYRDQLDFENIARQFGIFEWKDGVPRAAAYKGVVFRYQTSRRVFLVG
 HDSLQQLGIEDT*

[0286]

[0287]

서열 17 GnT II (EC 2.4.1.143)

ATGCGCTTTCGTATCTATAAACGTAAGTGCTGATCCTGACA
 CTGGTTGTTGCCGCTTGTGGTTTTGTTCTGTGGAGCAGTAATGGTCGTCAG
 CGTAAAAATGAAGCCCTGGCACCTCCTCTGCTGGATGCTGAACCGGCACG
 TGGTGCTGGCGGTCGTGGTGGTGATCATCCGTCTGTTGCCGTTGGTATTCCG
 TCGTGTGAGCAATGTTTCGGCTGCCTCTCTGGTCCC GGCTGTTCCCTCAACC
 TGAAGCTGATAACCTGACCCTGCGCTATCGCTCTCTGGTGTATCAACTGAA
 CTTTCGATCAAACCTCTGCGTAACTGGATAAAGCAGGCACATGGGCTCCTC
 GTGAACTGGTACTGGTAGTCCAGGTCCATAATCGTCCGGAATATCTGCGT
 CTGCTGCTGGATTCTCTGCGCAAAGCTCAAGGCATCGATAATGTCCTGGTC
 ATCTTCTCTCATGATTTCTGGAGCACGGAGATTAACCAGCTGATTGCCGGC
 GTGAATTTTTGTCTGTGCTGCAGGTGTTTTTCCGTTTTCTATCCAACGT
 ATCCGAACGAATTTCCGGGTTCTGATCCTCGTGATTGTCCTCGTGATCTGC
 CTA AAAATGCCGCTCTGAACTGGGCTGTATTAATGCCGAGTATCCTGATT
 CTTTTGGCCACTATCGTGAGGCGAAATTTCTCAGACCAAACATCATTGGT
 GGTGGAACTGCATTTTCGTGTGGGAACGTGTGAAAATCCTGCGCGACTAT
 GCTGGCCTGATTCTGTTTCTGGAAGAAGATCACTATCTGGCTCCGGACTTT
 TATCATGTGTTCAAAAAATGTGGAACTGAAACAGCAGGAATGTCCAGA
 ATGTGATGTGCTGTCACTGGGCACCTATAGTGCTTCTCGCTCCTTCTATGG
 TATGGCCGACAAAGTGGACGTTAAAACATGGAATCCACCGAGCACAAAC
 ATGGGTCTGGCACTGACTCGTAATGCCTATCAAAAACCTGATTGAGTGTAC
 CGACACCTTTTGTACGTATGATGACTATAACTGGGACTGGACCCTGCAAT
 ATCTGACCGTGAGCTGTCTGCCAAAATTTGGAAAGTTCTGGTGCCTCAGA
 TTCCTCGTATCTTTCATGCTGGCGACTGTGGTATGCACCATAAAAAA ACTT
 GCCGTCCGTCAACACAATCTGCTCAGATCGAGTCGCTGCTGAATAATAAC
 AAACAGTATATGTTCCCGGAGACTCTGACAATTTCTGAAAAATTCACCGT
 GGTCGCCATTTCTCCGCCTCGTAAAAATGGAGGTTGGGGCGATATCCGTG
 ACCATGAACTGTGTA AAAAGCTATCGTCGTCTGCAGTGA

[0288]

[0289]

서열 18 GnT II (EC 2.4.1.143)

MRFRIYKRKVLILTLVVAACGFVLWSSNGRQRKNEALAPPLD
 AEPARGAGGRGGDHPVAVGIRRVSNVSAASLVPAVPQPEADNLTLYRSLV
 YQLNFDQTLRNVDKAGTWA PRELVLVVQVHNRPEYLRLLLSLRKAQGIDN
 VLVIFSHDFWSTEINQLIAGVNFPCVLQVFFPFSIQLYPNEFPDPRDCPRDLP
 KNAALKLGCINA EYPDSFGHYREAKFSQTKHHWWWKLFHW ERVKILRDY
 AGLILFLEEDHYLAPDFYHVFKMWK LKQQECPECDVLSLGTYSASRSFYGM
 ADKVDVKTWKSTEHN MGLALTRNAYQKLECTDTFCTYDDYNWDWTLQYL
 TVSCLPKFWKVLVPQIPRIFHAGDCGMHKKTCRPSTQSAQIESLLNNNKQY
 MFPETLTISEKFTVVAISP PRKNGGWGDIRDHELCKSYRRLQ*

[0290]

[0291]

서열 19 GnTIV (EC 2.4.1.145)

TTGAAAGAACTGACGTCCAAAAAGAGCTTGCAAGTCCCGTCCATCTACTA
 TCCTTGCCGCACTGCTGCAAAACGAGGGCTCTTTGCAACCGGCAGTTC
 AGATCGGCAATGGTCGACCGGCGTGAGCATTGTTATGGGTATCCCGACC
 GTGAAACGTGAAGTGAAAAGCTATCTGATTGAAACGCTGCATAGCCTGAT
 CGATAACCTGTACCCGGAAGAAAAACTGGACTGCGTGATTGTCGTTTTCA
 TTGGTGAAACCGACACGGATTATGTGAATGGCGTTGTTGCCAATCTGGAA
 AAAGAGTTTCAGCAAAGAGATCAGCAGCGGCCTGGTTGAGATCATTCTCC
 GCCGGAGAGCTATTACCCGGATCTGACGAACCTGAAAGAAACCTTCGGTG
 ATAGCAAAGAGCGTGTCCGTTGGCGCACTAAGCAGAACCTGGACTATTGT
 TTTCTGATGATGTACGCGCAAGAAAAGGGTACGTATTACATCCAACTGGA
 GGACGACATTATTGTGAAGCAAACTACTTCAACACCATTAAGAACTTCG
 CGCTGCAGCTGAGCAGCGAAGAGTGGATGATTCTGGAGTTCAGCCAGCTG
 GGCTTCATTGGCAAGATGTTTCAGGCACCGGACTTGACCCTGATCGTGGA
 GTTTATCTTTATGTTCTACAAAGAGAAAACCGATCGATTGGCTGCTGGATCA
 TATCCTGTGGGTCAAGGTCTGCAATCCGAAAAAGATGCCAAGCATTGTG
 ACCGCCAGAAAAGCAATCTGCGTATTGTTTTCGTCCTAGCCTGTTCCAAC
 ACGTGGGTCTGCACAGCTCTCTGACCGTAAGATCCAAAAGCTGACCGAC
 AAAGATTACATGAAAACCGCTGCTGCTGAAGATCCATGTCAACCCGCCAGC
 AGAGGTGAGCACCTCGCTGAAAGTCTACCAGGGTCACACTCTGGAGAAAA
 CCTATATGGGCGAGGACTTCTTTTGGGCGATTACGCCTGTTGCGGGTGACT
 ATATCTTGTTAAGTTTGACAAGCCGGTTAATGTAGAGAGCTACTTGTTTC
 ATAGCGGTAACCAGGATCACCCAGGTGACATTCTGCTGAACACCACCGTT
 GAAGTGTGCCGCTGAAAAGCGAAGGTCTGGATATTCGAAAGAAACGA
 AGGATAAGCGTCTGGAGGATGGTTACTTCCGTATCGGCAAGTTCGAGAAT
 GGCGTGGCTGAAGGTATGGTCGACCCGAGCCTGAACCCGATTTCCGCATT
 TCGCCTGTCCGTCATCCAGAATAGCGCGGTTTGGGCTATCCTGAATGAGAT
 TCACATCAAAAAGGTTACGAATTA

[0292]

[0293]

서열 20 GnTIV (EC 2.4.1.145)

ILKELTSKKSQVPSIYYHLPHELLQNEGSLQPAVQIGNGRGTGVS
 VMGIPTVKREVKSYLETLHSLIDNLYPEEKLDVCVIVVFIGETDIDYVNGVVA
 NLEKEFSKEISSGLVEIISPESYYPDLTNLKETFGDSKERVRWRKQNLDYCF
 LMMYAQEKGTYYIQLEDDIIVKQNYFNTIKNFALQLSSEWMILEFSQLGFIG
 KMFQAPDLTLIVEFIFMFYKEKPIDWLLDHILWVKVCNPEKDAKHCDRQKAN
 LRIRFRPSLFQHVGLHSSLTGKIQLTDKDYMKPLLLKIHVNPPAEVSTSLKVY
 QGHTLEKTYMGEDFFWAITPVAGDYILFKFDKPVNVESYLFHSGNQDHPGDI
 LLNTTVEVLPLKSEGLDISKETKDKRLEDGYFRIGKFENGVAEGMVDPSLNPI
 SAFRLSVIQNSAVWAILNEIHIKVTN*

[0294]

[0295]

서열 21 GalT (EC 2.4.1.38)

ATGCGTGTCTTTATTATCAGTCTGAACCAGAAAGTGTGTGAC
AAATTCGGCCTGGTGTTCGTGATACCACAACCCTGCTGAATAACATCAAT
GCCACCCGCCACAAAGCACAGATTTTTGACGCCGTCTATAGCAAAACGTT
CGAAGGTGGGCTGCATCCACTGGTGAAAAAACATCTGCACCCGTATTCA
TTACCCAGAACATCAAAGACATGGGCATTACCACCAACCTGATTAGCGGT
GTATCCAAATTCTATTATGCTCTGAAATATCACGCCAAATTCATGAGCCTG
GGCGAACTGGGCTGTTATGCCAGCCATTATAGCCTGTGGGAGAAATGTAT
TGAGCTGAACGAGGCCATTTGTATCCTGGAAGATGACATTACGCTGAAAG
AAGATTTCAAAGAGGGCCTGGATTTCTGGAAAAACACATTCAGGAGCTG
GGCTATGTTTCGTCTGATGCATCTGCTGTATGATGCCTCCGTTAAAAGCGAA
CCTCTGTCCATAAAAACCACGAGATTCAGAGCGTGTCTGGGATCATTAA
AGCTTATAGTCACGGTGTGGCACTCAGGGATATGTGATTACTCCGAAAA
TTGCCAAAGTGTTCAAAAAATGCTCCCGTAAATGGGTTGTTCCGGTGGAT
ACGATCATGGATGCCACGTTTATTCATGGGGTGAAAAACCTGGTACTGCA
ACCGTTTGTGATTGCCGATGATGAGCAAATTTCCACGATTGTCCGTAAG
AGGAGCCGATTCCCCTAAAATTGCCCTGATGCGCGAACTGCACTTCAAA
TATCTGAAATATTGGCAGTTTGTGTGA

[0296]

[0297]

서열 22 GalT (EC 2.4.1.38)

MRVFIISLNQKVCDFGLVFRDTTTTLLNNINATRHKAQIFDAVY
SKTFEGLHPLVKKHLHPYFITQNIKDMGITNLISGVSKFYALVYHAKFMS
LGELGCYASHYSLWEKCIELNEAICILEDITLKEDFKEGLDFLEKHIQELGYV
RLMHLLYDASVKSEPLSHKNHEIQERVGIKAYSHGVTQGYVITPKIAKVKF
KCSRKWVVPVDTIMDATFIHGKLVLPFVIADDEQISTIVRKEEPYSPKIAL
MRELHFYLYKYWQFV*

[0298]

[0299]

서열 23 manB (EC 5.4.2.8)

ATGAAAAAATTAACCTGCTTTAAAGCCTATGATATTCGCGGGAA
 ATTAGGCGAAGAACTGAATGAAGATATCGCCTGGCGCATTGGTCGCGCCT
 ATGGCGAATTTCTCAAACCGAAAAACCATTGTGTTAGGCGGTGATGTCCGC
 CTCACCAGCGAAACCTTAAAACTGGCGCTGGCGAAAGGTTTACAGGATGC
 GGGCGTTGACGTGCTGGATATTGGTATGTCCGGCACCGAAGAGATCTATT
 TCGCCACGTTCCATCTCGGCGTGGATGGCGGCATTGAAGTTACCGCCAGC
 CATAATCCGATGGATTATAACGGCATGAAGCTGGTTTCGCGAGGGGGCTCG
 CCCGATCAGCGGAGATACCGGACTGCGCGACGTCCAGCGTCTGGCTGAAG
 CCAACGACTTTCTCCCCTCGATGAAACCAAACGCGGTCTGCTATCAGCAA
 ATCAACCTGCGTGACGCTTACGTTGATCACCTGTTTCGGTTATATCAATGTC
 AAAAACCTCACGCCGCTCAAGCTGGTGATCAACTCCGGGAACGGCGCAGC
 GGGTCCGGTGGTGGACGCCATTGAAGCCCGCTTTAAAGCCCTCGGCGCGC
 CCGTGGAATTAATCAAAGTGCACAACACGCCGGACGGCAATTTCCCAAC
 GGTATTCTTAACCCACTACTGCCGGAATGCCGCGACGACACCCGCAATGC
 GGTATCAAACACGGCGCGGATATGGGCATTGCTTTTGATGGCGATTTTG
 ACCGCTGTTTCTGTTTGACGAAAAAGGGCAGTTTATTGAGGGCTACTAC
 ATTGTCGGCCTGTTGGCAGAAGCATTCTCGAAAAAATCCCGGCGCGAA
 GATCATCCACGATCCACGTCTCTCCTGGAACACCGTTGATGTGGTACTGC
 CGCAGGTGGCAGCCGGTAATGTCGAAAACCGGACACGCCTTATTAAAG
 AACGTATGCGCAAGGAAGACGCCATCTATGGTGGCGAAATGAGCGCCCA
 CCATTACTTCCGTGATTTTCGCTTACTGCGACAGCGGCATGATCCCGTGGCT
 GCTGGTGCCCGAACTGGTGTGCCTGAAAGATAAAACGCTGGGCGAACTGG
 TACGCGACCGGATGGCGGCGTTTCCGGCAAGCGGTGAGATCAACAGCAA
 ACTGGCGCAACCCGTTGAGGCGATTAACCGCGTGGAACAGCATTTTAGCC
 GTGAGGCGCTGGCGGTGGATCGCACCGATGGCATCAGCATGACCTTTGCC
 GACTGGCGCTTTAACCTGCGCACCTCCAATACCGAACCGGTGGTGGCGCT
 GAATGTGGAATCGCGCGGTGATGTGCCGCTGATGGAAGCGCGAACGCGA
 ACTCTGCTGACGTTGCTGAACGAGTAA

[0300]

[0301]

서열 24 manB (EC 5.4.2.8)

MKKLTCFKAYDIRGKLGEEFNEDIAWRIGRAYGEFLKPKTIVL
 GGDVRLTSETLKLALAKGLQDAGVDVLDIGMSGTEEIYFATFHLGVDGGIEV
 TASHNPMDYNGMKLVREGARPISGDTGLRDVQRLAEANDFPPVDETKRGRY
 QQINLRDAYVDHLFGYINVKNLTPLKLVINSNGNAAGPVVDAIEARFKALGA
 PVELIKVHNTPDGNFPNGIPNLLPECRDDTRNAVIKHGADMGIAFDGFDRC
 FLFDEKQFIEGYYIVGLLAEAFLEKNPGAKIIHDPRLSWNTVDVVTAAGGTP
 VMSKTGHAFIKERMKEDAIYGGEMSAHHYFRDFAYCDSGMIPWLLVAELV
 CLKDKTLGELVRDRMAAFPASGEINSKLAQPVEAINRVEQHFSREALAVDRT
 DGISMTFADWRFNLRNTEPVVRLNVESRGDVPLMEARTRLLTLLNE*

[0302]

[0303]

서열 25 manC (EC 2.7.7.13)

ATGGCGCAGTCGAAACTCTATCCAGTTGTGATGGCAGGTGGCTCCGGTAG
 CCGCTTATGGCCGCTTTCCCGCTACTTTATCCCAAGCAGTTTTTATGCCT
 GAAAGGCGATCTCACCATGCTGCAAACCACCATCTGCCGCCTGAACGGCG
 TGGAGTGCGAAAGCCCGGTGGTGATTTGCAATGAGCAGCACCGCTTTATT
 GTCGCGGAACAGCTGCGTCAACTGAACAACTTACCGAGAACATTATTCT
 CGAACGGCAGGGCGAAACACGGCACCTGCCATTGCGCTGGCGGCGCTG
 GCGGCAAAACGTCATAGCCCGGAGAGCGACCCGTTAATGCTGGTATTGGC
 GGCGGATCATGTGATTGCCGATGAAGACGCGTTCCGTGCCGCCGTGCGTA
 ATGCCATGCCATATGCCGAAGCGGGCAAGCTGGTGACCTTCGGCATTGTG
 CCGGATCTACCAGAAACCGGTTATGGCTATATTCGTGCGGGTGAAGTGTC
 TGCGGGTGAGCAGGATATGGTGGCCTTTGAAGTGGCGCAGTTTGTGCGAAA
 AACCGAATCTGGAACCGCTCAGGCCTATGTGGCAAGCGGCGAATATTAC
 TGGAACAGCGGTATGTTCTGTTCCGCGCCGGACGCTATCTCGAAGAACT
 GAAAAAATATCGCCCGGATATCCTCGATGCCTGTGAAAAAGCGATGAGCG
 CCGTCGATCCGGATCTCAATTTTATTCGCGTGGATGAAGAAGCGTTTCTCG
 CCTGCCCGGAAGAGTCGGTGGATTACGCGGTCATGGAACGTACGGCAGAT
 GCTGTTGTGGTGCCGATGGATGCGGGCTGGAGCGATGTTGGCTCCTGGTC
 TTCATTATGGGAGATCAGCGCCCACACCGCCGAGGGCAACGTTTGCCACG
 GCGATGTGATTAATCACAAACTGAAAACAGCTATGTGTATGCTGAATCT
 GGCCTGGTCACCACCGTCGGGGTGAAAAGATCTGGTAGTGGTGCAGACCAA
 AGATGCGGTGCTGATTGCCGACCGTAACGCGGTACAGGATGTGAAAAAA
 GTGGTCGAGCAGATCAAAGCCGATGGTCGCCATGAGCATCGGGTGCATCG
 CGAAGTGTATCGTCCGTGGGGCAAATATGACTCTATCGACGCGGGCGACC
 GCTACCAGGTGAAACGCATCACCGTGAAACCGGGCGAGGGCTTGTGCGTA
 CAGATGCACCATCACCGCGCGGAACACTGGGTGGTTGTCGCGGGAACGGC
 AAAAGTCACCATTGATGGTGATATCAAACCTGTTGGTGAAAACGAGTCCA
 TTTATATTCCGCTGGGGGCGACGCATTGCCTGGAACCCGGGGAAAATT
 CCGCTCGATTTAATTGAAGTGCCTCCGGCTTATCTCGAAGAGGATGAT
 GTGGTGCGTTTTGCGGATCGCTACGGACGGGTGTAA

[0304]

[0305]

서열 26 manC (EC 2.7.7.13)

MAQSKLYPVVMAGGSGSRLWPLSRVLYPKQFLCLKGDLTMLQ
 TTICRLNGVECESPVVICNEQHRFIVAEQLRQLNKLTENIILEPAGRNTAPAIAL
 AALAAKRHSPESDPLMLVLAADHVIADDAFRAAVRNAMPYAEAGKLVTFG
 IVPDLPETGYGYIRRGEVSAGEQDMVAFEVAQFVEKPNLETAQAYVASGEYY
 WNSGMFLFRAGRYLEELKKYRPDILDACEKAMSAVDPDLNFIRVDEEAFAC
 PEESVDYAVMERTADAVVPMMDAGWSDVGSWSSLWEISAHTAEGNVCHGD
 VINHKTENSYVVAESGLVTTVGKDLVVVQTKDAVLIADRNAVQDVKKVVE
 QIKADGRHEHRVHREVYPWGKYDSIDAGDRYQVKRITVKPGEGLSVQMHH
 HRAEHVVVAGTAKVTIDGDIKLLGENESIYIPLGATHCLENPGKIPLDLIEVR
 SGSYLEEDDVVRFADRYGRV*

[0306]

[0307]

서열 27 g1mS (EC 2.6.1.16)

ATGTGTGGAATTGTTGGCGGATCGCGCAACGTGATGTAGC
 AGAAATCCTTCTTGAAGTTTACGTCTGTTGGAATACCGCGGATATGACTC
 TGCCGGTCTGGCCGTTGTTGATGCAGAAAGTCATATGACCCGCTGCGTC
 GCCTCGGTAAAGTCCAGATGCTGGCACAGGCAGCGGAAGAACATCCTCTG
 CATGGCGGCACTGGTATTGCTCACACTCGCTGGGCGACCCACGGTGAACC
 TTCAGAAGTGAATGCGCATCCGCATGTTTCTGAACACATTGTGGTGGTGC
 ATAACGGCATCATCGAAAACCATGAACCGCTGCGTGAAGAGCTAAAAGC
 GCGTGGCTATACCTTCGTTTCTGAAACCGACACCGAAGTGATTGCCCATCT
 GGTGAACTGGGAGCTGAAACAAGGCGGGACTCTGCGTGAGGCCGTTCTGC
 GTGCTATCCCGCAGCTGCGTGGTGCGTACGGTACAGTGATCATGGACTCC
 CGTCAACCCGATAACCTGCTGGCGGCACGTTCTGGTAGTCCGCTGGTGATT
 GGCCTGGGGATGGGCGAAAACCTTTATCGCTTCTGACCAGCTGGCGCTGTT
 GCCGGTGACCCGCTCGCTTTATCTTCCTTGAAGAGGGCGATATTGCGGAAA

[0308]

TCACTCGCCGTTTCGGTAAACATCTTCGATAAAACTGGCGCGGAAGTAAAA
 CGTCAGGATATCGAATCCAATCTGCAATATGACGCGGGCGATAAAGGCAT
 TTACCGTCACTACATGCAGAAAGAGATCTACGAACAGCCGAACGCGATCA
 AAAACACCCTTACCGGACGCATCAGCCACGGTCAGGTTGATTTAAGCGAG
 CTGGGACCGAACGCCGACGAACTGCTGTGCAAGGTTGAGCATATTCAGAT
 CCTCGCCTGTGGTACTTCTTATAACTCCGGTATGGTTTCCCGCTACTGGTTT
 GAATCGCTAGCAGGATTTCCGTGCGACGTCGAAATCGCCTCTGAATTCCG
 CTATCGCAAATCTGCCGTGCGTCGTAACAGCCTGATGATCACCTTGTCACA
 GTCTGGCGAAACCGCGGATAACCTGGCTGGCCTGCGTCTGTGAAAAGAGC
 TGGGTTACCTTGGTTCCTGGAATCTGTAACGTTCCGGGTTCTTCTCTGG
 TGCGCGAATCCGATCTGGCGCTAATGACCAACGCGGGTACAGAAATCGGC
 GTGGCATCCACTAAAGCATTCACTCAGTAACTGTGCTGTTGATGCTG
 GTGGCGAAGCTGTCTCGCCTGAAAGGTCTGGATGCCTCCATTGAACATGA
 CATCGTGCATGGTCTGCAGGCGCTGCCGAGCCGTATTGAGCAGATGCTGT
 CTCAGGACAAACGCATTGAAGCGCTGGCAGAAGATTTCTCTGACAAACAT
 CACGCGCTGTTCTGGGCCGTGGCGATCAGTACCCAATCGCGCTGGAAGG
 CGCATTGAAGTTGAAAGAGATCTCTTACATTCACGCTGAAGCCTACGCTG
 CTGGCGAACTGAAACACGGTCCGCTGGCGCTAATTGATGCCGATATGCCG
 GTTATTGTTGTTGCACCGAACCAACGAATTGCTGGAAAACTGAAATCCAA
 CATTGAAGAAGTTCGCGCGCGTGGCGGTGAGTTGTATGTCTTCGCCGATC
 AGGATGCGGGTTTTGTAAGTAGCGATAACATGCACATCATCGAGATGCCG
 CATGTGGAAGAGGTGATTGCACCGATCTTCTACACCGTTCCGCTGCAGCT
 GCTGGCTTACCATGTGCGCTGATCAAAGGCACCGACGTTGACCAGCCGC
 GTAACCTGGCAAAATCGGTTACGGTTGAGTAA

[0309]

[0310]

서열 28 gImS (EC 2.6.1.16)

MCGIVGAIQRDVAEILLEGLRRLEYRGYDSAGLAVVDAEGH
MTRLRRLGKVVQMLAQAAEEHPLHGGTGIAHTRWATHGEPSEVNAHPHVSEH
IVVVHNGIIEHNEPLREELKARGYTFVSETDTEVIAHLVNWELKQGGTLREAV
LRAIPQLRGAYGTVIMDSRHPDTLLAARSGSPLVIGLGMGENFIASDQLALLP
VTRRFIFLEEGDIAEITRRSVNIFDKTGAEVKRQDIESNLQYDAGDKGIYRHYM
QKEIYEQPNAIKNTLTGRISHGQVDLSELGPNADLLSKVEHIQILACGTSYNS
GMVSRYWFESLAGIPCDVEIASEFRYRKSAVRRNSLMITLSQSGETADTLAGL
RLSKELGYLGS LAICNVPGSSLVRESLALMTNAGTEIGVASTKAFTTQLTVL
LMLVAKLSRLKGLDASIEHDIVHGLQALPSRIEQLMSQDKRIEALAEDFSDKH
HALFLGRGDQYPIALEGALKLKEISYIHAEAYAAGELKHGPLALIDADMPVIV
VAPNNELLEKLKSNIEVRARGGQLYVFADQDAGFVSSDNMHIIEMPHVEEVI
APIFYTVPLQLLAYHVALIKGTDVDQPRNLAKSVTVE*

[0311]

[0312]

서열 29 (EC 2.4.99.1) ST6

ATGAAAAAATCCTGACCGTGCTGTCCATCTTTATCCTGTCTGC
CTGTAATAGCGACAATACCAGCCTGAAAGAGACTGTTAGCAGCAATTCAG
CGGATGTTGTGAAACCGAAACTTATCAACTGACGCCGATCGATGCTCCT
TCTTCGTTCTGAGCCATTCTTGGGAACAGACCTGTGGTACACCAATTCTG
AACGAGTCCGACAAACAGGCCATTTCTTCGATTTTGTTGCCCGGAACTG
AAACAAGACGAGAAATATTGCTTCACCTTCAAAGGCATTACCGGTGATCA
TCGTTATATCACGAACACCCTCTGACTGTCGTAGCACCGACACTGGAAG
TGTATATCGACCATGCCAGCCTGCCTAGTCTGCAGCAACTGATCCATATTA
TCCAGGCGAAAGACGAATATCCGAGCAACCAGCGTTTTGTGAGCTGGAAA
CGTGTACTGTGGATGCCGACAACGCCAATAAACTGAACATTACACCTA
TCCTCTGAAAGGCAATAACACCAGCCCTGAGATGGTAGCGGCGATTGATG
AGTATGCCAGAGCAAAAACCGTCTGAACATTGAGTTCTATACCAATACG
GCCCACGTGTTAATAACCTGCCGCAATCATTCAACCTCTGTATAACAAC
GAGAAAGTGAAATCAGCCACATTTCTGCTGTATGATGATGGCAGTAGCGA
GTATGTTAGCCTGTATCAGTGGAAAAGACACCCCGAATAAAATCGGAGACTC
TGGAGGGTGAAGTTTCTCTGCTGGCCAACTATCTGGCCGGTACAAGTCCT
GATGCTCCGAAAGGGATGGGTAACCGCTATAATTGGCACAAACTGTATGA
CACCGACTATTATTTCTGCGCGAGGATTATCTGGACGTGGAAGCCAATCT
GCATGATCTGCGCGATTATCTGGGTTCTAGCGCCAAACAAATGCCGTGGG
ATGAATTTGCTAAACTGTCCGATTCTCAGCAAACCCTGTTCTGGACATCG
TTGGCTTTGATAAAGAGCAGCTGCAACAGCAGTATAGCCAGTCACCGCTG
CCGAACCTCATTTTACTGGCACCACCACATGGGCAGGGGGTGAGACAAA
AGAGTATTATGCTCAACAACAGGTGAACGTCATCAACAATGCCATTAACG
AAACCTCCCCATATTATCTGGGTAAAGACTATGACCTGTTCTTTAAAGGCC
ATCCGGCTGGAGGAGTGATTAATGATATTATCCTGGGCTCCTTTCTGACA
TGATTAACATTCGGCGGAAAATCTCATTTGAGGTGCTGATGATGACTGAT

[0313]

ATGCTGCCGGATACCGTTGCTGGAATTGCCTCTTCCCTGTATTTACCATT
CCTGCCGACAAAAGTGAACCTCATCGTGTTCACCAGCAGTGATACCATTAC
AGACCGTGAAGAAGCGCTGAAATCTCCTCTGGTTCAGGTGATGCTGACAC
TGGGTATCGTGAAAGAAAAAGACGTCCTGTTTTGGGCCGACCATAAAGTG
AATAGCATGGAGGTGGCCATCGACGAAGCGTGTACTCGTATTATCGCCAA
ACGTCAGCCTACCGCTCAGATCTGCGTCTGGTTATCGCCATTATCAAAAC
GATCACCGATCTGGAGCGTATTGGAGATGTTGCCGAAAGCATTGCCAAAG
TTGCCCTGGAGAGCTTTTCTAACAAACAGTATAATCTGCTGGTCAGCCTGG
AATCTCTGGGTCAACACACCGTTCGTATGCTGCATGAAGTGCTGGATGCTT
TTGCCCGTATGGATGTGAAAGCAGCCATTGAAGTCTATCAGGAGGATGAC
CGTATCGATCAGGAATATGAGAGCATTGTCCGTCAACTGATGGCCCATAT
GATGGAAGATCCGTCTAGCATTCCGAATGTGATGAAAGTGATGTGGGCAG
CTCGTAGTATTGAACGTGTGGGTGACCGCTGCCAGAACATTTGTGAGTAT
ATCATCTATTTCTGAAAAGGCAAAGATGTTGCCACACCAAACCGGATGA
CTTCGGTACTATGCTGGACTGA

[0314]

[0315]

서열 30 (EC 2.4.99.1) ST6

MKKILTVLSIFILSACNSDNTSLKETVSSNSADVETETYQLTPIDAP
SSFLSHSWEQTCGPILNESDKQAISDFVAPELKQDEKYCFTFKGITGDHRYI
TNTTLTVVAPTLEVYIDHASLPSLQLIHIIQAKDEYPSNQRFVSWKRVTVDA
DNANKLNIHTYPLKGNNTSPEMVAAIDEYAQSKNRLNIEFYTNTAHVFNNLP
PIIQPLYNNEKVKISHISLYDDGSSEYVSLYQWKDTPNKIETLEGEVSSLANYL
AGTSPDAPKGMGNRYNWHKLYDTDYFLREDYLDVEANLHDLRDYLGSSA
KQMPWDEFKLSDSQQLFLDIVGFDKEQLQQYSQSPLPNFIFTGTTWAG
GETKEYYAQQQVNVINAINETSPYYLGKDYDLFFKGHPAGGVINDIILGSFP
DMINIPAKISFEVLMMDMLPDTVAGIASSLYFTIPADKVNFIIFTSSDTITDRE
EALKSPLVQVMLTLGIVKEKDVLFWADHKVNSMEVAIDEACTRIIAKRQPTA
SDLRLVIAIITITDLERIGDVAESIAKVALESFSNKQYNLLVSLESLGQHTVRM
LHEVLDAFARMDVKAIEVYQEDDRIDQEYESIVRQLMAHMMEDPSSIPNV
MKVMWAARSIERVGDRCQNICEYIIYFVKGKDVHRHTKPDDFGTMLD*

[0316]

[0317]

서열 31 MBP-GnTI 융합체

ATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCG
ATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT

[0318]

ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATT
CCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAC
ACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCC
CGGACAAAAGCGTTCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTA
CGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCG
CTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGA
GATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTG
ATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGAC
GGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGT
GGGCGTGGATAACGCTGGCGCGAAAAGCGGGTCTGACCTTCTGGTTGACC
TGATTA AAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAA
GCTGCCTTAATAAAGGCGAAAACAGCGATGACCATCAACGGCCCGTGGGC
ATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAAACGGTACTGC
CGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCA
GGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCGAAAAGATTCTCGA
AAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGAT
CCACGTATTGCCGCCACCATGGAAAACGCCCAGAAAGGTGAAATCATGCC
GAACATCCCGCAGATGTCCGCTTCTGGTATGCCGTGCGTACTGCGGTGAT
CAACGCCGCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGC
AGACTCGTATACCAAGGCGACACAGTCAGAATATGCAGATCGCCTTGCT
GCTGCAATTGAAGCAGAAAATCATTGTACAAGCCAGACCAGATTGCTTAT
TGACCAGATTAGCCTGCAGCAAGGAAGAATAGTTGCTCTTGAAGAACAAA
TGAAGCGTCAGGACCAGGAGTGCCGACAATTAAGGGCTCTTGTTTCAGGAT
CTTGAAAGTAAGGGCATAAAAAAGTTGATCGGAAATGTACAGATGCCAGT
GGCTGCTGTAGTTGTTATGGCTTGCAATCGGGCTGATTACCTGGAAAAGA
CTATTAATCCATCTTAAAATACCAAATATCTGTTGCGTCAAATATCCTC
TTTTCATATCCCAGGATGGATCACATCCTGATGTCAGGAAGCTTGCTTTGA
GCTATGATCAGCTGACGTATATGCAGCACTTGGATTTTGAACCTGTGCATA
CTGAAAGACCAGGGGAGCTGATTGCATACTACAAAATTGCACGTCATTAC
AAGTGGGCATTGGATCAGCTGTTTTACAAGCATAATTTAGCCGTGTTATC
ATACTAGAAGATGATATGGAAATTGCCCTGATTTTTTTGACTTTTTTGAG

[0319]

GCTGGAGCTACTCTTCTTGACAGAGACAAGTCGATTATGGCTATTTCTTCT
TGGAATGACAATGGACAAATGCAGTTTGTCCAAGATCCTTATGCTCTTTAC
CGCTCAGATTTTTTCCCGGTCTTGGATGGATGCTTTCAAAATCTACTTGG
GACGAATTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAG
ACTCAAAGAGAATCACAGAGGTCGACAATTTATTGCCCAGAAGTTTGCA
GAACATATAATTTTGGTGAGCATGGTTCTAGTTTGGGGCAGTTTTTCAAGC
AGTATCTTGAGCCAATTAATACTAAATGATGTCCAGGTTGATTGGAAGTCA
ATGGACCTTAGTTACCTTTTGGAGGACAATTACGTGAAACACTTTGGTGAC
TTGGTTAAAAAGGCTAAGCCCATCCATGGAGCTGATGCTGTCTTGAAAGC
ATTTAACATAGATGGTGATGTGCGTATTCAGTACAGAGATCAACTAGACT
TTGAAAATATCGCACGGCAATTTGGCATTTTGAAGAATGGAAGGATGGT
GTACCACGTGCAGCATATAAAGGAATAGTAGTTTTCCGGTACCAAACGTC
CAGACGTGTATTCTTGTGGCCATGATTTCGCTTCAACAACCTCGGAATTGA
AGATACTTAA

[0320]

서열 32 MBP-GnTII 융합체

[0321]

ATGAAAATCGAAGAAGGTAATACTGGATTAACGGCG
ATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT
ACCGGAATTAAGTACCGTTGAGCATCCGATAAACTGGAAGAGAAATT
CCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAC
ACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCC
CGGACAAAGCGTTCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTA
CGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCG
CTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGA
GATCCCGGCGCTGGATAAAGAAGTAAAGCGAAAGGTAAGAGCGCGCTG
ATGTTCAACCTGCAAGAACCGTACTTACCTGGCCGCTGATTGCTGCTGAC
GGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGT
GGGCGTGGATAACGCTGGCGGAAAAGCGGGTCTGACCTTCTGGTTGACC
TGATTA AAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAA
GCTGCCTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGC
ATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGC
CGACCTTCAAGGGTCAACCATCAAACCGTTCGTTGGCGTGCTGAGCGCA
GGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCGAAAGAGTTCTCTCGA

[0322]

AAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGAT
CCACGTATTGCCGCCACCATGGAAAACGCCAGAAAAGGTGAAATCATGCC
GAACATCCCCGAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGAT
CAACGCCGCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGC
AGACTCGTATCACCAAGCGTCAGCGTAAAAATGAAGCCCTGGCACCTCCT
CTGCTGGATGCTGAACCGGCACGTGGTGGTGGCGGTGCTGGTGGTGGTCA
TCCGTCTGTTGCCGTTGGTATTCTGTCGTGTGAGCAATGTTTCGGCTGCCTC
TCTGGTCCCGGCTGTTCCCTCAACCTGAAGCTGATAACCTGACCCTGCGCTA
TCGCTCTCTGGTGTATCAACTGAACTTCGATCAAACTCTGCGTAACTGGA
TAAAGCAGGCACATGGGCTCCTCGTGAAGTGGTACTGGTAGTCCAGGTCC
ATAATCGTCCGGAATATCTGCGTCTGCTGCTGGATTCTCTGCGCAAAGCTC
AAGGCATCGATAATGTCCTGGTCATCTTCTCTCATGATTTCTGGAGCACGG
AGATTAACCAGCTGATTGCCGGCGTGAATTTTTGTCCTGTGCTGCAGGTGT
TTTTCCGTTTTCTATCCAAGTATCCGAACGAATTTCCGGGTTCTGATCC
TCGTGATTGTCTCGTATCTGCCTAAAAATGCCGCTCTGAAACTGGGCTG
TATTAATGCCGAGTATCCTGATTTCTTTGGCCACTATCGTGAGGCGAAATT
TTCTCAGACCAAACATCATTGGTGGTGGAAACTGCATTTCTGTGTGGGAAC
GTGTGAAAATCCTGCGCGACTATGCTGGCCTGATTCTGTTTCTGGAAGAA
GATCACTATCTGGCTCCGGACTTTTATCATGTGTTCAAAAAATGTGGAAA
CTGAAACAGCAGGAATGTCCAGAATGTGATGTGCTGTCACTGGGCACCTA
TAGTGCTTCTCGCTCCTTCTATGGTATGGCCGACAAAGTGGACGTAAAAC
ATGAAAATCCACCGAGCACAAACATGGGTCTGGCACTGACTCGTAATGCCT
ATCAAAAACTGATTGAGTGTACCGACACCTTTTGTACGTATGATGACTATA
ACTGGGACTGGACCCTGCAATATCTGACCGTGAGCTGTCTGCCAAAATTT
GGAAAGTTCTGGTGCCTCAGATTCTCTCGTATCTTTCATGCTGGCGACTGTG
GTATGCACCATAAAAAAATTGCCGTCCGTCAACACAATCTGCTCAGATC
GAGTCGCTGCTGAATAATAACAAACAGTATATGTTCCCGGAGACTCTGAC
AATTTCTGAAAAATTCACCGTGGTCGCCATTTCTCCGCCTCGTAAAAATGG
AGTTGGGGCGATATCCGTGACCATGAACTGTGTAAAAGCTATCGTCGTC
TGCAGTGA

[0323]

[0324]

서열 33 MBP-GnTIV 융합체

ATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCG
ATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT
ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATT
CCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAC
ACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCC
CGGACAAAGCGTTCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTA
CGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCG
CTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGA
GATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTG
ATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGAC
GGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGT
GGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACC
TGATTAATAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAA
GCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGC
ATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAAACGGTACTGC
CGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCA
GGTATTAACGCCCGCAGTCCGAACAAAGAGCTGGCGAAAGAGTTCTCTCGA
AACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGAT
CCACGTATTGCCGCCACCATGGAAAACGCCAGAAAGGTGAAATCATGCC
GAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGAT
CAACGCCGCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGC
AGACTCGTATACCAAGATTTTGAAAGAACTGACGTCCAAAAAGAGCTTG
CAAGTCCCGTCCATCTACTATCACTTGCCGCACTTGCTGCAAAACGAGGG
CTCTTTGCAACCGGCAGTTCAGATCGGCAATGGTCGCACCGGCGTGAGCA
TTGTTATGGGTATCCCGACCGTGAACGTGAAGTGAAGGCTATCTGATT
GAAACGCTGCATAGCCTGATCGATAACCTGTACCCGGAAGAAAACTGGA
CTGCGTGATTGTCGTTTTATTGGTGAACCGACACGGATTATGTGAATGG
CGTTGTTGCCAATCTGGA AAAAGAGTTCAGCAAAGAGATCAGCAGCGGCC
TGGTTGAGATCATTTCTCCGCCGAGAGCTATTACCCGGATCTGACGAAC
CTGAAAGAAACCTTCGGTGATAGCAAAGAGCGTGTCCGTTGGCGCACTAA
GCAGAACCTGGACTATTGTTTTCTGATGATGTACGCGCAAGAAAAGGGTA

[0325]

CGTATTACATCCAACCTGGAGGACGACATTATTGTGAAGCAAACTACTTC
AACACCATTAAGAACTTCGCGCTGCAGCTGAGCAGCGAAGAGTGGATGAT
TCTGGAGTTCAGCCAGCTGGGCTTCATTGGCAAGATGTTTCAGGCACCGG
ACTTGACCCTGATCGTGGAGTTTATCTTTATGTTCTACAAAGAGAAAACCGA
TCGATTGGCTGCTGGATCATATCCTGTGGGTCAAGGTCTGCAATCCGGAA
AAAGATGCCAAGCATTGTGACCGCCAGAAAAGCGAATCTGCGTATTCTGTTT
TCGTCCTAGCCTGTTCCAACACGTGGGTCTGCACAGCTCTCTGACCGGTAA
GATCCAAAAGCTGACCGACAAAAGATTACATGAAAACCGCTGCTGCTGAAGA
TCCATGTCAACCCGCCAGCAGAGGTGAGCACCTCGCTGAAAAGTCTACCAG
GGTCACACTCTGGAGAAAACCTATATGGGCGAGGACTTCTTTTGGGCGAT
TACGCCTGTTGCGGGTGACTATATCTTGTTTAAGTTTGACAAGCCGGTTAA
TGTAGAGAGCTACTTGTTTCATAGCGGTAACCAGGATCACCCAGGTGACA
TTCTGCTGAACACCACCGTTGAAGTGTGCCGCTGAAAAGCGAAGGTCTG
GATATTTGAAAAGAAAACGAAGGATAAGCGTCTGGAGGATGGTACTTCCG
TATCGGCAAGTTCGAGAATGGCGTGGCTGAAGGTATGGTCGACCCGAGCC
TGAACCCGATTTCCGCATTTCCGCTGTCCGTCATCCAGAATAGCGCGGTTT
GGGCTATCCTGAATGAGATTACATCAAAAAGGTTACGAATTA

[0326]

[0327]

서열 34 GST-*alg11* 융합체

ATGAAATTGTTCTACAAACCGGGTGCCTGCTCTCTCGCTTCCCA
TATCACCTGCGTGAGAGCGGAAAGGATTTACCCCTCGTCAGTGTGGATTT
AATGAAAAACGTCTCGAAAACGGTGACGATTACTTTGCCGTTAACCTA
AGGGGCAGGTGCCTGCATTGCTGCTGGATGACGGTACTTTGCTGACGGAA
GGCGTAGCGATTATGCAGTATCTTGCCGACAGCGTCCCCGACCGCCAGTT
GCTGGCACCGGTAAACAGTATTTCCCGCTATAAAAACCATCGAATGGCTGA
ATTACATCGCCACCGAGCTGCATAAAGGTTTACACCTCTGTTTCGCCCTG
ATACACCGGAAGAGTACAAACCGACAGTTTCGCGCGCAGCTGGAGAAGAA
GCTGCAATATGTGAACGAGGCACTGAAGGATGAGCACTGGATCTGCGGGC
AAAGATTTACAATTGCTGATGCCTATCTGTTTACGGTTCTGCGCTGGGCAT
ACGCGGTGAACTGAATCTGGAAGGGTTAGAGCACATTGCAGCATTATG
CAACGTATGGCTGAACGTCGGAAAGTACAAGACGCGCTGTCAGCGGAAG
GCTTAAAGGGCAGTGCTTGGACAACTACAATTTTGAAGAGGTTAAGTCT
CATTTTGGGTTCAAAAAATATGTTGTATCATCTTAGTACTAGTGTATGGA

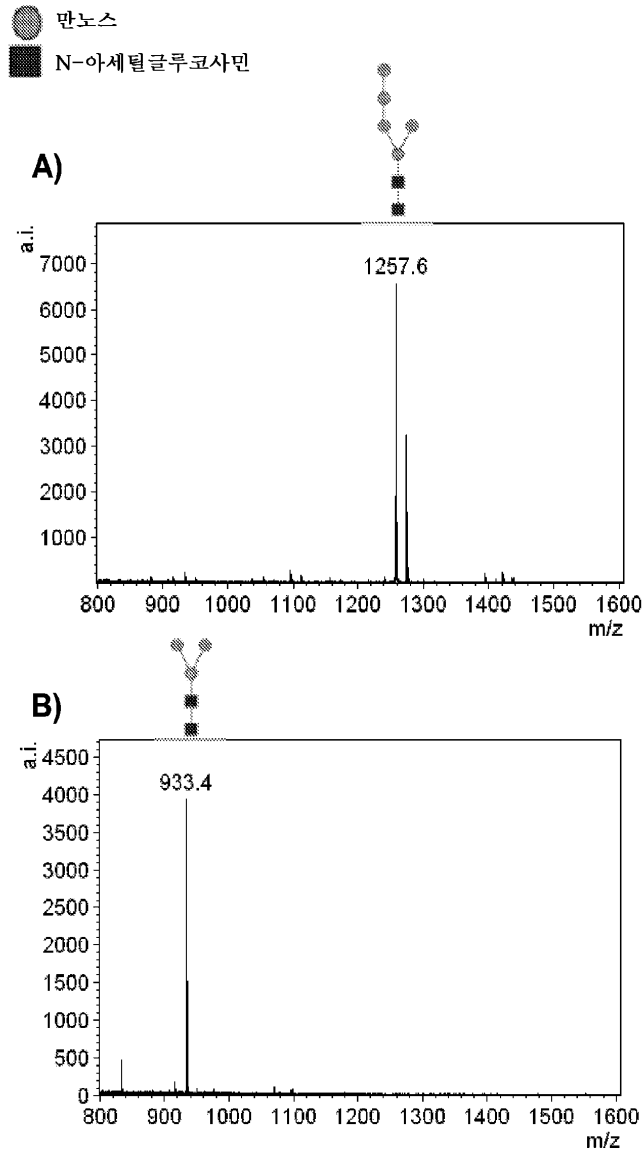
[0328]

CTAATTAAGGTTCTCACGTGGATCTTCCGTCAATGGGTGTATTCCAGCTTG
AATCCGTTCTCCAAAAATCTTCATTACTGAACAGAGCAGTTGCCTCCTGT
GGTGAGAAGAATGTGAAAGTTTTTGGTTTTTTTCATCCGTATTGTAATGCT
GGTGGTGGTGGGGAAAAAGTCTCTGGAAAGCTGTAGATATCACTTTGAG
AAAAGATGCTAAGAACGTTATTGTCATTTATTCAGGGGATTTTGTGAATG
GAGAGAATGTTACTCCGGAGAATATTCTAAATAATGTGAAAGCGAAGTTC
GATTACGACTTGGATTCCGATAGAATATTTTTTCATTTTCATTGAAGCTAAGA
TACTTGGTGGATTCTTCAACATGGAAGCATTTCACGTTGATTGGACAAGCA
ATTGGATCAATGATTCTCGCATTTGAATCCATTATTCAGTGTCCACCTGAT
ATATGGATTGATACAATGGGGTACCCTTTCAGCTATCCTATTATTGCTAGG
TTTTTGAGGAGAATTCCTATCGTCACATATACGCATTATCCGATAATGTCA
AAAGACATGTTAAATAAGCTGTTCAAATGCCCAAGAAGGGTATCAAAGT
TTACGGTAAAATATTACTGGAAAGTTTTTATGTTAATTTATCAATCCAT
TGGTTCTAAAATTGATATTGTAATCACAAACTCAACATGGACAAAATAACC
ACATAAAGCAAATTTGGCAATCCAATACGTGTAATAATTATATATCCTCCA
TGCTCTACTGAGAAATTAGTAGATTGGAAGCAAAAGTTTGGTACTGCAAA
GGTGAGAGATTAATCAAGCAATTGTGTTGGCACAATTTTCGTCCTGAGA
AACGTCATAAGTTAATCATTGAGTCCTTTGCAACTTTCTTGAAAAATTTAC
CGGATTCTGTATCGCCAATTAATGATAATGGCGGGTCCACTAGATCC
AAGCAAGATGAAAATTATGTTAAAAGTTTACAAGACTGGTCAGAAAATGT
ATTAATAATTCCTAAACATTTGATATCATTGAAAAAAATCTGCCCTTCGA
TAAGATTGAAATATTACTAAACAAATCTACTTTCGGTGTTAATGCCATGTG
GAATGAGCACTTTGGAATTGCAGTTGTAGAGTATATGGCTTCCGGTTTGAT
CCCCATAGTTTATGCCTCGGCGGGCCCATTTGTTAGATATAGTTACTCCATG
GGATGCCAACGGGAATATCGGAAAAGCTCCACCACAATGGGAGTTACAA
AAGAAATATTTGCAAAACTCGAAGATGATGGTGAAACTACTGGATTTTT
CTTTAAAGAGCCGAGTGATCCTGATTATAACACAACCAAAGATCCTCTGA
GATACCCTAATTTGTCCGACCTTTTCTTACAAATTACGAAACTGGACTATG
ACTGCCTAAGGGTGATGGGCGCAAGAAACCAGCAGTATTCATTGTATAAA
TTCTCTGATTTGAAGTTTGATAAAGATTGGGAAAACCTTGTACTGAATCCT
ATTTGTAAATTATTAGAAGAGGAGGAAAGGGGCTGA

[0329]

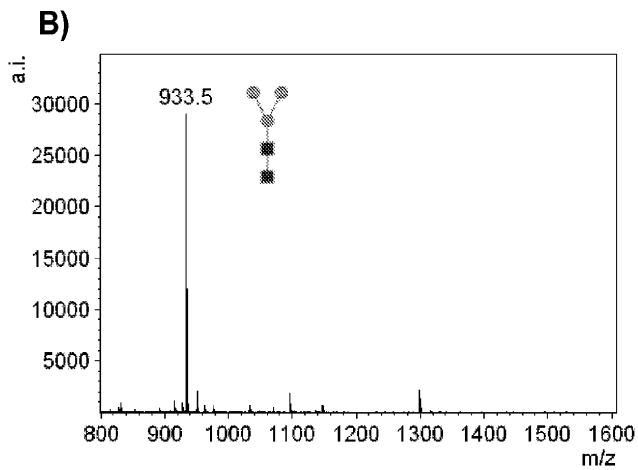
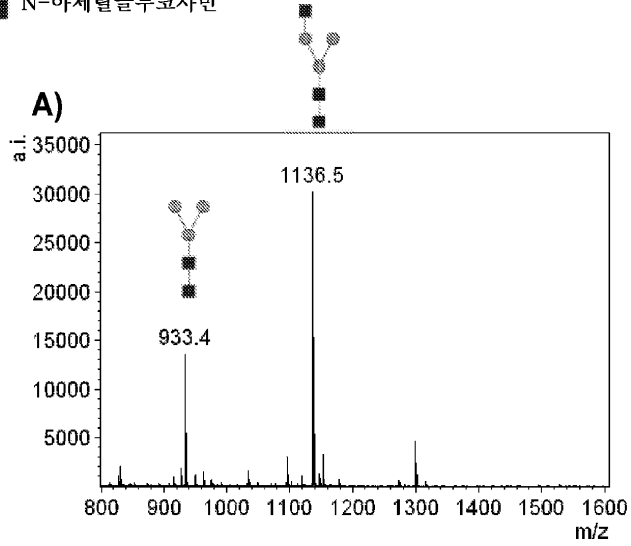
도면

도면1

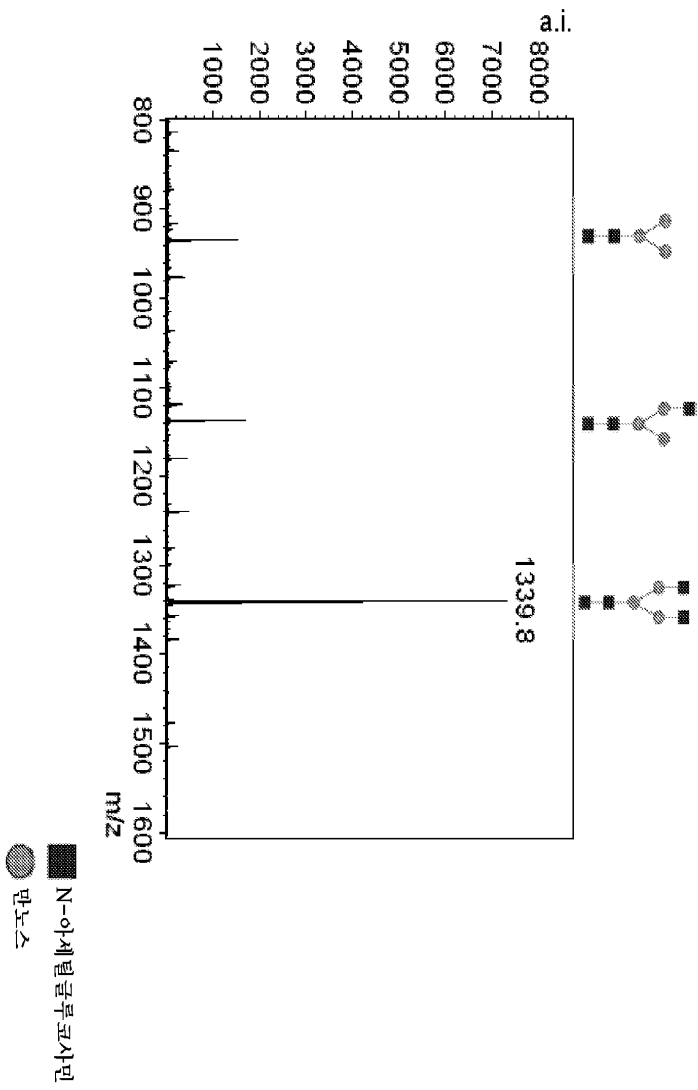


도면2

● 만노스
■ N-아세틸글루코사민

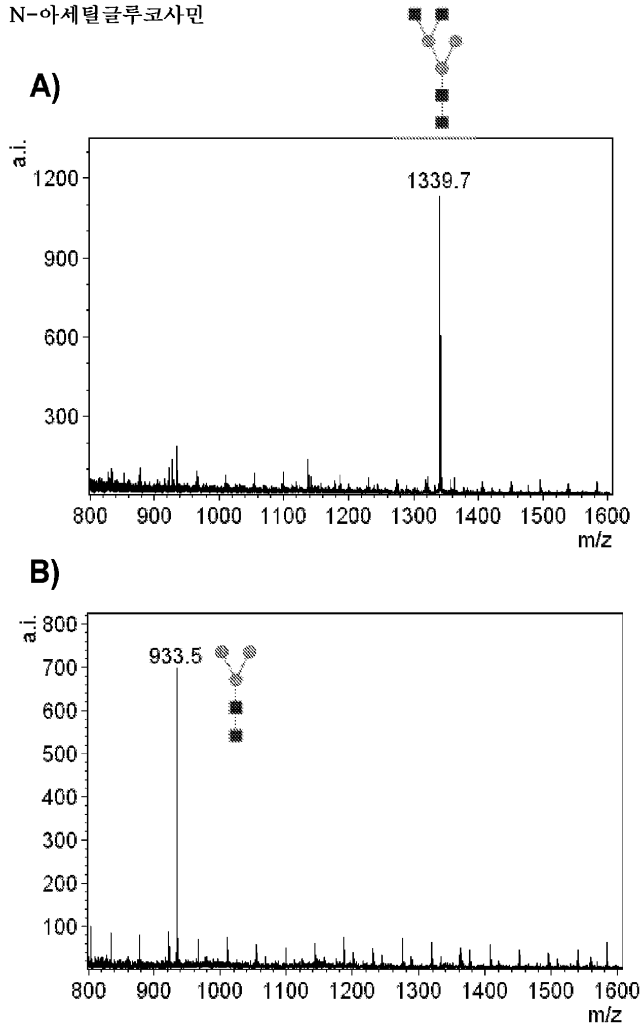


도면3

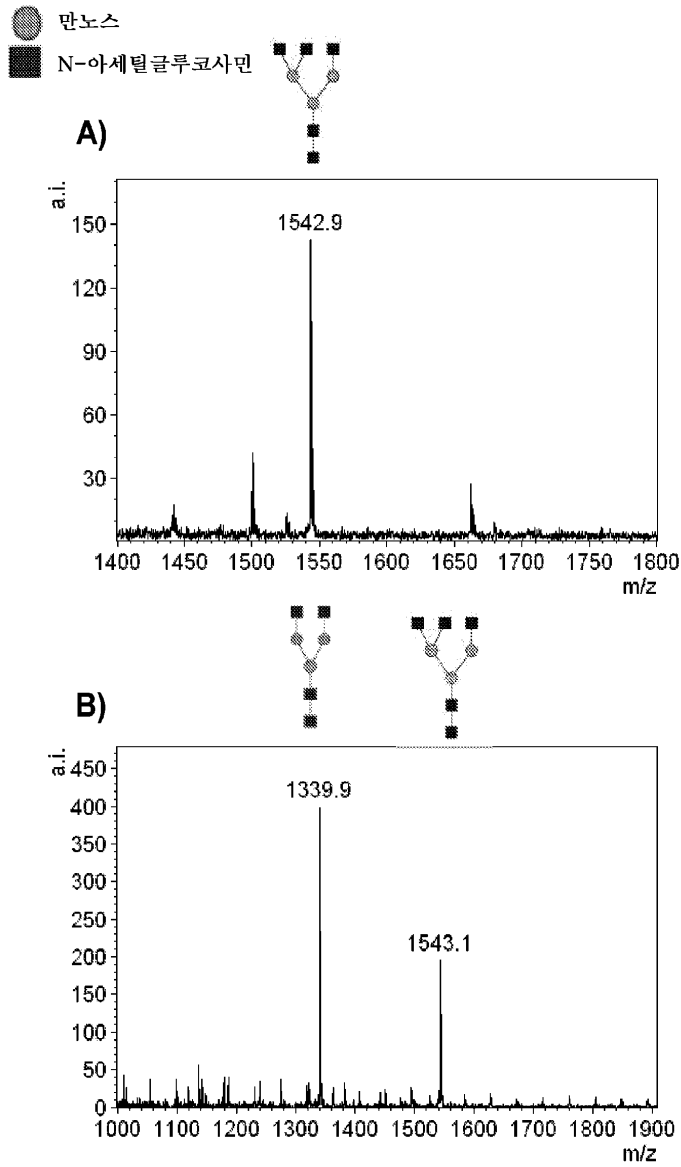


도면4

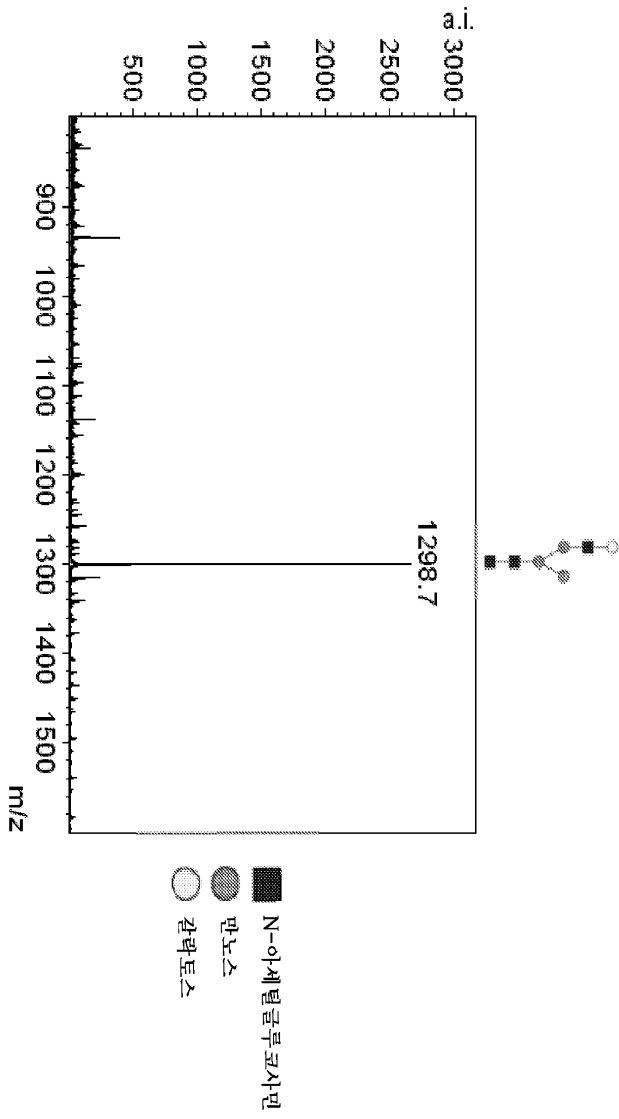
- 만노스
- N-아세틸글루코사민



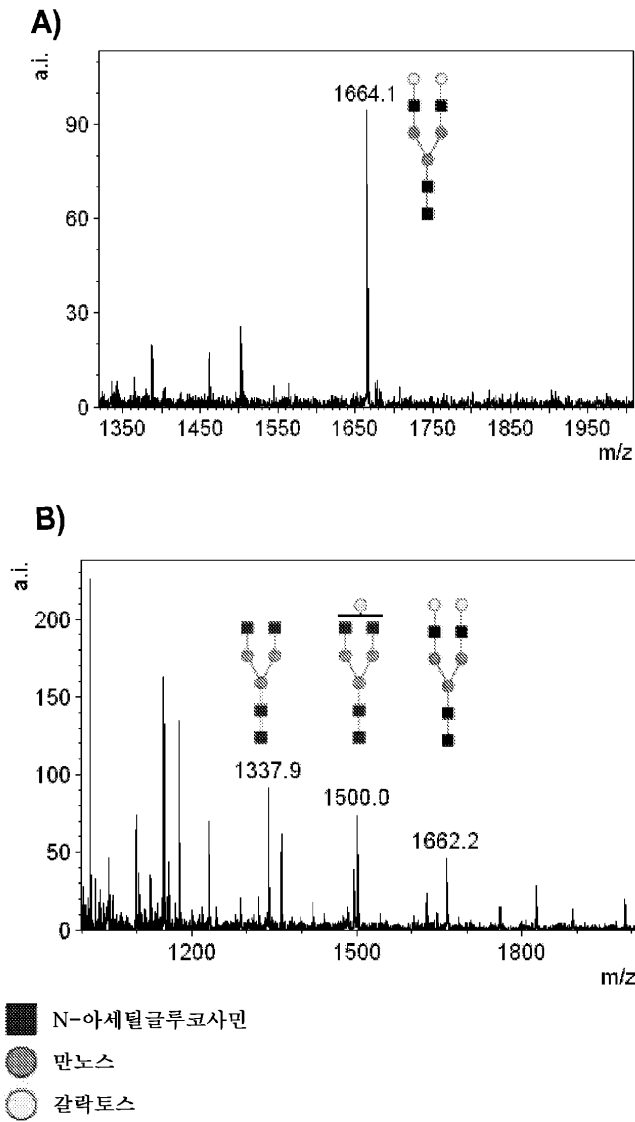
도면5



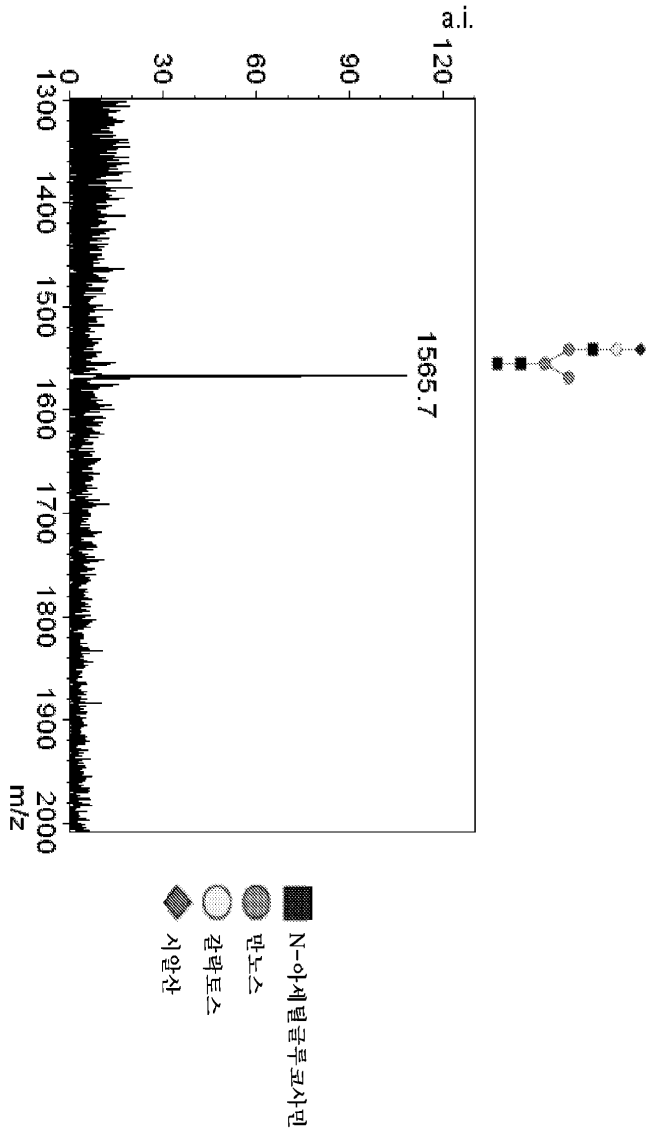
도면6



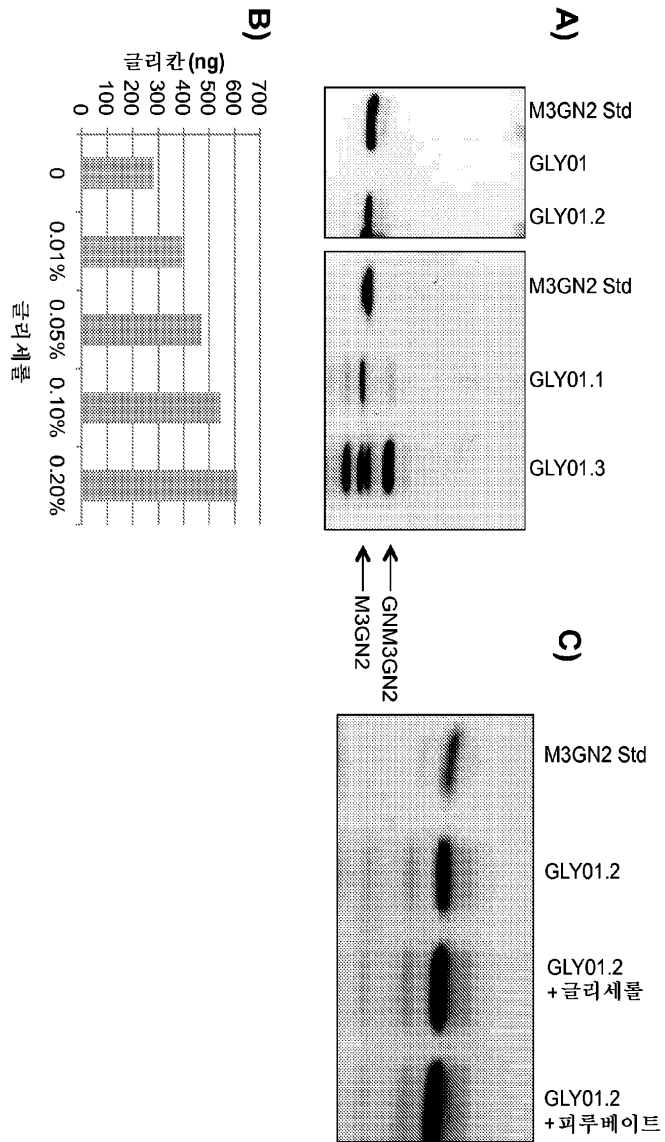
도면7



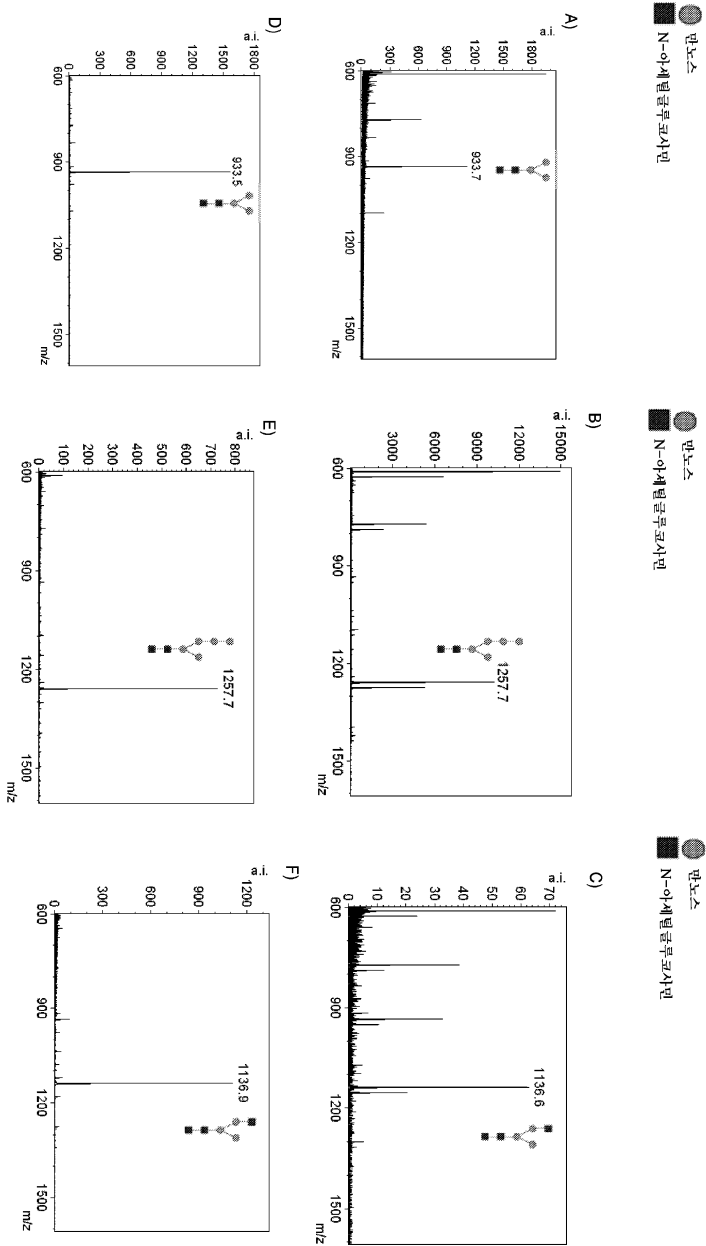
도면8



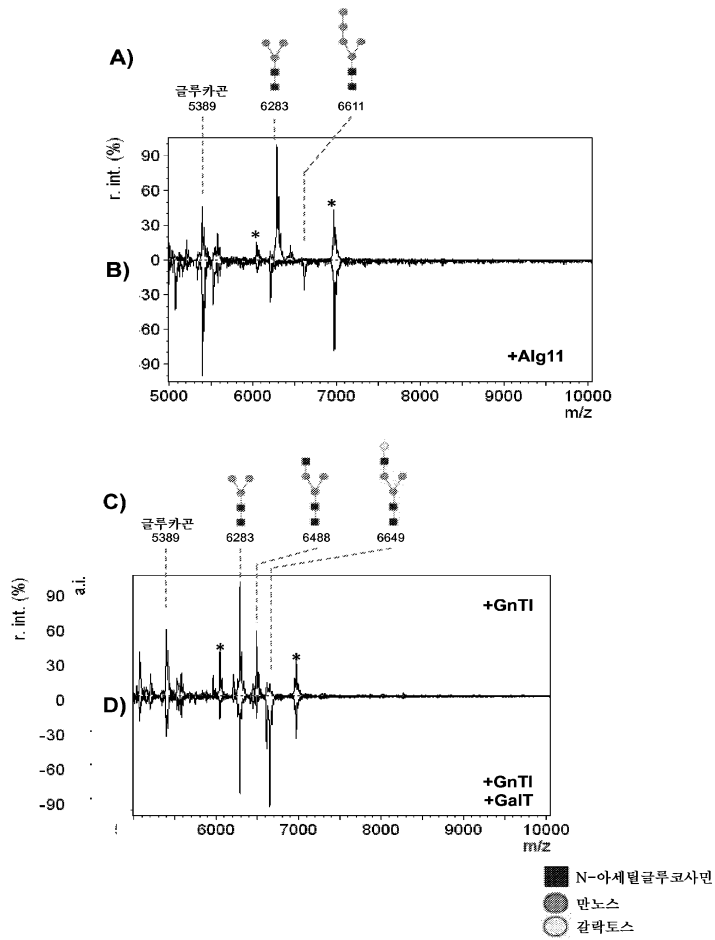
도면9



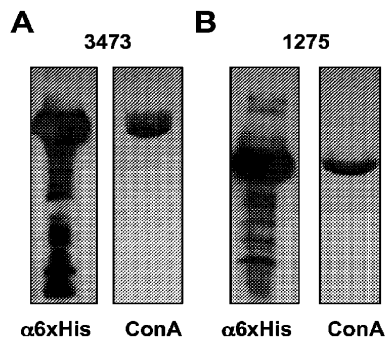
도면10



도면11



도면12



서열목록

SEQUENCE LISTING

<110> GLYCOBIA, INC.

<120> OLIGOSACCHARIDE COMPOSITIONS, GLYCOPROTEINS AND METHODS TO
PRODUCE THE SAME IN PROKARYOTES

<130> GLY-103 PCT

<140> PCT/US2014/026990

<141> 2014-03-14

<150> 61/785,586

<151> 2013-03-14

<160> 36

<170> PatentIn version 3.5

<210> 1

<211> 609

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 1

atgggtatca tcgaagaaaa agctctgttc gttacctgcg gtgctaccgt tccgttcccg 60

aaactggttt cttgcgttct gtctgacgaa ttctgccagg aactgatcca gtaacggtttc 120

gttcgtctga tcatccagtt cggctgtaac tactcttctg aattcgaaca cctgggttcag 180

gaacgtggtg gtcagcgtga atctcagaaa atcccgatcg accagttcgg ttgcggtgac 240

accgctcgtc agtacgttct gatgaacggg aaactgaaag ttatcggttt cgacttctct 300

accaaatagc agtctatcat ccgtgactac tctgacctgg ttatctctca cgctggtacc 360

ggttctatcc tggactctct gcgtctgaac aaaccgctga tcgtttgcgt taacgactct 420

ctgatggaca accaccagca gcagatcgtc gacaaattcg ttgaactggg ttacgtttgg 480

tcttgctctc cgaccgaaac cggctctgac gctggctctgc gtgcttctca gaccgaaaaa 540

ctgaaaccgt tcccgggttc tcacaaccgg tctttcgaac gtctgctggt tgaaaccatc 600

tactcttaa 609

<210> 2

<211> 202

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 2

Met Gly Ile Ile Glu Glu Lys Ala Leu Phe Val Thr Cys Gly Ala Thr

1 5 10 15
 Val Pro Phe Pro Lys Leu Val Ser Cys Val Leu Ser Asp Glu Phe Cys
 20 25 30

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

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

 Ser Cys Ala Pro Thr Glu Thr Gly Leu Ile Ala Gly Leu Arg Ala Ser
 165 170 175
 Gln Thr Glu Lys Leu Lys Pro Phe Pro Val Ser His Asn Pro Ser Phe
 180 185 190
 Glu Arg Leu Leu Val Glu Thr Ile Tyr Ser
 195 200

<210> 3

<211> 714

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 3

atgaaaaccg cttactggc ttctctgggt ctgatcggtt ctaccgctta cgttatccgt 60

ctgatcgcta tcctgccgtt cttccacacc caggctggta ccgaaaaaga caccaaagac 120

ggtgttaacc tgctgaaaat cgttaaatct tctaaaaaac cgctgaaaat cttegttttc 180

ctgggttctg gtggtcacac cggtgaaatg atccgtctgc tggaaaacta ccaggacctg 240

ctgctgggta aatctatcgt ttacctgggt tactctgacg aagcttctcg tcagcgtttc 300

gctcacttca tcaaaaaatt cggtcactgc aaagttaaat actacgaatt catgaaagct 360

cgtgaagtta aagctaccct gctgcagtct gttaaaacca tcatcggtac cctggttcag 420

tctttcgttc acgttggtcg taccggttc gctatgtcgc gttctccgca cctgttctcg 480

ctgaacggtc cgggtacctg ctgcatcctc tctttctggc tgaaaatcat ggaactgctg 540

ctgccgtgctc tgggttcttc tcacatcgtt tacgttgaat ctctggctcg tatcaacacc 600

ccgctctga ccggtaaaat cctgtactgg gttgttgacg aattcatcgt tcagtggcag 660

gaactgctg acaactacct gccgcgttct aaatggttcg gtatcctggt ttaa 714

<210> 4

<211> 237

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 4

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

1 5 10 15

Tyr Val Ile Arg Leu Ile Ala Ile Leu Pro Phe Phe His Thr Gln Ala

20 25 30

Gly Thr Glu Lys Asp Thr Lys Asp Gly Val Asn Leu Leu Lys Ile Arg

35 40 45

Lys Ser Ser Lys Lys Pro Leu Lys Ile Phe Val Phe Leu Gly Ser Gly

50 55 60

Gly His Thr Gly Glu Met Ile Arg Leu Leu Glu Asn Tyr Gln Asp Leu

65 70 75 80

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

85 90 95

Arg Gln Arg Phe Ala His Phe Ile Lys Lys Phe Gly His Cys Lys Val

100 105 110

Lys Tyr Tyr Glu Phe Met Lys Ala Arg Glu Val Lys Ala Thr Leu Leu
 115 120 125

Gln Ser Val Lys Thr Ile Ile Gly Thr Leu Val Gln Ser Phe Val His
 130 135 140

Val Val Arg Ile Arg Phe Ala Met Cys Gly Ser Pro His Leu Phe Leu
 145 150 155 160

Leu Asn Gly Pro Gly Thr Cys Cys Ile Ile Ser Phe Trp Leu Lys Ile
 165 170 175

Met Glu Leu Leu Leu Pro Leu Leu Gly Ser Ser His Ile Val Tyr Val
 180 185 190

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

Tyr Trp Val Val Asp Glu Phe Ile Val Gln Trp Gln Glu Leu Arg Asp
 210 215 220

Asn Tyr Leu Pro Arg Ser Lys Trp Phe Gly Ile Leu Val
 225 230 235

<210> 5

<211> 1350

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 5

atgttctctgg aaatcccgcg ttgctgctg gctctgatca tctgtacct gtctatcccg 60

ctggttgttt actacgttat cccgtacctg ttctacgta acaaatctac caaaaaacgt 120

atcatcatct tegtctctggg tgacgttggc cactctccgc gtatctgcta ccacgtatc 180

tctttcteta aactgggttg gcaggttgaa ctgtcgggtt acgttgaaga cacctgcccg 240

aaaatcatct cttctgacct gaacatcacc gttcaccaca tgtctaacct gaaacgtaaa 300

ggtgggtgta cctctgttat cttcatgggt aaaaaagttc tgttccaggt tctgtctatc 360

ttcaaactgc tgtgggaact gcgtggttct gactacatcc tggttcagaa cccgccgtct 420

atcccgatcc tgccgatcgc tgttctgtac aaactgaccg gttgcaaact gatcatcgac 480

tggcacaacc tggcttactc tctcctgcag ctgaaattca aaggtaactt ctaccaccg 540

ctggttctga tctcttacat ggttgaaatg atcttctcta aattcgtga ctacaacctg 600

accgttaccg aagctatcgg taaataacctg atccagtctt tccacctgaa cccgaaacgt 660

tgcgctgttc tgtacgaccg tccggcttct cagttccagc cgctggctgg tgacatctct 720

cgtcagaaaag ctctgaccac caaagctttc atcaaaaact acatccgtga cgacttcgac 780

accgaaaaag gtgacaaaat catcgttacc tctacctctt tcaccccgga cgaagacatc 840

ggtatcctgc tgggtgctct gaaaatctac gaaaactctt acgttaaatt cgactcttct 900

ctgccgaaaa tctctgtgctt catcaccggt aaaggtccgc tgaagaaaa atacatgaaa 960

caggttgaag aatacgactg gaaacgttgc cagatcgaat tcgtttggct gtctgctgaa 1020

gactaccgga aactgctgca gctgtgacgac tacgggtgtt ctctgcacac ctcttcttct 1080

ggtctggacc tgccgatgaa aatcctggac atgttcggtt ctggtctgcc ggttatcgtt 1140

atgaactacc cggttctgga cgaactggtt cagcacaacg ttaacggtct gaaattcgtt 1200

gaccgtcgtg aactgcacga atctctgac ttcgctatga aagacgctga cctgtaccag 1260

aaactgaaaa aaaacgttac ccaggaagct gaaaaccgtt ggcagtctaa ctgggaacgt 1320

accatgctg acctgaaact gatccactaa 1350

<210> 6

<211> 449

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 6

Met Phe Leu Glu Ile Pro Arg Trp Leu Leu Ala Leu Ile Ile Leu Tyr

1 5 10 15

Leu Ser Ile Pro Leu Val Val Tyr Tyr Val Ile Pro Tyr Leu Phe Tyr

20 25 30

Gly Asn Lys Ser Thr Lys Lys Arg Ile Ile Ile Phe Val Leu Gly Asp

35 40 45

Val Gly His Ser Pro Arg Ile Cys Tyr His Ala Ile Ser Phe Ser Lys

50 55 60

Leu Gly Trp Gln Val Glu Leu Cys Gly Tyr Val Glu Asp Thr Leu Pro

65 70 75 80

Lys Ile Ile Ser Ser Asp Pro Asn Ile Thr Val His His Met Ser Asn

Leu Ser Ala Glu Asp Tyr Pro Lys Leu Leu Gln Leu Cys Asp Tyr Gly
 340 345 350

Val Ser Leu His Thr Ser Ser Ser Gly Leu Asp Leu Pro Met Lys Ile
 355 360 365

Leu Asp Met Phe Gly Ser Gly Leu Pro Val Ile Ala Met Asn Tyr Pro
 370 375 380

Val Leu Asp Glu Leu Val Gln His Asn Val Asn Gly Leu Lys Phe Val
 385 390 395 400

Asp Arg Arg Glu Leu His Glu Ser Leu Ile Phe Ala Met Lys Asp Ala
 405 410 415

Asp Leu Tyr Gln Lys Leu Lys Lys Asn Val Thr Gln Glu Ala Glu Asn
 420 425 430

Arg Trp Gln Ser Asn Trp Glu Arg Thr Met Arg Asp Leu Lys Leu Ile
 435 440 445

His

<210> 7

<211> 1512

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 7

atgatcgaaa aagacaaaacg taccatcgct ttcacccacc cggacctggg tatcgggtgt	60
gctgaacgtc tgggtgttga cgctgctctg ggtctgcagc agcagggtca ctctgttatac	120
atctacacct ctcaactgca caaatctcac tgcttcgaag aagttaaaaa cggtcagctg	180
aaagttgaag tttacggtga cttctgccc accaacttcc tgggtcgttt cttcatcgtt	240
ttcgtacca tccgtcagct gtacctggtt atccagctga tcctgcagaa aaaagttaac	300
gcttaccagc tgatcatcat cgaccagctg tctacctgca tcccgtgct gcacatcttc	360
tctctgcta ccctgatgtt ctactgccac ttcccggacc agctgctggc tcagcgtgct	420
ggtctgctga aaaaaatcta ccgtctgccg ttcgacctga tcgaacagtt ctctgtttct	480
gctgctgaca ccgttgtgtg taactctaac ttcacaaaa acaccttcca ccagaccttc	540

aaatacctgt ctaacgaccc ggacgttatac taccctgctg ttgacctgtc taccatcgaa 600
 atcgaagaca tcgacaaaaa attcttcaaa accgttttca acgaaggatga ccgtttctac 660

 ctgtctatca accgtttcga aaaaaaaaa gacgttgctc tggtatcaa agctttcgtc 720
 ctgtctgaag accagatcaa cgacaacgtt aaactgggta tctgcgggtg ttacgacgaa 780
 cgtgttgctg aaaacgttga atacctgaaa gaactgcagt ctctggctga cgaatacgaa 840
 ctgtctcaca ccaccatcta ctaccaggaa atcaaactg tttctgacct ggaatctttc 900
 aaaaccaaca actctaaat catcttctg acctctatct cttcttctct gaaagaactg 960
 ctgtctggaac gtaccgaaat gctgctgtac accccggctt acgaacactt cggtatcgtt 1020
 ccgctggaag ctatgaaact gggtaaaccg gttctggtc ttaacaacgg tggccgctg 1080

 gaaacatca aatcttacgt tgctgggtaa aacgaatctt ctgctaccgg ttggctgaaa 1140
 ccgctgttcc cgatccagtg ggctaccgt atcgacgaat ctcgtaaaat cctgcagAAC 1200
 ggttctgtta acttcgaacg taacgggccg ctgctgttga aaaaatactt ctctcgtgaa 1260
 gctatgaccc agtctttcga agaaaacgtt gaaaaagtta tctggaaaga aaaaaaac 1320
 taccctggg aaatcttcgg tatctctttc tctaacttca tctgcacat ggctttcacc 1380
 aaaatctgc cgaacaaccc gtggccgttc ctggtcatgg ctaccttcat ggttctgtac 1440
 ttcaaaaact acctgtgggg tatctactgg gctttcgttt tcgctctgtc ttaccctac 1500

 gaagaaatct aa 1512
 <210> 8
 <211> 503
 <212> PRT
 <213> *Saccharomyces cerevisiae*
 <400> 8
 Met Ile Glu Lys Asp Lys Arg Thr Ile Ala Phe Ile His Pro Asp Leu
 1 5 10 15
 Gly Ile Gly Gly Ala Glu Arg Leu Val Val Asp Ala Ala Leu Gly Leu
 20 25 30
 Gln Gln Gln Gly His Ser Val Ile Ile Tyr Thr Ser His Cys Asp Lys
 35 40 45

 Ser His Cys Phe Glu Glu Val Lys Asn Gly Gln Leu Lys Val Glu Val
 50 55 60
 Tyr Gly Asp Phe Leu Pro Thr Asn Phe Leu Gly Arg Phe Phe Ile Val

Leu Leu Glu Arg Thr Glu Met Leu Leu Tyr Thr Pro Ala Tyr Glu His
 325 330 335

Phe Gly Ile Val Pro Leu Glu Ala Met Lys Leu Gly Lys Pro Val Leu
 340 345 350

Ala Val Asn Asn Gly Gly Pro Leu Glu Thr Ile Lys Ser Tyr Val Ala
 355 360 365

Gly Glu Asn Glu Ser Ser Ala Thr Gly Trp Leu Lys Pro Ala Val Pro
 370 375 380

Ile Gln Trp Ala Thr Ala Ile Asp Glu Ser Arg Lys Ile Leu Gln Asn
 385 390 395 400

Gly Ser Val Asn Phe Glu Arg Asn Gly Pro Leu Arg Val Lys Lys Tyr
 405 410 415

Phe Ser Arg Glu Ala Met Thr Gln Ser Phe Glu Glu Asn Val Glu Lys
 420 425 430

Val Ile Trp Lys Glu Lys Lys Tyr Tyr Pro Trp Glu Ile Phe Gly Ile
 435 440 445

Ser Phe Ser Asn Phe Ile Leu His Met Ala Phe Ile Lys Ile Leu Pro
 450 455 460

Asn Asn Pro Trp Pro Phe Leu Phe Met Ala Thr Phe Met Val Leu Tyr
 465 470 475 480

Phe Lys Asn Tyr Leu Trp Gly Ile Tyr Trp Ala Phe Val Phe Ala Leu
 485 490 495

Ser Tyr Pro Tyr Glu Glu Ile
 500

<210> 9

<211> 1647

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 9

atgggcagtg cttggacaaa ctacaatttt gaagagggtta agtctcattt tgggttcaaa 60

aaatatgttg tatcatcttt agtactagtg tatggactaa ttaaggttct cacgtggatc 120

ttccgtcaat ggggtgtattc cagcttgaat ccgttctcca aaaaatcttc attactgaac 180

agagcagttg cctcctgtgg tgagaagaat gtgaaagttt ttggtttttt tcatccgtat 240
 tgtaatgctg gtgggtgggg ggaaaaagtg ctctggaaag ctgtagatat cactttgaga 300

 aaagatgcta agaacgttat tgtcatttat tcaggggatt ttgtgaatgg agagaatgtt 360
 actccggaga atattctaaa taatgtgaaa gcgaagttag attacgactt ggattcggat 420
 agaatatttt tcatttcatt gaagctaaga tacttgggtgg attcttcaac atggaagcat 480
 ttcacgttga ttggacaagc aattggatca atgattctcg catttgaatc cattattcag 540
 tgtccacctg atatatggat tgatacaatg gggtagcctt tcagctatcc tattattgct 600
 aggtttttga ggagaattcc tatcgtcaca tatacgctt atccgataat gtcaaaagac 660
 atgttaaata agctgttcaa aatgcccag aagggtatca aagtttacgg taaaatatta 720

 tactggaaag tttttatgtt aatttatcaa tccattggtt ctaaaattga tattgtaatc 780
 acaactcaa catggacaaa taaccacata aagcaaattt ggcaatcaa tacgtgtaaa 840
 attatatac ctccatgctc tactgagaaa ttagtagatt ggaagcaaaa gtttggtagt 900
 gcaaagggtg agagattaaa tcaagcaatt gtgtggcac aatttcgtcc tgaaaaact 960
 cataagttaa tcattgagtc ctttgcaact ttcttgaaaa atttaccgga ttctgtatcg 1020
 ccaatfaaa tgataatggc ggggtccact agatccaagc aagatgaaaa ttatgttaaa 1080
 agtttacaag actggtcaga aaatgtatta aaaattccta aacatttgat atcattcgaa 1140

 aaaaatctgc ccttcgataa gattgaaata ttactaaaca aatctacttt cgggtttaat 1200
 gccatgtgga atgagcactt tggaattgca gttgtagagt atatggcttc cggtttgatc 1260
 cccatagttc atgcctcggc gggccatttg ttagatatag ttactccatg ggatgccaac 1320
 gggaaatcgc gaaaagctcc accacaatgg gaggtaaaaa agaaatattt tgcaaaactc 1380
 gaagatgatg gtgaaactac tggatttttc tttaaagagc cgagtgatcc tgattataac 1440
 acaaccaaag atcctctgag ataccctaat ttgtccgacc ttttcttaca aattacgaaa 1500
 ctggactatg actgcctaag ggtgatgggc gcaagaaacc agcagtattc attgtataaa 1560

 ttctctgatt tgaagtttga taaagattgg gaaaactttg tactgaatcc tatttgtaaa 1620
 ttattagaag aggaggaaag gggctga 1647

 <210> 10
 <211> 548
 <212> PRT
 <213> *Saccharomyces cerevisiae*
 <400> 10
 Met Gly Ser Ala Trp Thr Asn Tyr Asn Phe Glu Glu Val Lys Ser His

1 5 10 15
 Phe Gly Phe Lys Lys Tyr Val Val Ser Ser Leu Val Leu Val Tyr Gly
 20 25 30
 Leu Ile Lys Val Leu Thr Trp Ile Phe Arg Gln Trp Val Tyr Ser Ser
 35 40 45
 Leu Asn Pro Phe Ser Lys Lys Ser Ser Leu Leu Asn Arg Ala Val Ala
 50 55 60
 Ser Cys Gly Glu Lys Asn Val Lys Val Phe Gly Phe Phe His Pro Tyr
 65 70 75 80
 Cys Asn Ala Gly Gly Gly Glu Lys Val Leu Trp Lys Ala Val Asp
 85 90 95
 Ile Thr Leu Arg Lys Asp Ala Lys Asn Val Ile Val Ile Tyr Ser Gly
 100 105 110
 Asp Phe Val Asn Gly Glu Asn Val Thr Pro Glu Asn Ile Leu Asn Asn
 115 120 125
 Val Lys Ala Lys Phe Asp Tyr Asp Leu Asp Ser Asp Arg Ile Phe Phe
 130 135 140
 Ile Ser Leu Lys Leu Arg Tyr Leu Val Asp Ser Ser Thr Trp Lys His
 145 150 155 160
 Phe Thr Leu Ile Gly Gln Ala Ile Gly Ser Met Ile Leu Ala Phe Glu
 165 170 175
 Ser Ile Ile Gln Cys Pro Pro Asp Ile Trp Ile Asp Thr Met Gly Tyr
 180 185 190
 Pro Phe Ser Tyr Pro Ile Ile Ala Arg Phe Leu Arg Arg Ile Pro Ile
 195 200 205
 Val Thr Tyr Thr His Tyr Pro Ile Met Ser Lys Asp Met Leu Asn Lys
 210 215 220
 Leu Phe Lys Met Pro Lys Lys Gly Ile Lys Val Tyr Gly Lys Ile Leu
 225 230 235 240
 Tyr Trp Lys Val Phe Met Leu Ile Tyr Gln Ser Ile Gly Ser Lys Ile
 245 250 255

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

500	505	510	
Asn Gln Gln Tyr Ser Leu Tyr Lys Phe Ser Asp Leu Lys Phe Asp Lys			
515	520	525	
Asp Trp Glu Asn Phe Val Leu Asn Pro Ile Cys Lys Leu Leu Glu Glu			
530	535	540	
Glu Glu Arg Gly			
545			
<210> 11			
<211> 1113			
<212> DNA			
<213> Escherichia coli			
<400> 11			
aaaatcgaag aaggtaaact ggtaatctgg attaacggcg ataaaggcta taacggtctc			60
gctgaagtcg gtaagaaatt cgagaaagat accggaatta aagtcaccgt tgagcatccg			120
gataaactgg aagagaaatt cccacaggtt gcgcaactg gcgatggccc tgacattatc			180
ttctgggcac acgaccgctt tgggtgctac gctcaatctg gcctgttggc tgaatcacc			240
ccggacaaag cgttccagga caagctgtat ccgtttacct gggatgccgt acgttacaac			300
ggcaagctga ttgcttacc gatcgtgtt gaagcgttat cgctgattta taacaaagat			360
ctgctgccga acccgccaaa aacctgggaa gagatcccgg cgctggataa agaactgaaa			420
gcgaaaggta agagcgcgct gatgttcaac ctgcaagaac cgtacttcac ctggccgctg			480
attgctgctg acggggggtta tgcgttcaag tatgaaaacg gcaagtacga cattaagac			540
gtggcgctgg ataacgctgg cgcgaaagcg ggtctgacct tcctggttga cctgattaaa			600
aacaacaca tgaatgcaga caccgattac tccatcgag aagctgcctt taataaaggc			660
gaaacagcga tgaccatcaa cggcccgtgg gcatggtcca acatcgacac cagcaaagtg			720
aattatggtg taacggtact gccgacctc aagggtcaac catccaaacc gttcgttggc			780
gtgctgagcg caggtattaa cgccgccagt ccgaacaaag agctggcgaa agagttcctc			840
gaaaactatc tgctgactga tgaaggctctg gaagcgggta ataaagacaa accgctgggt			900
gccgtagcgc tgaagtctta cgaggaagag ttggcgaaag atccacgtat tgccgccacc			960
atgaaaaacg cccagaaagg tgaatcatg ccgaacatcc gcgagatgac cgctttctgg			1020
tatgccgtgc gtactcgggt gatcaacgcc gccagcggtc gtcagactgt cgatgaagcc			1080
ctgaaagacg cgcagactcg taccaccaag taa			1113

<210> 12

<211> 370

<212> PRT

<213> Escherichia coli

<400> 12

Lys Ile Glu Glu Gly Lys Leu Val Ile Trp Ile Asn Gly Asp Lys Gly

1 5 10 15

Tyr Asn Gly Leu Ala Glu Val Gly Lys Lys Phe Glu Lys Asp Thr Gly

 20 25 30

Ile Lys Val Thr Val Glu His Pro Asp Lys Leu Glu Glu Lys Phe Pro

 35 40 45

Gln Val Ala Ala Thr Gly Asp Gly Pro Asp Ile Ile Phe Trp Ala His

 50 55 60

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

65 70 75 80

Pro Asp Lys Ala Phe Gln Asp Lys Leu Tyr Pro Phe Thr Trp Asp Ala

 85 90 95

Val Arg Tyr Asn Gly Lys Leu Ile Ala Tyr Pro Ile Ala Val Glu Ala

 100 105 110

Leu Ser Leu Ile Tyr Asn Lys Asp Leu Leu Pro Asn Pro Pro Lys Thr

 115 120 125

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

 130 135 140

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

145 150 155 160

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

 165 170 175

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

 180 185 190

Thr Phe Leu Val Asp Leu Ile Lys Asn Lys His Met Asn Ala Asp Thr

 195 200 205

Asp Tyr Ser Ile Ala Glu Ala Ala Phe Asn Lys Gly Glu Thr Ala Met

<210> 14

<211> 110

<212> PRT

<213> Bacillus sp.

<400> 14

Met Phe Cys Thr Phe Phe Glu Lys His His Arg Lys Trp Asp Ile Leu

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

65 70 75 80
 Met Tyr Gly Gln Glu Lys Leu Asn Glu Lys Leu Asn Thr Ile Ile Lys
 85 90 95
 Gln Ile Leu Ser Ile Ser Val Ser Glu Glu Gly Glu Lys Glu
 100 105 110

<210> 15

<211> 1254

<212> DNA

<213> Nicotiana tabacum

<400> 15

gcgacacagt cagaatatgc agatgcctt gctgctgcaa ttgaagcaga aatcattgt 60
 acaagccaga ccagattgct tattgaccag attagcctgc agcaaggaag aatagttgct 120

 cttgaagaac aaatgaagcg tcaggaccag gagtgccgac aattaagggc tcttgttcag 180
 gatcttgaaa gtaagggcat aaaaaagttg atcggaaatg tacagatgcc agtggctgct 240
 gtagttgita tggcttgcaa tcgggctgat tacctggaaa agactattaa atccatctta 300
 aaatacaaaa tatctgttgc gtcaaaatat cctcttttca tatcccagga tggatcacat 360
 cctgatgtca ggaagcttgc tttgagctat gatcagctga cgtatatgca gcaacttgat 420
 tttgaacctg tgcatactga aagaccaggg gagctgattg catactacaa aattgcacgt 480

cattacaagt gggcattgga tcagctgttt tacaagcata attttagccg tgttatcata 540

ctagaagatg atatggaat tgcccctgat ttttttgact tttttgaggc tggagctact 600

cttcttgaca gagacaagtc gattatggct atttcttctt ggaatgacaa tggacaaatg 660

cagtttgacc aagatcctta tgccttttac cgctcagatt ttttcccg tcttggatgg 720

atgctttcaa aatctacttg ggacgaatta tctccaaagt ggccaaaggc ttactgggac 780

gactggctaa gactcaaaga gaatcacaga ggtcgacaat ttattcgccc agaagtttgc 840

agaacatata attttggatg gcatggttct agtttggggc agtttttcaa gcagtatctt 900

gagccaatta aactaaatga tgtccaggtt gattggaagt caatggacct tagttacctt 960

ttggaggaca attacgtgaa acactttggt gacttgggta aaaaggctaa gcccatccat 1020

ggagctgatg ctgcttgaa agcatttaac atagatggtg atgtgcgtat tcagtacaga 1080

gatcaactag actttgaaaa tatgcacagg caatttggca tttttgaaga atggaaggat 1140

ggtgtaccac gtgcagcata taaaggaata gtagttttcc ggtaccaaac gtccagacgt 1200

gtattccttg ttggccatga ttcgcttcaa caactcggaat tgaagatac ttaa 1254

<210> 16

<211> 417

<212> PRT

<213> Nicotiana tabacum

<400> 16

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

1 5 10 15

Glu Asn His Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser

20 25 30

Leu Gln Gln Gly Arg Ile Val Ala Leu Glu Glu Gln Met Lys Arg Gln

35 40 45

Asp Gln Glu Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser

50 55 60

Lys Gly Ile Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala

65 70 75 80

Val Val Val Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile

85 90 95

Lys Ser Ile Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu

Gly Asp Val Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile
 355 360 365

Ala Arg Gln Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg
 370 375 380

Ala Ala Tyr Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg

385 390 395 400

Val Phe Leu Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp
 405 410 415

Thr

<210> 17

<211> 1344

<212> DNA

<213> Homo sapiens

<400> 17

atgcgctttc gtatctataa acgtaaagtg ctgacccctga cactggttgt tgccgcttgt 60
 ggttttgttc tgtggagcag taatggctgt cagcgtaaaa atgaagccct ggcacctcct 120
 ctgctggatg ctgaaccggc acgtggctgt ggcggctcgt gtggtgatca tccgtctgtt 180
 gccgttggta ttcgtcgtgt gagcaatggt tcggctgcct ctctggtccc ggctgttctt 240

caacctgaag ctgataacct gaccctgcgc tatecgtctc tgggtgatca actgaacttc 300
 gatcaaacct tgcgtaacct ggataaagca ggcacatggg ctctcctgta actggtactg 360
 gtagtccagg tccataatcg tccggaatat ctgcgtctgc tgcctggattc tctgcgcaaa 420
 gctcaaggca tcgataatgt cctggctcgc ttctctcatg atttctggag cacggagatt 480
 aaccagctga ttgccggcgt gaatttttgt cctgtgctgc aggtgttttt tccgttttct 540
 atccaactgt atccgaacga atttccgggt tctgacccctc gtgattgtcc tcgtgatctg 600
 cctaaaaatg ccgctctgaa actgggctgt attaatagccg agtatcctga ttcttttggc 660

cactatcgtg aggcgaaatt ttctcagacc aaacatcatt ggtggaggaa actgcatttc 720
 gtgtgggaac gtgtgaaaat cctgcgcgac tatgctggcc tgattctgtt tctggaagaa 780
 gatcactatc tggctccgga cttttatcat gtgttcaaaa aaatgtgaa actgaaacag 840
 caggaatgct cagaatgtga tgtgctgtca ctgggcacct atagtcttc tcgctccttc 900
 tatggtatgg ccgacaaagt ggacgttaaa acatggaaat ccaccgagca caacatgggt 960
 ctggcactga ctcgtaatgc ctatcaaaaa ctgattgagt gtaccgacac cttttgtacg 1020

tatgatgact ataactggga ctggaccctg caatatctga ccgtgagctg tctgcaaaa 1080

ttttggaaag ttctggtgcc tcagattcct cgtatcttcc atgctggcga ctgtggtatg 1140

caccataaaa aaacttgccg tccgtcaaca caatctgctc agatcgagtc gctgctgaat 1200

aatacaaac agtatatggt cccggagact ctgacaattt ctgaaaaatt cacctgggtc 1260

gccatttctc cgcctcgtaa aaatggaggt tggggcgata tccgtgacca tgaactgtgt 1320

aaaagctatc gtcgtctgca gtga 1344

<210> 18

<211> 447

<212> PRT

<213> Homo sapiens

<400> 18

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

1 5 10 15

Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg

20 25 30

Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg

35 40 45

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

50 55 60

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

65 70 75 80

Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr

85 90 95

Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Thr

100 105 110

Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro

115 120 125

Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile

130 135 140

Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile

145 150 155 160

Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
 165 170 175
 Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
 180 185 190
 Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
 195 200 205
 Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
 210 215 220
 Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
 225 230 235 240
 Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu
 245 250 255
 Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
 260 265 270
 Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
 275 280 285
 Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
 290 295 300
 Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320
 Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335
 Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350
 Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365
 Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380
 Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400
 Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys

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

<210> 19

<211> 1329

<212> DNA

<213> Bos taurus

<

400> 19

```

ttgaaagaac tgacgtccaa aaagagcttg caagtcccgt ccatctacta tcaacttgccg      60
cacttgctgc aaaacgaggg ctctttgcaa cggcagttc agatcggcaa tggctgcacc      120
ggcgtgagca ttgttatggg tatcccgacc gtgaaactgt aagtaaaag ctatctgatt      180
gaaacgctgc atagcctgat cgataacctg taccgggaag aaaaactgga ctgcgtgatt      240
gtcgttttca ttggtgaaac cgacacggat tatgtgaatg gcgttgttgc caatctggaa      300
aaagatttca gcaaagatat cagcagcggc ctggttgaga tcattttctcc gccggagagc      360
tattaccggg atctgacgaa cctgaaagaa accttcggtg atagcaaaga gcgtgtccgt      420

tggcgcaact agcagaacct ggactattgt tttctgatga tgtacgcga agaaaagggt      480
acgtattaca tccaactgga ggacgacatt attgtgaagc aaaactactt caacaccatt      540
aagaacttgc cgctgcagct gagcagcga gagtggatga ttctggagtt cagccagctg      600
ggcttcattg gcaagatggt tcaggcaccg gacttgacc tgatcgtgga gtttatcttt      660
atgttctaca aagagaaacc gatcgattgg ctgctggatc atacctgtg ggtcaaggtc      720
tgcaatccgg aaaaagatgc caagcattgt gaccgccaga aagcgaatct gcgtattcgt      780
tttctgctca gcctgttcca acacgtgggt ctgcacagct ctctgaccgg taagatccaa      840

aagctgaccg acaaagatta catgaaaccg ctgctgctga agatccatgt caaccgcca      900
gcagaggatg gcacctcgt gaaagtctac cagggtcaca ctctggagaa aacctatag      960
ggcaggact tcttttgggc gattacgct gttcgggtg actatatctt gtttaagtt      1020
gacaagccgg ttaatgtaga gagctacttg tttcatagcg gtaaccagga tcaccaggt      1080
gacattctgc tgaacaccac cgttgaagtg ttgccgctga aaagcgaagg tctggatatt      1140
tcgaaagaaa cgaaggataa gcgtctggag gatggttact tccgtatcgg caagttcgag      1200
aatggcgtgg ctgaaggtat ggtcgaccgg agcctgaacc cgatttccgc atttcgctg      1260

```

tccgctcatcc agaatagcgc ggtttgggct atcctgaatg agattcacat caaaaagggtt 1320

acgaattaa 1329

<210> 20

<211> 443

<212> PRT

<213> Bos taurus

<400> 20

Ile Leu Lys Glu Leu Thr Ser Lys Lys Ser Leu Gln Val Pro Ser Ile

1 5 10 15

Tyr Tyr His Leu Pro His Leu Leu Gln Asn Glu Gly Ser Leu Gln Pro

20 25 30

Ala Val Gln Ile Gly Asn Gly Arg Thr Gly Val Ser Ile Val Met Gly

35 40 45

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

50 55 60

His Ser Leu Ile Asp Asn Leu Tyr Pro Glu Glu Lys Leu Asp Cys Val

65 70 75 80

Ile Val Val Phe Ile Gly Glu Thr Asp Thr Asp Tyr Val Asn Gly Val

85 90 95

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

100 105 110

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

115 120 125

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

130 135 140

Lys Gln Asn Leu Asp Tyr Cys Phe Leu Met Met Tyr Ala Gln Glu Lys

145 150 155 160

Gly Thr Tyr Tyr Ile Gln Leu Glu Asp Asp Ile Ile Val Lys Gln Asn

165 170 175

Tyr Phe Asn Thr Ile Lys Asn Phe Ala Leu Gln Leu Ser Ser Glu Glu

180 185 190

Trp Met Ile Leu Glu Phe Ser Gln Leu Gly Phe Ile Gly Lys Met Phe

<210> 21

<211> 822

<212> DNA

<213> Helicobacter pylori

<400> 21

```

atgcgtgtct ttattatcag tctgaaccag aaagtgtgtg acaaattcgg cctggtgttt      60
cgtgatacca caaccctgct gaataacatc aatgccaccc gccacaaagc acagattttt      120
gacgccgtct atagcaaaac gttcgaaggt gggctgcatc cactggtgaa aaaacatctg      180
caccctgatt tcattacca gaacatcaaa gacatgggca ttaccaccaa cctgattagc      240

ggtgtatcca aattctatta tgctctgaaa tatcacgcca aattcatgag cctgggcgaa      300
ctgggctgtt atgccagcca ttatagcctg tgggagaaat gtattgagct gaacgaggcc      360
atgtgtatcc tggaagatga cattacgctg aaagaagatt tcaaagaggg cctggatttc      420
ctggaaaaac acattcagga gctgggctat gttcgtctga tgcatctgct gtatgatgcc      480
tccgttaaaa gcgaacctct gtccataaaa aaccacgaga tcaagagcg tgcgggatac      540
attaaagctt atagtcacgg tgttggcact cagggatatg tgattactcc gaaaattgcc      600
aaagtgttca aaaaatgctc ccgtaaatgg gttgttccgg tggatacgat catggatgcc      660

acgtttattc atggggtgaa aaacctggta ctgcaaccgt ttgtgattgc cgatgatgag      720
caaatttcca cgattgtccg taaagaggag ccgtattccc ctaaattgc cctgatgcgc      780
gaactgcact tcaaatatct gaaatattgg cagtttgtgt ga                          822
    
```

<210> 22

<211> 273

<212> PRT

<213> Helicobacter pylori

<400> 22

```

Met Arg Val Phe Ile Ile Ser Leu Asn Gln Lys Val Cys Asp Lys Phe
1           5           10          15
Gly Leu Val Phe Arg Asp Thr Thr Thr Leu Leu Asn Asn Ile Asn Ala
           20           25           30

Thr Arg His Lys Ala Gln Ile Phe Asp Ala Val Tyr Ser Lys Thr Phe
           35           40           45
Glu Gly Gly Leu His Pro Leu Val Lys Lys His Leu His Pro Tyr Phe
           50           55           60
    
```


<400> 23

atgaaaaaat taacctgctt taaagcctat gatattcgcg ggaaattagg cgaagaactg 60

aatgaagata tgcctggcg cattggtcgc gcctatggcg aatttctcaa accgaaaacc 120

attgtgttag gcggtgatgt ccgcctcacc agcgaaacct taaaactggc gctggcgaaa 180

ggtttacagg atgcgggctg tgacgtgctg gatattggta tgtccggcac cgaagagatc 240

tatttcgcca cgttccatct cggcgtggat ggcgccattg aagtaccgc cagccataat 300

ccgatggatt ataacggcat gaagctggtt cgcgaggggg ctcgccgat cagcggagat 360

accggactgc gcgacgtcca gcgtctggct gaagccaacg actttcctcc cgtcgatgaa 420

accaaacgcg gtcgctatca gcaaatcaac ctgcgtgacg cttacgttga tcacctgttc 480

ggttataatca atgcaaaaa cctcacgccg ctcaagctgg tgatcaactc cgggaacggc 540

gcagcgggtc cggtggtgga cgccattgaa gcccgcttta aagccctcgg cgcgccctg 600

gaattaatca aagtcacaa cacgccggac ggcaatttcc ccaacggtat tcctaacca 660

ctactgccgg aatgccgca cgacaccgc aatgcggtca tcaaacacgg cgcggatag 720

ggcattgctt ttgatggcga ttttgaccgc tgtttcctgt ttgacgaaaa agggcagttt 780

attgagggct actacattgt cggcctgttg gcagaagcat tcctcgaaaa aaatcccggc 840

gcgaagatca tccacgatcc acgtctctcc tggaacaccg ttgatgtggt gactgccgca 900

ggtggcacgc cggtaatgtc gaaaaccgga cacgccttta ttaaagaacg tatgcgcaag 960

gaagacgcca tctatggtgg cgaaatgagc gccaccatt acttccgtga tttcgcttac 1020

tgcgacagcg gcatgatccc gtgctgctg gtcgccgaac tgggtgctcct gaaagataaa 1080

acgtgggctg aactggtacg cgaccggatg gcggcgttcc cggcaagcgg tgagatcaac 1140

agcaaaactg cgcaaccctg tgaggcgatt aaccgcgtgg aacagcattt tagccgtgag 1200

gcgctggcgg tggatcgcac cgatggcatc agcatgacct ttgccgactg gcgctttaac 1260

ctgcgcacct ccaataccga accggtggtg cgcctgaatg tggaatcgcg cggatgatg 1320

ccgctgatgg aagcgcgaac gcgaactctg ctgacgttgc tgaacgagta a 1371

<210> 24

<211> 456

<212> PRT

<213> Escherichia coli

<400> 24

Met Lys Lys Leu Thr Cys Phe Lys Ala Tyr Asp Ile Arg Gly Lys Leu

1 5 10 15

Gly Glu Glu Leu Asn Glu Asp Ile Ala Trp Arg Ile Gly Arg Ala Tyr
 20 25 30
 Gly Glu Phe Leu Lys Pro Lys Thr Ile Val Leu Gly Gly Asp Val Arg
 35 40 45

 Leu Thr Ser Glu Thr Leu Lys Leu Ala Leu Ala Lys Gly Leu Gln Asp
 50 55 60
 Ala Gly Val Asp Val Leu Asp Ile Gly Met Ser Gly Thr Glu Glu Ile
 65 70 75 80
 Tyr Phe Ala Thr Phe His Leu Gly Val Asp Gly Gly Ile Glu Val Thr
 85 90 95
 Ala Ser His Asn Pro Met Asp Tyr Asn Gly Met Lys Leu Val Arg Glu
 100 105 110

 Gly Ala Arg Pro Ile Ser Gly Asp Thr Gly Leu Arg Asp Val Gln Arg
 115 120 125
 Leu Ala Glu Ala Asn Asp Phe Pro Pro Val Asp Glu Thr Lys Arg Gly
 130 135 140
 Arg Tyr Gln Gln Ile Asn Leu Arg Asp Ala Tyr Val Asp His Leu Phe
 145 150 155 160
 Gly Tyr Ile Asn Val Lys Asn Leu Thr Pro Leu Lys Leu Val Ile Asn
 165 170 175

 Ser Gly Asn Gly Ala Ala Gly Pro Val Val Asp Ala Ile Glu Ala Arg
 180 185 190
 Phe Lys Ala Leu Gly Ala Pro Val Glu Leu Ile Lys Val His Asn Thr
 195 200 205
 Pro Asp Gly Asn Phe Pro Asn Gly Ile Pro Asn Pro Leu Leu Pro Glu
 210 215 220
 Cys Arg Asp Asp Thr Arg Asn Ala Val Ile Lys His Gly Ala Asp Met
 225 230 235 240

 Gly Ile Ala Phe Asp Gly Asp Phe Asp Arg Cys Phe Leu Phe Asp Glu
 245 250 255
 Lys Gly Gln Phe Ile Glu Gly Tyr Tyr Ile Val Gly Leu Leu Ala Glu

260	265	270	
Ala Phe Leu Glu Lys Asn Pro Gly Ala Lys Ile Ile His Asp Pro Arg			
275	280	285	
Leu Ser Trp Asn Thr Val Asp Val Val Thr Ala Ala Gly Gly Thr Pro			
290	295	300	
Val Met Ser Lys Thr Gly His Ala Phe Ile Lys Glu Arg Met Arg Lys			
305	310	315	320
Glu Asp Ala Ile Tyr Gly Gly Glu Met Ser Ala His His Tyr Phe Arg			
	325	330	335
Asp Phe Ala Tyr Cys Asp Ser Gly Met Ile Pro Trp Leu Leu Val Ala			
	340	345	350
Glu Leu Val Cys Leu Lys Asp Lys Thr Leu Gly Glu Leu Val Arg Asp			
	355	360	365
Arg Met Ala Ala Phe Pro Ala Ser Gly Glu Ile Asn Ser Lys Leu Ala			
	370	375	380
Gln Pro Val Glu Ala Ile Asn Arg Val Glu Gln His Phe Ser Arg Glu			
385	390	395	400
Ala Leu Ala Val Asp Arg Thr Asp Gly Ile Ser Met Thr Phe Ala Asp			
	405	410	415
Trp Arg Phe Asn Leu Arg Thr Ser Asn Thr Glu Pro Val Val Arg Leu			
	420	425	430
Asn Val Glu Ser Arg Gly Asp Val Pro Leu Met Glu Ala Arg Thr Arg			
	435	440	445
Thr Leu Leu Thr Leu Leu Asn Glu			
	450	455	
<210> 25			
<211> 1437			
<212> DNA			
<213> Escherichia coli			
<400> 25			
atggcgcagt cgaaactcta tccagttgtg atggcaggtg gtcctcggtag ccgettatgg			60
ccgctttccc gcgtacttta tccaagcag tttttatgcc tgaaaggcga tctcaccatg			120

ctgcaaacca ccatctgccg cctgaacggc gtggagtgcg aaagcccggg ggtgatttgc 180
aatgagcagc accgctttat tgtcgcggaa cagctgcgtc aactgaacaa acttaccgag 240

aacattattc tcgaaccggc agggcgaaac acggcacctg ccattgcgct ggccggcgtg 300
gcgcaaaac gtcatagccg ggagagcgac ccgttaatgc tggatttggc ggccggatcat 360
gtgattgccg atgaagacgc gttccgtgcc gccgtgcgta atgccatgcc atatgccgaa 420
gcccccaagc tggtagcctt cggcattgtg ccggatctac cagaaccggg ttatggctat 480
attcgtcgcg gtgaagtgtc tgcgggtgag caggatatgg tggcctttga agtggcgcag 540
tttgtcgaac aaccgaatct ggaaaccgct caggcctatg tggcaagcgg cgaatattac 600
tggaaacagc gtatgttctt gttccgcgcc ggacgctatc tcgaagaact gaaaaaatat 660

cgccccgata tctcctgatc ctgtgaaaaa gcgatgagcg ccgtcgtacc ggatctcaat 720
tttattcgcg tggatgaaga agcgtttctc gcctgcccgg aagagtcggt ggattacgag 780
gtcatggaac gtacggcaga tgctgttgtg gtgccgatgg atgccccgct gagcgtggtt 840
ggctcctggt cttcattatg ggagatcagc gcccacaccg ccgagggcaa cgtttgccac 900
ggcgtatgta ttaatcacia aactgaaaac agctatgtgt atgctgaatc tggcctggtc 960
accaccgtcg gggtagaaga tctggtagtg gtgcagacca aagatgcggt gctgattgcc 1020
gaccgtaacg cggtagagga tgtgaaaaaa gtggtcgagc agatcaaagc cgatggtcgc 1080

catgagcadc gggtagatcg cgaagtgtat cgtccgtggg gcaaatatga ctctatcgac 1140
gccccgcgacc gctaccaggt gaaacgcadc accgtgaaac cgggcgaggg cttgtcggta 1200
cagatgcacc atcaccgcgc ggaacactgg gtggttgtcg cgggaacggc aaaagtcacc 1260
attgatggtg atatacaact gcttggtaga aacagatcca tttatattcc gctgggggag 1320
acgcattgcc tggaaaaccg ggggaaaatt ccgctcgatt taattgaagt gcgctccggc 1380
tcttatctcg aagaggatga tgtggtgcgt ttcgaggatc gctacggagc ggtgtaa 1437

<210> 26

<211> 478

<212> PRT

<213> Escherichia coli

<400> 26

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

1 5 10 15

Ser Arg Leu Trp Pro Leu Ser Arg Val Leu Tyr Pro Lys Gln Phe Leu

20 25 30

Cys Leu Lys Gly Asp Leu Thr Met Leu Gln Thr Thr Ile Cys Arg Leu
 35 40 45

Asn Gly Val Glu Cys Glu Ser Pro Val Val Ile Cys Asn Glu Gln His
 50 55 60

Arg Phe Ile Val Ala Glu Gln Leu Arg Gln Leu Asn Lys Leu Thr Glu
 65 70 75 80

Asn Ile Ile Leu Glu Pro Ala Gly Arg Asn Thr Ala Pro Ala Ile Ala
 85 90 95

Leu Ala Ala Leu Ala Ala Lys Arg His Ser Pro Glu Ser Asp Pro Leu
 100 105 110

Met Leu Val Leu Ala Ala Asp His Val Ile Ala Asp Glu Asp Ala Phe
 115 120 125

Arg Ala Ala Val Arg Asn Ala Met Pro Tyr Ala Glu Ala Gly Lys Leu
 130 135 140

Val Thr Phe Gly Ile Val Pro Asp Leu Pro Glu Thr Gly Tyr Gly Tyr
 145 150 155 160

Ile Arg Arg Gly Glu Val Ser Ala Gly Glu Gln Asp Met Val Ala Phe
 165 170 175

Glu Val Ala Gln Phe Val Glu Lys Pro Asn Leu Glu Thr Ala Gln Ala
 180 185 190

Tyr Val Ala Ser Gly Glu Tyr Tyr Trp Asn Ser Gly Met Phe Leu Phe
 195 200 205

Arg Ala Gly Arg Tyr Leu Glu Glu Leu Lys Lys Tyr Arg Pro Asp Ile
 210 215 220

Leu Asp Ala Cys Glu Lys Ala Met Ser Ala Val Asp Pro Asp Leu Asn
 225 230 235 240

Phe Ile Arg Val Asp Glu Glu Ala Phe Leu Ala Cys Pro Glu Glu Ser
 245 250 255

Val Asp Tyr Ala Val Met Glu Arg Thr Ala Asp Ala Val Val Val Pro
 260 265 270

Met Asp Ala Gly Trp Ser Asp Val Gly Ser Trp Ser Ser Leu Trp Glu

ggtcatatga cccgectgcg tcgcctcggg aaagtccaga tgctggcaca ggcagcggaa 180

 gaacatcctc tgcattggcg cactggtatt gctcacactc gctgggcgac ccacggtgaa 240
 ccttcagaag tgaatgcgca tccgatggtt tctgaacaca ttgtggtggt gcataacggc 300
 atcatcgaaa accatgaacc gctgcgtgaa gagctaaaag cgcgtggcta taccttcgtt 360
 tctgaaaccg acaccgaagt gattgcccat ctggtgaact gggagctgaa acaaggcggg 420
 actctgcgtg aggccgttct gcgtgctatc ccgcagctgc gtggtgcgta cggtagcagtg 480
 atcatggact cccgtcaccg ggataccctg ctggcggcac gttctggtag tccgctggtg 540
 attggcctgg ggatggcgca aaactttatc gcttctgacc agctggcgct gttgccggtg 600

 acccgtcgtt ttatcttctc tgaagagggc gatattgcgg aaatcactcg ccgttcggta 660
 aacatcttcg ataaaactgg cgcggaagta aaacgtcagg atatcgaatc caatctgcaa 720
 tatgacgcgg gcgataaagg cattaccgt cactacatgc agaaagagat ctacgaacag 780
 ccgaacgcga tcaaaaacac ccttaccgga cgcacagcc acggtcaggt tgatttaagc 840
 gagctgggac cgaacgccga cgaactgctg tcgaaggttg agcatattca gatcctgcc 900
 tgtggtactt cttataactc cggtaggtt tcccgtact ggtttgaatc gctagcaggt 960
 attccgtgcg acgtcgaat cgcctctgaa tccgctatc gcaaatctgc cgtgcgtcgt 1020

 aacagcctga tgatcacctt gtcacagtct ggcgaaaccg cggataccct ggcggcctg 1080
 cgtctgtcga aagagctggg ttaccttggg tcaactggca tctgtaactg tccgggttct 1140
 tctctggtgc gcgaatccga tctggcgcta atgaccaacg cgggtacaga aatcggcgtg 1200
 gcatcacta aagcattcac cactcagtta actgtgctgt tgatgctggt ggccaagctg 1260
 tctcgcctga aaggtctgga tgcctccatt gaacatgaca tcgtgcatgg tctgcaggcg 1320
 ctgccgagcc gtattgagca gatgctgtct caggacaacg gcattgaagc gctggcagaa 1380
 gatttctctg acaaacatca cgcgctgttc ctgggccgtg gcgatcagta cccaatcgcg 1440

 ctggaaggcg cattgaagtt gaaagagatc tcttaccatc acgctgaagc ctacgctgct 1500
 ggccaactga aacacggtcc gctggcgcta attgatgccg atatgccggt tattgttgtt 1560
 gcaccgaaca acgaattgct ggaaaaactg aaatccaaca ttgaagaagt tcgcgcgct 1620
 ggccgtcagt tgtatgtctt cgccgatcag gatcggggtt ttgtaagtag cgataacatg 1680
 cacatcatcg agatgccgca tgtggaagag gtgattgcac cgatcttcta caccgttccg 1740
 ctgcagctgc tggettacca tgtcgcgctg atcaaaggca ccgacgttga ccagccgct 1800
 aacctggcaa aatcggttac ggttgagtaa 1830

<210> 28

<211> 609

<212> PRT

<213> Escherichia coli

<400> 28

Met Cys Gly Ile Val Gly Ala Ile Ala Gln Arg Asp Val Ala Glu Ile

1 5 10 15

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

 20 25 30

Gly Leu Ala Val Val Asp Ala Glu Gly His Met Thr Arg Leu Arg Arg

 35 40 45

Leu Gly Lys Val Gln Met Leu Ala Gln Ala Ala Glu Glu His Pro Leu

 50 55 60

His Gly Gly Thr Gly Ile Ala His Thr Arg Trp Ala Thr His Gly Glu

65 70 75 80

Pro Ser Glu Val Asn Ala His Pro His Val Ser Glu His Ile Val Val

 85 90 95

Val His Asn Gly Ile Ile Glu Asn His Glu Pro Leu Arg Glu Glu Leu

 100 105 110

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

 115 120 125

Ala His Leu Val Asn Trp Glu Leu Lys Gln Gly Gly Thr Leu Arg Glu

 130 135 140

Ala Val Leu Arg Ala Ile Pro Gln Leu Arg Gly Ala Tyr Gly Thr Val

145 150 155 160

Ile Met Asp Ser Arg His Pro Asp Thr Leu Leu Ala Ala Arg Ser Gly

 165 170 175

Ser Pro Leu Val Ile Gly Leu Gly Met Gly Glu Asn Phe Ile Ala Ser

 180 185 190

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

 195 200 205

Glu Gly Asp Ile Ala Glu Ile Thr Arg Arg Ser Val Asn Ile Phe Asp

210 215 220
 Lys Thr Gly Ala Glu Val Lys Arg Gln Asp Ile Glu Ser Asn Leu Gln
 225 230 235 240
 Tyr Asp Ala Gly Asp Lys Gly Ile Tyr Arg His Tyr Met Gln Lys Glu

 245 250 255
 Ile Tyr Glu Gln Pro Asn Ala Ile Lys Asn Thr Leu Thr Gly Arg Ile
 260 265 270
 Ser His Gly Gln Val Asp Leu Ser Glu Leu Gly Pro Asn Ala Asp Glu
 275 280 285
 Leu Leu Ser Lys Val Glu His Ile Gln Ile Leu Ala Cys Gly Thr Ser
 290 295 300
 Tyr Asn Ser Gly Met Val Ser Arg Tyr Trp Phe Glu Ser Leu Ala Gly

 305 310 315 320
 Ile Pro Cys Asp Val Glu Ile Ala Ser Glu Phe Arg Tyr Arg Lys Ser
 325 330 335
 Ala Val Arg Arg Asn Ser Leu Met Ile Thr Leu Ser Gln Ser Gly Glu
 340 345 350
 Thr Ala Asp Thr Leu Ala Gly Leu Arg Leu Ser Lys Glu Leu Gly Tyr
 355 360 365
 Leu Gly Ser Leu Ala Ile Cys Asn Val Pro Gly Ser Ser Leu Val Arg

 370 375 380
 Glu Ser Asp Leu Ala Leu Met Thr Asn Ala Gly Thr Glu Ile Gly Val
 385 390 395 400
 Ala Ser Thr Lys Ala Phe Thr Thr Gln Leu Thr Val Leu Leu Met Leu
 405 410 415
 Val Ala Lys Leu Ser Arg Leu Lys Gly Leu Asp Ala Ser Ile Glu His
 420 425 430
 Asp Ile Val His Gly Leu Gln Ala Leu Pro Ser Arg Ile Glu Gln Met

 435 440 445
 Leu Ser Gln Asp Lys Arg Ile Glu Ala Leu Ala Glu Asp Phe Ser Asp
 450 455 460

Lys His His Ala Leu Phe Leu Gly Arg Gly Asp Gln Tyr Pro Ile Ala
 465 470 475 480
 Leu Glu Gly Ala Leu Lys Leu Lys Glu Ile Ser Tyr Ile His Ala Glu
 485 490 495
 Ala Tyr Ala Ala Gly Glu Leu Lys His Gly Pro Leu Ala Leu Ile Asp
 500 505 510
 Ala Asp Met Pro Val Ile Val Val Ala Pro Asn Asn Glu Leu Leu Glu
 515 520 525
 Lys Leu Lys Ser Asn Ile Glu Glu Val Arg Ala Arg Gly Gly Gln Leu
 530 535 540
 Tyr Val Phe Ala Asp Gln Asp Ala Gly Phe Val Ser Ser Asp Asn Met
 545 550 555 560
 His Ile Ile Glu Met Pro His Val Glu Glu Val Ile Ala Pro Ile Phe
 565 570 575
 Tyr Thr Val Pro Leu Gln Leu Leu Ala Tyr His Val Ala Leu Ile Lys
 580 585 590
 Gly Thr Asp Val Asp Gln Pro Arg Asn Leu Ala Lys Ser Val Thr Val
 595 600 605
 Glu

<210> 29

<211> 2028

<212> DNA

<213> Photobacterium damsela

<400> 29

atgaaaaaaaa tctgaccgt gctgtccatc tttatcctgt ctgcctgtaa tagcgacaat 60
 accagcctga aagagactgt tagcagcaat tcagcggatg ttgtggaaac cgaacttat 120

 caactgacgc cgatcgatgc tccttcttcg ttcctgagcc attcttggga acagacctgt 180
 ggtacaccaa ttcigaacga gtccgacaaa caggccattt ccttcgattt tgttgccccg 240
 gaactgaac aagacgagaa atattgcttc accttcaaag gcattaccgg tgatcatcgt 300
 tatatcacga acaccactct gactgtcgtg gcaccgacac tggaagtgtg tatcgaccat 360
 gccagcctgc ctagtctgca gcaactgatc catattatcc aggcgaaaga cgaatatccg 420

agcaaccagc gttttgtgag ctggaaacgt gttactgtgg atgccgacaa cgccaataaa 480
 ctgaacattc acacctatcc tctgaaaggc aataacacca gccctgagat ggtagcggcg 540

 attgatgagt atgcccagag caaaaacgt ctgaacattg agttctatac caatacggcc 600
 cacgtgttta ataacctgcc gccaatcatt caacctctgt ataacaacga gaaagtgaaa 660
 atcagccaca tttcgctgta tgatgatggc agtagcagat atgttagcct gtatcagtgg 720
 aaagacacc cgaataaaat cgagactctg gaggtgaag tttctctgct ggccaactat 780
 ctggccggta caagtctga tgctccgaaa gggatgggta accgctataa ttggcacaaa 840
 ctgtatgaca ccgactatta ttttctgcgc gaggattatc tggacgtgga agccaatctg 900
 catgatctgc gcgattatct gggttctagc gccaaacaaa tgccgtggga tgaatttgct 960

 aaactgtccg attctcagca aacctgttc ctggacatcg ttggctttga taaagagcag 1020
 ctgcaacagc agtatagcca gtcaccgctg ccgaacttca ttttactgg caccaccaca 1080
 tgggcagggg gtgagacaaa agagtattat gctcaacaac aggtgaacgt catcaacaat 1140
 gccattaacg aaacctcccc atattatctg ggtaaagact atgacctgtt ctttaaaggc 1200
 catccggctg gaggagtgat taatgatatt atcctgggct cctttcctga catgattaac 1260
 attccggcga aaatctcatt tgagggtctg atgatgactg atatgctgcc ggataccgtt 1320
 gctggaattg cctcttcct gtatttcacc attcctgccg acaaagtga cttcatctg 1380

 ttcaccagca gtgataccat tacagaccgt gaagaagcgc tgaatctcc tctggttcag 1440
 gtgatgctga cactgggtat cgtgaaagaa aaagacgtcc tgttttgggc cgaccataaa 1500
 gtgaatagca tggaggtggc catcgacgaa gcgtgtactc gtattatcgc caaacgtcag 1560
 cctaccgctt cagatctcgc tctggttate gccattatca aaacgatcac cgatctggag 1620
 cgtattggag atgttgccga aagcattgcc aaagttgcc tggagagctt ttctaacaaa 1680
 cagtataatc tgctggctag cctggaatct ctgggtcaac acaccgttcg tatgctgcat 1740
 gaagtgctgg atgcttttgc ccgtatggat gtgaaagcag ccattgaagt ctatcaggag 1800

 gatgaccgta tcgatcagga atatgagagc attgtccgtc aactgatggc ccatatgatg 1860
 gaagatccgt ctagcattcc gaatgtgatg aaagtgatgt gggcagctcg tagtattgaa 1920
 cgtgtgggtg accgctgcca gaacatttgt gagtatatca tctatttctg aaaaggcaaa 1980
 gatgttcccc acaccaaacc ggatgacttc ggtactatgc tggactga 2028

 <210> 30
 <211> 675
 <212> PRT
 <213> Photobacterium damselaе

<400> 30

Met Lys Lys Ile Leu Thr Val Leu Ser Ile Phe Ile Leu Ser Ala Cys
 1 5 10 15

Asn Ser Asp Asn Thr Ser Leu Lys Glu Thr Val Ser Ser Asn Ser Ala
 20 25 30

Asp Val Val Glu Thr Glu Thr Tyr Gln Leu Thr Pro Ile Asp Ala Pro
 35 40 45

Ser Ser Phe Leu Ser His Ser Trp Glu Gln Thr Cys Gly Thr Pro Ile
 50 55 60

Leu Asn Glu Ser Asp Lys Gln Ala Ile Ser Phe Asp Phe Val Ala Pro
 65 70 75 80

Glu Leu Lys Gln Asp Glu Lys Tyr Cys Phe Thr Phe Lys Gly Ile Thr
 85 90 95

Gly Asp His Arg Tyr Ile Thr Asn Thr Thr Leu Thr Val Val Ala Pro
 100 105 110

Thr Leu Glu Val Tyr Ile Asp His Ala Ser Leu Pro Ser Leu Gln Gln
 115 120 125

Leu Ile His Ile Ile Gln Ala Lys Asp Glu Tyr Pro Ser Asn Gln Arg
 130 135 140

Phe Val Ser Trp Lys Arg Val Thr Val Asp Ala Asp Asn Ala Asn Lys
 145 150 155 160

Leu Asn Ile His Thr Tyr Pro Leu Lys Gly Asn Asn Thr Ser Pro Glu
 165 170 175

Met Val Ala Ala Ile Asp Glu Tyr Ala Gln Ser Lys Asn Arg Leu Asn
 180 185 190

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

Ile Ile Gln Pro Leu Tyr Asn Asn Glu Lys Val Lys Ile Ser His Ile
 210 215 220

Ser Leu Tyr Asp Asp Gly Ser Ser Glu Tyr Val Ser Leu Tyr Gln Trp
 225 230 235 240

atgaaaatcg aagaaggtaa actggtaatc tggattaacg gcgataaagg ctataacggt 60
ctcgctgaag tccgtaagaa attcagagaaa gataccggaa ttaaagtcac cgttgagcat 120
ccggataaac tggaaagaaa attcccacag gttcgggcaa ctggcgatgg ccctgacatt 180

atcttctggg cacacgaccg ctttgggtggc tacgctcaat ctggcctggt ggctgaaatc 240
accccggaca aagcgttcca ggacaagctg tatccgttta cctgggatgc cgtacgttac 300
aacggcaagc tgattgctta cccgatcgtc gttgaagcgt tatcgtgat ttataacaaa 360
gatctgctgc cgaaccgcc aaaaacctgg gaagagatcc cggcgtgga taaagaactg 420
aaagcgaag gtaagagcgc gctgatgttc aacctgcaag aacctgactt cacctggccg 480
ctgattgctg ctgacggggg ttatgcttc aagtatgaaa acggcaagta cgacattaaa 540
gacgtgggcg tggataacgc tggcgcgaaa gccggctctga ccttcctggt tgacctgatt 600

aaaaacaac acatgaatgc agacaccgat tactccatcg cagaagctgc ctttaataaa 660
ggcgaacag cgatgacat caacggcccg tggcatggt ccaacatcga caccagcaaa 720
gtgaattatg gtgtaacggt actgccgacc ttcaagggtc aaccatcaa accgttcgtt 780
ggcgtgctga gcgaggat taacgcccc agtccgaaca aagagctggc gaaagagttc 840
ctcgaaaact atctgctgac tgatgaaggt ctggaagcgg ttaataaaga caaacctg 900
ggtgccgtag cgctgaagtc ttacaggaa gaggttggcga aagatccacg tattgccgcc 960
acatggaac acgcccagaa aggtgaaatc atgccgaaca tcccgcagat gtccgctttc 1020

tggtatgccg tgcgtactgc ggtgatcaac gcccccagcg gtcgtcagac tgcgatgaa 1080
gccctgaaag acgcgagac tcgtatcacc aaggcgacac agtcagaata tgcagatcgc 1140
cttctgctg caattgaagc agaaaatcat tgtacaagcc agaccagatt gcttattgac 1200
cagattagcc tgcagcaagg aagaatagtt gctcttgaag aacaaatgaa gcgtcaggac 1260
caggagtccc gacaattaag ggctcttgtt caggatcttg aaagtaaggg cataaaaaag 1320
ttgatcggaa atgtacagat gccagtggct gctgtagtgt ttatggcttg caatcgggct 1380
gattacctgg aaaagactat taaatccatc ttaaataacc aaatatctgt tgcgtcaaaa 1440

tatctcttt tcatatccca ggatggatca catcctgatg tcaggaagct tgctttgagc 1500
tatgatcagc tgacgtatat gcagcaactg gattttgaa ctgtgcatac tgaagacca 1560
ggggagctga ttgcatacta caaaattgca cgtcattaca agtgggcatt ggatcagctg 1620
ttttacaagc ataattttag ccgtgttacc atactagaag atgatatgga aattgccct 1680
gattttttg acttttttga ggctggagct actctcttg acagagacaa gtcgattatg 1740
gctatttctt cttggaatga caatggacaa atgcagtttg tccaagatcc ttatgctctt 1800

taccgctcag attttttcc cggctcttga tggatgcttt caaaatctac ttgggacgaa 1860

ttatctccaa agtggccaaa ggcttactgg gacgactggc taagactcaa agagaatcac 1920

agaggtcgac aatttattcg cccagaagtt tgcagaacat ataattttgg tgagcatggt 1980

tctagtttgg ggcagttttt caagcagtat cttgagccaa ttaactaaa tgatgtccag 2040

gttgattgga agtcaatgga ccttagttac cttttggagg acaattacgt gaaacacttt 2100

ggtgacttgg ttaaaaaggc taagcccatc catggagctg atgctgtctt gaaagcattt 2160

aacatagatg gtgatgtcgc tattcagtac agagatcaac tagactttga aaatcgcga 2220

cggcaatttg gcatttttga agaatggaag gatggtgtac cacgtgcagc atataaagga 2280

atagtagttt tccgtacca aacgtccaga cgtgtattcc ttgttgcca tgattcgett 2340

caacaactcg gaattgaaga tacttaa 2367

<210> 32

<211> 2370

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 polynucleotide

<400> 32

atgaaaatcg aagaaggtaa actggtaatc tggattaacg gcgataaagg ctataacggt 60

ctcgtgaag tccgtaagaa attcgagaaa gataccggaa ttaaagtcac cgttgagcat 120

ccggataaac tggaaagaaa attcccacag gttcggcga cttggcgatgg ccctgacatt 180

atcttctggg cacacgaccg ctttgggtggc tacgctcaat ctggcctggt ggctgaaatc 240

accccggaca aagcgttcca ggacaagctg taccggttta cctgggatgc cgtacgttac 300

aacggcaagc tgattgctta cccgatcgt gttgaagcgt tatcgctgat ttataacaaa 360

gatctgctgc cgaaccgcc aaaaacctgg gaagatcc cggcgtgga taaagaactg 420

aaagcgaag gtaagagcgc gctgatgttc aacctgcaag aacctgactt cacctggccg 480

ctgattgctg ctgacggggg ttatgcgttc aagtatgaaa acggcaagta cgacattaaa 540

gacgtgggcg tggataacgc tggcgcgaaa gcggtctga ccttcctggt tgacctgatt 600

aaaaacaac acatgaatgc agacaccgat tactccatcg cagaagctgc ctttaataaa 660

ggcgaacag cgatgacat caacggccc tgggcatggt ccaacatcga caccagcaaa 720

gtgaattatg gtgtaacggt actgccgacc ttcaagggtc aaccatcaa accgttcgtt 780

ggcgtgctga gcgcaggtat taacgcccc agtccgaaca aagagctggc gaaagagttc 840

ctcgaaaact atctgctgac tgatgaaggt ctggaagcgg ttaataaaga caaacccctg 900
 ggtgccgtag cgctgaagtc ttacgaggaa gagtggcgga aagatccacg tattgccgcc 960
 accatggaaa acgcccagaa aggtgaaatc atgccgaaca tcccgcagat gtcgcgtttc 1020

tggtatgccg tgcgtactgc ggtgatcaac gccgccagcg gtcgtcagac tgtcgatgaa 1080
 gccctgaaag acgcgcagac tcgtatcacc aagcgtcagc gtaaaaatga agccctggca 1140
 cctcctctgc tggatgctga accggcacgt ggtgctggcg gtcgtggtgg tgatcatccg 1200
 tctgttgccg ttggtattcg tcgtgtgagc aatgtttcgg ctgcctctct ggtcccggct 1260
 gttctcaac ctgaagctga taacctgacc ctgctctatc gctctctggt gtatcaactg 1320
 aacttcgac aaactctgcg taactggat aaagcaggca catgggctcc tcgtgaactg 1380
 gtactggtag tccaggcca taatcgctcc gaatatctgc gctgctgct ggattctctg 1440

cgcaaagctc aaggcatcga taatgtctg gtcactctct ctcattgatt ctggagcacg 1500
 gagattaacc agctgattgc cggcgtgaat tttgtcctg tgctgcaggt gttttttccg 1560
 tttctatcc aactgtatcc gaacgaattt ccgggttctg atcctcgtga ttgtcctcgt 1620
 gatctgccta aaaatgccgc tctgaaactg ggctgtatta atgccgagta tctgattct 1680
 tttggccact atcgtgaggc gaaatctct cagaccaaac atcattggtg gtggaaactg 1740
 catttcgtgt gggaactgt gaaaatctg cgcgactatg ctggcctgat tctgtttctg 1800
 gaagaagatc actatctggc tccggacttt tatcatgtgt tcaaaaaat gtggaaactg 1860

aaacagcagg aatgtccaga atgtgatgtg ctgtcactgg gcacctatag tgcttctcgc 1920
 tccttctatg gtatggccga caaagtggac gttaaaacat ggaatccac cgagcacaac 1980
 atgggtctgg cactgactcg taatgcctat caaaaactga ttgagtgtac cgacaccttt 2040
 tgtacgtatg atgactataa ctgggactgg acctgcaat atctgacctg gagctgtctg 2100
 ccaaaatctt ggaaagtctt ggtgcctcag attctcgtga tctttcatgc tggcgactgt 2160
 ggtatgcacc ataaaaaac ttgccgtccg tcaacacaat ctgctcagat cgagtcgctg 2220
 ctgaataata acaaacagta tatgttcccg gagactctga caatttctga aaaattcacc 2280

gtggtcgcca tttctccgc tcgtaaaaat ggaggttggg gcgatatccg tgacctgaa 2340
 ctgtgtaaaa gctatcgtcg tctgcagtga 2370

<210> 33
 <211> 2445
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic

polynucleotide

<400> 33

atgaaaatcg aagaaggtaa actggtaatc tggattaacg gcgataaagg ctataacggt	60
ctcgtgaag tccgtaagaa attcgagaaa gataccggaa ttaaagtcac cgttgagcat	120
ccggataaac tggaagagaa attcccacag gttgcggcaa ctggcgatgg ccctgacatt	180
atcttctggg cacacgaccg ctttgggtggc tacgctcaat ctggcctgtt ggctgaaatc	240
accccggaca aagcgttcca ggacaagctg tatccgttta cctgggatgc cgtacgttac	300
aacggcaagc tgattgctta cccgatcgt gttgaagcgt tatcgctgat ttataacaaa	360
gatctgctgc cgaacccgcc aaaaacctgg gaagagatcc cggcgtgga taaagaactg	420
aaagcgaag gtaagagcgc gctgatgttc aacctgcaag aacctactt cacctggccg	480
ctgattgctg ctgacggggg ttatgcttc aagatgaaa acggcaagta cgacattaaa	540
gacgtgggcg tggataacgc tggcgcgaaa gcggtctga ccttctggt tgacctgatt	600
aaaaacaac acatgaatgc agacaccgat tactccatcg cagaagctgc ctttaataaa	660
ggcgaacag cgatgacat caacggcccg tggcatggt ccaacatcga caccagcaaa	720
gtgaattatg gtgtaacggt actgccgacc tccaagggtc aaccatcca accgttcgtt	780
ggcgtgctga gcgcaggat taacgcccc agtccgaaca aagagctggc gaaagagttc	840
ctcgaaaact atctgctgac tgatgaaggt ctggaagcgg ttaataaaga caaacctg	900
ggtgccgtag cgctgaagtc ttacgaggaa gattggcga aagatccacg tattgccgcc	960
accatggaac acgcccagaa aggtgaaatc atgccgaaca tcccgcagat gtccgcttc	1020
tggtatgccg tgcgtactgc ggtgatcaac gccccagcg gtcgtcagac tgtcgatgaa	1080
gcctgaaag acgcgagac tcgtatcacc aagatcttga aagaactgac gtccaaaaag	1140
agcttgcaag tcccgtccat ctactatcac ttgccgcact tgctgcaaaa cgagggtct	1200
ttgcaaccgg cagttcagat cggcaatggt cgcaccggcg tgagcattgt tatgggtatc	1260
ccgacctga aacgtgaagt gaaaagctat ctgattgaaa cgctgcatag cctgatcgat	1320
aacctgtacc cggaagaaaa actggactgc gtgattgtcg ttttcattgg tgaaccgac	1380
acggattatg tgaatggcgt tgttgccaat ctggaaaaag agttcagcaa agagatcagc	1440
agcggcctgg ttgagatcat ttctccgccg gagagctatt acccgatct gacgaacctg	1500
aaagaaacct tccgtgatag caaagagcgt gtccgttggc gcaactaagca gaacctggac	1560
tattgttttc tgatgatgta cgcgcaagaa aagggtacgt attacatcca actggaggac	1620
gacattatg tgaagcaaaa ctacttcaac accattaaga acttcgcgt gcagctgagc	1680
agcgaagagt ggatgattct ggagttcagc cagctgggct tcattggcaa gatgtttcag	1740

gcaccggact tgaccctgat cgtggagttt atctttatgt tctacaaaga gaaaccgatc 1800
gattggctgc tggatcatat cctgtgggtc aaggtctgca atccggaaaa agatgccaag 1860

cattgtgacc gccagaaagc gaatctgctg attcgttttc gtctagcct gttccaacac 1920
gtgggtctgc acagctctct gaccggtaag atccaaaagc tgaccgacaa agattacatg 1980
aaaccgctgc tgctgaagat ccatgtcaac ccgccagcag aggtgagcac ctgctgaaa 2040
gtctaccagg gtcacactct ggagaaaacc tatatgggcg aggacttctt ttgggcgatt 2100
acgcctgttg cgggtgacta tatcttgttt aagtttgaca agccggttaa tgtagagagc 2160
tacttgtttc atagcggtaa ccaggatcac ccaggtgaca ttctgctgaa caccaccgtt 2220
gaagtgttgc cgctgaaaag cgaaggtctg gatatttcca aagaaacgaa ggataagcgt 2280

ctggaggatg gttacttccg tatcggcaag ttcgagaatg gcgtggctga aggtatggtc 2340
gaccggagcc tgaaccgat ttccgcatth cgcctgtccg tcatccagaa tagcgcggtt 2400
tggcctatcc tgaatgagat tcacatcaaa aaggttacga attaa 2445

<210> 34
<211> 2247
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
polynucleotide
<400> 34

atgaaattgt tctacaaacc ggggtcctgc tctctcgtt cccatatcac cctgcgtgag 60
agcggaaagg attttacct cgtcagtggt gatttaatga aaaaacgtct cgaaaacggt 120

gacgattact ttgccgttaa ccctaagggg caggtgcctg cattgctgct ggatgacggt 180
actttgctga cggaaggcgt agcagattatg cagtatcttg ccgacagcgt ccccgaccgc 240
cagttgctgg caccggtaaa cagtatttcc cgctataaaa ccatcgaatg gctgaattac 300
atcgccaccg agctgcataa aggtttcaca cctctgtttc gccctgatac accggaagag 360
tacaaccga cagttcgcgc gcagctggag aagaagctgc aatatgtgaa cgaggcactg 420
aaggatgagc actggatctg cgggcaaaga tttacaattg ctgatgccta tctgtttacg 480
gttctgcgct gggcatacgc ggtgaaactg aatctggaag ggtagagca cattgcagca 540

tttatgcaac gtatggctga acgtccggaa gtacaagacg cgctgtcagc ggaaggctta 600
aagggcagtg cttggacaaa ctacaatttt gaagaggta agtctcattt tgggttcaaa 660
aaatatgttg tatcatcttt agtactagtg tatggactaa ttaaggttct cacgtggatc 720

ttccgtcaat ggggtattc cagcttgaat ccgttctcca aaaaatcttc attactgaac 780
 agagcagttg cctcctgtgg tgagaagaat gtgaaagttt ttggtttttt tcatccgtat 840
 tgtaatgctg gtgggtgggg ggaaaaagtg ctctggaaag ctgtagatat cactttgaga 900
 aaagatgcta agaacgttat tgtcatttat tcaggggatt ttgtgaatgg agagaatggt 960

 actccggaga atattctaaa taatgtgaaa gcgaagtctg attacgactt ggattcggat 1020
 agaataatatt tcatcttatt gaagctaaga tacttgggtg attcttcaac atggaagcat 1080
 ttcacgttga ttggacaagc aattggatca atgattctcg catttgaatc cattattcag 1140
 tgtccacctg atatatggat tgatacaatg ggggtaccctt tcagctatcc tattattgct 1200
 aggtttttga ggagaattcc tatcgtcaca tatacgcatt atccgataat gtcaaaagac 1260
 atgttaaata agctgttcaa aatgccaagc aagggtatca aagtttacgg taaaatatta 1320
 tactggaaag tttttatggt aatttatcaa tccattgggt ctaaaattga tattgtaatc 1380

 acaaaactcaa catggacaaa taaccacata aagcaaattt ggcaatccaa tacgtgtaaa 1440
 attatataat ctccatgctc tactgagaaa ttagtagatt ggaagcaaaa gtttggact 1500
 gcaaagggtg agagattaaa tcaagcaatt gtgttggcac aatttcgtcc tgagaaactg 1560
 cataagttaa tcattgagtc ctttgcaact ttcttgaaaa atttaccgga ttctgtatcg 1620
 ccaattaaat tgataatggc ggggtccact agatccaagc aagatgaaaa ttatgttaaa 1680
 agtttacaag actggtcaga aaatgtatta aaaattccta aacatttgat atcattcgaa 1740
 aaaaatctgc ccttcgataa gattgaaata ttactaaca aatctacttt cgggtttaat 1800

 gccatgtgga atgagcactt tggaattgca gttgtagagt atatggcttc cggtttgatc 1860
 cccatagttc atgctcggc gggccattg ttagatatag ttactccatg ggatgccaac 1920
 gggaatatcg gaaaagctcc accacaatgg gagttacaaa agaaatattt tgcaaaactc 1980
 gaagatgatg gtgaaactac tggatttttc tttaaagagc cgagtgatcc tgattataac 2040
 acaaccaaag atcctctgag ataccctaat ttgtccgacc ttttcttaca aattacgaaa 2100
 ctggactatg actgcctaag ggtgatgggc gcaagaaacc agcagtattc attgtataaa 2160
 ttctctgatt tgaagtttga taaagattgg gaaaactttg tactgaatcc tatttgtaaa 2220

 ttattagaag aggaggaaaag gggctga 2247
 <210> 35
 <211> 6
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic

6xHis tag

<400> 35

His His His His His His

1 5

<210> 36

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 36

Asp Gln Asn Ala Thr

1 5