



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2023년06월08일

(11) 등록번호 10-2540615

(24) 등록일자 2023년06월01일

(51) 국제특허분류(Int. Cl.)

C07H 15/256 (2006.01) A23L 2/60 (2006.01)

A23L 27/30 (2016.01) C12P 19/18 (2006.01)

C12P 19/44 (2006.01)

(52) CPC특허분류

C07H 15/256 (2013.01)

A23L 2/60 (2013.01)

(21) 출원번호 10-2019-7013306

(22) 출원일자(국제) 2017년10월13일

심사청구일자 2020년10월13일

(85) 번역문제출일자 2019년05월09일

(65) 공개번호 10-2019-0064627

(43) 공개일자 2019년06월10일

(86) 국제출원번호 PCT/US2017/056457

(87) 국제공개번호 WO 2018/071744

국제공개일자 2018년04월19일

(30) 우선권주장

62/408,179 2016년10월14일 미국(US)

62/555,809 2017년09월08일 미국(US)

(56) 선행기술조사문헌

W02016073740 A1

(73) 특허권자

코나겐 인크.

미국 01730 매사추세츠주 베드포드 던젤로 드라이브 15

(72) 발명자

마오, 구오홍

미국 01803 매사추세츠주 버링턴 포브스 애비뉴 25

빅, 제이콥, 에드워드

미국 02140 매사추세츠주 캠브리지 #3 엘름 스트리트 154

(뒷면에 계속)

(74) 대리인

양영준, 김영

전체 청구항 수 : 총 9 항

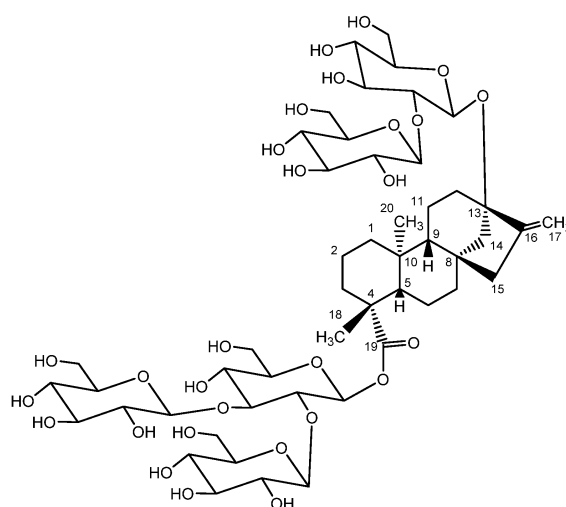
심사관 : 변진석

(54) 발명의 명칭 스테비올 글리코시드의 생합성 제조 및 이를 위한 방법

(57) 요약

본 발명은 스테비올 글리코시드 레바우디오시드 D4, WB1 및 WB2의 제조 및 Reb D4로부터의 레바우디오시드 M의 제조에 관한 것이다.

대표도 - 도1



(52) CPC특허분류

A23L 27/30 (2016.08)

C12P 19/18 (2013.01)

C12P 19/44 (2013.01)

C12Y 204/01 (2013.01)

C12Y 204/01013 (2013.01)

(72) 발명자

배튼, 마이클

미국 01886 매사추세츠주 웨스트포드 메인 스트리트 86

변, 데이비드

미국 01730 매사추세츠주 베드포드 와일드우드 드라이브 25

루오, 양

중국 214111 우시 지양수 신우 디스트릭트 춘후아 로드 8 꼬우신스지아 푸위엔 82-1602

우, 이린

중국 214000 우시 지양수 후이샹 디스트릭트 티앤이 스트리트 밍파쉬엔 위위엔 96-1002

장, 베이후아

중국 214000 우시 지양수 창 지양 노쓰 로드 빌딩 34 타이후 가든

위, 샤오단

미국 02421 매사추세츠주 렉싱턴 렉싱턴 리지 드라이브 6231

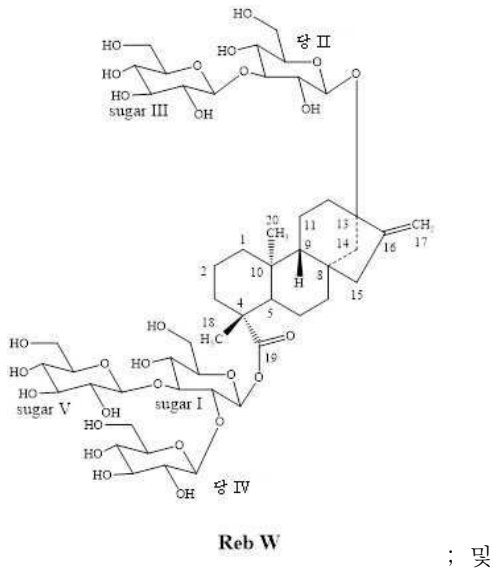
명세서

청구범위

청구항 1

다음 단계를 포함하는 스테비올 글리코시드 조성물의 제조 방법:

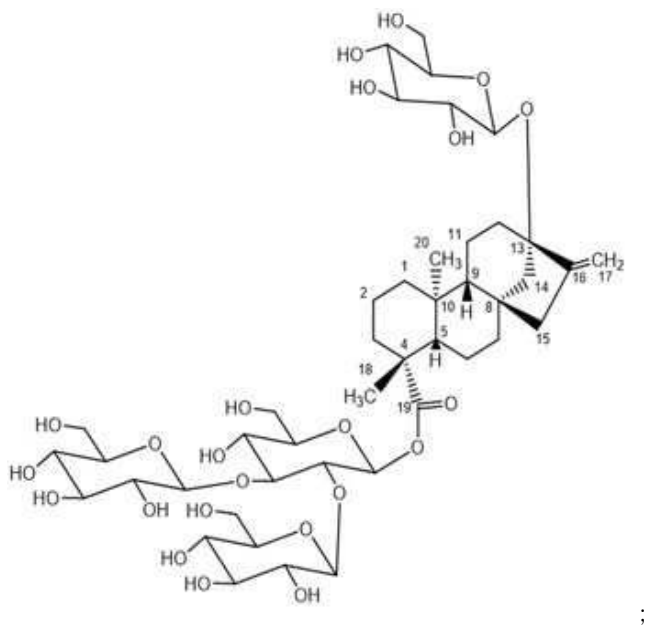
(a) 하기 구조를 갖는 레바우디오시드 W:



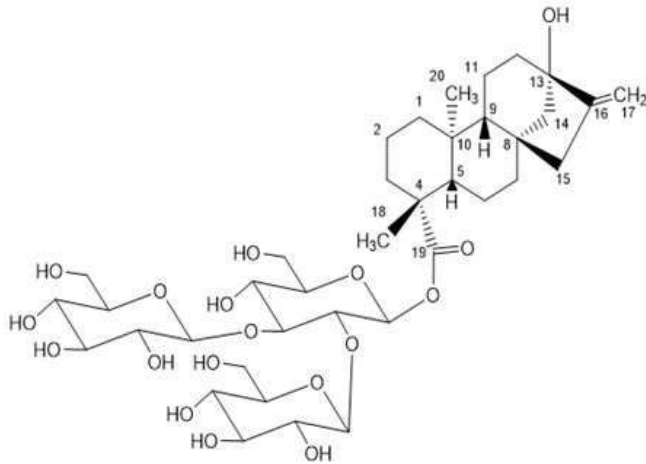
(b) 베타 글루코시다제

를 포함하는 반응 혼합물을 제조하는 단계; 및

레바우디오시드 WB1, 레바우디오시드 WB2, 또는 그의 혼합을 포함하는 스테비올 글리코시드 조성물을 제조하기에 충분한 시간 동안 상기 반응 혼합물을 인큐베이션하는 단계이며, 여기서 레바우디오시드 WB1은 하기 구조를 가지고:



레바우디오시드 WB2는 하기 구조를 갖는 것인 단계:



청구항 2

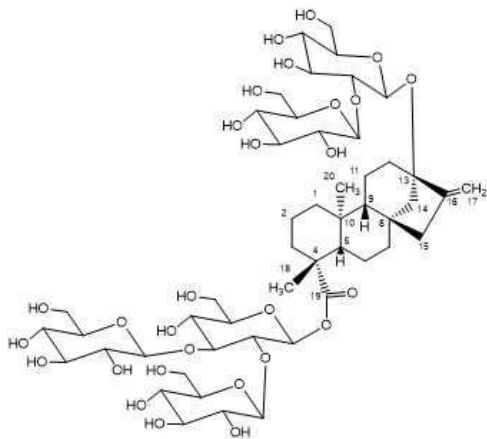
제1항에 있어서, 베타-글루코시다제가 SEQ ID NO:5의 아미노산 서열을 갖는 피치아 파스토리스(*Pichia pastoris*) 베타-글루코시다제인 방법.

청구항 3

제1항 또는 제2항에 있어서, 레바우디오시드 WB1을 포함하는 미정제 산물을 수득하는 단계를 추가로 포함하는 방법.

청구항 4

제1항에 있어서, 레바우디오시드 D4를 제조하기에 충분한 시간 동안 레바우디오시드 WB1을 UDP-글리코실트랜스퍼라제와 인큐베이션하는 단계를 추가로 포함하며, 여기서 UDP-글리코실트랜스퍼라제는 HV1 UGT 효소이고, 레바우디오시드 D4는 하기 구조를 갖는 것인 방법:



청구항 5

제4항에 있어서, 상기 UDP-글리코실트랜스퍼라제가 수크로스, 유리딘 디포스페이트(UDP) 및 유리딘 디포스페이트-글루코스(UDP-글루코스)의 존재 하에 수크로스 합성효소와 인큐베이션되는 것인 방법.

청구항 6

제4항에 있어서, 레바우디오시드 M을 제조하기에 충분한 시간 동안 레바우디오시드 D4를 UDP-글리코실트랜스퍼라제와 인큐베이션하는 단계를 추가로 포함하며, 여기서 UDP-글리코실트랜스퍼라제는 SEQ ID NO:1의 아미노산 서열을 갖는 UGT76G1인 방법.

청구항 7

제4항에 있어서, 레바우디오시드 M을 제조하기에 충분한 시간 동안 레바우디오시드 D4를 UDP 글리코실트랜스퍼라제와 인큐베이션하는 단계를 추가로 포함하며, 여기서 UDP 글리코실트랜스퍼라제는 SEQ ID NO:3의 아미노산 서열과 적어도 80% 동일성을 갖는 아미노산 서열을 포함하는 것인 방법.

청구항 8

제4항에 있어서, 레바우디오시드 M을 제조하기에 충분한 시간 동안 레바우디오시드 D4를 UDP 글리코실트랜스퍼라제와 인큐베이션하는 단계를 추가로 포함하며, 여기서 UDP 글리코실트랜스퍼라제는 SEQ ID NO:3의 아미노산 서열을 갖는 효소의 돌연변이체이며, 상기 돌연변이체는 SEQ ID NO:3의 위치 3, 6, 90, 91, 93, 181, 183, 184, 185, 350, 389, 410, 418, 450, 451, 452 및 454로부터 선택되는 위치에 상응하는 아미노산 잔기 위치에 서 적어도 하나의 돌연변이를 포함하는 것인 방법.

청구항 9

제6항 내지 제8항 중 어느 한 항에 있어서, 감미제로서 사용하기 위해 반응 혼합물로부터 제조되는 레바우디오시드 M을 수득하는 단계를 추가로 포함하는 방법.

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

발명의 설명

기술 분야

[0001] 관련 출원에 대한 교차-참조

[0002] 본 출원은 2016년 10월 14일에 출원된 미국 가출원 제62/408,179호 및 2017년 9월 8일에 출원된 미국 가출원 제 62/555,809호에 대한 우선권을 청구하며, 각각의 내용은 이의 전문이 본원에 참조로 포함된다.

[0003] 발명의 분야

[0004] 본 발명의 분야는 특정 스테비올 글리코시드의 제조에 유용한 방법 및 공정에 관한 것이다. 보다 구체적으로, 본 개시는 효소 전환을 통해 레바우디오시드 M("Reb M")으로 전환될 수 있는 이전에 알려지지 않은 레바우디오시드인 레바우디오시드 D4("Reb D4")의 제조를 제공한다. 본 개시는 또한 이전에 알려지지 않은 레바우디오시드인 레바우디오시드 WB1("Reb WB1") 및 레바우디오시드 WB2("Reb WB2")의 제조를 제공한다.

배경 기술

[0005] 본 개시는 신규한 스테비오시드 Reb D4, Reb WB1 및 Reb WB2의 제조 및 Reb D4의 Reb M으로의 전환에 초점을 맞춘다. 특히, 본 개시는 Reb D4의 합성 및 Reb M의 제조에서의 그 결과적 용도에 관한 것이다.

[0006] 스테비올 글리코시드는 스테비아 레바우디아나(*Stevia rebaudiana*) 잎에서 단리된 천연 산물이며, 식품, 사료 및 음료에서 고강도, 저-칼로리 감미제로서 널리 사용된다. 천연 발생 스테비올 글리코시드는 동일한 염기 디테르펜 구조(스테비올)를 갖지만, 스테비올 골격의 C13 및 C19 위치에서 탄수화물 잔기 변형(예컨대 글루코스, 람노스 및 자일로스 잔기)의 수 및 구조가 상이하다. 구조가 알려진 스테비올 글리코시드에는 스테비오시드, 레바우디오시드 A, 레바우디오시드 B, 레바우디오시드 C, 레바우디오시드 D, 레바우디오시드 E, 레바우디오시드 F, 레바우디오시드 M 및 둘코시드 A가 포함된다. 상업적 이용의 측면에서, 레바우디오시드 M은 자체가 일반적으로 안전한 것으로 간주되었다('GRAS' 상태).

[0007] 건조 중량 기준으로, 스테비오시드, 레바우디오시드 A, 레바우디오시드 C 및 둘코시드 A는 야생형 스테비아(*Stevia*) 잎에서 각각 스테비올 글리코시드의 총 중량의 9.1%, 3.8%, 0.6% 및 0.30%를 차지하는 반면, 다른 스테비올 글루코시드, 예컨대 Reb M은 상당히 더 소량으로 존재한다. 스테비아 레바우디아나(*Stevia rebaudiana*) 식물로부터의 추출물은 상업적으로 이용 가능하며, 이러한 추출물은 전형적으로 주 성분으로서 스테비오시드 및 레바우디오시드 A를 함유한다. 다른 알려진 스테비올 글리코시드는 전형적으로 소량 또는 미량 성분으로서 스테비아(*stevia*) 추출물에 존재한다. 예를 들어, 상업적 조제물 중 레바우디오시드 A의 양은 전체 스테비올 글리코시드 함량의 약 20% 내지 90% 초과로 다를 수 있지만, 레바우디오시드 B의 양은 전형적으로 전체 스테비올 글리코시드의 약 1% 내지 2%이며, 레바우디오시드 C의 양은 전체 스테비올 글리코시드의 약 7% 내지 15%일 수 있고, 레바우디오시드 D의 양은 전체 스테비올 글리코시드의 약 2%일 수 있다. 이러한 추출물에서, 레바우디오시드 M은 극히 소량으로만 존재한다. 흥미롭게도, 레바우디오시드 E도 전체 글리코시드의 0.5% 미만을 차지하는, 스테비아 레바우디아나(*Stevia rebaudiana*) 식물 품종에 존재하는 가장 덜 풍부한 스테비올 글리코시드 중 하나이다.

[0008] 천연 감미제로서, 상이한 스테비올 글리코시드는 평가되는 각각의 레바우디오시드 종에 연관된 상이한 단맛 정도, '구감' 및 특정한 뒷맛을 갖는다. 설탕(즉, 수크로스)에 비해, 스테비올 글리코시드의 단맛은 상당히 더 크다. 예를 들어, 스테비오시드는 수크로스보다 100배 내지 150배 더 달지만 맛 평가에서 주지되는 바와 같이 더 쓴 뒷맛을 갖는 반면, 레바우디오시드 A 및 E는 수크로스보다 250배 내지 450배 더 달고 뒷맛은 스테비오시드보다 훨씬 더 낮지만, 주지 가능한 뒷맛이 여전히 존재한다. 따라서, 임의의 스테비아(*stevia*) 추출물의 맛 프로파일은 추출물 중 스테비올 글리코시드의 상대 함량에 의해 현저히 영향받으며, 이는 다시 기저 식물이 겪는 환경 조건 및 사용되는 추출 방법에 의해 영향받을 수 있다. 이러한 식물 제조, 기후 조건 및 추출 조건의 변화는 스테비아(*stevia*) 추출물 중 스테비올 글리코시드의 일관성 없는 조성을 생성할 수 있어서, 이러한 맛 프로파일은 상이한 추출 산물 배치에 따라 크게 변한다.

- [0009] 스테비아(*stevia*) 추출물의 맛 프로파일은 또한 추출 공정 후 산물에 잔류하는 식물-유래 또는 환경-유래 오염물(예컨대 색소, 지질, 단백질, 페놀류 및 당류)에 의해 영향받을 수 있다. 이러한 오염물은 전형적으로 소비재에서 감미제로서의 스테비아(*stevia*) 추출물의 용도를 위해 바람직하지 못한 이들의 고유한 이취를 갖는다. 또한, 스테비아(*stevia*) 추출물 중 풍부하지 않은 개별 스테비올 레바우디오시드 또는 이의 특정 조합의 단리 비용은 고가이며 자원이 제한적이다. 일부 특정 스테비올 글리코시드의 제한된 양 및 이용 가능성에 기반하여, 천연 효소 또는 특정 미생물이 필요한 효소를 운반하고 관심 글리코시드의 제조를 특이적으로 증가시키기 위해 상업적으로 유의미한 발효 공정을 사용하도록 변형될 수 있는, 바이오-전환에 의해 상업적 공급이 더 잘 해결될 수 있다. 예를 들어, 변형된 미생물로부터 수득된 효소를 사용하는 스테비오시드의 Reb E로의 바이오-전환이 이전에 보고되었다(예컨대, PCT 출원 공개 제WO/2015/065650호 및 제WO/2015/171555호를 참고한다). 대안적으로, 다른 비-생물학적 합성 수단이 관심 스테비올 글리코시드를 개발하기 위해 사용될 수 있다.
- [0010] 생물학적 관점에서, 모든 스테비올 글리코시드는 스테비올의 일련의 글리코실화 반응에 의해 형성되며, 이는 전형적으로 당 모이어티의 공여체로서 유리된 5'-디포스포글루코스(UDP-글루코스)를 사용하는 UDP-글리코실트랜스퍼라제(UGT) 효소에 의해 촉매된다. 식물에서, UGT는 UDP-글루코스로부터 스테비올로 글루코스 잔기를 전달하는 널리 발산된 효소 그룹이다. 이러한 반응에서, 스테비오시드는 종종 다양한 레바우디오시드 화합물의 생합성에서의 중간체이다. 예를 들어, 스테비오시드의 C-13-O-글루코스에서 C-3'에서의 스테비오시드의 글리코실화는 레바우디오시드 A를 산출한다; 반면 스테비오시드의 19-O-글루코스 위치에서 C-2'에서의 글리코실화는 레바우디오시드 E를 산출한다.
- [0011] 본원에 기재된 바와 같이, 레바우디오시드 E의 특이적이고 지정된 글리코실화(C-19-O-글루코스에서의 글리코실화)는 레바우디오시드 Reb D4를 제조할 수 있고, UGT 효소에 의한 Reb D4의 추가 글리코실화는 레바우디오시드 M을 제조한다. 그러나, 본 개시 전에는 효소적인 D4의 제조를 위한 합성 단계가 보고된 바 없었다.
- [0012] 본 개시에 따르면, 스테비아(*stevia*) 추출물의 맛 품질을 개선하기 위한 실제적인 접근은 일반적으로 더 바람직한 맛 특징을 갖는 레바우디오시드 화합물의 수율을 증가시키는 것이며, 보다 생산적인 합성 경로를 통해 이를 수행하는 것이다. 평가된 스테비올 글리코시드 중에서, Reb M이 식품 및 음료에서의 사용을 위해 가장 바람직한 맛 및 화학적 특징을 갖는다고 믿는 이들이 많다. 그러나, 상기 언급된 바와 같이, 식물은 그 외에 극히 소량의 상기 화합물이 존재하며 이에 따라 상기 글리코시드의 대규모 제조를 위해서 뿐만 아니라 식품 및 음료 산업에 대안적인 감미제를 제공하기 위해 대안적인 생합성이 개발되어야 한다.
- [0013] 따라서, 더 우수하고 더 일관된 맛 프로파일을 갖는 스테비올 글리코시드가 상업적 제품으로서 개발될 필요가 있고, 이러한 스테비올 글리코시드에 있어서 비교적 일반적인 출발 기질, 예컨대 더 풍부한 스테비올 글리코시드를 출발 분자로서 이용하여, 이러한 바람직한 글리코시드의 제조가 상업적으로 최대한 비용 효과적일 수 있도록 하는 것이 필요하다. 본 개시는 이전에 알려지지 않은 스테비올 글리코시드, Reb D4로부터의 레바우디오시드 M의 제조 방법뿐만 아니라 Reb D4, Reb WB1 및 Reb WB2의 제조 방법을 제공한다.
- [0014] 더 나아가, 식물로부터의 추출 공정은 전형적으로 스테비올 글리코시드 회수를 위해 헥산, 클로로포름 및 에탄올과 같은 용매를 사용하는 고체-액체 추출 기법을 채택한다(Catchpole et al., 2003). 그러나, 용매 추출은 그 자체가 에너지 집약적이고, 독성 폐기물 처분 문제를 야기하며, 식물 스스로가 성장하기 위해 광범위한 면적을 필요로 하고, 회수될 소량 구성성분을 위한 추가 정제를 필요로 하는 산물을 산출한다. 따라서, 스테비올 글리코시드 제조 비용을 감소시키고 대규모 배양 및 가공의 환경적 영향을 경감시키기 위한 새로운 제조 방법이 또한 필요하다(Yao et al., 1994). 하나의 이러한 잠재적 해결책은 상업을 위해 이용 가능한 요망되는 스테비올 글리코시드의 선택성, 풍부도 및 순도를 증가시키는 소정 미생물 종에서의 제조를 허용하는 발효 바이오-전환 기술의 사용이다.
- [0015] 이에 더하여, 소비자는 식품, 사료, 풍미 또는 의학 성분을 위한 천연 및 생물 자원을 받아들이고 적극적으로 추구하는 반면, 조달, 일관된 맛 프로파일 및 환경적으로 지속 가능한 제조에 관해서도 우려한다. 이러한 상황에서, 본 개시의 미생물 발효 및 제조 방법은 다양한 산업 및 연구에 유용한 양으로 Reb M을 제공하면서 무기 합성 또는 현재의 식물 추출 기법에 비해 보다 천연 방식으로 이를 수행한다.
- [0016] 따라서, 인간 및 동물 소비를 추가로 구현하기 위해 경제적이고 편리하게 Reb M을 제조하는 신규한 방법을 개발할 필요가 있다.

발명의 내용

[0017] **발명의 요약**

[0018] 본 개시의 양태는 스테비올 글리코시드, 스테비올 글리코시드의 제조 방법 및 스테비올 글리코시드를 포함하는 조성물에 관한 것이다. 일부 양태에서, 본 개시는 이전에 보고되지 않은 스테비올 글리코시드 Reb D4로부터 Reb M을 제조하는 방법을 포괄한다.

[0019] 특히, 본 개시는 (13-[(2-O-β-D-글루코피라노실-β-D-글루코피라노실)옥시]엔트-카우르-16-엔-19-오산-[(2-O-β-D-글루코피라노실-3-O-β-D-글루코피라노실-β-D-글루코피라노실)에스테르])로서 확인되는 스테비올 글리코시드 레바우디오시드 D4 "Reb D4"의 제조 및 특정 UDP-글리코실트랜스퍼라제에 의한 Reb M으로의 그의 전환을 제공한다(도 1 참고). 본 개시는 또한 본원에 기재된 바와 같은 Reb WB1 및 Reb WB2의 제조를 제공한다.

[0020] 본원에 기재된 본 발명의 방법은 합성 경로를 사용하는 특정 스테비올 글리코시드의 합성을 위한 접근을 제공한다.

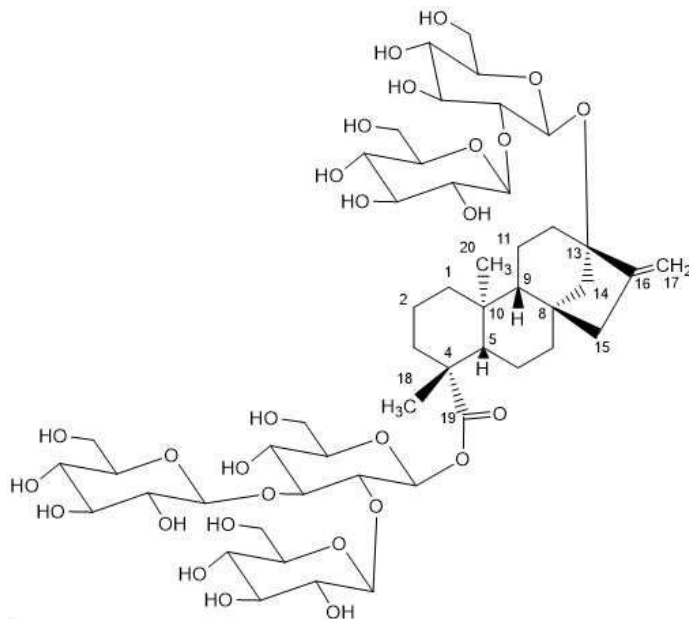
[0021] 대안적인 구현에는 RebWB1을 통한 경로를 이용하는 레바우디오시드 W로부터의 레바우디오시드 D4의 제조이다.

[0022] 추가 구현에는 레바우디오시드 D4로부터의 레바우디오시드 M의 제조이다.

[0023] 본 개시의 하나의 구현예에서, Reb WB2, Reb WB1 및 Reb D4를 통한 경로를 사용하는 Reb M의 제조를 가능하게 하는 방법이 제공된다.

[0024] 대안적인 구현예에서, Reb W의 Reb WB1로의 효소적 바이오전환을 촉매하기 위해 베타 글루코시다제가 사용되며, 도 14를 참고한다.

[0025] 일부 양태에서, 본 개시는 다음 구조를 갖는 스테비올 글리코시드 Reb D4를 제공한다:

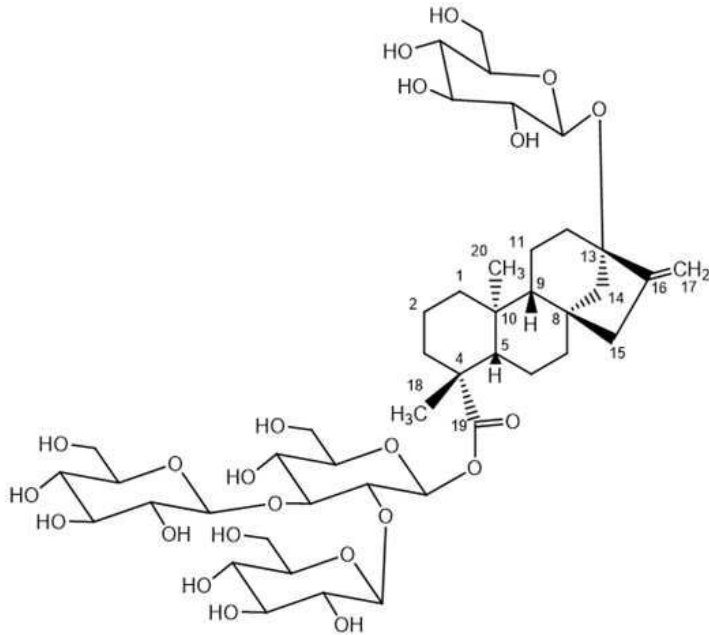


[0026]

[0027] 일부 구현예에서, 본 개시는 Reb D4를 포함하는 조성물을 제공하며, 선택적으로 조성물 중 상기 Reb D4 함량은 적어도 70%(예컨대, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100%) 순수하다. 일부 구현예에서, 본 개시는 감미량의 Reb D4를 포함하는 소비재를 제공한다. 일부 구현예에서, 소비재는 음료, 과자류, 베이커리 제품, 쿠키 및 츄잉 껌으로 구성되는 군으로부터 선택된다.

[0028] 다른 양태에서, 본 개시는 Reb D4 및 Reb M의 혼합물을 포함하는 조성물을 제공한다. 일부 구현예에서, 본 개시는 감미량의 Reb D4 및 Reb M의 혼합물을 포함하는 소비재를 제공한다. 일부 구현예에서, 소비재는 음료, 과자류, 베이커리 제품, 쿠키 및 츄잉 껌으로 구성되는 군으로부터 선택된다.

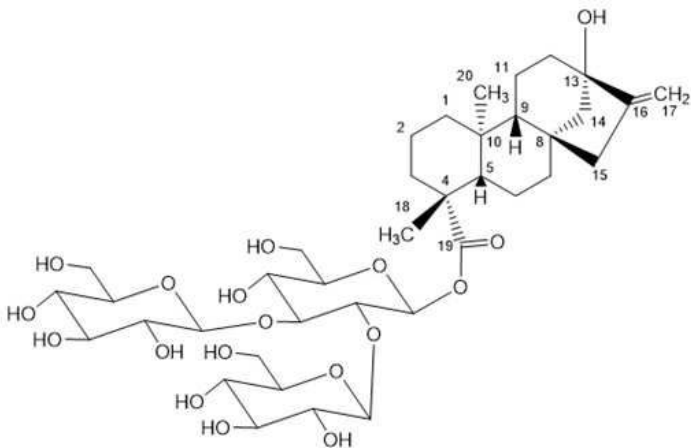
[0029] 다른 양태에서, 본 개시는 다음 구조를 갖는 스테비올 글리코시드 Reb WB1을 제공한다:



[0030]

[0031] 일부 구현예에서, 본 개시는 Reb WB1을 포함하는 조성물을 제공하며, 선택적으로 조성물 중 상기 Reb WB1 함량은 적어도 70%(예컨대, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100%) 순수하다. 일부 구현예에서, 본 개시는 감미량의 Reb WB1을 포함하는 소비재를 제공한다. 일부 구현예에서, 소비재는 음료, 과자류, 베이커리 제품, 쿠키 및 츄잉 껌으로 구성되는 군으로부터 선택된다.

[0032] 다른 양태에서, 본 개시는 다음 구조를 갖는 스테비올 글리코시드 Reb WB2를 제공한다:



[0033]

[0034] 일부 구현예에서, 본 개시는 Reb WB2를 포함하는 조성물을 제공하며, 선택적으로 조성물 중 상기 Reb WB2 함량은 적어도 70%(예컨대, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100%) 순수하다. 일부 구현예에서, 본 개시는 감미량의 Reb WB2를 포함하는 소비재를 제공한다. 일부 구현예에서, 소비재는 음료, 과자류, 베이커리 제품, 쿠키 및 츄잉 껌으로 구성되는 군으로부터 선택된다.

[0035] 일부 양태에서, 본 개시는 배지 내에서 성장하는 형질전환된 세포계에 의해 제조되는 관심 스테비올 글리코시드를 제공한다. 일부 구현예에서, 상기 형질전환된 세포계는 효모, 비-스테비올 글리코시드 제조 식물, 조류 및 박테리아로 구성되는 군으로부터 선택된다. 일부 구현예에서, 상기 세포계는 박테리아이며 에스케리치아 (*Escherichia*); 살모넬라(*Salmonella*); 바실러스(*Bacillus*); 아시네토박터(*Acinetobacter*); 스트렙토마이세스(*Streptomyces*); 코리네박테리움(*Corynebacterium*); 메틸로시누스(*Methylosinus*); 메틸로모나스

(*Methylomonas*); 로도코쿠스(*Rhodococcus*); 슈도모나스(*Pseudomonas*); 로도박터(*Rhodobacter*); 시네키시스티스(*Synechocystis*); 사카로마이세스(*Saccharomyces*); 자이코사카로마이세스(*Zygosaccharomyces*); 클루이베로마이세스(*Kluyveromyces*); 칸디다(*Candida*); 한센룰라(*Hansenula*); 데바리오마이세스(*Debaryomyces*); 무코르(*Mucor*); 피치아(*Pichia*); 토룰롭시스(*Torulopsis*); 아스페르길루스(*Aspergillus*); 아르트로보틀리스(*Arthrobotrys*); 브레비박테리아(*Brevibacterium*); 마이크로박테리움(*Microbacterium*); 아르트로박터(*Arthrobacter*); 시트로박터(*Citrobacter*); 클렙시엘라(*Klebsiella*); 판토에아(*Pantoea*); 살모넬라(*Salmonella*); 코리네박테리움(*Corynebacterium*); 클로스트리디움(*Clostridium*); 및 클로스트리디움 아세토부틸리쿰(*Clostridium acetobutylicum*)으로 구성되는 군으로부터 선택된다. 일부 구현예에서, 상기 세포계는 에스 케리치아 콜라이(*E. coli*)이다. 일부 구현예에서, 상기 스테비올 글리코시드는 Reb D4이다. 일부 구현예에서, 상기 스테비올 글리코시드는 Reb WB1이다. 일부 구현예에서, 상기 스테비올 글리코시드는 Reb WB2이다. 일부 구현예에서, 원료 물질은 스테비올이다. 일부 구현예에서, 상기 스테비올 글리코시드 함량은 적어도 70% 순도이다. 일부 구현예에서, 제조 방법은 i) 미정제 산물을 정제하는 단계; 및 ii) 진공 하에 용매를 제거하여 농축 산물을 제공하는 단계를 추가로 포함한다. 일부 구현예에서, 상기 미정제 산물은 칼럼 크로마토그래피에 의해 정제된다. 일부 구현예에서, 상기 미정제 산물은 산-염기 추출에 의해 정제된다. 일부 구현예에서, 상기 미정제 산물은 진공 증류에 의해 정제된다. 일부 구현예에서, 제조 방법은 반-분취용 HPLC를 사용하여 상기 스테비올 글리코시드를 정제하는 단계를 추가로 포함한다. 다른 양태에서, 본 개시는 감미량의 스테비올 글리코시드를 포함하는 소비재를 제공한다. 일부 구현예에서, 소비재는 음료, 과자류, 베이커리 제품, 쿠키 및 츄잉 껌으로 구성되는 군으로부터 선택된다.

[0036] 다른 양태에서, 본 개시는 SEQ ID NO:3과 적어도 80%(예컨대, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100%) 동일성을 갖는 DNA 서열을 포함하는 CP1 재조합 폴리펩타이드를 제공한다. 일부 구현예에서, CP1 재조합 폴리펩타이드의 아미노산 서열은 SEQ ID NO:4와 적어도 80%(예컨대, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99% 또는 100%) 동일성을 갖는다. 일부 구현예에서, CP1 재조합 폴리펩타이드는 표 2에 기재된 하나 이상의 위치에서 하나 이상의 돌연변이를 갖는다.

[0037] 다른 양태에서, 본 개시는 CP1 효소를 형질전환된 세포계에서 발현하는 단계; 세포계를 기질을 함유하는 배지 중에서 성장시키는 단계; 및 관심 스테비올 글리코시드를 제조하는 단계를 포함하는 관심 스테비올 글리코시드의 생합성 제조 방법을 제공한다. 일부 구현예에서, 방법은 재조합 수크로스 합성효소를 기질과 인큐베이션하는 단계를 추가로 포함한다. 일부 구현예에서, 방법은 재조합 UDP 글리코실트랜스퍼라제 UGT85C2를 수크로스 합성효소, 기질 및 CP1 재조합 폴리펩타이드와 인큐베이션하는 단계를 추가로 포함한다. 일부 구현예에서, 방법은 베타 글루코시다제 효소를 반응 혼합물에 첨가하는 단계를 추가로 포함한다. 일부 구현예에서, 수크로스 합성효소는 아라비도시스(*Arabidopsis*) 수크로스 합성효소 1, 아라비도시스(*Arabidopsis*) 수크로스 합성효소 3 및 비그나 라디에이트(*Vigna radiate*) 수크로스 합성효소로 구성되는 군으로부터 선택된다. 일부 구현예에서, 수크로스 합성효소는 아라비도시스 탈리아나(*Arabidopsis thaliana*) 수크로스 합성효소 1이다. 일부 구현예에서, 제조되는 스테비올 글리코시드는 Reb D4 및 Reb M의 혼합물이다. 일부 구현예에서, 방법은 i) 미정제 산물을 정제하는 단계; 및 ii) 진공 하에 용매를 제거하여 농축 산물을 제공하는 단계를 추가로 포함한다. 일부 구현예에서, 상기 미정제 산물은 칼럼 크로마토그래피에 의해 정제된다. 일부 구현예에서, 상기 미정제 산물은 산-염기 추출에 의해 정제된다. 일부 구현예에서, 상기 미정제 산물은 진공 증류에 의해 정제된다. 일부 구현예에서, 방법은 반-분취용 HPLC를 사용하여 상기 스테비올 글리코시드를 정제하는 단계를 추가로 포함한다. 일부 구현예에서, 상기 스테비올 글리코시드는 Reb WB1이다. 일부 구현예에서, 상기 스테비올 글리코시드는 Reb WB2이다. 일부 구현예에서, 상기 스테비올 글리코시드는 Reb D4이다. 일부 구현예에서, 상기 스테비올 글리코시드는 Reb M이다. 일부 구현예에서, 방법은 HV1(SEQ ID NO:9)의 사용을 추가로 포함한다. 일부 구현예에서, 방법은 UGT76G1(SEQ ID NO:1)의 사용을 추가로 포함한다.

[0038] 다른 양태에서, 본 개시는 재조합 세포가 스테비올 글리코시드를 제조하는 능력을 나타내는 적합한 성장 조건 하에 상기 재조합 세포를 배양하는 단계를 포함하는 레바우디오시드 M의 제조 방법을 제공하며, 방법은 상기 재조합 세포를 스테비올 글리코시드, 수크로스 합성효소 및 수크로스를 함유하는 반응 조성물과 접촉시키는 단계를 포함하고; 상기 재조합 세포는 상기 스테비올 글리코시드 기질을 사용하여 레바우디오시드 E를 제조할 수 있는 제1 UDP-글리코실트랜스퍼라제(UGT) 또는 이의 촉매 활성부를 발현하고; 상기 재조합 세포는 상기 레바우디오시드 E를 사용하여 레바우디오시드 D4를 제조할 수 있는 제2 UDP-글리코실트랜스퍼라제(UGT) 또는 이의 촉매 활성부를 발현하고; 상기 재조합 세포는 상기 레바우디오시드 D4를 사용하여 레바우디오시드 M을 제조할 수 있는 제3 UDP-글리코실트랜스퍼라제(UGT) 또는 이의 촉매 활성부를 발현한다. 일부 구현예에서, 방법은 상기 재조합 세포에서

발현되는 수크로스 합성효소 유전자 또는 이의 촉매 활성부를 추가로 포함한다. 일부 구현예에서, 방법은 반응 조성물에 첨가되는 수크로스 합성효소를 추가로 포함한다.

[0039] 일부 양태에서, 본 개시는 상기 단락에 기재되거나 본원에서 달리 개시되는 방법에 의해 제조되는 Reb M을 제공한다.

[0040] 다른 양태에서, 본 개시는 Reb M을 제조하기 위한 생합성 경로(예컨대, Reb D4의 Reb M으로의 전환을 통해 또는 Reb E에서 Reb D4에서 Reb M으로의 전환을 통해)를 발현하는 재조합 세포를 제공한다. 일부 구현예에서, 세포는 상기 스테비오시드 기질을 사용하여 레바우디오시드 E를 제조할 수 있는 제1 UDP-글리코실트랜스퍼라제(UGT) 또는 이의 촉매 활성부, 상기 레바우디오시드 E를 사용하여 레바우디오시드 D4를 제조할 수 있는 제2 UDP-글리코실트랜스퍼라제(UGT) 또는 이의 촉매 활성부 및 상기 레바우디오시드 D4를 사용하여 레바우디오시드 M을 제조할 수 있는 제3 UDP-글리코실트랜스퍼라제(UGT) 또는 이의 촉매 활성부 중 하나 이상을 발현한다. 일부 구현예에서, 세포는 효모 세포이다. 일부 구현예에서, 세포는 박테리아 세포이다. 일부 구현예에서, 세포는 식물 세포이다.

[0041] 다른 양태에서, 본 개시는 도 14에 기재된 효소 및 기질 또는 이의 하위세트(예컨대, Reb W, Reb WB1 또는 Reb D4로 시작하고/하거나 UGT76G1, CP1 또는 CR1을 이용함)를 사용하는 Reb M의 제조 방법을 제공한다. 일부 구현예에서, Reb M은 도 14에 기재된 효소 및 기질 또는 이의 하위세트(예컨대, Reb W, Reb WB1 또는 Reb D4로 시작하고/하거나 UGT76G1, CP1 또는 CR1을 이용함)를 함유하는 시험관내 반응 혼합물을 사용하여 제조된다. 일부 구현예에서, Reb M은 도 14에 기재된 효소 또는 이의 하위세트(예컨대, UGT76G1, CP1 또는 CR1)를 발현하는 세포에서 생체내 제조되며, 여기서 세포는 도 14에 기재된 기질(예컨대, Reb W, Reb WB1 또는 Reb D4)과 인큐베이션된다. 일부 구현예에서, 세포는 효모 세포이다. 일부 구현예에서, 세포는 박테리아 세포이다. 일부 구현예에서, 세포는 식물 세포이다.

[0042] 제품/상업적 유용성의 관점에서, 미국 내 시장에는 스테비올 글리코시드를 함유하는 수십 여 제품이 존재하며, 식품뿐만 아니라 진통제에서부터 해충 퇴치제에 이르기까지 모든 것에서, 그리고 식이 보충물질로서 사용될 수 있다. 스테비올 글리코시드를 함유하는 제품은 에어로졸, 액체 또는 과립 제형물일 수 있다.

[0043] 구현예에서의 세포계에 있어서, 일부 구현예에서, 이는 선택된 유전자로의 유전적 형질전환 및 이후 스테비올로부터 요망되는 스테비올 글리코시드의 생합성 제조를 허용할 박테리아, 효모 및 이의 조합 또는 임의의 세포계로 구성되는 군으로부터 선택된다. 가장 바람직한 미생물 시스템에서, 요망되는 스테비올 글리코시드 화합물을 제조하기 위해 에스케리치아 콜라이(*E. coli*)가 사용된다.

[0044] 본 개시에는 다양한 변형 및 대안적 형태가 적용되기 쉽지만, 이의 구체적 구현예를 도면에서 예로서 나타내며, 본원에서 상세히 설명될 것이다. 그러나, 본원에서 제시되는 도면 및 상세한 설명은 본 개시를 개시된 특정 구현예로 제한하려는 것이 아니라, 반대로, 첨부되는 청구범위에 의해 정의되는 본 개시의 정신 및 범위 내에 속하는 모든 변형, 균등부 및 대안을 커버하기 위한 것임이 이해되어야 한다.

[0045] 본 발명의 다른 속성 및 장점은 첨부 도면을 참조로 하여, 본 발명의 바람직한 구현예의 하기 상세한 설명에서 보다 명확해질 것이다.

도면의 간단한 설명

[0046] **도 1.** 레바우디오시드 D4(13-[(2-O-β-D-글루코피라노실-β-D-글루코피라노실)옥시]헵트-카우르-16-엔-19-오산-[(2-O-β-D-글루코피라노실-3-O-β-D-글루코피라노실-β-D-글루코피라노실)에스테르])의 구조를 나타낸다.

도 2. 레바우디오시드 WB1 및 WB2의 구조를 나타낸다.

도 3. 레바우디오시드 W 가수분해 산물의 HPLC 프로필을 나타낸다. 레바우디오시드 W는 B-glu1 효소에 의해 가수분해된다. A: 레바우디오시드 W의 표준물질("W"); B 내지 D: 레바우디오시드 W는 1시간(B), 6시간(C) 및 24시간(D)째에 재조합 B-glu1 효소에 의해 가수분해되었다.

도 4. UGT85C2에 의한 레바우디오시드 WB2("WB2")의 레바우디오시드 WB1("WB1")로의 바이오전환의 HPLC 프로필을 나타낸다. 레바우디오시드 WB2는 0시간(A), 2시간(B), 6시간(C) 및 18시간(D)째에 UGT85C2 효소와 인큐베이션되었다.

도 5. 레바우디오시드 WB1 및 WB2의 LC-MS 분석을 나타낸다.

도 6. HV1 효소에 의한 레바우디오시드 WB1의 레바우디오시드 D4로의 바이오전환의 HPLC 프로필을 나타낸다. A:

레바우디오시드 WB1의 표준물질("WB1"); B 내지 C: 레바우디오시드 WB1은 2시간(B) 및 6시간(C)째에 HV1효소에 의해 전환되었다.

도 7a 및 7b. Reb D4 분자 주위 구조 및 LC-MS 데이터를 나타낸다.

도 8. UDP 효소 UGT71G1의 구조를 나타낸다. 왼쪽에 히스티딘이 배치되고 오른쪽에 UDP가 배치된 표준 배향을 나타낸다.

도 9. UGT76G1 구조에서 알파 나선 및 베타 시트를 강조하는, UGT76G1 효소의 구조를 나타낸다.

도 10. CP1 및 UGT76G1 효소의 비교를 나타낸다. UGT76G1 결정 구조는 회색으로 표시되는 반면, CP1 모델은 검은색으로 표시된다.

도 11. UGT76G1 결정 구조 및 CP1 분자와의 상호작용을 나타낸다. 상기 결정 구조는 CP1 모델에서 베타 시트의 부재를 강조한다. UGT76G1 결정 구조는 회색으로 표시되는 반면, CP1 모델은 검은색으로 표시된다.

도 12. 효소 CP1의 반응 중심에서의 레바우디오시드 D4를 나타낸다. 이미지의 하부에 배치된 진회색 분자가 레바우디오시드 D4이다.

도 13a 및 도 13b. 재조합 UGT76G1 폴리펩타이드, 재조합 CP1 및 돌연변이체(CR1)의 조합에 의해 촉매되는 Reb D4로부터 Reb M의 시험관내 제조를 나타낸다. A: 레바우디오시드 D의 표준물질("D") 및 레바우디오시드 M의 표준물질("M")을 나타낸다. B: 레바우디오시드 D의 표준물질("D4")을 나타낸다. 30분(C) 및 1 hr(F)째에 UGT76G1에 의해 효소적으로 제조된 Reb M, 30분(D) 및 1 hr(G)째에 CP1에 의해 효소적으로 제조된 Reb M, 30분(E) 및 1 hr(H)째에 CR1에 의해 효소적으로 제조된 Reb M.

도 14. 레바우디오시드 W로부터의 레바우디오시드 M 생합성 경로에 대한 합성 경로를 나타낸다.

도 15 Reb WB2의 핵심 GHMBC 상관관계를 나타낸다.

도 16 Reb WB1의 핵심 GHMBC 상관관계를 나타낸다.

도 17 Reb D4의 핵심 TOCSY 및 GHMBC 상관관계를 나타낸다.

발명을 실시하기 위한 구체적인 내용

본원에서 사용되는 용어의 설명:

스테비올 글리코시드는 남아메리카 식물 스테비아 레바우디아나(*Stevia rebaudiana*(아스테라세아(*Asteraceae*))의 잎의 단 맛에 관여하는 화학적 화합물의 한 클래스이며, 식품, 사료 및 음료에서 감미제로서 사용될 수 있다.

정의:

"세포계"는 이소성 단백질의 발현을 제공하는 임의의 세포이다. 박테리아, 효모, 식물 세포 및 동물 세포가 포함된다. 원핵 및 진핵 세포가 모두 포함된다. 또한 세포 성분, 예컨대 리보솜에 기반하는 단백질의 시험관내 발현이 포함된다.

"코딩 서열"은 당업자에게 일반적이고 통상적인 의미로 주어지며 비제한적으로 특정 아미노산 서열을 인코딩하는 DNA 서열을 나타내기 위해 사용된다.

"세포계의 성장". 성장에는 세포가 증식하고 분열할 수 있도록 하는 적절한 배지의 제공이 포함된다. 또한 세포 또는 세포 성분이 재조합 단백질을 번역하고 제조할 수 있도록 하는 자원의 제공이 포함된다.

"단백질 발현". 단백질 제조는 유전자 발현 후 일어날 수 있다. 이는 DNA가 메신저 RNA(mRNA)로 전사된 후의 단계로 구성된다. 이어서 mRNA는 폴리펩타이드쇄로 번역되며, 궁극적으로 단백질로 폴딩된다. DNA는 전달감염(핵산을 세포 내로 의도적으로 도입하는 공정)을 통해 세포에 존재한다. 이 용어는 종종 진핵 세포에서 비-바이러스 방법에 대해 사용된다. 이는 또한 다른 방법 및 세포 유형을 나타낼 수 있지만, 다른 용어가 선호된다: "형질전환"은 박테리아, 식물 세포를 포함하는 비-동물 진핵 세포에서의 비-바이러스 DNA 전달을 설명하기 위해 더 자주 사용된다. 형질전환은 이들 세포에서 암성 상태로의 진행(발암)을 나타내기 위해서도 사용되므로, 동물 세포에서는 전달감염이 선호되는 용어이다. 형질도입은 바이러스-매개 DNA 전달을 설명하기 위해 종종 사용된다. 형질전환, 형질도입 및 바이러스 감염이 본 출원에 있어서 전달감염의 정의 하에 포함된다.

"효모". 본 개시에 따르면, 본원에서 청구되는 효모는 진균계 구성원으로서 분류되는 진핵, 단세포 미생물이다.

효모는 다세포 선조로부터 진화한 단세포 유기체이지만, 본 개시에 유용한 일부 종은 가성 균사 또는 위 균사로 알려진 연결된 발아 세포의 줄을 형성하여 다세포 특징을 전개하는 능력을 갖는 것들이다.

[0055] "UGT 효소". 본 개시에서 사용되는 UGT 효소의 명칭은 패밀리 번호, 서브패밀리를 표시하는 문자 및 개별 유전자 번호의 조합에 의해 UGT 유전자를 분류하는 UGT 명명 위원회에서 채택되는 명명 체계와 일치한다(Mackenzie et al., "The UDP glycosyltransferase gene super family: recommended nomenclature updated based on evolutionary divergence," PHARMACOGENETICS, 1997, vol. 7, pp. 255-269). 예를 들어, 명칭 "UGT76G1"은 UGT 패밀리 번호 76(식물 기원임), 서브패밀리 G 및 유전자 번호 1에 속하는 유전자에 의해 인코딩되는 UGT 효소를 나타낸다.

[0056] 구조 용어:

[0057] 본원에서 사용되는 단수 형태 부정관사 및 정관사("a, an" 및 "the")에는 내용 상 명확히 달리 나타내지 않는 한, 복수의 참조물이 포함된다.

[0058] 용어 "포함된다", "갖는다" 등이 설명 또는 청구범위에서 사용되는 정도까지, 이러한 용어는 "포함한다"라는 용어가 청구범위에서의 연결부로서 사용되는 것과 유사하게 포괄적인 것으로 의도된다.

[0059] 단어 "예시적인"은 본원에서 "실시예, 예 또는 예시로서 작용하는" 것을 의미하기 위해 사용된다. "예시적인" 것으로서 본원에 기재된 임의의 구현예가 반드시 다른 구현예에 비해 바람직하거나 유리한 것으로 간주되어야 하는 것은 아니다.

[0060] 용어 "상보적인"은 당업자에게 일반적이고 통상적인 의미로 주어지며, 비제한적으로 서로 혼성화할 수 있는 뉴클레오타이드 염기 간 관계를 설명하기 위해 사용된다. 예를 들어, DNA에 대해, 아데노신은 티민과 상보적이며 시토신은 구아닌과 상보적이다. 따라서, 종속 기술에는 실질적으로 유사한 핵산 서열뿐만 아니라 첨부된 서열 목록에서 보고되는 전체 서열과 상보적인 단리된 핵산 단편 또한 포함된다.

[0061] 용어 "핵산" 및 "뉴클레오타이드"는 당업자에게 이의 각각의 일반적이고 통상적인 의미로 주어지며, 비제한적으로 단일체 또는 이중체 형태의 데옥시리보뉴클레오타이드 또는 리보뉴클레오타이드 및 이의 중합체를 나타내기 위해 사용된다. 구체적으로 제한되지 않는 한, 이 용어는 참조 핵산과 유사한 결합 특성을 가지며 천연 발생 뉴클레오타이드와 유사한 방식으로 대사되는 천연 뉴클레오타이드의 알려진 유사체를 함유하는 핵산을 포괄한다. 달리 나타내지 않는 한, 특정 핵산 서열은 명시적으로 나타낸 서열뿐만 아니라, 묵시적으로 보존적으로 변형된 서열 또는 이의 축퇴성 변이체(예컨대, 축퇴성 코돈 치환) 및 상보적인 서열 또한 포괄한다.

[0062] 용어 "단리된"은 당업자에게 일반적이고 통상적인 의미로 주어지며, 단리된 핵산 또는 단리된 폴리펩타이드의 맥락에서 사용되는 경우, 비제한적으로 인간의 손에 의해, 그 원상태 환경으로부터 벗어나 존재하고 이에 따라 천연 산물이 아닌 핵산 또는 폴리펩타이드를 나타내기 위해 사용된다. 단리된 핵산 또는 폴리펩타이드는 정제된 형태로 존재할 수 있거나, 비-원상태 환경에서, 예컨대 트랜스제닉 숙주 세포에서 존재할 수 있다.

[0063] 본원에서 사용되는 용어 "인큐베이션하는" 및 "인큐베이션"은 2개 이상의 화학적 또는 생물학적 실체(예컨대 화학적 화합물 및 효소)를 혼합하고, 이들이 스테비올 글리코시드 조성물을 제조하기 바람직한 조건 하에 상호작용하도록 두는 공정을 의미한다.

[0064] 용어 "축퇴성 변이체"는 하나 이상의 축퇴성 코돈 치환에 의해 참조 핵산 서열과 상이한 잔기 서열을 갖는 핵산 서열을 나타낸다. 축퇴성 코돈 치환은 하나 이상의 선택된(또는 모든) 코돈의 제3 위치가 혼합된 염기 및/또는 데옥시이노신 잔기로 치환되는 서열을 생성하여 달성될 수 있다. 핵산 서열 및 이의 모든 축퇴성 변이체는 동일한 아미노산 또는 폴리펩타이드를 발현할 것이다.

[0065] 용어 "폴리펩타이드", "단백질" 및 "펩타이드"는 당업자에게 이의 각각의 일반적이고 통상적인 의미로 주어진다; 3개 용어는 때때로 상호 교환적으로 사용되며, 그 크기 또는 기능과 무관하게 비제한적으로 아미노산 또는 아미노산 유사체의 중합체를 나타내기 위해 사용된다. "단백질"은 종종 상대적으로 큰 폴리펩타이드를 나타내는 데 사용되며 "펩타이드"는 종종 작은 폴리펩타이드를 나타내는 데 사용되고, 당분야에서 이들 용어의 사용은 중첩되며 변한다. 본원에서 사용되는 용어 "폴리펩타이드"는 달리 언급되지 않는 한, 펩타이드, 폴리펩타이드 및 단백질을 나타낸다. 용어 "단백질", "폴리펩타이드" 및 "펩타이드"는 폴리뉴클레오타이드 산물을 나타내는 경우 본원에서 상호 교환적으로 사용된다. 따라서, 예시적인 폴리펩타이드에는 폴리뉴클레오타이드 산물, 천연 발생 단백질, 동족체, 오르소로그, 파라로그, 단편 및 이의 다른 등가물, 변이체 및 유사체가 포함된다.

[0066] 참조 폴리펩타이드에 대해 사용되는 경우, 용어 "폴리펩타이드 단편" 및 "단편"은 당업자에게 이의 일반적이고

통상적인 의미로 주어지며, 비제한적으로 아미노산 잔기가 참조 폴리펩타이드 자체에 비해 결실되지만, 잔여 아미노산 서열이 보통 참조 폴리펩타이드에서의 대응 위치와 동일한 폴리펩타이드를 나타내기 위해 사용된다. 이러한 결실은 참조 폴리펩타이드의 아미노-말단 또는 카복시-말단 또는 대안적으로 둘 다에서 일어날 수 있다.

[0067] 용어 폴리펩타이드 또는 단백질의 "기능적 단편"은 전장 폴리펩타이드 또는 단백질의 일부이며, 전장 폴리펩타이드 또는 단백질과 실질적으로 동일한 생물학적 활성을 갖거나 실질적으로 동일한 기능을 수행하는(예컨대, 동일한 효소 반응을 수행하는) 펩타이드 단편을 나타낸다.

[0068] 상호 교환적으로 사용되는 용어 "변이체 폴리펩타이드", "변형된 아미노산 서열" 또는 "변형된 폴리펩타이드"는 하나 이상의 아미노산에 의해, 예컨대, 하나 이상의 아미노산 치환, 결실 및/또는 부가에 의해 참조 폴리펩타이드와 상이한 아미노산 서열을 나타낸다. 하나의 양태에서, 변이체는 참조 폴리펩타이드의 일부 또는 모든 능력을 보유하는 "기능적 변이체"이다.

[0069] 용어 "기능적 변이체"에는 보존적으로 치환된 변이체가 추가로 포함된다. 용어 "보존적으로 치환된 변이체"는 하나 이상의 보존적 아미노산 치환에 의해 참조 펩타이드와 상이하며 참조 펩타이드의 일부 또는 모든 활성을 유지하는 아미노산 서열을 갖는 펩타이드를 나타낸다. "보존적 아미노산 치환"은 아미노산 잔기의 기능적으로 유사한 잔기로의 치환이다. 보존적 치환의 예에는 하나의 비극성(소수성) 잔기, 예컨대 이소류신, 발린, 류신 또는 메티오닌의 또 다른 비극성(소수성) 잔기로의 치환; 하나의 하전된 또는 극성(친수성) 잔기의 또 다른 극성(친수성) 잔기, 예컨대 아르기닌 및 라이신 간, 글루타민 및 아스파라긴 간, 트레오닌 및 세린 간 치환; 하나의 염기성 잔기, 예컨대 라이신 또는 아르기닌의 또 다른 염기성 잔기로의 치환; 또는 하나의 산성 잔기, 예컨대 아스파르트산 또는 글루탐산의 또 다른 산성 잔기로의 치환; 또는 하나의 방향족 잔기, 예컨대 페닐알라닌, 티로신 또는 트립토판의 또 다른 방향족 잔기로의 치환이 포함된다. 이러한 치환은 단백질 또는 폴리펩타이드의 겔보기 분자량 또는 등전점에 대한 효과를 거의 또는 전혀 갖지 않을 것으로 예상된다. 어구 "보존적으로 치환된 변이체"에는 생성 펩타이드가 본원에 기재된 바와 같은 참조 펩타이드의 일부 또는 모든 활성을 유지하는 한, 잔기가 화학적으로 유도체화된 잔기로 대체되는 펩타이드가 또한 포함된다.

[0070] 주제 기술의 폴리펩타이드에 관한 용어 "변이체"에는 참조 폴리펩타이드의 아미노산 서열과 적어도 75%, 적어도 76%, 적어도 77%, 적어도 78%, 적어도 79%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 심지어 100% 동일한 아미노산 서열을 갖는 기능적으로 활성이 있는 폴리펩타이드가 추가로 포함된다.

[0071] 모든 문법적 형태 및 철자 변형을 포함한 용어 "상동성"은 수퍼 패밀리로부터의 폴리뉴클레오타이드 또는 폴리펩타이드 및 상이한 종으로부터의 상동성 폴리뉴클레오타이드 또는 단백질을 포함하는, "공통 진화 기원"을 공유하는 폴리뉴클레오타이드 또는 폴리펩타이드 간 관계를 나타낸다(Reeck et al., CELL 50:667, 1987). 이러한 폴리뉴클레오타이드 또는 폴리펩타이드는 보존된 위치에서의 동일성 백분율 또는 특정 아미노산 또는 모티브의 존재의 관점과 무관하게, 이의 서열 유사성으로 반영되는 서열 상동성을 갖는다. 예를 들어, 2개의 상동성 폴리펩타이드는 적어도 75%, 적어도 76%, 적어도 77%, 적어도 78%, 적어도 79%, 적어도 80%, 적어도 81%, 적어도 82%, 적어도 83%, 적어도 84%, 적어도 85%, 적어도 86%, 적어도 87%, 적어도 88%, 적어도 89%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 심지어 100% 동일한 아미노산 서열을 가질 수 있다.

[0072] "적합한 조절 서열"은 당업자에게 일반적이고 통상적인 의미로 주어지며, 비제한적으로 코딩 서열의 상류(5' 비-코딩 서열), 코딩 서열 내 또는 하류(3' 비-코딩 서열)에 위치하고 연관된 코딩 서열의 전사, RNA 가공 또는 안정성 또는 번역에 영향을 미치는 뉴클레오타이드 서열을 나타내기 위해 사용된다. 조절 서열에는 프로모터, 번역 리더 서열, 인트론 및 폴리아데닐화 인식 서열이 포함될 수 있다.

[0073] "프로모터"는 당업자에게 일반적이고 통상적인 의미로 주어지며 비제한적으로 코딩 서열 또는 기능적 RNA의 발현을 제어할 수 있는 DNA 서열을 나타내기 위해 사용된다. 일반적으로, 코딩 서열은 프로모터 서열의 3'에 위치한다. 프로모터는 이의 전체가 원상태 유전자로부터 유래될 수 있거나, 자연에서 확인되는 상이한 프로모터로부터 유래되는 상이한 요소로 이루어질 수 있거나, 심지어 합성 DNA 절편을 포함할 수 있다. 당업자에게는 상이한 프로모터가 상이한 조직 또는 세포 유형에서 또는 상이한 발달 단계에서 또는 상이한 환경 조건에 반응하여 유전자의 발현을 지시할 수 있음이 이해된다. 유전자가 대부분의 시기에 대부분의 세포 유형에서 발현되도록 유도하는 프로모터는 일반적으로 "구성적 프로모터"로 언급된다. 대부분의 경우, 조절 서열의 정확한 경계가 완벽하게 정의되지 않았으므로, 상이한 길이의 DNA 단편이 동일한 프로모터 활성을 가질 수 있음이 추가로 인식된다.

- [0074] 용어 "작동 가능하게 연결된"은 하나의 기능이 다른 기능에 의해 영향받도록 하는 단일 핵산 단편 상에서의 핵산 서열의 연합을 나타낸다. 예를 들어, 프로모터는 이것이 해당 코딩 서열(즉, 프로모터의 전사 제어 하에 있는 코딩 서열)의 발현에 영향을 미칠 수 있는 경우 코딩 서열과 작동 가능하게 연결된 것이다. 코딩 서열은 센스 또는 안티센스 배향으로 조절 서열에 작동 가능하게 연결될 수 있다.
- [0075] 본원에서 사용되는 용어 "발현"은 당업자에게 일반적이고 통상적인 의미로 주어지며 비제한적으로 주체 기술의 핵산 단편으로부터 유래되는 센스(mRNA) 또는 안티센스 RNA의 전사 및 안정한 축적을 나타내기 위해 사용된다. "과발현"은 정상 또는 형질전환되지 않은 유기체에서의 제조 수준을 초과하는 트랜스제닉 또는 재조합 유기체에서의 유전자 산물의 제조를 나타낸다.
- [0076] "형질전환"은 당업자에게 일반적이고 통상적인 의미로 주어지며, 비제한적으로 표적 세포에 의한 추가 발현을 위한 해당 세포 내로의 폴리뉴클레오타이드의 전달을 나타내기 위해 사용된다. 전달된 폴리뉴클레오타이드는 표적 세포의 게놈 또는 염색체 DNA 내로 혼입되어, 유전적으로 안정한 유전을 야기할 수 있거나, 숙주 염색체와 무관하게 복제할 수 있다. 형질전환된 핵산 단편을 함유하는 숙주 유기체는 "트랜스제닉" 또는 "재조합" 또는 "형질전환된" 유기체로 언급된다.
- [0077] 숙주 세포에 관해 본원에서 사용되는 경우, 용어 "형질전환된", "트랜스제닉" 및 "재조합"은 당업자에게 이의 각각의 일반적이고 통상적인 의미로 주어지며, 비제한적으로 이중성 핵산 분자가 도입된 숙주 유기체의 세포, 예컨대 식물 또는 미생물 세포를 나타내기 위해 사용된다. 핵산 분자는 숙주 세포의 게놈 내로 안정적으로 통합될 수 있거나, 핵산 분자는 염색체의 분자로서 존재할 수 있다. 이러한 염색체의 분자는 자가-복제할 수 있다. 형질전환된 세포, 조직 또는 대상체는 형질전환 공정의 최종 산물뿐만 아니라, 이의 트랜스제닉 자손도 포괄하는 것으로 이해된다.
- [0078] 폴리뉴클레오타이드에 관해 본원에서 사용되는 경우, 용어 "재조합", "이중성" 및 "외인성"은 당업자에게 이의 일반적이고 통상적인 의미로 주어지며, 비제한적으로 특정 숙주 세포에 대한 외래 출처에서 기인하는, 또는 동일한 출처로부터인 경우, 그 원래 형태로부터 변형된 폴리뉴클레오타이드(예컨대, DNA 서열 또는 유전자)를 나타내기 위해 사용된다. 따라서, 숙주 세포 내의 이중성 유전자에는 특정 숙주 세포에 내인성인, 그러나 예를 들어, 부위-지정 돌연변이화 또는 다른 재조합 기법의 사용을 통해 변형된 유전자가 포함된다. 이 용어에는 또한 천연 발생 DNA 서열의 여러 비-천연 발생 카피가 포함된다. 따라서, 이 용어는 세포에 대해 외래이거나 이중성인, 또는 세포와 상동성이지만 요소가 보통 확인되지 않는 숙주 세포 내의 위치에 있거나 형태인 DNA 절편을 나타낸다.
- [0079] 유사하게, 폴리펩타이드 또는 아미노산 서열에 관해 본원에서 사용되는 경우, 용어 "재조합", "이중성" 및 "외인성"은 특정 숙주 세포에 대해 외래 출처에서 기인하는, 또는 동일한 출처로부터인 경우, 그 원래 형태로부터 변형된 폴리펩타이드 또는 아미노산 서열을 의미한다. 따라서, 재조합 DNA 절편은 숙주 세포에서 발현되어 재조합 폴리펩타이드를 제조할 수 있다.
- [0080] 용어 "플라스미드", "벡터" 및 "카세트"는 당업자에게 이의 각각의 일반적이고 통상적인 의미로 주어지며, 비제한적으로 종중 세포의 중심 대사의 일부가 아닌 유전자를 운반하고 보통 원형 이중쇄 DNA 분자 형태인 염색체외 요소를 나타내기 위해 사용된다. 이러한 요소는 임의의 출처로부터 유래되는, 단일쇄 또는 이중쇄 DNA 또는 RNA의 선형 또는 원형, 자가 복제 서열, 게놈 통합 서열, 파지 또는 뉴클레오타이드 서열일 수 있고, 여기서 여러 뉴클레오타이드 서열이 세포 내로 적절한 3' 미번역 서열을 따라 프로모터 단편 및 선택된 유전자 산물에 대한 DNA 서열을 도입할 수 있는 고유한 구성으로 연결되거나 재조합되었다. "형질전환 카세트"는 외래 유전자를 함유하며 특정 숙주 세포의 형질전환을 촉진하는 외래 유전자에 부가하는 요소를 갖는 특정 벡터를 나타낸다. "발현 카세트"는 외래 유전자를 함유하며 외래 숙주에서 해당 유전자의 발현 증가를 허용하는 외래 유전자에 부가하는 요소를 갖는 특정 벡터를 나타낸다.
- [0081] **상세한 설명**
- [0082] 본 개시는 관심 스테비올 글리코시드, Reb D4의 제조에 관한 것이며, 이어서 Reb D4 글리코시드가 Reb M으로 전환될 수 있도록 하는 UGT 효소의 사용에 관한 것이다. 본 개시는 또한 다른 관심 스테비올 글리코시드, Reb WB1 및 Reb WB2의 제조에 관한 것이다. 주체 기술은 UDP 글리코실트랜스퍼라제 활성, 예컨대 스테비올 글리코시드를 합성하기 위한 1,2-13-O-글루코스 글리코실화 활성 및 1,3-13-O-글루코스 글리코실화 활성을 갖는 재조합 폴리펩타이드를 제공한다. 주체 기술의 재조합 폴리펩타이드는 스테비올 글리코시드 화합물의 생합성에 유용하다. 본 개시에서, UDP-글리코실트랜스퍼라제(UGT)는 활성화된 공여체 분자(전형적으로 UDP-글루코스)로부터 수신체

분자로 당 잔기를 전달하는 효소를 나타낸다. 1,3-13-0-글루코스 글리코실화 활성은 레바우디오시드 D4의 13-0 글루코스 모이어티의 C-3'으로 당 모이어티를 전달하여 Reb M을 제조하는 효소 활성을 나타낸다(도 14). 주제 기술은 또한 스테비올 글리코시드의 합성을 위한 베타-글루코시다제 활성을 갖는 재조합 폴리펩타이드를 제공한다.

[0083] 합성 생물학

[0084] 본원에서 사용되는 표준 재조합 DNA 및 분자 클로닝 기법은 당분야에 널리 알려져 있으며, 예를 들어, 문헌 [Sambrook, J., Fritsch, E. F. and Maniatis, T. MOLECULAR CLONING: A LABORATORY MANUAL, 2nd ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., 1989 (이후 "Maniatis"); 및 Silhavy, T. J., Bannan, M. L. and Enquist, L. W. EXPERIMENTS WITH GENE FUSIONS; Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y., 1984; 및 Ausubel, F. M. et al., IN CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, published by GREENE PUBLISHING AND WILEY-INTERSCIENCE, 1987; (각각의 전문이 본원에 참조로 포함됨)]에 기재되어 있다.

[0085] 달리 정의되지 않는 한, 본원에서 사용되는 모든 기술 및 과학 용어는 본 개시가 속하는 분야의 당업자가 일반적으로 이해하는 바와 동일한 의미를 갖는다. 본원에 기재된 것과 유사하거나 동등한 임의의 방법 및 물질이 본 개시의 실시 또는 평가에서 사용될 수 있지만, 바람직한 물질 및 방법은 아래에 기재된다.

[0086] 본 개시는 하기 비제한적 실시예의 고려 시 보다 완전히 이해될 것이다. 이들 실시예는 주제 기술의 바람직한 구현예를 시사하지만, 단지 예시로서 주어짐이 이해되어야 한다. 상기 논의 및 이러한 실시예로부터, 당업자는 주제 기술의 본질적 특징을 확인할 수 있고, 이의 정신 및 범위에서 벗어나지 않고 다양한 용도 및 조건에 맞게 주제 기술을 다양하게 변화 및 변형시킬 수 있다.

[0087] 글리코실화는 종종 생활성 및 식물 천연 산물의 저장을 제어하는 도처에서 일어나는 반응으로 간주된다. 소분자의 글리코실화는 지금까지 연구된 대부분의 식물 종에서 트랜스퍼라제 수퍼패밀리에 의해 촉매된다. 이들 글리코실트랜스퍼라제(GT)는 60개를 초과하는 패밀리로 분류되었다. 이들 중에서, UDP 글리코실트랜스퍼라제(UGT)로도 알려진 패밀리의 1 GT 효소는 UDP-활성화된 당 모이어티를 특정 수신체 분자로 전달한다. 이들은 스테비올 글리코시드에서 이러한 당 모이어티를 전달하여 다양한 레바우디오시드를 생성하는 분자이다. 각각의 이러한 UGT는 이의 고유한 활성 프로필 및 이들이 이의 활성화된 당 모이어티를 전달하는 선호되는 구조 위치를 갖는다.

[0088] 제조 시스템

[0089] 원핵생물에서 단백질의 발현은 가장 흔하게는 융합 또는 비-융합 단백질의 발현을 지시하는 구성적 또는 유도성 프로모터를 함유하는 벡터로 박테리아 숙주 세포에서 수행된다. 융합 벡터는 보통 재조합 단백질의 아미노 말단에, 여기에서 인코딩되는 단백질에 여러 아미노산을 부가한다. 이러한 융합 벡터는 전형적으로 1) 재조합 단백질의 발현을 증가시키고; 2) 재조합 단백질의 용해도를 증가시키고; 3) 친화도 정제에서 리간드로서 작용하여 재조합 단백질의 정제를 보조하는 3가지 목적을 위해 작용한다. 종종, 단백질 분해 절단 부위가 융합 모이어티 및 재조합 단백질의 접합부에 도입되어 융합 단백질의 정제 이후 융합 모이어티로부터 재조합 단백질의 분리를 가능하게 한다. 이러한 벡터는 본 개시의 범위 내에 있다.

[0090] 하나의 구현예에서, 발현 벡터에는 박테리아 세포에서 재조합 폴리펩타이드의 발현을 위한 유전적 요소가 포함된다. 박테리아 세포에서 전사 및 번역을 위한 요소에는 프로모터, 단백질 복합체에 대한 코딩 영역 및 전사 종결인자가 포함될 수 있다.

[0091] 당업자는 발현 벡터의 제조를 위해 이용 가능한 분자 생물학 기법을 인지할 것이다. 상술된 바와 같이, 주제 기술의 발현 벡터 내로의 혼입을 위해 사용되는 폴리뉴클레오타이드는 일상적 기법, 예컨대 폴리머라제 연쇄 반응(PCR)에 의해 제조될 수 있다.

[0092] 상보적인 응집성 말단을 통해 DNA를 벡터에 작동 가능하게 연결하기 위해 여러 분자 생물학 기법이 개발되었다. 하나의 구현예에서, 상보적인 단독중합체 트랙이 벡터 DNA 내로 삽입될 핵산 분자에 부가될 수 있다. 이어서 벡터 및 핵산 분자는 상보적인 단독중합체 꼬리 사이의 수소 결합 형성에 의해 연결되어 재조합 DNA 분자를 형성한다.

[0093] 대안적인 구현예에서, 제공되는 하나 이상의 제한효소 부위를 함유하는 합성 링커가 주제 기술의 폴리뉴클레오타이드를 발현 벡터에 작동 가능하게 연결하기 위해 사용된다. 하나의 구현예에서, 폴리뉴클레오타이드는 제한효소 엔도뉴클레아제 소화에 의해 생성된다. 하나의 구현예에서, 핵산 분자는 이의 3'-5'-엑소뉴클레오타이드 분해 활성으로 돌출된 3'-단일쇄 말단을 제거하고 이의 중합 활성으로 오목해진 3'-말단을 채워서 무딘-말단 DNA

절편을 생성하는 효소인 박테리오파지 T4 DNA 폴리머라제 또는 에스케리치아 콜라이(*E. coli*) DNA 폴리머라제 I 로 처리된다. 이어서 무딘-말단 절편은 무딘-말단 DNA 분자의 결찰을 촉매할 수 있는 효소, 예컨대 박테리오파지 T4 DNA 리가제의 존재 하에 큰 몰 과량의 링커 분자와 인큐베이션된다. 이에 따라, 반응 산물은 그 말단에 중합체성 링커 서열을 운반하는 폴리뉴클레오타이드이다. 이어서 이들 폴리뉴클레오타이드는 적절한 제한 효소로 절단되고 폴리뉴클레오타이드의 말단과 호환성인 말단을 생성하는 효소로 절단된 발현 벡터에 결찰된다.

[0094] 대안적으로, 결찰-무관 클로닝(LIC) 부위를 갖는 벡터가 채택될 수 있다. 이어서 요구되는 PCR 증폭 폴리뉴클레오타이드가 제한효소 소화 또는 결찰 없이 LIC 벡터 내로 클로닝될 수 있다(본원과 일치하는 정도까지 본원에 참조로 포함되는, 문헌[Aslanidis and de Jong, NUCLEIC ACID RES. 18 6069-74, (1990), Haun, et al, BIOTECHNIQUES 13, 515-18 (1992)]).

[0095] 하나의 구현예에서, 선택된 플라스미드 내로의 삽입을 위해 관심 폴리뉴클레오타이드를 단리하고/하거나 변형하기 위해, PCR을 사용하는 것이 적합하다. 핵산 분자의 요구되는 코딩 영역을 단리하고, 제한효소 엔도뉴클레아제 또는 LIC 부위를 부가하고, 요망되는 해독틀에 코딩 영역을 배치하기 위해 서열의 PCR 제조에서 사용하기 위해 적절한 프라이머가 설계될 수 있다.

[0096] 하나의 구현예에서, 주제 기술의 발현 벡터 내로의 혼입을 위한 폴리뉴클레오타이드는 적절한 올리고뉴클레오타이드 프라이머를 사용하는 PCR을 이용하여 제조된다. 프라이머 자체가 증폭된 서열 산물 내로 혼입되는 동안 코딩 영역이 증폭된다. 하나의 구현예에서, 증폭 프라이머는 제한효소 엔도뉴클레아제 인식 부위를 함유하며, 이는 증폭된 서열 산물이 적절한 벡터 내로 클로닝될 수 있도록 한다.

[0097] 발현 벡터는 통상적인 형질전환 또는 전달감염 기법에 의해 식물 또는 미생물 숙주 세포 내로 도입될 수 있다. 주제 기술의 발현 벡터로의 적절한 세포의 형질전환은 당분야에 알려진 방법에 의해 달성되며 전형적으로 벡터 및 세포 유형 모두에 의존한다. 적합한 기법에는 칼슘 포스페이트 또는 칼슘 클로라이드 공-침전, DEAE-텍스트란 매개 전달감염, 리포펙션, 화학천공 또는 전기천공이 포함된다.

[0098] 성공적으로 형질전환된 세포, 즉, 발현 벡터를 함유하는 세포는 당분야에 널리 알려진 기법에 의해 확인될 수 있다. 예를 들어, 주제 기술의 발현 벡터로 전달감염된 세포를 배양하여 본원에 기재된 폴리펩타이드를 제조할 수 있다. 세포는 당분야에 널리 알려진 기법에 의해 발현 벡터 DNA의 존재에 대해 조사될 수 있다.

[0099] 숙주 세포는 이전에 기재된 발현 벡터의 단일 카피 또는 대안적으로 발현 벡터의 복수 카피를 함유할 수 있다.

[0100] 일부 구현예에서, 형질전환된 세포는 동물 세포, 곤충 세포, 식물 세포, 조류 세포, 진균 세포 또는 효모 세포이다. 일부 구현예에서, 세포는 유채 식물 세포, 평지씨 식물 세포, 야자 식물 세포, 해바라기 식물 세포, 목화 식물 세포, 옥수수 식물 세포, 땅콩 식물 세포, 아마 식물 세포, 참깨 식물 세포, 대두 식물 세포 및 페튜니아 식물 세포로 구성되는 군으로부터 선택되는 식물 세포이다.

[0101] 외래 단백질의 고수준 발현을 지시하는 조절 서열을 함유하는 미생물 숙주 세포 발현 시스템 및 발현 벡터는 당업자에게 널리 알려져 있다. 이들 중 임의의 것이 미생물 숙주 세포에서 주제 기술의 재조합 폴리펩타이드의 발현을 위한 벡터를 구축하기 위해 사용될 수 있다. 이어서 이들 벡터는 주제 기술의 재조합 폴리펩타이드의 고수준 발현을 허용하기 위해 형질전환을 통해 적절한 미생물 내로 도입될 수 있다.

[0102] 적합한 미생물 숙주 세포의 형질전환에 유용한 벡터 또는 카세트는 당분야에 널리 알려져 있다. 전형적으로 벡터 또는 카세트는 관련 폴리뉴클레오타이드의 전사 및 번역을 지시하는 서열, 선별 마커 및 자가 복제 또는 염색체 통합을 허용하는 서열을 함유한다. 적합한 벡터는 전사 개시 제어부를 보유하는 폴리뉴클레오타이드의 5' 영역 및 전사 종결을 제어하는 DNA 단편의 3' 영역을 포함한다. 이러한 제어 영역이 숙주로서 선택된 특정 종에 대해 원상태인 유전자로부터 유래될 필요가 없음이 이해되지만, 두 제어 영역이 모두 형질전환된 숙주 세포와 상동성인 유전자로부터 유래되는 것이 바람직하다.

[0103] 요망되는 미생물 숙주 세포에서 재조합 폴리펩타이드의 발현을 유도하기 유용한 개시 제어 영역 또는 프로모터는 여럿이며 당업자에게 친숙하다. 비제한적으로 CYCI, HIS3, GALI, GALIO, ADHI, PGK, PHO5, GAPDH, ADCI, TRPI, URA3, LEU2, ENO, TPI(사카로마이세스(*Saccharomyces*))에서의 발현에 유용함); AOXI(피치아(*Pichia*))에서의 발현에 유용함); 및 lac, trp, JPL, IPR, T7, tac 및 trc(에스케리치아 콜라이에서의 발현에 유용함)를 포함하는, 이들 유전자를 유도할 수 있는 실질적으로 모든 프로모터가 주제 기술을 위해 적합하다.

[0104] 종결 제어 영역은 또한 미생물 숙주에 대해 원상태인 다양한 유전자로부터 유래될 수 있다. 본원에 기재된 미생물 숙주를 위해 종결 부위가 선택적으로 포함될 수 있다.

- [0105] 식물 세포에서, 주제 기술의 발현 벡터에는 요망되는 발생 단계에 요망되는 조직에서 주제 기술의 재조합 폴리펩타이드의 발현을 지시할 수 있는 프로모터에 작동 가능하게 연결된 코딩 영역이 포함될 수 있다. 편의 상 이유로, 발현될 폴리뉴클레오타이드는 동일한 폴리뉴클레오타이드로부터 유래된 번역 리더 서열 및 프로모터 서열을 포함할 수 있다. 전사 종결 신호를 인코딩하는 3' 비-코딩 서열이 또한 존재할 것이다. 발현 벡터는 폴리뉴클레오타이드 발현을 촉진하기 위해 하나 이상의 인트론을 또한 포함할 수 있다.
- [0106] 식물 숙주 세포에 있어서, 코딩 영역의 발현을 유도할 수 있는 임의의 프로모터 및 임의의 종결인자의 임의의 조합이 주제 기술의 벡터 서열에서 사용될 수 있다. 프로모터 및 종결인자의 일부 적합한 예에는 노팔린 합성효소(nos), 옥토파인 합성효소(ocs) 및 칼리플라워 모자이크 바이러스(CaMV) 유전자로부터의 것들이 포함된다. 사용될 수 있는 효율적인 식물 프로모터의 하나의 유형은 고수준 식물 프로모터이다. 주제 기술의 발현 벡터와 작동 가능하게 연결된 이러한 프로모터는 벡터의 발현을 촉진할 수 있을 것이다. 주제 기술에서 사용될 수 있는 고수준 식물 프로모터에는 예를 들어 대두로부터의 리볼로스-1,5-바이포스페이트 카복실라제의 소형 서브유닛(들)의 프로모터(이것이 본원과 일치하는 정도까지 그 전문이 본원에 포함되는, 문헌[Berry-Lowe et al., J. MOLECULAR AND APP. GEN., 1:483 498 (1982)]) 및 엽록소 결합 단백질의 프로모터가 포함된다. 이들 두 프로모터는 식물 세포에서 광-유도되는 것으로 알려져 있다(예를 들어, 이들이 본원과 일치하는 정도까지 각각이 본원에 참조로 포함되는, 문헌[Genetic Engineering of Plants, an Agricultural PerspectiveENETIC ENGINEERING OF PLANTS, AN AGRICULTURAL PERSPECTIVE, A. Cashmore, Plenum, N.Y. (1983), pages 29 38; Coruzzi, G. et al., The Journal of Biological ChemistryCHEMISTRY, 258: 1399(1983) 및 Dunsmuir, P. et al., JOURNAL OF MOLECULAR AND APPLIED GENETICSournal of Molecular and Applied Genetics, 2:285(1983)] 참고).
- [0107] Reb D4에 대한 전구체 합성
- [0108] 전술된 바와 같이, 스테비올 글리코시드는 남아메리카 식물 스테비아 레바우디아나(*Stevia rebaudiana*(아스테라세아(*Asteraceae*)) 및 식물 루부스 칭기이(*Rubus chingii*(로사세아(*Rosaceae*))에서 단맛에 관여하는 화학적 화합물이다. 이들 화합물은 글리코실화된 디테르펜이다. 특히, 이의 분자는 하이드록실 수소 원자가 글루코스 분자에 의해 대체되어 에스테르를 형성하고 하이드록실 수소가 글루코스 및 람노스 조합으로 아세탈을 형성하는, 스테비올 분자로 간주될 수 있다.
- [0109] 본 개시의 관심 화합물의 하나의 제조 방법은 보편적인 또는 저렴한 전구체, 예컨대 화학적으로 유도되거나 조작된 미생물, 예컨대 박테리아 및/또는 효모에서 생합성을 통해 제조되는 스테비올 또는 루보수시드를 취하여, 알려져 있거나 저렴한 방법을 통해 표적화된 스테비올 글리코시드, 예컨대 Reb D4를 합성하는 것이다.
- [0110] 본 개시의 양태는 스테비올을 제조할 수 있는 미생물 시스템에서 재조합적으로 효소를 발현하는 것이 수반되는 방법에 관한 것이다. 일반적으로, 이러한 효소에는 코팔릴 디포스페이트 합성효소(CPS), 카우렌 합성효소(KS) 및 게라닐게라닐 디포스페이트 합성효소(GGPPS) 효소가 포함될 수 있다. 이는 내인성 이소프레노이드 합성 경로, 예컨대 비-메발로네이트(MEP) 경로 또는 메발론산 경로(MVA)를 발현하는 미생물 균주에서 일어난다. 일부 구현예에서 세포는 에스케리치아 콜라이를 포함하는 박테리아 세포 또는 효모 세포, 예컨대 사카로마이세스(*Saccharomyces*) 세포, 피치아(*Pichia*) 세포 또는 야로위아(*Yarrowia*) 세포이다. 일부 구현예에서, 세포는 조류 세포 또는 식물 세포이다.
- [0111] 이후, 전구체는 화학적 합성에서의 사용을 위해 발효 배양으로부터 회수된다. 전형적으로, 이는 스테비올이지만, 세포 배양으로부터의 카우렌 또는 스테비올 글리코시드일 수 있다. 일부 구현예에서, 스테비올, 카우렌 및/또는 스테비올 글리코시드는 기상으로부터 회수되는 반면, 다른 구현예에서는 유기층 또는 중합체 수지가 세포 배양에 첨가되고 카우렌, 스테비올 및/또는 스테비올 글리코시드가 유기층 또는 중합체 수지로부터 회수된다. 일부 구현예에서, 스테비올 글리코시드는 레바우디오시드 A, 레바우디오시드 B, 레바우디오시드 C, 레바우디오시드 D, 레바우디오시드 E, 레바우디오시드 F 또는 둘코시드 A로부터 선택된다. 일부 구현예에서, 제조되는 테르페노이드는 스테비오바이오시드 또는 스테비오시드이다. 일부 구현예에서, 적어도 하나의 효소적 단계, 예컨대 하나 이상의 글리코실화 단계가 생체의 수행됨이 또한 이해되어야 한다.
- [0112] 본 발명의 일부는 이후 Reb M으로의 추가적인 효소적 전환을 거치는 Reb D4 스테비올 글리코시드의 제조이다. 본 개시에 따르면, 미생물에 의해 제조된 스테비올을 요망되는 스테비올 글리코시드(여기서는 Reb D4)로 전환하기 위한 생합성은, 디테르페노이드 스테비올이 당 모이어티의 스테비올 골격 내로의 다단계 화학 어셈블리를 사용해서 스테비오시드 및 레바우디오시드 A로 전환되는 경우 일어난다. 보다 구체적으로 화학적 합성은 하기 단계로 구성된다: 1) 스테비올의 C19 COOH기에서 보호된 트리메틸실릴(TMS)이 출발 전구체 스테비올로부터 합성된다. 보호된 β -Glc- β -Glc(2 \rightarrow 1)- β -Glc(3 \rightarrow 1)기를 사용해서 스테비올의 C13-OH 위치에서 트리-글루코실화가 수

행된다. 여기에 TMS의 탈보호 및 보호된 모노 β -Glc-Br 모이어티의 커플링이 뒤따른다. 최종 탈보호로 모든 보호기를 제거하여 레바우디오시드 D4를 제조한다.

[0113] 스테비올 글리코시드의 생합성

[0114] 본원에 기재된 바와 같이, 본 발명의 기술의 제조합 폴리펩타이드는 UDP-글리코실트랜스퍼라제 활성을 가지며, 천연 출처에 존재하지 않거나 전형적으로 풍부도가 낮은 스테비올 글리코시드, 예컨대 레바우디오시드 D4 및 레바우디오시드 M 각각의 생합성 제조 방법의 개발에 유용하다. 본 발명의 기술의 제조합 폴리펩타이드는 베타-글루코시다제 또는 UDP-글리코실트랜스퍼라제 활성을 가지며, 신규한 스테비올 글리코시드, 예컨대 레바우디오시드 D4의 생합성 제조 방법의 개발에 유용하고, 레바우디오시드 M의 제조에 유용하다.

[0115] 기질은 하나 이상의 UDP 글리코실트랜스퍼라제에 의해 촉매되는 반응에서 스테비올 글리코시드 화합물로 전환될 수 있는 임의의 천연 또는 합성 화합물일 수 있다. 예를 들어, 기질은 천연 스테비아(*stevia*) 추출물, 스테비올, 스테비올-13-O-글루코시드, 스테비올-19-O-글루코시드, 스테비올-1,2-바이오시드, 루부소시드, 스테비오시드, 레바우디오시드 A, 레바우디오시드 G 또는 레바우디오시드 E일 수 있다. 기질은 순수한 화합물 또는 상이한 화합물의 혼합물일 수 있다. 바람직하게는, 기질에는 루부소시드, 스테비오시드, 스테비올, 레바우디오시드 A, 레바우디오시드 E 및 이의 조합으로 구성되는 군으로부터 선택되는 화합물이 포함된다.

[0116] 본원에 기재된 방법은 또한 본원에 기재된 제조합 펩타이드가 스테비올 글리코시드 화합물의 전반적 생합성의 효율을 개선하거나 결과를 변형하기 위해 하나 이상의 추가 효소와 조합되어 기능할 수 있도록 하는 커플링 반응계를 제공한다. 예를 들어, 추가 효소는 글리코실화 반응으로부터 생성된 UDP를 다시 UDP-글루코스로 전환시켜(예를 들어, 글루코스 잔기의 공여체로서 수크로스를 사용하여) 글리코실화 반응에 필요한 UDP-글루코스를 재생시킬 수 있고, 이에 따라 글리코실화 반응의 효율을 개선한다.

[0117] 또 다른 구현예에서, 주체 기술의 방법에는 제조합 UDP-글리코실트랜스퍼라제를 본원에 기재된 제조합 수크로스 합성효소, 기질 및 제조합 폴리펩타이드와 인큐베이션하는 단계가 추가로 포함된다. 제조합 UDP-글리코실트랜스퍼라제는 주체 기술의 제조합 폴리펩타이드에 의해 촉매되는 것과 상이한 글리코실화 반응을 촉매할 수 있다.

[0118] 적합한 UDP-글리코실트랜스퍼라제에는 스테비올 글리코시드 화합물, 예컨대 UGT85C2, UGT74G1, UGT76G1 또는 이의 기능적 동족체의 생합성에서 하나 이상의 반응을 촉매할 수 있는 당분야에 알려진 임의의 UGT가 포함된다.

[0119] 전형적으로, 주체 기술의 시험관내 방법에서, UDP-글루코스는 완충액에 약 0.2 mM 내지 약 5 mM, 바람직하게는 약 0.5 mM 내지 약 2 mM, 보다 바람직하게는 약 0.7 mM 내지 약 1.5 mM의 농도로 포함된다. 하나의 구현예에서, 제조합 수크로스 합성효소가 반응에 포함되는 경우, 수크로스는 또한 완충액에 약 100 mM 내지 약 500 mM, 바람직하게는 약 200 mM 내지 약 400 mM, 보다 바람직하게는 약 250 mM 내지 약 350 mM의 농도로 포함된다.

[0120] 전형적으로, 주체 기술의 시험관내 방법에서, 건조 중량 기준으로, 제조합 폴리펩타이드 대 기질의 중량 비는 약 1:100 내지 약 1:5, 바람직하게는 약 1:50 내지 약 1:10, 보다 바람직하게는 약 1:25 내지 약 1:15이다.

[0121] 전형적으로, 시험관내 방법의 반응 온도는 약 20°C 내지 약 40°C, 적합하게는 25°C 내지 약 37°C, 보다 적합하게는 28°C 내지 약 32°C이다.

[0122] 당업자는 본원에 기재된 방법에 의해 제조되는 스테비올 글리코시드 조성물이 요망되는 풍미제 또는 감미제 조성물을 획득하기 위해 추가 정제되고 다른 스테비올 글리코시드, 풍미제 또는 감미제와 혼합될 수 있음을 인식할 것이다. 예를 들어, 본원에 기재된 바와 같이 제조되는 레바우디오시드 D4가 농축된 조성물은 레바우디오시드 A를 우세한 스테비올 글리코시드로서 함유하는 천연 스테비아(*stevia*) 추출물과, 또는 요망되는 감미제 조성물을 제조하기 위해 다른 합성 또는 천연 스테비올 글리코시드 제품과 혼합될 수 있다. 대안적으로, 본원에 기재된 스테비올 글리코시드 조성물로부터 수득되는 실질적으로 정제된 스테비올 글리코시드(예컨대, 레바우디오시드 D4)는 다른 감미제, 예컨대 수크로스, 말토덱스트린, 아스파탐, 수크랄로스, 네오타ם, 아세설팜 칼륨 및 사카린과 조합될 수 있다. 다른 감미제 대비 스테비올 글리코시드의 양은 당분야에 알려진 바와 같이, 요망되는 맛을 수득하기 위해 조정될 수 있다. 본원에 기재된 스테비올 글리코시드 조성물(레바우디오시드 D, 레바우디오시드 E, 레바우디오시드 D4, 레바우디오시드 M 또는 이의 조합을 포함함)은 식품 제품(예컨대 음료, 소프트 드링크, 아이스크림, 유제품, 과자류, 시리얼, 츄잉 껌, 베이킹 제품 등),식이 보충물질, 의학적 영양분뿐만 아니라 약학 제품에 포함될 수 있다.

[0123] 당업자는 본원에 기재된 방법에 의해 제조되는 스테비올 글리코시드 조성물이 요망되는 풍미제 또는 감미제 조성물을 획득하기 위해 추가 정제되고 다른 스테비올 글리코시드, 풍미제 또는 감미제와 혼합될 수 있음을 인식

할 것이다. 예를 들어, 본원에 기재된 바와 같이 제조된 레바우디오시드 D4가 농축된 조성물은 레바우디오시드 A를 우세한 스테비올 글리코시드로서 함유하는 천연 스테비아(*stevia*) 추출물과, 또는 요망되는 감미제 조성물을 제조하기 위해 다른 합성 또는 천연 스테비올 글리코시드 제품과 혼합될 수 있다. 대안적으로, 본원에 기재된 스테비올 글리코시드 조성물로부터 수득되는 실질적으로 정제된 스테비올 글리코시드(예컨대, 레바우디오시드 D4)는 다른 감미제, 예컨대 수크로스, 말토덱스트린, 아스파탐, 수크랄로스, 네오타, 아세설팜 칼륨 및 사카린과 조합될 수 있다. 다른 감미제 대비 스테비올 글리코시드의 양은 당분야에 알려진 바와 같이, 요망되는 맛을 수득하기 위해 조정될 수 있다. 본원에 기재된 스테비올 글리코시드 조성물(레바우디오시드 D, 레바우디오시드 E, 레바우디오시드 D4, 레바우디오시드 WB1, 레바우디오시드 WB2, 레바우디오시드 M 또는 이의 조합을 포함함)은 식품 제품(예컨대 음료, 소프트 드링크, 아이스크림, 유제품, 과자류, 시리얼, 츄잉 껌, 베이킹 제품 등),식이 보충물질, 의학적 영양분뿐만 아니라 약학 제품에 포함될 수 있다.

[0124] 동일성 스코어링을 사용하는 서열 유사성 분석

[0125] 본원에서 사용되는 "서열 동일성"은 2개의 최적 정렬된 폴리뉴클레오타이드 또는 펩타이드 서열이 성분, 예컨대, 뉴클레오타이드 또는 아미노산의 정렬 윈도우에 걸쳐 변함없는 정도를 나타낸다. 평가 서열 및 참조 서열의 정렬된 절편에 대한 "동일성 비율"은, 2개의 정렬된 서열이 공유하는 동일한 성분의 수를, 참조 서열 절편에서 성분의 총 수, 즉 전체 참조 서열 또는 참조 서열의 더 작은 정의된 부분으로 나눈 것이다.

[0126] 본원에서 사용되는 용어 "서열 동일성 백분율" 또는 "동일성 백분율"은 2개 서열이 최적으로 정렬되는 경우(비교 윈도우에 걸쳐 참조 서열의 총 20% 미만인 적절한 뉴클레오타이드 삽입, 결실 또는 갭을 포함) 평가("대상체") 폴리뉴클레오타이드 분자(또는 그 상보체) 대비 참조("쿼리") 폴리뉴클레오타이드 분자(또는 그 상보체)의 선형 폴리뉴클레오타이드 서열에서 동일한 뉴클레오타이드의 백분율을 나타낸다. 비교 윈도우를 정렬하기 위한 서열의 최적 정렬은 당업자에게 잘 알려져 있고 Smith 및 Waterman의 국소 상동성 알고리즘, Needleman 및 Wunsch의 상동성 정렬 알고리즘, Pearson 및 Lipman의 유사성 검색 방법과 같은 도구에 의해, 바람직하게는 이러한 알고리즘의 전산화된 구현, 예컨대 GCG® Wisconsin Package®(Accelrys Inc., Burlington, MA)의 일환으로 이용 가능한 GAP, BESTFIT, FASTA 및 TFASTA에 의해 수행될 수 있다. 평가 서열 및 참조 서열의 정렬된 절편에 대한 "동일성 비율"은, 2개의 정렬된 서열이 공유하는 동일한 성분의 수를, 참조 서열 절편에서 성분의 총 수, 즉 전체 참조 서열 또는 참조 서열의 더 작은 정의된 부분으로 나눈 것이다. 서열 동일성 백분율은 동일성 비율에 100을 곱한 것으로 나타난다. 하나 이상의 폴리뉴클레오타이드 서열의 비교는 전장 폴리뉴클레오타이드 서열 또는 이의 일부 또는 더 긴 폴리뉴클레오타이드 서열에 대한 것일 수 있다. 본 개시의 목적을 위해 "동일성 백분율"은 또한 번역된 뉴클레오타이드 서열에 대해 BLASTX 버전 2.0을 그리고 폴리뉴클레오타이드 서열에 대해 BLASTN 버전 2.0을 사용하여 결정될 수 있다.

[0127] 서열 동일성 백분율은 바람직하게는 Sequence Analysis Software Package™(버전 10; Genetics Computer Group, Inc., Madison, WI)의 "Best Fit" 또는 "Gap" 프로그램을 사용하여 결정된다. "Gap"은 매칭 수를 최대화하고 갭 수를 최소화하는 2개 서열의 정렬을 찾기 위해 Needleman 및 Wunsch의 알고리즘(Needleman and Wunsch, JOURNAL OF MOLECULAR BIOLOGY, Journal of Molecular Biology 48:443-453, 1970)을 이용한다. "Best Fit"은 Smith 및 Waterman의 국소 상동성 알고리즘(Smith and Waterman, ADVANCES IN APPLIED MATHEMATICS, 2:482-489, 1981, Smith et al., Nucleic Acids Research 11:2205-2220, 1983)을 사용하여 2개 서열 간 최대 유사성 절편의 최적 정렬을 수행하고 매칭 수를 최대화하기 위해 갭을 삽입한다. 동일성 백분율은 가장 바람직하게는 "Best Fit" 프로그램을 사용하여 결정된다.

[0128] 서열 동일성을 결정하기 유용한 방법은 또한 National Library of Medicine(National Institute of Health, Bethesda, Md. 20894)의 NCBI(National Center Biotechnology Information)에서 공개적으로 이용 가능한 BLAST(Basic Local Alignment Search Tool) 프로그램에 개시되어 있다; 문헌[BLAST Manual, Altschul et al., NCBI, NLM, NIH; Altschul et al., J. MOL. BIOL. 215:403-410 (1990)]을 참고한다; 버전 2.0 이상의 BLAST 프로그램은 갭(결실 및 삽입)의 정렬 내 도입을 허용한다; 펩타이드 서열에 있어서 BLASTX가 서열 동일성을 결정하기 위해 사용될 수 있으며; 폴리뉴클레오타이드 서열에 있어서 BLASTN이 서열 동일성을 결정하기 위해 사용될 수 있다.

[0129] 본원에서 사용되는 용어 "상당한 서열 동일성 백분율"은 적어도 약 70% 서열 동일성, 적어도 약 80% 서열 동일성, 적어도 약 85% 동일성, 적어도 약 90% 서열 동일성 또는 훨씬 더 큰 서열 동일성, 예컨대 약 98% 또는 약 99% 서열 동일성의 서열 동일성 백분율을 나타낸다. 따라서, 본 개시의 하나의 구현에는 본원에 기재된 폴리뉴클레오타이드 서열과 적어도 약 70% 서열 동일성, 적어도 약 80% 서열 동일성, 적어도 약 85% 동일성, 적어도

약 90% 서열 동일성 또는 훨씬 더 큰 서열 동일성, 예컨대 약 98% 또는 약 99% 서열 동일성을 갖는 폴리뉴클레오타이드 분자이다. 본 개시의 Blu1 및 CP1 유전자의 활성을 갖는 폴리뉴클레오타이드 분자는 다양한 스테비올 글리코시드의 제조를 지시할 수 있고 본원에 제공된 폴리뉴클레오타이드 서열과 상당한 서열 동일성 백분율을 가지며 본 개시의 범위 내에 포괄된다.

[0130] 동일성 및 유사성

[0131] 동일성은 서열 정렬(서열 정보 또는 구조 정보 또는 일부 다른 정보만을 사용하여 수행될 수 있지만, 보통 서열 정보에만 기반함) 후 한 쌍의 서열 간에 동일한 아미노산의 비율이며, 유사성은 일부 유사성 매트릭스를 사용하는 정렬에 기반하여 할당되는 스코어이다. 유사성 지수는 하기 BLOSUM62, PAM250 또는 GONNET 또는 단백질의 서열 정렬을 위해 당업자가 사용하는 임의의 매트릭스 중 하나일 수 있다.

[0132] 동일성은 2개의 하위-서열 간에 대응하는 정도이다(서열 간 갭 없음). 25% 이상은 동일성은 기능의 유사성을 시사하는 반면, 18% 내지 25%는 구조 또는 기능의 유사성을 시사한다. 2개의 완전 무관한 또는 무작위 서열(100개 잔기 초과임)이 20% 초과 동일성을 가질 수 있음을 유념한다. 유사성은 2개 서열이 비교되는 경우 이들 간 닮음의 정도이다. 이는 이의 동일성에 의존한다.

[0133] 소비재

[0134] 또 다른 양태에서, 본 개시는 본원에 기재된 바와 같은 레바우디오시드, 예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M을 포함하는 소비재에 대한 것이다. 일부 구현예에서, 소비재는 본원에 기재된 바와 같은 레바우디오시드, 예컨대, 감미량의 Reb W1, Reb W2, Reb D4 및/또는 Reb M을 포함한다. 일부 구현예에서, 소비재는 음료 제품, 식품 제품, 영양제, 약제, 식이 보충물질, 치과 위생 조성물, 식품 겔 조성물, 미용 제품 및 식탁용 풍미제로 구성되는 군으로부터 선택된다.

[0135] 일부 구현예에서, 소비재는 약 1%(w/v-%) 내지 약 4%(w/v-%) 수크로스 용액과 동등한 단맛 강도를 가질 수 있다.

[0136] 일부 구현예에서, 본원에 기재된 바와 같은 레바우디오시드, 예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M은 경구 소비재에서 유일한 감미제이다.

[0137] 일부 구현예에서, 소비재는 또한 적어도 하나의 추가 감미제를 가질 수 있다. 적어도 하나의 추가 감미제는, 예를 들어 천연 고강도 감미제일 수 있다. 추가 감미제는 스테비아(*stevia*) 추출물, 스테비올 글리코시드, 스테비오시드, 레바우디오시드 A, 레바우디오시드 B, 레바우디오시드 C, 레바우디오시드 D, 레바우디오시드 D2, 레바우디오시드 E, 레바우디오시드 F, 돌코시드 A, 루부소시드, 스테비올바이오시드, 수크로스, 고프룩토스 옥수수 시럽, 프룩토스, 글루코스, 자일로스, 아라비노스, 람노스, 에리트리톨, 자일리톨, 만니톨, 소르비톨, 이노시톨, AceK, 아스파탐, 네오타, 수크랄로스, 사카린, 나링긴 디하이드로칼콘(NarDHC), 네오헤스페리딘 디하이드로칼콘(NDHC), 모그로시드 IV, 시아메노시드 I, 모그로시드 V, 모나틴, 타우마틴, 모넨린, 브라제인, L-알라닌, 글리신, Lo Han Guo, 헤르난돌신, 펠로돌신, 트릴롭타인 및 이의 조합으로부터 선택될 수 있다.

[0138] 일부 구현예에서, 소비재는 또한 적어도 하나의 첨가제를 가질 수 있다. 첨가제는, 예를 들어, 탄수화물, 폴리올, 아미노산 또는 이의 염, 폴리아미노산 또는 이의 염, 당 산 또는 이의 염, 뉴클레오타이드, 유기 산, 무기 산, 유기 염, 유기 산 염, 유기 염기 염, 무기 염, 더 쓴 화합물, 풍미료, 풍미 성분, 수렴성 화합물, 단백질, 단백질 가수분해물, 계면활성제, 유화제, 플라보노이드, 알코올, 중합체 및 이의 조합일 수 있다.

[0139] 하나의 양태에서, 본 개시는 본원에 기재된 바와 같은 감미량의 레바우디오시드, 예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M을 포함하는 음료 제품에 대한 것이다.

[0140] 음료 제품은, 예를 들어, 탄산 음료 제품 또는 비-탄산 음료 제품일 수 있다. 음료 제품은 또한, 예를 들어, 소프트 드링크, 파운틴(fountain) 음료, 프로즌 음료, 즉석 음료, 즉석 프로즌 음료, 커피, 차, 유제품 음료, 분말화된 소프트 드링크, 액체 농축액, 풍미가 가미된 물, 강화수, 과일 주스, 과일 주스 풍미가 가미된 음료, 스포츠 드링크 또는 에너지 드링크일 수 있다.

[0141] 일부 구현예에서, 본 개시의 음료 제품에는 하나 이상의 음료 성분, 예컨대 신맛 가미제, 과일 주스 및/또는 야채 주스, 펄프 등, 풍미제, 착색제, 보존제, 비타민, 미네랄, 전해질, 에리트리톨, 타가토스, 글리세린 및 이산화탄소가 포함될 수 있다. 이러한 음료 제품은 임의의 적합한 형태로, 예컨대 음료 농축액 또는 즉석 탄산 음료로 제공될 수 있다.

- [0142] 소정 구현예에서, 본 개시의 음료 제품은 임의의 여러 상이한 특정 제형 또는 구성을 가질 수 있다. 본 개시의 음료 제품의 제형은 제품의 의도되는 시장 구획, 이의 요망되는 영양 특징, 풍미 프로파일 등과 같은 요인에 따라, 소정 정도까지 변할 수 있다. 예를 들어, 소정 구현예에서, 일반적으로 특정 음료 제품의 제형에 추가 성분을 첨가하는 것이 옵션일 수 있다. 예를 들어, 추가 감미제가 첨가될 수 있고, 풍미제, 전해질, 비타민, 과일 주스 또는 다른 과일 제품, 테이스턴트(tastent), 차폐제 등, 풍미 증강제 및/또는 탄산이 전형적으로 맛, 구감, 영양 특징 등을 변화시키기 위해 임의의 이러한 제형에 첨가될 수 있다. 구현예들에서, 음료 제품은 물, 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M), 신맛 가미제 및 풍미제를 함유하는 콜라 음료일 수 있다. 예시적인 풍미제는, 예를 들어, 콜라 풍미제, 시트러스 풍미제 및 스파이스 풍미제일 수 있다. 일부 구현예에서, 이산화탄소 형태의 탄산이 발포(effervescence)를 위해 첨가될 수 있다. 다른 구현예에서, 다른 성분, 제조 기법, 요망되는 보관 기간 등에 따라 보존제가 첨가될 수 있다. 소정 구현예에서, 카페인이 첨가될 수 있다. 일부 구현예에서, 음료 제품은 특징적으로 탄산수, 감미제, 콜라 너트 추출물 및/또는 다른 풍미제, 카라멜 착색제, 하나 이상의 산 및 선택적으로 다른 성분을 함유하는 콜라-풍미 탄산 음료일 수 있다.
- [0143] 또 다른 양태에서, 본 개시는 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)를 포함하는 소비재에 대한 것이며, 여기서 소비재는 식품 제품, 영양제, 약제, 식이 보충물질, 치과 위생 조성물, 식용 겔 조성물, 미용 제품 또는 식탁용 풍미제이다. 일부 구현예에서, 레바우디오시드는 감미량으로 존재한다.
- [0144] 본원에서 사용되는 "식이 보충물질(들)"은 식이를 보충하고 식이에서 누락될 수 있거나 충분한 양을 소비하지 못할 수 있는 영양분, 예컨대 비타민, 미네랄, 섬유소, 지방산, 아미노산 등을 제공하기 위한 화합물을 나타낸다. 당분야에 알려진 임의의 적합한 식이 보충물질이 사용될 수 있다. 적합한 식이 보충물질의 예는, 예를 들어, 영양분, 비타민, 미네랄, 섬유소, 지방산, 허브, 식물, 아미노산 및 대사물질일 수 있다.
- [0145] 본원에서 사용되는 "영양제(들)"는 질환 또는 장애(예컨대, 피로, 불면증, 노화의 효과, 기억력 손실, 기분 장애, 심혈관 질환 및 혈중 고수준의 콜레스테롤, 당뇨병, 골다공증, 염증, 자가면역 장애 등)의 예방 및/또는 치료를 포함하는, 의학적 또는 건강 이익을 제공할 수 있는 임의의 식품 또는 식품의 일부를 포함하는 화합물을 나타낸다. 당분야에 알려진 임의의 적합한 영양제가 사용될 수 있다. 일부 구현예에서, 영양제는 식품 및 음료에 대한 보충물질로서 그리고 고체 제형, 예컨대 캡슐 또는 정제, 또는 액체 제형, 예컨대 용액 또는 현탁액일 수 있는 장 또는 비경구 적용을 위한 약학 제형으로서 사용될 수 있다.
- [0146] 일부 구현예에서, 식이 보충물질 및 영양제는 보호 하이드로콜로이드(예컨대 검, 단백질, 변형 전분), 결합제, 제막제, 캡슐화제/물질, 벽/셀 물질, 매트릭스 화합물, 코팅, 유화제, 표면 활성제, 가용화제(오일, 지방, 왁스, 레시틴 등), 흡착제, 담체, 충전제, 공동-화합물, 분산제, 수화제, 가공 보조제(용매), 유동제, 맛-차폐제, 증량제, 젤리화제, 겔-형성제, 향산화제 및 향균제를 추가로 함유할 수 있다.
- [0147] 본원에서 사용되는 "겔"은 입자의 네트워크가 액체 배지 부피에 걸쳐 존재하는 콜로이드성 시스템을 나타낸다. 겔은 주로 액체로 이루어지며 이에 따라 액체와 유사한 밀도를 나타내지만, 겔은 액체 배지에 걸쳐 존재하는 입자의 네트워크로 인해 고체의 구조적 응집을 갖는다. 상기 이유로 인해, 겔은 일반적으로 고체인 젤리형 물질로 나타난다. 겔은 여러 적용에서 사용될 수 있다. 예를 들어, 겔은 식품, 도료 및 접착제에서 사용될 수 있다. 먹을 수 있는 겔은 "식용 겔 조성물"로서 언급된다. 식용 겔 조성물은 전형적으로 스낵으로, 디저트로, 주식의 일부로 또는 주식과 함께 먹게 된다. 적합한 식용 겔 조성물의 예는, 예를 들어, 겔 디저트, 푸딩, 잼, 젤리, 페이스트, 트라이플, 아스픽, 마시멜로우, 구미 캔디 등일 수 있다. 일부 구현예에서, 식용 겔 혼합물은 일반적으로 유체가 첨가되어 식용 겔 조성물을 형성할 수 있는 분말화 또는 과립 고체이다. 적합한 유체의 예는, 예를 들어 물, 유제품액, 유제품 유사액, 주스, 알코올, 알코올계 음료 및 이의 조합일 수 있다. 적합한 유제품액의 예는, 예를 들어 우유, 배양유, 크림, 유장 및 이의 혼합물일 수 있다. 적합한 유제품 유사액의 예는, 예를 들어 두유 및 비-유제품 커피 프림일 수 있다.
- [0148] 본원에서 사용되는 용어 "겔화 성분"은 액체 배지 내에 콜로이드성 시스템을 형성할 수 있는 임의의 물질을 나타낸다. 적합한 겔화 성분의 예는, 예를 들어, 젤라틴, 알기네이트, 카라기난, 검, 펙틴, 곤약, 한천, 식품 산, 렌넷, 전분, 전분 유도체 및 이의 조합일 수 있다. 식용 겔 혼합물 또는 식용 겔 조성물에서 사용되는 겔화 성분의 양이 여러 요인, 예컨대 사용되는 특정 겔화 성분, 사용되는 특정 유체 기재 및 요망되는 겔 특성에 따라 상당히 변할 수 있음은 당업자에게 널리 알려져 있다.
- [0149] 본 개시의 겔 혼합물 및 겔 조성물은 당분야에 알려진 임의의 적합한 방법에 의해 제조될 수 있다. 일부 구현예

에서, 본 개시의 식용 겔 혼합물 및 식용 겔 조성물은 겔화제에 더하여 다른 성분을 사용하여 제조될 수 있다. 다른 적합한 성분의 예는, 예를 들어, 식품 산, 식품 산의 염, 완충 시스템, 벌크화제, 격리제, 가교제, 하나 이상의 풍미제, 하나 이상의 착색제 및 이의 조합일 수 있다.

[0150] 당분야에 알려진 임의의 적합한 약학 조성물이 사용될 수 있다. 일부 구현예에서, 본 개시의 약학 조성물은 생물학적 효과를 발휘하는 하나 이상의 활성 제제를 함유하는 약학 약물을 제형화하기 위해 사용될 수 있다. 따라서, 일부 구현예에서, 본 개시의 약학 조성물은 생물학적 효과를 발휘하는 하나 이상의 활성 제제를 함유할 수 있다. 적합한 활성 제제는 당분야에 널리 알려져 있다(예컨대, 문헌[The Physician's Desk Reference]). 이러한 조성물은, 예를 들어, 문헌[Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa., USA]에 기재된 바와 같이 당분야에 널리 알려진 절차에 따라 제조될 수 있다.

[0151] 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)는 당분야에 알려진 임의의 적합한 치과 및 구강 위생 조성물과 함께 사용될 수 있다. 적합한 치과 및 구강 위생 조성물의 예는, 예를 들어, 치약(toothpaste), 치아 연마제, 치실, 구강세척제, 구강 헥검제, 치약(dentifrice), 구강 스프레이, 구강 리프레쉬너, 치석 헥검제, 치통 완화제 등일 수 있다.

[0152] 본원에서 사용되는 "식품 제품"은 영양가를 가질 수 있고 인간 및 동물에 의한 소비를 위한 것일 수 있지만 반드시 그래야 하는 것은 아닌, 임의의 고체 또는 액체의 소화 가능한 물질을 나타낸다.

[0153] 적합한 식품 제품의 예는, 예를 들어, 과자류 조성물, 예컨대 캔디, 민트, 과일맛 드롭스, 코코아 제품, 초콜렛 등; 양념, 예컨대 케첩, 머스타드, 마요네즈 등; 휴잉 검; 시리얼 조성물; 베이킹 제품, 예컨대 빵, 케이크, 파이, 쿠키 등; 유제품, 예컨대 우유, 치즈, 크림, 아이스크림, 사워 크림, 요거트, 셔벗 등; 식탁용 감미제 조성물; 수프; 스튜; 간편식; 고기, 예컨대 햄, 베이컨, 소시지, 육포 등; 젤라틴 및 젤라틴 유사 제품, 예컨대 잼, 젤리, 설탕절임 등; 과일; 야채; 계란 제품; 아이싱; 당밀을 포함하는 시럽; 스낵; 견과육 및 견과 제품; 및 동물 사료일 수 있다.

[0154] 식품 제품은 또한 허브, 향신료 및 시즈닝, 천연 및 합성 풍미제 및 풍미 증강제, 예컨대 1나트륨 글루타메이트 일 수 있다. 일부 구현예에서, 식품 제품은, 예를 들어, 제조되고 포장된 제품, 예컨대 다이어트용 감미제, 액체 감미제, 과립화 풍미제 혼합물, 애완동물 식품, 가축 사료, 담배 및 베이킹 용도를 위한 물질, 예컨대 빵, 쿠키, 케이크, 팬케이크, 도넛의 제조를 위한 분말화된 베이킹 혼합물 등일 수 있다. 다른 구현예에서, 식품 제품은 또한 수크로스를 거의 또는 전혀 함유하지 않는 다이어트 및 저칼로리 식품 및 음료일 수 있다.

[0155] 임의의 이전 구현예와 조합될 수 있는 소정 구현예에서, 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)는 유일한 감미제이며, 선택적으로 제품은 약 1% 내지 약 4%(w/v-%) 수크로스 용액과 동등한 단맛 강도를 갖는다. 임의의 이전 구현예와 조합될 수 있는 소정 구현예에서, 소비재 및 음료 제품에는 추가 감미제가 추가로 포함될 수 있고, 선택적으로 제품은 약 1% 내지 약 10%(w/v-%) 수크로스 용액과 동등한 단맛 강도를 갖는다. 임의의 이전 구현예와 조합될 수 있는 소정 구현예에서, 제품 내의 모든 감미 성분은 고강도 감미제이다. 임의의 이전 구현예와 조합될 수 있는 소정 구현예에서, 제품 내의 모든 감미 성분은 천연 고강도 감미제일 수 있다. 임의의 이전 구현예와 조합될 수 있는 소정 구현예에서, 추가 감미제는 스테비아(*stevia*) 추출물, 스테비올 글리코시드, 스테비오시드, 레바우디오시드 A, 레바우디오시드 B, 레바우디오시드 C, 레바우디오시드 D, 레바우디오시드 D2, 레바우디오시드 F, 돌코시드 A, 루부소시드, 스테비올바이오시드, 수크로스, 고프룩토스 옥수수 시럽, 프룩토스, 글루코스, 자일로스, 아라비노스, 람노스, 에리트리톨, 자일리톨, 만니톨, 소르비톨, 이노시톨, AceK, 아스파탐, 네오탐, 수크랄로스, 사카린, 나링긴 디하이드로칼콘(NarDHC), 네오헤스페리딘 디하이드로칼콘(NDHC), 루부소시드, 모그로시드 IV, 시아메노시드 I, 모그로시드 V, 모나틴, 타우마틴, 모넨린, 브라제인, L-알라닌, 글리신, Lo Han Guo, 헤르난둘신, 필로둘신, 트릴롭타인 및 이의 조합으로부터 선택되는 하나 이상의 감미제를 함유한다. 임의의 이전 구현예와 조합될 수 있는 소정 구현예에서, 소비재 및 음료 제품에는 탄수화물, 폴리올, 아미노산 또는 이의 염, 폴리아미노산 또는 이의 염, 당 산 또는 이의 염, 뉴클레오타이드, 유기 산, 무기 산, 유기 염, 유기 산 염, 유기 염기 염, 무기 염, 더 쓴 화합물, 풍미료, 풍미 성분, 수렴성 화합물, 단백질, 단백질 가수분해물, 계면활성제, 유화제, 플라보노이드, 알코올, 중합체 및 이의 조합으로부터 선택되는 하나 이상의 첨가제가 추가로 포함될 수 있다. 임의의 이전 구현예와 조합될 수 있는 소정 구현예에서, 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)는 이것이 제품 내로 첨가되기 전에 약 50 중량% 내지 약 100 중량%의 순도를 갖는다.

[0156] 일부 구현예에서, 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의

조합, 예컨대 Reb D4 및 Reb M)는 충전제, 벌크화제 및 응집방지제 중 한 가지 이상을 추가로 포함하는 조성물에 제공된다. 적합한 충전제, 벌크화제 및 응집방지제는 당분야에 알려져 있다.

[0157] 소정 구현예에서, 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)는 소비재 및 음료 제품을 달게 하고/하거나 그 단맛을 증강시키기 충분한 최종 농도로 포함되고/되거나 첨가될 수 있다. 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)의 "최종 농도"는 최종 소비재 및 음료 제품에(즉, 모든 성분 및/또는 화합물이 첨가되어 소비재 및 음료 제품을 제조한 후) 존재한다. 따라서 소정 구현예에서, 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)는 소비재 및 음료 제품을 제조하기 위해 사용되는 화합물 또는 성분에 포함되고/되거나 첨가된다. 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)는 단일 화합물 또는 성분, 또는 여러 화합물 및 성분 존재할 수 있다. 일부 구현예에서, 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)는 소비재 및 음료 제품에 포함되고/되거나 첨가된다.

[0158] 소정 구현예에서, 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M)는 소비재 및 음료 제품에 포함되고/되거나 첨가되는 유일한 감미제이다. 일부 구현예에서, 레바우디오시드를 포함하는 소비재 및 음료 제품은 약 1% 내지 약 4%(w/v-%) 수크로스 용액, 약 1% 내지 약 3%(w/v-%) 수크로스 용액 또는 약 1% 내지 약 2%(w/v-%) 수크로스 용액과 동등한 단맛 강도를 갖는다. 대안적으로, 소비재 및 음료 제품은 약 1% 내지 약 4%(w/v-%) 수크로스 용액, 약 2% 내지 약 4%(w/v-%) 수크로스 용액, 약 3% 내지 약 4%(w/v-%) 수크로스 용액 또는 약 4%(w/v-%) 수크로스 용액과 동등한 단맛 강도를 갖는다. 예를 들어, 소비재 및 음료 제품은 이들 값 사이의 임의의 범위를 포함하여, 약 1%, 약 2%, 약 3% 또는 약 4%(w/v-%) 수크로스 용액과 동등한 단맛 강도를 가질 수 있다.

[0159] 본 개시의 소비재 및 음료 제품에는 본원에 기재된 바와 같은 레바우디오시드(예컨대, Reb W1, Reb W2, Reb D4, Reb M 또는 이의 조합, 예컨대 Reb D4 및 Reb M) 및 본 개시의 하나 이상의 감미제의 혼합물이 바람직한 단맛 강도, 영양 특징, 맛 프로필, 구감 또는 다른 감각자극 요인을 달성하기 충분한 비로 포함될 수 있다.

[0160] 상기 설명으로부터 명백하듯이, 본 개시의 소정 양태는 본원에 예시되는 실시예의 특정 세부사항에 의해 제한되지 않으며, 이에 따라 다른 변형 및 적용 또는 이의 균등부가 당업자에게 생겨날 것임이 고려된다. 따라서 청구 범위는 본 개시의 정신 및 범위에서 벗어나지 않는 모든 이러한 변형 및 적용을 커버하려는 것이다.

[0161] 또한 달리 정의되지 않는 한, 본원에 사용된 모든 기술 및 과학 용어는 본 개시가 속하는 당업자에 의해 일반적으로 이해되는 바와 동일한 의미를 갖는다. 바람직한 방법 및 물질이 상술되어 있지만, 본원에 기재된 것들과 유사하거나 동등한 임의의 방법 및 물질이 본 개시의 실시 또는 평가에서 사용될 수 있다.

[0162] 상기 발명이 이해의 목적을 위해 예시 및 예로서 일부 상세하게 설명되었으나, 당업자에게는 소정 변화 및 변형이 실시될 수 있음이 명백할 것이다. 따라서, 설명 및 실시예는 첨부되는 청구범위에 의해 서술되는 본 발명의 범위를 제한하는 것으로 간주되어서는 안 된다.

[0163] 실시예 1: Reb D4의 효소적 합성

[0164] Reb D4의 몇몇 효소적 제조 방법이 존재한다. Reb W로부터 시작하는 하나의 방법이 여기에 제시된다.

[0165] 이전에, 본 발명자들은 Reb V로부터 Reb W의 제조를 실증하였다(WO2016054540). 여기에서, 본 발명자들은 Reb W가 피치아 파스토리스(*Pichia pastoris*)로부터의 베타-글루코시다제(B-glu1, SEQ:5)에 의해 가수분해되어 우리가 "레바우디오시드 WB1"로 부르는 신규한 스테비올 글리코시드를 제조할 수 있음을 확인하였다. 제조된 레바우디오시드 WB1은 다시 B-glu1에 의해 가수분해되어 레바우디오시드 WB2를 제조할 수 있다(도 14 참고).

[0166] 보다 구체적으로, B-glu1(SEQ ID NO: 6) 유전자의 전장 DNA 단편을 합성하였다. 구체적으로, 에스케리치아 콜라이(*E. coli*) 발현을 위해 cDNA를 코돈 최적화하였다(Genscript, Piscataway, NJ). 합성된 DNA를 박테리아 발현 벡터 pETite N-His SUMO Kan 벡터(Lucigen) 내로 클로닝하였다. B-glu1(SEQ ID NO:5 참고)을 인코딩하는 뉴클레오타이드 서열(SEQ ID NO: 6)을 프레임에 맞춰 삽입하였다.

[0167] 발현 작제물을 에스케리치아 콜라이 BL21(DE3) 내로 형질전환하고, 이후 OD600이 0.8 내지 1.0에 도달할 때까지 37°C에서 50 $\mu\text{g/ml}$ 카나마이신을 함유하는 LB 배지에서 성장시켰다. 1 mM 이소프로필 β -D-1-티오갈락토피라노시드(IPTG)의 첨가에 의해 단백질 발현을 유도하고, 배양을 22 hr 동안 16°C에서 추가 성장시켰다. 세포를 원심

분리($3,000 \times g$; 10분; 4°C)에 의해 수확하였다. 세포 펠렛을 수집하고, 즉시 사용하거나 -80°C 에서 보관하였다.

[0168] 세포 펠렛을 용해 완충액(50 mM 인산칼륨 완충액, pH 7.2, $25 \mu\text{g/ml}$ 라이소자임, $5 \mu\text{g/ml}$ DNase I, 20 mM 이미다졸, 500 mM NaCl, 10% 글리세롤 및 0.4% TRITON X-100) 중에 재현탁하였다. 세포를 4°C 에서 초음파 분쇄에 의해 손상시키고, 세포 파편을 원심분리($18,000 \times g$; 30분)에 의해 청정화하였다. 상청액을 평형화된(평형 완충액: 50 mM 인산칼륨 완충액, pH 7.2, 20 mM 이미다졸, 500 mM NaCl, 10% 글리세롤) Ni-NTA(Qiagen) 친화도 칼럼에 로딩하였다. 단백질 샘플의 로딩 후, 칼럼을 평형 완충액으로 세척하여 미결합 오염 단백질을 제거하였다. His-태그가 붙은 B-glu1 재조합 폴리펩타이드를 250 mM 이미다졸을 함유하는 평형 완충액에 의해 용출하였다.

[0169] 재조합 B-glu1($10 \mu\text{g}$)을 $200 \mu\text{l}$ 시험관내 반응 시스템에 첨가하였다. 반응 시스템은 50 mM 인산칼륨 완충액, pH 7.2 및 1 mg/ml 레바우디오시드 W를 기질로서 함유하였다. 반응을 37°C 에서 수행하고 $200 \mu\text{l}$ 의 1-부탄올을 첨가하여 종결시켰다. 샘플을 $200 \mu\text{l}$ 의 1-부탄올로 3회 추출하였다. 풀링된 분획을 건조하고 고성능 액체 크로마토그래피(HPLC) 분석을 위해 $70 \mu\text{l}$ 의 80% 메탄올 중에 용해시켰다.

[0170] 사차 펌프, 온도 제어 칼럼 구획, 자동 샘플수집장치 및 UV 흡광도 검출기를 포함하는 Dionex UPLC ultimate 3000 시스템(Sunnyvale, CA)을 사용하여 HPLC 분석을 수행하였다. 가드 칼럼을 포함하는 Synergi Hydro-RP 칼럼을 스테비올 글리코시드의 특성규명을 위해 사용하였다. 수중 아세토니트릴을 HPLC 분석에서 용출을 위해 사용하였다. 검출 파장은 210 nm 였다.

[0171] 도 3에 나타난 바와 같이, B-glu1은 레바우디오시드 W 기질을 가수분해하여 1시간째에 레바우디오시드 WB1을 제조하였다(도 3의 B). 제조된 레바우디오시드 WB1은 이후 반응 시점에 레바우디오시드 WB2로 추가 전환될 수 있다(도 3의 C 및 도 3의 D).

[0172] 결론적으로, 레바우디오시드 W가 B-glu1에 의해 가수분해되어 WB1을 제조하였으며, 제조된 WB1은 B-glu1에 의해 추가 가수분해되어 WB2를 제조하였다.

[0173] 상기 중간체 레바우디오시드 WB2는 UGT85C2 효소와의 인큐베이션에 의해 레바우디오시드 WB1로 다시 전환될 수 있다(도 4). 재조합 UGT85C2 효소($10 \mu\text{g}$)를 $200 \mu\text{l}$ 시험관내 반응 시스템에서 평가하였다. 반응 시스템은 50 mM 인산칼륨 완충액, pH 7.2, 3 mM MgCl_2 , 0.5 mg/ml 레바우디오시드 WB2 기질 및 3 mM UDP-글루코스를 함유하였다. 도 3에 나타난 바와 같이, 레바우디오시드 WB2는 UGT85C2(SEQ ID NO: 7, 도 4)에 의해 레바우디오시드 WB1로 전환될 수 있다. UGT85C2 효소는 글루코스를 C4 카복실의 C-13에 부가하여 스테비올로부터 스테비올-13-모노사이드를 형성하는 활성을 갖는다. 이들 결과는 B-glu1이 레바우디오시드 WB1의 C13 위치로부터 글루코스를 가수분해하여 레바우디오시드 WB2를 제조함을 시사하였다. 레바우디오시드 WB1 및 레바우디오시드 WB2의 예측 구조를 도 2에 나타낸다. 구조들은 LC-MS 분석에 의해 확인되었다(도 5).

[0174] 상기 중간체 Reb WB1은 HV1 UGT 효소(WO/2015/065650)와의 인큐베이션에 의해 Reb D4로 전환될 수 있다. 재조합 HV1($10 \mu\text{g}$)을 $200 \mu\text{l}$ 시험관내 반응 시스템에서 평가하였다. 반응 시스템은 50 mM 인산칼륨 완충액, pH 7.2, 3 mM MgCl_2 , 0.5 mg/ml 의 레바우디오시드 WB1 기질 및 3 mM UDP-글루코스를 함유하였다. 도 4에 나타난 바와 같이, 레바우디오시드 WB1은 6시간째에 HV1에 의해 레바우디오시드 D4로 완전 전환될 수 있다(도 6의 C).

[0175] 상기 효소 반응에 따르면, D4의 구조는 (13-[(2-O- β -D-글루코피라노실- β -D-글루코피라노실)옥시]엔트-카우르-16-엔-19-오산-[(2-O- β -D-글루코피라노실-3-O- β -D-글루코피라노실- β -D-글루코피라노실)에스테르]로 예측되었다(도 7a). Reb D4의 구조는 LC-MS에 의해 확인되었다(도 7b). 질량 분광 분석은 예측 구조와 동일한 질량 [$(\text{M}+\text{Na})$ 1151.47 m/z]을 나타내었다.

[0176] 실시예 2: 야생형 효소에 의한 Reb D4의 Reb M으로의 전환

[0177] 본 발명자들은 Reb D4가 UGT76G1에 의해 Reb M으로 추가 전환될 수 있음을 발견하였다. UGT76G1(SEQ ID NO: 2)의 전장 DNA 단편을 합성하였다. cDNA를 에스케리치아 콜라이(*E. coli*) 발현을 위해 코돈 최적화하였다(Genscript). 합성된 DNA를 박테리아 발현 벡터 pETite N-His SUMO Kan 벡터(Lucigen) 내로 클로닝하였다. 76G1을 인코딩하는 뉴클레오타이드 서열을 프레임에 맞춰 삽입하였다.

[0178] 발현 작제물을 에스케리치아 콜라이 BL21(DE3) 내로 형질전환하고, 이후 OD600이 0.8 내지 1.0에 도달할 때까지 37°C 에서 $50 \mu\text{g/ml}$ 카나마이신을 함유하는 LB 배지에서 성장시켰다. 0.5 mM IPTG의 첨가에 의해 단백질 발현을 유도하고, 배양을 22 hr 동안 16°C 에서 추가 성장시켰다. 세포를 원심분리($3,000 \times g$; 10분; 4°C)에 의해 수확

하였다. 세포 펠렛을 수집하고, 즉시 사용하거나 -80°C 에서 보관하였다.

[0179] 세포 펠렛을 상술된 바와 같은 용해 완충액 중에 재현탁하였다. 세포를 4°C 에서 초음파 분쇄에 의해 손상시키고 세포 파편을 원심분리($18,000 \times g$; 30분)에 의해 청정화하였다. 상청액을 상술된 바와 같이 평형화된 Ni-NTA(Qiagen) 친화도 칼럼에 로딩하였다. 단백질 샘플의 로딩 후, 칼럼을 평형 완충액으로 세척하여 미결합 오염 단백질을 제거하였다. His-태그가 붙은 76G1 재조합 폴리펩타이드를 250 mM 이미다졸을 함유하는 평형 완충액에 의해 용출하였다.

[0180] 재조합 UGT76G1($10 \mu\text{g}$)을 $200 \mu\text{l}$ 시험관내 반응 시스템에 첨가하였다. 반응 시스템은 50 mM 인산칼륨 완충액, pH 7.2, 보조인자로서 1 mM UDPG 및 기질로서 1 mg/ml Reb D4를 함유하였다. 반응을 37°C 에서 수행하고 $200 \mu\text{l}$ 의 1-부탄올을 첨가하여 종결시켰다. 샘플을 $200 \mu\text{l}$ 의 1-부탄올로 3회 추출하였다. 폴링된 분획을 건조하고 고성능 액체 크로마토그래피(HPLC) 분석을 위해 $70 \mu\text{l}$ 의 80% 메탄올 중에 용해시켰다.

[0181] 사차 펌프, 온도 제어 칼럼 구획, 자동 샘플수집장치 및 UV 흡광도 검출기를 포함하는 Dionex UPLC ultimate 3000 시스템(Sunnyvale, CA)을 사용하여 HPLC 분석을 수행하였다. 가드 칼럼을 포함하는 Synergi Hydro-RP 칼럼을 스테비올 글리코시드의 특성규명을 위해 사용하였다. 수증 아세토니트릴을 HPLC 분석에서 용출을 위해 사용하였다. 검출 파장은 210 nm였다.

[0182] 도 13에 나타난 바와 같이, UGT76G1은 Reb D4를 Reb M으로 전환시킬 수 있다(도 13의 C; 및 도 13의 F).

[0183] 실시예 3: Reb M으로의 Reb D4를 촉매하는 UGT 효소의 반응 중심 분해

[0184] UDP-글루코실 트랜스퍼라제의 화학 공정을 보다 효율적으로 결정하기 위해, 본 발명자들은 야생형 스테비올 UDP-글루코실 트랜스퍼라제 UGT76G1의 결정 구조를 획득하였다. 상기 효소는 스테비올 글리코시드 바이오-전환을 수행하는 것으로 최초 보고된 효소이다. 본 발명자들은 Reb D4에서 Reb M으로의 전환을 보다 완벽하게 이해하기 위해 이를 위한 효소를 설계하고, 이후 상기 지식을 사용하여 Reb D4에서 Reb M으로의 바이오전환에 유용할 수 있는 효소를 보다 효율적으로 찾거나 설계하기 위해 반응 중심 및 기질 결합 부위의 구조 정보를 입수하고자 하였다.

[0185] 셀레노메티오닌(SeMet)-치환 단백질의 제조를 위해, 에스케리치아 콜라이 BL21(DE3) 세포를 pET-28a-UGT76G1 벡터로 형질전환하고, A600nm가 약 0.8이 될 때까지 37°C (250 rpm)에서 $50 \mu\text{g/ml}$ 카나마이신을 함유하는 SeMet가 보충된 M9 최소 배지(Doublie, 2007)에서 성장시켰다. 이소프로필 1-티오- β -D-갈락토피라노시드(0.8 mM 최종)의 첨가로 단백질 발현을 유도하고, 세포를 하룻밤 동안 성장시켰다(16°C). 세포 펠렛을 원심분리($10,000 \times g$; 10분)에 의해 수확하고, 용해 완충액(50 mM 트리스, pH 8.0, 500 mM NaCl, 20 mM 이미다졸, 1 mM β -메르캅토에탄올(β -ME), 10%(v/v) 글리세롤 및 1%(v/v) Tween-20) 중에 현탁하였다. 초음파 분쇄에 의한 용해 후, 세포 파편을 원심분리($30,000 \times g$; 45분)에 의해 제거하고 상청액을 세척 완충액(용해 완충액에서 Tween-20 제외)으로 평형화된 Ni²⁺-니트릴로아세트산(NTA; Qiagen) 칼럼으로 통과시켰다. 로딩 후, 칼럼을 10배 칼럼 부피의 세척 완충액으로 세척하였다. 결합된 융합 단백질을 용출 완충액(세척 완충액에 250 mM 이미다졸 추가)으로 용출하고 수집하였다. 추가 정제를 위해, 50 mM 트리스, pH 8.0, 25 mM NaCl, 1 mM 트리스(2-카복시에틸)포스핀(TCEP)으로 평형화된 Superdex-200 26/60 HiLoad FPLC 칼럼 상에서 크기-배제 크로마토그래피를 수행하였다. 피크 분획을 수집하고 원심분리 농축기(Amicon)를 사용해서 농축하고, 소 혈청 알부민을 표준물질로 하여 브래드포드 검정을 사용해서 단백질 농도를 결정하였다. 정제된 단백질을 액체 질소 중에 급속 냉동하고 -80°C 에서 보관하였다.

[0186] 정제된 UGT76G1을 10 mg/ml 로 농축하고 $2 \mu\text{l}$ 방울(1:1의 농축 단백질 및 결정화 조건)로 행잉-드랍 증기-확산 방법을 사용해서 결정화하였다. 20%(w/v) PEG-4000, 20% 2-프로판올(v/v) 및 100 mM 나트륨 시트레이트 3염기성 2수화물 완충액(pH 5.6)으로 4°C 에서 회절 품질 결정을 획득하였다. 동결보호제로 25% 글리세롤을 함유하는 모액으로 개별 결정을 액체 질소 중에 급속 냉동하였다. Argonne National Laboratory Advanced Photon Source 19-ID 빔라인($\lambda=0.98 \text{ \AA}$)에서 회절 데이터(100 K)를 수집하였다. HKL3000(Otwinowski & Minor, 1997)을 사용하여 회절 데이터의 인덱스를 얻고, 적분하고, 스케일링하였다. 단파장 변칙 회절(SAD) 위상처리에 의해 SeMet-치환 UGT76G1의 구조를 결정하였다. SHELX(Sheldrick, 2008)를 사용하여 SeMet 위치를 결정하고 피크 파장 데이터 세트로부터 초기 위상을 산정하였다. SeMet 위치 및 파라미터의 미세화(refinement)는 MLPHARE(Terwilliger, 2000)로 수행하였다. ARP/wARP(Morris et al., 2003)로 구현되는 밀도 변형을 사용하는 용매 평탄화를 채택하여 초기 모델을 구축하였다. 환산-칭동-스크리닝 파라미터 미세화를 포함하는 수동 모델 구축 및 미세화의 후속 반복 회에서는 각각 COOT(Emsley et al., 2010) 및 PHENIX(Adams et al., 2007)를 사용

하였다. 데이터 수집 및 미세화 데이터를 표 1에 요약한다.

표 1

결정학 통계의 요약

데이터 수집	AtGH3.5•AMP•IAA
공간 그룹	P2 ₁
셀 치수	$a = 91.61 \text{ \AA}$, $b = 143.5 \text{ \AA}$, $c = 102.3 \text{ \AA}$; $\beta = 114.7^\circ$
파장(Å)	0.979
해상도(Å) (최고 셀)	42.2 - 2.20 (2.25 - 2.20)
반사(전체/고유)	254,788 / 108,566
완결성(최고 셀)	89.5% (77.3%)
$\langle I/\sigma \rangle$ (최고 셀)	11.2 (2.0)
R _{sym} (최고 셀)	9.3% (51.8%)
미세화	
R _{cryst} / R _{free}	20.5% / 24.5%
단백질 원자 수	18,186
물 수	1,002
리간드 원자 수	154
R.m.s.d., 결합 길이	0.008
R.m.s.d., 결합 각(°)	1.167
평균 B-인자(Å ²): 단백질, 물, 리간드	30.7, 28.3, 26.4
입체 화학: 가장 선호됨, 허용됨, 허용되지 않음	97.3, 2.5, 0.2%

[0187]

[0188]

UGT71G1의 구조는 유사한 Rossmann-유형 폴딩을 갖는 N-말단 도메인 및 C-말단 도메인으로 구성되며, 예측된 바와 같이, GT-B 폴딩에 속한다(도 8). UGT76G1 결정 구조의 표준 배향을 도 8에 나타낸다. N-말단 도메인은 8개의 α 나선이 인접한 중심의 7-쇄 평행 β 시트를 포함한다. 도메인은 또한 촉매성 히스티딘을 포함한다. C-말단 도메인은 7개의 α 나선이 인접한 6-쇄 β 시트를 포함한다(도 9). 2개 도메인은 매우 친밀하게 패킹되며 UDP 분자가 결합되는 깊은 틈새를 형성한다.

[0189]

실시예 4: 돌연변이체의 논리적 설계

[0190]

UGT76G1 구조에 기반하여, 본 발명자들은 원형 순열(PLoS computational Biology, 2012, 8(3) e1002445; BIOINFORMATICS, 2015, (3)) 및 돌연변이 세트를 설계할 수 있었다. 원형 순열 분석은 유용하거나 가치가 있는 효소를 개발하기 위한 강력한 도구이다. 몇몇 원형 순열 버전의 평가 후, 본 발명자들은 매우 높은 활성을 갖는 한 버전의 원형 돌연변이 "원형 순열 1"("CP1")이 가장 높은 활성을 가짐을 확인하였다. 본 개시에 따라, 본 발명자들은 CP1(SEQ ID NO: 3) 효소의 활성 및 Reb D4의 Reb M으로의 전환을 보조하는 능력을 연구하였다.

[0191]

UGT76G1 및 CP1의 구조를 도 10에서 비교한다. UGT76G1 결정 구조의 구조적 속성의 대부분은 UGT 및 CP1 간에서와 같이 유사하지만, CP1은 그 구조의 "베타 시트" 부분에서 상당히 상이한 서열 및 구조를 갖는다(도 11).

[0192]

본 발명의 효소의 촉매 활성을 예측하기 위해, 본 발명자들은 Reb D4를 CP1의 반응 중심 내로 도킹하였다. 본 발명자들은 효소와 Reb D4의 상호작용에 초점을 두고 반응 중심을 강조하였다. 도킹된 레바우디오시드 D4 리간

드는 촉매성 히스티딘 및 결합된 UDP 대비 바람직한 위치에 있다(도 12). 상기 도킹 실험에 기반하여, 본 발명자들은 돌연변이화 연구를 통해 활성 평가에 가치가 있는 특정 잔기를 찾을 수 있었다.

[0193] CP1 모델링 분석에 기반하여, 본 발명자들은 효소 활성을 증가시키기 위한 CP1의 여러 돌연변이 부위를 선택하고 평가하였다. 마지막으로, 본 발명자들은 레바우디오시드 D4의 레바우디오시드 M으로의 바이오전환에 관련된 몇몇 돌연변이 부위를 확인하였다(표 2). CR1은 표 2에서의 적어도 하나의 돌연변이 부위를 포함하는 일종의 CP1 돌연변이체이다.

표 2

CP1의 돌연변이 부위의 요약.

위치	아미노산
3	W-L
6	L-A, L-G
90	T-A; T-G
91	S-G; S-L
93	V-A; V-G
181	S-G
183	F-V; F-A; F-G
184	G-A
185	L-A
350	G-A
389	L-V
410	S-G
418	H-V
450	T-A; T-G
451	K-A
452	D-A
454	K-L; K-V

[0195] 실시예 5: 돌연변이체 UGT의 사용에 의한 Reb D4의 Reb M으로의 전환.

[0196] 본 실시예에서, Reb D4의 레바우디오시드 M으로의 시험관내 전환을 확인하기 위해, 스테비올 글리코시드 기질로서 Reb D4를 사용하여 UGT76G1, CP1 및 효소 돌연변이체를 검정하였다. 제조합 폴리펩타이드(10 μ g)를 200 μ l 시험관내 반응 시스템에서 평가하였다. 반응 시스템은 50 mM 인산칼륨 완충액, pH 7.2, 3 mM $MgCl_2$, 1 mg/ml 스테비올 글리코시드 기질 및 1 mM UDP-글루코스를 함유하였다. 반응을 30°C에서 수행하고 200 μ l의 1-부탄올을 첨가하여 종결시켰다. 샘플을 200 μ l의 1-부탄올로 3회 추출하였다. 풀링된 분획을 건조하고 고성능 액체 크로마토그래피(HPLC) 분석을 위해 70 μ l의 80% 메탄올 중에 용해시켰다. 레바우디오시드 D4를 기질로 사용하였다. 사차 펌프, 온도 제어 칼럼 구획, 자동 샘플수집장치 및 UV 흡광도 검출기를 포함하는 Dionex UPLC ultimate 3000 시스템(Sunnyvale, CA)을 사용하여 HPLC 분석을 수행하였다. 가드 칼럼을 포함하는 Synergi Hydro-RP 칼럼을 스테비올 글리코시드의 특성규명을 위해 사용하였다. 수중 아세토니트릴을 HPLC 분석에서 용출을 위해 사용하였다. 검출 파장은 210 nm였다.

[0197] 도 13에 나타난 바와 같이, UGT76G1, CP1 및 CR1 돌연변이체는 하나의 글루코스 분자를 Reb D4로 전달하여 Reb M을 형성할 수 있다. 그러나, CP1 및 CR1은 UGT76G1 효소보다 유의미하게 높은 효소 활성을 갖는다.

[0198] 실시예 6: NMR에 의해 분석된 Reb WB2의 구조.

[0199] Reb WB2의 특성규명을 위해 사용되는 물질을 Reb W의 효소적 전환을 사용하여 제조하고 HPLC에 의해 정제하였다. 표준 펄스 시퀀스를 사용하여 Agilent VNMR 500 MHz instrument 기기 상에서 NMR 스펙트럼을 입수하였다. 1D(1H 및 ^{13}C) 및 2D(TOCSY, ASAPHMQC, GCOSY 및 GHMBC) NMR 스펙트럼을 CD3OD 중에 수행하였다.

[0200] m/z 827.3671에서 $[M+Na]^+$ 에 대응하는 부가물 이온을 나타낸 포지티브 고해상(HR) 질량 분광측정에 기반하여 Reb WB2의 분자식은 $C_{38}H_{60}O_{18}$ 로 추정되었다; 이 조성은 NMR 스펙트럼 데이터에 의해 뒷받침되었다.

[0201] Reb WB2의 NMR 스펙트럼 데이터는 엘트-카우란 디테르페노이드의 기본 골격을 드러내었고, GHMBC, COSY 및

TOCSY 실험에 의해 추가 확인되었다. APT 평가를 사용해서 탄소 다중도를 확인하였다. ^{13}C NMR은 3개의 아노머 탄소(δ 102.8, 101.7 및 92.46)뿐만 아니라 δ 62.2, 61.14 및 60.88에서 3개의 $-\text{CH}_2\text{OH}$ 신호를 나타내어, 3개의 당 단위를 확인시켜 주었다. 또한, δ 177.1에서 하나의 카보닐 공명 및 δ 152.2 및 104.4에서 2개의 알켄 탄소가 존재하였다. H21에서 C19로의 GHMBC 상관관계는 디테르페노이드 코어 구조에 대한 당의 부착점을 확인시켜 주었다. δ 79.4에서 C13의 화학적 이동은 상기 탄소에 부착된 산소를 시사한다. Reb WB2에 대한 ^1H 및 ^{13}C NMR 값을 TOCSY, HMQC 및 HMBC 데이터에 기반하여 할당하였고 표 3에 제공한다.

표 3

[0202] Reb WB2^{a-c}에 대한 ^1H 및 ^{13}C NMR 스펙트럼 데이터(화학적 이동 및 커플링 상수). ^a TOCSY, ASAPHMQC 및 GHMBC 상관관계에 기반하여 할당됨; ^b 화학적 이동 값 δ (ppm); ^c 커플링 상수 Hz.

C#	C13 δ	^1H δ (ppm), multp, J(Hz)
1	40.3	1.89m / 0.85 dt, $J_{12.8}$, 4.0
2	19.0	1.96 m / 1.43 m
3	37.2	2.39 bd, $J_{13.7}$ / 1.02 dd, $J_{13.6}$, 4.2
4	43.8	—
5	57.2	1.09 dd, $J_{10.6}$, 3.8
6	21.4	1.87 m
7	41.3	1.56 m / 1.45 m
8	41.3	—
9	53.9	0.99 m
10	39.2	—
11	39.2	1.76 m / 1.47 m
12	19.9	1.73 m / 1.60 m
13	79.4	—
14	46.2	2.05 s / 1.30 m
15	47.4	2.18 m / 2.08 m
16	155.5	—
17	102.1	4.95 s / 4.78 s
18	28.0	1.24 s
19	176.1	—
20	15.5	0.92 s
21	92.5	5.58 d, $J_{8.1}$
22	77.0	3.43 m
23	68.2	3.56 m
24	86.6	3.86 m
25	75.7	4.03 t, $J_{8.5}$
26	60.9	3.83 dd, $J_{12.7}$, 2.7 / 3.72 dd, $J_{12.1}$, 4.6
27	102.8	4.69 d, $J_{7.8}$
28	70.2	3.30 m
29	74.0	3.28 m
30	76.6	3.37 m
31	76.8	3.38 m
32	61.1	3.91 m / 3.64 m
33	101.7	5.01 d, $J_{7.9}$
34	71.2	3.18 m
35	76.7	3.36 m
36	76.7	3.33 m
37	74.4	3.12 m
38	62.1	3.89 m / 3.67 m

[0203] H33과 C25, H27과 C24 간 핵심 GHMBC 상관관계는 3개 당 분자의 연결성을 확인시켜 주었다. 모든 관찰된 2D 상관관계 및 화학적 이동 신호에 기반하여, Reb WB2의 구조는 도 15에 나타난 바와 같았다. Reb WB2의 구조는 13-

하이드록시-엔트-카우르-16-엔-19-오산-[(2-O-β-D-글루코피라노실-3-O-β-D-글루코피라노실-β-D-글루코피라노실)에스테르]로 추정되었다.

[0204] 실시예 7: NMR에 의해 분석된 Reb WB1의 구조.

[0205] Reb WB1의 특성규명을 위해 사용된 물질을 Reb W의 효소적 전환을 사용하여 제조하고 HPLC에 의해 정제하였다. 표준 펄스 시퀀스를 사용하여 Agilent VNMRS 500 MHz instrument 기기 상에서 NMR 스펙트럼을 입수하였다. 1D(¹H 및 ¹³C) 및 2D(TOCSY, ASAPHMQC, GCOSY 및 GHMBC) NMR 스펙트럼을 80% CD3OD - 20% D2O 중에 수행하였다.

[0206] m/z 989.4206에서 [M+Na]⁺에 대응하는 부가물 이온을 나타낸 포지티브 고해상(HR) 질량 분광측정에 기반하여 Reb WB1의 분자식은 C₄₄H₇₀O₂₃으로 추정되었다; 이 조성은 NMR 스펙트럼 데이터에 의해 뒷받침되었다.

[0207] Reb WB1의 NMR 스펙트럼 데이터는 엔트-카우란 디테르페노이드의 기본 골격을 드러내었고, GHMBC, COSY 및 TOCSY 실험에 의해 추가 확인되었다. APT 평가를 사용해서 탄소 다중도를 확인하였다. ¹³C NMR은 4개의 아노머 탄소(δ 102.6, 101.6 97.6 및 92.61뿐만 아니라 δ 62.0, 61.1, 61.0 및 60.8에서 4개의 -CH2OH 신호를 나타내어, 4개의 당 단위를 확인시켜 주었다. 또한, δ 177.0에서 하나의 카보닐 공명 및 δ 152.5 및 104.4에서 2개의 알켄 탄소가 존재하였다. H21에서 C19로의 및 H39에서 C13으로의 GHMBC 상관관계는 디테르페노이드 코어 구조에 대한 당의 부착점을 확인시켜 주었다. Reb WB1에 대한 ¹H 및 ¹³C NMR 값을 TOCSY, HMQC 및 HMBC 데이터에 기반하여 할당하였고 표 4에 제공한다.

표 4

[0208] Reb WB1^{a-c}에 대한 ¹H 및 ¹³C NMR 스펙트럼 데이터(화학적 이동 및 커플링 상수). ^a TOCSY, ASAPHMQC 및 GHMBC 상관관계에 기반하여 할당됨; ^b 화학적 이동 값 δ(ppm); ^c 커플링 상수 Hz.

C#	C13 δ	¹ H δ(ppm), multp, J(Hz)
1	40.2	1.86 m / 0.83 m
2	19.0	1.93 m / 1.41 m
3	37.1	2.35 d, J 13.8 / 1.00 m
4	44.0	-
5	51.1	1.09 m
6	21.5	1.89 m
7	41.1	1.57 m / 1.44 m
8	41.7	-
9	53.7	0.97 m
10	39.2	-
11	19.9	1.81 m / 1.62 m
12	37.3	1.97 m / 1.50 m
13	86.7	-
14	43.9	2.18 m / 1.54 m
15	47.6	2.15 m / 2.05 m
16	152.5	-
17	104.4	5.17 s / 4.88 s
18	28.1	1.23 s
19	177.0	-
20	15.5	0.91 s
21	92.6	5.55d, J7.9
22	85.9	3.72 m
23	68.2	3.57 m
24	85.9	3.91 m
25	75.9	4.05 m
26	61.1	3.90 m / 3.72 dd, J12.3, 4.7
27	102.6	4.72 d, J8.0
28	76.4	3.42 m
29	76.6	3.45 m
30	76.6	3.48 m

31	70.0	3.34 m
32	61.9	3.90 m / 3.64 m
33	101.6	4.96 m
34	74.2	3.16 m
35	76.5	3.43 m
36	76.1	3.23 m
37	73.8	3.30 m
38	60.8	3.83 m / 3.65 m
39	97.6	4.54 d, <i>J</i> 7.8
40	76.8	3.50 m
41	73.8	3.31 m
42	76.4	3.40 m
43	71.0	3.21 m
44	61.0	3.79 m / 3.68 m

[0209] H33과 C25, H27과 C24(및 반대) 간 다른 핵심 GHMBC 상관관계는 3개 당 분자의 연결을 확인시켜 주었다. 모든 관찰된 2D 상관관계 및 화학적 이동 신호에 기반하여, Reb WB1의 구조는 도 16에 나타낸 바와 같았다. Reb WB1의 구조는 13-β-D-글루코피라노실옥시-엔트-카우르-16-엔-19-오산-[(2-O-β-D-글루코피라노실-3-O-β-D-글루코피라노실-3-O-β-D-글루코피라노실-β-D-글루코피라노실)에스테르로 추정되었다.

[0210] **실시예 8: NMR에 의해 분석된 Reb D4의 구조.**

[0211] Reb D4의 특성규명을 위해 사용된 물질을 Reb WB1의 효소적 전환을 사용하여 제조하고 HPLC에 의해 정제하였다. 표준 펄스 시퀀스를 사용하여 Agilent VNMR5 500 MHz instrument 기기 상에서 NMR 스펙트럼을 입수하였다. 1D(¹H 및 ¹³C) 및 2D(TOCSY, ASAPHMQC, GCOSY 및 GHMBC) NMR 스펙트럼을 80% CD3OD 및 20% D2O 중에 수행하였다.

[0212] *m/z* 1151.4728에서 [M+Na]⁺에 대응하는 부가물 이온을 나타낸 포지티브 고해상(HR) 질량 분광측정에 기반하여 Reb D4의 분자식은 C₅₀H₈₀O₂₈로 추정되었다; 이 조성은 NMR 스펙트럼 데이터에 의해 뒷받침되었다.

[0213] Reb D4의 ¹H NMR 스펙트럼 데이터는 δ 1.24 및 0.92에서 2개의 메틸 단일선, 엑소사이클릭 이중 결합의 δ 5.20 및 4.86에서의 단일선으로서의 2개의 올레핀계 양성자의 존재를 나타내었다. 엔트-카우란 디테르페노이드의 기본 골격은 GHMBC, COSY 및 TOCSY 실험에 의해 뒷받침되었다. APT 평가를 사용해서 탄소 다중도를 확인하였다. 또한, ¹³C NMR은 5개의 아노머 탄소(δ 103.5, 102.5, 101.8, 95.6 및 92.8)를 나타내어 5개의 당 단위를, δ 177.1에서 하나의 카보닐 및 δ 152.2 및 104.4에서 2개의 알켄 탄소를 확인시켜 주었다. H40에서 C12로의 및 H22에서 C19로의 GHMBC 상관관계는 디테르페노이드 코어 구조에 대한 당의 부착점을 확인시켜 주었다. Reb D4에 대한 ¹H 및 ¹³C NMR 값을 TOCSY, HMQC 및 HMBC 데이터에 기반하여 할당하였고 표 5에 제공한다.

표 5

[0214] Reb D4^{a,c}에 대한 ¹H 및 ¹³C NMR 스펙트럼 데이터(화학적 이동 및 커플링 상수). ^a TOCSY, ASAPHMQC 및 GHMBC 상관관계에 기반하여 할당됨; ^b 화학적 이동 값 δ(ppm); ^c 커플링 상수 Hz.

C#	C13 δ	1H δ(ppm), multp, J(Hz)
1	19.0	1.94, m/ 1.44, m
2	37.0	2.31, m/ 1.04, m
3	43.9	-
4	57.0	1.07, m
5	39.2	-
6	40.0	1.83, m/ 0.84, m
7	21.9	1.89, m
8	41.3	1.56, m/ 1.40, m
9	41.3	-
10	53.6	0.97, m

11	43.9	2.22, m/ 1.50, m
12	87.2	-
13	37.0	1.99, m/ 1.53, m
14	19.7	1.62, m
15	46.7	2.14, m/ 2.04, m
16	152.2	-
17	104.4	5.20, s/ 4.86, s
18	27.9	1.24, s
19	177.1	-
22	92.8	5.57, d, $J_{7.6\text{Hz}}$
24	76.7	3.92, m
25	63.4	3.53, m
26	85.9	4.05, m
27	75.9	4.01, m
28	101.8	4.99, d, $J_{7.9\text{Hz}}$
30	70.1	3.32, m
31	71.1	3.21, m
32	69.8	3.39, m
33	70.6	3.18, m
34	102.5	4.76, d, $J_{7.9\text{Hz}}$
36	76.5	3.72, m
37	74.6	3.25, m
38	76.2	3.66, m
39	73.8	3.34, m
40	103.5	4.64, m
42	76.9	3.69, m
43	76.4	3.65, m
44	76.6	3.42, m
45	76.5	3.45, m
46	95.6	4.64, m
48	76.2	3.59, m
49	74.4	3.30, m
50	76.6	3.68, m
51	80.5	3.48, m
59	60.9	3.78, m
61	61.0	3.86, m
66	62.0	3.91, m
74	61.8	3.84, m
76	61.0	3.67, m
78	15.9	0.92, s

[0215] H45와 C46, H28과 C27 및 H26과 C34(및 반대) 간 다른 핵심 GHMBC 상관관계는 5개 당 분자의 연결을 확인시켜 주었다. 모든 관찰된 2D 상관관계 및 화학적 이동 신호에 기반하여, Reb D4의 구조는 도 17에 나타낸 바와 같았다. Reb D4의 구조는 13-[(2-O-β-D-글루코피라노실-β-D-글루코피라노실)옥시]엔트-카우르-16-엔-19-오산-[(2-O-β-D-글루코피라노실-3-O-β-D-글루코피라노실-β-D-글루코피라노실)에스테르]로 추정되었다.

[0216] 실시예 9: 레바우디오시드의 맛 평가

[0217] 수크로스를 대조군으로 사용하여 레바우디오시드의 감각 평가를 수행하였다. 수크로스 샘플을 Sigma-Aldrich로부터 구매하고 사용하여 실온에서 병에 넣은 수증 1.0%, 3.0% 및 6.0%(w/v) 수크로스의 3개의 상이한 농도로 대조군 샘플을 제조하였다. 1000 ml의 병에 넣은 수증으로 해당 물질을 첨가하여 감각 평가를 위해 300 ppm에서 레바우디오시드를 제조하였다. 혼합물을 실온에서 교반한 후 스테비올 글리코시드 샘플을 13명의 자원자 인간 대상체 패널에 의해 몇몇 대조군 수크로스 샘플 대비 1.0%, 3.0% 및 6.0%에서 평가하였다. 감각 평가 결과를 표 6에 나타낸다.

표 6

[0218]

수크로스 대비 Reb W1, Reb W2 및 Reb D4의 감각 평가

레바우디오시드	수크로스 대비 단맛 증가 배율
Reb WB1	133
Reb WB2	79
Reb D4	109

[0219]

산업적 적용 가능성/기술 분야의 진술

[0220]

본 개시는 식품, 사료, 음료 및 약리 산업에서 적용 가능성을 갖는다. 본 개시는 일반적으로 변형된 미생물 균주를 통한 스테비올 글리코시드의 생합성 제조 방법에 관한 것이다.

[0221]

참고문헌으로 인용되고 포함되는 문헌:

[0222]

1. Brandle, J. E. et al., (1998). *Stevia Rebaudiana: Its Agricultural, Biological, and Chemical Properties*, Canadian J. Plant Science. 78 (4): 527-36.

[0223]

2. Ceunen, S., and J.M. C. Geuns, *Steviol Glycosides: Chemical Diversity, Metabolism, and Function*, J. Nat. Prod., 2013, 76 (6), pp 1201-28 (2013).

[0224]

3. Du J et al., (2011), *Engineering microbial factories for synthesis of value-added products*, J Ind Microbiol Biotechnol. 38: 873-90.

[0225]

4. GRAS Notices, USA Food and Drug Administration, United States Health & Human Services. (2016) (relevant to steviol glycosides & polyglycosides).

[0226]

5. Häusler A, and Münch T., (1997), *Microbial production of natural flavors*, ASM News 63:551-59.

[0227]

6. Prakash I., et al.; *Isolation and Characterization of a Novel Rebaudioside M Isomer from a Bioconversion Reaction of Rebaudioside A and NMR Comparison Studies of Rebaudioside M Isolated from Stevia rebaudiana Bertoni and Stevia rebaudiana Morita*, Biomolecules, 2014 Jun; 4(2): 374-89. (Published online 2014 Mar 31. 2014).

[0228]

7. Prakash I., et al., *Development of Next Generation Stevia Sweetener: Rebaudioside M*, Foods, 2014, 3:162-175.

[0229]

8. Shockey JM. Et a., (2003), *Arabidopsis contains a large superfamily of acyl-activating enzymes: phylogenetic and biochemical analysis reveals a new class of acyl-coenzyme A synthetases*. Plant Physiol 132 1065-76.

[0230]

관심 서열:

[0231]

UGT76G1 서열:

[0232]

아미노산 서열:(SEQ ID NO: 1)

MENKTETTIVRRRRRIILFPVPFQGHINPILQLANVLYSKGFSITIFHTNFNPKPKTSNYPHF
TFRFILDNDPQDERISNLPTHGPLAGMRIPIINEHGADELRRLELLMLASEEDEEVSLI
TDALWYFAQSVADSLNLRRLVLMTSSLFNFHAHVSLPQFDELGYLDPDDKTRLEEQA
SGFPMKLVKDIKSAYSNWQILKEILGKMIKQTKASSGVIWNSFKELESELETVIREIPA
PSFLIPLPKHLTASSSSLLDHDRTVFQWLDQPPSSVLYVSFGSTSEVDEKDFLEIARGL
VDSKQSFLWVVRPGFVKGSTWVEPLPDGFLGERGRIVKWVPQQEVLAHGAIGAFWT
HSGWNSTLESVCEGVPMIFSDFGLDQPLNARYMSDVLKVGVYLENGWERGEIANAIR
RVMVDEEGEYIRQNARVLKQKADVSLMKGGSSYESLESLSYISSL

[0233]

[0234] DNA 서열: (SEQ ID NO: 2)

ATGGAGAATAAGACAGAAACAACCGTAAGACGGAGGCGGAGGATTATCTTGTTCC
CTGTACCATTTTCAGGGCCATATTAATCCGATCCTCCAATTAGCAAACGTCCTCTAC
TCCAAGGGATTTTCAATAACAATCTTCCATACTAACTTTAACAAGCCTAAAACGAG
TAATTATCCTCACTTTACATTCAGGTTTATTCTAGACAACGACCCTCAGGATGAGC
GTATCTCAAATTTACCTACGCATGGCCCCCTTGGCAGGTATGCGAATACCAATAATC
AATGAGCATGGAGCCGATGAACTCCGTCGCGAGTTAGAGCTTCTCATGCTCGCAA
GTGAGGAAGACGAGGAAGTTTCGTGCCTAATAACTGATGCGCTTTGGTACTTCGCC
CAATCAGTCGCAGACTCACTGAATCTACGCCGTTTGGTCCTTATGACAAGTTCATT
ATTCAACTTTTCACGCACATGTATCACTGCCGCAATTTGACGAGTTGGGTTACCTGG
ACCCGGATGACAAAACGCGATTGGAGGAACAAGCGTCGGGCTTCCCCATGCTGAA
AGTCAAAGATATTAAGAGCGCTTATAGTAATTGGCAAATTCTGAAAGAAATTCTC
GGAAAAATGATAAAGCAAACCAAAGCGTCCTCTGGAGTAATCTGGAACCTCCTTCA
AGGAGTTAGAGGAATCTGAACTGAAACGGTCATCAGAGAAATCCCCGCTCCCTC
GTTCTTAATTCCACTACCCAAGCACCTTACTGCAAGTAGCAGTTCCCTCCTAGATC
ATGACCGAACCCTGTTTCAGTGGCTGGATCAGCAACCCCCGTCGTCAGTTCTATAT
GTAAGCTTTGGGAGTACTTCGGAAGTGGATGAAAAGGACTTCTTAGAGATTGCGC
GAGGGCTCGTGGATAGCAAACAGAGCTTCCCTGTGGGTAGTGAGACCGGGATTCTGT
TAAGGGCTCGACGTGGGTCGAGCCGTTGCCAGATGGTTTTCTAGGGGAGAGAGGG
AGAATCGTGAAATGGGTTCCACAGCAAGAGGTTTTGGCTCACGGAGCTATAGGGG

[0235]

CCTTTTGGACCCACTCTGGTTGGAATTCTACTCTTGAAAGTGTCTGTGAAGGCGTTC
CAATGATATTTTCTGATTTTGGGCTTGACCAGCCTCTAAACGCTCGCTATATGTCTG
ATGTGTTGAAGGTTGGCGTGTACCTGGAGAATGGTTGGGAAAGGGGGGAAATTGC
CAACGCCATACGCCGGGTAATGGTGGACGAGGAAGGTGAGTACATACGTCAGAAC
GCTCGGGTTTTAAACAAAAAGCGGACGTCAGCCTTATGAAGGGAGGTAGCTCCT
ATGAATCCCTAGAAATCCTTGGAAGCTATATATCTTCGTTATAA

[0236]

[0237] CP1 서열:

[0238] 아미노산: (SEQ ID NO:3)

MNWQILKEILGKMIKQTKASSGVIWNSFKELEESELETVIREIPAPSFLIPLPKHLTASSS
SLLDHDRVTFQWLDQPPSSVLYVSFGSTSEVDEKDFLEIARGLVDSKQSFLWVVRPG
FVKGSTWVEPLPDGFLGERGRIVKWVPQQEVLHGAIGAFWTHSGWNSTLESVCEGV
PMIFSDFGLDQPLNARYMSDVLKVG VYLENGWERGEIANAIRRMVDEEGEYIRQNA
RVLKQKADVSLMKGGSSYESLESLSYISSLENKTETT VRRRRRIILFPVPFQGHINPIL
QLANVLYSKGFSITIFHTNFNPKTSNYPHFTFRFILDNDPQDERISNLPHTGPLAGMRI
PIINEHGADELRRLELLMLASEEDEEVSLITDALWYFAQSVADSLNLRRLVLMTSSL
FNFHAHVSLPQFDELGYLDPDDKTRLEEQASGF PMLKV KDIKSAYS

[0239]

[0240] DNA 서열: (SEQ ID NO:4)

ATGAACTGGCAAATCCTGAAAGAAATCCTGGGTAAAATGATCAAACAAACCAAAG
CGTCGTCGGGCGTTATCTGGAACCTCTTCAAAGAACTGGAAGAATCAGAACTGGA
AACCGTTATTCGCGAAATCCCGGCTCCGTCGTTCTGATTCCGCTGCCGAAACATC
TGACCGCGAGCAGCAGCAGCCTGCTGGATCACGACCGTACGGTCTTTCAGTGGCT
GGATCAGCAACCGCCGTCATCGGTGCTGTATGTTTCATTTCGGTAGCACCTCTGAAG
TCGATGAAAAAGACTTTTCTGGAATCGCTCGCGGCCTGGTGGATAGTAAACAGTC
CTTCCTGTGGGTGGTTCGTCCGGGTTTTGTGAAAGGCAGCACGTGGGTGAACCGC
TGCCGGATGGCTTCCTGGGTGAACGCGGCCGTATTGTCAAATGGGTGCCGCAGCA
AGAAGTGCTGGCACATGGTGCTATCGGCGCGTTTTGGACCCACTCTGGTTGGAACA
GTACGCTGGAATCCGTTTGCGAAGGTGTCCCGATGATTTTCAGCGATTTTGGCCTG
GACCAGCCGCTGAATGCCCCTATATGTCTGATGTTCTGAAAGTCGGTGTGTACCT
GGAAAACGGTTGGGAACGTGGCGAAATTGCGAATGCCATCCGTCGCGTTATGGTC
GATGAAGAAGGCGAATACATTGCCAGAACGCTCGTGTCTGAAACAAAAAGCGG

[0241]

ACGTGAGCCTGATGAAAGGCGGTAGCTCTTATGAATCACTGGAATCGCTGGTTAG
CTACATCAGTTCCCTGGAAAATAAAACCGAAACCACGGTGCGTCGCCGTCGCCGT
ATTATCCTGTTCCCGGTTCCGTTTCAGGGTCATATTAACCCGATCCTGCAACTGGC
GAATGTTCTGTATTCAAAGGCTTTTCGATCACCATCTTCCATACGAACTTCAACA
AACCGAAAACAGTAACTACCCGCACTTTACGTTCCGCTTTATTCTGGATAACGAC
CCGCAGGATGAACGTATCTCCAATCTGCCGACCCACGGCCCGCTGGCCGGTATGC
GCATTCCGATTATCAATGAACACGGTGCAGATGAACTGCGCCGTGAACTGGAAC
GCTGATGCTGGCCAGTGAAGAAGATGAAGAAGTGTCTGTCTGATCACCGACGCA
CTGTGGTATTTGCCCCAGAGCGTTGCAGATTCTCTGAACCTGCGCCGTCTGGTCCT
GATGACGTCATCGCTGTTCAATTTTCATGCGCACGTTTCTCTGCCGCAATTTGATGA
ACTGGGCTACCTGGACCCGGATGACAAAACCCGTCTGGAAGAACAAGCCAGTGGT
TTTCCGATGCTGAAAGTCAAAGACATTAAATCCGCCTATTCGTAA

[0242]

[0243] B-glul 서열:

[0244]

아미노산: (SEQ ID NO:5)

MTQLDVESLIQELTLNEKVQLLSGSDFWHTTPVRRLGIPKMRLSDGPNGVRGTKFFNG
VPTACFPCGTGLGATFDKELLKEAGSLMADEAKAKAASVVLGPTANIARGPNNGRGF
ESFGEDPVVNLSSAAMINGLQGGKYAATMKHYVCNDLEMDRNCIDAQVSHRALRE
VYLLPFQIAVRDANPRAIMTAYNKANGEHVSQSKFLLDEVLRKEWGWGGLMSDW
GVYDAKSSITNGLDLEMPGPPQCRVHSATDHAINSGEIHINDVDERVRSLLSLINYCHQ
SGVTEEDPETSNDNTPETIEKLRKISRESIVLLKDDDRNRSILPLKKSCKIAVIGNNAKQ
AA YCGGGSASVLSYHTTTPFDSIKSRLEDSNTPAYTIGADAYKNLPPLGPQMTDSDGK
PGFDAKFFVGSPTSKDRKLIDHFQLTNSQVFLVDYYNEQIPENKEFYVDVEGQFIPEED
GTYNFGLTVFGTGRLFVDDKLVSQSSQNQTPGDSFFGLAAQEVIGSIHLVKGKAYKIK
VLYGSSVTRTYEIAASVAFEGGAFTFGAAKQRNEDEEIARAVEIAKANDKVVLCIGLN
QDFESEGFDRPDIPGATNKMVSAVLKANPNTVIVNQTGTPVEMPWASDAPVILQA
WFGGSEAGTAIADVLFGDYNPSGKLTVTFLRFEDNPAYLNFQSNKQACWYGEDVYV
GYRYYETIDRPVLFPGHGLSFTEFDFTDMFVRLEENLEVEVVVRNTGKYDGAEEVQ
LYVAPVSPSLKRPIKELKEYAKIFLASGEAKTVHLSVPIKYATSFDEYQKKWCSEKGE
YTILLGSSSADIKVSQSITLEKTTFWKGL

[0245]

[0246]

DNA: (SEQ ID NO:6)

ATGACCCAACTGGATGTGGAGAGCCTGATTCAAGAGCTGACCCTGAACGAAAAGG
 TGCAACTGCTGAGCGGTAGCGACTTCTGGCATAACACCCCGGTTTCGTCTGCTGGGC
 ATCCCGAAGATGCGTCTGAGCGACGGTCCGAACGGCGTTTCGTGGTACCAAATTCTT
 TAACGGTGTTCGACCGCGTGCTTCCCGTGCGGTACCGGTCTGGGCGCGACCTTTG
 ACAAGGAACTGCTGAAAGAGGCGGGTAGCCTGATGGCGGATGAAGCGAAAGCGA
 AAGCGGCGAGCGTGGTTCTGGGTCCGACCGCGAACATTGCGCGTGGTCCGAACGG
 TGGCCGTGGCTTCGAGAGCTTCGGCGAGGACCCGGTGGTTAACGGTCTGAGCAGC
 GCGGCGATGATCAACGGCCTGCAGGGCAAGTACATTGCGGCGACCATGAAACACT
 ATGTTTGCAACGATCTGGAAATGGACCGTAACTGCATTGACGCGCAAGTTAGCCA
 CCGTGCGCTGCGTGAGGTGTACCTGCTGCCGTTCCAAATCGCGGTGCGTGATGCGA
 ACCCGCGTGCGATTATGACCGCGTATAACAAGGCGAACGGCGAACACGTTAGCCA
 GAGCAAATTCCTGCTGGACGAAGTGCTGCGTAAGGAGTGGGGCTGGGATGGTCTG
 CTGATGAGCGACTGGTTTGGTGTTCACGATGCGAAAAGCAGCATCACCAACGGCC
 TGGACCTGGAGATGCCGGGTCCGCCGAGTGCCGTGTGCACAGCGCGACCGATCA
 CGCGATCAACAGCGGCGAAATCCACATTAACGATGTTGACGAGCGTGTGCGTAGC
 CTGCTGAGCCTGATTAATACTACTGCCACCAAAGCGGTGTTACCGAGGAAGATCCGG
 AAACCAGCGACAACAACACCCCGGAAACCATCGAGAAGCTGCGTAAAATCAGCC
 GTGAGAGCATTGTGCTGCTGAAGGACGATGACCGTAACCGTAGCATTCTGCCGCT
 GAAGAAAAGCGACAAAATCGCGGTTATTGGTAACAACGCGAAACAAGCGGCGTA
 TTGCGGTGGCGGTAGCGCGAGCGTGCTGAGCTATCACACCACCACCCCGTTTCGAC
 AGCATCAAGAGCCGTCTGGAAGATAGCAACACCCCGGCGTACACCATTGGTGCGG
 ACGCGTATAAAAACCTGCCGCCGCTGGGTCCGCAAATGACCGATAGCGACGGCAA
 GCCGGGTTTTGATGCGAAATCTTTGTTGGCAGCCCGACCAGCAAGGATCGTAAAC
 TGATCGACCACTTCCAGCTGACCAACAGCCAAGTTTTTCTGGTGGACTACTATAAC
 GAACAGATCCCGGAAAACAAGGAGTTCTACGTTGACGTGGAGGGTCAATTTATTC
 CGGAGGAAGATGGCACCTATAACTTCGGTCTGACCGTGTTTGGTACCGGCCGTCTG
 TTCGTTGATGACAACTGGTTAGCGACAGCAGCCAGAACCAAACCCCGGGCGATA
 GCTTCTTTGGTCTGGCGGCGCAGGAAGTGATCGGCAGCATTACCTGGTGAAGGGT
 AAAGCGTACAAGATCAAAGTTCTGTATGGCAGCAGCGTGACCCGTACCTACGAAA
 TTGCGGCGAGCGTTGCGTTTGAGGGCGGTGCGTTACCTTTGGTGCGGCGAAACAG
 CGTAACGAAGACGAGGAAATCGCGCGTGCGGTGGAGATTGCGAAGGCGAACGAC
 AAAGTGTTCTGTGCATCGGCCTGAACCAAGATTTTCGAAAGCGAGGGTTTTGATCG

[0247]

TCCGGACATCAAGATTCCGGGCGCGACCAACAAAATGGTTAGCGCGGTGCTGAAG
GCGAACCCGAACACCGTTATTGTGAACCAGACCGGTACCCCGGTTGAGATGCCGT
GGGCGAGCGATGCGCCGGTGATCCTGCAAGCGTGGTTTGGCGGTAGCGAGGCGGG
TACCGCGATTGCGGATGTTCTGTTTGGCGACTACAACCCGAGCGGCAAGCTGACC
GTGACCTTCCCGCTGCGTTTTGAGGATAACCCGGCGTACCTGAACTTCCAGAGCAA
CAAACAAGCGTGCTGGTATGGCGAAGACGTTTACGTGGGTTATCGTTACTATGAG
ACCATCGATCGTCCGGTGCTGTTCCCGTTTTGGTCACGGCCTGAGCTTCACCGAGTT
CGATTTTACCGACATGTTTGTTCGTCTGGAGGAAGAGAACCTGGAAGTTGAGGTGG
TTGTGCGTAACACCGGCAAGTACGACGGTGCGGAAGTGGTGCAGCTGTATGTTGC
GCCGGTTAGCCCGAGCCTGAAACGTCCGATCAAGGAACTGAAAGAGTACGCGAAA
ATTTTCCTGGCGAGCGGTGAAGCGAAGACCGTTCACCTGAGCGTGCCGATCAAAT
ACGCGACCAGCTTCTTTGATGAGTATCAAAAGAAATGGTGCAGCGAAAAGGGCGA
GTATACCATTCTGCTGGGTAGCAGCAGCGCGGACATCAAAGTTAGCCAAAGCATC
ACCCTGGAAAAAACCACCTTCTGGAAAGGTCTGTAA

[0248]

[0249]

UGT85C2 서열:

[0250]

아미노산: (SEQ ID NO: 7)

MDAMATTEKKPHVIFIPFAQSHIKAMLKLAQLLHHKGLQITFVNTDFIHNQFLESSGP
HCLDGAPGFRFETIPDGVSHSPEASIPIRESLLRSIETNFLDRFIDLVTCLPDPPTCIISDGF
LSVFTIDAAKKLIPVMYWTAAACGFMGFYHIHSLIEKGFAPLKDASYLTNGYLDTV
IDWVPGMEGIRLKDFPLDWSTDLNDKVLMTTEAPQRSHKVSHHIFHTFDELEPSIIKT
LSLRNYHIYTIGPLQLLLDQIPEEKKQTGITS LHGYSLVKEEPECFQWLQSKEPNSVYVY
NFGSTTVMSLEDMTFGWGLANSNHYFLWIIRSNLVIGENAVLPPELEEHIKKRGFIAS
WCSQEKVLKHPSVGGFLTHCGWGSTIESLSAGVPMICWPYSWDQLTNCRYICKEWEV
GLEMGTQVKRDEVKRLVQELMGEKGHKMRNKAADWKEKARIAIAPNGSSSLNIDKM
VKEITVLARN

[0251]

[0252]

DNA(SEQ ID NO: 8)

ATGGACGCTATGGCCACGACCGAAAAGAAACCGCACGTTATCTTTATTCCGTTCCC
GGCACAGAGTCACATCAAGGCTATGCTGAAGCTGGCCCAACTGCTGCATCACAAA
GGCCTGCAAATTACCTTTGTGAACACGGATTTTCATCCATAATCAGTTTCTGGAAAG
CTCTGGCCCGCACTGCCTGGATGGTGCGCCGGGTTTTCGCTTCGAAACCATCCCCGG

[0253]

ATGGTGTCTCGCATAGCCCGGAAGCCTCTATTCCGATCCGTGAATCGCTGCTGCGC
 AGCATTGAAACCAACTTTCTGGATCGTTTCATCGACCTGGTGACGAAACTGCCGGA
 CCCGCCGACGTGCATTATCTCCGACGGCTTTCTGTCAGTTTTACCATTTGATGCGG
 CCAAAAAGCTGGGTATCCCGGTCATGATGTATTGGACGCTGGCAGCTTGTGGCTTT
 ATGGGTTTTCTACCATATTTCACTCACTGATCGAAAAAGGCTTTGCACCGCTGAAGGA
 TGCTAGTTATCTGACCAACGGCTATCTGGATACGGTCATTGACTGGGTGCCGGGCA
 TGGAAGGTATCCGTCTGAAAGATTTCCCGCTGGACTGGAGCACCGATCTGAATGA
 CAAAGTGCTGATGTTTACCACGGAAGCGCCGCAGCGCTCTATAAAGTTAGTCATC
 ACATTTTTTCACACCTTCGATGAACTGGAACCGTCGATTATCAAAACCCTGAGCCTG
 CGTTATAATCATATTTACACCATTTGGCCCGCTGCAACTGCTGCTGGACCAAATCCC
 GGAAGAAAAGAAACAAACCGGCATCACGTCGCTGCACGGTTATAGCCTGGTGAAA
 GAAGAACCAGGAATGCTTCCAGTGGCTGCAATCTAAGGAACCGAACAGTGTGGTTT
 ACGTGAATTTTGGTTCCACCACGGTTATGTCACTGGAAGATATGACCGAATTTGGC
 TGGGGTCTGGCAAACCTCTAACCATTTTCTGTGGATCATCCGTAGTAACCTGGT
 CATTTGGCGAAAATGCAGTGCTGCCGCCGGAACCTGGAAGAACACATTAAAAAGCGC
 GGTTTCATCGCTTCCTGGTGTTACAGGAAAAAGTTCTGAAGCATCCGTCCGTCCG
 CGTTTTCTGACCCACTGCGGCTGGGGTAGCACGATTGAATCTCTGAGTGCTGGTG
 TTCCGATGATTTGCTGGCCGTATAGCTGGGATCAACTGACCAACTGCCGCTACATC
 TGTAAGAATGGGAAGTCGGCCTGGAAATGGGTACGAAAGTGAAGCGTGACGAA
 GTTAAACGCCTGGTCCAAGAACTGATGGGCGAAGGCGGTCATAAAATGCGTAACA
 AAGCGAAGGATTGGAAAGAAAAGGCCCGCATTGCGATTGCGCCGAACGGCAGCA
 GCAGCCTGAACATTGACAAAATGGTGAAGGAAATCACCGTTCTGGCGCGTAATTA
 A

[0254]

[0255]

HV1 서열:

[0256]

아미노산 서열: (SEQ ID NO: 9)

MDGNSSSSPLHVVICPWLALGHLLPCLDIAERLASRGHRVSFVSTPRNIARLPPLRPV
 APLVDFVALPLPHVDGLPEGAESTNDVPYDKFELHRKAFFDGLAAPFSEFLRAACAEGA
 GSRPDWLIVDTFHHWAAAAAVENKVPVCMMLLGAATVIAGFARGVSEHAAA AVGKE
 RPAAEAPSFETERKLMTTQNASGMTVAERYFLTLMRSDLVAIRSCAEWEPESVAALT
 TLAGKPVVPLGLLPPSPEGGRGVSKEDAAVRWLDAQPAKSVVYVALGSEVPLRAEQV
 HELALGLELSGARFLWALRKPTDAPDAAVLPPGFEERTRGRGLVVTGWVPQIGVLAH
 GAVAAFLTHCGWNSTIEGLLFHGPLIMLPISSDQGPNARLMEGRKVGMQVPRDESDGS
 FRREDVAATVRAVAVEEDGRRVFTANAKKMQEIVADGACHERCIDGFIQQLRSYKA

[0257]

[0258]

[0259]

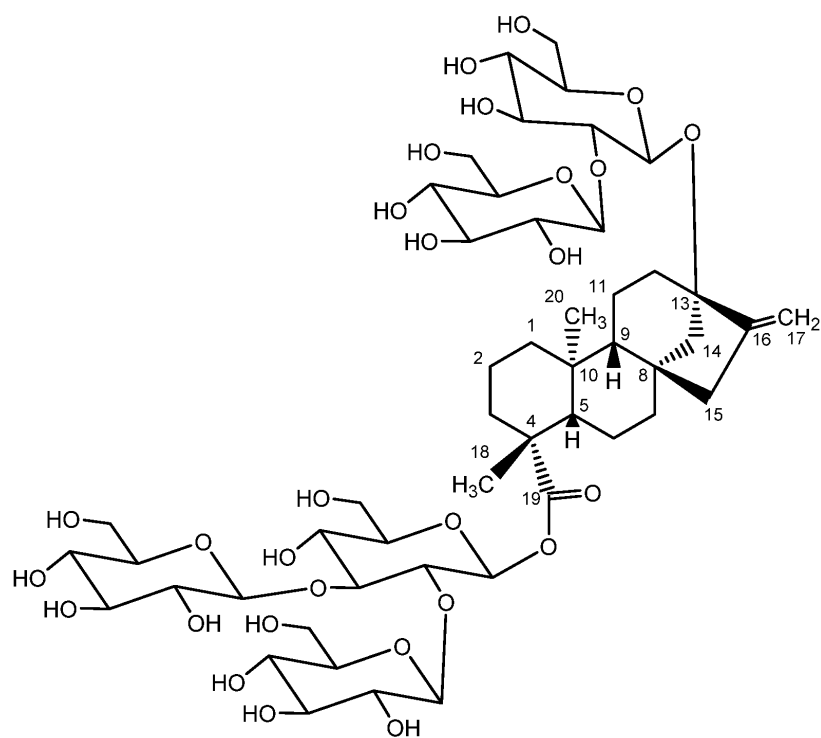
DNA 서열: (SEQ ID NO: 10)

ATGGATGGTAACTCCTCCTCCTCGCCGCTGCATGTGGTCATTTGTCCGTGGCTGGC
TCTGGGTACCTGCTGCCGTGTCTGGATATTGCTGAACGTCTGGCGTCACGCGGCC
ATCGTGTCAAGTTTTGTGTCCACCCCGCGCAACATTGCCCGTCTGCCGCCGCTGCGT
CCGGCTGTTGCACCGCTGGTTGATTTTCGTCGCACTGCCGCTGCCGCATGTTGACGG
TCTGCCGGAGGGTGCAGGAATCGACCAATGATGTGCCGTATGACAAATTTGAACTG
CACCGTAAGGCGTTTCGATGGTCTGGCGGCCCGTTTAGCGAATTTCTGCGTGCAGC
TTGCGCAGAAGGTGCAGGTTCTCGCCCGGATTGGCTGATTGTGGACACCTTTTCATC
ACTGGGCGGCGGCGGCGGCGGTGGAAAACAAAGTGCCGTGTGTTATGCTGCTGCT
GGGTGCAGCAACGGTGATCGCTGGTTTCGCGCGTGGTGTAGCGAACATGCGGCG
GCGGCGGTGGGTAAAGAACGTCCGGCTGCGGAAGCCCCGAGTTTGAACCGAAC
GTCGCAAGCTGATGACCACGCAGAATGCCTCCGGCATGACCGTGGCAGAACGCTA
TTTCCTGACGCTGATGCGTAGCGATCTGGTTGCCATCCGCTCTTGCGCAGAATGGG
AACCGGAAAGCGTGGCAGCACTGACCACGCTGGCAGGTAAACCGGTGGTTCCGCT
GGGTCTGCTGCCGCCGAGTCCGGAAGGCGGTCTGGCGTTTCCAAAGAAGATGCT
GCGGTCCGTTGGCTGGACGCACAGCCGGCAAAGTCAGTCGTGTACGTCGCACTGG
GTTTCGGAAGTGCCGCTGCGTGCAGAACAAAGTTCACGAACTGGCACTGGGCCTGGA
ACTGAGCGGTGCTCGCTTTCTGTGGGCGCTGCGTAAACCGACCGATGCACCGGAC
GCCGCAGTGCTGCCGCCGGGTTTCGAAGAACGTACCCGCGGCCGTGGTCTGGTTGT
CACGGGTTGGGTGCCGCAGATTGGCGTTCTGGCTCATGGTGCGGTGGCTGCGTTTC
TGACCCACTGTGGCTGGAACCTCTACGATCGAAGGCCTGCTGTTTCGGTCATCCGCTG
ATTATGCTGCCGATCAGCTCTGATCAGGGTCCGAATGCGCGCCTGATGGAAGGCC
GTAAAGTCGGTATGCAAGTGCCGCGTGATGAATCAGACGGCTCGTTTCGTCGCGA
AGATGTTGCCGCAACCGTCCGCGCCGTGGCAGTTGAAGAAGACGGTCGTCGCGTC
TTCACGGCTAACGCGAAAAAGATGCAAGAAATTGTGGCCGATGGCGCATGCCACG
AACGTTGTATTGACGGTTTTATCCAGCAACTGCGCAGTTACAAGGCGTGA

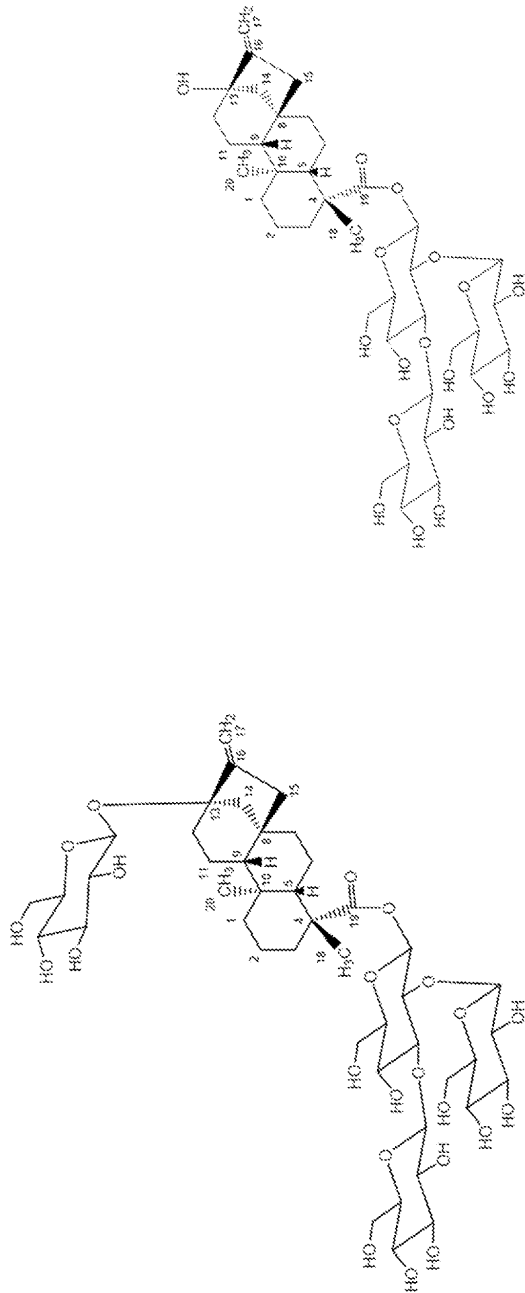
[0260]

도면

도면1



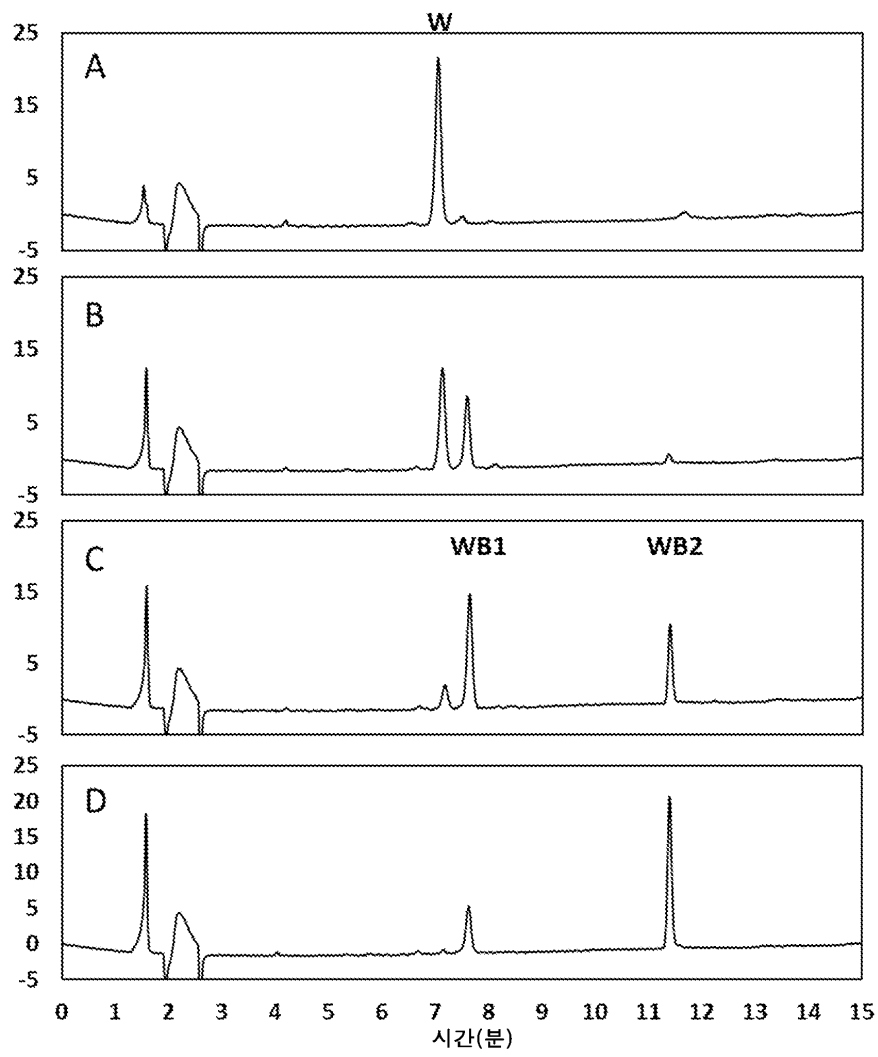
도면2



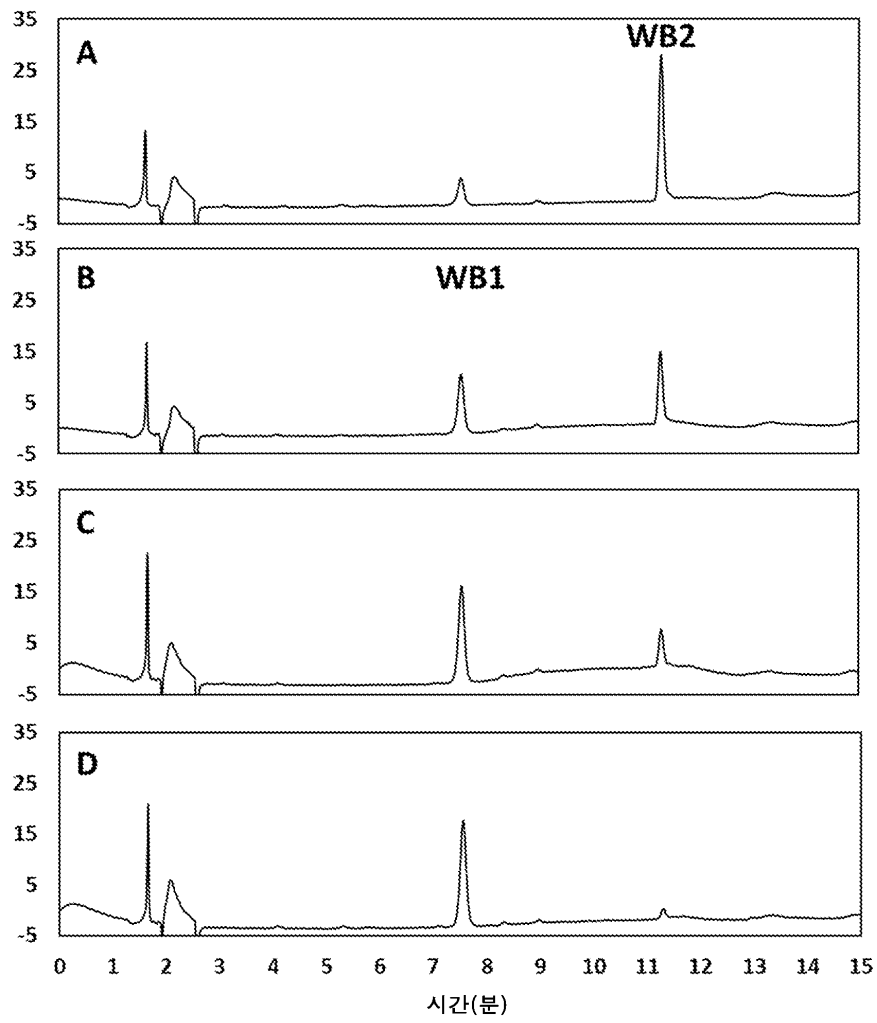
레바우디오시드 WB2

레바우디오시드 WB1

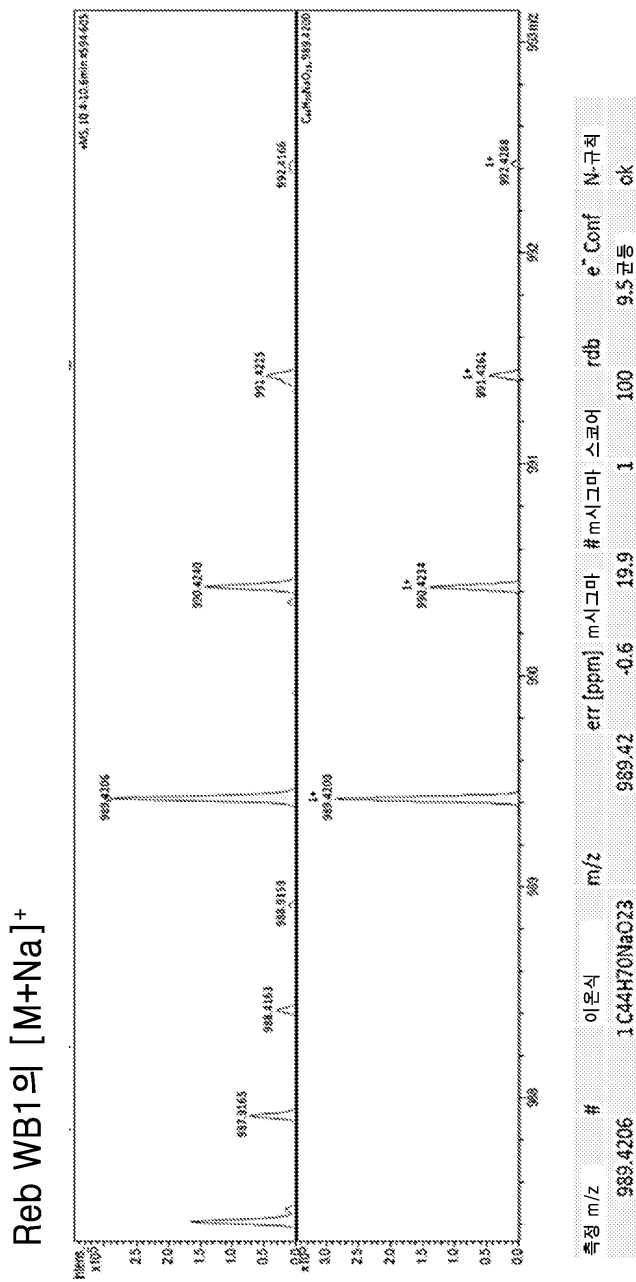
도면3



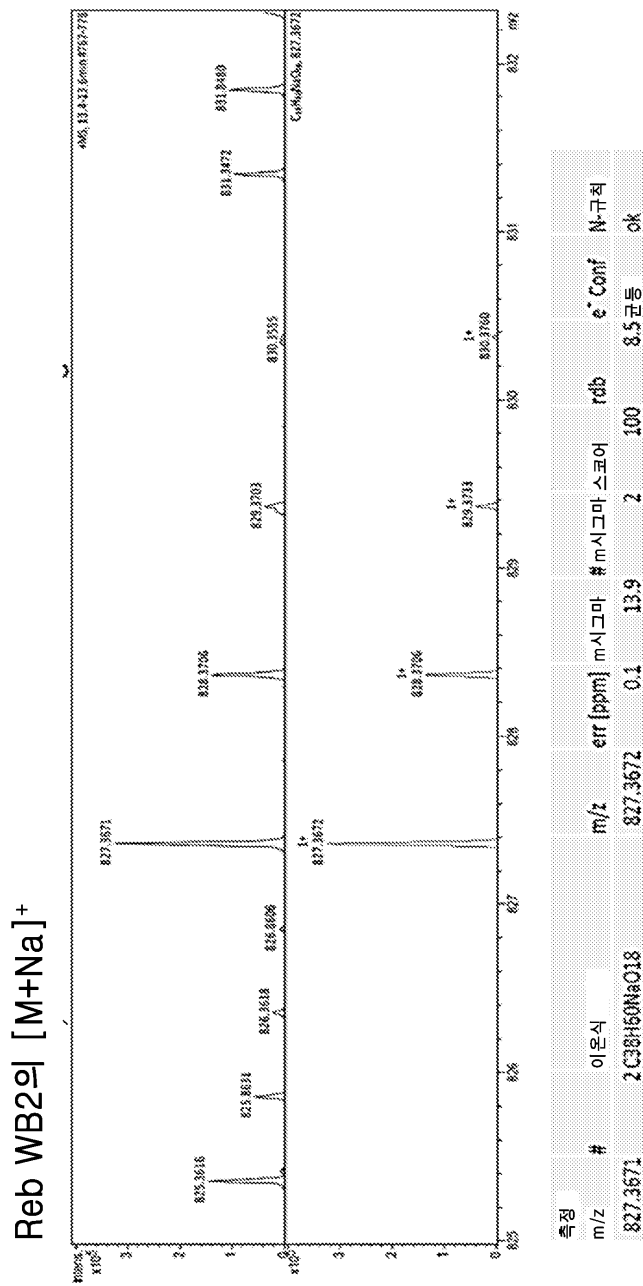
도면4



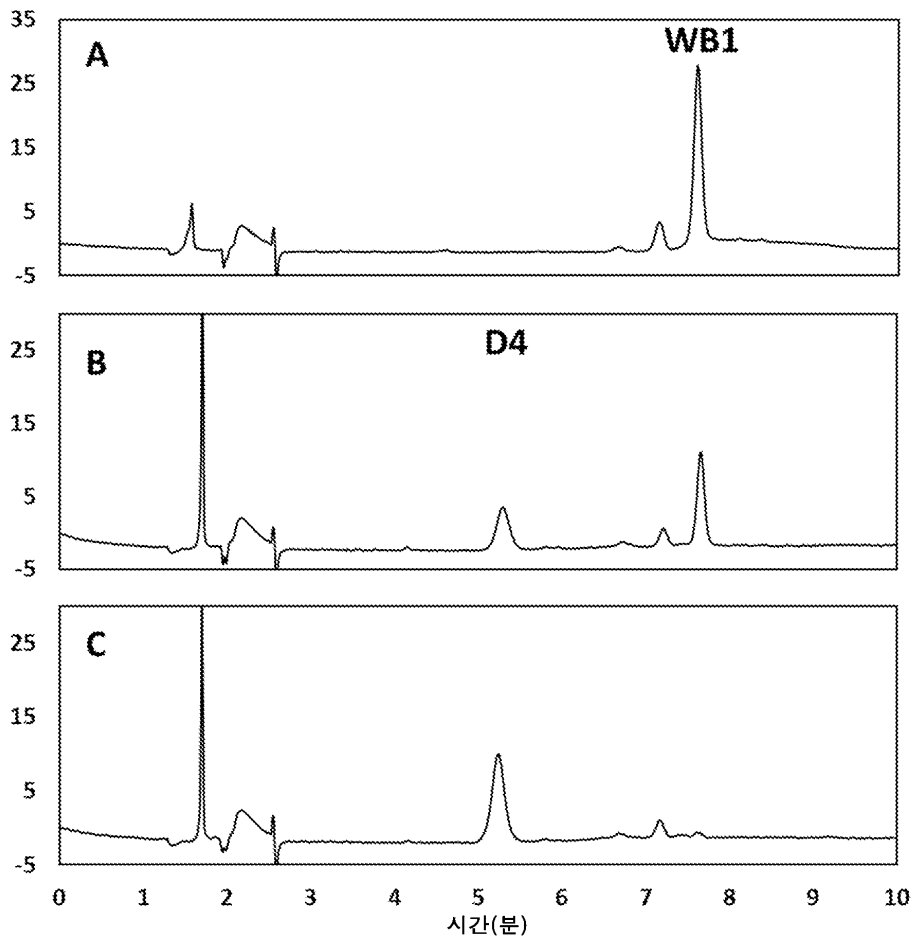
도면5a



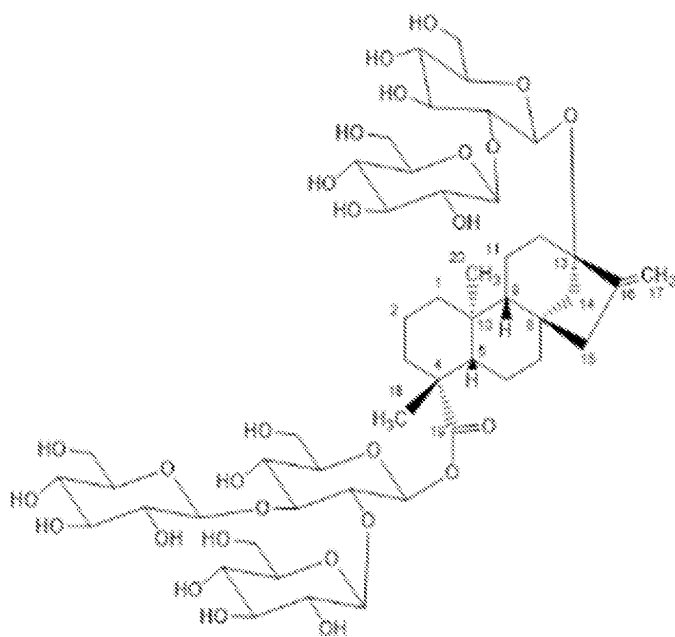
도면5b



도면6

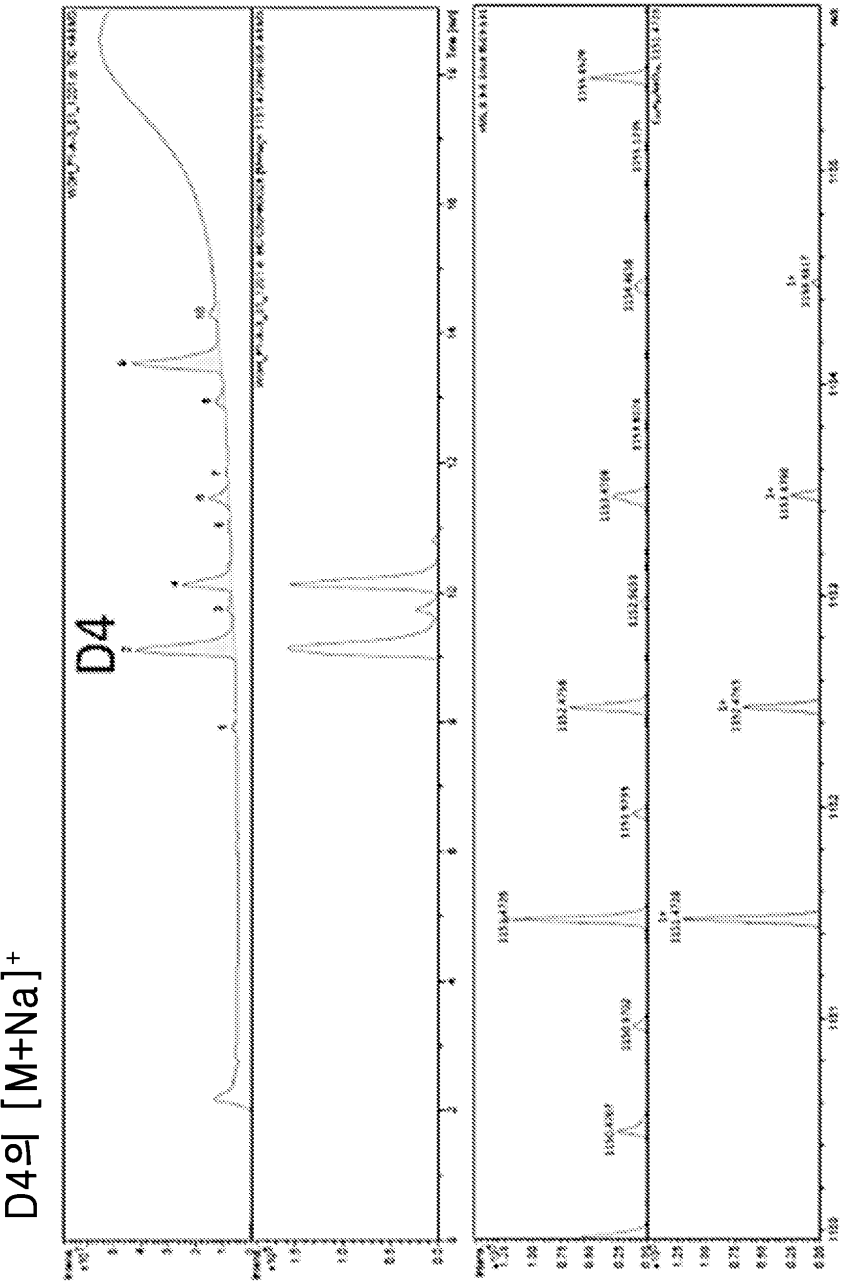


도면7a

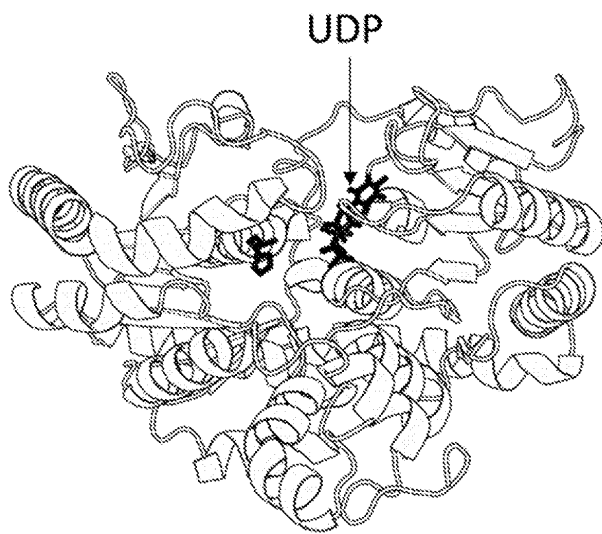


레바우디오시드 D4

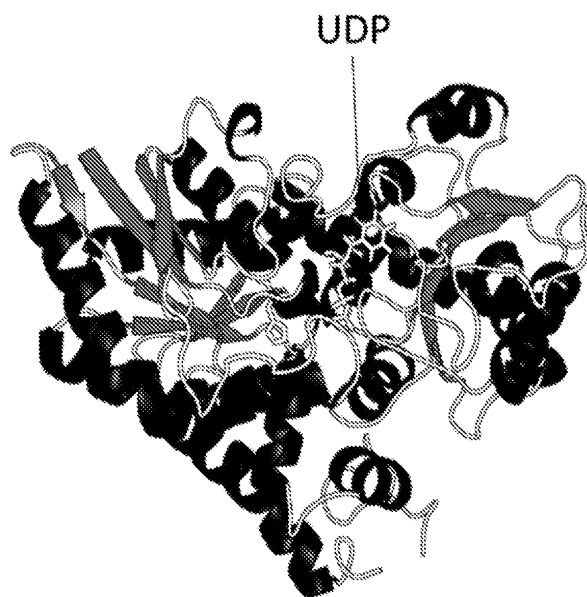
도면7b



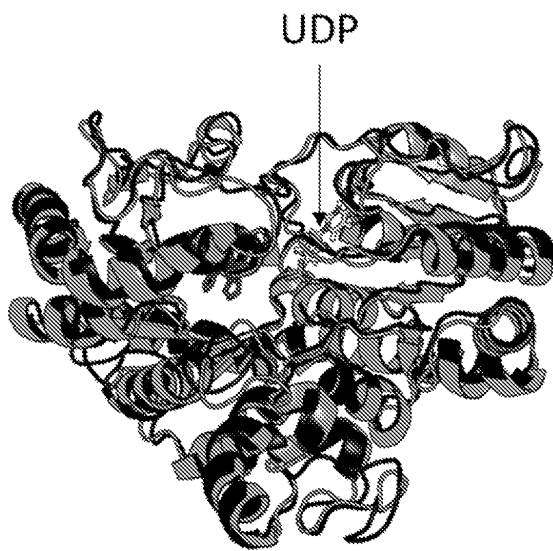
도면8



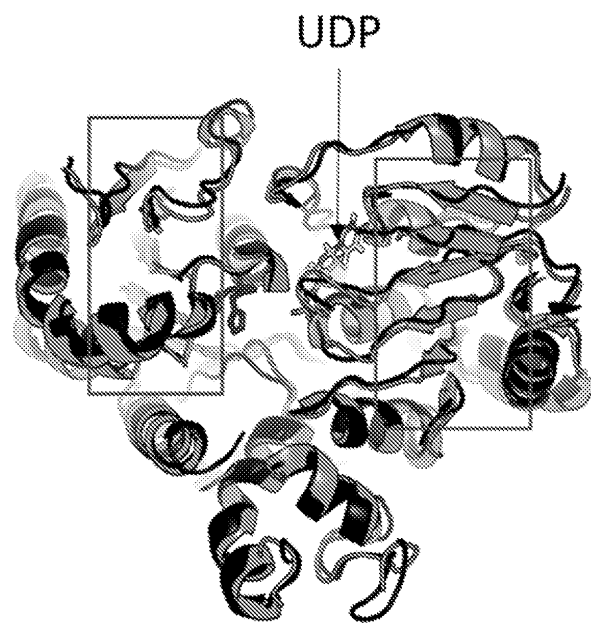
도면9



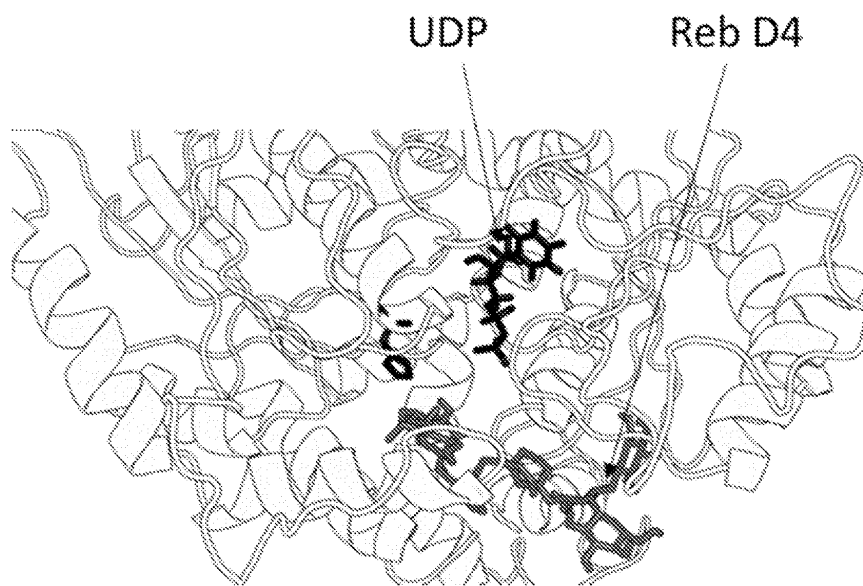
도면10



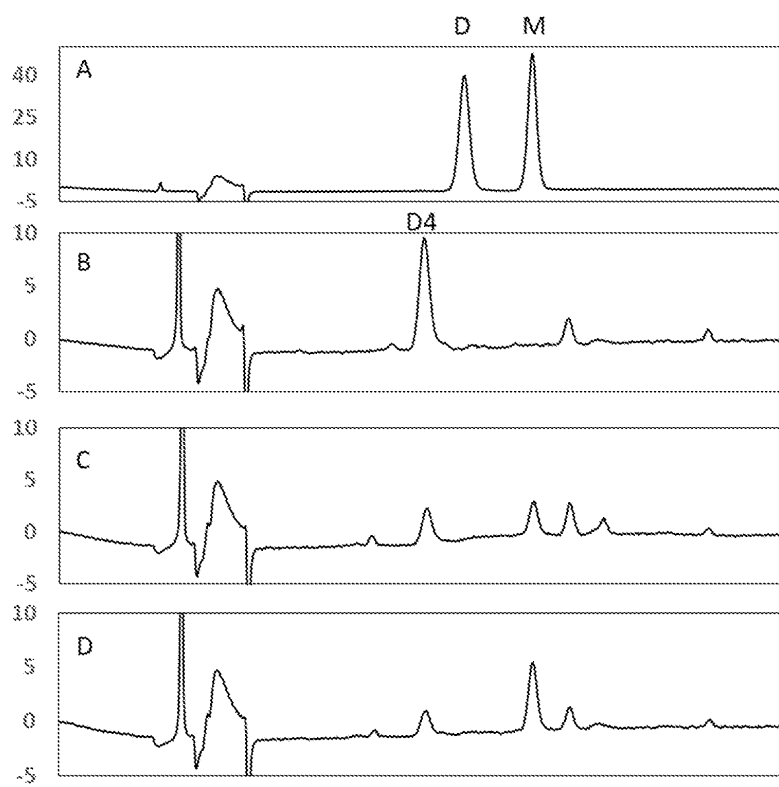
도면11



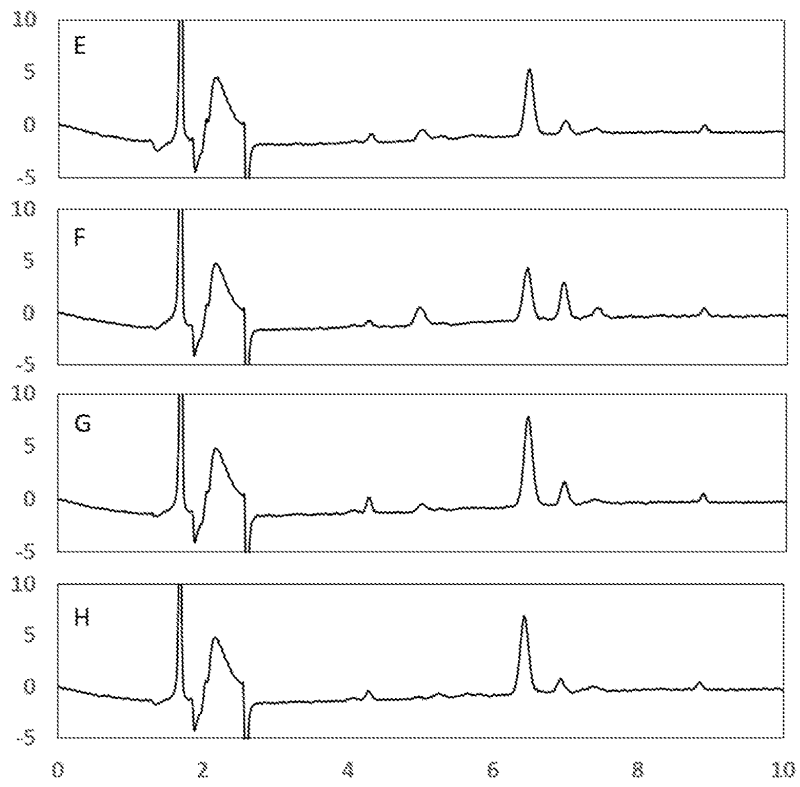
도면12



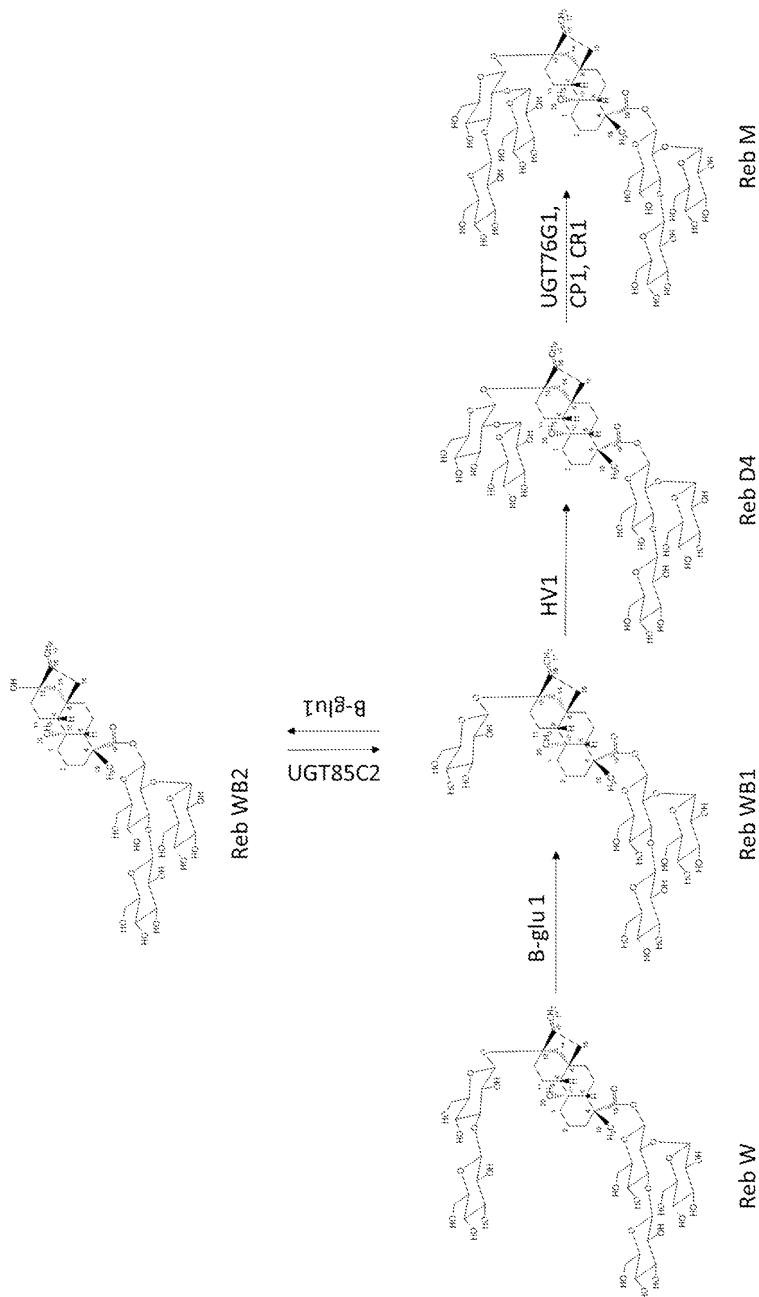
도면13a



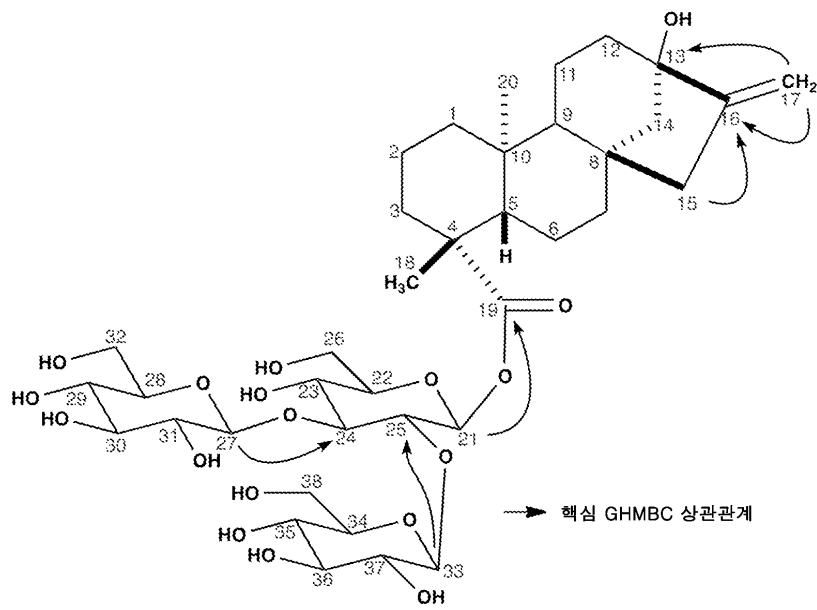
도면13b



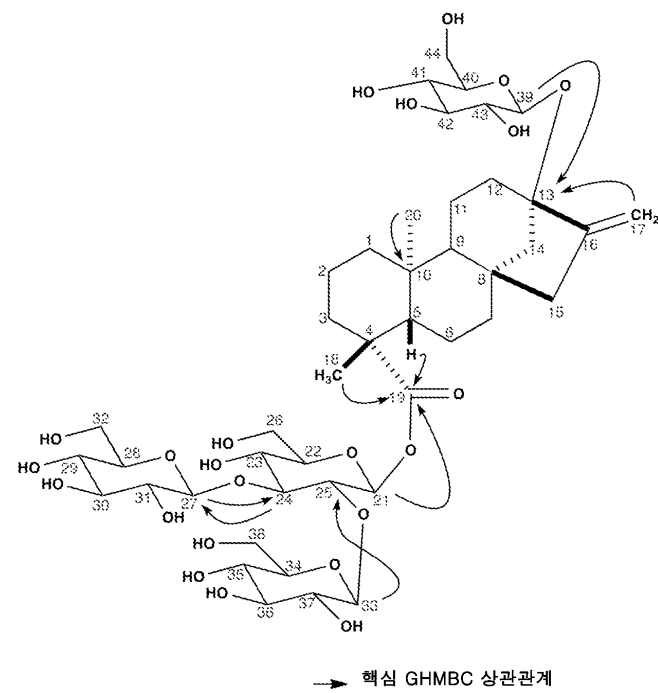
도면14



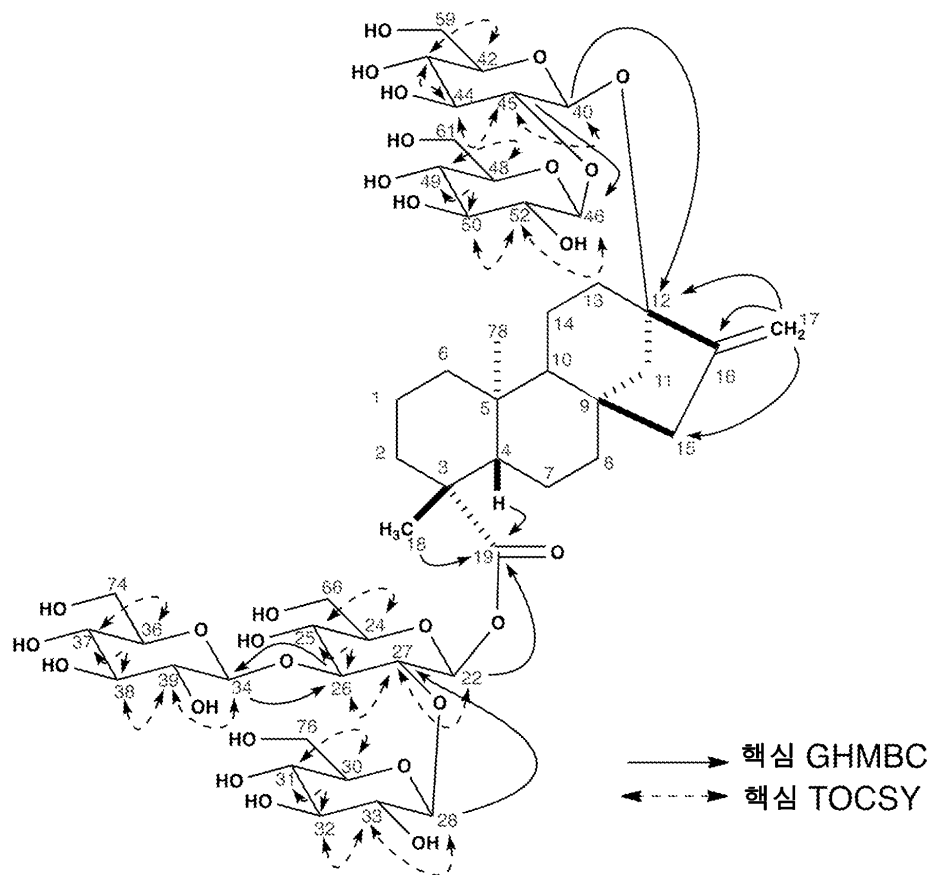
도면15



도면16



도면17



서열 목록

SEQUENCE LISTING

<110> Conagen Inc.

<120> BIOSYNTHETIC PRODUCTION OF STEVIOL GLYCOSIDES AND PROCESSES
THEREFORE

<130> C1497.70023W000

<150> US 62/408,179

<151> 2016-10-14

<150> US 62/555,809

<151> 2017-09-08

<160> 10

<170> PatentIn version 3.5

<210> 1

<211> 458

<212> PRT

<213> S. rebaudiana

<400> 1

Met Glu Asn Lys Thr Glu Thr Thr Val Arg Arg Arg Arg Arg Ile Ile

1 5 10 15

Leu Phe Pro Val Pro Phe Gln Gly His Ile Asn Pro Ile Leu Gln Leu

20 25 30

Ala Asn Val Leu Tyr Ser Lys Gly Phe Ser Ile Thr Ile Phe His Thr

35 40 45

Asn Phe Asn Lys Pro Lys Thr Ser Asn Tyr Pro His Phe Thr Phe Arg

50 55 60

Phe Ile Leu Asp Asn Asp Pro Gln Asp Glu Arg Ile Ser Asn Leu Pro

65 70 75 80

Thr His Gly Pro Leu Ala Gly Met Arg Ile Pro Ile Ile Asn Glu His

85 90 95

Gly Ala Asp Glu Leu Arg Arg Glu Leu Glu Leu Leu Met Leu Ala Ser

100 105 110

Glu Glu Asp Glu Glu Val Ser Cys Leu Ile Thr Asp Ala Leu Trp Tyr

115 120 125

Phe Ala Gln Ser Val Ala Asp Ser Leu Asn Leu Arg Arg Leu Val Leu

130 135 140

Met Thr Ser Ser Leu Phe Asn Phe His Ala His Val Ser Leu Pro Gln

145 150 155 160

Phe Asp Glu Leu Gly Tyr Leu Asp Pro Asp Asp Lys Thr Arg Leu Glu

165 170 175

Glu Gln Ala Ser Gly Phe Pro Met Leu Lys Val Lys Asp Ile Lys Ser

180 185 190

Ala Tyr Ser Asn Trp Gln Ile Leu Lys Glu Ile Leu Gly Lys Met Ile

195 200 205

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

210 215 220

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

225 230 235 240

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

<210> 2

<211> 1377

<212> DNA

<213> S. rebaudiana

<400> 2

atggagaata agacagaaac aaccgtaaga cggaggcgga ggattatctt gttccctgta	60
ccatttcagg gccatattaa tccgatactc caattagcaa acgtcctcta ctccaaggga	120
ttttcaataa caatcttcca tactaacttt aacaagccta aaacagagtaa ttatcctcac	180
tttacattca ggttcattct agacaacgac cctcaggatg agcgtatctc aaatttacct	240
acgcatggcc ccttggcagg tatgcgaata ccaataatca atgagcatgg agccgatgaa	300
ctccgtcgcg agttagagct tctcatgctc gcaagtgagg aagacgagga agtttcgtgc	360
ctaataactg atgcgctttg gtacttcgcc caatcagtcg cagactcact gaatctacgc	420
cgtttgggcc ttatgacaag ttcatatttc aactttcacg cacatgtatc actgccgcaa	480
tttgacgagt tgggttacct ggaccgggat gacaaaacgc gattggagga acaagcgtcg	540
ggcttcccca tgcgaaagt caaagatatt aagagcgctt atagtaattg gcaaattctg	600
aaagaaattc tcggaaaaat gataaagcaa accaaagcgt cctctggagt aatctggaac	660
tccttcaagg agttagagga atctgaactt gaaacggcca tcagagaaat ccccgctccc	720
tcgttcttaa ttccactacc caagcacctt actgcaagta gcagttccct cctagatcat	780
gaccgaaccg tgtttcagtg gctggatcag caacccccgt cgtcagttct atatgtaagc	840
tttgggagta cttcggaagt ggatgaaaag gacttcttag agattgcgcg agggctcgtg	900
gatatcaaac agagcttctt gtgggtagtg agaccgggat tcgttaaggg ctgcacgtgg	960
gtcagaccgt tgcagatgg ttttctaggg gagagaggga gaatcgtgaa atgggttcca	1020
cagcaagagg ttttggctca cggagctata ggggcctttt ggaccactc tggttggaat	1080
tctactcttg aaagtgtctg tgaaggcgtt ccaatgatat tttctgattt tgggcttgac	1140
cagcctctaa acgtcgccta tatgtctgat gtgttgaagg ttggcgtgta cctggagaat	1200
ggtttgggaaa ggggggaaat tgccaacgcc atacgccggg taatggtgga cgaggaaggt	1260
gagtacatac gtcagaacgc tcgggtttta aaacaaaaag cggacgtcag cttatgaag	1320
ggaggtagct cctatgaatc cctagaatcc ttgtaagct atatatcttc gttataa	1377

<210> 3

<211> 458

<212> PRT

<213> Artificial sequence

<220><223> CP1 amino acid sequence

<400> 3

Met Asn Trp Gln Ile Leu Lys Glu Ile Leu Gly Lys Met Ile Lys Gln

1 5 10 15
 Thr Lys Ala Ser Ser Gly Val Ile Trp Asn Ser Phe Lys Glu Leu Glu
 20 25 30
 Glu Ser Glu Leu Glu Thr Val Ile Arg Glu Ile Pro Ala Pro Ser Phe

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

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

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

 225 230 235 240
 Asp Val Ser Leu Met Lys Gly Gly Ser Ser Tyr Glu Ser Leu Glu Ser
 245 250 255

Leu Val Ser Tyr Ile Ser Ser Leu Glu Asn Lys Thr Glu Thr Thr Val
 260 265 270
 Arg Arg Arg Arg Arg Ile Ile Leu Phe Pro Val Pro Phe Gln Gly His
 275 280 285
 Ile Asn Pro Ile Leu Gln Leu Ala Asn Val Leu Tyr Ser Lys Gly Phe
 290 295 300
 Ser Ile Thr Ile Phe His Thr Asn Phe Asn Lys Pro Lys Thr Ser Asn
 305 310 315 320
 Tyr Pro His Phe Thr Phe Arg Phe Ile Leu Asp Asn Asp Pro Gln Asp
 325 330 335
 Glu Arg Ile Ser Asn Leu Pro Thr His Gly Pro Leu Ala Gly Met Arg
 340 345 350
 Ile Pro Ile Ile Asn Glu His Gly Ala Asp Glu Leu Arg Arg Glu Leu
 355 360 365
 Glu Leu Leu Met Leu Ala Ser Glu Glu Asp Glu Glu Val Ser Cys Leu
 370 375 380
 Ile Thr Asp Ala Leu Trp Tyr Phe Ala Gln Ser Val Ala Asp Ser Leu
 385 390 395 400
 Asn Leu Arg Arg Leu Val Leu Met Thr Ser Ser Leu Phe Asn Phe His
 405 410 415
 Ala His Val Ser Leu Pro Gln Phe Asp Glu Leu Gly Tyr Leu Asp Pro
 420 425 430
 Asp Asp Lys Thr Arg Leu Glu Glu Gln Ala Ser Gly Phe Pro Met Leu
 435 440 445
 Lys Val Lys Asp Ile Lys Ser Ala Tyr Ser
 450 455
 <210> 4
 <211> 1377
 <212> DNA
 <213> Artificial sequence
 <220><223> CP1 nucleotide sequence
 <400> 4

atgaactggc aaatcctgaa agaaatcctg ggtaaaatga tcaaaaaaac caaagcgtcg 60
tcgggcgtta tctggaactc cttcaaagaa ctggaagaat cagaactgga aaccgttatt 120

cgcgaaatcc cggtccgtc gttcctgatt ccgctgccga aacatctgac cgcgagcagc 180
agcagcctgc tggatcacga ccgtacggtc ttccagtggc tggatcacga accgccgtca 240
tcggtgctgt atgtttcatt cggtagcacc tctgaagtcg atgaaaaaga ctttctggaa 300
atcgctcgcg gcctgggtga tagtaaacag tccttctgtt gggtgggttcg tccgggtttt 360
glgaaaggca gcacgtgggt tgaaccgtcg ccggtgggtt tcctgggtga acgcggccgt 420
attgtcaaat gggtgccga gcaagaagtg ctggcacatg gtgctatcgg cgcgttttgg 480
accactctg gttggaacag tacgttgga tccgttttgc aagggtgtcc gatgattttc 540

agcgattttg gcctggacca gccgtgaat gcccgctata tgtctgatgt tctgaaagtc 600
ggtgtgtacc tggaaaacgg ttgggaacgt ggcgaaattg cgaatgcat ccgtcgcgtt 660
atggtcgatg aagaaggcga atacattcgc cagaacgtc gtgtcctgaa acaaaaagcg 720
gacgtgagcc tgatgaaagg cggtagctct tatgaatcac tggaatcgct ggtagctac 780
atcagttccc tggaaaataa aaccgaaacc acggtgcgtc gccgtgcgcg tattatcctg 840
ttcccgggtc cgtttcaggg tcatattaac ccgacatcgc aactggcgaa tgttctgtat 900
tcaaaaggct tttgatcac catcttccat acgaacttca acaaaccgaa aaccagtaac 960

taccgcact ttacgttccg ctttattctg gataacgacc cgcaggatga acgtatctcc 1020
aatctgccga cccacggccc gctggccggt atgcgcattc cgattatcaa tgaacacggt 1080
gcagatgaac tgcgcctga actggaactg ctgatgctgg ccagtgaaga agatgaagaa 1140
gtgtcctgtc tgatcacga cgcactgtgg tatttcgccc agagcgttgc agattctctg 1200
aacctgcgcc gtctggtcct gatgacgtca tcgtgttca atttcatgc gcacgtttct 1260
ctgccgaat ttgatgaact gggctacctg gaccggatg acaaaaccg tctggaagaa 1320
caagccagtg gttttccgat gctgaaagtc aaagacatta aatccgccta ttcgtaa 1377

<210> 5

<211> 839

<212> PRT

<213> Artificial sequence

<220><223> B-glul amino acid sequence

<400> 5

Met Thr Gln Leu Asp Val Glu Ser Leu Ile Gln Glu Leu Thr Leu Asn

1

5

10

15

Glu Lys Val Gln Leu Leu Ser Gly Ser Asp Phe Trp His Thr Thr Pro
 20 25 30
 Val Arg Arg Leu Gly Ile Pro Lys Met Arg Leu Ser Asp Gly Pro Asn
 35 40 45

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

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

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

 Asp Leu Glu Met Pro Gly Pro Pro Gln Cys Arg Val His Ser Ala Thr
 245 250 255
 Asp His Ala Ile Asn Ser Gly Glu Ile His Ile Asn Asp Val Asp Glu

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

Val Lys Gly Lys Ala Tyr Lys Ile Lys Val Leu Tyr Gly Ser Ser Val
515 520 525

Thr Arg Thr Tyr Glu Ile Ala Ala Ser Val Ala Phe Glu Gly Gly Ala
530 535 540

Phe Thr Phe Gly Ala Ala Lys Gln Arg Asn Glu Asp Glu Glu Ile Ala
545 550 555 560

Arg Ala Val Glu Ile Ala Lys Ala Asn Asp Lys Val Val Leu Cys Ile
565 570 575

Gly Leu Asn Gln Asp Phe Glu Ser Glu Gly Phe Asp Arg Pro Asp Ile
580 585 590

Lys Ile Pro Gly Ala Thr Asn Lys Met Val Ser Ala Val Leu Lys Ala
595 600 605

Asn Pro Asn Thr Val Ile Val Asn Gln Thr Gly Thr Pro Val Glu Met
610 615 620

Pro Trp Ala Ser Asp Ala Pro Val Ile Leu Gln Ala Trp Phe Gly Gly
625 630 635 640

Ser Glu Ala Gly Thr Ala Ile Ala Asp Val Leu Phe Gly Asp Tyr Asn
645 650 655

Pro Ser Gly Lys Leu Thr Val Thr Phe Pro Leu Arg Phe Glu Asp Asn
660 665 670

Pro Ala Tyr Leu Asn Phe Gln Ser Asn Lys Gln Ala Cys Trp Tyr Gly
675 680 685

Glu Asp Val Tyr Val Gly Tyr Arg Tyr Tyr Glu Thr Ile Asp Arg Pro
690 695 700

Val Leu Phe Pro Phe Gly His Gly Leu Ser Phe Thr Glu Phe Asp Phe
705 710 715 720

Thr Asp Met Phe Val Arg Leu Glu Glu Glu Asn Leu Glu Val Glu Val
725 730 735

Val Val Arg Asn Thr Gly Lys Tyr Asp Gly Ala Glu Val Val Gln Leu
740 745 750

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

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

 Ser Ser Ser Ala Asp Ile Lys Val Ser Gln Ser Ile Thr Leu Glu Lys
 820 825 830
 Thr Thr Phe Trp Lys Gly Leu
 835

<210> 6

<211> 2520

<212> DNA

<213> Artificial sequence

<220><223> B-glul nucleotide sequence

<400> 6

atgacccaac tggatgtgga gagcctgatt caagagctga cccigaacga aaagggtgcaa	60
ctgctgagcg gtagcgactt ctggcatacc accccggttc gtcgtctggg catcccgaag	120
atgcgtctga gcgacgggtcc gaacggcggtt cgtggtacca aattctttaa cgggtgttccg	180
accgctgct tcccgtagcg taccggtctg ggcgcgacct ttgacaagga actgctgaaa	240
gaggcgggta gcctgatggc ggatgaagcg aaagcgaaag cggcgagcgt ggttctgggt	300
ccgaccgca acattgcgcg tgggtccgaac ggtggccgtg gcttcgagag cttcggcgag	360
gacccggtgg ttaacggtct gacgagcgcg gcgatgatca acggcctgca gggcaagtac	420
attgcggcga ccatgaaaca ctatgtttgc aacgatctgg aaatggaccg taactgcatt	480
gacgcgcaag ttagccaccg tgcgtgcgt gaggtgtacc tgctgccgtt ccaaatacgcg	540
gtgcgtgatg cgaacccgcg tgcgattatg accgcgtata acaaggcgaa cggcgaacac	600
gttagccaga gcaaattcct gctggacgaa gtgctgcgta aggagtgggg ctgggatggt	660
ctgctgatga gcgactggtt tgggtgtttac gatgcgaaaa gcagcatcac caacggcctg	720
gacctggaga tgccgggtcc gccgcagtgc cgtgtgcaca gcgcgaccga tcacgcgatc	780
aacagcggcg aaatccacat taacgatgtt gacgagcgtg tgcgtagcct gctgagcctg	840
attaactact gccaccaaag cgggtgttacc gaggaagatc cggaaccag cgacaacaac	900

accccggaac ccacgagaa gctgcgtaaa atcagccgtg agagcattgt gctgctgaag 960
 gacgatgacc gtaaccgtag cattctgccg ctgaagaaaa gcgacaaaat cgcggttatt 1020

 ggtaacaacg cgaacaacgc ggcgtattgc ggtggcggta gcgcgagcgt gctgagctat 1080
 cacaccacca ccccggttca cagcatcaag agccgtctgg aagatagcaa cccccggcg 1140
 tacaccattg gtgcggacgc gtataaaaac ctgccgccgc tgggtccgca aatgaccgat 1200
 agcgacggca agccgggttt tgatgcgaaa tttttgttg gcagcccgac cagcaaggat 1260
 cgtaaacgta tcgaccattt ccagctgacc aacagccaag ttttctggt ggactactat 1320
 aacgaacaga tcccggaaaa caaggagttc tacgttgacg tggagggtca atttattccg 1380
 gaggaagatg gcacctataa cttcggtctg accgtgtttg gtaccggccg tctgttcgtt 1440

 gatgacaaac tggttagcga cagcagccag aaccaaacc cgggcgatag ctcttttgg 1500
 ctggcggcgc aggaagtgat cggcagcatt cacctggtga agggtaaagc gtacaagatc 1560
 aaagtctgt atggcagcag cgtgaccctg acctacgaaa ttgcggcgag cgttgctgtt 1620
 gagggcggtg cgttcacctt tgggtcggcg aaacagcgta acgaagacga ggaaatcgcg 1680
 cgtgcggtgg agattgcgaa ggcaaacgac aaagtgttc tgtgcatcgg cctgaaccaa 1740
 gatttcgaaa gcgaggggtt tgatcgtccg gacatcaaga ttccgggcgc gaccaacaaa 1800
 atggttagcg cgggtgctgaa ggcaaacccg aacaccgtta ttgtgaacca gaccggtacc 1860

 ccggttgaga tgccgtgggc gagcgtatcg ccggtgatcc tgcaagcgtg gtttggcgg 1920
 agcgaggcgg gtaccgcat tgccgatgtt ctgtttggcg actacaacc gagcggcaag 1980
 ctgaccgtga cttcccgct gcgttttgag gataaccgg cgtacctgaa ctccagagc 2040
 aacaacaag cgtgctggta tggcgaagac gtttacgtgg gttatcgta ctatgagacc 2100
 atcgatcgtc cgggtgctgt cccgtttggt cagggcctga gcttcaccga gttcgatatt 2160
 accgacatgt ttgttcgtct ggaggaagag aacctggaag ttgaggtggt tgtgcgtaac 2220
 accggcaagt acgacgggtc ggaagtgtg cagctgtatg ttgcgccgt tagcccagc 2280

 ctgaaacgtc cgatcaagga actgaaagag tacgcgaaaa ttttctggc gagcgtgaa 2340
 gcgaagaccg ttcacctgag cgtgccgatc aaatacgca ccagcttctt tgatgagtat 2400
 caaaagaaat ggtgcagcga aaagggcgag tataccattc tgctgggtag cagcagcgcg 2460
 gacatcaaag ttagccaaag catcacctg gaaaaacca cttcttgaa aggtctgtaa 2520

 <210> 7
 <211> 481
 <212> PRT
 <213> S. rebaudiana

<400> 7

Met Asp Ala Met Ala Thr Thr Glu Lys Lys Pro His Val Ile Phe Ile

1 5 10 15

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

20 25 30

Leu Leu His His Lys Gly Leu Gln Ile Thr Phe Val Asn Thr Asp Phe

35 40 45

Ile His Asn Gln Phe Leu Glu Ser Ser Gly Pro His Cys Leu Asp Gly

50 55 60

Ala Pro Gly Phe Arg Phe Glu Thr Ile Pro Asp Gly Val Ser His Ser

65 70 75 80

Pro Glu Ala Ser Ile Pro Ile Arg Glu Ser Leu Leu Arg Ser Ile Glu

85 90 95

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

100 105 110

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

115 120 125

Asp Ala Ala Lys Lys Leu Gly Ile Pro Val Met Met Tyr Trp Thr Leu

130 135 140

Ala Ala Cys Gly Phe Met Gly Phe Tyr His Ile His Ser Leu Ile Glu

145 150 155 160

Lys Gly Phe Ala Pro Leu Lys Asp Ala Ser Tyr Leu Thr Asn Gly Tyr

165 170 175

Leu Asp Thr Val Ile Asp Trp Val Pro Gly Met Glu Gly Ile Arg Leu

180 185 190

Lys Asp Phe Pro Leu Asp Trp Ser Thr Asp Leu Asn Asp Lys Val Leu

195 200 205

Met Phe Thr Thr Glu Ala Pro Gln Arg Ser His Lys Val Ser His His

210 215 220

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

225 230 235 240

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

<210> 8

<211> 1446

<212> DNA

<213> S. rebaudiana

<400> 8

```
atggacgcta tggccacgac cgaaaagaaa ccgcacgtta tctttattcc gttcccgga      60
cagagtcaca tcaaggctat gctgaagctg gcccaactgc tgcatacaca aggcctgcaa      120
attacctttg tgaacacgga ttatcatccat aatcagtttc tggaaagctc tggcccgcac      180
tgcctggatg gtgcgccggg ttttcgcttc gaaaccatcc cggatgggtg ctgcgcatagc      240
ccggaagcct ctattccgat ccgtgaatcg ctgctgcgca gcattgaaac caactttctg      300
```

```
gatcgtttca tcgacctggt gacgaaactg ccggaccgcg cgacgtgcat tatctccgac      360
ggctttctgt cagttttcac cattgatgcg gccaaaaagc tgggtatccc ggtcatgatg      420
tattggacgc tggcagcttg tggctttatg ggtttctacc atattcactc actgatcgaa      480
aaaggctttg caccgctgaa ggatgctagt tatctgacca acggctatct ggatacggtc      540
attgactggg tgccgggcat ggaaggtatc cgtctgaaag atttcccgtt ggactggagc      600
accgatctga atgacaaagt gctgatgttt accacggaag cgccgcagcg ctctcataaa      660
gtagtcatc acatttttca caccttcgat gaactggaac cgtcgattat caaaaccctg      720
```

```
agcctgcgtt ataatcatat ttacaccatt ggcccgtgc aactgctgct ggaccaaadc      780
ccggaagaaa agaacaacac cggcatacag tcgctgcacg gttatagcct ggtgaaagaa      840
gaaccggaat gcttcagtg gctgcaatct aaggaaaccg acagtggtgt ttacgtgaat      900
tttggttcca ccacggttat gtcactggaa gatatgaccg aatttggttg gggctctggca      960
aactctaacc attattttct gtggatcatc cgtagtaacc tggtcattgg cgaaaatgca      1020
gtgctgccgc cggaactgga agaacacatt aaaaagcgcg gtttcacgcg ttcctgggtg      1080
tcacaggaaa aagtcttgaa gcatccgtcc gtcggcgggt ttctgacca ctgcggctgg      1140
```

```
ggtagcacga ttgaatctct gagtgtgtgt gttccgatga tttgctggcc gtatagctgg      1200
gatcaactga ccaactgccg ctacatctgt aaagaatggg aagtcggcct ggaaatgggt      1260
acgaaagtga agcgtgacga agttaaacgc ctggtccaag aactgatggg cgaaggcggg      1320
cataaaatgc gtaacaaagc gaaggattgg aaagaaaagg cccgcattgc gattgcgccg      1380
aacggcgaca gcagcctgaa cattgacaaa atggtgaagg aaatcacctg tctggcgcgt      1440
aattaa
```

<210> 9

<211> 459

<212> PRT

<213> H. vulgare

<400> 9

Met Asp Gly Asn Ser Ser Ser Ser Pro Leu His Val Val Ile Cys Pro

1 5 10 15

Trp Leu Ala Leu Gly His Leu Leu Pro Cys Leu Asp Ile Ala Glu Arg

20 25 30

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

35 40 45

Ile Ala Arg Leu Pro Pro Leu Arg Pro Ala Val Ala Pro Leu Val Asp

50 55 60

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

65 70 75 80

Glu Ser Thr Asn Asp Val Pro Tyr Asp Lys Phe Glu Leu His Arg Lys

85 90 95

Ala Phe Asp Gly Leu Ala Ala Pro Phe Ser Glu Phe Leu Arg Ala Ala

100 105 110

Cys Ala Glu Gly Ala Gly Ser Arg Pro Asp Trp Leu Ile Val Asp Thr

115 120 125

Phe His His Trp Ala Ala Ala Ala Ala Val Glu Asn Lys Val Pro Cys

130 135 140

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

145 150 155 160

Gly Val Ser Glu His Ala Ala Ala Ala Val Gly Lys Glu Arg Pro Ala

165 170 175

Ala Glu Ala Pro Ser Phe Glu Thr Glu Arg Arg Lys Leu Met Thr Thr

180 185 190

Gln Asn Ala Ser Gly Met Thr Val Ala Glu Arg Tyr Phe Leu Thr Leu

195 200 205

Met Arg Ser Asp Leu Val Ala Ile Arg Ser Cys Ala Glu Trp Glu Pro

210	215	220	
Glu Ser Val Ala Ala Leu Thr Thr Leu Ala Gly Lys Pro Val Val Pro			
225	230	235	240
Leu Gly Leu Leu Pro Pro Ser Pro Glu Gly Gly Arg Gly Val Ser Lys			
245	250	255	
Glu Asp Ala Ala Val Arg Trp Leu Asp Ala Gln Pro Ala Lys Ser Val			
260	265	270	
Val Tyr Val Ala Leu Gly Ser Glu Val Pro Leu Arg Ala Glu Gln Val			
275	280	285	
His Glu Leu Ala Leu Gly Leu Glu Leu Ser Gly Ala Arg Phe Leu Trp			
290	295	300	
Ala Leu Arg Lys Pro Thr Asp Ala Pro Asp Ala Ala Val Leu Pro Pro			
305	310	315	320
Gly Phe Glu Glu Arg Thr Arg Gly Arg Gly Leu Val Val Thr Gly Trp			
325	330	335	
Val Pro Gln Ile Gly Val Leu Ala His Gly Ala Val Ala Ala Phe Leu			
340	345	350	
Thr His Cys Gly Trp Asn Ser Thr Ile Glu Gly Leu Leu Phe Gly His			
355	360	365	
Pro Leu Ile Met Leu Pro Ile Ser Ser Asp Gln Gly Pro Asn Ala Arg			
370	375	380	
Leu Met Glu Gly Arg Lys Val Gly Met Gln Val Pro Arg Asp Glu Ser			
385	390	395	400
Asp Gly Ser Phe Arg Arg Glu Asp Val Ala Ala Thr Val Arg Ala Val			
405	410	415	
Ala Val Glu Glu Asp Gly Arg Arg Val Phe Thr Ala Asn Ala Lys Lys			
420	425	430	
Met Gln Glu Ile Val Ala Asp Gly Ala Cys His Glu Arg Cys Ile Asp			
435	440	445	
Gly Phe Ile Gln Gln Leu Arg Ser Tyr Lys Ala			
450	455		

<210> 10

<211> 1380

<212> DNA

<213> H. vulgare

<400> 10

```

atggatggta actcctcctc ctgccgctg catgtggtca tttgtccgtg gctggctctg      60
ggtcacctgc tgccgtgtct ggatattgct gaacgtctgg cgtcacgcgg ccatcgtgtc      120
agttttgtgt ccaccccgcg caacattgcc cgtctgccgc cgtgcgtcc ggctgttgca      180
ccgctggttg atttcgtcgc actgccgctg ccgcatgttg acggtctgcc ggagggtgcg      240
gaatcgacca atgatgtgcc gtatgacaaa ttgaaactgc accgtaaggc gttcgatgg      300

ctggcggccc cgtttagcga atttctgctg gcagcttgcg cagaagggtc aggttctcgc      360
ccggaattggc tgattgtgga cacctttcat cactgggcgg cggcggcggc ggtggaaaac      420
aaagtgccgt gtgttatgct gctgctgggt gcagcaacgg tgatcgctgg tttcgcgct      480
ggtgttagcg aacatgcggc ggccggcggtg ggtaaagaac gtccggctgc ggaagccccg      540
agttttgaaa ccgaacgtcg caagctgatg accacgcaga atgcctccgg catgaccgtg      600
gcagaacgct atttctgac gctgatgcgt agcgatctgg ttgcatccg ctcttgcgca      660
gaatgggaac cggaaagcgt ggccagcactg accacgtcgg caggtaaacc ggtggttccg      720

ctgggtctgc tgccgccgag tccggaaggc ggtcgtggcg tttcaaaga agatgctgcg      780
gtccgttggc tggacgcaca gccggcaaag tcagtcgtgt acgtcgact gggttcggaa      840
gtgccgtgc gtgcggaaca agttcacgaa ctggcactgg gcctggaact gagcgggtgct      900
cgctttctgt gggcgctcgc taaaccgacc gatgcaccgg acgccgcagt gctgccgccg      960
ggtttcgaag aacgtaccgc cggccgtggt ctggttgta cgggttgggt gccgcagatt      1020
ggcgttctgg ctcatggtgc ggtggctcgc tttctgacc actgtggctg gaactctacg      1080
atcgaaggcc tgctgttcgg tcatccgctg attatgctgc cgatcagctc tgatcagggt      1140

ccgaatgcgc gcctgatgga aggccgtaaa gtcggtatgc aagtgccgcg tgatgaatca      1200
gacggctcgt ttcgtcgcga agatgttgcc gcaaccgtcc gcgccgtggc agttgaagaa      1260
gacggctcgc gcgtcttcac ggctaacgcg aaaaagatgc aagaaattgt ggccgatggc      1320
gcatgccacg aacgttgtat tgacggtttt atccagcaac tgcgcagtta caaggcgtga      1380

```

【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 6

【변경전】

제4항에 있어서, 레바우디오시드 M을 제조하기에 충분한 시간 동안 레바우디오시드 D4를 UDP 글리코실트랜스퍼라제와 인큐베이션하는 단계를 추가로 포함하며, 여기서 UDP 글리코실트랜스퍼라제는 SEQ ID NO:1의 아미노산 서열을 갖는 UGT76G1 또는 이의 기능적 동족체인 방법.

【변경후】

제4항에 있어서, 레바우디오시드 M을 제조하기에 충분한 시간 동안 레바우디오시드 D4를 UDP 글리코실트랜스퍼라제와 인큐베이션하는 단계를 추가로 포함하며, 여기서 UDP 글리코실트랜스퍼라제는 SEQ ID NO:1의 아미노산 서열을 갖는 UGT76G1인 방법.