



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0091345  
(43) 공개일자 2018년08월16일

(51) 국제특허분류(Int. Cl.)

G01N 33/532 (2006.01) G01N 33/543 (2006.01)

G01N 33/68 (2006.01)

(52) CPC특허분류

G01N 33/532 (2013.01)

C12Y 111/01 (2013.01)

(21) 출원번호 10-2017-0016359

(22) 출원일자 2017년02월06일

심사청구일자 2017년02월06일

(71) 출원인

울산과학기술원

울산광역시 울주군 언양읍 유니스트길 50

(72) 발명자

이현우

울산광역시 울주군 언양읍 유니스트길 50

이송이

울산광역시 울주군 언양읍 유니스트길 50, 303동 2201호

(74) 대리인

유미특허법인

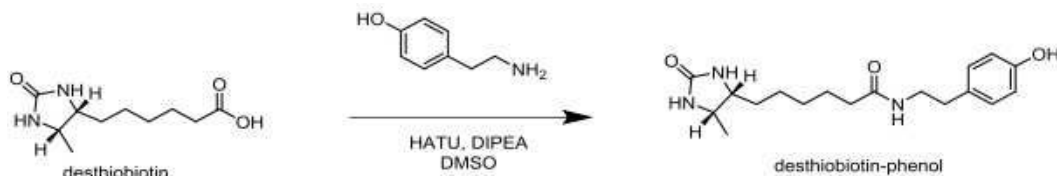
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 단백질 표지화용 페놀 화합물을 이용한 단백질을 탐지하는 방법

### (57) 요약

본 발명은 페놀 화합물, 및 이의 제조방법에 관한 것으로서, 구체적으로 본 발명은 데스티오비오틴-페놀 (Desthiobiotin-Phenol; DBP), 이의 제조방법, 상기 데스티오비오틴-페놀 (Desthiobiotin-Phenol; DBP)을 이용한 단백질 표지화용, 및 단백질을 탐지하는 용도에 관한 것이다.

대표도 - 도1



(52) CPC특허분류

**G01N 33/54326** (2013.01)

**G01N 33/6848** (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711029419

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 신진연구자지원

연구과제명 효소화학반응을 이용한 살아있는 세포의 단백질체 레이블링 시스템

기 여 율 1/1

주관기관 울산과학기술원

연구기간 2015.11.01 ~ 2016.10.31

---

## 명세서

### 청구범위

#### 청구항 1

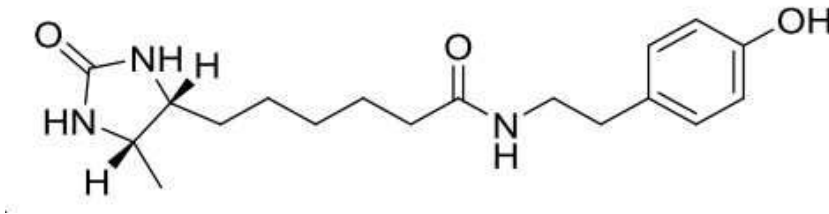
탐지를 위한 세포 내 목적 단백질이 존재하는 세포 소기관으로 표적화 서열을 코딩하는 유전자와 과산화효소를 코딩하는 유전자를 포함하는 폴리뉴클레오타이드를 세포에 도입하는 단계;

상기 세포에 과산화수소와 하기 화학식 1의 구조를 갖는 페놀 화합물, 이의 입체 이성질체 (Stereoisomers), 호변 이성질체 (Tautomers) 또는 이들의 염을 처리하여 상기 페놀 화합물로 세포 내 단백질을 표지화하는 단계;

상기 세포의 과쇄물에 스트렙타비딘 (Streptavidin: SA) 비드를 첨가하여, 상기 스트렙타비딘 비드에 결합된 표지 단백질을 수득하는 단계; 및

상기 수득물에서 목적 단백질을 확인하는 단계를 포함하는 단백질의 탐지 방법:

[화학식 1]



#### 청구항 2

제1항에 있어서, 상기 수득물에서 목적 단백질을 확인하는 단계는, 상기 스트렙타비딘 비드에 결합된 표지 단백질에 효소를 처리하여 단백질을 절단하고, 상기 스트렙타비딘 비드에 결합된 절단 단편을 분리하고, 상기 분리된 절단 단편을 분석하는 단계로 수행되는 것인, 단백질의 탐지 방법.

#### 청구항 3

제2항에 있어서, 상기 절단 단편을 분리하는 것은 탈염처리 (desalting) 하여 수행하는 것인, 단백질의 탐지 방법.

#### 청구항 4

제2항에 있어서, 상기 효소는 트립신 (trypsin)인, 단백질의 탐지방법.

#### 청구항 5

제1항에 있어서, 상기 폴리뉴클레오타이드를 세포에 도입하는 단계에 앞서, 목적 단백질의 세포내 위치를 분석하는 단계를 수행하는 것인 단백질의 탐지방법.

#### 청구항 6

제1항에 있어서, 상기 페놀 화합물은 과산화효소에 의해서 라디칼 (Radical)을 생성하는 것인, 단백질의 탐지 방법.

#### 청구항 7

제6항에 있어서, 상기 과산화효소는 대두 아스코르베이트 과산화효소(Soybean ascorbate peroxidase), 애기장대 아스코르베이트 과산화효소 (Arabidopsis ascorbate peroxidase), 시금치 아스코르베이트 과산화효소 (Spinach ascorbate peroxidase), 사카로미세스 세레비시에 시토크롬 c 과산화효소 (Saccharomyces cerevisiae cytochrome c peroxidase), 파네로카에테 크리소스포리움 과산화효소 PCL1 (Phanerochaete chrysosporium

PCL1), 파네로카에테 크리소스포리움 과산화효소 PCM (Phanerochaete chrysosporium PCM), 코프리놉시스 시네레아 과산화효소 (Coprinopsis cinerea peroxidase), 땅콩 과산화효소 (Peanut Peroxidase), 애기장대 과산화효소 ATPA2 (Arabidopsis ATPA2), 서열번호 1의 아미노산 서열로 이루어진 효소, 서열번호 3의 아미노산 서열로 이루어진 효소 및 서열번호 5의 아미노산 서열로 이루어진 효소로 이루어진 군에서 선택된 1 이상의 효소인, 단백질의 탐지 방법.

#### 청구항 8

제6항에 있어서, 상기 생성된 라디칼은 단백질 또는 펩타이드 표면에 존재하는 타이로신 (Tyrosine)기와 공유결합을 형성하여 상기 단백질 또는 펩타이드에 결합하는 것인, 단백질의 탐지 방법.

#### 청구항 9

제4항에 있어서, 상기 라디칼이 단백질 또는 펩타이드에 결합하는 거리는 5 내지 50 nm인, 단백질의 탐지 방법.

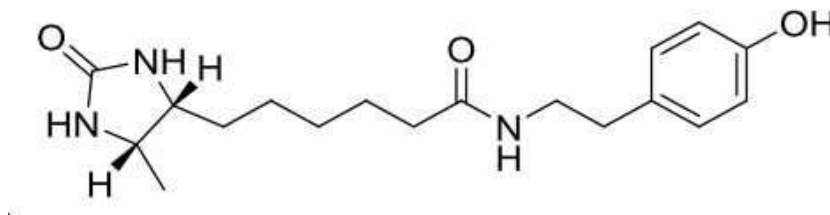
#### 청구항 10

제1항에 있어서, 상기 세포 소기관은 미토콘드리아, 소포체, 핵, 세포질 및 골지체로 이루어진 군에서 선택되는 것인, 단백질의 탐지 방법.

#### 청구항 11

제1항에 있어서, 상기 페놀 화합물의 처리는, 하기 하기 화학식 1의 구조를 가지는 화합물을 0.1  $\mu$ M 내지 1 mM 농도로 10 내지 60분 동안 처리하는 것인, 단백질의 탐지 방법:

[화학식 1]



#### 청구항 12

제1항에 있어서, 상기 과산화수소의 처리는, 과산화수소를 0.1 내지 2mM의 농도로 1 내지 10분 동안 처리하는 것인, 단백질의 탐지 방법.

#### 청구항 13

제1항에 있어서, 상기 세포 소기관으로 표적화 서열을 코딩하는 유전자는 미토콘드리아 기질 발현 유전자, 세포질 발현 유전자, 미토콘드리아 이중막 사이 공간 발현 유전자, 핵 발현 유전자, 세포 외막 발현 유전자, 소포체 내막 발현 유전자, 소포체 외막 발현 유전자, 또는 세포 내막 발현 유전자인, 단백질의 탐지 방법.

#### 청구항 14

제1항에 있어서, 상기 세포는 신장 세포, 암 세포, 피부 세포, 난소 세포, 활막세포, 말초혈액단핵구 세포, 섬유아세포, 섬유세포, 신경세포, 상피세포, 각질세포, 조혈세포, 멜라닌 세포, 연골세포, 대식세포, 근육세포, 혈액 세포, 골수 세포, 림프구 세포, 단핵세포, 폐 세포, 위 세포, 췌장 세포, 간 세포, 위 세포, 장 세포, 심장 세포, 방광 세포, 요도 세포, 배아 생식세포 또는 난구세포인, 단백질의 탐지 방법.

#### 청구항 15

제1항에 있어서, 상기 스트렙타아비딘 비드는 스트렙타아비딘 코팅된 자성 비드인 것인, 단백질의 탐지 방법.

#### 청구항 16

제1항에 있어서, 상기 목적 단백질을 확인하는 단계는 질량 분석, 웨스턴 블롯, 형광 현미경 분석, 닷 블롯, 또

는 ELISA인, 단백질의 탐지 방법.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 단백질 표지화용 페놀 화합물을 포함하는 단백질 탐지용 조성물, 및 상기 페놀 화합물을 이용한 단백질 탐지하는 방법에 관한 것이다. 구체적으로, 본 발명은 데스티오비오틴-페놀 (Desthiobiotin-Phenol; DBP)을 이용하여 단백질을 표지하여 단백질을 탐지하는 방법에 관한 것이다.

### 배경 기술

[0002] 사람 세포에 존재한다고 알려져 있는 18,000여개의 단백질들은 각자 적절한 세포 소기관에 특유의 구조로 존재하면서 기능을 수행한다.

[0003] 단백질은 직선의 1차 구조 (Primary structure)에서 나선 또는 병풍구조로 바뀌는 2차 구조 (Secondary structure)를 거쳐 수소결합, 이온결합, 소수성 결합 등을 통해 3차 구조 (Tertiary structure)를 형성하며, 여러 개의 3차 구조가 모여 4차 구조 (Quaternary structure)를 형성하기도 한다. 단백질의 구조는 매우 복잡하고 단백질이 존재하는 주변 환경에 따라서 그 구조가 유동적으로 변화하기 때문에 단백질의 정확한 구조를 알아 내기가 쉽지 않다.

[0004] 과거 생물학자들은 반응성이 있는 바이오틴 프로브를 사용하여 막단백질 (Membrane protein)의 토폴로지 (Topology)를 확인하거나 단백질 표면의 용제 접근 가능 잔기 (Residue)를 찾곤 하였다. 하지만 바이오틴-에스터 또는 바이오틴 말레이마이드는 수용액 상에서 수명이 매우 길고 ( $t_{1/2} = 1 \text{ hr}$ ) 정제된 단백질이나 세포막에만 적용할 수 있다는 단점이 있어 적용상의 한계가 있었다.

[0005] 최근, 과산화효소 (Peroxidase)를 미토콘드리아 기질 (Mitochondrial Matrix)과 막사이공간에 발현시켜 바이오틴 페놀로 주변단백질을 표지하여 미토콘드리아 기질과 막사이공간을 맵핑한 연구가 있었다. 하지만 이 방법은 페놀기가 표지되지 않은 펩타이드를 분석하는 것이기 때문에 위양성 (False positive)의 결과를 얻을 가능성이 존재하는 문제가 있었다. 또한, 종래의 경우에는 바이오틴의 황원자의 산화 작용 때문에 단일한 최종산물이 아니라 다양한 종류의 최종산물이 혼재하게 되어, 직접적으로 페놀기가 표지된 펩타이드를 질량분석 할 수가 없는 문제점이 있었다.

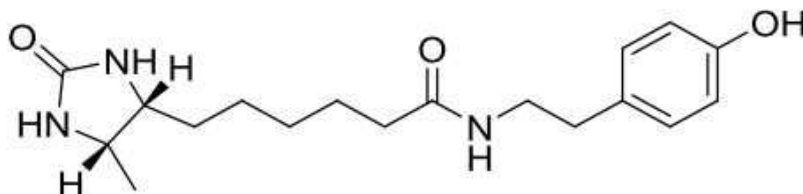
## 발명의 내용

### 해결하려는 과제

[0006] 이에, 상기와 같은 문제점을 해결하기 위해서 과산화효소를 이용한 단백질 표지를 하기 위한 프로브를 개발하여 본 발명을 완성하였다.

[0007] 본 발명의 목적은 화합물내에 황원자가 존재하지 않기 때문에 표지된 단백질 또는 펩타이드가 한 종류만 존재하여, 표지된 단백질 또는 펩타이드의 질량분석이 가능하고, 상기 질량분석 결과를 통해 단백질의 종류를 확인하고, 그 구조정보도 확인이 가능한, 하기 화학식 1을 갖는 페놀 화합물, 및 이의 제조방법에 관한 것이다.

[0008] [화학식 1]



[0009]

[0010] 본 발명의 또 다른 목적은, 상기의 화학식 1의 페놀 화합물, 이의 입체 이성질체, 호변 이성질체, 또는 이들의 염을 포함하는 단백질 또는 펩타이드 표지용 조성물을 제공하는 것이다.

[0011] 본 발명의 또 다른 목적은, 상기의 화학식 1의 페놀 화합물, 이의 입체 이성질체, 호변 이성질체, 또는 이들의

염을 포함하는 단백질 또는 펩타이드의 표지화하는 방법에 관한 것이다.

[0012] 또한, 본 발명의 목적은 화합물내에 황원자가 존재하지 않기 때문에 표지된 단백질 또는 펩타이드가 한 종류만 존재하여, 표지된 단백질 또는 펩타이드의 질량분석이 가능하고, 상기 질량분석 결과를 통해 단백질의 종류를 확인할 수 있는, 상기 화학식 1을 갖는 페놀 화합물을 이용하여 단백질을 표지하여 단백질을 탐지하는 방법에 관한 것이다.

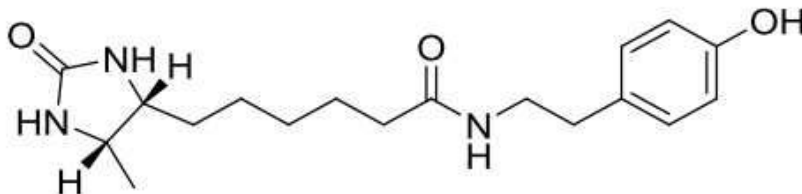
[0013] 본 발명의 또 다른 목적은 상기 화학식 1을 갖는 페놀 화합물을 포함하는 단백질 탐지용 조성물을 제공하는 것이다.

### 과제의 해결 수단

[0014] 이에, 상기와 같은 문제점을 해결하기 위해서 과산화효소를 이용한 단백질 표지를 하기 위한 프로브를 개발하여 본 발명을 완성하였다.

[0015] 본 발명의 목적은 화합물내에 황원자가 존재하지 않기 때문에 표지된 단백질 또는 펩타이드가 한 종류만 존재하여, 표지된 단백질 또는 펩타이드의 질량분석이 가능하고, 상기 질량분석 결과를 통해 단백질의 종류를 확인할 수 있고, 그 구조정보도 확인이 가능한, 하기 화학식 1을 갖는 페놀 화합물, 및 이의 제조방법에 관한 것이다.

[0016] [화학식 1]



[0017] 본 발명의 또 다른 목적은, 상기의 화학식 1의 페놀 화합물, 이의 입체 이성질체, 호변 이성질체, 또는 이들의 염을 포함하는 단백질 또는 펩타이드 표지용 조성물을 제공하는 것이다.

[0019] 본 발명의 또 다른 목적은, 상기의 화학식 1의 페놀 화합물, 이의 입체 이성질체, 호변 이성질체, 또는 이들의 염을 포함하는 단백질 또는 펩타이드의 표지화하는 방법에 관한 것이다.

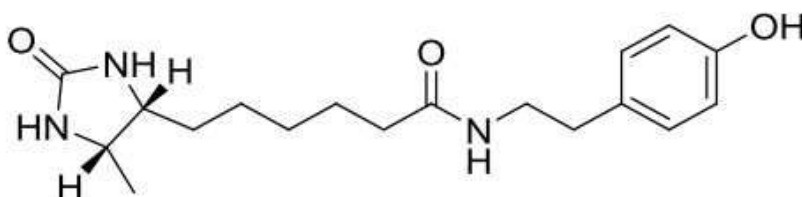
[0020] 또한, 본 발명의 목적은 화합물내에 황원자가 존재하지 않기 때문에 표지된 단백질 또는 펩타이드가 한 종류만 존재하여, 표지된 단백질 또는 펩타이드의 질량분석이 가능하고, 상기 질량분석 결과를 통해 단백질의 종류를 확인할 수 있는, 상기 화학식 1을 갖는 페놀 화합물을 이용하여 단백질을 표지하여 단백질을 탐지하는 방법에 관한 것이다.

[0021] 본 발명의 또 다른 목적은 상기 화학식 1을 갖는 페놀 화합물을 포함하는 단백질 탐지용 조성물을 제공하는 것이다.

[0022] 본 발명에 따른 페놀 화합물인 데스티오비오틴-페놀 (Desthiobiotin-Phenol; DBP)은 황원자를 포함하지 않기 때문에 최종산물이 한 종류만 존재하여, 이를 이용하여 단백질을 표시할 경우, 표지된 단백질을 용이하게 탐지할 수 있다. 자세하게는, 본 발명에 따른 데스티오비오틴-페놀은 단백질의, 타이로신 잔기에 결합하여, 표지하며, 그 단백질의 구조에 관한 정보도 확인할 수 있다.

[0023] 본 발명의 일예는 하기의 화학식 1의 구조를 가지는 페놀 화합물, 이의 입체 이성질체 (Stereoisomers), 호변 이성질체 (Tautomers) 또는 이들의 염을 제공하는 것이다.

[0024] [화학식 1]

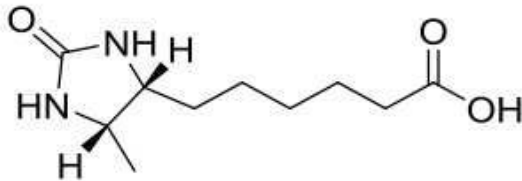


[0025]

[0026] 본 발명의 일예는 하기 화학식 2의 화합물과 티라민 (Tyramine)을 아마이드 결합으로 연결하는 단계를 포함하는, 상기 화학식 1의 화합물을 제조하는 방법을 제공하는 것이다.

[0027] [화학식 2]

[0028]



[0029] 본 발명의 일예는 상기의 화학식 1의 페놀 화합물, 이의 입체 이성질체, 호변 이성질체, 또는 이들의 염을 포함하는 단백질 또는 펩타이드 표지용 조성물을 제공하는 것이다.

[0030] 본 발명의 일예는 표지를 위한 세포 내 목적 단백질이 존재하는 세포 소기관으로 표적화 서열을 코딩하는 유전자와 과산화효소를 코딩하는 유전자를 포함하는 폴리뉴클레오타이드를 세포에 도입하는 단계;

[0031] 상기 화합물 및 과산화수소를 상기 세포에 처리하는 단계;

[0032] 상기 세포를 파쇄하여 세포 파쇄물을 수득하는 단계; 및

[0033] 상기 세포 파쇄물에 스트렙타비딘 (Streptavidin: SA) 비드를 넣어 표지된 단백질을 수득하는 단계를 포함하는 단백질의 표지 방법을 제공하는 것이다.

[0034] 본 발명자들은 새로운 화학적 프로브 (a new chemical probe)를 디자인하였다. 본 발명의 새로운 화학적 프로브는 과산화효소에 의해서 단백질을 표지하는 역할을 할 수 있다.

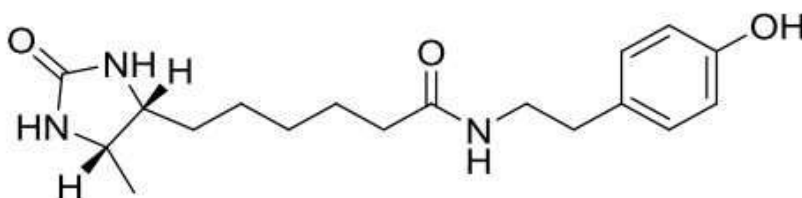
[0035] 본 발명에 있어서, 과산화효소를 세포에 처리하고 상기 화학적 프로브 (테비오틴이 결합된 페놀 화합물) 및 과산화수소를 처리하면, 세포 안에서 발현된 상기 APEX의 작용으로 테스티오비오틴이 결합된 페놀가 테스티오비오틴이 결합된 페녹시 라디칼 (biotin-phenoxy radical)로 전환되고, 상기 테스티오비오틴이 결합된 페녹시 라디칼은 반경 약 5 내지 50 nm 내에 존재하는 단백질들에 존재하는 타이로신 (Tyrosine)기와 공유결합을 형성함으로써 상기 단백질들에 테스티오비오틴이 결합하게 된다. 또한, 생성된 테스티오비오틴이 결합된 페녹시 라디칼은 인지질 막을 건널 수 없을 만큼 짧은 시간 동안만 존재하기 때문에, 막으로 둘러싸인 각각의 세포 소기관내에서 비오틴이 결합된 페녹시 라디칼이 생성되는 경우 각 세포 소기관내의 고유의 단백질들에만 비오틴이 결합할 수 있게 된다.

[0036] 이하 본 발명을 더욱 자세히 설명하고자 한다.

[0037] 본 발명은 하기의 화학식 1의 구조를 가지는 신규한 페놀 화합물, 이의 입체 이성질체 (Stereoisomers), 호변 이성질체 (Tautomers) 또는 이들의 염에 관한 것이다.

[0038] [화학식 1]

[0039]



[0040] 상기 화학식 1의 구조를 가지는 신규한 화합물은 테스티오비오틴-페놀(Desthiobiotin-Phenol; 이하 DBP)이다.

[0041] 본 명세서에서, 용어 "입체 이성질체"는 분자식 및 구성원자의 연결 방법도 같으나 원자 사이의 공간적 배치가 다른 것을 말하며, 그 결합방법의 차이에 따라 부분입체 이성질체 및 거울상 이성질체로 나뉜다.

[0042] 본 명세서에서, 용어 "호변 이성질체"는 유기물이 쉽게 변환이 가능한 구조 이성질체를 의미하며, 두 개의 호변 이성질체 간의 변환은 대개는 수소원자의 결합 위치가 변함으로써 일어나는 것이 알려져 있다.

[0043] 상기 신규한 화합물, 이의 입체 이성질체, 또는 호변 이성질체에 대한 염은 산 또는 염기의 부가염 및 그의 입

체화학적 이성체 형태를 모두 포함하며, 예컨대, 유기산 또는 무기산의 부가염일 수 있다. 상기 염에는 투여 대상에서 모화합물(parent compound)의 활성을 유지하며 바람직하지 못한 효과를 유발하지 않는 염이라면 어느 것이든 포함되는 것으로, 특별히 제한되는 것이 아니다.

[0044] 이러한 염에는 무기염과 유기염이 포함되며, 예를 들어, 아세트산, 질산, 아스파르트산, 술폰산, 설푸릭산, 말레산, 글루탐산, 포름산, 숙신산, 인산, 프탈산, 탄닌산, 타르타르산, 히드로브롬산, 프로피온산, 벤젠술폰산, 벤조산, 스테아르산, 락트산(lactic acid), 비카르본산, 비설푸릭산, 비타르타르산, 옥살산, 부틸산, 칼슘 이데트, 카르보닉산, 클로로벤조산, 시트르산, 이데트산, 톨루엔술폰산, 푸마르산, 글루세프산, 에실린산, 파모익산, 글루코닉산, 메틸니트릭산, 말론산, 히드로클로린산, 히드로요도익산, 히드록시나프톨산, 이세티온산, 락토비오닉산, 만델산, 점액산, 나프틸릭산, 뮤코닉산, p-니트로메탄설폰산, 헥사믹산, 판토테닉산, 모노히드로젠인산, 디히드로젠인산, 살리실산, 술파민산, 술파닐린산, 메탄술폰산일 수 있다.

[0045] 또한, 상기 염의 형태로는, 암모늄염, 리튬염, 소듐염, 포타슘염, 마그네슘염 및 칼슘염과 같은 알칼리 및 알칼리 토금속의 염, 예를 들어, 벤자틴, N-메틸-D-글루카민, 하이드라바민 염과 같은 유기 염기를 갖는 염, 및 아르기닌, 리신과 같은 아미노산을 갖는 염을 포함한다. 또한, 상기 염 형태는 적당한 염기 또는 산으로 처리함으로써 유리 형태로 전환될 수도 있다.

[0046] 상기 페놀 화합물은 과산화효소에 의해서 라디칼을 생성하며, 상기 과산화효소 자체에서도 라디칼이 생성될 수 있다. 상기 페놀 화합물 또는 과산화효소에 생성된 라디칼은 전파(propagation)될 수도 있다. 상기 페놀 화합물에 라디칼이 생성되는 위치 및 상기 라디칼이 전파되는 과정은 도 30에 나타나 있다.

[0047] 상기 생성된 라디칼은 단백질 또는 펩타이드에 존재하는 타이로신(Tyrosine)기와 공유결합을 형성하여 상기 단백질 또는 펩타이드에 결합하는 것일 수 있으며, 상기 라디칼이 단백질 또는 펩타이드에 결합하는 거리는 5 내지 50 nm일 수 있으나, 이에 제한되는 것은 아니다.

[0048] 상기 과산화효소는 대두 아스코르베이트 과산화효소(Soybean ascorbate peroxidase), 애기장대 아스코르베이트 과산화효소(Arabidopsis ascorbate peroxidase), 시금치 아스코르베이트 과산화효소(Spinach ascorbate peroxidase), 사카로미세스 세레비시에 시토크롬 c 과산화효소(Saccharomyces cerevisiae cytochrome c peroxidase), 파네로카에테 크리소스포리움 과산화효소 PCL1(Phanerochaete chrysosporium PCL1), 파네로카에테 크리소스포리움 과산화효소 PCM(Phanerochaete chrysosporium PCM), 코프린옵시스 시네레아 과산화효소(Coprinopsis cinerea peroxidase), 땅콩 과산화효소(Peanut Peroxidase), 애기장대 과산화효소 ATPA2(Arabidopsis ATPA2), 서열번호 1의 아미노산 서열로 이루어진 효소, 서열번호 3의 아미노산 서열로 이루어진 효소 및 서열번호 5의 아미노산 서열로 이루어진 효소로 이루어진 군에서 선택된 1 이상의 효소일 수 있으나, 이에 제한되는 것은 아니다.

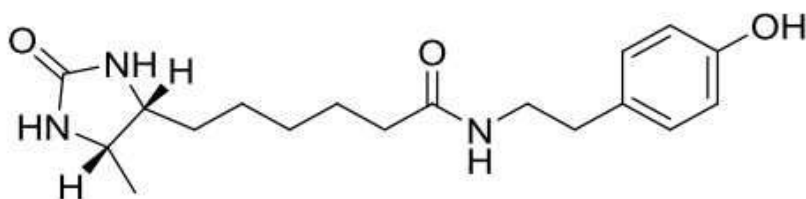
[0049] 상기 서열번호 1의 아미노산 서열을 포함하는 아스코르베이트 과산화효소는 서열번호 2의 염기 서열에 의하여 코딩되고, 서열번호 3의 아미노산 서열을 포함하는 아스코르베이트 과산화효소는 서열번호 4의 염기 서열에 의하여 코딩될 수 있다.

[0050] 본 발명에서, 서열번호 5의 아미노산 서열로 이루어진 효소는 호스라디쉬 과산화효소(Horseradish Peroxidase, HRP)이다. 상기 호스라디쉬 과산화효소는 산화 환경을 갖춘 분비 경로(Secretory Pathway)에서 활성을 나타내는 것으로, 본 발명에서는 자연적으로 산화환경이 조성되어있는 분비 경로상의 세포 내 공간인 소포체나 골지체 등에서 서열번호 1 또는 서열번호 3으로 이루어진 아스코르베이트 과산화효소와 동일한 효과를 낼 수 있다. 서열번호 5의 아미노산 서열을 포함하는 호스라디쉬 과산화효소는 서열번호 6의 염기 서열에 의하여 코딩될 수 있다.

[0051] 본 발명은 하기 화학식 1의 화합물을 제조하는 방법을 제공하는 것이다.

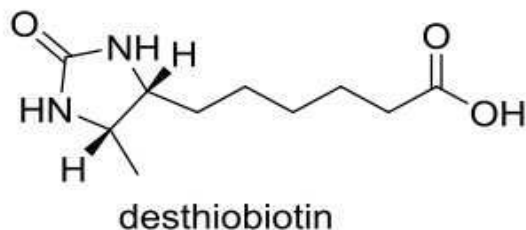
[0052] 본 발명에서 하기 화학식 1의 화합물은 하기의 화학식 2의 화합물을 시작물질로 사용하여 반응식 1의 반응에 의해 제조될 수 있다.

[0053] [화학식 1]



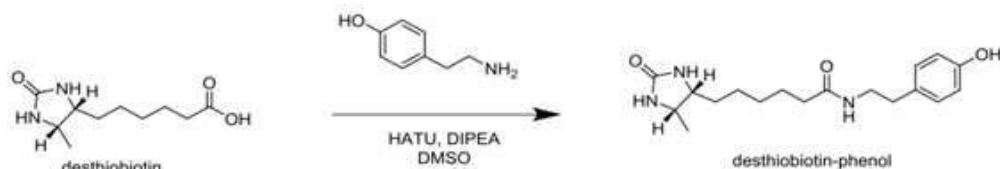
[0054]

[0055] [화학식 2]



[0056]

[0057] [반응식 1]



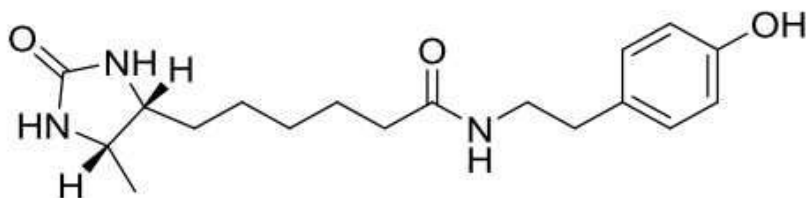
[0058]

[0059] 구체적으로, 상기 화학식 2의 화합물과 티라민 (Tyramine)을 아미드 결합으로 연결할 수 있고, 상기 아미드 결합은 상기 화학식 2의 화합물을 유기용매에 용해하고 커플링제와 티라민을 첨가하여 형성할 수 있으며, 상기 커플링제는 2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU)과 N,N-diisopropylethylamine (DIPEA)일 수 있으나, 이에 제한되는 것은 아니다.

[0060] 본 발명은 하기의 화학식 1의 페놀 화합물을 포함하는 단백질 또는 펩타이드 표지용 조성물을 제공하는 것이다.

[0061] 본 발명에서 상기 단백질 표지용 조성물은 하기의 화학식 1의 페놀 화합물, 이의 입체 이성질체, 호변 이성질체, 또는 이들의 염을 포함하는 것일 수 있으나, 이에 제한되는 것은 아니다:

[0062] [화학식 1]



[0063]

[0064] 상기 페놀 화합물은 과산화효소에 의해서 라디칼을 생성하며, 상기 과산화효소 자체에서도 라디칼이 생성될 수 있다. 상기 페놀 화합물 또는 과산화효소에 생성된 라디칼은 전파 (propagation)될 수도 있다. 상기 페놀 화합물에 라디칼이 생성되는 위치 및 상기 라디칼이 전파되는 과정은 도 30에 나타나 있다.

[0065] 상기 생성된 라디칼은 단백질 또는 펩타이드에 존재하는 타이로신 (Tyrosine)기와 공유결합을 형성하여 상기 단백질 또는 펩타이드에 결합하는 것일 수 있으며, 상기 라디칼이 단백질 또는 펩타이드에 결합하는 거리는 5 내지 50 nm일 수 있으나, 이에 제한되는 것은 아니다.

[0066] 상기 과산화효소는 대두 아스코르베이트 과산화효소 (Soybean ascorbate peroxidase), 애기장대 아스코르베이트 과산화효소 (Arabidopsis ascorbate peroxidase), 시금치 아스코르베이트 과산화효소 (Spinach ascorbate

peroxidase), 사카로미세스 세레비시에 시토크롬 c 과산화효소 (Saccharomyces cerevisiae cytochrome c peroxidase), 파네로카에테 크리소스포리움 과산화효소 PCL1 (Phanerochaete chrysosporium PCL1), 파네로카에테 크리소스포리움 과산화효소 PCM (Phanerochaete chrysosporium PCM), 코프리놉시스 시네레아 과산화효소 (Coprinosia cinerea peroxidase), 땅콩 과산화효소 (Peanut Peroxidase), 애기장대 과산화효소 ATPA2 (Arabidopsis ATPA2), 서열번호 1의 아미노산 서열로 이루어진 효소, 서열번호 3의 아미노산 서열로 이루어진 효소 및 서열번호 5의 아미노산 서열로 이루어진 효소로 이루어진 군에서 선택된 1 이상의 효소일 수 있으나, 이에 제한되는 것은 아니다.

[0067] 본 발명에서 상기 단백질 표지용 조성물은 표지를 위한 세포 내 목적 단백질을 표지하는데 사용될 수 있다.

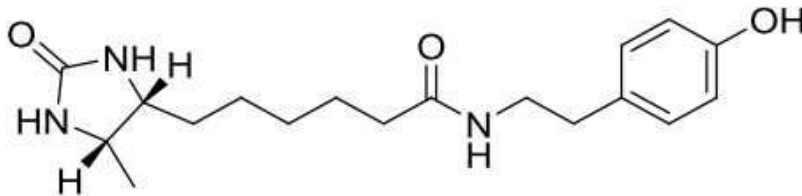
[0068] 본 발명은 하기 화학식 1의 페놀 화합물을 이용하여 단백질을 표지하는 방법을 제공하는 것이다.

[0069] 상기 단백질을 표지하는 방법은 하기와 같은 방법을 통해 수행될 수 있으나, 이에 제한되는 것은 아니다.

[0070] 구체적으로, 본 발명은 표지를 위한 세포 내 목적 단백질이 존재하는 세포 소기관으로 표적화 서열을 코딩하는 유전자와 과산화효소를 코딩하는 유전자를 포함하는 폴리뉴클레오타이드를 세포에 도입하는 단계; 및

[0071] 상기 세포에 하기 화학식 1의 페놀 화합물 및 과산화수소를 처리하여 상기 페놀 화합물로 세포 내 단백질을 표지화하는 단계를 포함하는 단백질의 표지 방법을 제공하는 것이다.

[0072] [화학식 1]



[0073]

[0074] 상기 폴리뉴클레오타이드를 세포에 도입하는 단계 이전에 상기 폴리뉴클레오타이드를 제조하는 과정을 추가로 포함할 수 있다.

[0075] 상기 폴리뉴클레오타이드를 제조하는 과정은 상기 목적 단백질이 존재하는 세포 소기관으로 표적화 서열을 코딩하는 유전자와 과산화효소를 코딩하는 유전자가 연결되도록 융합된 폴리뉴클레오타이드를 제조한 것을 의미할 수 있으나, 이에 제한되는 것은 아니다.

[0076] 상기 폴리뉴클레오타이드를 제조하는 과정에서, 상기 세포 소기관은 미토콘드리아, 소포체, 핵, 세포질 및 골지체로 이루어진 군에서 선택되는 것일 수 있으며, 상기 세포 소기관으로 표적화 서열을 코딩하는 유전자에는 미토콘드리아 기질 발현 유전자, 세포질 발현 유전자, 미토콘드리아 이중막 사이 공간 발현 유전자, 핵 발현 유전자, 세포 외막 발현 유전자, 소포체 내막 발현 유전자, 소포체 외막 발현 유전자, 또는 세포 내막 발현 유전자일 수 있으나, 이에 제한되는 것은 아니다. 상기 세포 소기관으로 표적화 서열을 코딩하는 유전자는 하기 표 1에 나타나 있다.

표 1

[0077]

| 플라스미드     | 특징                                                               | Promoter/ Vector | 세부사항                                                                           |
|-----------|------------------------------------------------------------------|------------------|--------------------------------------------------------------------------------|
| mito-APEX | NotI-mito-BamHI-V5-APEX-Stop-XhoI<br>미토콘드리아 기질 발현                | CMV/pCDNA3       | mito matrix targeting sequence: MLATRVFSLVGKRAISTSVCVRAH<br>V5: GKPIPNLLGLDST  |
| APEX-NES  | NotI-Flag-APEX-NES-Stop-XhoI<br>세포질 발현                           | CMV/pCDNA3       | NES: LQLPPLRLTLD                                                               |
| IMS-APEX  | NotI-LACTB(1-68)-BamHI-V5-APEX-Stop-XhoI<br>미토콘드리아 이중 막 사이 공간 발현 | CMV/pCDNA3       | LACTB(1-68): MYRLSSVTARAAATAGPAWDGGRRGAHRRPGLPVLGLGWAGGLGLGLALGAKLVVGLRGAVPIQS |

|               |                                                         |              |                                                                                                                                                                     |
|---------------|---------------------------------------------------------|--------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| APEX-NLS      | NotI-Flag-APEX-EcoRI-NLS-Stop-XhoI<br>핵 발현              | CMV/pCDNA3   | NLS: PKKKRKVDPKKKRKVDPKKKRKV                                                                                                                                        |
| APEX-TM       | EcoRI-ss-HA-ApaI-APEX-SacII-myc-TM-Stop-NotI<br>세포외막 발현 | CMV/pDisplay | ss: Igk chain signal sequence in pDisplay (Invitrogen).<br>HA: YPYDVPDYA<br>myc: EQKLISEEDL<br>TM: transmembrane domain of PDGF receptor from pDisplay (Invitrogen) |
| APEX-KDEL     | EcoRI-ss-BglIII-APEX-V5-KDEL-Stop-NotI<br>소포체 내막 발현     | CMV/pDisplay | KDEL: ER retention motif                                                                                                                                            |
| C1(1-29)-APEX | HindIII-C1(1-29)-NotI-APEX-V5-Stop-XhoI<br>소포체 외막 발현    | CMV/pCDNA3   | C1(1-29):MDPVVVLGLCLSCLLLSLWKQSYGGGKL<br>(endoplasmic reticulum membrane anchor)                                                                                    |
| W41AAPX-CAAX  | NotI-Flag-W41AAPX-EcoRI-CAAX-Stop-XhoI<br>세포내막 발현       | CMV/pCDNA3   | CAAX: RSKLNPPDESGPGCMSCCKVLS                                                                                                                                        |
| HRP-TM        | EcoRI-ss-HA-ApaI-HRP-SacII-myc-TM-Stop-NotI<br>세포외막 발현  | CMV/pDisplay | The HRP gene was a gift from Frances Arnold (Caltech)                                                                                                               |
| HRP-KDEL      | EcoRI-ss-BglIII-HRP-V5-KDEL-Stop-NotI<br>소포체 내막 발현      | CMV/pDisplay |                                                                                                                                                                     |

[0078] 상기 폴리뉴클레오타이드를 세포에 도입하는 단계는 형질 전환 (Transfection) 또는 형질 도입 (Transduction) 을 통해 매개될 수 있으나, 이에 제한되지 않는다. 또한 위의 APEX 폴리뉴클레오타이드 자리에 퍼옥시테이즈 활성이 더욱 향상된 APEX2의 폴리뉴클레오타이드 (서열번호 3의 아미노산 및 서열번호 4의 염기서열)를 사용할수 있으나 이에 제한되지 않는다.

[0079] 상기 페놀 화합물 및 과산화수소를 처리하는 단계에서, 상기 화합물의 처리는, 상기 화학식 1의 구조를 가지는 화합물을 0.1  $\mu$ M 내지 1 mM의 농도로 10 내지 60분 동안 처리하는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0080] 상기 페놀 화합물 및 과산화수소를 처리하는 단계에서, 과산화효소는 대두 아스코르베이트 과산화효소(Soybean ascorbate peroxidase), 애기장대 아스코르베이트 과산화효소 (Arabidopsis ascorbate peroxidase), 시금치 아스코르베이트 과산화효소 (Spinach ascorbate peroxidase), 사카로미세스 세레비시에 시토크롬 c 과산화효소 (Saccharomyces cerevisiae cytochrome c peroxidase), 파네로카에테 크리소스포리움 과산화효소 PCL1 (Phanerochaete chrysosporium PCL1), 파네로카에테 크리소스포리움 과산화효소 PCM (Phanerochaete chrysosporium PCM), 코프리놉시스 시네레아 과산화효소 (Coprinopsis cinerea peroxidase), 땅콩 과산화효소 (Peanut Peroxidase), 애기장대 과산화효소 ATPA2 (Arabidopsis ATPA2), 서열번호 1의 아미노산 서열로 이루어진 효소, 서열번호 3의 아미노산 서열로 이루어진 효소 및 서열번호 5의 아미노산 서열로 이루어진 효소로 이루어진 군에서 선택된 1 이상의 효소일 수 있고, 상기 과산화수소의 처리는 과산화수소를 0.1 내지 2 mM의 농도로 1 내지 10분 동안 처리하는 것일 수 있으나, 이에 제한되는 것은 아니다.

[0081] 상기 세포는 제한이 없고, 모든 동물 세포, 식물 세포, 미생물 등을 포함할 수 있다. 예를 들어, 형질 감염 (transfection)이 가능한 동물 세포 (Mammalian cell)일 수 있다. 구체적인 예로, 신장 세포, 암 세포, 피부 세포, 난소 세포, 활막세포, 말초혈액단핵구 세포, 섬유아세포, 섬유세포, 신경세포, 상피세포, 각질세포, 조혈 세포, 멜라닌 세포, 연골세포, 대식세포, 근육세포, 혈액 세포, 골수 세포, 림프구 세포, 단핵세포, 폐 세포, 췌장 세포, 간 세포, 위 세포, 장 세포, 심장 세포, 방광 세포, 요도 세포, 배아 생식세포 또는 난구세포 등을 포함할 수 있으나, 이에 제한되는 것은 아니다. 더욱 구체적으로, 상기 동물 세포는 신장세포 (HEK293T 세포, HEK293 세포 등), 암 세포 (HeLa 세포 등), 원숭이 콩팥 세포 (COS-7 세포 등), 골육종 세포 (U2OS 세포 등), 피부세포 (NIH3T3 세포 등), 난소 세포 (CHO 세포 등) 등일 수 있으나, 이에 제한되지 않는다.

[0082] 본 발명에서 과산화효소를 세포에 주입하고 페놀 화합물 및 과산화수소를 처리하면, 세포 안에서 발현된 상기 효소의 작용으로 페놀가 페녹시 라디칼로 전환되고, 상기 페녹시 라디칼은 반경 약 5 내지 50 nm 내에 존재하는

단백질들의 존재하는 타이로신 (Tyrosine)기와 공유결합을 형성함으로써 상기 단백질들에 결합하게 된다. 또한, 생성된 페녹시 라디칼은 인지질 막을 건널 수 없을 만큼 짧은 시간 동안만 존재하기 때문에, 막으로 둘러싸인 각각의 세포 소기관내에서 페녹시 라디칼이 생성되는 경우 각 세포 소기관내의 고유의 단백질들에만 페놀 화합물이 결합할 수 있게 된다.

[0083] 상기 화합물은 과산화효소에 의해서 라디칼 (Radical)을 생성하는 것이고, 상기 생성된 라디칼은 단백질 또는 펩타이드에 존재하는 타이로신 (Tyrosine)기와 공유결합을 형성하여 상기 단백질 또는 펩타이드에 결합하는 것이며, 상기 라디칼이 단백질 또는 펩타이드에 결합하는 거리는 5 내지 50 nm일 수 있으나, 이에 제한되는 것은 아니다.

[0084] 각 세포 소기관에는 고유의 단백질체가 존재하기 때문에, 상기와 같이 데스티오비오틴이 결합되는 단백질들은 세포 소기관마다 상이하고 과산화효소를 발현시킨 공간에 존재하는 단백질들이 표지된다. 상기 표지된 세포를 용해 후, 데스티오비오틴과 결합할 수 있는 스트렙타비딘 (Streptavidin; SA)-비드에 상기 표지된 단백질들을 붙여두고 트립신 처리를 하여 표지되지 않은 펩타이드들을 잘라낼 수 있다. 그리고 나서, 포름아마이드 (Formamide)를 비드에 결합된 표지된 단백질에 처리하고 가열하여 표지된 펩타이드를 얻어내며, 이를 질량분석하여 그 단백질을 동정하고 그 구조 정보를 확인할 수 있다.

[0085] 상기 세포 소기관 (Organelle)은 세포의 내부에 존재하며 세포의 여러 가지 기능을 분업으로 하고 있는 구조단위를 의미한다. 바람직하게, 본 발명에 있어서 세포 소기관은 세포 내막에 싸인 구조체로, 다른 공간과는 구획한 세포 내 구획을 형성하고 있으며 고유의 미세환경을 갖는 소기관을 의미한다. 이러한 세포 소기관은 고유의 기능을 가지며, 예를 들어, 미토콘드리아, 소포체, 핵, 세포질, 골지체, 리소솜, 파고솜, 퍼옥시솜 등 일 수 있으나 이에 제한되지 않으며, 식물세포의 경우 엽록체, 액포 등을 포함할 수 있다.

[0086] 상기 페놀 화합물로 세포 내 단백질을 표지화 하는 단계 이후에 세포를 파쇄하여 세포 파쇄물을 수득하고, 상기 세포의 파쇄물에 스트렙타비딘 (Streptavidin: SA) 비드를 첨가하여, 상기 스트렙타비딘 비드에 결합된 표지 단백질을 수득하는 단계를 추가로 포함할 수 있다.

[0087] 상기 스트렙타비딘 비드에 결합된 표지 단백질을 수득하는 단계에서 상기 세포 파쇄물에 스트렙타비딘 비드를 첨가하면 페놀 화합물, 구체적으로는 DBP가 결합된 표지 단백질이 스트렙타비딘 비드와 결합할 수 있다.

[0088] 본 발명은 상기 페놀 화합물을 이용하여 단백질을 표지하는 방법을 제공하고, 상기 표지된 단백질을 확인하여 단백질을 탐지하는 방법을 제공한다.

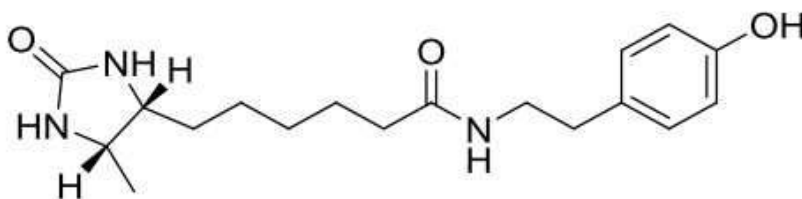
[0089] 구체적으로, 탐지를 위한 세포 내 목적 단백질이 존재하는 세포 소기관으로 표적화 서열을 코딩하는 유전자와 과산화효소를 코딩하는 유전자를 포함하는 폴리뉴클레오타이드를 세포에 도입하는 단계;

[0090] 상기 세포에 하기 화학식 1의 구조를 갖는 페놀 화합물 및 과산화수소를 처리하여 상기 페놀 화합물로 세포 내 단백질을 표지화하는 단계;

[0091] 상기 세포의 파쇄물에 스트렙타비딘 (Streptavidin: SA) 비드를 첨가하여, 상기 스트렙타비딘 비드에 결합된 표지 단백질을 수득하는 단계; 및

[0092] 상기 수득물에서 목적 단백질을 확인하는 단계를 포함하는 단백질의 탐지 방법을 통해서 표지된 목적 단백질을 탐지할 수 있다.

[0093] [화학식 1]



[0094]

[0095] 상기 폴리뉴클레오타이드를 세포에 도입하는 단계 이전에 상기 폴리뉴클레오타이드를 제조하는 과정을 추가로 포함할 수 있다.

[0096] 상기 폴리뉴클레오타이드를 제조하는 과정은 상기 목적 단백질이 존재하는 세포 소기관으로 표적화 서열을 코딩

하는 유전자와 과산화효소를 코딩하는 유전자가 연결되도록 융합된 폴리뉴클레오타이드를 제조한 것을 의미할 수 있으나, 이에 제한되는 것은 아니다.

- [0097] 상기 폴리뉴클레오타이드를 세포에 도입하는 단계에서, 상기 세포 소기관은 미토콘드리아, 소포체, 핵, 세포질 및 골지체로 이루어진 군에서 선택되는 것일 수 있으며, 상기 세포 소기관으로 표적화 서열을 코딩하는 유전자에는 미토콘드리아 기질 발현 유전자, 세포질 발현 유전자, 미토콘드리아 이중막 사이 공간 발현 유전자, 핵 발현 유전자, 세포 외막 발현 유전자, 소포체 내막 발현 유전자, 소포체 외막 발현 유전자, 또는 세포 내막 발현 유전자일 수 있으나, 이에 제한되는 것은 아니다. 상기 세포 소기관으로 표적화 서열을 코딩하는 유전자는 상기 표 1에 나타나 있다.
- [0098] 상기 폴리뉴클레오타이드를 세포에 도입하는 단계는 형질 전환 (Transfection) 또는 형질 도입 (Transduction) 을 통해 매개될 수 있으나, 이에 제한되지 않는다.
- [0099] 상기 페놀 화합물 및 과산화수소의 처리에서, 상기 페놀 화합물의 처리는 페놀 화합물을 0.1 내지 1 mM의 농도로 10 내지 60분 동안 처리하는 것일 수 있으나, 이에 제한되는 것은 아니다.
- [0100] 상기 화합물은 과산화효소에 의해서 라디칼 (Radical)을 생성하는 것이고, 상기 생성된 라디칼은 단백질 또는 펩타이드에 존재하는 타이로신 (Tyrosine)기와 공유결합을 형성하여 상기 단백질 또는 펩타이드에 결합하는 것이며, 상기 라디칼이 단백질 또는 펩타이드에 결합하는 거리는 5 내지 50 nm일 수 있으나, 이에 제한되는 것은 아니다.
- [0101] 상기 과산화효소는 대두 아스코르베이트 과산화효소(Soybean ascorbate peroxidase), 애기장대 아스코르베이트 과산화효소 (Arabidopsis ascorbate peroxidase), 시금치 아스코르베이트 과산화효소 (Spinach ascorbate peroxidase), 사카로미세스 세레비시에 시토크롬 c 과산화효소 (Saccharomyces cerevisiae cytochrome c peroxidase), 파네로카에테 크리소스포리움 과산화효소 PCL1 (Phanerochaete chrysosporium PCL1), 파네로카에테 크리소스포리움 과산화효소 PCM (Phanerochaete chrysosporium PCM), 코프리놉시스 시네레아 과산화효소 (Coprinopsis cinerea peroxidase), 땅콩 과산화효소 (Peanut Peroxidase), 애기장대 과산화효소 ATPA2 (Arabidopsis ATPA2), 서열번호 1의 아미노산 서열로 이루어진 효소, 서열번호 3의 아미노산 서열로 이루어진 효소 및 서열번호 5의 아미노산 서열로 이루어진 효소로 이루어진 군에서 선택된 1 이상의 효소일 수 있으나, 이에 제한되는 것은 아니다.
- [0102] 상기 세포 소기관은 미토콘드리아, 소포체, 핵, 세포질 및 골지체로 이루어진 군에서 선택되는 것일 수 있으나, 이에 제한되는 것은 아니다.
- [0103] 상기 페놀 화합물로 세포 내 단백질을 표지화 하는 단계 이후에 세포를 파쇄하여 세포 파쇄물을 수득하고, 상기 세포 파쇄물에 스트렙타아비딘 (Streptavidin: SA) 비드를 넣어 표지된 단백질을 수득할 수 있다.
- [0104] 상기 스트렙타비딘 비드에 결합된 표지 단백질을 수득하는 단계에서 상기 세포 파쇄물에 스트렙타비딘 비드를 첨가하면 페놀 화합물, 구체적으로는 DBP가 결합된 표지 단백질이 스트렙타비딘 비드와 결합할 수 있다.
- [0105] 상기 상기 수득물에서 목적 단백질을 확인하는 단계는, 상기 스트렙타비딘 비드에 결합된 표지 단백질에 효소를 처리하여 단백질을 절단하고, 상기 스트렙타비딘 비드에 결합된 절단 단편을 분리하여 분리된 절단 단편을 얻고, 상기 분리된 절단 단편을 분석하는 단계로 수행되는 것일 수 있으나, 이에 제한되는 것은 아니다.
- [0106] 상기 절단 단편을 분리하는 것은 탈염처리 (desalting)하여 수행하는 것일 수 있으나, 이에 제한되는 것은 아니다.
- [0107] 상기 효소는 트립신 (trypsin)일 수 있으나, 이에 제한되는 것은 아니다.
- [0108] 상기 폴리뉴클레오타이드를 세포에 도입하는 단계에 앞서, 목적 단백질의 세포내 위치를 분석하는 단계를 수행하는 것일 수 있다.
- [0109] 상기 목적 단백질의 세포내 위치를 분석하는 단계는 과산화효소를 이용하여 세포 내 단백질 위치를 확인하는 방법을 사용할 수 있으며, 이는 대한민국 공개특허공보 제2016-0071960호에 자세히 기재되어 있다.
- [0110] 상기 스트렙타아비딘 비드는 스트렙타아비딘 코팅된 자성 비드일 수 있으나, 이에 제한되는 것은 아니다.
- [0111] 상기 수득물에서 목적 단백질을 확인하는 단계에서 목적 단백질을 확인하는 방법은 질량 분석, 웨스턴 블롯, 형광 현미경 분석, 닷 블롯, 또는 ELISA일 수 있으나, 이에 제한되는 것은 아니다.

## 발명의 효과

[0112] 본 발명은 페놀 화합물을 이용하여 단백질을 탐지하는 방법에 관한 것으로서, 구체적으로 본 발명은 데스티오비오틴-페놀 (Desthiobiotin-Phenol; DBP)을 이용하여 단백질을 탐지하는 방법에 관한 것이며, 본 발명의 단백질을 탐지하는 방법을 통해 표지된 단백질을 확인할 수 있다.

## 도면의 간단한 설명

[0113] 도 1은 본 발명의 일 실시예에 따른 데스티오비오틴-페놀 (Desthiobiotin-Phenol; DBP)을 제조하기 위한 반응식이다.

도 2는 본 발명의 일 실시예에 따른 비오틴-페놀과 데스티오비오틴-페놀로 표지하여 펩타이드를 검출하였을 때, 비오틴-페놀은 산화된 비오틴-페놀이 검출되어 질량분석에서 더 약한 신호 세기를 보여주는 것을 나타내는 도식도이다.

도 3의 왼쪽은 본 발명의 일 실시예에 따른 데스티오비오틴-페놀에 의해 표지되는 자리를 질량분석 신호 세기에 따라 다양한 색으로 표시한 것이고, 오른쪽은 단백질 구조 및 질량분석 신호 세기와 용제 접근 면적의 상관관계 그래프를 나타낸 것이다.

도 4는 본 발명의 일 실시예에 따라 비오틴-페놀, 산화된 비오틴-페놀 및 데스티오비오틴-페놀로 표지된 펩타이드의 MS/MS 스펙트럼 결과를 보여주는 그래프이다.

도 5는 본 발명의 일 실시예에 따른 표지된 펩타이드의 펩타이드 점수(Peptide score) (Sequest X corr value)를 나타내는 표이다.

도 6은 본 발명의 일 실시예에 따라 비오틴-페놀로 표지한 APEX가 트랜스펙션된 전체 세포 용해물의 Streptavidin (SA)-HRP 웨스턴 블랏 (WB) 결과를 보여주는 사진이다.

도 7은 본 발명의 일 실시예에 따라 데스티오비오틴-페놀로 표지한 APEX가 트랜스펙션된 전체 세포 용해물의 Streptavidin(SA)-HRP 웨스턴 블랏 (WB) 결과를 보여주는 사진이다.

도 8은 본 발명의 일 실시예에 따른 DBP 또는 BP로 표지한 단백질의 Sco1-APEX2를 사용한 SA-비드 농축(enrichment)을 확인한 것이다.

도 9는 본 발명의 일 실시예에 따른 DBP 또는 BP로 표지된 단백질의 Sco1-APEX2를 사용한 이미징 어세이(Imaging assay) 결과이다.

도 10은 본 발명의 일 실시예에 따른 mitochondrial matrix-APEX2, Lactb-APEX2 또는 Sco1-APEX2가 트랜스펙션된 세포, 및 트랜스펙션되지 않은 세포에서의 215개의 DBP에 의해 표지된 자리를 나타내는 그래프이다.

도 11은 본 발명의 일 실시예에 따른 그룹 I 내지 V의 미토콘드리아 특이성을 보여주는 그래프이다.

도 12는 본 발명의 일 실시예에 따른 동정된 단백질과 기존의 알려진 단백질의 숫자를 나타낸 것이다.

도 13은 본 발명의 일 실시예에 따른 Lactb-APEX2 또는 Sco1-APEX2에 의해 표지된 단백질을 보여주는 그림이다.

도 14는 본 발명의 일 실시예에 CYP51A1의 결정 구조 및 이의 DBP에 의해 표지되는 티로신 잔기(녹색) 및 헴(heme, 빨간색)을 보여주는 그림이다.

도 15는 본 발명의 일 실시예에 따른 TMEM261-V5-APEX2를 트랜스펙션한 세포를 DBP로 표지하여 관찰한 결과이다.

도 16은 본 발명의 일 실시예에 따른 미토콘드리아에서의 TMEM261의 막 위상(Membrane topology)을 보여주는 그림이다.

도 17은 본 발명의 일 실시예에 따른 미토피린(Mitofilin; IMMT) 단백질 위상을 보여주는 그림이다.

도 18은 본 발명의 일 실시예에 따른 IMM-TM 단백질의 위상을 보여주는 그림이다.

도 19는 본 발명의 일 실시예에 따라 표지된 사이트 결과에 따라 결정된 IMM-TM 위상을 보여주는 그림이다.

도 20은 본 발명의 일 실시예에 따라 표지된 사이트 결과에 따라 결정된 LETM1 및 HIGD2A의 막 위상을 보여주는 그림이다. matrix-APEX2에 의해 DBP 표지된 자리는 녹색 동그라미로 표시하였으며, IMS-APEX2에 의해 DBP 표지

된 자리는 파란 동그라미로 표시하였다.

도 21은 본 발명의 일 실시예의 결과에 따른 주변 막 단백질 바인딩 face를 보여주는 그림이다. matrix-APEX2에 의해 DBP 표지된 자리는 녹색 동그라미, IMS-APEX2에 의해 DBP 표지된 자리는 파란 동그라미로 표시하였다.

도 22는 본 발명의 일 실시예에 따른 IMM에서의 표지된 사이트에 의거한 몇몇 표면 막 단백질의 막 바인딩 face를 보여주는 그림이다. matrix-APEX2에 의해 DBP 표지된 자리는 녹색 동그라미, IMS-APEX2에 의해 DBP 표지된 자리는 파란 동그라미로 표시하였다.

도 23은 본 발명의 일 실시예에 따라 DBP로 표지된 자리를 mitochondrial ATP synthase complex (PDB ID: 2WSS)에 맵핑한 그림이다. matrix-APEX2에 의해 DBP 표지된 티로신 잔기(녹색) 및 IMS-APEX2에 의해 DBP 표지된 자리 (파란색)으로 표시하였다.

도 24는 본 발명의 일 실시예에 따라 DBP로 표지된 자리를 mitochondrial complex II (PDB ID: 1BCC) (왼쪽) 및 complex IV (PDB ID 1OCC) (오른쪽)에 맵핑한 그림이다. matrix-APEX2에 의해 DBP 표지된 티로신 잔기 (녹색) 및 IMS-APEX2에 의해 DBP 표지된 자리 (파란색)를 표시하였다.

도 25는 본 발명의 일 실시예에 따라 DBP로 표지된 자리를 mitochondrial chaperone complex (right: PDB ID: 4PJ1)에 맵핑한 그림이다. matrix-APEX2에 의해 DBP 표지된 티로신 잔기 (녹색)를 표시하였다.

도 26은 본 발명의 일 실시예에 따라 합성한 DBP에 대하여  $^1\text{H}$ -NMR를 수행한 결과를 나타낸다.

도 27은 본 발명의 일 실시예에 따라 합성한 DBP에 대하여  $^{13}\text{C}$ -NMR를 수행한 결과를 나타낸다.

도 28은 BP (Biotin-Phenol)를 이용하여 단백질 탐지하는 방법을 나타내는 모식도이다.

도 29는 DBP를 이용하여 단백질을 탐지하는 방법을 나타내는 모식도이며, 상기 방법을 이용하여 단백질의 구조 탐색이 가능하다.

도 30은 본 발명의 일 실시예에 따라 DBP에 라디칼이 생성되는 위치 및 상기 라디칼이 전파되는 과정을 나타낸 모식도이다.

### 발명을 실시하기 위한 구체적인 내용

[0114] 이하, 본 발명을 하기의 실시예를 통하여 더욱 상세히 설명한다. 그러나 이들 실시예는 본 발명을 예시하기 위한 것일 뿐이며, 본 발명의 범위가 이들 실시예에 의하여 한정되는 것은 아니다.

#### [0116] 실시예 1. Desthiobiotin-Phenol의 합성 및 정제

[0117] 실온 (25℃)에서 데스티오비오틴 (Desthiobiotin, Sigma, 카탈로그 번호 D1411) 100 mg을 2 mL의 DMSO (dimethyl sulfoxide, Sigma-Aldrich, USA)에 용해시켜 0.233 M의 용액을 만든 후, 1.1 당량의 HATU (2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, 0.257 M)과 2.0 당량의 DIPEA (N,N-diisopropylethylamine, 0.467 M)를 첨가하고 실온에서 10분간 교반하여 혼합 용액을 제조하였다. 그 다음, 상기 혼합 용액에 2.0 당량의 티라민 (Tyramine, 0.467 M)을 첨가하고, 실온에서 하룻밤 동안 교반하여 반응을 완료시켰다. 상기 합성 방법은 도 1에 도식화했다.

[0118] 상기 반응으로 제조된 Desthiobiotin-phenol을 C18 역상컬럼 (Agilent Dynamax 250 X 21.4 mm, microsorb 300-5 C18)을 사용한 HPLC (영린 Semi-prep HPLC YL9100)로 5-95% 그라디언트의 아세토나이트릴/물로 40분간 정제하였다. HPLC 작동은 제조사의 프로토콜에 따라 수행되었고, 그 결과 얻은 화합물은 20분 근처에서 용출되었으며 수율은 약 70%였다.

[0119] HPLC로 정제하여 얻은 Desthiobiotin-phenol를 Bruker 400 MHz NMR 기계를 사용하여 제조사의 프로토콜에 따라서 NMR을 수행하였다.

[0120]  $^1\text{H}$ -NMR에 대한 결과는 도 26 및  $^{13}\text{C}$ -NMR에 대한 결과는 도 27에 나타나 있다. 도 26 및 도 27에 나타난  $^1\text{H}$ -NMR 및  $^{13}\text{C}$ -NMR 결과를 통해서, HPLC로 정제하여 얻은 물질은 Desthiobiotin-phenol임을 알 수 있었다.

[0121]  $^1\text{H}$ -NMR (400 MHz, d4-Methanol) :  $\delta$  1.10 (3H,d,J = 6.5 Hz), 1.30-1.46 (8H,m), 1.57 (2H,m), 2.14(2H,t,J

= 7.5 Hz), 2.68 (2H,t,J = 7.5 Hz), 3.35 (2H,m), 3.77 (1H,m), 3.81 (1H,m), 6.70 (2H, d, J = 8 Hz), 7.02 (2H, d, J = 8 Hz).

<sup>13</sup>C-NMR (100 MHz, d<sub>6</sub>-DMSO): δ 15.508, 25.212, 25.591, 28.710, 29.522, 34.439, 35.371, 40.433, 50.259, 55.016, 115.049, 129.450, 129.564, 155.581, 162.865, 171.963, HMRS m/z (C<sub>18</sub>H<sub>28</sub>N<sub>3</sub>O<sub>3</sub>): theoretical: 334.2131, observed: 334.2136.

## 실시예 2. 단백질 표지화

### 2-1. 형질 전환

ATCC에서 구입한 HEK293T 세포를 10% (v/v) 소태아혈청(Fetal Bovine Serum), 50 units/mL penicillin, 및 50 ug/mL streptomycin이 함유된 MEM (Minimum Essential Medium, Gibco) 배지에 37°C로 16 내지 24시간동안 5% 이하의 CO<sub>2</sub> incubator 에서 배양하였다. 세포는 두 개의 T75 플라스크에서 배양했다.

각기 다른 세포 소기관 (예를 들어, 미토콘드리아 기질 공간 (Mitochondria matrix), 미토콘드리아 이중막 사이 공간 (Intermembrane Space of Mitochondrial, IMS), 미토콘드리아의 외막 공간 (Outer mitochondrial membrane, OMM), 소포체, 핵, 세포질, 골지체 등)에 APEX를 도입하기 위하여, 각 세포 소기관으로 타겟팅하는 아미노산 서열을 코딩하는 유전자와 APEX2의 유전자 서열이 연결된 폴리뉴클레오타이드가 삽입된 벡터를 리포펙타민 2000 (Lipofectamine 2000, Thermo Fisher Scientific) transfection reagent를 사용하여 상기 배양된 HEK293T 세포에 트랜스펙션하였다. 상기 벡터는 타겟팅하고자 하는 세포 소기관에 따라 다른 종류의 것을 사용하였으며, 본 발명에서 사용한 벡터들의 구성 및 특징을 아래 표 2에 정리하였다.

표 2

| 항목 | 명명                                                            | 특징                                             | Promotor/<br>Vector | 세부사항                                                                                    |
|----|---------------------------------------------------------------|------------------------------------------------|---------------------|-----------------------------------------------------------------------------------------|
| 1  | Mito-V5-APEX2<br>(unprocessed: 32kDa, processed: 29kDa)       | NotI-Mito-BamHI-V5-APEX2-Stop-XhoI             | CMV/<br>pCDNA3      | Mito-: MLATRVFSLVGKRAISTSVCVRAH<br>(matrix targeting sequence, Fornuskova et al., 2010) |
| 2  | ScoI-V5-APEX2<br>(unprocessed: 62kDa, processed: 59kDa)       | KpnI-ScoI-BamHI-V5-APEX2-Stop-XhoI             | CMV/<br>pCDNA3      | ScoI (NM_004589)                                                                        |
| 3  | ScoI-V5-APEX2<br>(unprocessed: 62kDa, processed: 59kDa)       | HindIII-ScoI-BamHI-V5-APEX2-Stop-XhoI          | CMV/<br>pCDNA5      | ScoI (NM_004589)                                                                        |
| 4  | Tom20-V5-APEX2<br>(45.7 kDa)                                  | NotI-Tom20-10aa linker-NheI-V5-APEX2-Stop-XhoI | CMV/<br>pCDNA3      | 10aa linker : GGSGDPPVAT<br>Tom20 (NM_014765.2)                                         |
| 5  | TRMT61B-V5-APEX2<br>(81.5 kDa)                                | NotI-TRMT61B-NheI-V5-APEX2-Stop-XhoI           | CMV/<br>pCDNA3      | TRMT61B (NM_017910)                                                                     |
| 6  | ATP5J-V5-APEX2<br>(unprocessed: 41.2 kDa, processed: 37.2kDa) | NotI-ATP5J-BamHI-V5-APEX2-Stop-XhoI            | CMV/<br>pCDNA3      | ATP5J (NM_001685)                                                                       |
| 7  | MGST3-V5-APEX2<br>(45.1 kDa)                                  | NotI-MGST3-BamHI-V5-APEX2-Stop-XhoI            | CMV/<br>pCDNA3      | MGST3 (NM_004528)                                                                       |
| 8  | V5-APEX2-NES<br>(29.8 kDa)                                    | NotI-V5-APEX2-NES-Stop-XhoI                    | CMV/<br>pCDNA3      | NES: LQLPPLRLTLD<br>(nuclear exclusion signal)                                          |

|    |                                                                   |                                                      |                |                                                                                                                                                                |
|----|-------------------------------------------------------------------|------------------------------------------------------|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 9  | LACTB-V5-APEX2<br>(unprocessed: 35 kDa,<br>processed: 29 kDa)     | NotI-LACTB-<br>BamHI-V5-<br>APEX2-Stop-XhoI          | CMV/<br>pCDNA3 | LACTB (1-68):<br>MYRLLSSVTARAAATAGPAWDGGRGAHRRPGLPVLG<br>LGWAGGLGLGLGLALGAKLVVGLRGAVPIQS<br>(IMS targeting signal, Polianskyte et al., 2009)<br>Flag: DYKDDDDK |
| 10 | TMEM261-V5-APEX2<br>(unprocessed: 42.8 kDa)                       | HindIII-TMEM261-<br>BamHI-V5-<br>APEX2-Stop-XhoI     | CMV/<br>pCDNA3 | TMEM261 (NM_033428.1)                                                                                                                                          |
| 11 | Pdk1-V5-APEX2*<br>(unprocessed, 81.5 kDa,<br>processed: 76.0 kDa) | NotI-Pdk1-10aa<br>linker-NheI-V5-APEX<br>2-Stop-XhoI | CMV/<br>pCDNA3 | Pdk1 (NM_002610)                                                                                                                                               |

[0129] 2-2. 세포내 단백질 표지화

[0130] T75 플라스크 두 개에 HEK293T 세포를 배양하고, APEX2를 HEK293T 세포에 트랜스펙션하였다. APEX2를 HEK293T 세포에 트랜스펙션하고 12 내지 24 시간 후, Desthiobiotin-phenol이 250  $\mu$ M의 농도가 되도록 처리하고 37°C에서 30분 동안 배양했다. 그 후, H<sub>2</sub>O<sub>2</sub>를 1 mM의 농도가 되도록 넣고, 1분후 5 mM Trolox (Sigma Aldrich, 제품번호 238813), 10 mM sodium azide, 10 mM sodium ascorbate (이하 3가지 혼합물을 소광물질이라 함)이 들어있는 DPBS (Dulbecco's Phosphate Buffered Saline, Thermo Fisher Scientific)로 3번 씻어주었다. 세포를 20 mL DPBS로 파이펫팅하여 50 mL 튜브에 옮긴 후, 1500g 속도로 5분간 원심분리하여 세포 펠렛을 만들어 상층액은 버린다. 다시 DPBS로 파이펫팅 하여 1.5 mL 마이크로 튜브에 옮긴 후, 1500g 속도에서 5분간 원심분리하여 세포 펠렛을 만들어 상층액은 버린다. 1X 프로테아좀 저해제 각테일 (Protease inhibitor cocktail, Sigma, P8849), PMSF 1 mM (phenylmethylsulfonylfluoride, Biosesang, P1021), 및 소광물질 용액이 첨가된 RIPA 용해 버퍼 (lysis buffer) 3 mL (Biosesang, R2002)로 세포를 모두 용해 (Lysis)시켰다.

[0132] 2-3. 표지된 펩타이드의 수득

[0133] 상기 세포 용해물을 4°C의 온도에서 원심 분리기를 이용하여 20,000 rpm의 속도로 10분간 원심분리하여 상층액만을 수득하였다. 상층액은 Amicon filter (10K 원심필터, Merck Milipore, Amicon Ultra-15) 에 넣고 7500g의 속도로 4°C로 20분간 원심분리하였다. 그 후, PBS를 첨가하여 3회 더 원심분리 하였다. 남은 용액을 1.5 mL 마이크로튜브에 옮겨 PBS를 첨가해 1 mL로 맞추고 스트렙타비딘 (streptavidin; SA) 코팅된 자성 비드 300  $\mu$ L를 넣고 1시간 동안 상온(25°C)에서 천천히 돌려주었다. 마그네틱 스탠드에 꽂아서 상층액은 타겟 단백질이 남김없이 수득 (enrichment)가 잘 되었는지 테스트용으로 사용하기 위해 따로 보관했다. 이후 자성 비드는 PBS로 3회 세척하였다. PBS를 제거하고 100  $\mu$ L의 6 M Urea, 2 M Thiourea, 10 mM Hepes를 넣어 서스펜션 시켰다. 50 mM Ammonium bicarbonate (ABC)에 들어있는 100 mM DTT를 20  $\mu$ L 넣어주고 800rpm 속도로 56°C에서 1시간 원심분리했다. 50 mM ABC에 들어있는 300 mM Iodoacetamide 35  $\mu$ L를 넣고 800rpm 속도로 30분동안 원심분리했다. 50 mM ABC를 840  $\mu$ L 첨가한다. 1  $\mu$ g/ $\mu$ L 트립신 (Promega, V5280) 8  $\mu$ L를 넣어주었다. 37°C에서 6-18시간동안 800rpm 속도로 원심분리했다. 마그네틱 스탠드에 꽂아서 상층액은 모아두고 자성 비드를 PBS로 4회 세척했다. 95% (v/v) 포름아마이드 (Formamide), 10 mM EDTA, pH 8.2 용액을 250  $\mu$ L 첨가하고 95°C에서 10분간 가열했다. 그리고 나서 즉시 마그네틱 스탠드에 꽂은 후, 얼음에 두어 식혔다. 3% (v/v) Acetonitrile /0.1% (v/v) Formic acid 용액을 750  $\mu$ L 첨가했다.

[0135] 2-4. 펩타이드 절단 단편의 분리

[0136] 상기 스트렙타비딘에 결합된 펩타이드 절단 단편을 분리하기 위해서 Varian Bond ELUT (Agilent, 12109301)를 이용하여 하기와 같은 방법으로 디솔팅 (desalting)을 수행하였다.

[0137] 1 mL의 3% (v/v) Acetonitrile /0.1% (v/v) Formic acid, 1 mL의 100% acetonitrile, 그리고 2 mL 3% (v/v) Acetonitrile /0.1% (v/v) Formic acid를 흘려주어 컬럼을 활성화시킨다. 샘플을 컬럼에 천천히 흘려주어 펩타이드 절단 단편을 컬럼에 붙게했다. 2 mL의 3% (v/v) Acetonitrile /0.1% (v/v) Formic acid 용액으로 컬럼을

씻어주었다. 1.4 mL의 70% (v/v) Acetonitrile /0.1% (v/v) Formic acid 용액 과 500 uL의 100% acetonitrile 를 흘려주어 펩타이드를 추출했다.

[0138] 질량 분석을 위해서는 추출된 용액을 완전히 말린 후, 0.1% (v/v) Formic acid에 다시 녹이는 작업을 수행하였다.

[0140] **비교예 1. 비오틴-페놀을 처리한 경우**

[0141] 실시예 1 내지 2와 동일한 방법으로 실험을 수행하되, 실시예 2에서 APEX2를 HEK293T 세포에 트랜스펙션하고 12 내지 24 시간 후 페놀 화합물을 처리하는 단계에서, 테스티오비오틴-페놀 대신 500 uM 농도의 비오틴-페놀을 처리하였다.

[0143] **실험예 1. 질량 분석**

[0144] 본 발명에서 합성하여 사용한 테스티오비오틴-페놀에 의해서 표지된 펩타이드에는 산화된 형태가 존재하지 않음을 확인하기 위해서, 실시예 1 및 2를 수행하여 수득한 펩타이드 절단 단편 및 비교예 1을 수행하여 수득한 펩타이드 절단 단편에 대하여 하기 방법으로 질량분석을 수행하였다.

[0145] 펩타이드 혼합물을 분리하기 위해, 10 내지 24% (v/v) Acetonitrile /0.1% (v/v) Formic acid 구배의 300 nL/min의 유량에서 90분 동안 C18 역상 HPLC 컬럼(150 mm x 75 um)을 이용하였다.

[0146] 그리고 나서, MS/MS 분석을 위해 precursor ion scan MS spectra (m/z 400 내지 2000)를 Orbitrap spectrometer (Thermo Fisher Scientific, Germany)를 사용하였다. Orbitrap의 resolution은 400 m/z을 기준으로 60,000의 분해능 (resolution) 으로 맞추어졌으며 내부적인 internal lock mass 기능을 이용하였다. 20개의 가장 강한 이온을 CID (collisionally induced dissociation, 충돌에 의한 분자 쪼개짐 반응)모드로 분리하고 조각화하였다.

[0147] MS/MS 분석은 Sequest (Thermo Fisher Scientific, San Jose, CA, USA; version 1.4.1.14)와 X! Tandem (The GPM, thegpm.org; version CYCLONE (2010.12.01.1))를 이용하여 수행하였다.

[0148] 구체적으로, Sequest 및 X! Tandem은 호모 사피엔스(Homo sapiens) 단백질 서열 데이터베이스 (89113 entries, UniProt ([www.uniprot.org](http://www.uniprot.org)))을 트립신 효소에 의한 소화를 가정하여 검색하도록 셋팅하였다. Sequest 및 X! Tandem은 0.60 Da의 질량 허용 이온 조각(fragment ion mass tolerance)과 10.0 PPM의 상위 이온 허용 조각(parent ion tolerance)을 검색하였다.

[0149] 카바미도메틸 시스테인 (Carbamidomethyl cysteine)은 고정 수정(fixed modification)으로 Sequest 및 X! Tandem에서 지정하였으며, 하기의 조건에서 수행하였다:

[0150] Glu->pyro-Glu of n-terminus,

[0151] ammonia-loss of n-terminus, 및

[0152] gln->pyro-Glu of n-terminus.

[0153] 메티오닌의 산화 및 BP (delta monoisotopic mass: 361.146), 산화된 BP (delta monoisotopic mass: 377.1409), 티로신의 DBP (delta monoisotopic mass: 331.1896)는 X! Tandem 및 Sequest에서 변수 수정(variable modifications)으로 사용하였다.

[0154] 도 4 및 도 5에 나타난 것과 같이, 실시예 1 및 2의 방법으로 테스티오비오틴-페놀에 의해서 펩타이드를 표지한 경우에는 펩타이드의 티로신기에 테스티오비오틴-페놀 (Y+DBP)가 공유결합을 한 프로브의 mass 값인 M+331 Da만을 검출할 수 있었지만, 비교예 1의 방법으로 테스티오비오틴-페놀 대신 비오틴-페놀을 처리한 경우에는 펩타이드의 티로신기에 비오틴-페놀 (Y+BP)가 공유결합을 한 프로브의 mass 값인 M+361 Da 및 상기 비오틴 페놀이 산화된 형태 (Y+oxidized BP)의 프로브의 mass 값인 M+377 Da이 동시에 검출할 수 있었다. 상기 결과는 도 2에 도식화했다.

[0155] 또한, 상기 질량 분석 결과를 바탕으로 질량분석 신호 세기와 단백질의 용제 접근 면적간의 상관관계를 도 3에 나타냈다. 그 결과, 용제 접근 면적이 증가할수록 질량분석 신호 세기가 증가하는 양의 상관관계가 있음을 확인

할 수 있었다.

[0157] **실험예 2. 표지된 단백질의 확인**

[0158] 본 발명에서 합성하여 사용한 데스티오비오틴-페놀에 의해서 표지된 펩타이드를 확인하기 위해서, 실시예 2-1 내지 2-2를 수행하여 얻은 세포 용해물 및 비교예 1를 수행하여 얻은 세포 용해물에 대하여 하기와 같은 방법으로 웨스턴 블롯을 수행하였다.

[0159] 구체적으로, 실시예 2-2 및 비교예 1를 수행하여 얻은 세포 용해물을 SDS-PAGE를 위해서 샘플링하고, SDS-PAGE 겔에 150V에서 약 1시간 동안 전기영동을 실시하였다. 그 다음, SDS-PAGE 겔에 전기 영동된 단백질을 니트로 셀룰로소막 (Nitrocellulose membrane)으로 옮긴 후, 폰슈 염색(ponceau staining)을 통해 전체 단백질 양을 시각적으로 확인하고, TBST용액 (Tris-Buffered Saline 및 Tween 20의 혼합액)으로 폰슈를 제거하고, 2% (w/v)의 투석된 BSA 용액으로 1시간 동안 상온에서 교반하였다. 2% (w/v)의 투석된 BSA 용액에 SA-HRP를 10000:1로 희석하여 1시간동안 상온에서 교반하였다. 그 다음, 1회당 5분씩 총 4회 TBST로 막 (Membrane)을 씻었다. TBST로 세척한 막을 현상액으로 현상하고, 제조사의 프로토콜에 따라서 LAS 4000 장비 (GE, USA, LAS 4000)를 사용하여 이미지를 얻었다.

[0160] 도 6 및 도 7의 1번은 표 2의 1 (MitoMatrix-APEX2)이고, 2번은 표 2의 9 (LACTB-APEX2 (IMS))이고, 3번은 표 2의 4 (Tom20-APEX2 (OMM))이고, 4번은 표 2의 8 (APEX2-NES)이고, 5번은 표 2의 11 (Pdk1-APEX2 (matrix))이고, 6번은 표 2의 3 (SCO1-APEX2 (IMS))이고, 7번은 표 2의 5 (TRMT61B-APEX2)이고, 8번은 표 2의 6 (ATP5J-APEX2)이고, 9번은 표 2의 7 (MGST3-APEX2)이다.

[0161] 도 7에서 나타난 것과 같이, 각 세포소기관에서 desthiobiotin-phenol에 의해 표지된 단백질의 웨스턴 블롯 패턴을 확인할 수 있었고, 상기 표지된 단백질의 웨스턴 블롯 패턴은 도 6에서 나타난 것과 같이 biotin-phenol에 의해서 표지된 단백질의 웨스턴 블롯 패턴과 유사함을 확인할 수 있었다.

[0163] **실험예 3. 수득한 표지 펩타이드의 확인**

[0164] 실시예 2-3에서 스트렙타아비딘에 결합시켜 수득한 표지 펩타이드를 확인하기 위해서, 실시예 2-3 및 비교예 1을 수행하여 수득한 표지 펩타이드를 실험예 2와 같은 방법으로 웨스턴 블롯을 수행하였다.

[0165] DBP는 실시예 1 및 실시예 2의 방법을 수행한 것이고, BP는 비교예 1을 수행한 것이다. Input은 스트렙타아비딘에 결합하기 전의 샘플, Flow-through는 스트렙타아비딘과 결합시키고 마그네틱 스탠드에 쏘아 가라앉힌 후의 상층액, SA-Enriched는 스트렙타아비딘에 결합한 샘플을 나타낸다.

[0166] 도 8에 나타난 것과 같이, Input에 존재하는 DBP로 표지된 단백질이 스트렙타아비딘 비드에 특이적으로 결합함을 확인할 수 있었고, 상기 결과는 BP로 표지된 단백질이 스트렙타아비딘 비드에 특이적으로 결합하는 웨스턴 블롯 패턴과 유사함을 확인할 수 있었다.

[0168] **실험예 4. 형광 현미경 분석**

[0169] 실시예 1 및 2의 방법으로 DBP로 단백질을 표지한 세포 및 비교예 1의 방법으로 BP로 단백질을 표지한 세포들을 하기와 같은 방법으로 형광 현미경을 사용하여 분석하였다.

[0170] 형광 현미경 분석을 위해, 실시예 1 및 2의 방법으로 DBP로 단백질을 표지한 세포 및 비교예 1의 방법으로 BP로 단백질을 표지한 세포들을 4% (v/v) 파라포름알데하이드 (paraformaldehyde)를 포함한 DPBS 용액으로 실온에서 15분 동안 고정시켰다. 그 다음, 세포들을 DPBS (Dulbecco's Phosphate-Buffered Saline, Gibco)로 3번 세척하고, 4℃에서 보관한 메탄올을 -20℃에서 5분 동안 처리하여 세포를 투과시켰다 (Permeabilized).

[0171] 위와 같이 세포를 투과시킨 후, 세포들을 다시 DPBS로 3번 세척하고, 2% (w/v) BSA (Bovine serum albumin)를 포함하고 있는 DPBS (blocking buffer)로 실온에서 1시간 동안 블로킹하였다.

[0172] APEX2-융합 단백질의 발현을 탐지하기 위해, 세포들은 마우스 항-V5 항체 (Invitrogen, R960-25, 1:5000 희석)와 함께 1시간 동안 실온에서 배양하였다. 그 다음, 세포들을 TBST를 이용하여 5분씩 4회 세척 후, 동시에

이차 Alexa Fluor 488-염소 항-마우스 IgG (Invitrogen, A-11001, 1:1000 희석) 및 streptavidin-Alexa Fluor 568 IgG (Invitrogen, S11226, 1:1000 희석)의 존재 하에서 30분 동안 실온에서 배양하였다. 그 다음, 상기 세포들을 TBST를 사용하여 5분씩 4회 세척하고, EVOS Epi-fluorescence microscope (Life Technology)를 이용한 이미징을 위해 얼음 위에 보관한 DPBS에 놓았다. 형광 현미경을 사용한 이미징은 제조사의 프로토콜에 따라서 수행하였다. 그 결과를, 도 9 및 15에 나타내었다.

[0173] 도 9에서 확인할 수 있듯이, 실시예 1 및 2의 방법으로 DBP로 단백질을 표지한 세포 및 비교예 1의 방법으로 BP로 단백질을 표지한 세포의 형광 이미징한 결과는 큰 차이가 없이 동일했다.

[0174] 도 15는, U2OS 세포에 실시예 2-1 및 2-2와 같은 방법으로 Mito-BFP와 TMEM261-V5-APEX2를 발현시킨 후, DBP로 표지하였다. 녹색은 항-V5로 TMEM261-V5-APEX2의 발현 패턴을 관찰한 것이고, 빨간색은 SA-568, 즉 DBP 표지된 단백질의 패턴을 관찰한 것이고, 파란색은 Mito-BFP, 즉 미토콘드리아 패턴을 관찰한 것이다. 상기 3가지를 오버랩한 것이 merged이며, 녹색과 파란색은 잘 겹쳐지기 때문에 TMEM261-V5-APEX2가 미토콘드리아에 잘 발현되는 것을 의미하고, 빨간색이 잘 겹쳐지지 않는 것은 DBP 표지가 제한된 공간이 아닌 OMM (Outer mitochondrial membrane)이나 IMS (Intermembrane space)에서 일어나는 것을 의미한다. IMM (Inner mitochondrial membrane)과 다르게 OMM에는 pore가 존재하기 때문에 OMM이나 IMS는 matrix와는 다르게 제한되지 않은 공간이다.

#### [0176] 실험예 5. DBP로 표지한 특정 세포소기관 단백질의 탐지

[0177] 본 발명의 방법에 의해서 특정 세포세포 소기관에 존재하는 단백질을 탐지하기 위해서 실시예 1 및 2의 방법을 사용하여 실험을 수행하고, 얻어진 펩타이드 절단을 실험예 1과 같은 방법으로 질량분석을 수행하여 그 결과를 분석하였다.

[0178] 그 결과 도 10 내지 14에서 확인할 수 있는 바와 같이, 미토콘드리아 세포소기관의 이중막 사이공간 (IMS)에 존재하는 LACTB 단백질과 근접거리에 존재하는 세포질에 존재하는 ENO1 단백질이 표지되었음을 확인할 수 있었고, 미토콘드리아 세포소기관 이중막 사이공간에 존재하는 SCO1 단백질과 근접거리에 존재하는 IMS에 존재하는 HIGD1A 단백질이 표지되었음을 확인할 수 있었다. 이는 OMM은 투과 가능 (permeable)하고, IMM은 투과 가능하지 않기 (impermeable) 때문에, LACTB는 OMM에 가까운 위치에 존재하여 세포질에 존재하는 단백질을 DBP로 표지가 가능하지만, SCO1은 IMM 또는 매트릭스 (matrix)에 가까운 위치에 존재함에도 불구하고 매트릭스에 존재하는 단백질을 DBP로 표지가 불가능하기 때문이라고 판단된다.

[0179] 즉, LACTB와 SCO1은 둘다 IMS에 존재하는 단백질이지만, LACTB는 SCO1에 비해서 상대적으로 IMS 공간중에서 세포질쪽에 가까운 위치에 존재하고, SCO1은 LACTB에 비해서 상대적으로 IMS 공간중에서 IMM쪽에 가까운 위치에 때문에, LACTB-APEX2는 세포질에 존재하는 단백질을 표지가 가능하지만, SCO1-APEX2는 세포질에 존재하는 단백질을 표지할 수 없다.

#### [0181] 실험예 6. DBP로 특정 세포소기관 단백질을 표지하여 막방향성 구조 규명

[0182] 비교예 1의 biotin-phenol (BP)을 사용하여 표지한 단백질을 탐지하는 방법 (도 28)은 단백질 상에서 어떤 타이로신 잔기가 biotin-phenol과 반응했는지 알수 없지만, 본 발명의 실시예 1 및 2의 방법으로 desthiobiotin-phenol (DBP)을 사용하여 단백질을 표지하고, 상기 표지한 단백질을 탐지하는 방법 (도 29)은 단백질 상의 어떤 타이로신 잔기가 표지되었는지 질량분석으로 확인할 수 있으므로, 이를 통해서 DBP로 표지된 단백질의 구조적 정보도 확인할 수 있는지 여부를 하기와 같이 확인하였다.

[0183] 구체적으로, 실시예 1 및 2과 같은 방법으로 APEX를 미토콘드리아 이중막 사이공간 (IMS)에 발현시키고 DBP 처리후, 실험예 1과 같은 방법으로 질량분석을 수행하면, 도 17에 나타난 것과 같이 IMMT라고 하는 미토콘드리아 막단백질의 경우 이중막 사이공간에 노출된 단백질 도메인만 DBP로 표지됨을 확인할 수 있었다.

[0184] 또한, APEX를 미토콘드리아 기질공간 (Matrix)에 발현시킨 경우에는, 도 17에 나타난 것과 같이 상기 IMS에 발현시킨 경우와 다르게 미토콘드리아 기질 공간에 노출된 도메인이 표지됨을 확인할 수 있었다.

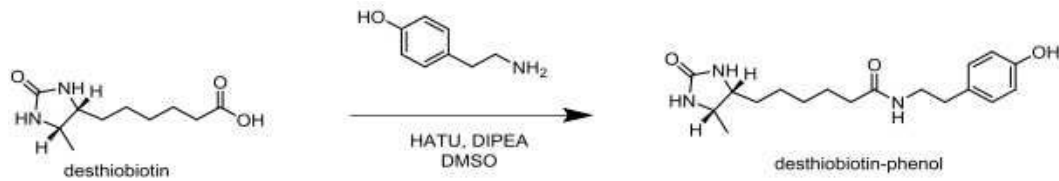
[0185] 상기 방법과 동일하게 미토콘드리아 이중막 사이공간에 발현시킨 SCO1-APEX2 (파란색표시) 및 미토콘드리아 기질공간에 발현시킨 Matrix-APEX2 (초록색 표시)을 이용하여 막방향성을 결정하였고, 그 결과는 도 16, 도 18 내지 22에 나타냈다.

[0186] 본 발명의 DBP를 이용한 특정 세포세포 소기관에 존재하는 단백질의 표지 및 상기 단백질의 확인 방법을 통해서 미토콘드리아 막단백질의 막방향성이 아직 확립되지 않은 다양한 미토콘드리아 막단백질의 막방향성을 명확히 규명할 수 있다.

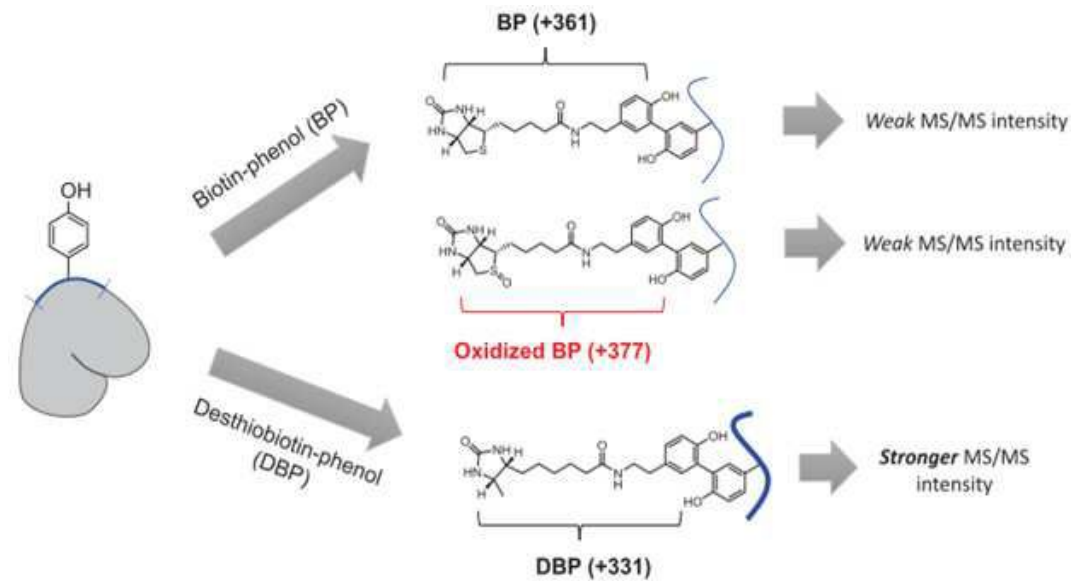
[0187] 상기와 같은 방법으로 DBP로 표지하여 mitochondrial ATP synthase complex (PDB ID: 2WSS)를 맵핑한 결과는 도 23, mitochondrial complex II (PDB ID: 1BCC) 및 complex IV (PDB ID 10CC)를 맵핑한 결과는 도 24, 및 mitochondrial chaperone complex (right: PDB ID: 4PJ1)를 맵핑한 결과를 도 25에 나타냈다.

## 도면

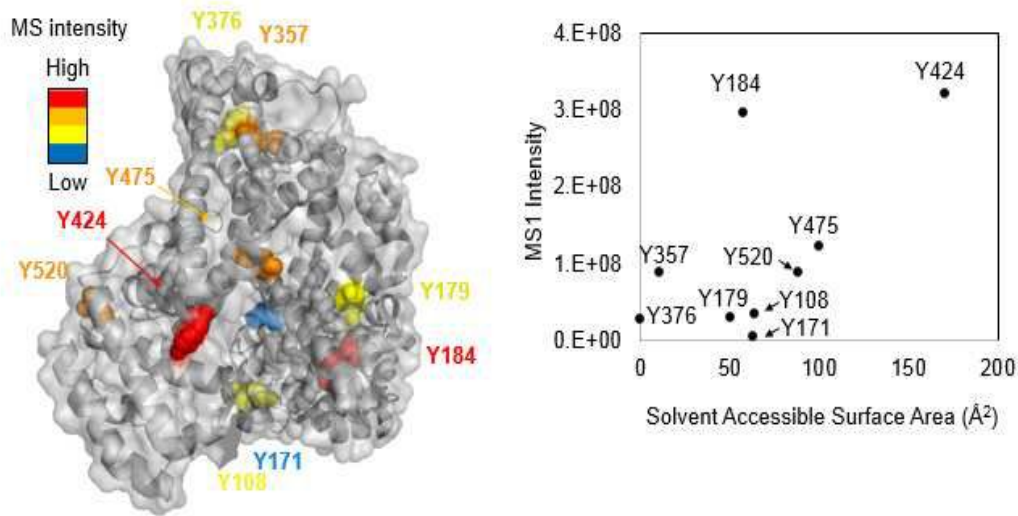
### 도면1



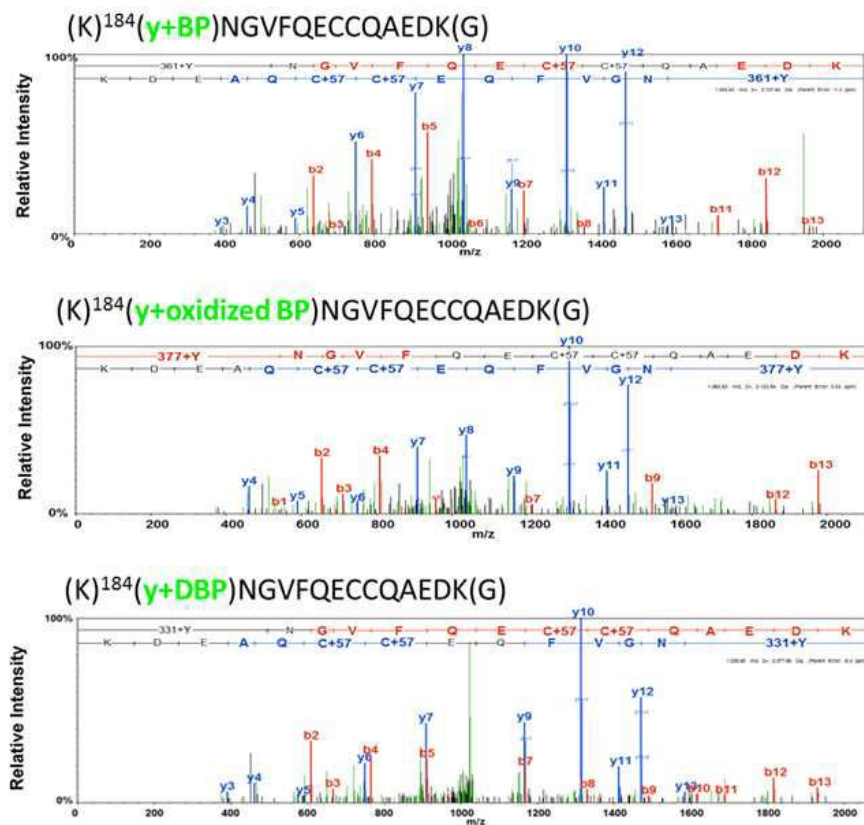
### 도면2



도면3



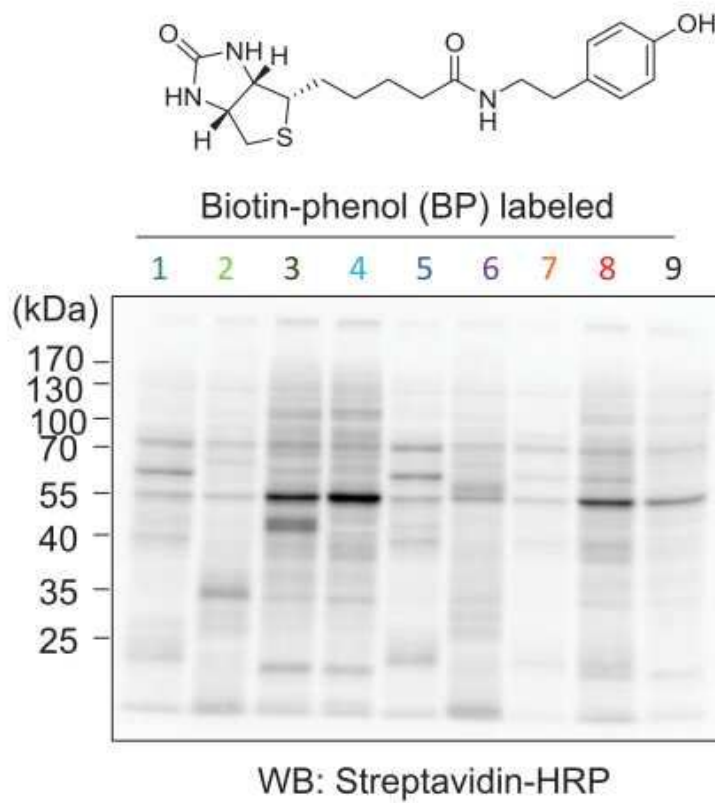
도면4



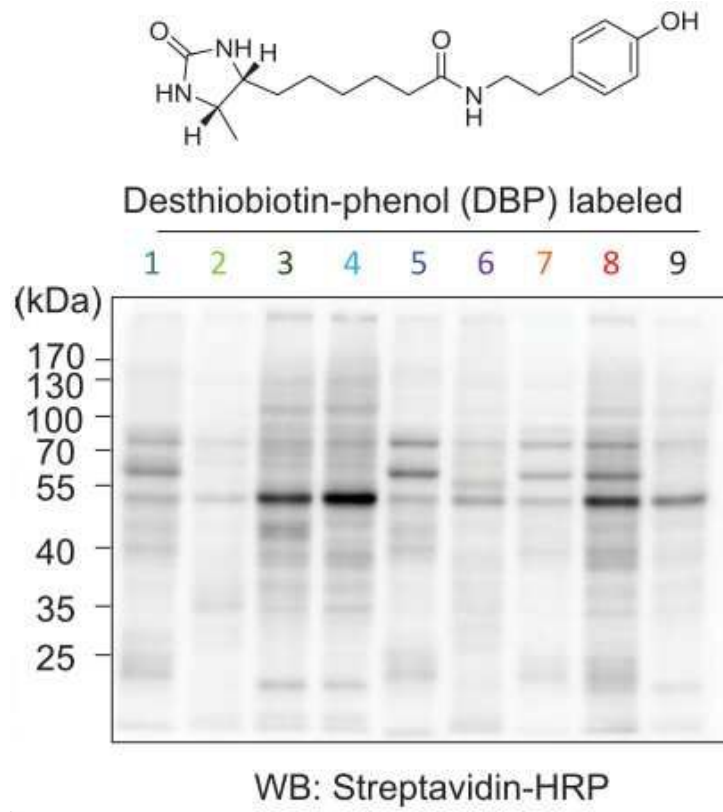
도면5

| Start | Stop | Sequence               | BP-labeled<br>(Y+361) |           | Oxidized<br>BP-labeled<br>(Y+377) | DBP-labeled<br>(Y+331) |           |
|-------|------|------------------------|-----------------------|-----------|-----------------------------------|------------------------|-----------|
|       |      |                        | SEQUEST<br>XCorr      | X! Tandem |                                   | SEQUEST<br>XCorr       | X! Tandem |
| 184   | 197  | (K)yNGVFQECCQAEDK(G)   | 3.56                  | 3.74      | 3.37                              | 4.24                   | 4.79      |
| 421   | 433  | (K)LGEyGFQNALIVR(Y)    | 3.12                  | 3.00      | 3.65                              | 3.73                   | 5.52      |
| 508   | 523  | (R)RPCFSALTPDETyVPK(A) | 2.43                  | 2.67      | 2.84                              | 3.62                   | 3.46      |

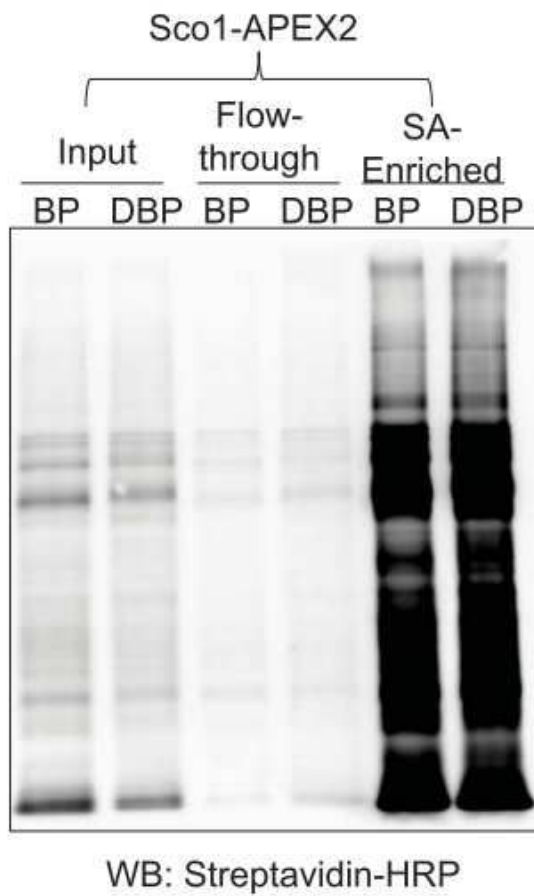
도면6



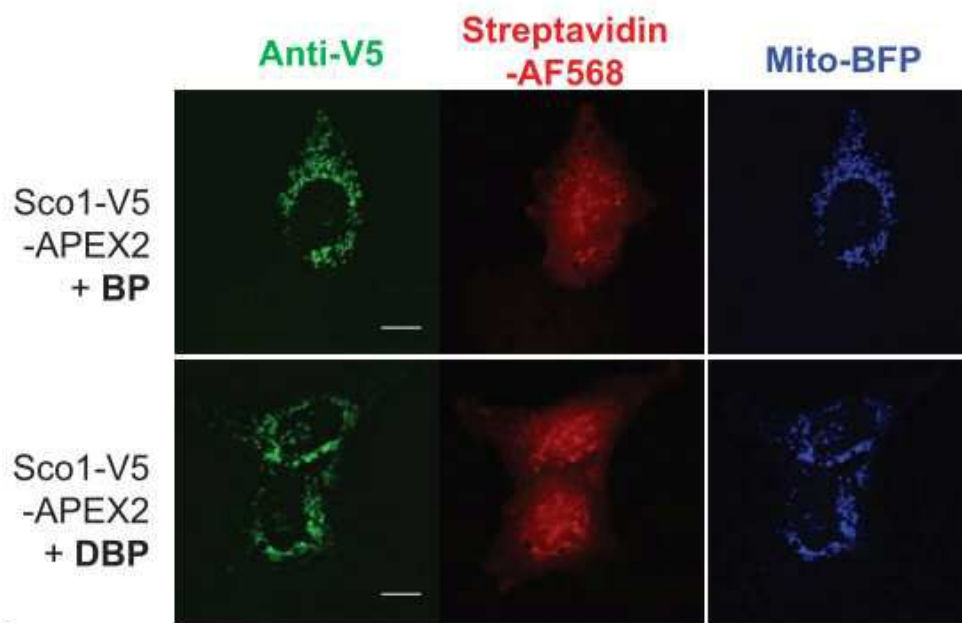
도면7



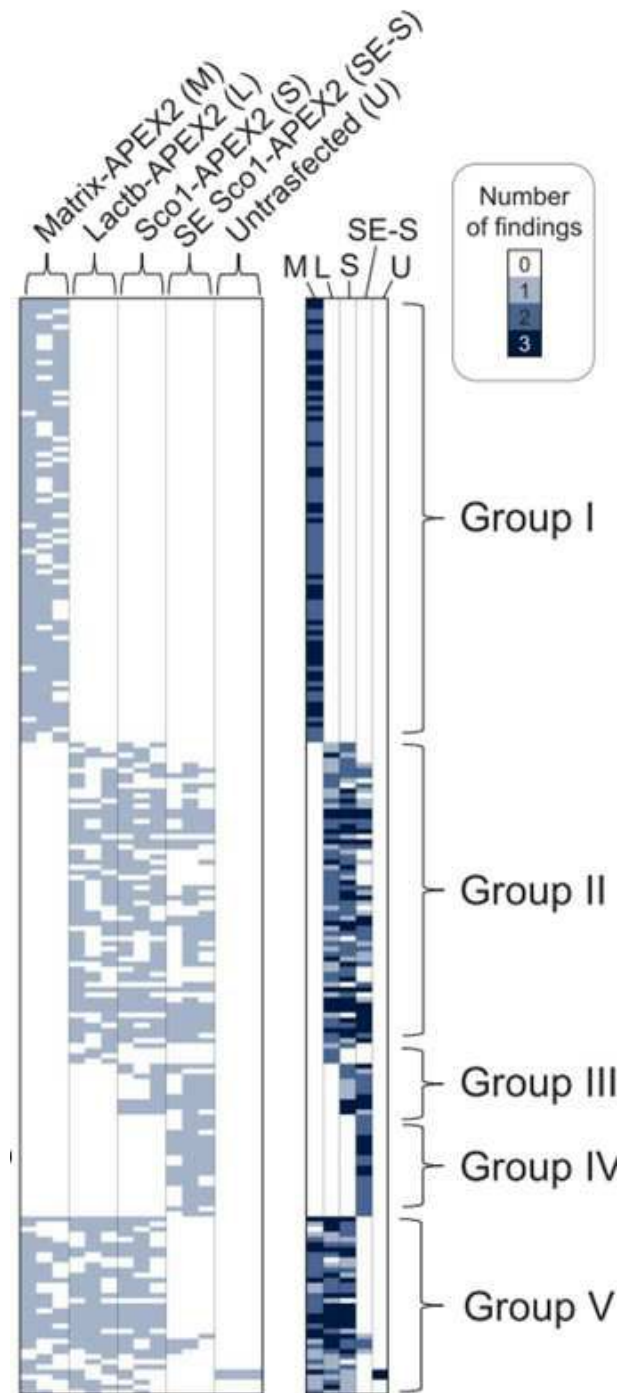
도면8



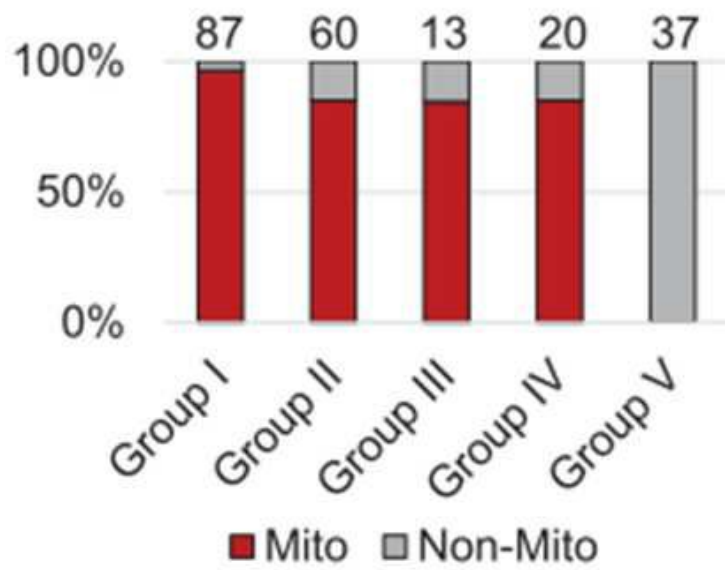
도면9



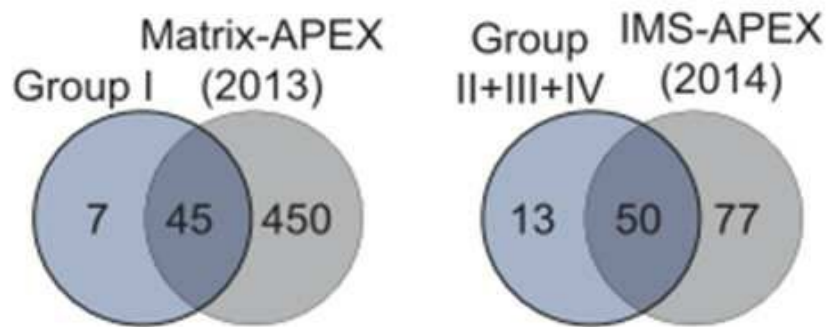
도면10



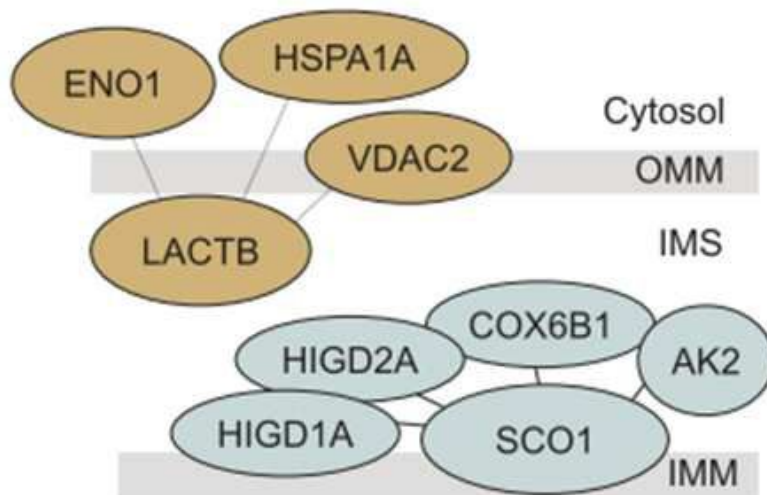
도면11



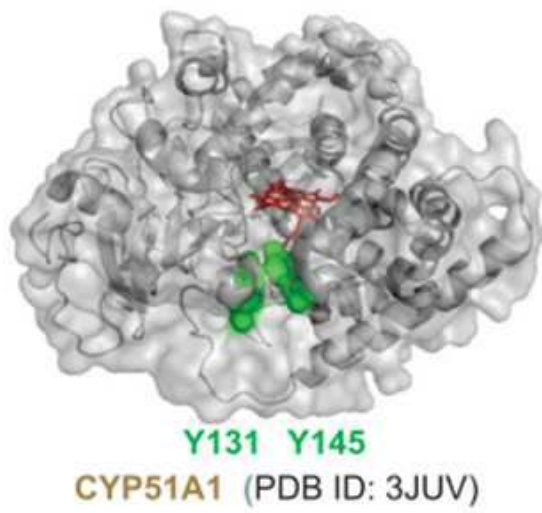
도면12



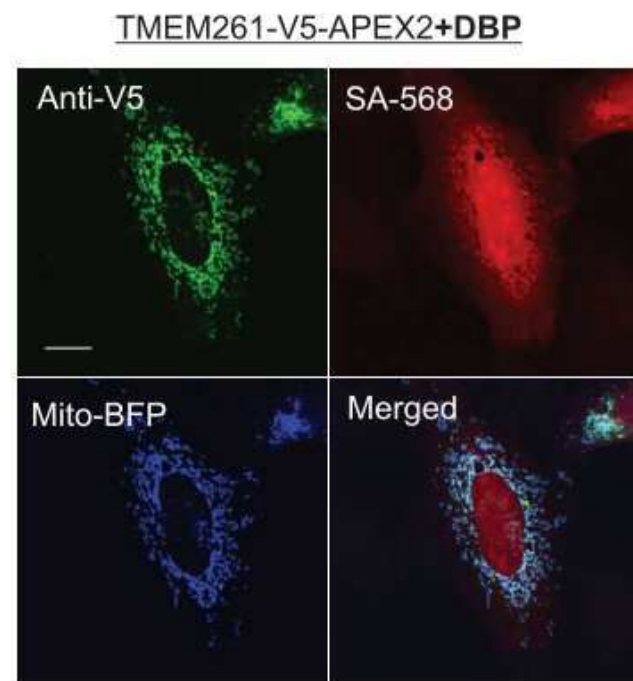
도면13



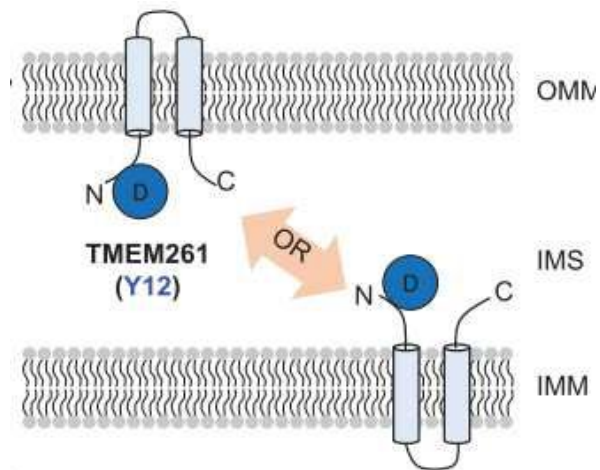
도면14



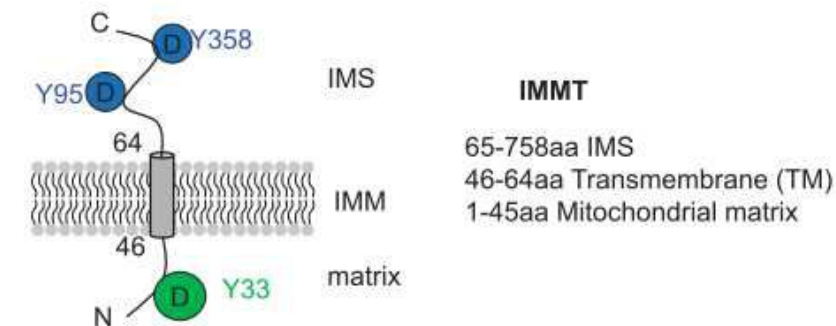
도면15



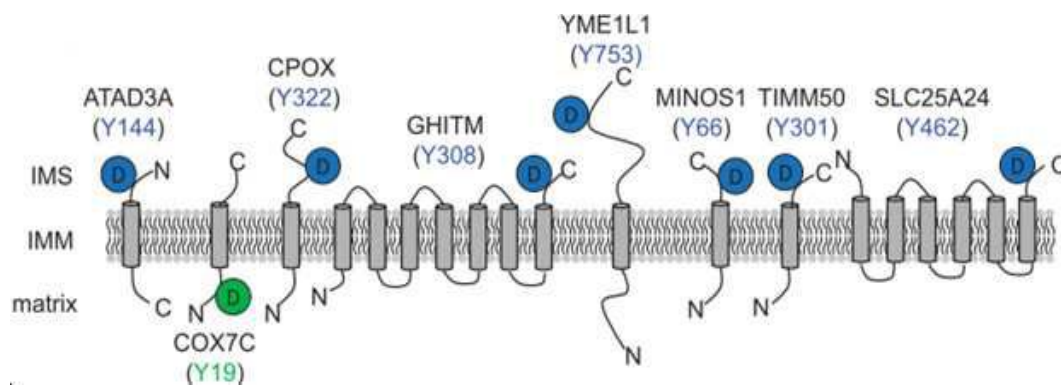
도면16



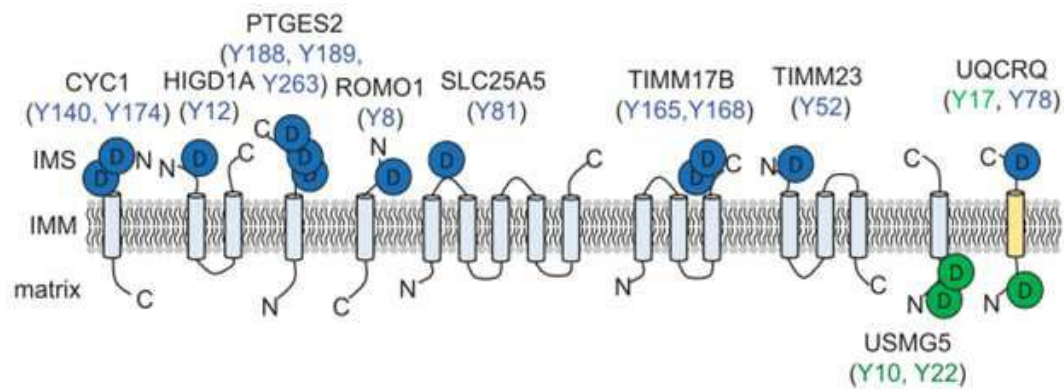
도면17



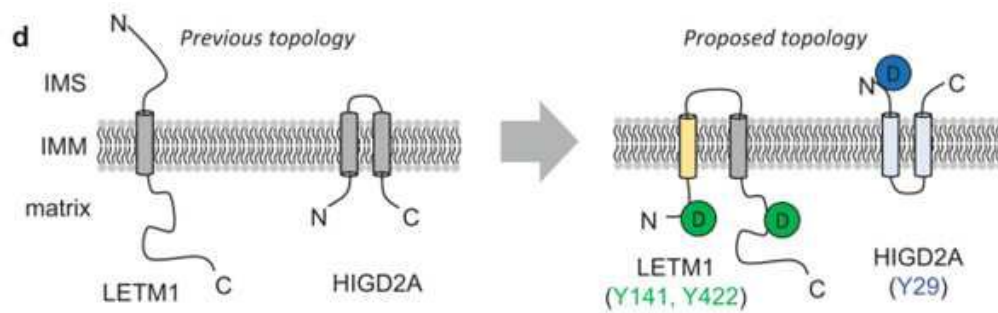
도면18



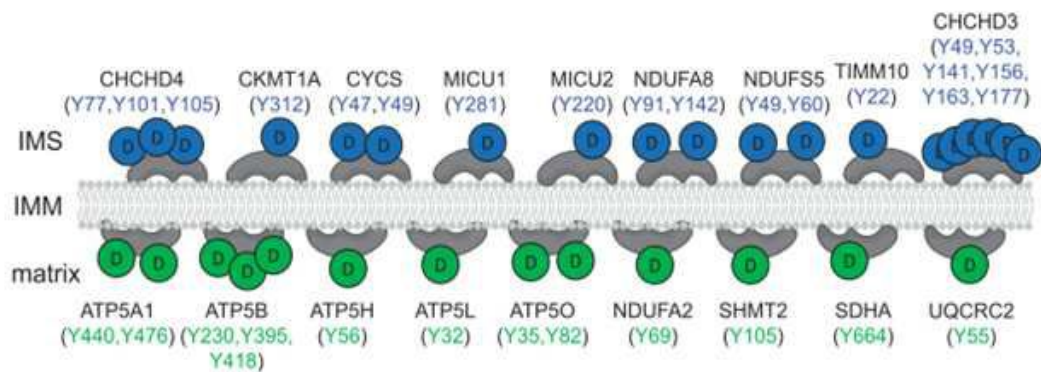
도면19



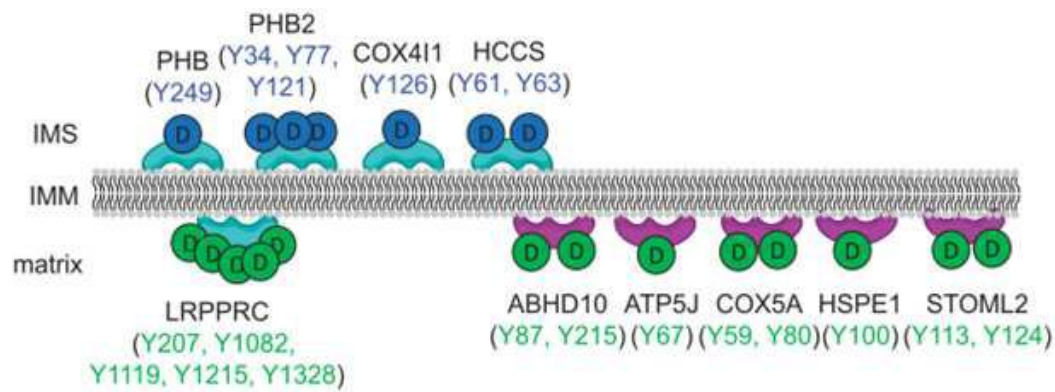
도면20



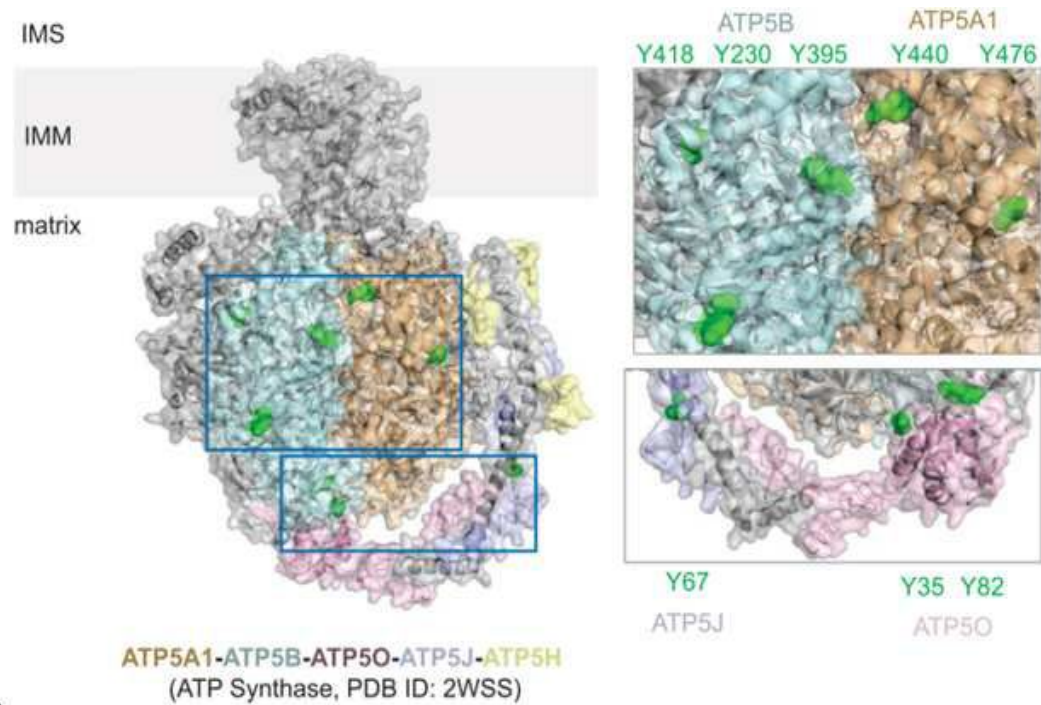
도면21



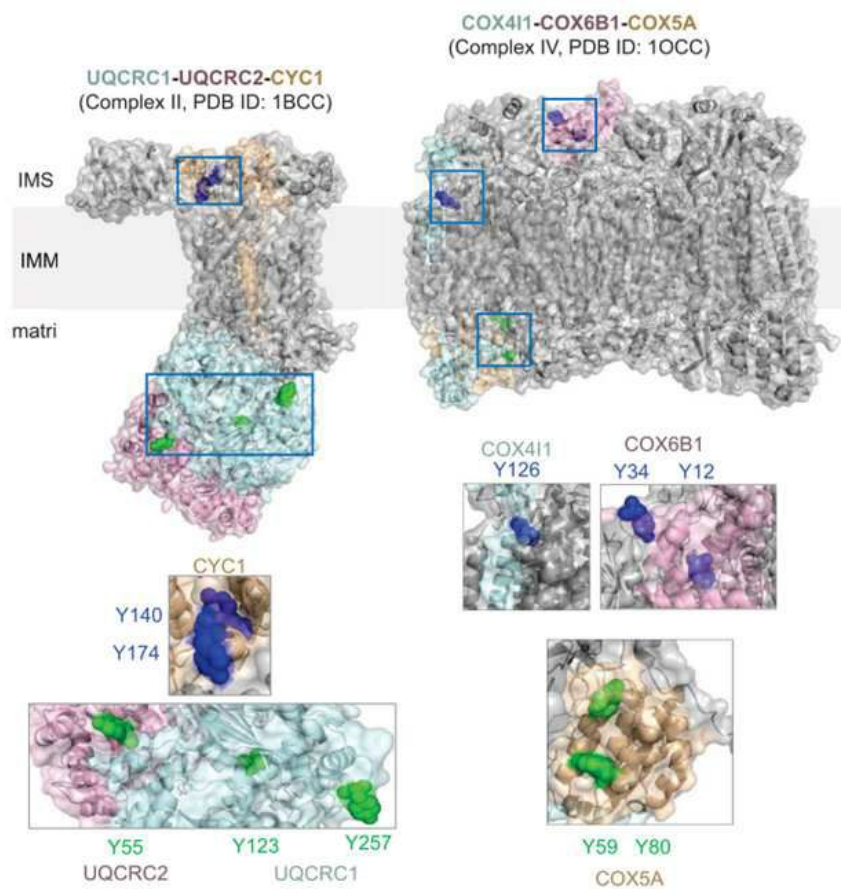
도면22



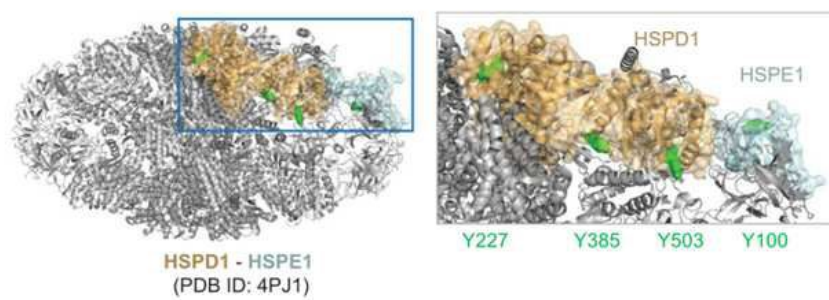
도면23



도면24



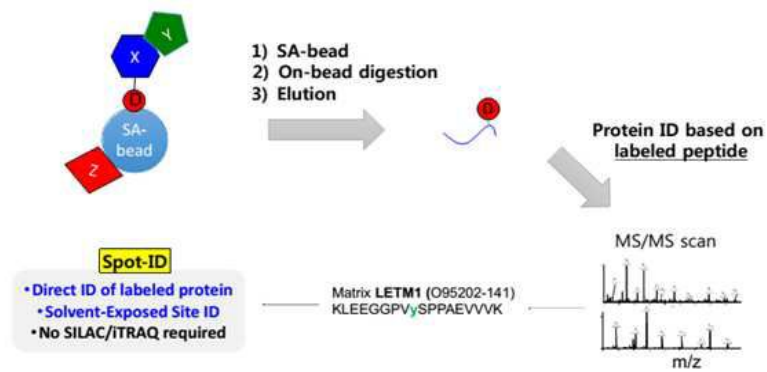
도면25



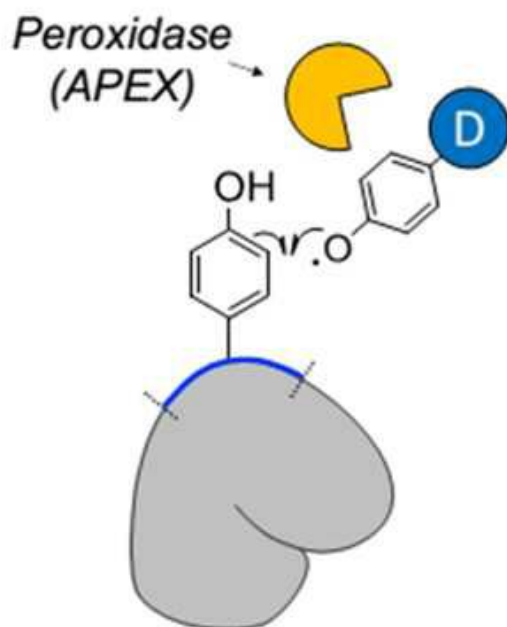




도면29



도면30



## 서열 목록

- <110> UNIST(ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY)
- <120> Methods for identification of proteins using phenolic compounds  
for protein labeling
- <130> DPP20160206KR
- <160> 6
- <170> KoPatent In 3.0
- <210> 1
- <211> 692
- <212> PRT
- <213> Artificial Sequence

<220><223> Engineered ascorbate peroxidase(APEX)

<400> 1

Met Glu Thr Gly Leu Tyr Leu Tyr Ser Ser Glu Arg Thr Tyr Arg Pro  
1 5 10 15

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

20 25 30

Thr Tyr Arg Gly Leu Asn Ala Ser Pro Ala Leu Ala Val Ala Leu Gly  
35 40 45

Leu Leu Tyr Ser Ala Leu Ala Leu Tyr Ser Leu Tyr Ser Leu Tyr Ser  
50 55 60

Leu Glu Ala Arg Gly Gly Leu Tyr Pro His Glu Ile Leu Glu Ala Leu  
65 70 75 80

Ala Gly Leu Leu Tyr Ser Ala Arg Gly Cys Tyr Ser Ala Leu Ala Pro  
85 90 95

Arg Leu Glu Met Glu Thr Leu Glu Ala Arg Gly Leu Glu Ala Leu Ala  
100 105 110

Pro His Glu His Ile Ser Ser Glu Arg Ala Leu Ala Gly Leu Tyr Thr  
115 120 125

His Arg Pro His Glu Ala Ser Pro Leu Tyr Ser Gly Leu Tyr Thr His  
130 135 140

Arg Leu Tyr Ser Thr His Arg Gly Leu Tyr Gly Leu Tyr Pro Arg Pro  
145 150 155 160

His Glu Gly Leu Tyr Thr His Arg Ile Leu Glu Leu Tyr Ser His Ile

165 170 175

Ser Pro Arg Ala Leu Ala Gly Leu Leu Glu Ala Leu Ala His Ile Ser  
180 185 190

Ser Glu Arg Ala Leu Ala Ala Ser Asn Ala Ser Asn Gly Leu Tyr Leu  
195 200 205

Glu Ala Ser Pro Ile Leu Glu Ala Leu Ala Val Ala Leu Ala Arg Gly  
210 215 220

Leu Glu Leu Glu Gly Leu Pro Arg Leu Glu Leu Tyr Ser Ala Leu Ala

225                      230                      235                      240  
 Gly Leu Pro His Glu Pro Arg Ile Leu Glu Leu Glu Ser Glu Arg Thr  
                          245                      250                      255  
 Tyr Arg Ala Leu Ala Ala Ser Pro Pro His Glu Thr Tyr Arg Gly Leu  
                          260                      265                      270  
 Asn Leu Glu Ala Leu Ala Gly Leu Tyr Val Ala Leu Val Ala Leu Ala  
                          275                      280                      285  
 Leu Ala Val Ala Leu Gly Leu Val Ala Leu Thr His Arg Gly Leu Tyr  
                          290                      295                      300  
 Gly Leu Tyr Pro Arg Leu Tyr Ser Val Ala Leu Pro Arg Pro His Glu  
  
 305                      310                      315                      320  
 His Ile Ser Pro Arg Gly Leu Tyr Ala Arg Gly Gly Leu Ala Ser Pro  
                          325                      330                      335  
 Leu Tyr Ser Pro Arg Gly Leu Pro Arg Pro Arg Pro Arg Gly Leu Gly  
                          340                      345                      350  
 Leu Tyr Ala Arg Gly Leu Glu Pro Arg Ala Ser Pro Ala Leu Ala Thr  
                          355                      360                      365  
 His Arg Leu Tyr Ser Gly Leu Tyr Ser Glu Arg Ala Ser Pro His Ile  
                          370                      375                      380  
  
 Ser Leu Glu Ala Arg Gly Ala Ser Pro Val Ala Leu Pro His Glu Gly  
 385                      390                      395                      400  
 Leu Tyr Leu Tyr Ser Ala Leu Ala Met Glu Thr Gly Leu Tyr Leu Glu  
                          405                      410                      415  
 Thr His Arg Ala Ser Pro Gly Leu Asn Ala Ser Pro Ile Leu Glu Val  
                          420                      425                      430  
 Ala Leu Ala Leu Ala Leu Glu Ser Glu Arg Gly Leu Tyr Gly Leu Tyr  
                          435                      440                      445  
 His Ile Ser Thr His Arg Ile Leu Glu Gly Leu Tyr Ala Leu Ala Ala  
  
                          450                      455                      460  
 Leu Ala His Ile Ser Leu Tyr Ser Gly Leu Ala Arg Gly Ser Glu Arg  
 465                      470                      475                      480

Gly Leu Tyr Pro His Glu Gly Leu Gly Leu Tyr Pro Arg Thr Arg Pro  
485 490 495

Thr His Arg Ser Glu Arg Ala Ser Asn Pro Arg Leu Glu Ile Leu Glu  
500 505 510

Pro His Glu Ala Ser Pro Ala Ser Asn Ser Glu Arg Thr Tyr Arg Pro  
515 520 525

His Glu Thr His Arg Gly Leu Leu Glu Leu Glu Ser Glu Arg Gly Leu  
530 535 540

Tyr Gly Leu Leu Tyr Ser Gly Leu Gly Leu Tyr Leu Glu Leu Glu Gly  
545 550 555 560

Leu Asn Leu Glu Pro Arg Ser Glu Arg Ala Ser Pro Leu Tyr Ser Ala  
565 570 575

Leu Ala Leu Glu Leu Glu Ser Glu Arg Ala Ser Pro Pro Arg Val Ala  
580 585 590

Leu Pro His Glu Ala Arg Gly Pro Arg Leu Glu Val Ala Leu Ala Ser  
595 600 605

Pro Leu Tyr Ser Thr Tyr Arg Ala Leu Ala Ala Leu Ala Ala Ser Pro  
610 615 620

Gly Leu Ala Ser Pro Ala Leu Ala Pro His Glu Pro His Glu Ala Leu  
625 630 635 640

Ala Ala Ser Pro Thr Tyr Arg Ala Leu Ala Gly Leu Ala Ala Leu Ala  
645 650 655

His Ile Ser Gly Leu Asn Leu Tyr Ser Leu Glu Ala Ser Glu Arg Gly  
660 665 670

Leu Ala Leu Glu Ala Gly Leu Tyr Pro His Glu Ala Leu Ala Ala Ser  
675 680 685

Pro Ala Leu Ala  
690

<210> 2

<211> 750

<212> DNA

<213> Artificial Sequence

<220><223> gene of Engineered ascorbate peroxidase(APEX)

<400> 2

atgggaaagt cttaccaac tgtgagtgt gattaccagg acgccgttga gaaggcgaag 60

aagaagctca gaggcctcat cgctgagaag agatgcgctc ctctaagtct ccgtttggca 120

ttccactctg ctggaacctt tgacaagggc acgaagaccg gtggacctt cggaaccatc 180

aagcacctg ccgaactggc tcacagcgt aacaacggtc ttgacatcgc tgttaggctt 240

ttggagccac tcaaggcgga gtccctatt ttgagctacg ccgatttcta ccagttggct 300

ggcgttgttg ccgttgaggt cacgggtgga cctaaggctt cattccaccc tggagagag 360

gacaagcctg agccaccacc agagggtcgc ttgccgatg ccactaaggg ttctgacct 420

ttgagagatg tgtttggcaa agctatgggg ctactgacc aagatctgt tgctctatct 480

gggggtcaca ctattggagc tgcacacaag gacgctctg gatttgaggg tccctggacc 540

tctaactctc ttattttcga caactcatac ttcacggagt tgttgagtgg tgagaaggaa 600

ggtctcttc agctaccttc tgacaaggct cttttgtctg accctgtatt ccgccctctc 660

gttgataaat atgcagcgga cgaagatgcc ttctttgtctg attacgtga ggctcaccaa 720

aagctttccg agcttgggtt tgctgatgcc 750

<210> 3

<211> 691

<212> PRT

<213> Artificial Sequence

<220><223> Engineered ascorbate peroxidase 2 (APEX2)

<400> 3

Met Glu Thr Gly Leu Tyr Leu Tyr Ser Ser Glu Arg Thr Tyr Arg Pro

1 5 10 15

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

20 25 30

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

35 40 45

Leu Leu Tyr Ser Ala Leu Ala Leu Tyr Ser Leu Tyr Ser Leu Tyr Ser

50 55 60

Leu Glu Ala Arg Gly Gly Leu Tyr Pro His Glu Ile Leu Glu Ala Leu

65 70 75 80

Ala Gly Leu Leu Tyr Ser Ala Arg Gly Cys Tyr Ser Ala Leu Ala Pro

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 85                                                              | 90  | 95  |     |
| Arg Leu Glu Met Glu Thr Leu Glu Ala Arg Gly Leu Glu Ala Leu Ala |     |     |     |
| 100                                                             | 105 | 110 |     |
| Pro His Glu His Ile Ser Ser Glu Arg Ala Leu Ala Gly Leu Tyr Thr |     |     |     |
| 115                                                             | 120 | 125 |     |
| His Arg Pro His Glu Ala Ser Pro Leu Tyr Ser Gly Leu Tyr Thr His |     |     |     |
| 130                                                             | 135 | 140 |     |
| Arg Leu Tyr Ser Thr His Arg Gly Leu Tyr Gly Leu Tyr Pro Arg Pro |     |     |     |
| 145                                                             | 150 | 155 | 160 |
| His Glu Gly Leu Tyr Thr His Arg Ile Leu Glu Leu Tyr Ser His Ile |     |     |     |
| 165                                                             | 170 | 175 |     |
| Ser Pro Arg Ala Leu Ala Gly Leu Leu Glu Ala Leu Ala His Ile Ser |     |     |     |
| 180                                                             | 185 | 190 |     |
| Ser Glu Arg Ala Leu Ala Ala Ser Asn Ala Ser Asn Gly Leu Tyr Leu |     |     |     |
| 195                                                             | 200 | 205 |     |
| Glu Ala Ser Pro Ile Leu Glu Ala Leu Ala Val Ala Leu Ala Arg Gly |     |     |     |
| 210                                                             | 215 | 220 |     |
| Leu Glu Leu Glu Gly Leu Pro Arg Leu Glu Leu Tyr Ser Ala Leu Ala |     |     |     |
| 225                                                             | 230 | 235 | 240 |
| Gly Leu Pro His Glu Pro Arg Ile Leu Glu Leu Glu Ser Glu Arg Thr |     |     |     |
| 245                                                             | 250 | 255 |     |
| Tyr Arg Ala Leu Ala Ala Ser Pro Pro His Glu Thr Tyr Arg Gly Leu |     |     |     |
| 260                                                             | 265 | 270 |     |
| Asn Leu Glu Ala Leu Ala Gly Leu Tyr Val Ala Leu Val Ala Leu Ala |     |     |     |
| 275                                                             | 280 | 285 |     |
| Leu Ala Val Ala Leu Gly Leu Val Ala Leu Thr His Arg Gly Leu Tyr |     |     |     |
| 290                                                             | 295 | 300 |     |
| Gly Leu Tyr Pro Arg Leu Tyr Ser Val Ala Leu Pro Arg Pro His Glu |     |     |     |
| 305                                                             | 310 | 315 | 320 |
| His Ile Ser Pro Arg Gly Leu Tyr Ala Arg Gly Gly Leu Ala Ser Pro |     |     |     |
| 325                                                             | 330 | 335 |     |

Leu Tyr Ser Pro Arg Gly Leu Pro Arg Pro Arg Pro Arg Gly Leu Gly  
 340 345 350  
 Leu Tyr Ala Arg Gly Leu Glu Pro Arg Ala Ser Pro Pro Arg Thr His  
 355 360 365  
 Arg Leu Tyr Ser Gly Leu Tyr Ser Glu Arg Ala Ser Pro His Ile Ser  
 370 375 380  
  
 Leu Glu Ala Arg Gly Ala Ser Pro Val Ala Leu Pro His Glu Gly Leu  
 385 390 395 400  
 Tyr Leu Tyr Ser Ala Leu Ala Met Glu Thr Gly Leu Tyr Leu Glu Thr  
 405 410 415  
 His Arg Ala Ser Pro Gly Leu Asn Ala Ser Pro Ile Leu Glu Val Ala  
 420 425 430  
 Leu Ala Leu Ala Leu Glu Ser Glu Arg Gly Leu Tyr Gly Leu Tyr His  
 435 440 445  
 Ile Ser Thr His Arg Ile Leu Glu Gly Leu Tyr Ala Leu Ala Ala Leu  
  
 450 455 460  
 Ala His Ile Ser Leu Tyr Ser Gly Leu Ala Arg Gly Ser Glu Arg Gly  
 465 470 475 480  
 Leu Tyr Pro His Glu Gly Leu Gly Leu Tyr Pro Arg Thr Arg Pro Thr  
 485 490 495  
 His Arg Ser Glu Arg Ala Ser Asn Pro Arg Leu Glu Ile Leu Glu Pro  
 500 505 510  
 His Glu Ala Ser Pro Ala Ser Asn Ser Glu Arg Thr Tyr Arg Pro His  
 515 520 525  
  
 Glu Thr His Arg Gly Leu Leu Glu Leu Glu Ser Glu Arg Gly Leu Tyr  
 530 535 540  
 Gly Leu Leu Tyr Ser Gly Leu Gly Leu Tyr Leu Glu Leu Glu Gly Leu  
 545 550 555 560  
 Asn Leu Glu Pro Arg Ser Glu Arg Ala Ser Pro Leu Tyr Ser Ala Leu  
 565 570 575  
 Ala Leu Glu Leu Glu Ser Glu Arg Ala Ser Pro Pro Arg Val Ala Leu  
 580 585 590

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

595 600 605  
 Leu Tyr Ser Thr Tyr Arg Ala Leu Ala Ala Leu Ala Ala Ser Pro Gly  
 610 615 620  
 Leu Ala Ser Pro Ala Leu Ala Pro His Glu Pro His Glu Ala Leu Ala  
 625 630 635 640  
 Ala Ser Pro Thr Tyr Arg Ala Leu Ala Gly Leu Ala Ala Leu Ala His  
 645 650 655  
 Ile Ser Gly Leu Asn Leu Tyr Ser Leu Glu Ala Ser Glu Arg Gly Leu  
 660 665 670

Ala Leu Glu Ala Gly Leu Tyr Pro His Glu Ala Leu Ala Ala Ser Pro  
 675 680 685  
 Ala Leu Ala  
 690

<210> 4  
 <211> 750  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> gene of Engineered ascorbate peroxidase 2 (APEX2)  
 <400> 4

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atgggaaagt cttaccaac tgtgagtgt gattaccagg acgccgttga gaaggcgaag   | 60  |
| aagaagctca gaggtttcat cgctgagaag agatgcgtc ctctaattgt ccgtttggca  | 120 |
| ttccactctg ctggaacctt tgacaagggc acgaagaccg gtggaccctt cggaaccatc | 180 |
|                                                                   |     |
| aagcaccttg ccgaactggc tcacagcgtt aacaacggtc ttgacatcgc ttttaggctt | 240 |
| ttggagccac tcaaggcgga gtccctatt ttgagctacg ccgatttcta ccagttggct  | 300 |
| ggcgtttgtt ccgttgaggt cacgggtgga cctaagggtt cattccaccc tggagagag  | 360 |
| gacaagcctg agccaccacc agagggtcgc ttgcccgatc ccactaagggt ttctgacct | 420 |
| ttgagagatg tttttggcaa agctatgggg cttactgacc aagatatcgt tgctctatct | 480 |
| gggggtcaca ctattggagc tgcacacaag gacggttctg gatttgaggg tccctggacc | 540 |
| tctaatectc ttattttcga caactcatac ttcacggagt tgttgagtgg tgagaaggaa | 600 |
|                                                                   |     |
| ggtctccttc agctaccttc tgacaaggct cttttgtctg accctgtatt ccgccctctc | 660 |

gttgataaat atgcagcgga cgaagatgcc ttctttgctg attacgtga ggctcaccaa 720  
aagctttccg agcttgggtt tgctgatgcc 750

<210> 5  
<211> 870  
<212> PRT  
<213> Artificial Sequence  
<220><223> Horseradish peroxidase  
<400> 5

Met Glu Thr Gly Leu Asn Leu Glu Thr His Arg Pro Arg Thr His Arg  
1 5 10 15  
Pro His Glu Thr Tyr Arg Ala Ser Pro Ala Ser Asn Ser Glu Arg Cys  
20 25 30  
Tyr Ser Pro Arg Ala Ser Asn Val Ala Leu Ser Glu Arg Ala Ser Asn  
35 40 45  
Ile Leu Glu Val Ala Leu Ala Arg Gly Ala Ser Pro Thr His Arg Ile  
50 55 60  
Leu Glu Val Ala Leu Ala Ser Asn Gly Leu Leu Glu Ala Arg Gly Ser  
65 70 75 80  
Glu Arg Ala Ser Pro Pro Arg Ala Arg Gly Ile Leu Glu Ala Leu Ala  
85 90 95  
Ala Leu Ala Ser Glu Arg Ile Leu Glu Leu Glu Ala Arg Gly Leu Glu  
100 105 110  
His Ile Ser Pro His Glu His Ile Ser Ala Ser Pro Cys Tyr Ser Pro  
115 120 125  
His Glu Val Ala Leu Ala Ser Asn Gly Leu Tyr Cys Tyr Ser Ala Ser  
130 135 140  
Pro Ala Leu Ala Ser Glu Arg Ile Leu Glu Leu Glu Leu Glu Ala Ser  
145 150 155 160  
Pro Ala Ser Asn Thr His Arg Thr His Arg Ser Glu Arg Pro His Glu  
165 170 175  
Ala Arg Gly Thr His Arg Gly Leu Leu Tyr Ser Ala Ser Pro Ala Leu  
180 185 190

Ala Pro His Glu Gly Leu Tyr Ala Ser Asn Ala Leu Ala Ala Ser Asn  
195 200 205

Ser Glu Arg Ala Leu Ala Ala Arg Gly Gly Leu Tyr Pro His Glu Pro  
210 215 220

Arg Val Ala Leu Ile Leu Glu Ala Ser Pro Ala Arg Gly Met Glu Thr  
225 230 235 240

Leu Tyr Ser Ala Leu Ala Ala Leu Ala Val Ala Leu Gly Leu Ser Glu  
245 250 255

Arg Ala Leu Ala Cys Tyr Ser Pro Arg Ala Arg Gly Thr His Arg Val  
260 265 270

Ala Leu Ser Glu Arg Cys Tyr Ser Ala Leu Ala Ala Ser Pro Leu Glu  
275 280 285

Leu Glu Thr His Arg Ile Leu Glu Ala Leu Ala Ala Leu Ala Gly Leu  
290 295 300

Asn Gly Leu Asn Ser Glu Arg Val Ala Leu Thr His Arg Leu Glu Ala  
305 310 315 320

Leu Ala Gly Leu Tyr Gly Leu Tyr Pro Arg Ser Glu Arg Thr Arg Pro  
325 330 335

Ala Arg Gly Val Ala Leu Pro Arg Leu Glu Gly Leu Tyr Ala Arg Gly  
340 345 350

Ala Arg Gly Ala Ser Pro Ser Glu Arg Leu Glu Gly Leu Asn Ala Leu  
355 360 365

Ala Pro His Glu Leu Glu Ala Ser Pro Leu Glu Ala Leu Ala Ala Ser  
370 375 380

Asn Ala Leu Ala Ala Ser Asn Leu Glu Pro Arg Ala Leu Ala Pro Arg  
385 390 395 400

Pro His Glu Pro His Glu Thr His Arg Leu Glu Pro Arg Gly Leu Asn  
405 410 415

Leu Glu Leu Tyr Ser Ala Ser Pro Ser Glu Arg Pro His Glu Ala Arg  
420 425 430

Gly Ala Ser Asn Val Ala Leu Gly Leu Tyr Leu Glu Ala Ser Asn Ala  
435 440 445

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

450 455 460

Ala Leu Ala Leu Glu Ser Glu Arg Gly Leu Tyr Gly Leu Tyr His Ile

465 470 475 480

Ser Thr His Arg Pro His Glu Gly Leu Tyr Leu Tyr Ser Ala Ser Asn

485 490 495

Gly Leu Asn Cys Tyr Ser Ala Arg Gly Pro His Glu Ile Leu Glu Met

500 505 510

Glu Thr Ala Ser Pro Ala Arg Gly Leu Glu Thr Tyr Arg Ala Ser Asn

515 520 525

Pro His Glu Ser Glu Arg Ala Ser Asn Thr His Arg Gly Leu Tyr Leu

530 535 540

Glu Pro Arg Ala Ser Pro Pro Arg Thr His Arg Leu Glu Ala Ser Asn

545 550 555 560

Thr His Arg Thr His Arg Thr Tyr Arg Leu Glu Gly Leu Asn Thr His

565 570 575

Arg Leu Glu Ala Arg Gly Gly Leu Tyr Leu Glu Cys Tyr Ser Pro Arg

580 585 590

Leu Glu Ala Ser Asn Gly Leu Tyr Ala Ser Asn Leu Glu Ser Glu Arg

595 600 605

Ala Leu Ala Leu Glu Val Ala Leu Ala Ser Pro Pro His Glu Ala Ser

610 615 620

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

625 630 635 640

Glu Pro His Glu Ala Ser Pro Ala Ser Asn Leu Tyr Ser Thr Tyr Arg

645 650 655

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

660 665 670

Leu Asn Leu Tyr Ser Gly Leu Tyr Leu Glu Ile Leu Glu Gly Leu Asn

675 680 685

Ser Glu Arg Ala Ser Pro Gly Leu Asn Gly Leu Leu Glu Pro His Glu

690                      695                      700  
 Ser Glu Arg Ser Glu Arg Pro Arg Ala Ser Asn Ala Leu Ala Thr His  
 705                      710                      715                      720  
 Arg Ala Ser Pro Thr His Arg Ile Leu Glu Pro Arg Leu Glu Val Ala  
                          725                      730                      735  
 Leu Ala Arg Gly Ser Glu Arg Pro His Glu Ala Leu Ala Ala Ser Asn

740                      745                      750  
 Ser Glu Arg Thr His Arg Gly Leu Asn Thr His Arg Pro His Glu Pro  
                          755                      760                      765  
 His Glu Ala Ser Asn Ala Leu Ala Pro His Glu Val Ala Leu Gly Leu  
                          770                      775                      780  
 Ala Leu Ala Met Glu Thr Ala Ser Pro Ala Arg Gly Met Glu Thr Gly  
 785                      790                      795                      800  
 Leu Tyr Ala Ser Asn Ile Leu Glu Thr His Arg Pro Arg Leu Glu Ala  
                          805                      810                      815

Thr His Arg Gly Leu Tyr Thr His Arg Gly Leu Asn Gly Leu Tyr Gly  
                          820                      825                      830  
 Leu Asn Ile Leu Glu Ala Arg Gly Leu Glu Ala Ala Ser Asn Cys Tyr  
                          835                      840                      845  
 Ser Ala Arg Gly Val Ala Leu Val Ala Leu Ala Ser Asn Ser Glu Arg  
                          850                      855                      860  
 Ala Ser Asn Ser Glu Arg  
 865                      870

<210> 6

<211> 927

<212> DNA

<213> Artificial Sequence

<220><223> gene of Horseradish peroxidase

<400> 6

atgcagttaa ccctacatt ctacacaat agctgtccca acgtgtccaa catcgttcgc 60

gacacaatcg tcaacgagct cagatccgat cccaggatcg ctgcttcaat attacgtctg 120

cattccatg actgcttcgt gaatggttgc gacgctagca tattactgga caacaccacc 180

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| agtttccgca ctgaaaagga tgcattcggg aacgctaaca gcgccagggg ctttccagt  | 240 |
| atcgatcgca tgaaggctgc cgttgagtca gcatgcccac gaacagtcag ttgtgcagac | 300 |
| ctgctgacta tagctgcgca acagagcgtg actcttgacg gcggaccgtc ctggagagt  | 360 |
| ccgctcggtc gacgtgactc cctacaggca ttcttagatc tggccaacgc caacttcct  | 420 |
|                                                                   |     |
| gctccattct tcacctgcc ccagctgaag gatagcttta gaaacgtggg tctgaatgc   | 480 |
| tcgagtgcac ttgtggctct gtccggagga cacacatttg gaaagaacca gtgtaggttc | 540 |
| atcatggata ggctctacaa tttagcaac actgggttac ctgacccac gctgaacact   | 600 |
| acgtatctcc agacactgag aggccttgctg cactgaatg gcaacctcag tgcactagt  | 660 |
| gactttgatc tgcggacccc aacctcttc gataacaagt actatgtgaa tctagaggag  | 720 |
| cagaaaggcc tgatacagag tgatcaagaa ctgtttagca gtccaaacgc cactgacacc | 780 |
| atccactgg tgagaagttt tgctaactct actcaaacct tctttaacgc ctctgtggaa  | 840 |
|                                                                   |     |
| gccatggacc gtatgggtaa cattaccct ctgacgggtg cccaaggcca gattcgtctg  | 900 |
| aactgcagag tggtaacag caactct                                      | 927 |