



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0108053
(43) 공개일자 2022년08월02일

(51) 국제특허분류(Int. Cl.)
G01N 33/68 (2006.01) G01N 33/543 (2006.01)
(52) CPC특허분류
G01N 33/6818 (2013.01)
G01N 33/54366 (2013.01)
(21) 출원번호 10-2022-7017589
(22) 출원일자(국제) 2020년10월28일
심사청구일자 없음
(85) 번역문제출일자 2022년05월25일
(86) 국제출원번호 PCT/US2020/057638
(87) 국제공개번호 WO 2021/086901
국제공개일자 2021년05월06일
(30) 우선권주장
62/926,897 2019년10월28일 미국(US)

(71) 출원인
퀀텀-에스아이 인코포레이티드
미국 06437 코네티컷주 길포드 올드 화이트필드
스트리트 530
(72) 발명자
리드, 브라이언
미국 06443 코네티컷주 매디슨 힐 로드 94
다이어, 매튜
미국 77386 텍사스주 스프링 카티지 파인즈 레인
3643
(74) 대리인
양영준, 김영

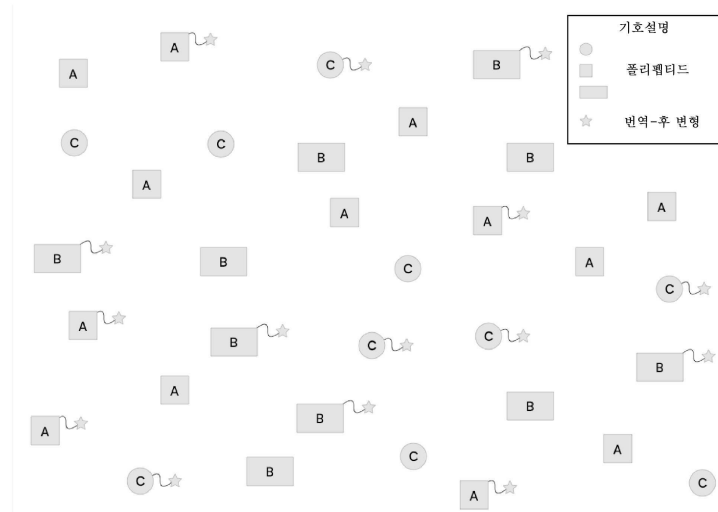
전체 청구항 수 : 총 58 항

(54) 발명의 명칭 폴리펩티드 서열분석을 위한 풍부화된 샘플을 제조하는 방법

(57) 요약

본 출원의 측면은 폴리펩티드 서열분석을 위한 풍부화된 샘플을 제조하는 방법, 및 그에 유용한 조성물, 키트 및 장치를 제공한다.

대표도 - 도1



(52) CPC특허분류
G01N 33/6842 (2013.01)

명세서

청구범위

청구항 1

- (i) 복수의 풍부화 분자를 사용하여 복수의 폴리펩티드로부터 폴리펩티드의 하위세트를 선택함으로써, 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생성하는 단계; 및
- (ii) 풍부화된 샘플에서 폴리펩티드를 병행 서열분석하는 단계를 포함하는 방법.

청구항 2

- (i) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시켜 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계; 및
- (ii) 풍부화된 샘플의 폴리펩티드를 병행 서열분석하는 단계를 포함하는 방법.

청구항 3

제1항 또는 제2항에 있어서, (i)이

- (a) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 결합된 하위세트 및 폴리펩티드의 비결합된 하위세트를 생성하는 것인 단계; 및
- (b) 폴리펩티드의 결합된 하위세트를 분리하여 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함하는 것인 방법.

청구항 4

제1항 또는 제2항에 있어서, (i)이

- (a) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 결합된 하위세트 및 폴리펩티드의 비결합된 하위세트를 생성하는 것인 단계; 및
- (b) 폴리펩티드의 비결합된 하위세트를 분리하여 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함하는 것인 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 복수의 풍부화 분자 중 풍부화 분자 각각이 항체, 압타머, 또는 효소를 포함하는 것인 방법.

청구항 6

제1항 내지 제4항 중 어느 한 항에 있어서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자가 항체, 압타머, 또는 효소를 포함하는 것인 방법.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 복수의 풍부화 분자 중 풍부화 분자 각각이 기관 상에 고정되는 것

인 방법.

청구항 8

제1항 내지 제6항 중 어느 한 항에 있어서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자가 기관 상에 고정되는 것인 방법.

청구항 9

제7항 또는 제8항에 있어서, (i)에서 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 것은 복수의 폴리펩티드를 포함하는 샘플을 기관과 접촉시킬 때 일어나는 것인 방법.

청구항 10

제6항 내지 제9항 중 어느 한 항에 있어서, 기관이 표면, 비드, 입자, 및 겔로 이루어진 군으로부터 선택되고, 임의로 여기서:

표면은 고체 표면이거나;

비드는 자기 비드이거나; 또는

입자는 자기 입자인

방법.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 복수의 풍부화 분자 중 풍부화 분자 각각이 상이한 아미노산 서열을 포함하는 2종 이상의 폴리펩티드에 결합하는 것인 방법.

청구항 12

제1항 내지 제10항 중 어느 한 항에 있어서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자가 상이한 아미노산 서열을 포함하는 2종 이상의 폴리펩티드에 결합하는 것인 방법.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서, 복수의 풍부화 분자 중 풍부화 분자 각각이 아미노산 번역-후 변형에 결합하는 것인 방법.

청구항 14

제1항 내지 제12항 중 어느 한 항에 있어서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자가 아미노산 번역-후 변형에 결합하는 것인 방법.

청구항 15

제13항 또는 제14항에 있어서, 번역-후 변형이 아세틸화, ADP-리보실화, 카스파제 절단, 시트룰린화, 포르밀화, 히드록실화, 메틸화, 미리스토일화, N-연결된 글리코실화, NEDD화, 니트로화, O-연결된 글리코실화, 산화, 팔미토일화, 인산화, 프레닐화, S-니트로실화, 황산화, SUMO화, 유비퀴틸화로 이루어진 군으로부터 선택된 것인 방법.

청구항 16

제13항 내지 제15항 중 어느 한 항에 있어서, 복수의 풍부화 분자의 제1 하위세트가 제1 번역-후 변형에 결합하고, 복수의 풍부화 분자의 풍부화 분자의 제2 하위세트가 제2 번역-후 변형에 결합하는 것인 방법.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, (i)이

(a) 복수의 폴리펩티드를 제1 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 제1 복수의 풍부화 분자 중

풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 제1 결합된 하위세트 및 폴리펩티드의 제1 비결합된 하위세트를 생성하는 것인 단계; 및

(b) (a)의 폴리펩티드의 제1 결합된 하위세트 또는 폴리펩티드의 제1 비결합된 하위세트를 분리하는 단계; 및

(c) 단계 (a) 및 (b)를 1종 이상의 추가의 복수의 풍부화 분자로 반복적으로 반복하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계

를 포함하는 것인 방법.

청구항 18

제17항에 있어서, 제1 복수의 풍부화 분자가 번역-후 변형에 결합하고, 제2 복수의 풍부화 분자가 1종 이상의 관심 폴리펩티드에 결합하는 것인 방법.

청구항 19

제17항에 있어서, 제1 복수의 풍부화 분자가 1종 이상의 관심 폴리펩티드에 결합하고, 제2 복수의 풍부화 분자가 번역-후 변형에 결합하는 것인 방법.

청구항 20

제1항 내지 제19항 중 어느 한 항에 있어서, 복수의 폴리펩티드 중 1종 이상의 폴리펩티드가 (i)에서 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키기 전에, 그와 공동으로, 또는 그 후에, 폴리펩티드를 변형제와 접촉시키는 것에 의해 시험관내 변형되는 것인 방법.

청구항 21

제20항에 있어서, 변형제가 변성제를 포함하고, 적어도 1개의 폴리펩티드가 변성에 의해 변형되는 것인 방법.

청구항 22

제20항 또는 제21항에 있어서, 변형제가 유리 카르복실레이트 기를 차단하고, 적어도 1개의 폴리펩티드가 폴리펩티드의 유리 카르복실레이트 기를 차단하는 것에 의해 변형되는 것인 방법.

청구항 23

제20항 내지 제22항 중 어느 한 항에 있어서, 변형제가 유리 티올 기를 차단하고, 적어도 1개의 폴리펩티드가 폴리펩티드의 유리 티올 기를 차단하는 것에 의해 변형되는 것인 방법.

청구항 24

제20항 내지 제23항 중 어느 한 항에 있어서, 변형제가 절단제를 포함하고, 적어도 1개의 폴리펩티드가 절단에 의해 변형되는 것인 방법.

청구항 25

제20항 내지 제24항 중 어느 한 항에 있어서, 적어도 1개의 폴리펩티드가 번역-후 변형의 부가에 의해 변형되는 것인 방법.

청구항 26

제25항에 있어서, 번역-후 변형이 아세틸화, ADP-리보실화, 카스파제 절단, 시트룰린화, 포르밀화, 히드록실화, 메틸화, 미리스토일화, N-연결된 글리코실화, NEDD화, 니트로화, O-연결된 글리코실화, 산화, 팔미토일화, 인산화, 프레닐화, S-니트로실화, 황산화, SUMO화, 유비퀴틸화로 이루어진 군으로부터 선택된 것인 방법.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서, (ii)가

(a) 풍부화된 샘플의 단일 폴리펩티드 분자를 1종 이상의 말단 아미노산 인식 분자와 접촉시키는 단계; 및

(b) 단일 폴리펩티드가 분해되는 동안 단일 폴리펩티드의 말단에 노출된 연속 아미노산과 1종 이상의 말단 아미노산 인식 분자의 회합을 나타내는 일련의 신호 펄스를 검출하여 단일 폴리펩티드 분자를 서열분석하는 단계를 포함하는 것인 방법.

청구항 28

제1항 내지 제26항 중 어느 한 항에 있어서, (ii)가

(a) 풍부화된 샘플의 단일 폴리펩티드 분자를 1종 이상의 말단 아미노산 인식 분자 및 절단 시약을 포함하는 조성물과 접촉시키는 단계; 및

(b) 절단 시약의 존재 하에 단일 폴리펩티드 분자의 말단과 1종 이상의 말단 아미노산 인식 분자의 회합을 나타내는 일련의 신호 펄스를 검출하는 단계이며, 여기서 일련의 신호 펄스는 절단 시약에 의한 말단 아미노산 절단의 결과로서 시간 경과에 따라 말단에 노출된 일련의 아미노산을 나타내는 것인 단계를

를 포함하는 것인 방법.

청구항 29

제1항 내지 제26항 중 어느 한 항에 있어서, (ii)가

(a) 풍부화된 샘플의 단일 폴리펩티드 분자의 말단에서 제1 아미노산을 확인하는 단계;

(b) 제1 아미노산을 제거하여 단일 폴리펩티드 분자의 말단에서 제2 아미노산을 노출시키는 단계, 및

(c) 단일 폴리펩티드 분자의 말단에서 제2 아미노산을 확인하는 단계를

를 포함하며, 여기서 (a)-(c)는 단일 반응 혼합물에서 수행되는 것인 방법.

청구항 30

제1항 내지 제26항 중 어느 한 항에 있어서, (ii)가

(a) 풍부화된 샘플의 단일 폴리펩티드 분자를 단일 폴리펩티드 분자에 결합하는 1종 이상의 아미노산 인식 분자와 접촉시키는 단계;

(b) 폴리펩티드 분해 조건 하에 단일 폴리펩티드 분자와 1종 이상의 아미노산 인식 분자의 회합을 나타내는 일련의 신호 펄스를 검출하는 단계; 및

(c) 일련의 신호 펄스에서의 제1 특징적인 패턴에 기초하여 단일 폴리펩티드 분자에서 제1 유형의 아미노산을 확인하는 단계를

를 포함하는 것인 방법.

청구항 31

제1항 내지 제26항 중 어느 한 항에 있어서, (ii)가

(a) 폴리펩티드 분해 과정 동안 데이터를 획득하는 단계;

(b) 데이터를 분석하여, 분해 과정 동안 폴리펩티드의 말단에서 순차적으로 노출되는 아미노산에 상응하는 데이터 부분을 결정하는 단계; 및

(c) 폴리펩티드를 나타내는 아미노산 서열을 출력하는 단계를

를 포함하는 것인 방법.

청구항 32

제1항 내지 제26항 중 어느 한 항에 있어서, (ii)가

(a) 풍부화된 샘플의 폴리펩티드를 폴리펩티드의 말단에서 1종 이상의 유형의 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약과 접촉시키는 단계; 및

(b) 폴리펩티드와 1종 이상의 표지된 친화성 시약의 상호작용을 검출함으로써 폴리펩티드의 말단에서 말단 아미노산을 확인하는 단계

를 포함하는 것인 방법.

청구항 33

제1항 내지 제26항 중 어느 한 항에 있어서, (ii)가

(a) 풍부화된 샘플 중의 폴리펩티드를 폴리펩티드의 말단에서 1종 이상의 유형의 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약과 접촉시키는 단계;

(b) 폴리펩티드와 1종 이상의 표지된 친화성 시약의 상호작용을 검출함으로써 폴리펩티드의 말단에서 말단 아미노산을 확인하는 단계;

(c) 말단 아미노산을 제거하는 단계; 및

(d) 폴리펩티드의 말단에서 (a)-(c)를 1회 이상 반복하여 폴리펩티드의 아미노산 서열을 결정하는 단계

를 포함하는 것인 방법.

청구항 34

제33항에 있어서,

(a) 후 및 (b) 전에, 말단 아미노산에 선택적으로 결합하지 않는 1종 이상의 표지된 친화성 시약 중 임의의 것을 제거하는 단계; 및/또는

(b) 후 및 (c) 전에, 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약 중 임의의 것을 제거하는 단계

를 추가로 포함하는 방법.

청구항 35

제33항에 있어서, (c)가 말단 아미노산을 이소티오시아네이트와 접촉시킴으로써 말단 아미노산을 변형시키는 단계, 및:

변형된 말단 아미노산을, 변형된 말단 아미노산에 특이적으로 결합하고 이를 제거하는 프로테아제와 접촉시키는 단계; 또는

변형된 말단 아미노산을, 변형된 말단 아미노산을 제거하기에 충분한 산성 또는 염기성 조건에 적용하는 단계

를 포함하는 것인 방법.

청구항 36

제33항에 있어서, 말단 아미노산을 확인하는 단계가

말단 아미노산을 1종 이상의 표지된 친화성 시약이 결합하는 1종 이상의 유형의 말단 아미노산 중 1종의 유형인 것으로서 확인하는 단계; 또는

말단 아미노산을 1종 이상의 표지된 친화성 시약이 결합하는 1종 이상의 유형의 말단 아미노산 이외의 유형인 것으로서 확인하는 단계

를 포함하는 것인 방법.

청구항 37

제33항에 있어서, 1종 이상의 표지된 친화성 시약이 1종 이상의 표지된 압타머, 1종 이상의 표지된 펩티다제, 1종 이상의 표지된 항체, 1종 이상의 표지된 분해 경로 단백질, 1종 이상의 아미노트랜스퍼라제, 1종 이상의 tRNA 신테타제, 또는 그의 조합을 포함하는 것인 방법.

청구항 38

제37항에 있어서, 1종 이상의 표지된 펩티다제가 절단 활성을 불활성화시키기 위해 변형되거나; 또는 1종 이상의 표지된 펩티다제가 (c)의 제거를 위한 절단 활성을 보유하는 것인 방법.

청구항 39

복수의 풍부화 분자를 포함하는, 제1항 내지 제38항 중 어느 한 항의 방법을 수행하기 위한 키트.

청구항 40

제39항에 있어서, 복수의 풍부화 분자 중 풍부화 분자 각각이 항체, 압타머, 또는 효소를 포함하는 것인 키트.

청구항 41

제39항에 있어서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자가 항체, 압타머, 또는 효소를 포함하는 것인 키트.

청구항 42

제39항 내지 제41항 중 어느 한 항에 있어서, 변형체를 추가로 포함하는 키트.

청구항 43

제42항에 있어서, 변형체가 폴리펩티드 단편화, 폴리펩티드 변성, 번역-후 변형의 부가, 및/또는 1개 이상의 관능기의 차단을 매개하는 것인 키트.

청구항 44

제39항 내지 제43항 중 어느 한 항에 있어서, 표지된 친화성 시약을 추가로 포함하는 키트.

청구항 45

제44항에 있어서, 표지된 친화성 시약이 1종 이상의 표지된 압타머, 1종 이상의 표지된 펩티다제, 1종 이상의 표지된 항체, 1종 이상의 표지된 분해 경로 단백질, 1종 이상의 아미노트랜스퍼라제, 1종 이상의 tRNA 신테타제, 또는 그의 조합을 포함하는 것인 키트.

청구항 46

적어도 1개의 하드웨어 프로세서; 및

적어도 1개의 하드웨어 프로세서에 의해 실행될 때 적어도 1개의 하드웨어 프로세서가 제1항 내지 제38항 중 어느 한 항의 방법을 수행하게 하는 프로세서-실행가능 명령을 저장한 적어도 1개의 비-일시적 컴퓨터-판독가능 저장 매체

를 포함하는 장치.

청구항 47

적어도 1개의 하드웨어 프로세서에 의해 실행될 때 적어도 1개의 하드웨어 프로세서가 제1항 내지 제38항 중 어느 한 항의 방법을 수행하게 하는 프로세서-실행가능 명령을 저장한 적어도 1개의 비-일시적 컴퓨터-판독가능 저장 매체.

청구항 48

1개 이상의 카트리지와 인터페이스되도록 구성된 샘플 제조 모듈을 포함하며, 각각의 카트리는 (a) 복합 샘플을 수용하도록 구성된 1개 이상의 저장소 또는 반응 용기; (b) 복수의 풍부화 분자를 포함하는 1개 이상의 서열 샘플 제조 시약; 및 (c) 1개 이상의 고정된 포획 프로브를 포함하는 매트릭스를 포함하는 것인 장치.

청구항 49

제48항에 있어서, 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 고정된 포획 프로브에 공유 부착된 것인 장치.

청구항 50

제48항 또는 제49항에 있어서, 풍부화 분자의 적어도 하위세트가 고정된 포획 프로브에 의해 결합될 수 있는 비드 또는 입자 상에 공유 부착된 것인 장치.

청구항 51

제48항 내지 제50항 중 어느 한 항에 있어서, 복수의 풍부화 분자 중 풍부화 분자 각각이 항체, 압타머, 또는 효소를 포함하는 것인 장치.

청구항 52

제48항 내지 제50항 중 어느 한 항에 있어서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자가 항체, 압타머, 또는 효소를 포함하는 것인 장치.

청구항 53

제48항 내지 제52항 중 어느 한 항에 있어서, 샘플 제조 시약이 변형제를 포함하는 것인 장치.

청구항 54

제53항에 있어서, 변형제가 폴리펩티드 단편화, 폴리펩티드 변성, 번역-후 변형의 부가, 및/또는 1개 이상의 관능기의 차단을 매개하는 것인 장치.

청구항 55

제48항 내지 제54항 중 어느 한 항에 있어서, 픽셀의 어레이를 포함하는 서열분석 모듈을 추가로 포함하며, 여기서 각각의 픽셀은 샘플 제조 모듈로부터 서열분석 샘플을 수용하도록 구성되고, (a) 샘플 웰; 및 (b) 적어도 1개의 광검출기를 포함하는 것인 장치.

청구항 56

제55항에 있어서, 서열분석 모듈이 서열분석 시약을 각각의 픽셀의 샘플 웰에 전달하도록 구성된 저장소 또는 반응 용기를 추가로 포함하는 것인 장치.

청구항 57

제56항에 있어서, 서열분석 시약이 표지된 친화성 시약을 포함하는 것인 장치.

청구항 58

제57항에 있어서, 표지된 친화성 시약이 1종 이상의 표지된 압타머, 1종 이상의 표지된 펩티다제, 1종 이상의 표지된 항체, 1종 이상의 표지된 분해 경로 단백질, 1종 이상의 아미노트랜스퍼라제, 1종 이상의 tRNA 신테타제, 또는 그의 조합을 포함하는 것인 장치.

발명의 설명

기술 분야

[0001]

관련 출원

[0002]

본 출원은 35 U.S.C. § 119(e) 하에 2019년 10월 28일에 출원된 미국 가출원 일련 번호 62/926,897의 출원일의 이익을 주장하며, 그의 전체 내용은 본원에 참조로 포함된다.

배경 기술

[0003]

프로테오믹스는 생물계 연구에서 게노믹스 및 트랜스크립토믹스에 대한 중요하고 필수적인 보완으로서 부상하였다. 전형적으로, 복합 샘플에서 폴리펩티드의 분획만이 관심대상이다 (예를 들어, 임상적으로 관련됨). 그러나, 고도로 멀티플렉스화된 방식으로 관심 폴리펩티드를 풍부화하기 위한 분자의 활용은 지금까지 제한적이었다.

발명의 내용

- [0004] 풍부화 분자 (및 풍부화 분자의 조합)를 활용하여 관심 폴리펩티드의 상대 존재비를 증가시키는, 폴리펩티드 서열분석을 위한 풍부화된 샘플을 제조하는 방법이 본원에 제공된다. 이들 방법은 고도로 멀티플렉스화된 방식으로 수행될 수 있다. 또한, 그에 유용한 조성물, 키트 및 장치가 본원에 제공된다.
- [0005] 일부 측면에서, 본 개시내용은 폴리펩티드 서열분석을 위한 풍부화된 샘플을 제조하는 방법에 관한 것이다. 일부 실시양태에서, 방법은 (i) 복수의 풍부화 분자를 사용하여 복수의 폴리펩티드로부터 폴리펩티드의 하위세트를 선택함으로써, 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생성하는 단계; 및 (ii) 풍부화된 샘플에서 폴리펩티드를 병행 서열분석하는 단계를 포함한다. 일부 실시양태에서, 방법은 (i) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시켜 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계; 및 (ii) 풍부화된 샘플의 폴리펩티드를 병행 서열분석하는 단계를 포함한다.
- [0006] 일부 실시양태에서, (i)은 (a) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 결합된 하위세트 및 폴리펩티드의 비결합된 하위세트를 생성하는 것인 단계; 및 (b) 폴리펩티드의 결합된 하위세트를 분리하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다.
- [0007] 일부 실시양태에서, (i)은 (a) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 결합된 하위세트 및 폴리펩티드의 비결합된 하위세트를 생성하는 것인 단계; 및 (b) 폴리펩티드의 비결합된 하위세트를 분리하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다.
- [0008] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 항체, 압타머, 또는 효소를 포함한다. 일부 실시양태에서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자는 항체, 압타머, 또는 효소를 포함한다.
- [0009] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 기관 상에 고정된다. 일부 실시양태에서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자는 기관 상에 고정된다. 일부 실시양태에서, (i)에서 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 것은 복수의 폴리펩티드를 포함하는 샘플을 기관과 접촉시킬 때 일어난다. 일부 실시양태에서, 기관은 표면, 비드, 입자, 및 겔로 이루어진 군으로부터 선택되며, 임의로 여기서 표면은 고체 표면이거나; 비드는 자기 비드이거나; 또는 입자는 자기 입자이다.
- [0010] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 상이한 아미노산 서열을 포함하는 2종 이상의 폴리펩티드에 결합한다. 일부 실시양태에서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자는 상이한 아미노산 서열을 포함하는 2종 이상의 폴리펩티드에 결합한다.
- [0011] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 아미노산 번역-후 변형에 결합한다. 일부 실시양태에서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자는 아미노산 번역-후 변형에 결합한다. 일부 실시양태에서, 번역-후 변형은 아세틸화, ADP-리보실화, 카스파제 절단, 시트룰린화, 포르밀화, 히드록실화, 메틸화, 미리스토일화, N-연결된 글리코실화, NEDD화, 니트로화, O-연결된 글리코실화, 산화, 팔미토일화, 인산화, 프레닐화, S-니트로실화, 황산화, SUMO화, 유비퀴틸화로 이루어진 군으로부터 선택된다. 일부 실시양태에서, 복수의 풍부화 분자의 제1 하위세트는 제1 번역-후 변형에 결합하고, 복수의 풍부화 분자의 풍부화 분자의 제2 하위세트는 제2 번역-후 변형에 결합한다.
- [0012] 일부 실시양태에서, (i)은 (a) 복수의 폴리펩티드를 제1 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 제1 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 제1 결합된 하위세트 및 폴리펩티드의 제1 비결합된 하위세트를 생성하는 것인 단계; (b) (a)의 폴리펩티드의 제1 결합된 하위세트 또는 폴리펩티드의 제1 비결합된 하위세트를 분리하는 단계; 및 (c) 단계 (a) 및 (b)를 1종 이상의 추가의 복수의 풍부화 분자로 반복적으로 반복하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다. 일부 실시양태에서, 제1 복수의 풍부화 분자는 번역-후 변형에 결합하고, 제2 복수의 풍부화 분자는 1종 이상의 관심 폴리펩티드에 결합한다. 일부 실시양태에서, 제1 복수의 풍부화 분자는 1종 이상의 관심 폴리펩티드에 결합하고, 제2 복수의 풍부화 분자는 번역-후 변형에 결합한다.

- [0013] 일부 실시양태에서, 복수의 폴리펩티드 중 1종 이상의 폴리펩티드는 (i)에서 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키기 전에, 그와 공동으로, 또는 그 후에, 폴리펩티드를 변형제와 접촉시키는 것에 의해 시험관내 변형된다. 일부 실시양태에서, 변형제는 변성제를 포함하고, 적어도 1개의 폴리펩티드는 변성에 의해 변형된다. 일부 실시양태에서, 변형제는 유리 카르복실레이트 기를 차단하고, 적어도 1개의 폴리펩티드는 폴리펩티드의 유리 카르복실레이트 기를 차단하는 것에 의해 변형된다. 일부 실시양태에서, 변형제는 유리 티올 기를 차단하고, 적어도 1개의 폴리펩티드는 폴리펩티드의 유리 티올 기를 차단하는 것에 의해 변형된다. 일부 실시양태에서, 변형제는 절단제를 포함하고, 적어도 1개의 폴리펩티드는 절단에 의해 변형된다. 일부 실시양태에서, 적어도 1개의 폴리펩티드는 번역-후 변형의 부가에 의해 변형된다. 일부 실시양태에서, 번역-후 변형은 아세틸화, ADP-리보실화, 카스파제 절단, 시트룰린화, 포르밀화, 히드록실화, 메틸화, 미리스토일화, N-연결된 글리코실화, NEDD화, 니트로화, O-연결된 글리코실화, 산화, 팔미토일화, 인산화, 프레닐화, S-니트로실화, 황산화, SUMO화, 유비퀴틸화로 이루어진 군으로부터 선택된다.
- [0014] 일부 실시양태에서, (ii)는 (a) 풍부화된 샘플의 단일 폴리펩티드 분자를 1종 이상의 말단 아미노산 인식 분자와 접촉시키는 단계; 및 (b) 단일 폴리펩티드가 분해되는 동안 단일 폴리펩티드의 말단에 노출된 연속 아미노산과 1종 이상의 말단 아미노산 인식 분자의 회합을 나타내는 일련의 신호 펄스를 검출하여 단일 폴리펩티드 분자를 서열분석하는 단계를 포함한다.
- [0015] 일부 실시양태에서, (ii)는 (a) 풍부화된 샘플의 단일 폴리펩티드 분자를 1종 이상의 말단 아미노산 인식 분자 및 절단 시약을 포함하는 조성물과 접촉시키는 단계; 및 (b) 절단 시약의 존재 하에 단일 폴리펩티드 분자의 말단과 1종 이상의 말단 아미노산 인식 분자의 회합을 나타내는 일련의 신호 펄스를 검출하는 단계이며, 여기서 일련의 신호 펄스는 절단 시약에 의한 말단 아미노산 절단의 결과로서 시간 경과에 따라 말단에 노출된 일련의 아미노산을 나타내는 것인 단계를 포함한다.
- [0016] 일부 실시양태에서, (ii)는 (a) 풍부화된 샘플의 단일 폴리펩티드 분자의 말단에서 제1 아미노산을 확인하는 단계; (b) 제1 아미노산을 제거하여 단일 폴리펩티드 분자의 말단에서 제2 아미노산을 노출시키는 단계, 및 (c) 단일 폴리펩티드 분자의 말단에서 제2 아미노산을 확인하는 단계를 포함하며, 여기서 (a)-(c)는 단일 반응 혼합물에서 수행된다.
- [0017] 일부 실시양태에서, (ii)는 (a) 풍부화된 샘플의 단일 폴리펩티드 분자를 단일 폴리펩티드 분자에 결합하는 1종 이상의 아미노산 인식 분자와 접촉시키는 단계; (b) 폴리펩티드 분해 조건 하에 단일 폴리펩티드 분자와 1종 이상의 아미노산 인식 분자의 회합을 나타내는 일련의 신호 펄스를 검출하는 단계; 및 (c) 일련의 신호 펄스에서의 제1 특징적인 패턴에 기초하여 단일 폴리펩티드 분자에서 제1 유형의 아미노산을 확인하는 단계를 포함한다.
- [0018] 일부 실시양태에서, (ii)는 (a) 폴리펩티드 분해 과정 동안 데이터를 획득하는 단계; (b) 데이터를 분석하여, 분해 과정 동안 폴리펩티드의 말단에서 순차적으로 노출되는 아미노산에 상응하는 데이터 부분을 결정하는 단계; 및 (c) 폴리펩티드를 나타내는 아미노산 서열을 출력하는 단계를 포함한다.
- [0019] 일부 실시양태에서, (ii)는 (a) 풍부화된 샘플의 폴리펩티드를 폴리펩티드의 말단에서 1종 이상의 유형의 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약과 접촉시키는 단계; 및 (b) 폴리펩티드와 1종 이상의 표지된 친화성 시약의 상호작용을 검출함으로써 폴리펩티드의 말단에서 말단 아미노산을 확인하는 단계를 포함한다.
- [0020] 일부 실시양태에서, (ii)는 (a) 풍부화된 샘플 중의 폴리펩티드를 폴리펩티드의 말단에서 1종 이상의 유형의 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약과 접촉시키는 단계; (b) 폴리펩티드와 1종 이상의 표지된 친화성 시약의 상호작용을 검출함으로써 폴리펩티드의 말단에서 말단 아미노산을 확인하는 단계; (c) 말단 아미노산을 제거하는 단계; 및 (d) 폴리펩티드의 말단에서 (a)-(c)를 1회 이상 반복하여 폴리펩티드의 아미노산 서열을 결정하는 단계를 포함한다. 일부 실시양태에서, 방법은 (a) 후 및 (b) 전에, 말단 아미노산에 선택적으로 결합하지 않는 1종 이상의 표지된 친화성 시약 중 임의의 것을 제거하는 단계; 및/또는 (b) 후 및 (c) 전에, 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약 중 임의의 것을 제거하는 단계를 추가로 포함한다. 일부 실시양태에서, (c)는 말단 아미노산을 이소티오시아네이트와 접촉시킴으로써 말단 아미노산을 변형시키는 단계, 및: 변형된 말단 아미노산을, 변형된 말단 아미노산에 특이적으로 결합하고 이를 제거하는 프로테아제와 접촉시키는 단계; 또는 변형된 말단 아미노산을, 변형된 말단 아미노산을 제거하기에 충분한 산성 또는 염기성 조건에 적용하는 단계를 포함한다. 일부 실시양태에서, 말단 아미노산을 확인하는 단계는, 말단 아미노산을 1종 이상의 표지된 친화성 시약이 결합하는 1종 이상의 유형의 말단 아미노산 중 1종의 유형인 것으로서 확인하는 단계; 또는 말단 아미노산을 1종 이상의 표지된 친화성 시약이 결합하는 1종 이상의 유

형의 말단 아미노산 이외의 유형인 것으로서 확인하는 단계를 포함한다.

- [0021] 일부 실시양태에서, 1종 이상의 표지된 친화성 시약은 1종 이상의 표지된 압타머, 1종 이상의 표지된 펩티다제, 1종 이상의 표지된 항체, 1종 이상의 표지된 분해 경로 단백질, 1종 이상의 아미노트랜스퍼라제, 1종 이상의 tRNA 신테타제, 또는 그의 조합을 포함한다. 일부 실시양태에서, 1종 이상의 표지된 펩티다제는 절단 활성을 불활성화시키기 위해 변형되거나; 또는 1종 이상의 표지된 펩티다제는 (c)의 제거를 위한 절단 활성을 보유한다.
- [0022] 일부 측면에서, 본 개시내용은 본원에 기재된 방법을 수행하는데 사용하기 위한 키트에 관한 것이다. 일부 실시양태에서, 키트는 복수의 풍부화 분자를 포함한다. 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 항체, 압타머, 또는 효소를 포함한다. 일부 실시양태에서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자는 항체, 압타머, 또는 효소를 포함한다.
- [0023] 일부 실시양태에서, 키트는 변형제를 포함한다. 일부 실시양태에서, 변형제는 폴리펩티드 단편화, 폴리펩티드 변성, 번역-후 변형의 부가, 및/또는 1개 이상의 관능기의 차단을 매개한다.
- [0024] 일부 실시양태에서, 키트는 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약은 1종 이상의 표지된 압타머, 1종 이상의 표지된 펩티다제, 1종 이상의 표지된 항체, 1종 이상의 표지된 분해 경로 단백질, 1종 이상의 아미노트랜스퍼라제, 1종 이상의 tRNA 신테타제, 또는 그의 조합을 포함한다.
- [0025] 일부 측면에서, 본 개시내용은, 적어도 1개의 하드웨어 프로세서에 의해 실행될 때, 적어도 1개의 하드웨어 프로세서가 본원에 기재된 바와 같은 풍부화 방법을 수행하게 하는 프로세서-실행가능 명령을 저장한 비-일시적 컴퓨터-판독가능 저장 매체에 관한 것이다.
- [0026] 일부 측면에서, 본 개시내용은 본원에 기재된 방법을 수행하기 위한 장치에 관한 것이다. 일부 실시양태에서, 장치는 적어도 1개의 하드웨어 프로세서; 및 적어도 1개의 하드웨어 프로세서에 의해 실행될 때 적어도 1개의 하드웨어 프로세서가 풍부화 방법을 수행하게 하는 프로세서-실행가능 명령을 저장한 적어도 1개의 비-일시적 컴퓨터-판독가능 저장 매체를 포함한다.
- [0027] 일부 실시양태에서, 장치는 (i) 1개 이상의 카트리지와 인터페이스되도록 구성된 것으로, 각각의 카트리는 (a) 복합 샘플을 수용하도록 구성된 1개 이상의 저장소 또는 반응 용기; (b) 복수의 풍부화 분자를 포함하는 1개 이상의 서열 샘플 제조 시약; 및 (c) 1개 이상의 고정된 포획 프로브를 포함하는 매트릭스를 포함하는 것인 샘플 제조 모듈; (ii) 픽셀의 어레이를 포함하는 것으로, 여기서 각각의 픽셀은 샘플 제조 모듈로부터 서열분석 샘플을 수용하도록 구성되고, (a) 샘플 웰; 및 (b) 적어도 1개의 광검출기를 포함하는 것인 서열분석 모듈을 포함한다.
- [0028] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트는 고정된 포획 프로브 상에 고정 (예를 들어, 공유 부착)된다. 일부 실시양태에서, 풍부화 분자의 적어도 하위세트는 고정된 포획 프로브에 의해 결합될 수 있는 비드 또는 입자 상에 고정 (예를 들어, 공유 부착)된다.
- [0029] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 항체, 압타머, 또는 효소를 포함한다. 일부 실시양태에서, 복수의 풍부화 분자의 하위세트 중 풍부화 분자는 항체, 압타머, 또는 효소를 포함한다.
- [0030] 일부 실시양태에서, 샘플 제조 시약은 변형제를 포함한다. 일부 실시양태에서, 변형제는 폴리펩티드 단편화, 폴리펩티드 변성, 번역-후 변형의 부가, 및/또는 1개 이상의 관능기의 차단을 매개한다.
- [0031] 일부 실시양태에서, 서열분석 모듈은 서열분석 시약을 각각의 픽셀의 샘플 웰에 전달하도록 구성된 저장소 또는 반응 용기를 추가로 포함한다.
- [0032] 일부 실시양태에서, 서열분석 시약은 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약은 1종 이상의 표지된 압타머, 1종 이상의 표지된 펩티다제, 1종 이상의 표지된 항체, 1종 이상의 표지된 분해 경로 단백질, 1종 이상의 아미노트랜스퍼라제, 1종 이상의 tRNA 신테타제, 또는 그의 조합을 포함한다.

도면의 간단한 설명

- [0033] 통상의 기술자는 본원에 기재된 도면이 단지 예시적 목적을 위한 것임을 이해할 것이다. 일부 경우에, 본 발명의 다양한 측면은 본 발명의 이해를 용이하게 하기 위해 과장되거나 확대되어 제시될 수 있는 것으로 이해되어야 한다. 도면에서, 유사 참조 문자는 일반적으로 다양한 도면 전반에 걸쳐 비슷한 특색, 기능적으로 유사한 및/또는 구조적으로 유사한 요소를 지칭한다. 도면은 반드시 축척에 맞는 것은 아니며, 대신 교시의 원리를 예

시할 때 강조된다. 도면은 어떠한 방식으로든 본 교시내용의 범주를 제한하는 것으로 의도되지 않는다.

본 발명의 특색 및 이점은 도면과 함께 하기 제시된 상세한 설명으로부터 보다 명백해질 것이다.

도면을 참조하여 실시양태를 기재할 때, 방향 언급 ("위", "아래", "상부", "하부", "좌측", "우측", "수평", "수직" 등)이 사용될 수 있다. 이러한 언급은 단지 통상의 배향으로 도면을 보는 독자를 돕기 위한 것이다. 이들 방향 언급은 구현된 장치의 바람직한 또는 유일한 배향을 기재하는 것으로 의도되지 않는다. 장치는 다른 배향으로 구현될 수 있다.

상세한 설명으로부터 명백한 바와 같이, 도면에 도시되고 본 출원 전반에 걸쳐 예시의 목적으로 추가로 기재된 예는 비제한적 실시양태를 기재하고, 일부 경우에 보다 명확한 예시의 목적으로 특정 과정을 단순화하거나 또는 특색 또는 단계를 생략할 수 있다.

도 1은 복합 샘플의 예시적인 예시를 제공한다. 복합 샘플은 많은 상이한 폴리펩티드로 구성되며, 이들 중 일부만이 관심대상일 수 있다. A (사각형), B (직사각형), 및 C (원형)는 상이한 폴리펩티드를 나타낸다. 별형은 변형 (예를 들어, 번역-후 변형)을 포함하는 폴리펩티드를 강조한다.

도 2는 풍부화 분자의 예시적인 예시를 제공한다. 풍부화 분자는 1종 이상의 폴리펩티드에 결합할 수 있다 (또는 그에 의해 결합될 수 있다). 예를 들어, 풍부화 분자 1은 폴리펩티드 C (원형)에 결합한다 (또는 그에 의해 결합된다). 대조적으로, 풍부화 분자 2는 폴리펩티드 A (사각형) 및 B (직사각형)에 결합하고 (또는 그에 의해 결합되고), 풍부화 분자 3은 폴리펩티드 변형 (여기서 폴리펩티드 A, B 및 C 상에서 발견됨)에 결합한다 (또는 그에 의해 결합된다). 또한, 풍부화 분자는 조합되어 사용될 수 있다.

도 3은 풍부화의 예시적 실시양태를 제공한다. 풍부화 분자는 복합 혼합물에 첨가되며, 표적 폴리펩티드에 결합한다 (또는 그에 의해 결합된다). 이어서, 비결합된 폴리펩티드 (여기서 관심 폴리펩티드)가 분리되고, 이는 표적 폴리펩티드에 비해 풍부화된다. 이러한 방식으로, 관심 폴리펩티드의 상대 존재비가 증가된다.

도 4는 풍부화의 예시적 실시양태를 제공한다. 풍부화 분자는 복합 혼합물에 첨가되며, 표적 폴리펩티드 (여기서 관심 폴리펩티드)에 결합한다 (또는 그에 의해 결합된다). 이어서, 결합된 폴리펩티드가 분리되고, 이는 비-표적 폴리펩티드에 비해 풍부화된다. 이러한 방식으로, 관심 폴리펩티드의 상대 존재비가 증가된다.

도 5는 풍부화된 샘플을 제조하는 예시적인 작업흐름을 도시한 예시를 제공한다.

도 6은 풍부화된 샘플을 제조하는 예시적인 작업흐름을 도시한 예시를 제공한다.

도 7은 풍부화된 샘플을 제조하는 예시적인 작업흐름을 도시한 예시를 제공한다.

도 8은 복합 샘플의 풍부화를 수행하기 위한 예시적인 기구를 도시한 예시를 제공한다.

발명을 실시하기 위한 구체적인 내용

[0034] 본원에 기재된 바와 같이, 본 발명자들은 차등 결합 상호작용이 폴리펩티드 서열분석에서 통상적인 표지 전략에 대한 추가적 또는 대안적 접근법을 제공할 수 있음을 인식하고 인지하였다. 통상적인 폴리펩티드 서열분석은 각각의 유형의 아미노산을 고유하게 확인가능한 표지로 표지하는 것을 수반할 수 있다. 이러한 과정은 수고스럽고 오류를 유발할 수 있는데, 이는 적어도 20종의 상이한 유형의 자연 발생 아미노산과, 또한 그의 수많은 번역-후 변이가 존재하기 때문이다. 일부 측면에서, 본 개시내용은 폴리펩티드의 아미노산 서열을 나타내는 검출 가능한 특징적인 서명을 생산하기 위해 상이한 유형의 아미노산과 차등 회합하는 아미노산 인식 분자를 사용하는 것을 수반하는 기술의 발견에 관한 것이다.

[0035] 일부 측면에서, 본 개시내용은 폴리펩티드 서열분석 반응이 단일 반응 혼합물만을 사용하여 (예를 들어, 반응 용기를 통한 반복적인 시약 순환을 필요로 하지 않으면서) 실시간으로 모니터링될 수 있다는 발견에 관한 것이다. 통상적인 폴리펩티드 서열분석 반응은 폴리펩티드를 상이한 시약 혼합물에 노출시켜 아미노산 검출 및 아미노산 절단 단계 사이를 순환시키는 것을 수반할 수 있다. 따라서, 일부 측면에서, 본 개시내용은 실시간으로 진행 중인 분해 반응 전반에 걸친 아미노산 검출에 의한 폴리펩티드의 분석을 가능하게 하는 차세대 서열분석의 진보에 관한 것이다.

[0036] 개별 유기체의 프로테옴 분석은 세포 과정 및 반응 패턴에 대한 통찰을 제공할 수 있으며, 이는 개선된 진단 및 치료 전략으로 이어진다. 그러나, 복합 샘플은 프로테옴 분석에 대한 다양한 도전과제를 제기한다. 여기서 특히 관련있는 것은, 복합 샘플은 막대한 수의 상이한 폴리펩티드를 포함하고, 전형적으로 폴리펩티드의 단지 작

은 분획만이 관심대상이라는 것이다 (예를 들어, 임상적으로 관련됨). 예를 들어, 혈액으로부터 유래된 복합 샘플에서, 단백질 내용물 중 대부분 (예를 들어, 99%)은 "하우스킵핑 단백질", 예컨대 알부민으로 구성된다. 따라서, 복합 샘플의 프로테옴 분석을 통해 생성된 데이터의 유의한 부분은 거의 가치가 없다. 생성된 데이터의 내용을 관심 영역에 집중시키지 않으면, 중요한 통찰이 과소인식되고/거나 완전히 누락될 수 있다. 본원에 기재된 바와 같이, 본 발명자들은 관심 폴리펩티드를 풍부화하는 능력이 복합 샘플의 프로테옴 분석의 효율을 증가시킬 것임을 인식하고 인지하였다.

[0037] 따라서, 일부 측면에서, 본 개시내용은 풍부화 분자 (및 풍부화 분자의 조합)를 활용하여 관심 폴리펩티드의 상대 존재비를 증가시키는, (예컨대, 본원에 개시된 방법에 의해) 폴리펩티드 서열분석을 위한 풍부화된 샘플을 제조하는 방법에 관한 것이다. 이들 방법은 고도로 멀티플렉스화된 방식으로 수행될 수 있으며, 그에 의해 복합 샘플의 프로테옴 분석의 효율을 증가시키고, 그와 연관된 비용을 감소시킬 수 있다. 또한, 그에 유용한 조성물, 키트 및 장치가 본원에 제공된다.

[0038] I. 복합 폴리펩티드 샘플을 제조하는 방법

[0039] 일부 측면에서, 본 개시내용은 복합 샘플 (예를 들어, 복합 폴리펩티드 샘플)을 제조하는 방법에 관한 것이다. 본원에 사용된 용어 "복합 샘플"은 복수의 분자 (예를 들어, 폴리펩티드, 폴리핵산, 대사물 등)를 포함하며, 그 중 적어도 2종은 화학적으로 고유한 것인 샘플을 지칭한다. 일부 실시양태에서, 복합 샘플은 복수의 폴리펩티드를 포함하고, 여기서 복수는 상이한 아미노산 서열을 포함하는 적어도 2종의 폴리펩티드를 포함한다.

[0040] 전형적으로, 복합 샘플은 세포 집단으로부터 유래된다 (예를 들어, 세포 집단에 의해 생산됨). 일부 실시양태에서, 세포 집단은 단일 세포로 이루어진다. 다른 실시양태에서, 세포 집단은 2종 이상의 세포를 포함한다.

[0041] 예를 들어, 일부 실시양태에서, 세포 집단은 적어도 5, 적어도 10, 적어도 20, 적어도 30, 적어도 40, 적어도 50, 적어도 60, 적어도 70, 적어도 80, 적어도 90, 적어도 100, 적어도 150, 적어도 200, 적어도 250, 적어도 300, 적어도 350, 적어도 400, 적어도 450, 적어도 500, 적어도 600, 적어도 700, 적어도 800, 적어도 900, 적어도 1×10^3 , 적어도 1×10^4 , 적어도 1×10^5 , 적어도 1×10^6 , 적어도 1×10^7 , 적어도 1×10^8 , 적어도 1×10^9 , 또는 적어도 1×10^{10} 개의 세포를 포함한다.

[0042] 일부 실시양태에서, 집단은 1-5, 1-10, 1-20, 1-30, 1-50, 1-60, 1-70, 1-80, 1-90, 1-100, 1-150, 1-200, 1-250, 1-300, 1-350, 1-400, 1-450, 1-500, 1-600, 1-700, 1-800, 1-900, 1- 1×10^3 , 1- 1×10^4 , 1- 1×10^5 , 1- 1×10^6 , 1- 1×10^7 , 1- 1×10^8 , 1- 1×10^9 , 1- 1×10^{10} , 100-150, 100-200, 100-250, 100-300, 100-350, 100-400, 100-450, 100-500, 100-600, 100-700, 100-800, 100-900, 100- 1×10^3 , 100- 1×10^4 , 100- 1×10^5 , 100- 1×10^6 , 100- 1×10^7 , 100- 1×10^8 , 100- 1×10^9 , 100- 1×10^{10} , 1×10^3 -1 $\times 10^4$, 1×10^3 -1 $\times 10^5$, 1×10^3 -1 $\times 10^6$, 1×10^3 -1 $\times 10^7$, 1×10^3 -1 $\times 10^8$, 1×10^3 -1 $\times 10^9$, 1×10^3 -1 $\times 10^{10}$, 1×10^4 -1 $\times 10^5$, 1×10^4 -1 $\times 10^6$, 1×10^4 -1 $\times 10^7$, 1×10^4 -1 $\times 10^8$, 1×10^4 -1 $\times 10^9$, 1×10^4 -1 $\times 10^{10}$, 1×10^5 -1 $\times 10^6$, 1×10^5 -1 $\times 10^7$, 1×10^5 -1 $\times 10^8$, 1×10^5 -1 $\times 10^9$, 또는 1×10^5 -1 $\times 10^{10}$ 개의 세포를 포함한다.

[0043] 세포 집단은 원핵 세포 및/또는 진핵 세포를 포함할 수 있다. 세포 집단은 복수의 동종 세포를 포함할 수 있다. 대안적으로, 세포 집단은 복수의 이종 세포를 포함할 수 있다.

[0044] 세포 집단은 대상체 (예를 들어, 다세포 또는 공생 유기체)로부터 단리될 수 있다. 일부 실시양태에서, 대상체는 마우스, 래트, 토끼, 기니 피그, 햄스터, 돼지, 양, 개, 영장류, 고양이, 또는 인간이다.

[0045] 세포 집단을 단리하는 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다. 예를 들어, 복합 샘플을 제조하는 방법은 생검, 절제 (예를 들어, 미세절제, 예컨대 레이저 포획), 제한 희석, 미세조작, 면역자기 세포 분리, 형광-활성화 세포 분류, 밀도 구배 원심분리, 면역밀도 세포 단리, 마이크로유체 세포 분류, 침강, 부착, 또는 그의 조합을 포함할 수 있다.

[0046] 일부 실시양태에서, 복합 샘플을 제조하는 방법은 세포 집단을 용해시켜 복수의 분자 (예를 들어, 폴리펩티드, 폴리핵산, 대사물 등)를 포함하는 용해 샘플을 생성하는 단계를 포함한다. 세포 집단을 용해시키는 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다. 일부 실시양태에서, 세포를 포함하는 샘플은 공지된 물리적 또는 화학적 방법론 중 어느 하나를 사용하여 용해되어, 상기 세포로부터 표적 분자가 방출된다. 일부 실시양

태에서, 샘플은 전해 방법, 효소적 방법, 세제-기반 방법, 및/또는 기계적 균질화를 사용하여 용해될 수 있다. 일부 실시양태에서, 샘플이 세포 또는 조직을 포함하지 않는 경우 (예를 들어, 정제된 폴리펩티드를 포함하는 샘플), 용해 단계는 생략될 수 있다.

[0047] 대안적으로 또는 추가로, 복합 샘플을 제조하는 방법은 세포하 분획화 (즉, 1개 이상의 세포 구획, 예컨대 엔도솜, 시냅소솜, 세포질, 핵질, 염색질, 미토콘드리아, 퍼옥시솜, 리소솜, 멜라노솜, 엑소솜, 골지체, 내형질 세포망, 중심체, 위축, 또는 그의 조합의 단리)를 포함할 수 있다.

[0048] 동일한 세포 집단으로부터 유래된 분자는 본원에서 동일한 "기원"을 갖는 것으로 기재된다.

[0049] II. 풍부화된 샘플을 제조하는 방법

[0050] 일부 측면에서, 본 개시내용은 샘플을 1종 이상의 관심 분자 (예를 들어, 1종 이상의 관심 폴리펩티드)에 대해 풍부화하는 방법에 관한 것이다. 특히, 일부 측면에서, 본 개시내용은 폴리펩티드 풍부화 방법에 관한 것이다. 본원에 사용된 용어 "폴리펩티드 풍부화"는 1종 이상의 관심 폴리펩티드의 존재비가 1종 이상의 참조 폴리펩티드 (예를 들어, 관심 대상이 아닌 복합 샘플 중의 폴리펩티드)의 존재비에 비해 증가되는 과정을 지칭한다. 본원에 사용된 용어 "관심 폴리펩티드"는 풍부화시키고자 하는 폴리펩티드를 지칭한다. 관심 폴리펩티드는 특이적 아미노산 서열을 포함할 수 있다. 대안적으로 또는 추가로, 관심 폴리펩티드는 특이적 폴리펩티드 변형 (예를 들어, 번역-후 변형)을 포함할 수 있다. 이들 방법은 많은 상이한 폴리펩티드로 구성되어 있고 이들 중 일부만이 관심 대상일 수 있는, 복합 샘플의 프로테옴 분석을 용이하게 한다 (도 1).

[0051] 일부 실시양태에서, 폴리펩티드 풍부화 방법은 복수의 풍부화 분자를 사용하여 복수의 폴리펩티드로부터 폴리펩티드의 하위세트를 선택함으로써, 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생성하는 단계를 포함한다. 일부 실시양태에서, 방법은 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시켜 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다.

[0052] 일부 실시양태에서, 폴리펩티드 풍부화 방법은 (a) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 결합된 하위세트 및 폴리펩티드의 비결합된 하위세트를 생성하는 것인 단계; 및 (b) 폴리펩티드의 결합된 하위세트를 단리하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다. 도 3에 예시된 폴리펩티드 풍부화 방법론은 이러한 실시양태의 예를 제공한다.

[0053] 일부 실시양태에서, 폴리펩티드 풍부화 방법은 (a) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 결합된 하위세트 및 폴리펩티드의 비결합된 하위세트를 생성하는 것인 단계; 및 (b) 폴리펩티드의 비결합된 하위세트를 단리하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다. 도 4에 예시된 폴리펩티드 풍부화 방법론은 이러한 실시양태의 예를 제공한다.

[0054] 상기 단락에 기재된 실시양태에서, 풍부화 분자의 폴리펩티드에 대한 결합은 폴리펩티드의 풍부화 분자에 대한 결합과 동등한 것으로 이해된다. 따라서, 상기 기재된 실시양태에서 단계 (a)는 (a) 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 의해 결합되어, 폴리펩티드의 결합된 하위세트 및 폴리펩티드의 비결합된 하위세트가 생성되는 것인 단계로 동등하게 기재될 수 있다.

[0055] 또한, 상기 기재된 실시양태의 단계 (a) 및 (b)는 추가의 풍부화된 샘플을 생산하기 위해 추가의 복수의 풍부화 분자를 사용하여 1회 이상 반복될 수 있는 것으로 이해된다. 예를 들어, 일부 실시양태에서, 방법은 (a) 복수의 폴리펩티드를 제1 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 제1 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 제1 결합된 하위세트 및 폴리펩티드의 제1 비결합된 하위세트를 생성하는 것인 단계; (b) (a)의 폴리펩티드의 제1 결합된 하위세트 또는 폴리펩티드의 제1 비결합된 하위세트를 단리하는 단계; 및 (c) 단계 (a) 및 (b)를 1종 이상의 추가의 복수의 풍부화 분자로 반복적으로 반복하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다. 일부 실시양태에서, 단계 (a) 및 (b)는 제2, 제3, 제4, 제5, 제6, 제7, 제8, 제9, 제10, 또는 임의의 서수의 추가의 복수의 풍부화 분자를 사용하여 반복된다.

[0056] 예를 들어, 일부 실시양태에서, 방법은 (a) 복수의 폴리펩티드를 제1 복수의 풍부화 분자와 접촉시키는 단계이

며, 여기서 제1 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 제1 결합된 하위세트 및 폴리펩티드의 제1 비결합된 하위세트를 생성하는 것인 단계; (b) (a)의 폴리펩티드의 제1 결합된 하위세트 또는 폴리펩티드의 제1 비결합된 하위세트를 단리하는 단계; (c) (b)의 단리된 폴리펩티드를 제2 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 제2 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 (b)에서 단리된 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 제2 결합된 하위세트 및 폴리펩티드의 제2 비결합된 하위세트를 생성하는 것인 단계; (d) (c)의 폴리펩티드의 제2 결합된 하위세트 또는 폴리펩티드의 제2 비결합된 하위세트를 단리하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다.

[0057] 대안적으로 또는 추가로, 풍부화 방법은 크로마토그래피 (예를 들어, 크기 배제, 이온 교환 등), 등전 포커싱, 막 여과, 분자체 여과, 농축, 침전 (예를 들어, 동결침전), 건조, 투석, 또는 그의 조합을 포함할 수 있다.

[0058] 일부 실시양태에서, 방법은 복합 샘플을 본원에 기재된 키트 또는 장치와 접촉시키는 단계를 포함한다. "샘플 제조를 위한 키트" 및 "샘플 제조 및 샘플 서열분석을 위한 장치"를 참조한다.

[0059] 일부 실시양태에서, 풍부화된 샘플 중의 폴리펩티드는 동일하다 (즉, 동일한 아미노산 서열을 함유한다). 일부 실시양태에서, 풍부화된 샘플은 적어도 2종의 고유한 폴리펩티드 (즉, 상이한 아미노산 서열을 가짐)를 포함한다. 예를 들어, 일부 실시양태에서, 풍부화된 샘플은 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 적어도 20, 적어도 25, 적어도 30, 적어도 40, 적어도 50, 적어도 60, 적어도 70, 적어도 80, 적어도 90, 또는 적어도 100종의 고유한 폴리펩티드를 포함한다. 일부 실시양태에서, 풍부화된 샘플은 1-2, 1-5, 1-10, 1-15, 1-20, 1-30, 1-40, 1-50, 1-60, 1-70, 1-80, 1-90, 1-100, 2-5, 2-10, 2-15, 2-20, 2-30, 2-40, 2-50, 2-60, 2-70, 2-80, 2-90, 2-100, 5-10, 5-15, 5-20, 5-30, 5-40, 5-50, 5-60, 5-70, 5-80, 5-90, 5-100, 10-15, 10-20, 10-30, 10-40, 10-50, 10-60, 10-70, 10-80, 10-90, 10-100, 15-20, 20-30, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 20-30, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 30-40, 30-50, 30-60, 30-70, 30-80, 30-90, 30-100, 40-50, 40-60, 40-70, 40-80, 40-90, 40-100, 50-60, 50-70, 50-80, 50-90, 또는 50-100종의 고유한 폴리펩티드를 포함한다.

[0060] 일부 실시양태에서, 풍부화된 샘플은 적어도 50%, 60%, 70%, 80%, 90%, 95%, 또는 99% 서열 동일성을 공유하는 폴리펩티드를 포함한다. 일부 실시양태에서, 풍부화된 샘플은 1종 이상의 폴리펩티드 변형 (예를 들어, 번역-후 변형)을 공유하는 폴리펩티드를 포함한다. 번역-후 변형의 예는 관련 기술분야의 통상의 기술자에게 공지되어 있고, 아세틸화, 아데닐릴화, ADP-리보실화, 알킬화 (예를 들어, 메틸화), 아미드화, 아르기닐화, 비오틴화, 부티릴화, 카르바밀화, 카르보닐화, 카르복실화, 시트룰린화, 탈아미드화, 엘리미널화, 포르밀화, 글리코실화 (예를 들어, N-연결된 글리코실화, O-연결된 글리코실화), 글리피아티온, 당화, 히드록실화, 아이오딘화, ISG화, 이소프레닐화, 리포일화, 말로닐화, 미리스토일화, NEDD화, 니트로화, 산화, 팔미토일화 PEG화, 인산화, 포스포판테테이닐화, 폴리글리실화, 폴리글루타밀화, 프레닐화, 프로피오닐화, 푸필화, S-글루타티오닐화, S-니트로실화, S-술페닐화, S-술피닐화, S-술폰화, 숙시닐화, 황산화, SUMO화, 및 유비퀴틴화를 포함하나, 이에 제한되지는 않는다.

[0061] A. 풍부화 분자

[0062] 본원에 사용된 용어 "풍부화 분자"는 1종 이상의 표적 폴리펩티드에 대해 (또는 그에 의해) 우선적 결합을 나타내는 분자를 지칭한다. 풍부화 분자는 표적 폴리펩티드의 아미노산 서열과의 직접 상호작용을 통해 표적 폴리펩티드에 결합할 수 있다 (또는 그에 의해 결합될 수 있다). 대안적으로 또는 추가로, 풍부화 분자는 표적 폴리펩티드의 변형 (예를 들어, 번역-후 변형)과의 상호작용을 통해 표적 폴리펩티드에 결합할 수 있다 (또는 그에 의해 결합될 수 있다). 표적 폴리펩티드에 대한 (또는 그에 의한) 풍부화 분자의 결합은 정전기적 상호작용, 소수성 상호작용, 상보적 형상, 또는 그의 조합을 통해 매개될 수 있다.

[0063] 일부 실시양태에서, 표적 폴리펩티드는 관심 폴리펩티드이다. 다른 실시양태에서, 표적 폴리펩티드는 관심 폴리펩티드가 아니다.

[0064] 1종 이상의 표적 폴리펩티드 (또는 표적 폴리펩티드 변이체)에 우선적으로 결합하는 예시적인 풍부화 분자는 이뮤노글로불린, 안티칼린, 리포칼린, DARP인, 압타머, 효소, 렉틴, 및 펩티드 상호작용 도메인을 포함한다.

[0065] 본원에 사용된 용어 "이뮤노글로불린"은 이뮤노글로불린 폴드를 갖는 것을 특징으로 하고 항체로서 기능하고 1종 이상의 기질 (예를 들어, 표적 폴리펩티드)에 결합하는 폴리펩티드를 지칭한다. 따라서, 용어 "이뮤노글로

불린"은 통상적인 이뮤노글로불린 (즉, IgA, IgD, IgE, IgG, 및 IgM), 단일-쇄 가변 단편 (scFv), 항원-결합 단편 (Fab), 아피바디, 및 단일 도메인 항체 (sdAb), 예컨대 나노바디, VHH 및 VNAR을 포괄한다.

- [0066] 본원에 사용된 용어 "압타머"는 1종 이상의 표적 분자 (예를 들어, 표적 폴리펩티드)에 우선적으로 결합하는 폴리핵산 (예를 들어, DNA 또는 RNA) 또는 폴리펩티드를 지칭한다. 자연에서 발견되는 예가 존재하지만, 압타머는 통상적으로 반복된 라운드의 시험관내 선택을 통해 조작된다.
- [0067] 본원에 사용된 용어 "효소"는 1종 이상의 기질 (예를 들어, 표적 폴리펩티드)에의 결합 시 화학 반응을 가속화하는 거대분자 생물학적 촉매를 지칭한다. 전형적으로, 효소는 화학 반응의 완료 후에 그의 기질을 방출할 것이다. 따라서, 풍부화 분자가 효소를 포함하는 것인 일부 실시양태에서, 효소가 기질에 결합된 상태로 유지되어 있을 가능성을 증가시키기 위해 효소는 촉매적으로 불활성화된다. 촉매 불활성화는 1종 이상의 효소적 보조인자 (즉, 촉매로서의 효소의 활성화에 필요한 비-단백질 화학적 화합물 또는 금속성 이온)의 돌연변이유발 및/또는 고갈을 통해 수행될 수 있다.
- [0068] 본원에 사용된 용어 "펩티드 상호작용 도메인"은 1종 이상의 폴리펩티드 (예를 들어, 표적 폴리펩티드)와 상호작용하는 폴리펩티드 (또는 폴리펩티드의 부분)를 지칭한다. 예를 들어, 펩티드 상호작용 도메인은 스캐폴드 단백질, 다중단백질 복합체의 폴리펩티드, 또는 그의 부분일 수 있다.
- [0069] 일부 실시양태에서, 풍부화 분자는 이뮤노글로불린, 압타머, 효소, 및/또는 펩티드 상호작용 도메인을 포함한다.
- [0070] 1종 이상의 표적 폴리펩티드에 의해 우선적으로 결합되는 예시적인 풍부화 분자는 올리고뉴클레오티드 (예를 들어, 이중-가닥 DNA, 단일-가닥 DNA, 이중-가닥 RNA, 단일-가닥 RNA 등), 올리고사카라이드 (또는 폴리사카라이드), 지질, 당단백질, 수용체 리간드, 수용체 효능제, 수용체 길항제, 효소 기질, 및 효소 보조인자를 포함한다.
- [0071] 일부 실시양태에서, 풍부화 분자는 올리고뉴클레오티드 (예를 들어, 이중-가닥 DNA, 단일-가닥 DNA, 이중-가닥 RNA, 단일-가닥 RNA 등), 올리고사카라이드, 지질, 수용체 리간드, 수용체 효능제, 수용체 길항제, 효소 기질, 및/또는 효소 보조인자를 포함한다.
- [0072] (i) 풍부화 분자는 높은 특이성을 나타낼 필요가 없고 (즉, 단일 표적 폴리펩티드에 인지가능한 수준으로만 결합하고 (또는 그에 의해 결합되고)); (ii) 풍부화 분자는 어느 정도의 오프-타겟 결합을 나타낼 수 있고 (즉, 오프-타겟 분자에 검출가능한 수준으로 결합하고 (또는 그에 의해 결합되고)); (iii) 풍부화 분자가 표적 폴리펩티드에 100% 효율로 결합할 필요가 없다는 것 (즉, 과량의 풍부화 분자의 존재 하에서도, 복합 샘플 중의 모든 표적 폴리펩티드가 반드시 결합될 필요는 없다는 것)을 강조하기 위해, 풍부화 분자를 특징화하는데 우선적 결합이 본원에서 사용된다.
- [0073] 일부 실시양태에서, 풍부화 분자는 단일 표적 폴리펩티드 (예를 들어, 도 2의 풍부화 분자 1)에 우선적으로 결합한다 (또는 그에 의해 우선적으로 결합된다). 그러나, 다른 실시양태에서, 풍부화 분자는 2종 이상의 표적 폴리펩티드 (예를 들어, 도 2의 풍부화 분자 2 및 3)에 우선적으로 결합한다 (또는 그에 의해 우선적으로 결합된다).
- [0074] 일부 실시양태에서, 풍부화 분자는 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 적어도 20, 적어도 25, 적어도 30, 적어도 40, 적어도 50, 적어도 60, 적어도 70, 적어도 80, 적어도 90, 또는 적어도 100, 적어도 200, 적어도 300, 적어도 400, 적어도 500, 적어도 600, 적어도 700, 적어도 800, 적어도 900, 적어도 1000, 적어도 2000, 적어도 3000, 적어도 4000, 적어도 5000, 또는 적어도 10,000종의 표적 폴리펩티드에 대해 우선적 결합을 나타낸다 (또는 그에 의해 우선적으로 결합된다).
- [0075] 일부 실시양태에서, 풍부화 분자는 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 또는 15종의 표적 폴리펩티드에 대해 우선적 결합을 나타낸다 (또는 그에 의해 우선적으로 결합된다).
- [0076] 일부 실시양태에서, 풍부화 분자는 1-2, 1-5, 1-10, 1-15, 1-20, 1-30, 1-40, 1-50, 1-60, 1-70, 1-80, 1-90, 1-100, 2-5, 2-10, 2-15, 2-20, 2-30, 2-40, 2-50, 2-60, 2-70, 2-80, 2-90, 2-100, 5-10, 5-15, 5-20, 5-30, 5-40, 5-50, 5-60, 5-70, 5-80, 5-90, 5-100, 10-15, 10-20, 10-30, 10-40, 10-50, 10-60, 10-70, 10-80, 10-90, 10-100, 15-20, 20-30, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 20-30, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 30-40, 30-50, 30-60, 30-70, 30-80, 30-90, 30-100, 40-50, 40-60, 40-

70, 40-80, 40-90, 40-100, 50-60, 50-70, 50-80, 50-90, 또는 50-100, 100-200, 100-300, 100-400, 100-500, 100-600, 100-700, 100-800, 100-900, 100-1000, 100-5000, 100-10,000, 500-600, 500-700, 500-800, 500-900, 500-1000, 500-5000, 500-10,000, 1000-5000, 또는 1000-10,000종의 표적 폴리펩티드에 대해 우선적 결합을 나타낸다 (또는 그에 의해 우선적으로 결합된다).

[0077] 일부 실시양태에서, 풍부화 분자는 적어도 50%, 60%, 70%, 80%, 90%, 95%, 또는 99% 서열 상동성을 공유하는 복수의 관련 표적 폴리펩티드 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50종, 또는 그 초과)의 관련 폴리펩티드에 대해 우선적 결합을 나타낸다 (또는 그에 의해 우선적으로 결합된다).

[0078] 일부 실시양태에서, 풍부화 분자는 번역-후 변형, 예컨대 아세틸화, 아데닐릴화, ADP-리보실화, 알킬화 (예를 들어, 메틸화), 아미드화, 아르기닐화, 비오틴화, 부티릴화, 카르바밀화, 카르보닐화, 카르복실화, 시트룰린화, 탈아미드화, 엘리미닐화, 포르밀화, 글리코실화 (예를 들어, N-연결된 글리코실화, O-연결된 글리코실화), 글리피타이온, 당화, 히드록실화, 아이오딘화, ISG화, 이소프레닐화, 리포일화, 말로닐화, 미리스토일화, NEDD화, 니트로화, 산화, 팔미토일화 PEG화, 인산화, 포스포판테테이닐화, 폴리글리실화, 폴리글루타밀화, 프레닐화, 프로피오닐화, 푸빌화, S-글루타티오닐화, S-니트로실화, S-술폰화, S-술폰화, S-술폰화, 숙시닐화, 황산화, SUMO화, 및 유비퀴틴화에 대해 우선적 결합을 나타낸다 (또는 그에 의해 우선적으로 결합된다).

[0079] 풍부화 분자는 기관 (예를 들어, "샘플 제조 및 샘플 서열분석을 위한 장치"에 기재된 바와 같은 포획 프로브) 상에 고정될 수 있다 (예를 들어, 그에 공유 부착될 수 있다). 기관은 표면 (예를 들어, 고체 표면), 비드 (예를 들어, 자기 비드), 입자 (예를 들어, 자기 입자), 또는 겔일 수 있다.

[0080] (i) 복수의 풍부화 분자

[0081] 전형적으로, 본원에 기재된 풍부화 방법은 복수의 풍부화 분자를 이용한다. 복수의 풍부화 분자는 화학적으로 동일할 수 있다 (즉, 복수는 1종의 풍부화 분자 "유형"을 가짐). 대안적으로, 복수의 풍부화 분자는 상이한 풍부화 분자의 조합을 함유할 수 있다 (즉, 2종 이상의 풍부화 분자 "유형"을 가질 수 있다).

[0082] 일부 실시양태에서, 복수의 풍부화 분자는 단일 풍부화 분자 유형을 함유한다. 다른 실시양태에서, 복수의 풍부화 분자는 2종 이상, 3종 이상, 4종 이상, 5종 이상, 6종 이상, 7종 이상, 8종 이상, 9종 이상, 10종 이상, 11종 이상, 12종 이상, 13종 이상, 14종 이상, 또는 15종 이상의 풍부화 분자 유형의 조합을 포함한다. 일부 실시양태에서, 복수의 풍부화 분자는 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 16, 적어도 17, 적어도 18, 적어도 19, 적어도 20, 적어도 25, 적어도 30, 적어도 40, 적어도 50, 적어도 60, 적어도 70, 적어도 80, 적어도 90, 또는 적어도 100, 적어도 200, 적어도 300, 적어도 400, 적어도 500종의 풍부화 분자 유형을 포함한다.

[0083] 일부 실시양태에서, 복수의 풍부화 분자는 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 또는 15종의 풍부화 분자 유형의 조합을 포함한다.

[0084] 일부 실시양태에서, 복수의 풍부화 분자는 1-2, 1-5, 1-10, 1-15, 1-20, 1-30, 1-40, 1-50, 1-60, 1-70, 1-80, 1-90, 1-100, 2-5, 2-10, 2-15, 2-20, 2-30, 2-40, 2-50, 2-60, 2-70, 2-80, 2-90, 2-100, 5-10, 5-15, 5-20, 5-30, 5-40, 5-50, 5-60, 5-70, 5-80, 5-90, 5-100, 10-15, 10-20, 10-30, 10-40, 10-50, 10-60, 10-70, 10-80, 10-90, 10-100, 15-20, 20-30, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 20-30, 20-40, 20-50, 20-60, 20-70, 20-80, 20-90, 20-100, 30-40, 30-50, 30-60, 30-70, 30-80, 30-90, 30-100, 40-50, 40-60, 40-70, 40-80, 40-90, 40-100, 50-60, 50-70, 50-80, 50-90, 또는 50-100, 100-200, 100-300, 100-400, 또는 100-500종의 풍부화 분자 유형의 조합을 함유한다.

[0085] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 단일 표적 폴리펩티드에 우선적으로 결합한다 (또는 그에 의해 우선적으로 결합된다). 다른 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자의 1종 이상 (예를 들어, 하위세트)은 2종 이상의 표적 폴리펩티드에 대해 우선적 결합을 나타낸다 (또는 그에 의해 우선적으로 결합된다). 또 다른 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 2종 이상의 표적 폴리펩티드에 대해 우선적 결합을 나타낸다 (또는 그에 의해 우선적으로 결합된다).

[0086] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자의 1종 이상 (예를 들어, 하위세트)은 번역-후 폴리펩티드 변형에 결합한다. 다른 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 2종 이상의 번역-후 폴리펩티드 변형에 대해 우선적 결합을 나타낸다.

- [0087] 일부 실시양태에서, 복수의 풍부화 분자 중 풍부화 분자 각각은 기관 (예를 들어, "샘플 제조 및 샘플 서열분석을 위한 장치"에 기재된 바와 같은 포획 프로브), 예컨대 표면 (예를 들어, 고체 표면), 비드 (예를 들어, 자기 비드), 입자 (예를 들어, 자기 입자, 또는 겔) 상에 고정된다 (예를 들어, 그에 공유 부착된다). 일부 실시양태에서, 복수의 풍부화 분자의 1종 이상 (예를 들어, 하위세트)이 기관 상에 고정된다 (예를 들어, 그에 공유 부착된다). 따라서, 일부 실시양태에서, 복수의 폴리펩티드를 복수의 풍부화 분자와 접촉시키는 것은 복수의 폴리펩티드를 포함하는 샘플을 기관과 접촉시킬 때 일어난다.
- [0088] 예를 들어, 일부 실시양태에서, 풍부화 분자는 겔 상에 고정되고 (예를 들어, 그에 공유 부착되거나 가교되고), 샘플은 겔을 통해 잡아당겨진다. 일부 실시양태에서, 풍부화 분자는 비드 (예를 들어, 자기 비드) 상에 고정되고 (예를 들어, 그에 공유 부착되고), 이어서 아래로 잡아당겨진다.
- [0089] (ii) 다중 풍부화 분자 복수물
- [0090] 상기 기재된 바와 같이, 일부 실시양태에서, 방법은 (a) 복수의 폴리펩티드를 제1 복수의 풍부화 분자와 접촉시키는 단계이며, 여기서 제1 복수의 풍부화 분자 중 풍부화 분자의 적어도 하위세트가 복수의 폴리펩티드 중 폴리펩티드의 하위세트에 결합하여, 폴리펩티드의 제1 결합된 하위세트 및 폴리펩티드의 제1 비결합된 하위세트를 생성하는 것인 단계; (b) (a)의 폴리펩티드의 제1 결합된 하위세트 또는 폴리펩티드의 제1 비결합된 하위세트를 분리하는 단계; 및 (c) 단계 (a) 및 (b)를 1종 이상의 추가의 복수의 풍부화 분자로 반복적으로 반복하여, 복수의 폴리펩티드 중 폴리펩티드의 하위세트를 포함하는 풍부화된 샘플을 생산하는 단계를 포함한다. 일부 실시양태에서, 단계 (a) 및 (b)는 제2, 제3, 제4, 제5, 제6, 제7, 제8, 제9, 제10, 또는 임의의 서수의 추가의 복수의 풍부화 분자를 사용하여 반복된다.
- [0091] 일부 실시양태에서, 폴리펩티드 풍부화 방법에 사용되는 복수의 풍부화 분자 각각은 고유하다 (즉, 각각은 상이한 복수의 풍부화 분자를 포함한다). 다른 실시양태에서, 복수물 중 2종 이상은 동일하다. 일부 실시양태에서, 복수의 풍부화 분자 중 적어도 1종은 번역-후 폴리펩티드 변형을 표적화하고, 복수의 풍부화 분자 중 적어도 1종은 번역-후 변형을 표적화하지 않는다.
- [0092] 예를 들어, 제1 풍부화 단계 (제1 복수의 풍부화 분자를 이용함)는 특정한 번역-후 폴리펩티드 변형을 풍부화할 수 있고, 제2 풍부화 단계 (제2 복수의 풍부화 분자를 이용함)는 특정한 폴리펩티드 (및 그러한 폴리펩티드의 변이체)를 풍부화할 수 있다. 대안적으로, 제1 풍부화 단계 (제1 복수의 풍부화 분자를 이용함)는 특정한 폴리펩티드 (및 그러한 폴리펩티드의 변이체)를 풍부화할 수 있고, 제2 풍부화 단계 (제2 복수의 풍부화 분자를 이용함)는 특정한 번역-후 변형을 풍부화할 수 있다.
- [0093] B. 폴리펩티드 변형
- [0094] 복합 샘플의 폴리펩티드 중 1종 이상은 상기 기재된 폴리펩티드 풍부화 전에, 그와 공동으로, 및/또는 그 후에 시험관내 변형될 수 있다. 예를 들어, 일부 실시양태에서, 복합 샘플은 폴리펩티드 풍부화의 수행 전에, 그와 공동으로, 및/또는 그 후에 변형제와 접촉된다. 특히, 변형제는 폴리펩티드 단편화, 폴리펩티드 변성, 번역-후 변형의 부가, 및/또는 1개 이상의 관능기의 차단을 매개할 수 있다.
- [0095] 일부 실시양태에서, 복합 샘플의 1종 이상의 폴리펩티드는 단편화에 의해 변형된다. 일부 실시양태에서, 단편화는 효소적 소화를 포함한다. 일부 실시양태에서, 소화는 폴리펩티드를 엔도펩티다제 (예를 들어, 트립신)와 소화 조건 하에 접촉시킴으로써 수행된다. 일부 실시양태에서, 단편화는 화학적 소화를 포함한다. 화학적 및 효소적 소화에 적합한 시약의 예는 관련 기술분야에 공지되어 있고, 제한 없이, 트립신, 케모트립신, Lys-C, Arg-C, Asp-N, Lys-N, BNPS-스카톨, CNBr, 카스파제, 포름산, 글루타밀 엔도펩티다제, 히드록실아민, 아이오도소벤조산, 호중구 엘라스타제, 펩신, 프로틴-엔도펩티다제, 프로테이나제 K, 스타필로코쿠스 펩티다제 I, 써모리신, 및 트롬빈을 포함한다.
- [0096] 일부 실시양태에서, 복합 샘플의 1종 이상의 폴리펩티드는 변성에 의해 (예를 들어, 열 및/또는 화학적 수단에 의해) 변형된다.
- [0097] 일부 실시양태에서, 복합 샘플의 1종 이상의 폴리펩티드는 시험관내 번역-후 변형에 의해, 예컨대 아세틸화, 아데닐화, ADP-리보실화, 알킬화 (예를 들어, 메틸화), 아미드화, 아르기닐화, 비오틴화, 부티릴화, 카르바미드화, 카르보닐화, 카르복실화, 시트룰린화, 탈아미드화, 엘리미널화, 포르밀화, 글리코실화 (예를 들어, N-연결된 글리코실화, O-연결된 글리코실화), 글리피타이온, 당화, 히드록실화, 아이오딘화, ISG화, 이소프레닐화, 리포일화, 말로닐화, 미리스토일화, NEDD화, 니트로화, 산화, 팔미토일화 PEG화, 인산화, 포스포판테테이닐화, 폴

리글리실화, 폴리글루타밀화, 프레닐화, 프로피오닐화, 푸필화, S-글루타티오닐화, S-니트로실화, S-술피닐화, S-술피닐화, S-술포닐화, 숙시닐화, 황산화, SUMO화, 또는 유비퀴틴화에 의해 변형된다.

[0098] 일부 실시양태에서, 복합 샘플의 1종 이상의 폴리펩티드는 1개 이상의 관능기 (예를 들어, 유리 카르복실레이트 기 및/또는 티올 기)의 차단에 의해 변형된다.

[0099] 일부 실시양태에서, 유리 카르복실레이트 기를 차단하는 것은 비변형된 카르복실레이트에 비해 화학적 반응성을 변경시키는 이들 기의 화학적 변형을 지칭한다. 적합한 카르복실레이트 차단 방법은 관련 기술분야에 공지되어 있고, 관능화된 폴리펩티드의 카르복시-말단 카르복실레이트 기와 화학적으로 상이하도록 측쇄 카르복실레이트 기를 변형시켜야 한다. 일부 실시양태에서, 유리 카르복실레이트 기를 차단하는 것은 폴리펩티드의 유리 카르복실레이트 기의 에스테르화 또는 아마이드화를 포함한다. 일부 실시양태에서, 유리 카르복실레이트 기를 차단하는 것은 예를 들어 폴리펩티드를 메탄올성 HCl과 반응시키는 것에 의한 폴리펩티드의 유리 카르복실레이트 기의 메틸 에스테르화를 포함한다. 유리 카르복실레이트 기를 차단하는데 유용한 시약 및 기술의 추가의 예는, 제한 없이, 4-술포-2,3,5,6-테트라플루오로페놀 (STP) 및/또는 카르보다이미드 예컨대 N-(3-디메틸아미노프로필)-N'-에틸카르보다이미드 히드로클로라이드 (EDAC), 우로늄 시약, 디아조메탄, 피셔 에스테르화를 위한 알콜 및 산, (잠재적으로 후속 에스테르 또는 아민 형성을 위한 중간체로서) N-히드록실숙신이미드 (NHS) 에스테르를 형성하기 위한 NHS의 사용, 또는 카르보닐디이미다졸 (CDI)과의 반응 또는 혼합 무수물의 형성, 또는 잠재적으로 에스테르 또는 아마이드의 형성을 통하는 카르복실산을 변형시키거나 차단하는 임의의 다른 방법을 포함한다.

[0100] 일부 실시양태에서, 유리 티올 기를 차단하는 것은 비변형된 티올에 비해 화학적 반응성을 변경시키는 이들 기의 화학적 변형을 지칭한다. 일부 실시양태에서, 유리 티올 기를 차단하는 것은 폴리펩티드의 유리 티올 기의 환원 및 알킬화를 포함한다. 일부 실시양태에서, 환원 및 알킬화는 폴리펩티드를 디티오프레이톨 (DTT), 및 아이오도아세트아미드 및 아이오도아세트산 중 하나 또는 둘 다와 접촉시키는 것에 의해 수행된다. 사용될 수 있는 추가의 및 대안적 시스테인-환원 시약의 예는 널리 공지되어 있고, 제한 없이, 2-메르캅토에탄올, 트리스 (2-카르복시에틸) 포스핀 히드로클로라이드 (TCEP), 트리부틸포스핀, 디티오부틸아민 (DTBA), 또는 티올 기를 환원시킬 수 있는 임의의 시약을 포함한다. 사용될 수 있는 추가의 및 대안적 시스테인-차단 (예를 들어, 시스테인-알킬화) 시약의 예는 널리 공지되어 있고, 제한 없이, 아크릴아미드, 4-비닐피리딘, N-에틸말레이미드 (NEM), N-ε-말레이미도카프로산 (EMCA), 또는 디술피드 결합 형성을 막도록 시스테인을 변형시키는 임의의 시약을 포함한다.

[0101] 일부 실시양태에서, 폴리펩티드의 N-말단 아미노산 또는 C-말단 아미노산이 변형된다.

[0102] 일부 실시양태에서, 폴리펩티드의 카르복시-말단은 (i) 폴리펩티드의 유리 카르복실레이트 기를 차단하고; (ii) 폴리펩티드를 (예를 들어, 열 및/또는 화학적 수단에 의해) 변형시키고; (iii) 폴리펩티드의 유리 티올 기를 차단하고; (iv) 폴리펩티드를 소화시켜 유리 C-말단 카르복실레이트 기를 포함하는 적어도 1개의 폴리펩티드 단편을 생산하고; (v) 관능성 모이어티를 유리 C-말단 카르복실레이트 기에 (예를 들어, 화학적으로) 접합시키는 것을 포함하는 방법으로 변형된다. 일부 실시양태에서, 방법은 (i) 후 및 (ii) 전에, 폴리펩티드를 포함하는 샘플을 투석하는 것을 추가로 포함한다.

[0103] 일부 실시양태에서, 폴리펩티드의 카르복시-말단은 (i) 폴리펩티드를 변형시키고 (예를 들어, 열 및/또는 화학적 수단에 의해); (ii) 폴리펩티드의 유리 티올 기를 차단하고; (iii) 폴리펩티드를 소화시켜 유리 C-말단 카르복실레이트 기를 포함하는 적어도 1개의 폴리펩티드 단편을 생산하고; (iv) 유리 C-말단 카르복실레이트 기를 차단시켜 차단된 C-말단 카르복실레이트 기를 포함하는 적어도 1개의 폴리펩티드 단편을 생산하고; (v) 관능성 모이어티를 차단된 C-말단 카르복실레이트 기에 (예를 들어, 효소적으로) 접합시키는 것을 포함하는 방법으로 변형된다. 일부 실시양태에서, 방법은 (iv) 후 및 (v) 전에, 폴리펩티드를 포함하는 샘플을 투석하는 것을 추가로 포함한다.

[0104] 일부 실시양태에서, 복합 샘플은 폴리펩티드 단편화, 폴리펩티드 변성, 번역-후 변형의 부가, 및/또는 1개 이상의 관능기의 차단을 매개하기 위해 풍부화 전에 변형제와 접촉된다. 대안적으로 또는 추가로, 일부 실시양태에서, 복합 샘플은 폴리펩티드 단편화, 폴리펩티드 변성, 번역-후 변형의 부가, 및/또는 1개 이상의 관능기의 차단을 매개하기 위해 풍부화와 공동으로 변형제와 접촉된다. 대안적으로 또는 추가로, 일부 실시양태에서, 복합 샘플 (또는 1종 이상의 관심 폴리펩티드를 포함하는, 그로부터 유래된 샘플)은 폴리펩티드 단편화, 폴리펩티드 변성, 번역-후 변형의 부가, 및/또는 1개 이상의 관능기의 차단을 매개하기 위해 풍부화 후에 변형제와 접촉된다.

열 정보를 조합하는 것은 컴퓨터 하드웨어 및 소프트웨어에 의해 수행될 수 있다. "샘플 제조 및 샘플 서열분석을 위한 장치"를 참조한다. 본원에 기재된 방법은 관련 폴리펩티드의 세트, 예컨대 유기체의 전체 프로테오미 서열분석되도록 할 수 있다. 일부 실시양태에서, 복수의 단일 분자 서열분석 반응은 본 출원의 측면에 따라 병행 (예를 들어, 단일 칩 상에서) 수행된다. 예를 들어, 일부 실시양태에서, 복수의 단일 분자 서열분석 반응은 각각 단일 칩 또는 어레이 상의 개별 샘플 웰에서 수행된다.

[0114] 일부 실시양태에서, 본원에 제공된 방법은 폴리펩티드의 복합 혼합물 또는 풍부화된 혼합물을 포함하는 샘플에서 개별 폴리펩티드를 서열분석하고 확인하는데 사용될 수 있다. 일부 실시양태에서, 본 출원은 폴리펩티드의 복합 혼합물 또는 풍부화된 혼합물에서 개별 폴리펩티드를 고유하게 확인하는 방법을 제공한다. 일부 실시양태에서, 개별 폴리펩티드는 혼합 샘플에서 폴리펩티드의 부분 아미노산 서열을 결정함으로써 검출된다. 일부 실시양태에서, 폴리펩티드의 부분 아미노산 서열은 대략 5 내지 50개 아미노산의 인접 스트레치 내에 있다.

[0115] 어떠한 특정한 이론에 얽매이는 것을 원하지는 않지만, 대부분의 인간 단백질은 불완전한 서열 정보를 사용하여 프로테오미 데이터베이스를 참조하여 확인될 수 있는 것으로 여겨진다. 예를 들어, 인간 프로테오미의 단순 모델링은 대략 98%의 단백질이 6 내지 40개의 아미노산의 스트레치 내에서 단지 4종의 유형의 아미노산을 검출하는 것에 의해 고유하게 확인될 수 있음을 제시하였다 (예를 들어, 문헌 [Swaminathan, et al., PLoS Comput Biol. 2015, 11(2):e1004080; 및 Yao, et al., Phys. Biol. 2015, 12(5):055003] 참조). 따라서, 폴리펩티드의 복합 혼합물 또는 풍부화된 혼합물은 대략 6 내지 40개의 아미노산의 짧은 폴리펩티드 단편으로 분해 (예를 들어, 화학적으로 분해, 효소적으로 분해)될 수 있고, 이 폴리펩티드 라이브러리의 서열분석은 원래의 복합 혼합물 또는 풍부화된 혼합물에 존재하는 각각의 폴리펩티드의 동일성 및 존재비를 밝혀낼 것이다. 선택적으로 아미노산 표지하고 부분 서열 정보를 결정함으로써 폴리펩티드를 확인하기 위한 조성물 및 방법은 2015년 9월 15일에 출원된 발명의 명칭 "SINGLE MOLECULE PEPTIDE SEQUENCING"의 미국 특허 출원 번호 15/510,962에 상세히 기재되어 있으며, 이는 그 전문이 참조로 포함된다.

[0116] 실시양태는 단일 폴리펩티드 분자를 높은 정확도, 예컨대 적어도 약 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.9%, 99.99%, 99.999%, 또는 99.9999%의 정확도로 서열분석할 수 있다. 일부 실시양태에서, 단일 분자 서열분석에 사용되는 표적 분자는 고체 지지체의 표면, 예컨대 샘플 웰의 하부 표면 또는 측면 표면 상에 고정된 폴리펩티드이다. 샘플 웰은 또한 본 출원에 따른 서열분석 반응에 필요한 다른 시약, 예컨대 1종 이상의 적합한 완충제, 보조-인자, 표지된 친화성 시약, 및 효소 (예를 들어, 발광 표지되거나 비표지될 수 있는, 촉매 활성 또는 불활성 엑소펩티다제 효소)를 함유할 수 있다.

[0117] 일부 측면에서, 본 출원에 따른 서열분석은 기관 (예를 들어, 고체 지지체, 예를 들어 칩, 예를 들어 본원에 기재된 바와 같은 집적 장치)의 표면 상에 폴리펩티드를 고정시키는 것을 수반할 수 있다. 일부 실시양태에서, 폴리펩티드는 기관 상의 샘플 웰의 표면 상에 (예를 들어, 샘플 웰의 하부 표면 상에) 고정될 수 있다. 일부 실시양태에서, 폴리펩티드의 N-말단 아미노산이 고정된다 (예를 들어, 표면에 부착된다). 일부 실시양태에서, 폴리펩티드의 C-말단 아미노산이 고정된다 (예를 들어, 표면에 부착된다). 일부 실시양태에서, 1개 이상의 비-말단 아미노산이 고정된다 (예를 들어, 표면에 부착된다). 고정된 아미노산(들)은 예를 들어 본 출원에 기재된 바와 같은 임의의 적합한 공유 또는 비-공유 연결을 사용하여 부착될 수 있다. 일부 실시양태에서, 복수의 폴리펩티드는, 예를 들어 기관 상의 샘플 웰의 어레이 내의 복수의 샘플 웰에 부착된다 (예를 들어, 각각의 샘플 웰의 표면, 예를 들어 하부 표면에 부착된 1종의 폴리펩티드로).

[0118] 일부 측면에서, 본 출원에 따른 서열분석은 단일 분자 분석을 허용하는 시스템을 사용하여 수행될 수 있다. 시스템은 서열분석 장치 및 서열분석 장치와 인터페이스되도록 구성된 기기를 포함할 수 있다. "샘플 제조 및 샘플 서열분석을 위한 장치"를 참조한다.

[0119] A. 표지된 친화성 시약 및 사용 방법

[0120] 일부 실시양태에서, 본원에 제공된 방법은 폴리펩티드를 1종의 유형의 말단 아미노산에 선택적으로 결합하는 표지된 친화성 시약 (또한 본원에서, 표지를 포함할 수 있거나 포함하지 않을 수 있는 아미노산 인식 분자로 지칭됨)과 접촉시키는 것을 포함한다. 본원에 사용된 일부 실시양태에서, 말단 아미노산은 폴리펩티드의 아미노-말단 아미노산 또는 폴리펩티드의 카르복시-말단 아미노산을 지칭할 수 있다. 일부 실시양태에서, 표지된 친화성 시약은 다른 유형의 말단 아미노산에 비해 하나의 유형의 말단 아미노산에 선택적으로 결합한다. 일부 실시양태에서, 표지된 친화성 시약은 하나의 유형의 내부 아미노산에 비해 동일한 유형의 말단 아미노산에 선택적으로 결합한다. 또 다른 실시양태에서, 표지된 친화성 시약은 폴리펩티드의 임의의 위치에서 하나의 유형의 아미노산에, 예를 들어 말단 아미노산 및 내부 아미노산과 동일한 유형의 아미노산에 선택적으로 결합한다.

- [0121] 본원에 사용된 바와 같이, 일부 실시양태에서, 아미노산의 유형은 20종의 자연 발생 아미노산 중 1종 또는 그의 유형의 하위세트를 지칭한다. 일부 실시양태에서, 아미노산의 유형은 20종의 자연 발생 아미노산 중 1종의 변형된 변이체 또는 그의 비변형된 및/또는 변형된 변이체의 하위세트를 지칭한다. 변형된 아미노산 변이체의 예는, 제한 없이, 번역-후-변형된 변이체 (예를 들어, 아세틸화, ADP-리보실화, 카스파제 절단, 시트룰린화, 포르밀화, N-연결된 글리코실화, O-연결된 글리코실화, 히드록실화, 메틸화, 미리스토일화, NEDD화, 니트로화, 산화, 팔미토일화, 인산화, 프레닐화, S-니트로실화, 황산화, SUMO화, 및 유비퀴틴화), 화학적으로 변형된 변이체, 비천연 아미노산, 및 단백질생성 아미노산, 예컨대 셀레노시스테인 및 피롤리신을 포함한다. 일부 실시양태에서, 아미노산의 유형의 하위세트는 1종 이상의 유사한 생화학적 특성을 갖는 1종 초과 20종 미만의 아미노산을 포함한다. 예를 들어, 일부 실시양태에서, 아미노산의 유형은 하전된 측쇄 (예를 들어, 양으로 및/또는 음으로 하전된 측쇄)를 갖는 아미노산, 극성 측쇄 (예를 들어, 극성 비하전된 측쇄)를 갖는 아미노산, 비극성 측쇄 (예를 들어, 비극성 지방족 및/또는 방향족 측쇄)를 갖는 아미노산, 및 소수성 측쇄를 갖는 아미노산으로부터 선택된 하나의 유형을 지칭한다.
- [0122] 일부 실시양태에서, 본원에 제공된 방법은 폴리펩티드를 1종 이상의 유형의 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약과 접촉시키는 것을 포함한다. 예시적이고 비제한적인 예로서, 4종의 표지된 친화성 시약이 본 출원의 방법에 사용되는 경우에, 임의의 1종의 시약은 다른 3종 중 임의의 것이 선택적으로 결합하는 또 다른 유형의 아미노산과 상이한 1종의 유형의 말단 아미노산에 선택적으로 결합한다 (예를 들어, 제1 시약은 제1 유형에 결합하고, 제2 시약은 제2 유형에 결합하고, 제3 시약은 제3 유형에 결합하고, 제4 시약은 제4 유형의 말단 아미노산에 결합함). 이러한 논의의 목적상, 본원에 기재된 방법과 관련하여 1종 이상의 표지된 친화성 시약은 대안적으로, 표지된 친화성 시약의 세트로서 지칭될 수 있다.
- [0123] 일부 실시양태에서, 표지된 친화성 시약의 세트는 적어도 1종 및 최대 6종의 표지된 친화성 시약을 포함한다. 예를 들어, 일부 실시양태에서, 표지된 친화성 시약의 세트는 1, 2, 3, 4, 5, 또는 6종의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 10종 이하의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 8종 이하의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 6종 이하의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 4종 이하의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 3종 이하의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 2종 이하의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 4종의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 적어도 2종 최대 20종 (예를 들어, 적어도 2종 최대 10종, 적어도 2종 최대 8종, 적어도 4종 최대 20종, 적어도 4종 최대 10종)의 표지된 친화성 시약을 포함한다. 일부 실시양태에서, 표지된 친화성 시약의 세트는 20종 초과 (예를 들어, 20 내지 25, 20 내지 30종)의 친화성 시약을 포함한다. 그러나, 목적하는 용도를 수용하기 위해 임의의 수의 친화성 시약이 본 출원의 방법에 따라 사용될 수 있다는 것이 인지되어야 한다.
- [0124] 본 출원에 따라, 일부 실시양태에서, 1종 이상의 유형의 아미노산은 표지된 친화성 시약 (예를 들어, 발광 표지를 포함하는 아미노산 인식 분자)의 발광을 검출함으로써 확인된다. 일부 실시양태에서, 표지된 친화성 시약은 1종의 유형의 아미노산에 선택적으로 결합하는 친화성 시약 및 친화성 시약과 회합되는 발광을 갖는 발광 표지를 포함한다. 이러한 방식으로, 발광 (예를 들어, 발광 수명, 발광 강도, 및 본원의 다른 곳에 기재된 다른 발광 특성)은 폴리펩티드의 아미노산을 확인하기 위해 친화성 시약의 선택적 결합과 회합될 수 있다. 일부 실시양태에서, 복수의 유형의 표지된 친화성 시약이 본 출원에 따른 방법에 사용될 수 있으며, 여기서 각각의 유형은 복수의 것 중에서 고유하게 확인가능한 발광을 갖는 발광 표지를 포함한다. 적합한 발광 표지는 발광 분자, 예컨대 형광단 염료를 포함할 수 있고, 이는 본원의 다른 곳에 기재되어 있다.
- [0125] 일부 실시양태에서, 1종 이상의 유형의 아미노산은 표지된 친화성 시약의 1종 이상의 전기적 특징을 검출하는 것에 의해 확인된다. 일부 실시양태에서, 표지된 친화성 시약은 1종의 유형의 아미노산에 선택적으로 결합하는 친화성 시약 및 친화성 시약과 회합되는 전도성 표지를 포함한다. 이러한 방식으로, 1종 이상의 전기적 특징 (예를 들어, 전하, 전류 진동 색, 및 다른 전기적 특징)은 폴리펩티드의 아미노산을 확인하기 위해 친화성 시약의 선택적 결합과 회합될 수 있다. 일부 실시양태에서, 복수의 유형의 표지된 친화성 시약이 본 출원에 따른 방법에 사용될 수 있으며, 여기서 각각의 유형은 복수의 것 중에서 고유하게 확인가능한 전기 신호의 변화 (예를 들어, 전도도의 변화, 예컨대 전도성의 진폭의 변화 및 특징적인 패턴의 전도성 전이)를 생산하는 전도성 표지를 포함한다. 일부 실시양태에서, 복수의 유형의 표지된 친화성 시약은 각각 상이한 수의 하전된 기 (예를 들어, 상이한 수의 음으로 및/또는 양으로 하전된 기)를 갖는 전도성 표지를 포함한다. 따라서, 일부 실시양태

에서, 전도성 표지는 전하 표지이다. 전하 표지의 예는 텐드리머, 나노입자, 핵산 및 다중 하전된 기를 갖는 다른 중합체를 포함한다. 일부 실시양태에서, 전도성 표지는 그의 알짜 전하 (예를 들어, 알짜 양전하 또는 알짜 음전하)에 의해, 그의 전하 밀도에 의해, 및/또는 그의 하전된 기의 수에 의해 고유하게 확인가능하다.

[0126] 일부 실시양태에서, 친화성 시약 (예를 들어, 아미노산 인식 분자)은 통상적으로 공지된 기술을 사용하여 관련 기술분야의 통상의 기술자에 의해 조작될 수 있다. 일부 실시양태에서, 바람직한 특성은 폴리펩티드의 말단 (예를 들어, N-말단 또는 C-말단)에 위치하는 경우에만 1종의 유형의 아미노산에 선택적으로 고친화도로 결합하는 능력을 포함할 수 있다. 또 다른 실시양태에서, 바람직한 특성은 폴리펩티드의 말단 (예를 들어, N-말단 또는 C-말단)에 위치하는 경우 및 폴리펩티드의 내부 위치에 위치하는 경우에 1종의 유형의 아미노산에 선택적으로 고친화도로 결합하는 능력을 포함할 수 있다.

[0127] 일부 실시양태에서, 본원에 사용된 용어 "선택적" 및 "특이적" (및 그의 변경, 예를 들어, 선택적으로, 특이적으로, 선택성, 특이성)은 우선적 결합 상호작용을 지칭한다. 예를 들어, 일부 실시양태에서, 하나의 유형의 아미노산에 선택적으로 결합하는 표지된 친화성 시약은 또 다른 유형의 아미노산에 비해 하나의 유형에 우선적으로 결합한다. 선택적 결합 상호작용은 하나의 유형의 아미노산 (예를 들어, 하나의 유형의 말단 아미노산)과 다른 유형의 아미노산 (예를 들어, 다른 유형의 말단 아미노산) 사이를, 전형적으로 약 10- 내지 100-배 또는 그 초과 (예를 들어, 약 1,000- 또는 10,000-배 초과)로 구별할 것이다. 따라서, 선택적 결합 상호작용은 다른 유형의 아미노산에 비해 하나의 유형의 아미노산에 대한 고유하게 확인가능한 임의의 결합 상호작용을 지칭할 수 있는 것으로 인지되어야 한다. 예를 들어, 일부 측면에서, 본 출원은 1종 이상의 아미노산 인식 분자와 폴리펩티드 분자의 회합을 나타내는 데이터를 수득하는 것에 의해 폴리펩티드 서열분석하는 방법을 제공한다. 일부 실시양태에서, 데이터는 폴리펩티드 분자의 아미노산과의 일련의 가역적 아미노산 인식 분자 결합 상호작용에 상응하는 일련의 신호 펄스를 포함하고, 데이터는 아미노산의 정체를 결정하는데 사용될 수 있다. 따라서, 일부 실시양태에서, "선택적" 또는 "특이적" 결합 상호작용은 하나의 유형의 아미노산과 다른 유형의 아미노산 사이를 구별하는 검출된 결합 상호작용을 지칭한다.

[0128] 일부 실시양태에서, 표지된 친화성 시약 (예를 들어, 아미노산 인식 분자)은 다른 유형의 아미노산에는 유의하게 결합하지 않으면서, 하나의 유형의 아미노산에 약 10^{-6} M 미만의 (예를 들어, 약 10^{-7} M 미만, 약 10^{-8} M 미만, 약 10^{-9} M 미만, 약 10^{-10} M 미만, 약 10^{-11} M 미만, 약 10^{-12} M 미만의, 10^{-16} M만큼 낮은) 해리 상수 (K_D)로 선택적으로 결합한다. 일부 실시양태에서, 표지된 친화성 시약은 1종의 유형의 아미노산 (예를 들어, 1종의 유형의 말단 아미노산)에 약 100 nM 미만, 약 50 nM 미만, 약 25 nM 미만, 약 10 nM 미만, 또는 약 1 nM 미만의 K_D 로 선택적으로 결합한다. 일부 실시양태에서, 표지된 친화성 시약은 하나의 유형의 아미노산에 약 50 nM 내지 약 50 μ M (예를 들어, 약 50 nM 내지 약 500 nM, 약 50 nM 내지 약 5 μ M, 약 500 nM 내지 약 50 μ M, 약 5 μ M 내지 약 50 μ M, 또는 약 10 μ M 내지 약 50 μ M)의 K_D 로 선택적으로 결합한다. 일부 실시양태에서, 아미노산 인식 분자는 1종의 유형의 아미노산에 약 50 nM의 K_D 로 결합한다.

[0129] 일부 실시양태에서, 표지된 친화성 시약 (예를 들어, 아미노산 인식 분자)은 2종 이상의 유형의 아미노산에 약 10^{-6} M 미만의 (예를 들어, 약 10^{-7} M 미만, 약 10^{-8} M 미만, 약 10^{-9} M 미만, 약 10^{-10} M 미만, 약 10^{-11} M 미만, 약 10^{-12} M 미만의, 10^{-16} M만큼 낮은) K_D 로 결합한다. 일부 실시양태에서, 아미노산 인식 분자는 2종 이상의 유형의 아미노산에 약 100 nM 미만, 약 50 nM 미만, 약 25 nM 미만, 약 10 nM 미만, 또는 약 1 nM 미만의 K_D 로 결합한다. 일부 실시양태에서, 아미노산 인식 분자는 2종 이상의 유형의 아미노산에 약 50 nM 내지 약 50 μ M (예를 들어, 약 50 nM 내지 약 500 nM, 약 50 nM 내지 약 5 μ M, 약 500 nM 내지 약 50 μ M, 약 5 μ M 내지 약 50 μ M, 또는 약 10 μ M 내지 약 50 μ M)의 K_D 로 결합한다. 일부 실시양태에서, 아미노산 인식 분자는 2종 이상의 유형의 아미노산에 약 50 nM의 K_D 로 결합한다.

[0130] 일부 실시양태에서, 표지된 친화성 시약 (예를 들어, 아미노산 인식 분자)은 적어도 1종의 유형의 아미노산에 적어도 0.1 s^{-1} 의 해리율 (k_{off})로 결합한다. 일부 실시양태에서, 해리율은 약 0.1 s^{-1} 내지 약 $1,000 \text{ s}^{-1}$ (예를 들어, 약 0.5 s^{-1} 내지 약 500 s^{-1} , 약 0.1 s^{-1} 내지 약 100 s^{-1} , 약 1 s^{-1} 내지 약 100 s^{-1} , 또는 약 0.5 s^{-1} 내지 약 50 s^{-1})이다. 일부 실시양태에서, 해리율은 약 0.5 s^{-1} 내지 약 20 s^{-1} 이다. 일부 실시양태에서, 해리율은 약 2 s^{-1} 내지 약 20 s^{-1} 이다. 일부 실시양태에서, 해리율은 약 0.5 s^{-1} 내지 약 2 s^{-1} 이다.

[0131] 일부 실시양태에서, K_D 또는 k_{off} 에 대한 값은 공지된 문헌 값일 수 있거나, 또는 값은 실험적으로 결정될 수

있다. 예를 들어, KD 또는 koff에 대한 값은 단일-분자 검정 또는 앙상블 검정에서 측정될 수 있다. 일부 실시양태에서, koff에 대한 값은 본원의 다른 곳에 기재된 바와 같은 단일-분자 검정에서 수득된 신호 펄스 정보에 기초하여 실험적으로 결정될 수 있다. 예를 들어, koff에 대한 값은 평균 펄스 지속기간의 역수에 의해 근사화될 수 있다. 일부 실시양태에서, 아미노산 인식 분자는 2종 이상의 유형의 아미노산에 2종 이상의 유형 각각에 대한 상이한 KD 또는 koff로 결합한다. 일부 실시양태에서, 제1 유형의 아미노산에 대한 제1 KD 또는 koff는 제2 유형의 아미노산에 대한 제2 KD 또는 koff와 적어도 10% (예를 들어, 적어도 25%, 적어도 50%, 적어도 100%, 또는 그 초과)만큼 상이하다. 일부 실시양태에서, KD 또는 koff에 대한 제1 및 제2 값은 약 10-25%, 25-50%, 50-75%, 75-100%, 또는 100% 초과, 예를 들어 약 2-배, 3-배, 4-배, 5-배, 또는 그 초과만큼 상이하다.

[0132] 일부 실시양태에서, 표지된 친화성 시약은 발광 표지 (예를 들어, 표지) 및 폴리펩티드의 1종 이상의 유형의 말단 아미노산에 선택적으로 결합하는 친화성 시약을 포함한다. 일부 실시양태에서, 친화성 시약은 아미노산의 1종의 유형 또는 아미노산의 유형의 하위세트 (예를 들어, 아미노산의 20종 미만의 통상적인 유형)에 대해 말단 위치에서 또는 말단 및 내부 위치 둘 다에서 선택적이다.

[0133] 본원에 기재된 바와 같이, 친화성 시약 (또한 "인식 분자"로 공지됨)은 또 다른 분자에 비해 하나의 분자에 (예를 들어, 본원에 언급된 "아미노산 인식 분자"에서와 같이, 또 다른 유형의 아미노산에 비해 하나의 유형의 아미노산에) 선택적으로 또는 특이적으로 결합할 수 있는 임의의 생체분자일 수 있다. 친화성 시약 (예를 들어, 인식 분자)은, 예를 들어 합성 또는 재조합될 수 있는 단백질 및 핵산을 포함한다. 일부 실시양태에서, 친화성 시약 또는 인식 분자는 항체 또는 항체의 항원-결합 부분, 또는 효소적 생체분자, 예컨대 펩티다제, 아미노트랜스퍼라제, 리보자임, 압타자임, 또는 tRNA 신테타제, 예컨대 아미노아실-tRNA 신테타제, 및 2016년 9월 2일에 출원된 발명의 명칭 "MOLECULES AND METHODS FOR ITERATIVE POLYPEPTIDE ANALYSIS AND PROCESSING"의 미국 특허 출원 번호 15/255,433에 기재된 관련 분자일 수 있다.

[0134] 일부 실시양태에서, 본 출원의 친화성 시약 또는 인식 분자는 분해 경로 단백질이다. 인식 분자로서 사용하기에 적합한 분해 경로 단백질의 예는, 제한 없이, N-단부 규칙 경로 단백질, 예컨대 Arg/N-단부 규칙 경로 단백질, Ac/N-단부 규칙 경로 단백질, 및 Pro/N-단부 규칙 경로 단백질을 포함한다. 일부 실시양태에서, 인식 분자는 Gid4 단백질, Ubr1 UBR 박스 단백질, 및 ClpS 단백질 (예를 들어, ClpS2)로부터 선택된 N-단부 규칙 경로 단백질이다.

[0135] 프로테아제 또는 프로테이나제로도 지칭되는 펩티다제는 펩티드 결합의 가수분해를 촉매하는 효소이다. 펩티다제는 폴리펩티드를 보다 짧은 단편으로 소화시키고, 일반적으로 각각 내부 및 말단에서 폴리펩티드 쇄를 절단하는 엔도펩티다제 및 엑소펩티다제로 분류될 수 있다. 일부 실시양태에서, 표지된 친화성 시약은 엑소펩티다제 또는 엔도펩티다제 활성을 불활성화시키기 위해 변형된 펩티다제를 포함한다. 이러한 방식으로, 표지된 친화성 시약은 또한 폴리펩티드로부터 아미노산을 절단하지 않으면서 선택적으로 결합한다. 또 다른 실시양태에서, 엑소펩티다제 또는 엔도펩티다제 활성을 불활성화시키기 위해 변형되지 않은 펩티다제가 사용될 수 있다. 예를 들어, 일부 실시양태에서, 표지된 친화성 시약은 표지된 엑소펩티다제를 포함한다.

[0136] 본 출원의 특정 실시양태에 따르면, 폴리펩티드 서열분석 방법은 폴리펩티드의 말단 단부에서의 반복 검출 및 절단을 포함할 수 있다. 일부 실시양태에서, 표지된 엑소펩티다제는 아미노산의 검출 및 절단 단계 둘 다를 수행하는 단일 시약으로서 사용될 수 있다. 일반적으로 도시된 바와 같이, 일부 실시양태에서, 표지된 엑소펩티다제는 폴리펩티드로부터 각각 N-말단 또는 C-말단 아미노산에 선택적으로 결합하고 이를 절단하도록 아미노펩티다제 또는 카르복시펩티다제 활성을 갖는다. 특정 실시양태에서, 표지된 엑소펩티다제는 본원에 기재된 바와 같이, 표지된 엑소펩티다제가 비-절단 표지된 친화성 시약으로서 사용하기 위한 선택적 결합 특성을 보유하도록 관련 기술분야의 통상의 기술자에 의해 촉매적으로 불활성화될 수 있다는 것이 인지되어야 한다.

[0137] 엑소펩티다제는 일반적으로 폴리펩티드 기질이 그의 아미노-말단에 유리 아미노 기 또는 그의 카르복시-말단에 유리 카르복실 기 중 적어도 1개를 포함할 것을 요구한다. 일부 실시양태에서, 본 출원에 따른 엑소펩티다제는 폴리펩티드의 말단에서 또는 그 근처에서 결합을 가수분해한다. 일부 실시양태에서, 엑소펩티다제는 폴리펩티드 말단으로부터 3개 이하의 잔기의 결합을 가수분해한다. 예를 들어, 일부 실시양태에서, 엑소펩티다제에 의해 촉매되는 단일 가수분해 반응은 폴리펩티드 말단 단부로부터 단일 아미노산, 디펩티드, 또는 트리펩티드를 절단한다.

[0138] 일부 실시양태에서, 본 출원에 따른 엑소펩티다제는 각각 아미노- 또는 카르복시-말단으로부터 단일 아미노산을 절단하는 아미노펩티다제 또는 카르복시펩티다제이다. 일부 실시양태에서, 본 출원에 따른 엑소펩티다제는 각

각 아미노- 또는 카르복시-말단으로부터 디펩티드를 절단하는 디펩티딜-펩티다제 또는 펩티딜-디펩티다제이다. 또 다른 실시양태에서, 본 출원에 따른 엑소펩티다제는 아미노-말단으로부터 트리펩티드를 절단하는 트리펩티딜-펩티다제이다. 각각의 부류 또는 그의 하위부류의 펩티다제 분류 및 활성은 널리 공지되어 있고, 문헌에 기재되어 있다 (예를 들어, 문헌 [Gurupriya, V. S. & Roy, S. C. *Proteases and Protease Inhibitors in Male Reproduction. Proteases in Physiology and Pathology* 195-216 (2017); 및 Brix, K. & Stoecker, W. *Proteases: Structure and Function. Chapter 1*] 참조).

[0139] 본 출원에 따른 엑소펩티다제는 서열분석 반응의 방향성에 기초하여 선택 또는 조작될 수 있다. 예를 들어, 폴리펩티드의 아미노-말단에서 카르복시-말단으로의 서열분석의 실시양태에서, 엑소펩티다제는 아미노펩티다제 활성을 포함한다. 반대로, 폴리펩티드의 카르복시-말단에서 아미노-말단으로의 서열분석의 실시양태에서, 엑소펩티다제는 카르복시펩티다제 활성을 포함한다. 표지된 엑소펩티다제로서 사용될 수 있거나 또는 본원에 기재된 비-절단 표지된 친화성 시약으로서 사용되도록 불활성화될 수 있는, 특이적 카르복시-말단 아미노산을 인식하는 카르복시펩티다제의 예는 문헌에 기재되어 있다 (예를 들어, 문헌 [Garcia-Guerrero, M.C., et al., (2018) *PNAS* 115(17)] 참조).

[0140] 절단 시약 및/또는 친화성 시약 (예를 들어, 인식 분자)으로서 사용하기에 적합한 펩티다제는 1종 이상의 유형의 아미노산에 선택적으로 결합하는 아미노펩티다제를 포함한다. 일부 실시양태에서, 아미노펩티다제 인식 분자는 아미노펩티다제 활성을 불활성화시키기 위해 변형된다. 일부 실시양태에서, 아미노펩티다제 절단 시약은 폴리펩티드의 말단 단부로부터 대부분의 또는 모든 유형의 아미노산을 절단하도록 비-특이적이다. 일부 실시양태에서, 아미노펩티다제 절단 시약은 폴리펩티드의 말단 단부의 다른 유형의 아미노산과 비교하여 폴리펩티드의 말단 단부로부터 하나 이상의 유형의 아미노산을 절단하는데 있어서 보다 효율적이다. 예를 들어, 본 출원에 따른 아미노펩티다제는 알라닌, 아르기닌, 아스파라긴, 아스파르트산, 시스테인, 글루타민, 글루탐산, 글리신, 히스티딘, 이소류신, 류신, 리신, 메티오닌, 페닐알라닌, 프롤린, 셀레노시스테인, 세린, 트레오닌, 트립토판, 티로신, 및/또는 발린을 특이적으로 절단한다. 일부 실시양태에서, 아미노펩티다제는 프롤린 아미노펩티다제이다. 일부 실시양태에서, 아미노펩티다제는 프롤린 아미노펩티다제이다. 일부 실시양태에서, 아미노펩티다제는 글루타메이트/아스파르테이트-특이적 아미노펩티다제이다. 일부 실시양태에서, 아미노펩티다제는 메티오닌-특이적 아미노펩티다제이다. 일부 실시양태에서, 아미노펩티다제는 표 1에 제시된 아미노펩티다제이다. 일부 실시양태에서, 아미노펩티다제 절단 시약은 표 1에 제시된 펩티드 기질을 절단한다.

[0141] 일부 실시양태에서, 아미노펩티다제는 비-특이적 아미노펩티다제이다. 일부 실시양태에서, 비-특이적 아미노펩티다제는 아연 메탈로프로테아제이다. 일부 실시양태에서, 비-특이적 아미노펩티다제는 표 2에 제시된 아미노펩티다제이다. 일부 실시양태에서, 비-특이적 아미노펩티다제는 표 2에 제시된 펩티드 기질을 절단한다.

[0142] 따라서, 일부 실시양태에서, 본 출원은 표 1 또는 표 2로부터 선택된 아미노산 서열을 갖는 (또는 표 1 또는 표 2로부터 선택된 아미노산 서열에 대해 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 80-90%, 90-95%, 95-99%, 또는 그 초과)의 아미노산 서열 동일성을 갖는 아미노산 서열을 갖는) 아미노펩티다제 (예를 들어, 아미노펩티다제 인식 분자, 아미노펩티다제 절단 시약)를 제공한다. 일부 실시양태에서, 아미노펩티다제는 표 1 또는 표 2에 열거된 아미노펩티다제에 대해 25-50%, 50-60%, 60-70%, 70-80%, 80-90%, 90-95%, 또는 95-99%, 또는 그 초과)의 아미노산 서열 동일성을 갖는다. 일부 실시양태에서, 아미노펩티다제는 변형된 아미노펩티다제이고, 표 1 또는 표 2에 제시된 서열에 비해 1개 이상의 아미노산 돌연변이를 포함한다.

[0143] 표 1. 아미노펩티다제의 비제한적 예.

명칭	SEQ ID NO:	서열
엘. 뉴모필라M1 아미노펩티다제 (Glu/Asp 특이적)	1	MGSSHHHHHHSSGLVPRGSHMMVKQGVFMKTDQSKVKKLSDYKSLDYF VIHVDLQIDLKKPVESKARLTVPVNLNVDSHSNDLVLDGENMTLVSLQ MNDNLKENEYELTKDSLIIKNIPQNTPTTIEMTSLLGENTDLFGLYETEGV ALVKAESGLRRVFYLPDRPDNLATYKTTIIANQEDYPVLLSNGVLEKKE LPLGLHSVTWLDVDPKPSYLFALVAGNLQRSVTYYQTKSGRELPIEFYVP PSATSKCDFAKEVLKEAMAWDERTFNLECALRQHMVAGVDKYASGASE PTGLNLFNTENLFASPETKTDLGILRVLEVVAHEFFHYWSGDRVTIRDWF NLPLKEGLTTFRAAMFREELFGTDLIRLLDGKNLDERAPRQSAytaVRSL YTAAYEKSADIFRMMMLFIGKEPFIEAVAKFFKDNDGGAVTLEDIESIS NSSGKDLRSFSLSWFTESGIPELIVTDELNPDTKYFLKIKTVNGRNRPIPI MGLLDSSGAEIVADKLLIVDQEEIEFQFENIQTPIPSLLRSFAPVHMKYE YSYQDLLLLMQFDTNLYNRCEAAKQLISALINDFCIGKKIELSPQFFAVYK ALLSDNSLNEWMLAELITLSLEELIENQDKPDFEKLNEGRQLIQNALANE LKTDYFNLLFRIQISGDDDKQKLKGFDLKQAGLRRKSKVCSYLLNVDFE KTKELILQFEDALGKNMTETALASMLCEINCEEADVALEDYHYWKN DPGAVNNWFSIQALAHSPDVIERVKKLMRHGDFDLNPNKVVALLGSFIK NPFGFHSVTGEGYQLVADAIFDLKINPTLAANLTEKFTYWDKYDVNRQ AMMISTLKIIYSNATSSDVRTMAKKGLDKVKEDLPLPIHLTFHGGSTMQD RTAQLIADGNKENAYQLH
이. 콜라이 메티오닌 아미노펩티다제 (Met 특이적)	2	MAHHHHHHMGTAISIKTPEDIEKMRVAGRLAAEVLEMIPEYVKPGVSTGE LDRICNDYIVNEQHAVSACLG YHGYPKSVCSINEVVCHGIPDDAKLLKD GDIVNIDVTVIKDGFGHDTSKMFIVGKPTIMGERLCRITQESLYALRMVK PGINLREIGAAIQKFVEAEGFSVVREYCGHGIGRGFHEEPQVLHYDSRETN VVLKPGMTFTIEMVNAAGKEIRTMKDGWTVKTKDRSLSAQYEHTIVVT DNGCEILTLRKDDTIPAIHSHD
엘. 스메그마티스 프롤린 아미노펩티다제 (Pro 특이적)	3	MAHHHHHHMGTLAANTNGPGSMLSRMPVSSRTVPFGDHETWVQVTTP NAQPHALPLIVLHGGPGMAHNYVANIALADETGRTVIHYDQVGCNGST HLDPADPFWTPQLFVDEFHAVCTALGIERYHVLGQSWGGMLGAELAVR QPSGLVSLAICNSPASMRLWSEAAGDLRAQLPAETRAALDRHEAAGTITH PDYQLAAAEFYRRHVCRVVPPTQDFADSVQMEAEPTVYHTMNGPNEF HVVGTLDGWSVIDRLPDVTAPVLVIAGEHDEATPKTWQPFVDHIPDVRS HVFPGTSHCTHLEKPEEFRAVVAQFLHQHDLAADARV
와이. 페스티스 프롤린 아미노펩티다제 (Pro 특이적)	4	MTQQEYQNNRQALLAKMAPGSAIIFAPEATRSADSEYPYRQNSDFSYL TGFNEPEAVLILVKSDETHNHSVLFNRIRDLTAELWFGRRLLGQEAAPT KLA VDRALPFDEINEQLYLLLNRLDVIYHAQQQYAYADNIVFAALEKLRH GFRKNLRAPATLTDWRPWLHEMRLFKSAEEIALLRRAGEISALAHTRAMEK CRPGMFEYQLEGEILHEFTRHGARYPAYNTIVGGGENGCILHYTENECEL RDGDLVLIDAGCEYRGYAGDITRTFPVNGKFTPAQRAVYDIVLAAINKSL TLFRPGTSIREVTEEVVRIMVVGLVELGILKGDIQLIAEQAHRPFFMHGL SHWLGMVDVHDVGDYGSDDRGRILEPGMVLTVEPGLYIAPDADVPPQYR GIGIRIEDDIVITATGNENLTASVVKDPDJDIEALMALNHAGENLYFQEH HHHHH
피. 푸리오수스 메티오닌 아미노펩티다제	5	MDTEKLMKAGEIAKKVREKAIKLARPGMLLELAESIEKMIMELGGKPA FPVNL.SINEIAAHYTPYKGDITVLKEGDYLLKIDVGVHIDGFIADTA VTVRVGMEEDELMEAAKEALNAAISVARAGVEIKELGKAIENEIRKRG FKPIVNL.SGHKIERYKLGAGISIPNIYRPHDNYVLKEGDVFAIEPFA TIGAGQVIEVPPTLIYMYVRDVPVRVAQARFLAKIKREYGTLPFA YRWLQNDMPEGQLKLALKTLEKAGAIYGYPVLKEIRNGIVAQFEHT IIVEKDSVIVTQDMINKSTLE
아에로모나스 소브리아 프롤린 아미노펩티다제	6	HMSSPLHYVLDGIHCEPHFFTVPDLHQPPDDEETITLFGRTL CRKDRLDDELPWLLYLQGGPGFGAPRPSANGGWIKRALQEFRVLLLD QRGTHGSTPIHAELL.AHLNPRQQADYLSHFRADSVIRDAELIREQL SPDHPWLLGQSGGGFCSLTLYLSLFPDSLHEVYLTGGVAPIGRSA DEVYRATYQRVADKNRAFFARFPHAQAIAANRLATHLQRH DVRLPNGQRLTVEQLQQQGLDLGASGAFEELYLLED AFIGEKLNP AFLYQVQAMQPFNTNPVFAILHELIYCEGAASHWAA ERVRGEFPALAWAQGKDF AFTGEMIFPWFMEQFRELIP LKEAAHLLAEKA

[0144]

		DWGPLYDPVQLARNKVPVACAVYAEDMYVEFDYSRETLKGLSNSRAWI TNEYEHNGLRVDGEQILDRLIRLNRDCLE
피로코쿠스 푸리오수스 프롤린 아미노펩티다제 (X-/Pro)	7	MKERLEKLVKFMDENSIDRVFIKPVNVYYFSGTSPLGGGYIIVDGDDEATL YVPELEYEMAKEESKLPVVKFKKFDEIYEILKNTETLGIEGTLSSYSMVNF KEKSNVKEFKKIDDVIKDLRIKTKEEIEIEKACEIADKAVMAAIEEITEGK REREVAAKVEYLMKMNGAEKPAFDTHASGHRSAHPGVASDKRIERGDL VVIDLGALYNHYNSDITRTIVVGGSPNEKQREIYEIVLEAQKRAVEAAKPG MTAKELDSIAREIIKEYGYGDYFIHSLGHGVGLEIHEWPRISQYDETVLKE GMVITIEPGIYIPKLGGVRIEDTVLITENGAKRLTKTERELL
엘리자베스킨기아 메닝고셉터카 프롤린 아미노펩티다제	8	MIPITTPVGNFKVWTKRFGTNPVKIKVLLHGGPAMTHEYMECFETFFQRE GFEFYEYDQLGSYYSDQPTDEKLWNIDRFVDEVEQVRKAHADKENFYV LGNSWGGILAMEYALKYQQNLKGLIVANMMASAPEYVKYAEVLSKQM KPEVLAEVRAIEAKKDYANPRYTELLFPNYAQHICRLKEWPDALNRSK HVNSTVYTLMQGPSELGMSSDARLAKWDIKNRLHEIATPTLMIGARYDT MDPKAMEEQSKLVQKGRYLCPNGSHLAMWDDQKVFMDGVKFIKDV DTKSFN
아에로모나스 소브리야 프롤린 아미노펩티다제	9	HMSSPLHYVLDGIHCEPHFFTVPDLHQQPDDEETITLFGRTLCKRDLDDDE LPWLLYLQGGPGFGAPRPSANGGWIKRALQEFRVLLLDQRGTGHSTPIHA ELLAHLNPRQQADYLSHFRAADSVRDAELIREQLSPDHPWSLLGQSFGGFC SLTYLSLFPDSLHEVYL TGGVAPIGRSADEVYRATYQRVADKNRAFFARF PHAQAIANRLATHLQRHDVRLPNGQRLTVEQLQQQGLDLGASGAFEELY YLLEDAFIGEKLNP AFLYQVQAMQPFNTNPVFAILHELIYCEGAASHWAA ERVRGEFPALAWAQGKDF AFTGEMIFPWMFEQFREL IPLKEAAHLLAEKA DWGPLYDPVQLARNKVPVACAVYAEDMYVEFDYSRETLKGLSNSRAWI TNEYEHNGLRVDGEQILDRLIRLNRDCLE
엔. 코노레아에 프롤린 이미노펩티다제	10	MYEIKQPFHSGYLVQVSEIHQIYWEESGNPDGVPVIFLHGGPGAGASPECRG FFNPDVFRIVIIDQRGCGRSHPYACAEDNTTWDLVADIEKVREMLGIGKW LVFGGSWGSTLSLAYAQTHPERVKGLVLRGIFLCRPSETAWLNEAGGVSR IYPEQWQKFVAPIAENRRNRLIEAYHGLLFHQDEEVCLSAKAWADWES YLIRFEPEGVDEDAYASLAARLENHYFVNGGWLQGDKAILNNIGKIRHIP TVIVQGRYDLCTPMQSAWELSKAFPEAELRVVQAGHCAFDPLADALVQ AVEDILPRL

[0145]

[0146] 표 2. 비-특이적 아미노펩티다제의 비제한적 예

명칭	SEQ ID NO:	서열
이. 콜라이 아미노펩티다제 N (아연 메탈로프로테아제) *	11	MGSSHHHHHHSSGENLYFQGHMTQQPQAKYRHDYRAPDYQITDIDLTFD LDAQKTVVTAVSQAVRHGASDAPLRLNGEDLKLVSVINDEPWTAWKE EEGALVISNLPERFTLKIINEISPAANTALEGLYQSGDALCTQCEAEGFRHIT YYLDRPDVLARFTTKIADKIKYPFLLSNGNRVAQGELENGRHWVQWQD FFPKPCYLFALVAGDFDVLDRDTFTTRSGREVALELYVDRGNLDRAPWAM TSLKNSMKWDEERFGLDYLDIYMIVA VDFNMGAMENKGLNIFNSKYV LARTDTATDKDYLDIERVIGHEYFHNWVTGNRVTCRDWFQLSLKEGLTVF RDQEFSSDLGSRVNRINNVRTMRGLQFAEDASPMAPHPDPDMVIEMNNF YTLTVYEKGAEVIRMIHTLLGEENFQKGMQLYFERHDGSAATCDDFVQA MEDASNVDLSHFRRWYSQSGTPIVTVKDDYNPETEQYTLTISQRTPTPTD QAEKQPLHIPFAIELYDNEGKVIPLQKGGHPVNSVLNVTQAEQTFVFDNV YFQPVALLCEFSAPVKLEYKWSQQQLTFLMRHARNDFSRWDAAQSLLA TYIKLNVARHQQQPLSLPVHVADAFRAVLLDEKIDPALAAEILTLPVNE MAELFDIIDPIAAEVREALTRTLATELADELLAIYNANYQSEYRVEHEDIA KRTLNRACLRLAFGETHLADVLVSKQFHEANNMTDALAALSAAVAAQL PCRDALMQEYDDKWHQNGLVMDKWFILQATSPAANVLETVRGLLQHRS FTMSNPNRIRSLIGAFAGSNPAAFHAEDGSGYLFVLEMLTDLNSRNPQVAS RLIEPLIRLKR YDAKRQEKMRAALEQLKGLLENLSGDLYEKITKALA
피. 팔시파룸 M1 아미노펩티다제**	12	PKIHYRKDYKPSGFIINQVTLNINIHDQETIVRSVLDMDISKHNVGEDLVFD GVGLKINEISINNKKLVEGEEYTYDNEFLTIFSKFVPKSKFAFSSEVIIHPET NYALTGLYKSKNIIVSQCEATGFRRTTFFIDRPDMMAKYDVTVTADKEYY PVLLSNGDKVNEFEIPGGRHGARFNDPPLKPCYLFVAVAGDLKHLSATYI TKYTKKKVELYVFSEKYYVSKLQWALECLKKSMAFDEDEYFGLDYDLSRL NLVAVSDFNVGAMENKGLNIFNANSLASKKNSIDFSYARILTVVGHEYF HQYTGNRVTLRDWFQLTLKEGLTVHRENLFSEEMTKTVTTRLSHVDLLR SVQFLEDSSPLSHPIRPESYVSMENFYTTTVYDKGSEVMRMVLTILGEEYY KKGFDIYIKKNDGNTATCEDFNYAMEQAYKMKKADNSANLNQYLLWFS QSGTPHVSFKYNYDAEKKQYSIHVNQYTKPDENQKEKKPLFIPISVGLINP ENGKEMISQTTLELTKESDTFVFNIAVKPIPSLFRGFSAPVYIEDQLTDEE RILLK YDSDAFVR YNSCTNIYMKQILMNYNEFLKAKNEKLESFQLTPVN AQFIDAIKYLLEDPHADAGFKSYIVSLPQDRYIINFVSNLDTDLADTKEYI YKQIGDKLNDVYYKMFKSLEAKADDLT YFNDESHVDFDQMNMRTLRNT LLSLLSKAQYPNILEIIEHHSKSPYPSNWLTSLSVSAFYFDKYFELYDKTYKL SKDDELLQEWLKT VSRSDRKDIYEILKKLENEVLKDSKNPNDIRAVYLPF TNLRLRFHDISGKG YKLI AE VITKTDKFNPMVATQLCEPFKLWNKLDTKR QELMLNEMNTMLQEPQISNNLKEYLLRLTNK
NPEPPS	13	MGSSHHHHHHSSGMWLAAPSLARRLLFLGPPPPPLLLL VFSRSSRRRL HSLGLAAMPEKRPFERLPADVSPINYSCLKPDLLDFTFEGKLEAAAQVR QATNQIVMNCADIDIITASYPEGDEEIHATGFNYQNEDEKVTLSPSTLQ TGTTGLKIDFVGELNDKMKGFYRSKYTTTPSGEVRYAAVTQFEATDARRA FPCWDEPAIKATFDISLVVPKDRVALSNMNVDRKPYPDENLVEVKFAR TPVMSTYLVAFVVGGEYDFVETRSGDGVCVRVYTPVGKAEQGGKFALEVA AKTLPFYKDYFNVPYPLPKIDLIAADFAAGAMENWGLVTYRETALLIDPK NSCSSSRQWVALVVGHEL AHQWFGNLVTMEWWTHLWLNEGFA SWIEY LCVDHCFPEYDIWTQFVSADYTRAQELDALDNSHPIEVSVGHPSEVDEIFD AISYSKGASVIRMLHDYIGDKDFKKGMNMYLTKFQQKNAATEDLWESLE NASGKPIAAVMNTWTKQMGFPLIYVEAEQVEDDRLLRLSQKKFCAGGSY VGEDCPQWMPITISTEDPNQAKLKILMDKPEMNVLKNVKPDQWVKL NLGTGVGFYRTQYSSAMLESLLPGIRDLSLPPVDRLGLQNDLFLSARAGIIST VEVLK VMEAFVNENPYTVWSDLSCLNGLSTLLSHTDFYEEIQEFVKD VFS PIGERLGWDPKPGEGHLDALLRGLVLGKLGKAGHKATLEEARRRFKDHV EGKQILSADLRSPVYLVTKHGDGTTLDIMLKLHKQADMQEEKNRIERVL GATLLPDLIQKVLTFALSEEVRPQD TVSVIGGVAGGSKHGRKAAWKFIKD NWEELYNRYQGGFLISRLIKLSVEGFAVDKMAGEVKAFFESHAPSAERTI QQCCENILLNAAWLKRDAESI HQYLLQRKASPTTV

[0147]

NPEPPS E366V	14	MGSSHHHHHHSSGMWLAAAAPSLARRLLFLGPPPPPLLLLVSRRSSRRRL HSLGLAAMPEKRPFERLPADVSPINYSCLKPDLLDFTFEGKLEAAAQVR QATNQIVMNCADIDIITASYAPEGDEEIHATGFNYQNEDEKVTLSFPSTLQ TGTGTLKIDFVGELENDKMGFYRSKYTTSPSGEVR YAAVTQFEATDARRA FPCWDEPAIKATFDISLVVPKDRVALSNMNVDRKPYDDENLVEVKFAR TPVMSTYLVAFVVGGEYDFVETRSGDGVVVRVYTPVGKAEQGGKFALEVA AKTLFPYKDYFNVPYPLPKIDLIAADFAAGAMENWGLVTYRETALLIDPK NSCSSSRQWVALVVGHVLAHQWFGNLVTMEWWTHLWLNEGFAWIEY LCVDHCFPEYDIWTQFVSADYTRAQELDALDNSHPIEVSVGHPSEVDEIFD AISYSKGASVIRMLHDYIGDKDFKKGMNMYLTKFQQKNAATEDLWESLE NASGKPIAAVMNTWTKQMGFPLIYVEAEQVEDDRLLRLSQKKFCAGGSY VGEDCPQWMVPITISTEDPNQAKLKILMDKPEMNVVLKNVCPDQWVKL NLGTGVGFYRTQYSSAMLESLLPGIRDLSLPPVDRLGLQNDLFLSARAGIIST VEVLKVMFAFVNENYTVWSDLSCLNLGILSTLLSHTDFYEEIQEFVKDVFVS PIGERLGDWPKPGEGHLDALLRGLVLGKLGKAGHKATLEEARRRFKDHV EGKQILSADLRSPVYLVTKHGDGTTLDIMLKLHKQADMQEEKNRIERVL GATLLPDLIQKVLTFALSEEVRPQDTSVIGGVAGGSKHGRKAAWKFIKD NWEELYNRYQGGFLISRLIKLSVEGFAVDKMAGEVKAFFESHAPAPSAERTI QQCCENILLNAAWLKRDAESIHQYLLQRKASPTTV
프란시셀라 툴라렌시스 아미노펩티다제 N	15	MIYEFVMTDPKIKYKDYKPSNYLIDETHLIFELDESKTRVTANLYIVANR ENRENNTLVLDGVELKLLSIKLNKHLSPAFAVNENQLIINNVPKEFVLQ TVVEINPSANTSLLEGLYKSGDVSTQCEATGFRKITYYLDPRDVMMAFTV KIIADKKKYPIILSNGDKIDSGDISDNQHFVAVKDPFKKPCYLFALVAGDL ASIKDTYITKSQRKVSLEIYAFKQDIDKCHYAMQAVKDSMKWDEDRFGL EYDLDTFMIVAVPDFNAGAMENKGLNIFNTK YIMASNKTATDKDFELVQ SVVGHEYFHNWTDGDRVTCRDWFQSLKEGLTVFRDQEFSTDLSNRDVKR IDDVRIIRSAQFAEDASPMSPHIRPESYIEMNNFYTVT VYNKGAEIIRMIHTL LGEEGFQKGMKLYFERHDGQAVTCDDFVNAMADANNRDFSLFKRWYA QSGTPNIKVSENYDASSQTYSLTLEQTTLPTADQKEKQALHIPVKMGLINP EGKNIAEQVIELKEQKQTYTFENIAAKPVASLFRDFSAPVKVEHKRSEKDL LHIVKYDNNAFNRWDSLQIATNIILNNADLNDEFLNAFKSILHDKDLDK ALISNALLIPIESTIAEAMRVIMVDDIVLSRKNVNVNQLADKLKDDWLAVY QQCNDNKPYSLSAEQIAKRKLKGVCLSYLMNASDQKVGTDLAQQLFDN ADNMTDQQTAFTELLKSNDKQVRDNAINEFYNRWRHEDLVVNKWLLSQ AQISHESALDIVKGLVNHPAYNPKNPNKVYSLIGGFANFLQYHCKDGLG YAFMADTVLALDKFNHQVAARMARNLMSWKRYDSRQAMMKNALEKI KASNPSKNVFEIVSKSLES
피로코쿠스 호르코쉬이 TET 아미노펩티다제	16	MGSSHHHHHHSSGMEVRNMVDYELLKKVVEAPGVSGYEFLGIRDVVIEE IKDYVDEVKVDKLGNVIAHKKGEQPKVMIAAHMDQIGLMVTHIEKNGFL RVAPIGGVDPKTLIAQRFKVWIDKGFYIGVGASVPPHIQKPEDRKKAPD WDQIFIDIGAESKEEAEDMGVKIGTVITWDGRLERL GKHRFVSIAFDDRIA VYTILEVAKQLKDAKADVYFVATVQEEVGLRGARTSAFGIEPDYGF AIDV TIAADIPGTPEHKQVTHLGKGTAIKIMDRSVICHPTIVRWLEELAKKHEIPY QLEILLGGGTDAIAHLTKAGVPTGALSVPARYIHSNTEVVDERDV DATV ELMTKALENIHELKI
티. 아쿠아티쿠스 아미노펩티다제 T	17	MDAFTENLNKLAE LAIRVGLNLEEGQEIVATAPIEAVDFVRLLA EKA YEN GASLFTVL YGDNLIARKRLAL VPEAHLDRAPAWLYEGMAKAFHEGAARL AVSGNDPKALEGLPPERVGRAQQAQSRAYRPTLSAITEFTVNWTIVPFAH PGWAKAVFPGLPEEEAVQRLWQAIFQATRVDQEDPVAAWEAHRNVLHA KVAFLNEKRFHALHFQGP GTDLTVGLAEGHLWQGGATPTKKGRLCNPNL PTEEVFTAPHRERVEGVVRASRPLALSGQLVEGLWARFEGGVAVEVGAE KGEEVLKLLDTDEGARRLGEVALVPADNPIAKTGLVFFDTLFDENAASH IAFGQAYAEENLEGRPSGEEFRRRGGNESMVHVVDWMIGSEEVDVDGLLED GTRVPLMRRGRWVI
바실루스 스테아로썬모필루 스 펩티다제 M28	18	MAKLDETLTMLKAL TDAKGVPGNEREARDVMKTYIAPYADEVTTDGLG SLIAKKEGKSGGPKVMIAAGHLDEVGFMVTQIDDKGFIRFQTLGGWWSQV MLAQRVTIVTKKGDITGVIGSKPPHILPSEARKKPVEIKDMFIDIGATSREE AMEWGVRPGDMIVPYFEFTVLNNEKMMLAKAWDNRIGCAVAIDVLKQL KGVDPNTVYGVGTVQEEVGLRGARTAAQFIQPDIAFAVDVGIAGDTPG VSEKEAMGKLGAGPHIVLYDATMVSHRGLREFVIEVAEELNIPHHFDAMP GVGTDAIAHLTGIGVPSLTIAIPTRYIHSHAAILHRDDYENTVKLLVEVIK RLDADKVKQLTFDE

[0148]

비브리오 콜레라 아미노펩티다제	19	MEDKVVWISMGADAVGSLNPALSESLPHSFASGSQVWIGEVAIDELAEALS HTMHQHNRCGGYMVHTSAQGAMAALMPESIANFTIPAPSQQDLVNA WLPQVSADQITNTIRALSSFNNRFTTTSGAQASDWLANEWRSLISSLPGS RIEQIKHSGYNQKSVVLTIQSEKPDDEWVIVGGHLDSTLGSHTNEQSIAPG ADDDASGLASLSEIRVLRDNNFRPKRSVALMAYAAEEVGLRGSQDLANQ YKAQGGKKVSVLQLDMTNYRGS AEDIVFITDYTDSNLTQFLTTLIDEYLP ELTYGYDRCGYACSDHASWHKAGFSAAMPFESKFKDYNPKIHTSQDTLA NSDPTGNHAKVFTKLGLAYVIEMANAGSSQVPDDSVLQDGTAKINLSGA RGTQKRFTFELSQSKPLTIQTYGGSGDVLVYKYGSAPSKSNWDCRPYQN GNRETCSFNNAQPGIYHVMLDGYTNYNDVALKASTQHIIHHHHH
포토박테리움 할로볼테란스 아미노펩티다제j	20	MEDKVVWISIGSDASQTVKSVMQSNARSLLPESLASNGPVWVGQVDYSQ AELSHHMHEDHQRCGGYMVHSSPESAIAASNMPQSLVAFSPEISQQDTV NAWL PQVNSQAITGTITSLTSFINRFTTTSGAQASDWLANEWRSL SASLP NASVRQVSHFGYNQKSVVLTIQSEKPDDEWVIVGGHLDSTIGSHTNEQSV APGADDDASGLASVTEIRVLSENFPKRSIAFMAYAAEEVGLRGSQDLA NQYKAEGKQVISALQLDMTNYKGSVEDIVFITDYTDSNLTTFLSQLVDEY LPSLTYGFDTCGYACSDHASWHKAGFSAAMPFEAKFNDYNPMIHTPNDT LQNSDPTASHAVKFTKLGLAYAIEMASTTGGTPPPTGNVLDKGVVNGLS GATGSQVHYSFELPAQKNLQISTAGSGDVLVYVSGSEATKQNWDCRP YRNGNNEVCTFAGATPGTYSIMLDGYRQFSGVTLKASTQHIIHHHHH
에르시니아 페스트스 아미노펩티다제N	21	MTQQPQAKYRHDYRAPDYTTITDIDLDFALDAQKTTVTAVSKVKRQGTDV TPLILNGEDLTLISVSDGQAWPHYRQQDNTLVIEQLPADFTLTIVNDIHPA TNSALEGLYLSGEALCTQCEAEGFRHITYYLRDPDVLARFTTRIVADKSR PYLLSNGNRVQGQELDDGRHWKWEDEFPKPSYLFALVAGDFDVLQDK FITRSGREVALEIFVDRGNLDRADWAMTSLKNSMKWDETRFGLDYLDI YMIVA VDDFNMGAMENKGLNVFNSKYVLAKAETATDKDYL NIEA VIGHE YFHNWTGNRVT CRDWFQLSLKEGLTVFRDQEFSSDLGSRSVNRIENVRV MRAAQF AEDASPMMAHAIRPDKVIEMNNFYTLTVYEKGSEVIRMMHTLLG EQQFQAGMRLYFERHDGSAATCDDFVQAMEDVSNVDLSLFRRWYSQSG TPLLVHDDYDVEKQQYHLFVSQKTLPTADQPEKLPLHIPLDIELYDSKGN VIPLQHNGPLVHHVNLVTEAEQTFTFDNVAQKPIPSLLREFSAPVKLDYYPY SDQQLTFLMQHARNEFSRWDAQSLLATYIKLVNAKYQQQPLSLPAHV ADAFRAILLDEHLDPALAAQILTLPSENEMAELFTTIDPQAISTVHEAITRC LAQELSDPELLAVYVANMTPVYRIEHGDIKRALRNTCLNYLAFGDEEFAN KLVSLQYHQADNMTDSLAALAAAVAAQLPCRDELLAAFDVVRWNHDGL VMDKWFALQATSPAANVLVQVRLLKHPAFSLSNPNRTRSLIGSFASGNP AAFHAADGSGYQFLVEILSDLNTRNPQVAARLIEPLIRLKRYDAGRQALM RKALEQLKTL DNLSGDLYEKIKALAAHHHHHHH
비브리오 양구일라룸 아미노펩티다제	22	MEEKVWISIGGDATQALRSGAQSLPENLINQTSVWVGQVPVSELATLS HEMHENHQRCGGYMVHPSAQSAMSVSAMPLNLNAFSAPETQTTVNA WLPSVSAQQTSTITTLTQFKNRFTTSTGAQASNWIADHWRSLSASLPAS KVEQITHSGYNQKSVMLTIQSEKPDDEWVIVGGHLDSTLGSRTNESSIAPG ADDDASGLAGVTEIRLLSEQNFRPKRSIAFMAYAAEEVGLRGSQDLANRF KAEGKKVMSVMQLDMTNYQGSREDIVFITDYTDSNFTQYLTQLLDEYLP SLTYGFDTCGYACSDHASWHAVGYPAAMPFESKFNNDYNPNIHSPQDTLQ NSDPTGFHAKVFTKLGLAYVVMGNASTPPTPSNQLKNGVPVNGLSASR NSKTYWYQFELQEAGNLSIVLSGGSGDADLVYKYQTDADLQQYDCRPYRS GNNETCQFSNAQPGRYSL LHGYNNYSNASLVANAQHIIHHHHH
살리니비브리오 spYCSC6 아미노펩티다제	23	MEDKKVWISIGADAQQTALSSGAQPLLAQSVAHNGQAWIGEVSESELA LSHEMHENHQRCGGYIVHSSAQSAMASNMPLSRASFIAPASQQALVTP WISQIDSALIVNTIDRLTDFPNRFTTTSGAQASDWIKQRWQSLSAGLAGA SVTQISHSGYNQASVMLTIEGSESPDEWVIVGGHLDSTIGSRTNEQSIAPG ADDDASGLAAVTEVIRVLAQNFPKRSIAFVAYAAEEVGLRGSQDVAN QFKQAGKDV RGVQLDMTNYQGS AEDIVFITDYTDNQLTQYLTQLLDEY LPTLNYGFDTCGYACSDHASWHQVGYPAAMPFEAKFNDYNPNIHPTQDT LANSDSEGAHAAKFTKLGLAYTVELANADSSPNPGNELKLGEPIGNSGA RGNEKYFNRYLDQSGELVIRTYGGSGDVLVYKANGDVSTGNWDCRPY RSGNDEVCRFDNATPGNYAVMLRGYRTYDNVSLIVEHHIIHHHHH
비브리오 프로테올리터쿠스 아미노펩티다제 I	24	GMPPITQQATVTAWL PQVDASQITGTISLESFTNRFYTTTSGAQASDWIA SEWQLSASLPNASVKQVSHSGYNQKSVVMITITGSEAPDEWVIVGGHLD TIGSHTNEQSVAPGADDDASGLAAVTEVIRVLSNNFPKRSIAFMAYAAE EVGLRGSQDLANQYKSEKGNVVSALQLDMTNYKGS AQDVVFITDYTDS

		NFTQYLTQLMDEYLPSTLYGFDTCGYACSDHASWHNAGYPAAMPFESKF NDYNPRIHTTQDTLANS DPTGSHAKKFTQLGLAYAIEMGSATGDTPTPGN QLEHHHHHHH
파. 푸리오수스 아미노펩티다제 I	25	MVDWELMKKIIIESPGVSGYEHLGIRDLVVDILKDVADDEVKIDKLG NVIAH FKGSAPKVMVAAHMDKIGLMVNHIIDKGYLRVVPVIGGVLPETLIAQKIRF FTEKGERYGVVGVLPPLRLREAKDQGGKIDWDSIIVDVGASSREAEEMG FRIGTIGEFAPNFTRLSEHRFATPYLDDRICLYAMIEAARQLGEHEADIIYV ASVQEEIGLRGARVASFAIDPEVGIAMDVTF AKQPNDKGKIVPELGKGPV MDVGPINPNKLRQFADEVAKKYEIPLQVEPSRPRTGTDANVMQINREGVA TAVLSIPRIYMHSQVELADARDVDNTIKLAKALLEELKPMDFTPLEHHHHH HH

*절단 효율 (최대에서 최소로): 아르기닌 > 리신 > 소수성 잔기 (알라닌, 류신, 메티오닌, 및 페닐알라닌 포함)
> 프롤린 (예를 들어, 문헌 [Matthews Biochemistry 47, 2008, 5303-5311] 참조).

**절단 효율 (최대에서 최소로): 류신 > 알라닌 > 아르기닌 > 페닐알라닌 > 프롤린; 글루타메이트 및 아스파르

테이트 다음은 절단하지 않음.

- [0153] 2개 이상의 아미노산 서열을 비교하기 위한 목적으로, 제1 아미노산 서열과 제2 아미노산 서열 사이의 "서열 동일성"의 백분율 (또한 본원에서 "아미노산 동일성"으로도 지칭됨)은 [제2 아미노산 서열 내의 상응하는 위치에 서의 아미노산 잔기와 동일한 제1 아미노산 서열 내의 아미노산 잔기의 수]를 [제1 아미노산 서열 내의 아미노산 잔기의 총수]로 나누고 [100]을 곱함으로써 계산될 수 있고, 여기서 제1 아미노산 서열과 비교하여 제2 아미노산 서열 내의 아미노산 잔기의 각각의 결실, 삽입, 치환 또는 부가는 단일 아미노산 잔기 (위치)에서의 차이로 간주된다. 대안적으로, 2개의 아미노산 서열 사이의 서열 동일성의 정도는 공지된 컴퓨터 알고리즘을 사용하여 (예를 들어, 문헌 [Smith and Waterman (1970) Adv. Appl. Math. 2:482c]의 국부 상동성 알고리즘에 의해, 문헌 [Needleman and Wunsch, J. Mol. Biol. (1970) 48:443]의 상동성 정렬 알고리즘에 의해, 문헌 [Pearson and Lipman, Proc. Natl. Acad. Sci. USA (1998) 85:2444]의 유사성 검색 방법에 의해, 또는 블라스트(Blast), 클러스탈 오메가(Clustal Omega), 또는 다른 서열 정렬 알고리즘으로서 이용가능한 알고리즘의 컴퓨터 구현에 의해), 및 예를 들어 표준 설정을 사용하여 계산될 수 있다. 통상적으로, 상기 약술된 계산 방법에 따라 2개의 아미노산 서열 사이의 "서열 동일성"의 백분율을 결정하기 위한 목적으로, 최대 개수의 아미노산 잔기를 갖는 아미노산 서열은 "제1" 아미노산 서열로 취급될 것이고, 다른 아미노산 서열은 "제2" 아미노산 서열로 취급될 것이다.
- [0154] 추가적으로 또는 대안적으로, 2개 이상의 서열은 서열 사이의 동일성에 대해 평가될 수 있다. 2개 이상의 핵산 또는 아미노산 서열과 관련하여 용어 "동일한" 또는 퍼센트 "동일성"은 동일한 2개 이상의 서열 또는 하위서열을 지칭한다. 2개의 서열은, 상기 서열 비교 알고리즘 중 1개를 사용하여 또는 수동 정렬 및 육안 검사에 의해 측정된 바와 같이, 비교 윈도우 또는 지정된 영역에 걸쳐 최대 상응도를 위해 비교 및 정렬된 경우에, 2개의 서열이 명시된 영역에 걸쳐 또는 전체 서열에 걸쳐 동일한 (예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, 99.7%, 99.8%, 또는 99.9% 동일한) 아미노산 잔기 또는 뉴클레오타이드의 명시된 백분율을 갖는 경우에 "실질적으로 동일하다". 임의로, 동일성은 적어도 약 25, 50, 75, 또는 100개 아미노산 길이인 영역에 걸쳐, 또는 100 내지 150, 150 내지 200, 100 내지 200, 또는 200개 또는 그 초과 아미노산 길이인 영역에 걸쳐 존재한다.
- [0155] 추가적으로 또는 대안적으로, 2개 이상의 서열이 서열 사이의 정렬에 대해 평가될 수 있다. 2개 이상의 핵산 또는 아미노산 서열과 관련하여 용어 "정렬" 또는 퍼센트 "정렬"은 동일한 2개 이상의 서열 또는 하위서열을 지칭한다. 2개의 서열은, 상기 서열 비교 알고리즘 중 1개를 사용하여 또는 수동 정렬 및 육안 검사에 의해 측정된 바와 같이, 비교 윈도우 또는 지정된 영역에 걸쳐 최대 상응도를 위해 비교 및 정렬된 경우에, 2개의 서열이 명시된 영역에 걸쳐 또는 전체 서열에 걸쳐 동일한 (예를 들어, 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, 99.7%, 99.8% 또는 99.9% 동일한) 아미노산 잔기 또는 뉴클레오타이드의 명시된 백분율을 갖는 경우에 "실질적으로 정렬된다". 임의로, 정렬은 적어도 약 25, 50, 75, 또는 100개 아미노산 길이인 영역에 걸쳐, 또는 100 내지 150, 150 내지 200, 100 내지 200, 또는 200개 또는 그 초과 아미노산 길이인 영역에 걸쳐 존재한다.
- [0156] 폴리펩티드 분자에 더하여, 핵산 분자는 본 출원에 따라 친화성 시약 (예를 들어, 아미노산 인식 분자)으로서 사용하기 위한 다양한 유리한 특성을 보유한다.
- [0157] 핵산 압타머는 목적하는 표적에 높은 친화도 및 선택성으로 결합하도록 조작된 핵산 분자이다. 따라서, 핵산 압타머는 관련 기술분야에 공지된 선택 및/또는 풍부화 기술을 사용하여 목적하는 유형의 아미노산에 선택적으로 결합하도록 조작될 수 있다. 따라서, 일부 실시양태에서, 친화성 시약은 핵산 압타머 (예를 들어, DNA 압타머, RNA 압타머)를 포함한다. 일부 실시양태에서, 표지된 친화성 시약은 하나의 유형의 말단 아미노산에 선택적으로 결합하는 표지된 압타머이다. 예를 들어, 일부 실시양태에서, 표지된 압타머는 본원에 기재된 바와 같이 폴리펩티드의 말단에서 하나의 유형의 아미노산 (예를 들어, 단일 유형의 아미노산 또는 아미노산의 유형의 하위세트)에 선택적으로 결합한다. 제시되지는 않았지만, 표지된 압타머는 본 출원의 방법에 따라 폴리펩티드의 임의의 위치에서 (예를 들어, 폴리펩티드의 말단 위치에서 또는 말단 및 내부 위치에서) 하나의 유형의 아미노산에 선택적으로 결합하도록 조작될 수 있다는 것이 인지되어야 한다.
- [0158] 일부 실시양태에서, 표지된 친화성 시약은 결합-유도된 발광을 갖는 표지를 포함한다. 예를 들어, 일부 실시양태에서, 표지된 압타머는 공여자 표지 및 수용자 표지 및 기능을 포함한다. 또 다른 실시양태에서, 표지된 압타머는 켄칭 모이어터를 포함하고, 분자 비콘과 유사하게 기능하며, 여기서 표지된 압타머의 발광은 유리 분자로서 내부적으로 켄칭되고, 선택적으로 결합된 분자로서 회복된다 (예를 들어, 문헌 [Hamaguchi, et al.,

(2001) Analytical Biochemistry 294, 126-131] 참조). 이론에 얽매는 것을 원하지는 않지만, 결합-유도된 발광에 대한 이들 및 다른 유형의 메카니즘은 유리하게는 배경 발광을 감소시키거나 제거하여 본원에 기재된 방법의 전체 민감도 및 정확도를 증가시킬 수 있는 것으로 생각된다.

[0159] 폴리펩티드의 말단 아미노산을 확인하는 방법에 더하여, 본 출원은 표지된 친화성 시약을 사용하여 폴리펩티드를 서열분석하는 방법을 제공한다. 일부 실시양태에서, 서열분석 방법은 폴리펩티드 말단을 말단 아미노산 검출 및 말단 아미노산 절단의 반복 사이클에 적용하는 것을 수반할 수 있다. 예를 들어, 일부 실시양태에서, 본 출원은 폴리펩티드를 본원에 기재된 1종 이상의 표지된 친화성 시약과 접촉시키고, 폴리펩티드를 에드만 분해에 적용하는 것을 포함하는, 폴리펩티드의 아미노산 서열을 결정하는 방법을 제공한다.

[0160] 통상적인 에드만 분해는 폴리펩티드의 말단 아미노산을 변형시키고 절단하는 반복 사이클을 수반하며, 여기서 각각의 연속적으로 절단된 아미노산을 확인하여 폴리펩티드의 아미노산 서열을 결정한다. 통상적인 에드만 분해의 예시적인 예로서, 폴리펩티드의 N-말단 아미노산을 페닐 이소티오시아네이트 (PITC)를 사용하여 변형시켜 PITC-유도체화된 N-말단 아미노산을 형성한다. 이어서, PITC-유도체화된 N-말단 아미노산을 산성 조건, 염기성 조건, 및/또는 승온을 사용하여 절단한다. 또한, PITC-유도체화된 N-말단 아미노산을 절단하는 단계는 중성 또는 거의 중성 pH에서 비교적 보다 온화한 절단 조건을 수반하는 원충 트리파노소마 크루지(*Trypanosoma cruzi*)로부터의 변형된 시스테인 프로테아제를 사용하여 효소적으로 달성될 수 있는 것으로 밝혀졌다. 유용한 효소의 비제한적 예는 2016년 9월 2일에 출원된 발명의 명칭 "MOLECULES AND METHODS FOR ITERATIVE POLYPEPTIDE ANALYSIS AND PROCESSING"의 미국 특허 출원 번호 15/255,433에 기재되어 있다.

[0161] 일부 실시양태에서, 에드만 분해에 의한 서열분석은 링커를 통해 고체 지지체의 표면에 고정된 (예를 들어, 샘플 웰의 하부 또는 측벽 표면에 고정된) 폴리펩티드를 제공하는 것을 포함한다. 일부 실시양태에서, 본원에 기재된 바와 같이, 폴리펩티드는 다른 말단이 말단 아미노산의 검출 및 절단을 위해 유리되도록 하나의 말단 (예를 들어, 아미노-말단 아미노산 또는 카르복시-말단 아미노산)에서 고정된다. 따라서, 일부 실시양태에서, 본원에 기재된 에드만 분해 방법에 사용되는 시약은 폴리펩티드의 비-고정된 (예를 들어, 유리) 말단에서 말단 아미노산과 우선적으로 상호작용한다. 이러한 방식으로, 폴리펩티드는 검출 및 절단의 반복 사이클에 걸쳐 고정된 상태로 유지된다. 이를 위해, 일부 실시양태에서, 링커는 예를 들어 화학적 절단 조건 하에 표면으로부터 폴리펩티드의 탈착을 제한하기 위해 검출 및 절단에 사용되는 목적하는 조건의 세트에 따라 설계될 수 있다. 폴리펩티드를 표면에 고정시키기 위한 적합한 링커 조성물 및 기술은 본원의 다른 곳에 상세히 기재되어 있다.

[0162] 본 출원에 따르면, 일부 실시양태에서, 에드만 분해에 의한 서열분석 방법은 (i) 폴리펩티드를 1종 이상의 유형의 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약과 접촉시키는 단계를 포함한다. 일부 실시양태에서, 표지된 친화성 시약은 말단 아미노산에 선택적으로 결합함으로써 폴리펩티드와 상호작용한다. 일부 실시양태에서, 단계 (i)은 폴리펩티드의 말단 아미노산 (예를 들어, 유리 말단 아미노산)에 선택적으로 결합하지 않는 1종 이상의 표지된 친화성 시약 중 임의의 것을 제거하는 것을 추가로 포함한다.

[0163] 일부 실시양태에서, 방법은 표지된 친화성 시약을 검출함으로써 폴리펩티드의 말단 아미노산을 확인하는 단계를 추가로 포함한다. 일부 실시양태에서, 검출은 표지된 친화성 시약으로부터 발광을 검출하는 것을 포함한다. 본원에 기재된 바와 같이, 일부 실시양태에서, 발광은 표지된 친화성 시약과 고유하게 회합되고, 발광은 그에 의해 표지된 친화성 시약이 선택적으로 결합하는 아미노산의 유형과 회합된다. 따라서, 일부 실시양태에서, 아미노산의 유형은 표지된 친화성 시약의 1종 이상의 발광 특성을 결정함으로써 확인된다.

[0164] 일부 실시양태에서, 에드만 분해에 의한 서열분석 방법은 (ii) 폴리펩티드의 말단 아미노산을 제거하는 단계를 포함한다. 일부 실시양태에서, 단계 (ii)는 폴리펩티드로부터 표지된 친화성 시약 (예를 들어, 말단 아미노산에 선택적으로 결합하는 1종 이상의 표지된 친화성 시약 중 임의의 것)을 제거하는 것을 포함한다. 일부 실시양태에서, 단계 (ii)는 말단 아미노산 (예를 들어, 유리 말단 아미노산)을 이소티오시아네이트 (예를 들어, PITC)와 접촉시킴으로써 폴리펩티드의 말단 아미노산을 변형시켜 이소티오시아네이트-변형된 말단 아미노산을 형성하는 것을 포함한다. 일부 실시양태에서, 이소티오시아네이트-변형된 말단 아미노산은 비변형된 말단 아미노산보다 절단 시약 (예를 들어, 화학적 또는 효소적 절단 시약)에 의한 제거에 더 감수성이다.

[0165] 일부 실시양태에서, 단계 (ii)는 폴리펩티드를 이소티오시아네이트-변형된 말단 아미노산에 특이적으로 결합하고 이를 절단하는 프로테아제와 접촉시킴으로써 말단 아미노산을 제거하는 것을 포함한다. 일부 실시양태에서, 프로테아제는 변형된 시스테인 프로테아제를 포함한다. 일부 실시양태에서, 프로테아제는 변형된 시스테인 프로테아제, 예컨대 트리파노소마 크루지로부터의 시스테인 프로테아제를 포함한다 (예를 들어, 문헌 [Borgo, et al., (2015) Protein Science 24:571-579] 참조). 또 다른 실시양태에서, 단계 (ii)는 폴리펩티드를 이소티오

시아네이트-변형된 말단 아미노산을 절단하기에 충분한 화학적 (예를 들어, 산성, 염기성) 조건에 적용함으로써 말단 아미노산을 제거하는 것을 포함한다.

- [0166] 일부 실시양태에서, 에드만 분해에 의한 서열분석 방법은 (iii) 말단 아미노산 절단 후에 폴리펩티드를 세척하는 단계를 포함한다. 일부 실시양태에서, 세척은 프로테아제를 제거하는 것을 포함한다. 일부 실시양태에서, 세척은 폴리펩티드를 (예를 들어, 산성 또는 염기성 조건에 의한 화학적 절단 후에) 중성 pH 조건으로 회복시키는 것을 포함한다. 일부 실시양태에서, 에드만 분해에 의한 서열분석 방법은 단계 (i) 내지 (iii)을 복수의 사이클 동안 반복하는 것을 포함한다.
- [0167] 일부 실시양태에서, 폴리펩티드의 복합 혼합물 또는 풍부화된 혼합물 (예를 들어, 폴리펩티드의 혼합물)을 함유하는 샘플은 통상의 효소를 사용하여 대략 6 내지 40개 아미노산의 짧은 폴리펩티드 단편으로 분해될 수 있다. 일부 실시양태에서, 본 출원의 방법에 따른 이러한 폴리펩티드 라이브러리의 서열분석은 원래의 복합 혼합물 또는 풍부화된 혼합물에 존재하는 폴리펩티드의 각각의 정제 및 존재비를 밝혀낼 것이다. 본원 및 문헌에 기재된 바와 같이, 6 내지 40개 아미노산의 크기 범위의 대부분의 폴리펩티드는 폴리펩티드 쇄 내의 단지 4개의 아미노산의 수 및 위치를 결정함으로써 고유하게 확인될 수 있다.
- [0168] 따라서, 일부 실시양태에서, 에드만 분해에 의한 서열분석 방법은 4종의 DNA 압타머 유형을 포함하는 표지된 압타머 세트를 사용하여 수행될 수 있으며, 각각의 유형은 상이한 N-말단 아미노산을 인식한다. 각각의 압타머 유형은 상이한 압타머 유형이 1종 이상의 발광 특성에 기초하여 구별될 수 있도록 상이한 발광 표지로 표지될 수 있다. 예시적 목적을 위해, 표지된 압타머의 예시적인 세트는 제1 발광 표지 ("염료 1")로 표지된 시스템인-특이적 압타머; 제2 발광 표지 ("염료 2")로 표지된 리신-특이적 압타머; 제3 발광 표지 ("염료 3")로 표지된 트립토판-특이적 압타머; 및 제4 발광 표지 ("염료 4")로 표지된 글루탐에이트-특이적 압타머를 포함한다.
- [0169] 일부 실시양태에서, 단계 (i) 전에, 폴리펩티드 라이브러리로부터의 단일 폴리펩티드 분자는 고체 지지체의 표면에, 예를 들어, 샘플 웰의 어레이의 샘플 웰의 하부 또는 측면 표면에 고정된다. 일부 실시양태에서, 본원의 다른 곳에 기재된 바와 같이, 표면 고정을 가능하게 하거나 (예를 들어, 비오틴) 또는 용해도를 개선시키는 (예를 들어, 올리고뉴클레오타이드) 모이어티가 폴리펩티드의 C-말단에 화학적으로 또는 효소적으로 부착될 수 있다. 각각의 폴리펩티드의 서열을 결정하기 위해, 일부 실시양태에서, 고정된 폴리펩티드는 N-말단 아미노산 검출 및 N-말단 아미노산 절단의 반복 사이클에 적용된다. 일부 실시양태에서, 과정은 자동화된 유체 시스템을 사용하여 검출 표면 위의 플로우셀 내로의 주입에 의해 수행되는 시약 첨가 및 세척 단계를 포함한다. 일부 실시양태에서, 단계 (i) 내지 (iv)는 표지된 압타머를 사용한 검출 및 절단의 1 사이클을 예시한다.
- [0170] 일부 실시양태에서, 에드만 분해에 의한 서열분석 방법은 (i) 4종의 직교 표지된 DNA 압타머의 혼합물 중에 유동시키고, 압타머가 N-말단에서 4종의 정확한 아미노산 중 1종을 함유하는 임의의 고정된 폴리펩티드 (예를 들어, 어레이의 샘플 웰 내에 고정된 폴리펩티드)에 결합하게 하도록 인큐베이션하는 단계를 포함한다. 일부 실시양태에서, 방법은 고정된 폴리펩티드를 세척하여 비결합된 압타머를 제거하는 것을 추가로 포함한다. 일부 실시양태에서, 방법은 고정된 폴리펩티드를 영상화하는 것 ("영상화 단계 (i)")을 추가로 포함한다. 일부 실시양태에서, 획득된 영상은 압타머-결합된 폴리펩티드의 위치 (예를 들어, 샘플 웰의 어레이 내의 위치) 및 4종의 압타머 중 어떠한 것이 각각의 위치에서 결합되는지를 결정하기에 충분한 정보를 함유한다. 일부 실시양태에서, 방법은 적절한 완충제를 사용하여 고정된 폴리펩티드를 세척하여 고정된 폴리펩티드로부터 압타머를 제거하는 것을 추가로 포함한다.
- [0171] 일부 실시양태에서, 서열분석 방법은 (ii) N-말단 아민 기를 특이적으로 변형시키는 반응성 분자 (예를 들어, 제시된 바와 같은 PITC)를 함유하는 용액 중에 유동시키는 단계를 포함한다. 일부 실시양태에서, 이소티오시아네이트 분자, 예컨대 PITC는 N-말단 아미노산을 변형된 프로테아제, 예컨대 트리파노소마 크루지로부터의 시스템인 프로테아제 크루자인에 의한 절단을 위한 기질로 변형시킨다.
- [0172] 일부 실시양태에서, 서열분석 방법은 (iii) 고정된 폴리펩티드를 세척한 후, 고정된 폴리펩티드로부터 변형된 N-말단 아미노산을 인식하고 절단하는 적합한 변형된 프로테아제 중에 유동시키는 단계를 포함한다.
- [0173] 일부 실시양태에서, 방법은 (iv) 효소적 절단 후에 고정된 폴리펩티드를 세척하는 단계를 포함한다. 일부 실시양태에서, 단계 (i) 내지 (iv)는 에드만 분해의 1 사이클을 도시한다. 따라서, 제시된 바와 같은 단계 (i')는 단계 (i) 내지 (iv)에 대해 상기 기재된 바와 같이 수행되는 단계 (i') 내지 (iv')로서 진행되는 다음 반응 사이클의 시작이다. 일부 실시양태에서, 단계 (i) 내지 (iv)는 대략 20-40 사이클 동안 반복된다.
- [0174] 일부 실시양태에서, 표지된 이소티오시아네이트 (예를 들어, 염료-표지된 PITC)를 사용하여 샘플 로딩을 모니터

링할 수 있다. 예를 들어, 일부 실시양태에서, 폴리펩티드 샘플을 서열분석 방법에 적용하기 전에, 폴리펩티드 샘플은 염료-표지된 PITC를 사용한 말단 단부의 변형에 의해 말단 단부에서 발광 표지로 사전-접합된다. 이러한 방식으로, 폴리펩티드 샘플의 샘플 웰의 어레이 내로의 로딩은 상기 기재된 단계 (i) 전에 표지로부터의 발광을 검출함으로써 모니터링될 수 있다. 일부 실시양태에서, 발광은 어레이 내의 샘플 웰의 단일 점유율 (예를 들어, 단일 폴리펩티드 분자를 함유하는 샘플 웰의 분율)을 결정하는데 사용되며, 이는 유리하게는 주어진 샘플에 대해 신뢰성 있게 취득되는 정보의 양을 증가시킬 수 있다. 목적하는 샘플 로딩 상태가 발광에 의해 결정되면, 단계 (i)로 진행하기 전에, 기재된 바와 같이 화학적 또는 효소적 절단이 수행될 수 있다.

[0175] 일부 실시양태에서, 표지된 이소티오시아네이트 (예를 들어, 염료-표지된 PITC)는 어레이에서 폴리펩티드 샘플에 대한 반응 진행을 모니터링하는데 사용될 수 있다. 예를 들어, 일부 실시양태에서, 단계 (ii)는 샘플 중의 폴리펩티드의 N-말단 아민 기를 특이적으로 변형시키고 표지하는 염료-표지된 PITC를 함유하는 용액 중에 유도시키는 것을 포함한다. 일부 실시양태에서, 단계 (ii) 동안 또는 그 후에 표지로부터의 발광을 검출하여 샘플 중의 폴리펩티드의 N-말단 PITC 변형을 평가할 수 있다. 따라서, 일부 실시양태에서, 발광은 단계 (ii)에서 단계 (iii)로 진행하는지 여부 또는 진행하는 시기를 결정하는데 사용된다. 일부 실시양태에서, 단계 (iii) 동안 또는 그 후에 표지로부터의 발광을 검출하여 샘플 중의 폴리펩티드의 N-말단 아미노산 절단을 평가할 수 있고-예를 들어, 단계 (iii)에서 단계 (iv)로 진행하는지 여부 또는 진행하는 시기를 결정할 수 있다.

[0176] 서열분석 방법은 폴리펩티드의 말단 아미노산을 검출하고 절단하기 위한 별개의 시약을 이용할 수 있다. 그럼에도 불구하고, 일부 측면에서, 본 출원은 펩티다제 (예컨대 상이한 유형의 말단 아미노산에 선택적으로 결합하고 이를 절단하는 표지된 엑소펩티다제)를 포함하는 단일 시약이 폴리펩티드의 말단 아미노산을 검출 및 절단하는데 사용될 수 있는 서열분석 방법을 제공한다.

[0177] 표지된 엑소펩티다제는 제1 발광 표지를 포함하는 리신-특이적 엑소펩티다제, 제2 발광 표지를 포함하는 글리신-특이적 엑소펩티다제, 제3 발광 표지를 포함하는 아스파르트레이트-특이적 엑소펩티다제, 및 제4 발광 표지를 포함하는 류신-특이적 엑소펩티다제를 포함할 수 있다. 본원에 기재된 특정 실시양태에 따르면, 각각의 표지된 엑소펩티다제는 단지 그의 각각의 아미노산이 폴리펩티드의 아미노- 또는 카르복시-말단에 있는 경우에만 그러한 아미노산에 선택적으로 결합하고 이를 절단한다. 따라서, 이러한 접근법에 의한 서열분석이 펩티드의 하나의 말단에서 다른 말단으로 진행함에 따라, 표지된 엑소펩티다제는 모든 시약 세트가 아미노펩티다제 또는 카르복시펩티다제 활성을 보유하도록 조작되거나 선택된다.

[0178] 일부 측면에서, 본 출원은 말단 아미노산과 표지된 아미노산 인식 분자 (예를 들어, 표지된 친화성 시약) 및 표지된 절단 시약 (예를 들어, 표지된 비-특이적 엑소펩티다제)의 결합 상호작용을 평가함으로써 실시간으로 폴리펩티드를 서열분석하는 방법을 제공한다. 이론에 얽매이는 것을 원하지는 않지만, 표지된 친화성 시약은 결합의 회합률 또는 "온" 레이트 (k_{on}) 및 결합의 해리율 또는 "오프" 레이트 (k_{off})에 의해 정의되는 결합 친화도 (K_D)에 따라 선택적으로 결합한다. 속도 상수 k_{off} 및 k_{on} 은 각각 펄스 지속기간 (예를 들어, 검출가능한 결합 사건에 상응하는 시간) 및 펄스간 지속기간 (예를 들어, 검출가능한 결합 사건 사이의 시간)의 중요한 결정인자이다. 일부 실시양태에서, 이들 속도는 최상의 서열분석 정확도를 제공하는 펄스 지속기간 및 펄스 속도 (예를 들어, 신호 펄스의 빈도)를 달성하도록 조작될 수 있다.

[0179] 서열분석 반응 혼합물은 표지된 친화성 시약의 것과 상이한 발광 표지를 포함하는 표지된 비-특이적 엑소펩티다제를 추가로 포함할 수 있다. 일부 실시양태에서, 표지된 비-특이적 엑소펩티다제는 표지된 친화성 시약의 농도보다 낮은 농도로 혼합물에 존재한다. 일부 실시양태에서, 표지된 비-특이적 엑소펩티다제는 대부분의 또는 모든 유형의 말단 아미노산을 절단하도록 광범위한 특이성을 나타낸다.

[0180] 일부 실시양태에서, 표지된 비-특이적 엑소펩티다제에 의한 말단 아미노산 절단은 신호 펄스를 발생시키고, 이들 사건은 표지된 친화성 시약의 결합 펄스보다 더 낮은 빈도로 발생한다. 이러한 방식으로, 폴리펩티드의 아미노산은 실시간 서열분석 과정에서 카운팅 및/또는 확인될 수 있다. 일부 실시양태에서, 각각 상응하는 말단 아미노산을 확인하는데 사용될 수 있는 진단 펄싱 패턴 (예를 들어, 특징적인 패턴)을 갖는 복수의 표지된 친화성 시약이 사용될 수 있다. 예를 들어, 일부 실시양태에서, 상이한 특징적인 패턴은 상이한 유형의 말단 아미노산과 1종 초과 표지된 친화성 시약의 회합에 상응한다. 본원에 기재된 바와 같이, 1종 초과 유형의 아미노산과 회합하는 단일 친화성 시약이 본 출원에 따라 사용될 수 있다는 것이 인지되어야 한다. 따라서, 일부 실시양태에서, 상이한 특징적인 패턴은 상이한 유형의 말단 아미노산과 1종의 표지된 친화성 시약의 회합에 상응한다.

- [0181] 상기 상술된 바와 같이, 실시간 서열분석 과정은 일반적으로 말단 아미노산 인식 및 말단 아미노산 절단의 사이클을 수반할 수 있고, 여기서 인식 및 절단의 상대적 발생은 표지된 친화성 시약 및 표지된 비-특이적 엑소캡티다제 사이의 농도 차이에 의해 제어될 수 있다. 일부 실시양태에서, 농도 차이는 개별 아미노산의 인식 동안 검출된 신호 펄스의 수가 확인을 위한 목적하는 신뢰 구간을 제공하도록 최적화될 수 있다. 예를 들어, 초기 서열분석 반응이 절단 사건 사이에 너무 적은 신호 펄스를 갖는 신호 데이터를 제공하여 목적하는 신뢰 구간을 갖는 특징적인 패턴의 결정을 허용하지 않는 경우에, 서열분석 반응은 친화성 시약에 비해 감소된 농도의 비-특이적 엑소캡티다제를 사용하여 반복될 수 있다. 본 발명자들은 기재된 바와 같은 농도 차이 접근법과 조합하여 또는 그에 대안적으로 사용될 수 있는, 실시간 서열분석 반응을 제어하기 위한 추가의 기술을 인식하였다.
- [0182] 일부 실시양태에서, 서열분석 반응은 온도-의존성 말단 아미노산 인식 및 말단 아미노산 절단 사이클을 수반한다. 서열분석 반응의 각각의 사이클은 2가지 온도 범위: 엑소캡티다제 활성화에 비해 친화성 시약 활성화에 최적인 (예를 들어, 말단 아미노산 인식을 촉진하기 위한) 제1 온도 범위 (" T_1 "), 및 친화성 시약 활성화에 비해 엑소캡티다제 활성화에 최적인 (예를 들어, 말단 아미노산 절단을 촉진하기 위한) 제2 온도 범위 (" T_2 ")에 걸쳐 수행될 수 있다. 서열분석 반응은 반응 혼합물 온도를 제1 온도 범위 T_1 (아미노산 인식을 개시하기 위함)과 제2 온도 범위 T_2 (아미노산 절단을 개시하기 위함) 사이에서 교대시킴으로써 진행될 수 있다. 따라서, 온도-의존성 서열분석 과정의 진행은 온도, 및 수동 또는 자동화 과정을 통해 수행될 수 있는 상이한 온도 범위 사이 (예를 들어, T_1 과 T_2 사이)의 교대에 의해 제어가능하다. 일부 실시양태에서, 제2 온도 범위 T_2 와 비교하여 제1 온도 범위 T_1 내에서의 친화성 시약 활성화 (예를 들어, 아미노산에 대한 결합 친화성 (K_D))은 적어도 10-배, 적어도 100-배, 적어도 1,000-배, 적어도 10,000-배, 적어도 100,000-배, 또는 그 초과만큼 증가된다. 일부 실시양태에서, 제1 온도 범위 T_1 과 비교하여 제2 온도 범위 T_2 내에서의 엑소캡티다제 활성화 (예를 들어, 절단 생성물로의 기질 전환 속도)은 적어도 2-배, 10-배, 적어도 25-배, 적어도 50-배, 적어도 100-배, 적어도 1,000-배, 또는 그 초과만큼 증가된다.
- [0183] 일부 실시양태에서, 제1 온도 범위 T_1 은 제2 온도 범위 T_2 보다 낮다. 일부 실시양태에서, 제1 온도 범위 T_1 은 약 15°C 내지 약 40°C (예를 들어, 약 25°C 내지 약 35°C, 약 15°C 내지 약 30°C, 약 20°C 내지 약 30°C)이다. 일부 실시양태에서, 제2 온도 범위 T_2 는 약 40°C 내지 약 100°C (예를 들어, 약 50°C 내지 약 90°C, 약 60°C 내지 약 90°C, 약 70°C 내지 약 90°C)이다. 일부 실시양태에서, 제1 온도 범위 T_1 은 약 20°C 내지 약 40°C (예를 들어, 대략 30°C)이고, 제2 온도 범위 T_2 는 약 60°C 내지 약 100°C (예를 들어, 대략 80°C)이다.
- [0184] 일부 실시양태에서, 제1 온도 범위 T_1 은 제2 온도 범위 T_2 보다 더 높다. 일부 실시양태에서, 제1 온도 범위 T_1 은 약 40°C 내지 약 100°C (예를 들어, 약 50°C 내지 약 90°C, 약 60°C 내지 약 90°C, 약 70°C 내지 약 90°C)이다. 일부 실시양태에서, 제2 온도 범위 T_2 는 약 15°C 내지 약 40°C (예를 들어, 약 25°C 내지 약 35°C, 약 15°C 내지 약 30°C, 약 20°C 내지 약 30°C)이다. 일부 실시양태에서, 제1 온도 범위 T_1 은 약 60°C 내지 약 100°C (예를 들어, 대략 80°C)이고, 제2 온도 범위 T_2 는 약 20°C 내지 약 40°C (예를 들어, 대략 30°C)이다.
- [0185] 일부 실시양태에서, 본 출원은 발광-활성화된 시약을 사용하는 발광-의존성 서열분석 과정을 제공한다. 일부 실시양태에서, 발광-의존성 서열분석 과정은 발광-의존성 아미노산 인식 및 절단의 사이클을 수반한다. 서열분석 반응의 각각의 사이클은 서열분석 반응 혼합물을 2가지 상이한 발광 조건: 엑소캡티다제 활성화에 비해 친화성 시약 활성화에 최적인 (예를 들어, 아미노산 인식을 촉진하기 위한) 제1 발광 조건, 및 친화성 시약 활성화에 비해 엑소캡티다제 활성화에 최적인 (예를 들어, 아미노산 절단을 촉진하기 위한) 제2 발광 조건에 노출시킴으로써 수행될 수 있다. 서열분석 반응은 반응 혼합물을 제1 발광 조건에 노출시키는 것 (아미노산 인식을 개시하기 위함)과 반응 혼합물을 제2 발광 조건에 노출시키는 것 (아미노산 절단을 개시하기 위함) 사이에서 교대시킴으로써 진행된다. 예로서 제한 없이, 일부 실시양태에서, 2종의 상이한 발광 조건은 제1 파장 및 제2 파장을 포함한다.
- [0186] 일부 측면에서, 본 출원은 1종 이상의 표지된 친화성 시약과 말단 및 내부 아미노산의 결합 상호작용 및 표지된 비-특이적 엑소캡티다제와 말단 아미노산의 결합 상호작용을 평가함으로써 실시간으로 폴리펩티드를 서열분석하는 방법을 제공한다. 일부 실시양태에서, 말단 및 내부 위치 둘 다에서 1종의 유형의 아미노산에 선택적으로 결합하고 그로부터 해리되는 표지된 친화성 시약이 사용된다. 선택적 결합은 신호 출력에서 일련의 펄스를 발생시킨다. 그러나, 이러한 접근법에서, 일련의 펄스는 폴리펩티드 전반에 걸쳐 아미노산 유형의 수에 의해 결

정되는 속도로 발생한다. 따라서, 일부 실시양태에서, 결합 사건에 상응하는 펄싱 속도는 폴리펩티드에 현재 존재하는 동족 아미노산의 수를 진단할 것이다.

[0187] 표지된 비-특이적 펩티다제는 표지된 친화성 시약보다 상대적으로 더 낮은 농도로 존재할 수 있어서, 예를 들어 절단 사건 사이에 최적의 시간 윈도우를 제공할 수 있다. 추가적으로, 특정 실시양태에서, 표지된 비-특이적 펩티다제의 고유하게 확인가능한 발광 표지는 절단 사건이 발생한 때를 나타낼 것이다. 폴리펩티드가 반복적 절단을 겪을 때, 표지된 친화성 시약에 의한 결합에 상응하는 펄싱 속도는 말단 아미노산이 표지된 비-특이적 펩티다제에 의해 절단될 때마다 단계적 방식으로 떨어질 것이다. 따라서, 일부 실시양태에서, 절단 사건 사이에서 검출되는 패턴 내에서 발생하는 펄싱 패턴 및/또는 펄싱 속도를 기초로 하는 이러한 접근법에서 아미노산이 확인될 수 있고 - 이에 의해 폴리펩티드가 서열분석될 수 있다.

[0188] B. 표지된 폴리펩티드의 분해에 의한 서열분석

[0189] 일부 측면에서, 본 출원은 공지된 폴리펩티드 서열에 상응하는 아미노산의 고유한 조합을 확인함으로써 폴리펩티드를 서열분석하는 방법을 제공한다. 일부 실시양태에서, 방법은 표지된 폴리펩티드의 선택적으로 표지된 아미노산을 검출하는 것을 포함한다. 일부 실시양태에서, 표지된 폴리펩티드는 상이한 아미노산 유형이 상이한 발광 표지를 포함하도록 선택적으로 변형된 아미노산을 포함한다. 달리 나타내지 않는 한, 본원에 사용된 표지된 폴리펩티드는 1개 이상의 선택적으로 표지된 아미노산 측쇄를 포함하는 폴리펩티드를 지칭한다. 표지된 폴리펩티드의 제조 및 분석에 관한 선택적 표지화 방법 및 세부사항은 관련 기술분야에 공지되어 있다 (예를 들어, 문헌 [Swaminathan, et al., PLoS Comput Biol. 2015, 11(2):e1004080] 참조).

[0190] 본원에 기재된 바와 같이, 일부 측면에서, 본 출원은 폴리펩티드 분해 과정 동안 데이터를 획득하고, 데이터를 분석하여 분해 과정 동안 폴리펩티드의 말단에서 순차적으로 노출되는 아미노산에 상응하는 데이터의 부분을 결정함으로써 폴리펩티드를 서열분석하는 방법을 제공한다. 일부 실시양태에서, 데이터의 부분은 (예를 들어, 분해 동안) 폴리펩티드의 말단에 노출된 연속 아미노산과 1종 이상의 아미노산 인식 분자의 회합을 나타내는 일련의 신호 펄스를 포함한다. 일부 실시양태에서, 일련의 신호 펄스는 분해 과정 동안 폴리펩티드의 말단에서의 일련의 가역적 단일 분자 결합 상호작용에 상응한다.

[0191] 일부 측면에서, 본원에 기재된 폴리펩티드 서열분석 기술은 폴리펩티드가 절단 수단 (예를 들어, 1종 이상의 절단 시약)에 의해 분해되는 동안 폴리펩티드가 결합 수단 (예를 들어, 1종 이상의 아미노산 인식 분자)과 어떻게 상호작용하는지를 나타내는 데이터를 생성한다. 상기 논의된 바와 같이, 데이터는 말단에서의 절단 사건 사이에 폴리펩티드의 말단에서의 회합 사건에 상응하는 일련의 특징적인 패턴을 포함할 수 있다. 일부 실시양태에서, 본원에 기재된 서열분석 방법은 단일 폴리펩티드 분자를 결합 수단 및 절단 수단과 접촉시키는 것을 포함하며, 여기서 결합 수단 및 절단 수단은 절단 사건 전에 적어도 10개의 회합 사건을 달성하도록 구성된다. 일부 실시양태에서, 수단은 2개의 절단 사건 사이에 적어도 10개의 회합 사건을 달성하도록 구성된다.

[0192] 본원에 기재된 바와 같이, 일부 실시양태에서, 복수의 단일-분자 서열분석 반응은 샘플 웰의 어레이에서 병행 수행된다. 일부 실시양태에서, 어레이는 약 10,000 내지 약 1,000,000개의 샘플 웰을 포함한다. 일부 구현에서, 샘플 웰의 부피는 약 10^{-21} 리터 내지 약 10^{-15} 리터일 수 있다. 샘플 웰이 작은 부피를 갖기 때문에, 단지 약 1개의 폴리펩티드가 임의의 주어진 시간에 샘플 웰 내에 있을 수 있기 때문에, 단일-분자 사건의 검출이 가능할 수 있다. 통계적으로, 일부 샘플 웰은 단일-분자 서열분석 반응을 함유하지 않을 수 있고, 일부는 1개 초과 이상의 폴리펩티드 분자를 함유할 수 있다. 그러나, 인지가능한 수의 샘플 웰이 각각 단일-분자 반응을 함유할 수 있으므로 (예를 들어, 일부 실시양태에서 적어도 30%), 다수의 샘플 웰에 대해 단일-분자 분석이 병행 수행될 수 있다. 일부 실시양태에서, 결합 수단 및 절단 수단은 단일-분자 반응이 일어나고 있는 샘플 웰의 적어도 10% (예를 들어, 10-50%, 50% 초과, 25-75%, 적어도 80%, 또는 그 초과)에서 절단 사건 전에 적어도 10회 회합 사건을 달성하도록 구성된다. 일부 실시양태에서, 결합 수단 및 절단 수단은 단일-분자 반응에서 폴리펩티드의 아미노산의 적어도 50% (예를 들어, 50%, 50-75% 초과, 적어도 80%, 또는 그 초과)에 대한 절단 사건 전에 적어도 10회의 회합 사건을 달성하도록 구성된다.

[0193] 일부 실시양태에서, 표지된 폴리펩티드는 고정되고, 여기원에 노출된다. 표지된 폴리펩티드로부터의 합쳐진 발광이 검출될 수 있고, 일부 실시양태에서, 시간 경과에 따른 발광에 대한 노출은 발광 표지 분해 (예를 들어, 광표백으로 인한 분해)로 인해 검출된 신호의 상실을 유발할 수 있다. 일부 실시양태에서, 표지된 폴리펩티드는 초기 검출된 신호를 생성하는 선택적으로 표지된 아미노산의 고유한 조합을 포함한다. 시간 경과에 따른 발광 표지의 분해는 광표백된 표지된 폴리펩티드에 대한 검출된 신호의 상응하는 감소를 유발한다. 일부 실시양태에서, 신호는 1종 이상의 발광 특성의 분석에 의해 디컨볼루션될 수 있다 (예를 들어, 발광 수명 분석에 의한

신호 디컨볼루션). 일부 실시양태에서, 표지된 폴리펩티드의 선택적으로 표지된 아미노산의 고유한 조합은-예를 들어 프로테오믹의 공지된 폴리펩티드 서열에 기초하여 컴퓨터로 사전산출되고 실험적으로 검증되었다. 일부 실시양태에서, 검출된 아미노산 표지의 조합을 유기체의 프로테오믹의 공지된 서열의 데이터베이스에 대해 비교하여 표지된 폴리펩티드에 상응하는 데이터베이스의 특정한 폴리펩티드를 확인한다.

[0194] 일부 실시양태에서, 최적 샘플 농도는 대규모 병행 분석에서 샘플링을 최대화하는 서열분석 반응을 수행하기 위해 결정된다. 일부 실시양태에서, 농도는 어레이의 샘플 웰의 목적하는 분율 (예를 들어, 30%)이 임의의 주어진 시간에 점유되도록 선택된다. 이론에 얽매이는 것을 원하지는 않지만, 폴리펩티드가 소정의 기간에 걸쳐 표백되는 동안, 동일한 웰이 추가 분석에 계속 이용가능한 것으로 생각된다. 확산을 통해, 어레이의 샘플 웰의 대략 30%가 3분마다 분석에 사용될 수 있다. 예시적인 예로서, 백만개의 샘플 웰 칩에서, 시간당 6,000,000개의 폴리펩티드, 또는 4시간의 기간에 걸쳐 24,000,000개의 폴리펩티드가 샘플링될 수 있다.

[0195] 일부 측면에서, 본 출원은 말단 아미노산 변형 및 절단의 반복 사이클에 적용되는 표지된 폴리펩티드의 발광을 검출함으로써 폴리펩티드를 서열분석하는 방법을 제공한다. 일부 실시양태에서, 방법은 일반적으로 에드만 분해에 의한 다른 서열분석 방법에 대해 본원에 기재된 바와 같이 진행된다.

[0196] 일부 실시양태에서, 방법은 (i) 표지된 폴리펩티드의 말단 아미노산을 변형시키는 단계를 포함한다. 본원의 다른 곳에 기재된 바와 같이, 일부 실시양태에서, 변형은 말단 아미노산을 이소티오시아네이트 (예를 들어, PITC)와 접촉시켜 이소티오시아네이트-변형된 말단 아미노산을 형성하는 것을 포함한다. 일부 실시양태에서, 이소티오시아네이트 변형은 말단 아미노산을 절단 시약 (예를 들어, 본원에 기재된 바와 같은 화학적 또는 효소적 절단 시약)에 의한 제거에 보다 감수성인 형태로 전환시킨다. 따라서, 일부 실시양태에서, 방법은 (ii) 에드만 분해에 대해 본원의 다른 곳에 상술된 화학적 또는 효소적 수단을 사용하여 변형된 말단 아미노산을 제거하는 단계를 포함한다.

[0197] 일부 실시양태에서, 방법은 단계 (i) 내지 (ii)를 복수의 사이클 동안 반복하는 것을 포함하며, 그 동안 표지된 폴리펩티드의 발광이 검출되고, 말단으로부터 표지된 아미노산의 제거에 상응하는 절단 사건은 검출되는 신호의 감소로서 검출될 수 있다. 일부 실시양태에서, 단계 (ii) 후에 신호에서 변화가 없는 것으로 미지의 유형의 아미노산을 확인한다. 따라서, 일부 실시양태에서, 부분 서열 정보는 단계 (ii) 후에, 검출된 신호에서의 변화에 기초하여 결정된 정체에 의해 아미노산 유형을 배정하거나 또는 검출된 신호에서 변화가 없는 것에 기초하여 아미노산 유형을 미지의 것으로서 확인하는 것에 의하는 각각의 순차적 라운드 동안 검출된 신호를 평가함으로써 결정될 수 있다.

[0198] 일부 측면에서, 본 출원에 따른 폴리펩티드를 서열분석하는 방법은 표지된 폴리펩티드의 진행되는 효소적 절단에 의한 서열분석을 포함한다. 일부 실시양태에서, 표지된 폴리펩티드는 말단 아미노산을 하나의 말단에서 또 다른 말단으로 연속적으로 절단하는 변형된 진행성 엑소펩티다제를 사용하는 분해에 적용된다. 엑소펩티다제는 본원의 다른 곳에 상세히 기재되어 있다. 일부 실시양태에서, 표지된 폴리펩티드는 고정된 진행성 엑소펩티다제에 의한 분해에 적용된다. 일부 실시양태에서, 고정된 표지된 폴리펩티드는 진행성 엑소펩티다제에 의한 분해에 적용된다.

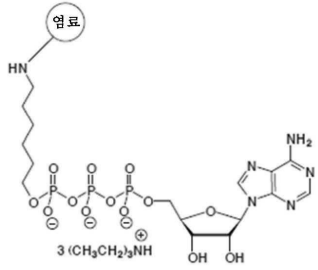
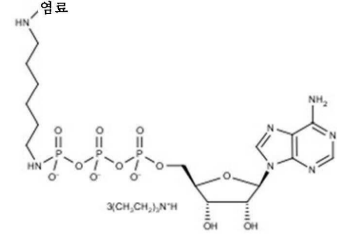
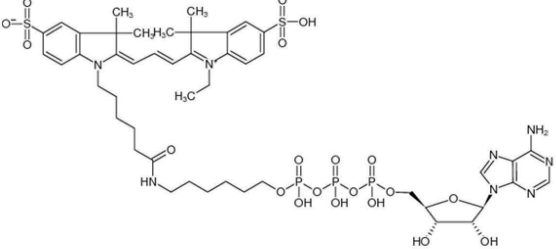
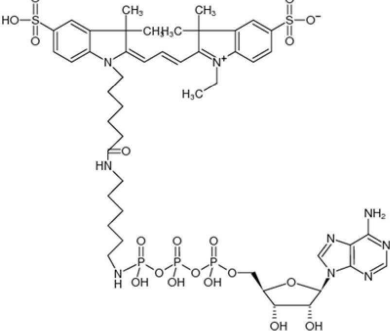
[0199] 일부 실시양태에서, 진행성 엑소펩티다제의 진행 속도는 공지되어 있으므로, 검출된 신호 감소 사이의 타이밍은 각각의 검출 사건 사이의 비표지된 아미노산의 수를 계산하는데 사용될 수 있다. 예를 들어, 40개 아미노산의 폴리펩티드가 아미노산이 매초 제거되는 방식으로 절단되는 경우, 3개의 신호를 갖는 표지된 폴리펩티드는 처음에는 모든 3개, 이어서 2개, 이어서 1개, 및 마지막으로 신호 없음을 나타낼 것이다. 이러한 방식으로, 표지된 아미노산의 순서가 결정될 수 있다. 따라서, 이들 방법은 예를 들어 폴리펩티드 단편 서열분석에 기초한 프로테오믹 분석을 위해 부분 서열 정보를 결정하는데 사용될 수 있다.

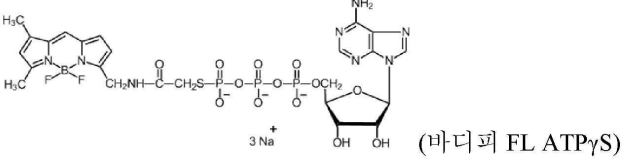
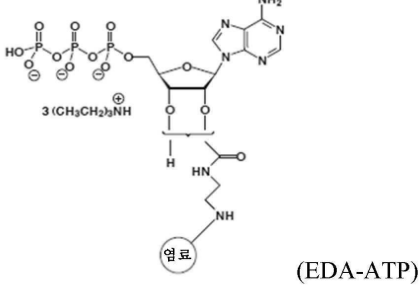
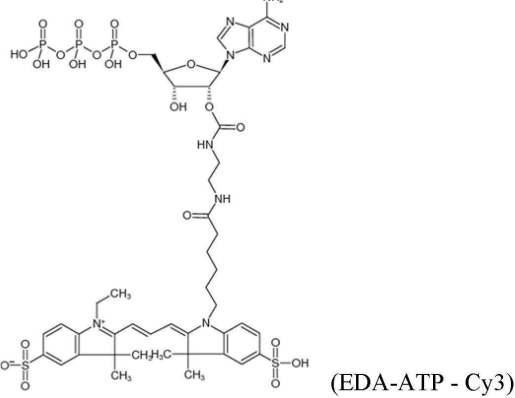
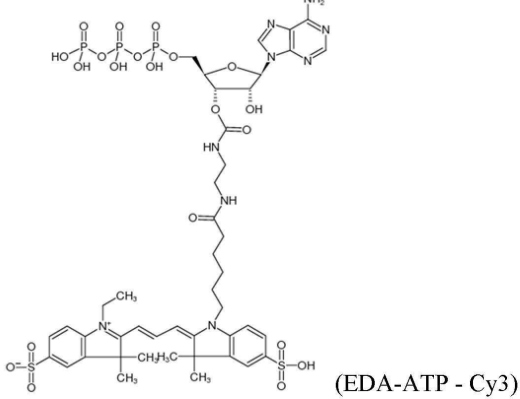
[0200] 일부 실시양태에서, 단일 분자 폴리펩티드 서열분석은 ATP-기반 포스터 공명 에너지 전달 (FRET) 스킴 (예를 들어, 1종 이상의 표지된 보조인자를 사용함)을 사용하여 달성될 수 있다. 일부 실시양태에서, 보조인자-기반 FRET에 의한 서열분석은 고정된 ATP-의존성 프로테아제, 공여자-표지된 ATP, 및 폴리펩티드 기질의 수용자-표지된 아미노산을 사용하여 수행될 수 있다. 일부 실시양태에서, 아미노산은 수용자로 표지될 수 있고, 1종 이상의 보조인자는 공여자로 표지될 수 있다.

[0201] 예를 들어, 일부 실시양태에서, 추출된 폴리펩티드는 변성되고, 시스테인 및 리신은 형광 염료로 표지된다. 일부 실시양태에서, 단백질 트랜스로카제의 조작된 버전 (예를 들어, 박테리아 ClpX)은 개별 기질 폴리펩티드에 결합하고, 이를 언폴딩시키고, 이를 그의 나노-채널을 통해 전위시키는데 사용된다. 일부 실시양태에서, 트랜

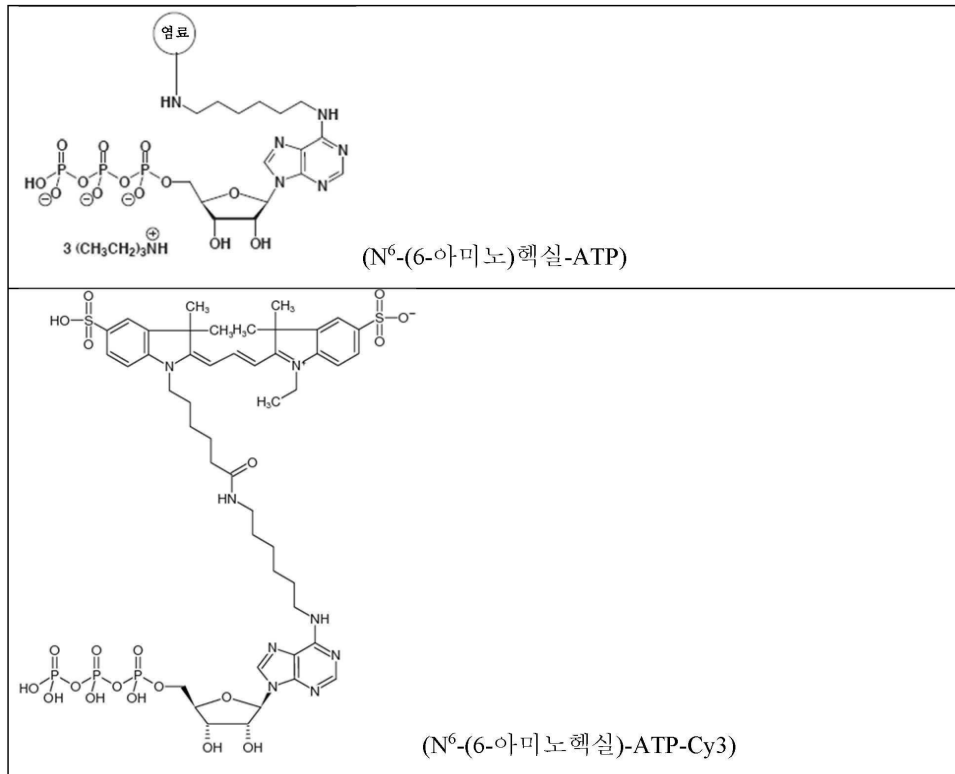
스로카제는 공여자 염료로 표지되고, FRET는 기질이 나노-채널을 통과할 때 트랜스로카제 상의 공여자와 기질 상의 2종 이상의 별개의 수용자 염료 사이에 발생한다. 이어서, 표지된 아미노산의 순서는 FRET 신호로부터 결정될 수 있다. 일부 실시양태에서, 표 3에 제시된 하기 비제한적 표지된 ATP 유사체 중 1종 이상이 사용될 수 있다.

표 3. 표지된 ATP 유사체의 비제한적 예

포스페이트-표지된 ATP:  (γ-[(6-아미노)헥실]-ATP)
 (γ-[(6-아미노)헥실]이미도]-ATP)
 (γ-(6-아미노)헥실)-ATP-Cy3)
 (γ-[(6-아미노)헥실]이미도]-ATP-Cy3)

 (바디 피 FL ATPγS)
리보스-표지된 ATP:
 (EDA-ATP)
 (EDA-ATP - Cy3)
 (EDA-ATP - Cy3)
염기-표지된 ATP:

[0204]



[0205]

[0206]

C. 서열분석을 위한 샘플의 제조

[0207]

폴리펩티드 샘플 (예를 들어, 풍부화된 폴리펩티드 샘플)은 서열분석 전에 변형될 수 있다.

[0208]

일부 실시양태에서, 폴리펩티드의 N-말단 아미노산 또는 C-말단 아미노산이 변형된다. 일부 실시양태에서, 폴리펩티드의 말단 단부는 표면 (예를 들어, 폴리펩티드 분석에 사용되는 칩 상의 샘플 웰의 표면)에 대한 고정을 가능하게 하는 모이어티로 변형된다. 일부 실시양태에서, 이러한 방법은 본 출원에 따라 분석될 표지된 폴리펩티드의 말단 단부를 변형시키는 것을 포함한다. 또 다른 실시양태에서, 이러한 방법은 본 출원에 따라 폴리펩티드 기질을 분해하거나 전위시키는 단백질 또는 효소의 말단 단부를 변형시키는 것을 포함한다.

[0209]

일부 실시양태에서, 폴리펩티드의 카르복시-말단은 (i) 폴리펩티드의 유리 카르복실레이트 기를 차단하고; (ii) 폴리펩티드를 (예를 들어, 열 및/또는 화학적 수단에 의해) 변성시키고; (iii) 폴리펩티드의 유리 티올 기를 차단하고; (iv) 폴리펩티드를 소화시켜 유리 C-말단 카르복실레이트 기를 포함하는 적어도 1개의 폴리펩티드 단편을 생산하고; (v) 관능성 모이어티를 유리 C-말단 카르복실레이트 기에 (예를 들어, 화학적으로) 접합시키는 것을 포함하는 방법으로 변형된다. 일부 실시양태에서, 방법은 (i) 후 및 (ii) 전에, 폴리펩티드를 포함하는 샘플을 투석하는 것을 추가로 포함한다.

[0210]

일부 실시양태에서, 폴리펩티드의 카르복시-말단은 (i) 폴리펩티드를 변성시키고 (예를 들어, 열 및/또는 화학적 수단에 의해); (ii) 폴리펩티드의 유리 티올 기를 차단하고; (iii) 폴리펩티드를 소화시켜 유리 C-말단 카르복실레이트 기를 포함하는 적어도 1개의 폴리펩티드 단편을 생산하고; (iv) 유리 C-말단 카르복실레이트 기를 차단시켜 차단된 C-말단 카르복실레이트 기를 포함하는 적어도 1개의 폴리펩티드 단편을 생산하고; (v) 관능성 모이어티를 차단된 C-말단 카르복실레이트 기에 (예를 들어, 효소적으로) 접합시키는 것을 포함하는 방법으로 변형된다. 일부 실시양태에서, 방법은 (iv) 후 및 (v) 전에, 폴리펩티드를 포함하는 샘플을 투석하는 것을 추가로 포함한다.

[0211]

일부 실시양태에서, 유리 카르복실레이트 기를 차단하는 것은 비변형된 카르복실레이트에 비해 화학적 반응성을 변경시키는 이들 기의 화학적 변형을 지칭한다. 적합한 카르복실레이트 차단 방법은 관련 기술분야에 공지되어 있고, 관능화된 폴리펩티드의 카르복시-말단 카르복실레이트 기와 화학적으로 상이하도록 측쇄 카르복실레이트 기를 변형시켜야 한다. 일부 실시양태에서, 유리 카르복실레이트 기를 차단하는 것은 폴리펩티드의 유리 카르복실레이트 기의 에스테르화 또는 아마이드화를 포함한다. 일부 실시양태에서, 유리 카르복실레이트 기를 차단하는 것은 예를 들어 폴리펩티드를 메탄올성 HCl과 반응시키는 것에 의한 폴리펩티드의 유리 카르복실레이트 기의 메틸 에스테르화를 포함한다. 유리 카르복실레이트 기를 차단하는데 유용한 시약 및 기술의 추가의 예는, 제한

없이, 4-술포-2,3,5,6-테트라플루오로페놀 (STP) 및/또는 카르보다이미드 예컨대 N-(3-디메틸아미노프로필)-N'-에틸카르보다이미드 히드로클로라이드 (EDAC), 우로늄 시약, 디아조메탄, 피셔 에스테르화를 위한 알콜 및 산, (잠재적으로 후속 에스테르 또는 아민 형성을 위한 중간체로서) N-히드록실숙신이미드 (NHS) 에스테르를 형성하기 위한 NHS의 사용, 또는 카르보닐디이미다졸 (CDI)과의 반응 또는 혼합 무수물의 형성, 또는 잠재적으로 에스테르 또는 아미드의 형성을 통하는 카르복실산을 변형시키거나 차단하는 임의의 다른 방법을 포함한다.

[0212] 일부 실시양태에서, 유리 티올 기를 차단하는 것은 비변형된 티올에 비해 화학적 반응성을 변경시키는 이들 기의 화학적 변형을 지칭한다. 일부 실시양태에서, 유리 티올 기를 차단하는 것은 폴리펩티드의 유리 티올 기의 환원 및 알킬화를 포함한다. 일부 실시양태에서, 환원 및 알킬화는 폴리펩티드를 디티오트레이톨 (DTT), 및 아이오도아세트아미드 및 아이오도아세트산 중 하나 또는 둘 다와 접촉시키는 것에 의해 수행된다. 사용될 수 있는 추가의 및 대안적 시스테인-환원 시약의 예는 널리 공지되어 있고, 제한 없이, 2-메르캅토에탄올, 트리스 (2-카르복시에틸) 포스핀 히드로클로라이드 (TCEP), 트리부틸포스핀, 디티오부틸아민 (DTBA), 또는 티올 기를 환원시킬 수 있는 임의의 시약을 포함한다. 사용될 수 있는 추가의 및 대안적 시스테인-차단 (예를 들어, 시스테인-알킬화) 시약의 예는 널리 공지되어 있고, 제한 없이, 아크릴아미드, 4-비닐피리딘, N-에틸말레이미드 (NEM), N-ε-말레이미도카프로산 (EMCA), 또는 디술포드 결합 형성을 막도록 시스테인을 변형시키는 임의의 시약을 포함한다.

[0213] 일부 실시양태에서, 소화는 효소적 소화를 포함한다. 일부 실시양태에서, 소화는 폴리펩티드를 엔도펩티다제 (예를 들어, 트립신)와 소화 조건 하에 접촉시킴으로써 수행된다. 일부 실시양태에서, 소화는 화학적 소화를 포함한다. 화학적 및 효소적 소화에 적합한 시약의 예는 관련 기술분야에 공지되어 있고, 제한 없이, 트립신, 케모트립신, Lys-C, Arg-C, Asp-N, Lys-N, BNPS-스카톨, CNBr, 카스파제, 포름산, 글루타밀 엔도펩티다제, 히드록실아민, 아이오도소벤조산, 호중구 엘라스타제, 펩신, 프로테인-엔도펩티다제, 프로테이나제 K, 스타필로코쿠스 펩티다제 I, 써모리신, 및 트롬빈을 포함한다.

[0214] 일부 실시양태에서, 기능적 모이어티는 비오틴 분자를 포함한다. 일부 실시양태에서, 관능성 모이어티는 반응성 화학적 모이어티, 예컨대 알키닐을 포함한다. 일부 실시양태에서, 기능적 모이어티의 접합은 관련 기술분야에 공지된 바와 같은, 카르복시펩티다제 Y에 의한 카르복시-말단 카르복시-메틸 에스테르 기의 비오틴화합을 포함한다.

[0215] 일부 실시양태에서, 가용화 모이어티가 폴리펩티드에 부가된다. 따라서, 일부 실시양태에서, 본원에 제공된 방법 및 조성물은 폴리펩티드의 말단 단부를 그의 용해도를 증가시키는 모이어티로 변형시키는데 유용하다. 일부 실시양태에서, 가용화 모이어티는 단편화 (예를 들어, 트립신을 사용한 효소적 단편화)로부터 생성되고 비교적 불용성인 소형 폴리펩티드에 유용하다. 예를 들어, 일부 실시양태에서, 폴리펩티드 풀 내의 짧은 폴리펩티드는 중합체 (예를 들어, 짧은 올리고, 당, 또는 다른 하전된 중합체)를 폴리펩티드에 접합시키는 것에 의해 가용화될 수 있다.

[0216] D. 발광 표지

[0217] 본원에 사용된 발광 표지는 1개 이상의 광자를 흡수하고, 후속적으로 1회 이상의 지속기간 후에 1개 이상의 광자를 방출할 수 있는 분자이다. 일부 실시양태에서, 용어는 문맥에 따라 "표지" 또는 "발광 분자"와 상호교환 가능하게 사용된다. 본원에 기재된 특정 실시양태에 따른 발광 표지는 표지된 친화성 시약의 발광 표지, 표지된 펩티다제 (예를 들어, 표지된 엑소펩티다제, 표지된 비-특이적 엑소펩티다제)의 발광 표지, 표지된 펩티드의 발광 표지, 표지된 보조인자의 발광 표지, 또는 본원에 기재된 또 다른 표지된 조성물을 지칭할 수 있다. 일부 실시양태에서, 본 출원에 따른 발광 표지는 1개 이상의 표지된 아미노산을 포함하는 표지된 폴리펩티드의 표지된 아미노산을 지칭한다.

[0218] 일부 실시양태에서, 발광 표지는 제1 및 제2 발색단을 포함할 수 있다. 일부 실시양태에서, 제1 발색단의 여기 상태는 제2 발색단으로의 에너지 전달을 통해 이완될 수 있다. 일부 실시양태에서, 에너지 전달은 포스터 공명 에너지 전달 (FRET)이다. 이러한 FRET 쌍은 발광 표지에 혼합물 중의 복수의 발광 표지와 보다 용이하게 구분되도록 하는 특성을 제공하는데 유용할 수 있다. 또 다른 실시양태에서, FRET 쌍은 제1 발광 표지의 제1 발색단 및 제2 발광 표지의 제2 발색단을 포함한다. 특정 실시양태에서, FRET 쌍은 제1 스펙트럼 범위에서 여기 에너지를 흡수하고 제2 스펙트럼 범위에서 발광을 방출할 수 있다.

[0219] 일부 실시양태에서, 발광 표지는 형광단 또는 염료를 지칭한다. 전형적으로, 발광 표지는 방향족 또는 헤테로 방향족 화합물을 포함하고, 피렌, 안트라센, 나프탈렌, 나프틸아민, 아크리딘, 스틸벤, 인돌, 벤즈인돌,

옥사졸, 카르바졸, 티아졸, 벤조티아졸, 벤족사졸, 페난트리딘, 페녹사진, 포르피린, 퀴놀린, 에티듐, 벤즈아미드, 시아닌, 카르보시아닌, 살리실레이트, 안트라닐레이트, 쿠마린, 플루오레세인, 로다민, 크산텐, 또는 다른 유사 화합물일 수 있다.

[0220]

일부 실시양태에서, 발광 표지는 하기 중 1종으로부터 선택된 염료를 포함한다: 5/6-카르복시로다민 6G, 5-카르복시로다민 6G, 6-카르복시로다민 6G, 6-TAMRA, 아베리오르(Abberior)® 스타 440SXP, 아베리오르® 스타 470SXP, 아베리오르® 스타 488, 아베리오르® 스타 512, 아베리오르® 스타 520SXP, 아베리오르® 스타 580, 아베리오르® 스타 600, 아베리오르® 스타 635, 아베리오르® 스타 635P, 아베리오르® 스타 레드, 알렉사 플루오르(Alexa Fluor)® 350, 알렉사 플루오르® 405, 알렉사 플루오르® 430, 알렉사 플루오르® 480, 알렉사 플루오르® 488, 알렉사 플루오르® 514, 알렉사 플루오르® 532, 알렉사 플루오르® 546, 알렉사 플루오르® 555, 알렉사 플루오르® 568, 알렉사 플루오르® 594, 알렉사 플루오르® 610-X, 알렉사 플루오르® 633, 알렉사 플루오르® 647, 알렉사 플루오르® 660, 알렉사 플루오르® 680, 알렉사 플루오르® 700, 알렉사 플루오르® 750, 알렉사 플루오르® 790, AMCA, 아토(ATTO) 390, 아토 425, 아토 465, 아토 488, 아토 495, 아토 514, 아토 520, 아토 532, 아토 542, 아토 550, 아토 565, 아토 590, 아토 610, 아토 620, 아토 633, 아토 647, 아토 647N, 아토 655, 아토 665, 아토 680, 아토 700, 아토 725, 아토 740, 아토 옥사(ATTO 0xa)12, 아토 로(ATTO Rho)101, 아토 로11, 아토 로12, 아토 로13, 아토 로14, 아토 로3B, 아토 로6G, 아토 티오(ATTO Thio)12, 비디 호리즌(BD Horizon)TM V450, 바디피(BODIPY)® 493/501, 바디피® 530/550, 바디피® 558/568, 바디피® 564/570, 바디피® 576/589, 바디피® 581/591, 바디피® 630/650, 바디피® 650/665, 바디피® FL, 바디피® FL-X, 바디피® R6G, 바디피® TMR, 바디피® TR, 칼 플루오르(CAL Fluor)® 골드 540, 칼 플루오르® 그린 510, 칼 플루오르® 오렌지 560, 칼 플루오르® 레드 590, 칼 플루오르® 레드 610, 칼 플루오르® 레드 615, 칼 플루오르® 레드 635, 캐스케이드(Cascade)® 블루, CFTM350, CFTM405M, CFTM405S, CFTM488A, CFTM514, CFTM532, CFTM543, CFTM546, CFTM555, CFTM568, CFTM594, CFTM620R, CFTM633, CFTM633-V1, CFTM640R, CFTM640R-V1, CFTM640R-V2, CFTM660C, CFTM660R, CFTM680, CFTM680R, CFTM680R-V1, CFTM750, CFTM770, CFTM790, 크로메오(Chromo)™ 642, 크로미스(Chromis) 425N, 크로미스 500N, 크로미스 515N, 크로미스 530N, 크로미스 550A, 크로미스 550C, 크로미스 550Z, 크로미스 560N, 크로미스 570N, 크로미스 577N, 크로미스 600N, 크로미스 630N, 크로미스 645A, 크로미스 645C, 크로미스 645Z, 크로미스 678A, 크로미스 678C, 크로미스 678Z, 크로미스 770A, 크로미스 770C, 크로미스 800A, 크로미스 800C, 크로미스 830A, 크로미스 830C, Cy®3, Cy®3.5, Cy®3B, Cy®5, Cy®5.5, Cy®7, 다이라이트(DyLight)® 350, 다이라이트® 405, 다이라이트® 415-Co1, 다이라이트® 425Q, 다이라이트® 485-LS, 다이라이트® 488, 다이라이트® 504Q, 다이라이트® 510-LS, 다이라이트® 515-LS, 다이라이트® 521-LS, 다이라이트® 530-R2, 다이라이트® 543Q, 다이라이트® 550, 다이라이트® 554-R0, 다이라이트® 554-R1, 다이라이트® 590-R2, 다이라이트® 594, 다이라이트® 610-B1, 다이라이트® 615-B2, 다이라이트® 633, 다이라이트® 633-B1, 다이라이트® 633-B2, 다이라이트® 650, 다이라이트® 655-B1, 다이라이트® 655-B2, 다이라이트® 655-B3, 다이라이트® 655-B4, 다이라이트® 662Q, 다이라이트® 675-B1, 다이라이트® 675-B2, 다이라이트® 675-B3, 다이라이트® 675-B4, 다이라이트® 679-C5, 다이라이트® 680, 다이라이트® 683Q, 다이라이트® 690-B1, 다이라이트® 690-B2, 다이라이트® 696Q, 다이라이트® 700-B1, 다이라이트® 700-B1, 다이라이트® 730-B1, 다이라이트® 730-B2, 다이라이트® 730-B3, 다이라이트® 730-B4, 다이라이트® 747, 다이라이트® 747-B1, 다이라이트® 747-B2, 다이라이트® 747-B3, 다이라이트® 747-B4, 다이라이트® 755, 다이라이트® 766Q, 다이라이트® 775-B2, 다이라이트® 775-B3, 다이라이트® 775-B4, 다이라이트® 780-B1, 다이라이트® 780-B2, 다이라이트® 780-B3, 다이라이트® 800, 다이라이트® 830-B2, 디오믹스(Dyomics)-350, 디오믹스-350XL, 디오믹스-360XL, 디오믹스-370XL, 디오믹스-375XL, 디오믹스-380XL, 디오믹스-390XL, 디오믹스-405, 디오믹스-415, 디오믹스-430, 디오믹스-431, 디오믹스-478, 디오믹스-480XL, 디오믹스-481XL, 디오믹스-485XL, 디오믹스-490, 디오믹스-495, 디오믹스-505, 디오믹스-510XL, 디오믹스-511XL, 디오믹스-520XL, 디오믹스-521XL, 디오믹스-530, 디오믹스-547, 디오믹스-547P1, 디오믹스-548, 디오믹스-549, 디오믹스-549P1, 디오믹스-550, 디오믹스-554, 디오믹스-555, 디오믹스-556, 디오믹스-560, 디오믹스-590, 디오믹스-591, 디오믹스-594, 디오믹스-601XL, 디오믹스-605, 디오믹스-610, 디오믹스-615, 디오믹스-630, 디오믹스-631, 디오믹스-632, 디오믹스-633, 디오믹스-634, 디오믹스-635, 디오믹스-636, 디오믹스-647, 디오믹스-647P1, 디오믹스-648, 디오믹스-648P1, 디오믹스-649, 디오믹스-649P1, 디오믹스-650, 디오믹스-651, 디오믹스-652, 디오믹스-654, 디오믹스-675, 디오믹스-676, 디오믹스-677, 디오믹스-678, 디오믹스-679P1, 디오믹스-680, 디오믹스-681, 디오믹스-682, 디오믹스-700, 디오믹스-701, 디오믹스-703, 디오믹스-704, 디오믹스-730, 디오믹스-731, 디오믹스-732, 디오믹스-734, 디오믹스-749, 디오믹스-749P1, 디오믹스-750, 디오믹스-751, 디오믹스-752, 디오믹스-754, 디오믹스-776, 디오믹스-777, 디오믹스-778, 디오믹스-780, 디오믹스-781,

디오믹스-782, 디오믹스-800, 디오믹스-831, 이플루오르(eFluor)® 450, 에오신, FITC, 플루오레세인, 하이라이트(HiLyte)™ 플루오르 405, 하이라이트™ 플루오르 488, 하이라이트™ 플루오르 532, 하이라이트™ 플루오르 555, 하이라이트™ 플루오르 594, 하이라이트™ 플루오르 647, 하이라이트™ 플루오르 680, 하이라이트™ 플루오르 750, IRDye® 680LT, IRDye® 750, IRDye® 800CW, 조(JOE), 라이트사이클러(LightCycler)® 640R, 라이트사이클러® 레드 610, 라이트사이클러® 레드 640, 라이트사이클러® 레드 670, 라이트사이클러® 레드 705, 리사민 로다민 B(Lissamine Rhodamine B), 나프토플루오레세인, 오레곤 그린(Oregon Green)® 488, 오레곤 그린® 514, 퍼시픽 블루(Pacific Blue)™, 퍼시픽 그린(Pacific Green)™, 퍼시픽 오렌지(Pacific Orange)™, PET, PF350, PF405, PF415, PF488, PF505, PF532, PF546, PF555P, PF568, PF594, PF610, PF633P, PF647P, 퀴사르(Quasar)® 570, 퀴사르® 670, 퀴사르® 705, 로다민(Rhodamine) 123, 로다민 6G, 로다민 B, 로다민 그린, 로다민 그린-X, 로다민 레드, ROX, 세타(Seta)™ 375, 세타™ 470, 세타™ 555, 세타™ 632, 세타™ 633, 세타™ 650, 세타™ 660, 세타™ 670, 세타™ 680, 세타™ 700, 세타™ 750, 세타™ 780, 세타™ APC-780, 세타™ PerCP-680, 세타™ R-PE-670, 세타™ 646, 세타우(SeTau) 380, 세타우 425, 세타우 647, 세타우 405, 스퀘어(Square) 635, 스퀘어 650, 스퀘어 660, 스퀘어 672, 스퀘어 680, 숄포로다민 101, TAMRA, TET, 텍사스 레드(Texas Red)®, TMR, TRITC, 야키마 옐로우(Yakima Yellow)™, 제논(Zenon)®, Zy3, Zy5, Zy5.5, 및 Zy7.

[0221] E. 발광

[0222] 일부 측면에서, 본 출원은 발광 표지의 1종 이상의 발광 특성에 기초한 폴리펩티드 서열분석 및/또는 확인에 관한 것이다. 일부 실시양태에서, 발광 표지는 발광 수명, 발광 강도, 휘도, 흡수 스펙트럼, 방출 스펙트럼, 발광 양자 수율, 또는 그의 2종 이상의 조합에 기초하여 확인된다. 일부 실시양태에서, 복수의 유형의 발광 표지는 상이한 발광 수명, 발광 강도, 휘도, 흡수 스펙트럼, 방출 스펙트럼, 발광 양자 수율, 또는 그의 2종 이상의 조합에 기초하여 서로 구별될 수 있다. 확인은 발광 표지와 회합된 1종의 유형의 아미노산 (예를 들어, 단일 유형 또는 유형의 하위세트)의 정확한 정제 및/또는 양을 배정하는 것을 의미할 수 있고, 또한 다른 유형의 아미노산 대비 폴리펩티드에서의 아미노산 위치를 배정하는 것을 의미할 수 있다.

[0223] 일부 실시양태에서, 발광은 발광 표지를 일련의 개별 광 펄스에 노출시키고 표지로부터 방출된 각각의 광자의 타이밍 또는 다른 특성을 평가함으로써 검출된다. 일부 실시양태에서, 표지로부터 순차적으로 방출된 복수의 광자에 대한 정보는 합쳐져 있고, 이를 평가하여 표지를 확인하고, 그에 의해 회합된 아미노산 유형을 확인한다. 일부 실시양태에서, 표지의 발광 수명은 표지로부터 순차적으로 방출되는 복수의 광자로부터 결정되고, 발광 수명을 사용하여 표지를 확인할 수 있다. 일부 실시양태에서, 표지의 발광 강도는 표지로부터 순차적으로 방출되는 복수의 광자로부터 결정되고, 발광 강도를 사용하여 표지를 확인할 수 있다. 일부 실시양태에서, 표지의 발광 수명 및 강도는 표지로부터 순차적으로 방출되는 복수의 광자로부터 결정되고, 발광 수명 및 강도를 사용하여 표지를 확인할 수 있다.

[0224] 본 출원의 일부 측면에서, 단일 폴리펩티드 분자는 복수의 개별 광 펄스에 노출되고, 일련의 방출된 광자가 검출 및 분석된다. 일부 실시양태에서, 일련의 방출된 광자는 실험 시간에 걸쳐 반응 샘플에 존재하고 변화되지 않은 단일 폴리펩티드 분자에 대한 정보를 제공한다. 그러나, 일부 실시양태에서, 일련의 방출된 광자는 (예를 들어, 반응 또는 과정이 진행됨에 따라) 반응 샘플에서 상이한 시간에 존재하는 일련의 상이한 분자에 대한 정보를 제공한다. 예로서 및 제한 없이, 이러한 정보는 본 출원에 따라 화학적 또는 효소적 분해에 적용된 폴리펩티드를 서열분석 및/또는 확인하는데 사용될 수 있다.

[0225] 특정 실시양태에서, 발광 표지는 1개의 광자를 흡수하고 시간 지속기간 후에 1개의 광자를 방출한다. 일부 실시양태에서, 표지의 발광 수명은 시간 지속기간을 측정함으로써 결정 또는 추정될 수 있다. 일부 실시양태에서, 표지의 발광 수명은 다중 펄스 사건 및 방출 사건에 대한 복수의 시간 지속기간을 측정함으로써 결정 또는 추정될 수 있다. 일부 실시양태에서, 표지의 발광 수명은 시간 지속기간을 측정하는 것에 의해 복수의 유형의 표지의 발광 수명과 구분될 수 있다. 일부 실시양태에서, 표지의 발광 수명은 다중 펄스 사건 및 방출 사건에 대한 복수의 시간 지속기간을 측정하는 것에 의해 복수의 유형의 표지의 발광 수명과 구분될 수 있다. 특정 실시양태에서, 표지는 표지의 발광 수명을 결정 또는 추정하는 것에 의해 복수의 유형의 표지 중에서 확인 또는 구분된다. 특정 실시양태에서, 표지는 표지의 발광 수명을 복수의 유형의 표지의 복수의 발광 수명과 구분하는 것에 의해 복수의 유형의 표지 중에서 확인 또는 구분된다.

[0226] 발광 표지의 발광 수명의 결정은 임의의 적합한 방법을 사용하여 (예를 들어, 적합한 기술을 사용하여 수명을 측정함으로써 또는 방출의 시간-의존성 특징을 결정함으로써) 수행될 수 있다. 일부 실시양태에서, 하나의 표

지의 발광 수명을 결정하는 것은 또 다른 표지에 대비하여 수명을 결정하는 것을 포함한다. 일부 실시양태에서, 표지의 발광 수명을 결정하는 것은 참조물에 대비하여 수명을 결정하는 것을 포함한다. 일부 실시양태에서, 표지의 발광 수명을 결정하는 것은 수명 (예를 들어, 형광 수명)을 측정하는 것을 포함한다. 일부 실시양태에서, 표지의 발광 수명을 결정하는 것은 수명을 나타내는 1종 이상의 시간적 특징을 결정하는 것을 포함한다. 일부 실시양태에서, 표지의 발광 수명은 여기 펄스에 비해 하나 이상의 시간-게이팅 윈도우에 걸쳐 발생한 복수의 방출 사건 (예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100개, 또는 그 초과)의 분포에 기초하여 결정될 수 있다. 예를 들어, 표지의 발광 수명은 여기 펄스에 대해 측정된 광자 도달 시간의 분포에 기초하여 상이한 발광 수명을 갖는 복수의 표지와 구별될 수 있다.

[0227] 발광 표지의 발광 수명은 표지가 여기 상태에 도달한 후에 방출된 광자의 타이밍을 나타내고, 표지는 광자의 타이밍을 나타내는 정보에 의해 구별될 수 있음이 인지되어야 한다. 일부 실시양태는 표지에 의해 방출된 광자와 연관된 시간을 측정함으로써 표지의 발광 수명에 기초하여 복수의 표지로부터 표지를 구별하는 것을 포함할 수 있다. 시간의 분포는 분포로부터 결정될 수 있는 발광 수명의 표지를 제공할 수 있다. 일부 실시양태에서, 표지는 시간의 분포에 기초하여, 예컨대 시간의 분포를 공지된 표지에 상응하는 참조 분포와 비교함으로써 복수의 표지와 구별가능하다. 일부 실시양태에서, 발광 수명에 대한 값은 시간의 분포로부터 결정된다.

[0228] 본원에 사용된 바와 같이, 일부 실시양태에서, 발광 강도는 펄스 여기 에너지의 전달에 의해 여기되는 발광 표지에 의해 방출되는 단위 시간당 방출된 광자의 수를 지칭한다. 일부 실시양태에서, 발광 강도는 펄스 여기 에너지의 전달에 의해 여기되는 표지에 의해 방출되고 특정한 센서 또는 센서 세트에 의해 검출되는, 단위 시간당 방출된 광자의 검출된 수를 지칭한다.

[0229] 본원에 사용된 바와 같이, 일부 실시양태에서, 휘도는 발광 표지당 평균 방출 강도를 보고하는 파라미터를 지칭한다. 따라서, 일부 실시양태에서, "방출 강도"는 일반적으로 1종 이상의 표지를 포함하는 조성물의 휘도를 지칭하는데 사용될 수 있다. 일부 실시양태에서, 표지의 휘도는 그의 양자 수율 및 흡광 계수의 곱과 동일하다.

[0230] 본원에 사용된 바와 같이, 일부 실시양태에서, 발광 양자 수율은 방출 사건으로 이어지는 주어진 파장에서의 또는 주어진 스펙트럼 범위 내에서의 여기 사건의 분율을 지칭하며, 전형적으로 1 미만이다. 일부 실시양태에서, 본원에 기재된 발광 표지의 발광 양자 수율은 0 내지 약 0.001, 약 0.001 내지 약 0.01, 약 0.01 내지 약 0.1, 약 0.1 내지 약 0.5, 약 0.5 내지 0.9, 또는 약 0.9 내지 1이다. 일부 실시양태에서, 표지는 발광 양자 수율을 결정 또는 추정함으로써 확인된다.

[0231] 본원에 사용된 바와 같이, 일부 실시양태에서, 여기 에너지는 광원으로부터의 광의 펄스이다. 일부 실시양태에서, 여기 에너지는 가시 스펙트럼 내에 있다. 일부 실시양태에서, 여기 에너지는 자외선 스펙트럼 내에 있다. 일부 실시양태에서, 여기 에너지는 적외선 스펙트럼 내에 있다. 일부 실시양태에서, 여기 에너지는 복수의 방출된 광자가 검출될 발광 표지의 흡수 최대점 또는 그 근처에 있다. 특정 실시양태에서, 여기 에너지는 약 500 nm 내지 약 700 nm (예를 들어, 약 500 nm 내지 약 600 nm, 약 600 nm 내지 약 700 nm, 약 500 nm 내지 약 550 nm, 약 550 nm 내지 약 600 nm, 약 600 nm 내지 약 650 nm, 또는 약 650 nm 내지 약 700 nm)이다. 특정 실시양태에서, 여기 에너지는 단색일 수 있거나 또는 스펙트럼 범위로 한정될 수 있다. 일부 실시양태에서, 스펙트럼 범위는 약 0.1 nm 내지 약 1 nm, 약 1 nm 내지 약 2 nm, 또는 약 2 nm 내지 약 5 nm의 범위를 갖는다. 일부 실시양태에서, 스펙트럼 범위는 약 5 nm 내지 약 10 nm, 약 10 nm 내지 약 50 nm, 또는 약 50 nm 내지 약 100 nm의 범위를 갖는다.

[0232] IV. 샘플 제조를 위한 키트

[0233] 일부 측면에서, 본 개시내용은 서열분석을 위한 폴리펩티드 샘플 (예를 들어, 풍부화된 샘플)을 제조하기 위한 키트에 관한 것이다. 키트는 서열분석을 위한 1종 이상의 폴리펩티드 샘플 (예를 들어, 풍부화된 샘플)을 제조하는데 충분할 수 있다. 일부 실시양태에서, 키트는 단일 폴리펩티드 샘플을 제조하는데 충분하다. 다른 실시양태에서, 키트는 적어도 2, 적어도 3, 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 적어도 10, 적어도 11, 적어도 12, 적어도 13, 적어도 14, 적어도 15, 적어도 20, 적어도 25, 적어도 30, 적어도 40, 적어도 50, 적어도 60, 적어도 70, 적어도 80, 적어도 90, 또는 적어도 100종의 폴리펩티드 샘플을 제조하는데 충분하다.

[0234] 일부 실시양태에서, 키트는 본원에 기재된 바와 같은 복수의 풍부화 분자를 포함하는 풍부화 성분을 포함한다. "폴리펩티드 풍부화 방법"을 참조한다. 일부 실시양태에서, 키트는 본원에 기재된 바와 같은 변형제를 포함한다.

다. "폴리펩티드 풍부화 방법"을 참조한다. 일부 실시양태에서, 키트는 본원에 기재된 바와 같은 친화성 시약을 포함한다. "폴리펩티드 서열분석 방법론"을 참조한다. 일부 실시양태에서, 키트는 본원에 기재된 바와 같은 표지된 펩티다제를 포함한다. "폴리펩티드 서열분석 방법론"을 참조한다.

[0235] 키트는 1종 이상의 유기체 (예를 들어, 1종 이상의 단세포 및/또는 다세포 유기체)에 특이적일 수 있다. 일부 실시양태에서, 키트는 1종 이상의 유기체의 폴리펩티드를 변형시키고, 그에 결합하고, 그에 의해 결합되는 등의 성분 (예를 들어, 풍부화 분자, 변형제, 또는 그의 조합)을 포함한다. 예를 들어, 일부 실시양태에서, 키트는 인간 프로테옴 내의 1종 이상의 공지된 폴리펩티드를 변형시키고, 그에 결합하고, 그에 의해 결합되는 등의 성분을 포함한다.

[0236] 일부 실시양태에서, 키트는 1종 이상의 질환 또는 상태에 특이적이다. 예를 들어, 키트는 종양학 키트, 심장학 키트, 유전성 질환 키트, 박테리아 병독성 인자 키트, 항생제 내성 키트, 또는 그의 조합일 수 있다.

[0237] 종양학 키트는 ABL1, ABL2, ACSL3, ACVR2A, ADAMTS20, ADGRA2, ADGRB3, ADGRL3, AFF1, AFF3, AKAP9, AKT1, AKT2, AKT3, ALK, AMER1, APC, AR, ARID1A, ARID2, ARNT, ASXL1, ATF1, ATM, ATR, ATRX, AURKA, AURKB, AURKC, AXL, BAP1, BCL10, BCL11A, BCL11B, BCL2, BCL2L1, BCL2L2, BCL3, BCL6, BCL7A, BCL9, BCR, BIRC2, BIRC3, BIRC5, BLM, BLNK, BMP1A, BRAF, BRCA1, BRCA2, BRD3, BRIP1, BTK, BUB1B, CACNA1D, CARD11, CASC5, CASP8, CBA2T3, CBF, CBL, CCND1, CCND2, CCNE1, CD79A, CD79B, CDC73, CDH1, CDH11, CDH2, CDH20, CDH5, CDK12, CDK4, CDK6, CDK8, CDKN2A, CDKN2B, CDKN2C, CEBPA, CHEK1, CHEK2, CIC, CKS1B, CMPK1, COL1A1, CRBN, CREB1, CREBBP, CRKL, CRLF2, CRTC1, CSF1R, CSMD3, CTNNA1, CTNNB1, CYLD, CYP2C19, CYP2D6, DAXX, DCC, DDB2, DDIT3, DDR2, DEK, DICER1, DNMT3A, DPYD, DST, EGFR, EML4, EP300, EP400, EPHA3, EPHA7, EPHB1, EPHB4, EPHB6, ERBB2, ERBB3, ERBB4, ERCC1, ERCC2, ERCC3, ERCC4, ERCC5, ERG, ESR1, ETS1, ETV1, ETV4, EXT1, EXT2, EZH2, FANCA, FANCC, FANCD2, FANCF, FANCG, FAS, FBXW7, FCGR2B, FGFR1, FGFR2, FGFR3, FGFR4, FH, FLCN, FLI1, FLT1, FLT3, FLT4, FN1, FOXA1, FOXL2, FOXO1, FOXO3, FOXO4, FOXO5, FOXO6, FOXO7, FOXO8, FOXO9, FOXO10, FOXO11, FOXO12, FOXO13, FOXO14, FOXO15, FOXO16, FOXO17, FOXO18, FOXO19, FOXO20, FOXO21, FOXO22, FOXO23, FOXO24, FOXO25, FOXO26, FOXO27, FOXO28, FOXO29, FOXO30, FOXO31, FOXO32, FOXO33, FOXO34, FOXO35, FOXO36, FOXO37, FOXO38, FOXO39, FOXO40, FOXO41, FOXO42, FOXO43, FOXO44, FOXO45, FOXO46, FOXO47, FOXO48, FOXO49, FOXO50, FOXO51, FOXO52, FOXO53, FOXO54, FOXO55, FOXO56, FOXO57, FOXO58, FOXO59, FOXO60, FOXO61, FOXO62, FOXO63, FOXO64, FOXO65, FOXO66, FOXO67, FOXO68, FOXO69, FOXO70, FOXO71, FOXO72, FOXO73, FOXO74, FOXO75, FOXO76, FOXO77, FOXO78, FOXO79, FOXO80, FOXO81, FOXO82, FOXO83, FOXO84, FOXO85, FOXO86, FOXO87, FOXO88, FOXO89, FOXO90, FOXO91, FOXO92, FOXO93, FOXO94, FOXO95, FOXO96, FOXO97, FOXO98, FOXO99, FOXO100, FOXO101, FOXO102, FOXO103, FOXO104, FOXO105, FOXO106, FOXO107, FOXO108, FOXO109, FOXO110, FOXO111, FOXO112, FOXO113, FOXO114, FOXO115, FOXO116, FOXO117, FOXO118, FOXO119, FOXO120, FOXO121, FOXO122, FOXO123, FOXO124, FOXO125, FOXO126, FOXO127, FOXO128, FOXO129, FOXO130, FOXO131, FOXO132, FOXO133, FOXO134, FOXO135, FOXO136, FOXO137, FOXO138, FOXO139, FOXO140, FOXO141, FOXO142, FOXO143, FOXO144, FOXO145, FOXO146, FOXO147, FOXO148, FOXO149, FOXO150, FOXO151, FOXO152, FOXO153, FOXO154, FOXO155, FOXO156, FOXO157, FOXO158, FOXO159, FOXO160, FOXO161, FOXO162, FOXO163, FOXO164, FOXO165, FOXO166, FOXO167, FOXO168, FOXO169, FOXO170, FOXO171, FOXO172, FOXO173, FOXO174, FOXO175, FOXO176, FOXO177, FOXO178, FOXO179, FOXO180, FOXO181, FOXO182, FOXO183, FOXO184, FOXO185, FOXO186, FOXO187, FOXO188, FOXO189, FOXO190, FOXO191, FOXO192, FOXO193, FOXO194, FOXO195, FOXO196, FOXO197, FOXO198, FOXO199, FOXO200, FOXO201, FOXO202, FOXO203, FOXO204, FOXO205, FOXO206, FOXO207, FOXO208, FOXO209, FOXO210, FOXO211, FOXO212, FOXO213, FOXO214, FOXO215, FOXO216, FOXO217, FOXO218, FOXO219, FOXO220, FOXO221, FOXO222, FOXO223, FOXO224, FOXO225, FOXO226, FOXO227, FOXO228, FOXO229, FOXO230, FOXO231, FOXO232, FOXO233, FOXO234, FOXO235, FOXO236, FOXO237, FOXO238, FOXO239, FOXO240, FOXO241, FOXO242, FOXO243, FOXO244, FOXO245, FOXO246, FOXO247, FOXO248, FOXO249, FOXO250, FOXO251, FOXO252, FOXO253, FOXO254, FOXO255, FOXO256, FOXO257, FOXO258, FOXO259, FOXO260, FOXO261, FOXO262, FOXO263, FOXO264, FOXO265, FOXO266, FOXO267, FOXO268, FOXO269, FOXO270, FOXO271, FOXO272, FOXO273, FOXO274, FOXO275, FOXO276, FOXO277, FOXO278, FOXO279, FOXO280, FOXO281, FOXO282, FOXO283, FOXO284, FOXO285, FOXO286, FOXO287, FOXO288, FOXO289, FOXO290, FOXO291, FOXO292, FOXO293, FOXO294, FOXO295, FOXO296, FOXO297, FOXO298, FOXO299, FOXO300, FOXO301, FOXO302, FOXO303, FOXO304, FOXO305, FOXO306, FOXO307, FOXO308, FOXO309, FOXO310, FOXO311, FOXO312, FOXO313, FOXO314, FOXO315, FOXO316, FOXO317, FOXO318, FOXO319, FOXO320, FOXO321, FOXO322, FOXO323, FOXO324, FOXO325, FOXO326, FOXO327, FOXO328, FOXO329, FOXO330, FOXO331, FOXO332, FOXO333, FOXO334, FOXO335, FOXO336, FOXO337, FOXO338, FOXO339, FOXO340, FOXO341, FOXO342, FOXO343, FOXO344, FOXO345, FOXO346, FOXO347, FOXO348, FOXO349, FOXO350, FOXO351, FOXO352, FOXO353, FOXO354, FOXO355, FOXO356, FOXO357, FOXO358, FOXO359, FOXO360, FOXO361, FOXO362, FOXO363, FOXO364, FOXO365, FOXO366, FOXO367, FOXO368, FOXO369, FOXO370, FOXO371, FOXO372, FOXO373, FOXO374, FOXO375, FOXO376, FOXO377, FOXO378, FOXO379, FOXO380, FOXO381, FOXO382, FOXO383, FOXO384, FOXO385, FOXO386, FOXO387, FOXO388, FOXO389, FOXO390, FOXO391, FOXO392, FOXO393, FOXO394, FOXO395, FOXO396, FOXO397, FOXO398, FOXO399, FOXO400, FOXO401, FOXO402, FOXO403, FOXO404, FOXO405, FOXO406, FOXO407, FOXO408, FOXO409, FOXO410, FOXO411, FOXO412, FOXO413, FOXO414, FOXO415, FOXO416, FOXO417, FOXO418, FOXO419, FOXO420, FOXO421, FOXO422, FOXO423, FOXO424, FOXO425, FOXO426, FOXO427, FOXO428, FOXO429, FOXO430, FOXO431, FOXO432, FOXO433, FOXO434, FOXO435, FOXO436, FOXO437, FOXO438, FOXO439, FOXO440, FOXO441, FOXO442, FOXO443, FOXO444, FOXO445, FOXO446, FOXO447, FOXO448, FOXO449, FOXO450, FOXO451, FOXO452, FOXO453, FOXO454, FOXO455, FOXO456, FOXO457, FOXO458, FOXO459, FOXO460, FOXO461, FOXO462, FOXO463, FOXO464, FOXO465, FOXO466, FOXO467, FOXO468, FOXO469, FOXO470, FOXO471, FOXO472, FOXO473, FOXO474, FOXO475, FOXO476, FOXO477, FOXO478, FOXO479, FOXO480, FOXO481, FOXO482, FOXO483, FOXO484, FOXO485, FOXO486, FOXO487, FOXO488, FOXO489, FOXO490, FOXO491, FOXO492, FOXO493, FOXO494, FOXO495, FOXO496, FOXO497, FOXO498, FOXO499, FOXO500, FOXO501, FOXO502, FOXO503, FOXO504, FOXO505, FOXO506, FOXO507, FOXO508, FOXO509, FOXO510, FOXO511, FOXO512, FOXO513, FOXO514, FOXO515, FOXO516, FOXO517, FOXO518, FOXO519, FOXO520, FOXO521, 또는 그의 임의의 조합에 결합하는 (또는 그에 의해 결합되는) 풍부화 분자를 포함할 수 있다.

[0238] 심장학 키트는 ABCG9, ABCG5, ABCG8, ACTA1, ACTA2, ACTC1, ACTN2, AKAP9, ALMS1, ANK2, ANKRD1, APOA4, APOA5, APOB, APOC2, APOE, BAG3, BRAF, CACNA1C, CACNA2D1, CACNB2, CALM1, CALR3, CASQ2, CAV3, CBL, CBS, CETP, COL3A1, COL5A1, COL5A2, COX15, CREB3L3, CRELD1, CRYAB, CSRP3, CTF1, DES, DMD, DNAJC19, DOLK, DPP6, DSC2, DSG2, DSP, DTNA, EFEMP2, ELN, EMD, EYA4, FBN1, FBN2, FHL1, FHL2, FKBP, FKTN, FXN, GAA, GATAD1, GCKR, GJA5, GLA, GPD1L, GPIHBP1, HADHA, HCN4, HFE, HRAS, HSPB8, ILK, JAG1, JPH2, JUP, KCNA5, KCND3, KCNE1, KCNE2, KCNE3, KCNH2, KCNJ2, KCNJ5, KCNJ8, KCNQ1, KLF10, KRAS, LAMA2, LAMA4, LAMP2, LDB3, LDLR, LDLRAP1, LMF1, LMNA, LPL, LTBP2, MAP2K1, MAP2K2, MIB1, MURC, MYBPC3, MYH11, MYH6, MYH7, MYL2, MYL3, MYLK, MYLK2, MYO6, MYOZ2, MYPN, NEXN, NKX2-5, NODAL, NOTCH1, NPPA, NRAS, PCSK9, PDLIM3, PKP2,

PLN, PRDM16, PRKAG2, PRKAR1A, PTPN11, RAF1, RANGRF, RBM20, RYR1, RYR2, SALL4, SCN1B, SCN2B, SCN3B, SCN4B, SCN5A, SCO2, SDHA, SEPN1, SGCB, SGCD, SGCG, SHOC, SLC25A4, SLC2A10, SMAD3, SMAD4, SNTA1, SOS1, SREBF2, TAZ, TBX20, TBX3, TBX5, TCAP, TGFB2, TGFB3, TGFBF1, TGFBF2, TMEM43, TMPO, TNNC1, TNNI3, TNNT2, TPM1, TRDN, TRIM63, TRPM4, TTN, TTR, TXNRD2, VCL, ZBTB17, ZHX3, 및/또는 ZIC3에 결합하는 (또는 그에 의해 결합되는) 풍부화 분자를 포함할 수 있다.

[0239] 유전성 질환 키트는 ABCA4, ABCC9, ABCD1, ACADVL, ACTA1, ACTC1, ACTN1, ADA, AIPL1, AIRE, AKAP9, ALPL, AMT, ANK2, APC, APP, APTX, ARL6, ARSA, ASL, ASPA, ATL1, ATM, ATP2A2, ATP7A, ATP7B, ATXN1, ATXN2, ATXN7, BAG3, BCKDHA, BCKDHB, BEST1, BMPR1A, BTD, BTK, CA4, CACNA1C, CACNB2, CALR3, CAPN3, CASQ2, CAV3, CCDC39, CCDC40, CDH23, CEP290, CERKL, CFTR, CHAT, CHD7, CHEK2, CHM, CHRNA1, CHRN1, CHRND, CHRNE, CLCN1, CNGB1, COL11A1, COL11A2, COL1A1, COL1A2, COL2A1, COL3A1, COL4A1, COL4A5, COL5A1, COL5A2, COL7A1, COL9A1, CRB1, CRX, CTDP1, CTNS, CYP27A1, DBT, DCX, DES, DHCR7, DKC1, DLD, DMD, DNAH11, DNAH5, DNAH9, DNAI1, DNAI2, DNM2, DOK7, DSC2, DSG2, DSP, DYSF, ELN, EMD, ENG, EXT1, EYA1, EYS, F8, F9, FANCA, FANCC, FANCF, FANCG, FBN1, FBXO7, FGFR1, FGFR3, FMO3, FOXL2, FRG1, FRMD7, FSCN2, FXN, GAA, GALT, GATA4, GBA, GBE1, GCSH, GDF5, GJB2, GJB3, GJB6, GLA, GLDC, GNE, GNPTAB, GPC3, GPD1L, GPR143, GUCY2D, HBA2, HBB, HCN4, HEXA, HFE, HIBCH, HMBS, HR, IDS, IDUA, IKBKAP, IL2RG, IMPDH1, ITGB4, JAG1, JUP, KCNE1, KCNE2, KCNE3, KCNH2, KCNJ2, KCNQ1, KCNQ4, KIAA0196, KLHL7, KRAS, KRT14, KRT5, L1CAM, LAMB3, LAMP2, LDB3, LMNA, LRAT, LRRK2, MAPT, MC1R, MECP2, MED12, MEN1, MERTK, MFN2, MLH1, MMAA, MMAB, MMACHC, MPZ, MSH2, MTM1, MUT, MYBPC3, MYH11, MYH6, MYH7, MYL2, MYL3, MYLK, MYO7A, MYOZ2, NF1, NF2, NIPBL, NKX2-5, NME8, NPC1, NPC2, NR2E3, NRAS, NSD1, OCA2, OCRL, OTC, PABPN1, PAFAH1B1, PAH, PAX3, PAX6, PCDH15, PEX1, PEX10, PEX13, PEX14, PEX19, PEX26, PEX3, PEX5, PINK1, PKD1, PKD2, PKHD1, PKP2, PLEC, PLN, PLOD1, PMM2, PMP22, POLG, PPT1, PRCD, PRKAG2, PROM1, PRPF31, PRPF8, PRPH2, PSEN1, PSEN2, PTCH1, PTPN11, RAF1, RAG1, RAG2, RAI1, RAPSN, RB1, RDH12, RET, RHO, ROR2, RP9, RPE65, RPGR, RPGRIP1, RPL11, RPL35A, RPS10, RPS19, RPS24, RPS26, RPS6KA3, RPS7, RS1, RSPH4A, RSPH9, RYR1, RYR2, SALL4, SCN1B, SCN3B, SCN4B, SCN5A, SCN9A, SEMA4A, SERPINA1, SERPING1, SGCD, SH3BP2, SIX1, SIX5, SLC25A13, SLC25A4, SLC26A4, SMAD3, SMAD4, SNCA, SNRNP200, SNTA1, SOD1, SOS1, SOX9, SPATA7, SPG7, STARD3, TAF1, TAZ, TBX5, TCOF1, TGFBR1, TGFBR2, TMEM43, TNNC1, TNNI3, TNNT1, TNNT2, TNXB, TOPORS, TP53, TPM1, TSC1, TSC2, TTPA, TTR, TULP1, TWIST1, TYR, USH1C, USH2A, VCL, VHL, WAS, WRN, WT1, 또는 그의 임의의 조합에 결합하는 (또는 그에 의해 결합되는) 풍부화 분자를 포함할 수 있다.

[0240] 박테리아 병독성 인자 키트는 <알파>-C 단백질, <알파>-용혈소, <베타>-C 단백질, <베타>-용혈소/세포용해소, <베타>-용혈소, <델타>-용혈소, <감마>-용혈소, <i>T2SS, AAFs, ACF, AI-2, ALO, AS, Ace, 산 포스파타제, 아시네토박틴, Acm, AcrAB, ActA, AdeFGH 유출 펌프, 접착성 펌브리아, Adr1, Adr2, AdsA, 에어로박틴, 에어로리신, Afa/Dr 패밀리, Agf, AhpC, Ail, AipA, 알기네이트, 알칼리성 프로테아제, 알란티온 이용, Ami, AnsP, 탄저병 독소, 항원 85, ArgP, AslA, Asp14, AtxA, 아우레오리신, 오토, 자가용해소, BFP, BSH, BabA, BadA/Vomp, Bap, BapC, BfmRS, BimA, 비오틴 합성, BoNT, BoaA, BoaB, BopD, Brk, Bsa T3SS, BslA, BtpA/Btp1/TcpB, BtpB, BvgAS, BvrR-BvrS, C2 독소, C3 독소, C5a 펩티다제, C<베타>G, CAI-1, CAMP 인자, CARDS 독소, CBPs, CDT, CHIPS, CNA, CNF-1, CNF<s>y</s>, CPAF, CPE, CT, CadF, CagA, 캡슐, 캡슐 I, CbpA/PspC, CcmC, CdpA, CdtB, Chu, CiaB, CiaC, Cif, ClpC, ClpE, ClpP, 응괴 인자, 콜리박틴, CsrA, Csu 펌브리아, Cya, CytK, 세포부착 소기관, 세포용해소, DNase, DT, DevRS, DipA, 디스페르신, Dnt, Dot/Icm, Dot/Icm T4SS, Dr 어드헤신, EAST1, ECP, EF-Tu, ESAT-6/CFP-10, ESX-1, ESX-3, ESX-5, Eap/Map, Ebp 선모, EbpS, EcbA, Efa-1/LifA, EfaA, Ent, 엔테로박틴, Erp, Esp, EspA, EspB, EspC, EspD, EspF, EspG, EspH, EspP, EtpA, Exe T2SS, 탈락 독소, ExoA, ExoS, ExoT, ExoU, ExoY, F1 항원, F1C 펌브리아, FBP, FHA, FadD33, FarAB, FbpA, FbpABC, FbsA, FbsB, FdeC, FeoAB, 펌브리아, 편모, Flp 유형 IV 선모, FmvB, FnBPs, FrgA, FsaP, Fsr, FupA, Fur, GGT, GRAB, GadC, 젤라티나제, GrvA, GspA, GtcA, HBL, HMW1/HMW2, HP-NAP, HSI-I, 혈구응집성 선모, Hap, HbhA, 열-불안정성 독소 (LT), 열-안정한 독소 (ST), 용혈소, Hgp, HhuA, Hia/Hsf, HitABC, HmbR, HopZ, Hpt, HpuAB, Hsp60, HspR, HspX, HxuABC, 히알루로네이트 리아제, 히알루론산 캡슐, 히알루로니다제, Ibes, IcsA (VirG), IcsP (SopA), IdeR, IdeS, IgA1 프로테아제, IleP, IlpA, InhA, InlA, InlB, InlC, InlF, InlJ, InlK, InlP, 세포간 부착 단백질, 인티민, 인바신, 인바신 B/Ifp, 인바신 C/Ilp, 인바신 D, IraAB, IroN, Isd, 이소시트레이트 리아제, JlpA, K1 캡슐, KatA, KatAB, KatG, LAM, LLO, LLS, LOS, LPS, Lap, LapB, LasA, LasB, 측면 편모, Lbp, 레지오박틴, Ler, LetA/S, 루이스 항원, LigA, LipF, 리파제, Lmb, LntA, LpeA, Lpf, LplA1, Lsp, 림포스타틴/LifA, M 단백질, MAM7, MARTX, MOMP, MSHA 선모, MSHA

유형 IV 선모, Map, MgtBC, MgtC, Mig-14, Mig-5, Mip, MisL, MmaA4, MntABC, Mpl, MprAB, MsrAB, MtrCDE, 미코박틴, Myf/pH6 항원, 뉴라미니다제, Nhe, 니트레이트 리덕타제, NleA/EspI, NleC, NleD, NspA, O-항원, OapA, OatA, OipA, OmpA, OmpU, Opa, Opc, P 펌브리아, P2 단백질, P44/Msp2 패밀리, P5 단백질, P97/P102 파라로그 패밀리, PDIM, PE/PE-PGRS, PEB1, PI-1, PI-2, PI-2a, PLC, PNAG, PVL, Paa, PanC/PanD, PavA, PavB, PbpG, PcaA, Pef, Per, 페르타틴, Pet, PfbA, PgdA, PhoP, PhoPQ, 포스포리파제 A2, 포스포리파제 C, 포스포리파제 D, Pht, Pic, Pili, Pla, PlcA, PlcB, Pld, 뉴모리신, 극성 편모, 포린, PrfA, PrsA2, PsaA, PspA, Ptx, 퓨린 생합성, 피오켈린, 피오시아닌, 피오베르딘, 피리미딘 생합성, 쿠오름 감지, 쿠오름 감지, 쿠오름-감지, RatB, Rck, RcsAB, RecN, RelA, 램노리피드, 리조페린, RicA, RickA, RipA, RmpA, RpoS, RtxA, Rvh T4SS, S 펌브리아, SCIN, SDr, SE, SIC, SLO, SLS, SMase, SabA, Sal, Sat, Sbi, Sca1, Sca2, Sca4, Scm, SgrA, ShET1, ShET2, ShdA, 시가 독소, Shu, SigA, SigE, SigF, SigH, SinH, SodA, SodB, SodC, SodCI, SpA, SpaP, SpeB, Spes, SprE, Spv, 스타포파인, 스타필로코아글라제, 스타필로키나제, StcE, 스트렙토키나제, Stx, 표면 지단백질, SvpA, T2SS, T3SS, T3SS1, T3SS2, T6SS, T6SS-1, T7SS, TCP, TCT, TDH, TRH, TSST-1, TTSS, TTSS(SPI-1 코딩), TTSS(SPI-2 코딩), Tap 유형 IV 선모, Tbp, TcdA, TcdB, TcfA, TcpC, TeNT, Tir, TlyC, ToxB, TraJ, Trw 유형 IV 분비 시스템, Tsh, 유형 1 펌브리아, 유형 3 펌브리아, 유형 I 펌브리아, 유형 I 선모, 유형 IV 선모, 유형 IV 분비 시스템, 유형 VII 분비 시스템, 우레아제, V8 프로테아제, VCC, VacA, Vi 항원, Vip, VirB 유형 IV 분비 시스템, VirB/VirD4 유형 IV 분비 시스템, VpadF, WhiB3, YadA, YapC, YapE, YapJ, YapK, YapV, YaxAB, Ybt, 예르시니아박틴, Ymt, Yst, Zot, 알파-클로스트리파인, 알파-독소 (CpPLC), 알파-독소 (노비이), 알파-독소 (셉티쿰), 베타-독소, 베타2-독소, enh 유전자좌, 엡실론-독소, fHbp, 이오타-독소, 카파-독소, 뮤-독소, p60, 필루스, pmiA, rOmpA/Sca0, rOmpB/Sca5, 시알리다제, 세타-독소/PF0, vWbp, xcp 분비 시스템, 또는 그의 임의의 조합에 결합하는 (또는 그에 의해 결합되는) 풍부화 분자를 포함할 수 있다.

[0241]

항생제 내성 키트는 AAC(1)-I, AAC(2')-IIa, AAC(2')-IIb, AAC(2')-Ia, AAC(2')-Ib, AAC(2')-Ic, AAC(2')-Id, AAC(2')-Ie, AAC(3)-IIIa, AAC(3)-IIIb, AAC(3)-IIIc, AAC(3)-IIa, AAC(3)-IIb, AAC(3)-IIc, AAC(3)-IId, AAC(3)-IIE, AAC(3)-IV, AAC(3)-IXa, AAC(3)-Ia, AAC(3)-Ib, AAC(3)-Ib/AAC(6')-Ib", AAC(3)-Ic, AAC(3)-Id, AAC(3)-VIIa, AAC(3)-VIIa, AAC(3)-VIa, AAC(3)-Xa, AAC(6')-29a, AAC(6')-29b, AAC(6')-30/AAC(6')-Ib' 융합 단백질, AAC(6')-31, AAC(6')-32, AAC(6')-33, AAC(6')-34, AAC(6')-I30, AAC(6')-IIa, AAC(6')-IIb, AAC(6')-IIc, AAC(6')-Ia, AAC(6')-Iaa, AAC(6')-Iad, AAC(6')-Iae, AAC(6')-Iaf, AAC(6')-Iag, AAC(6')-Iai, AAC(6')-Iaj, AAC(6')-Iak, AAC(6')-Ian, AAC(6')-Ib, AAC(6')-Ib', AAC(6')-Ib-항저우, AAC(6')-Ib-SK, AAC(6')-Ib-수조우, AAC(6')-Ib-cr, AAC(6')-Ib10, AAC(6')-Ib11, AAC(6')-Ib3, AAC(6')-Ib4, AAC(6')-Ib7, AAC(6')-Ib8, AAC(6')-Ib9, AAC(6')-Ic, AAC(6')-Ie-APH(2")-Ia, AAC(6')-If, AAC(6')-Ig, AAC(6')-Ih, AAC(6')-Ii, AAC(6')-Iid, AAC(6')-Iih, AAC(6')-Ij, AAC(6')-Ik, AAC(6')-Il, AAC(6')-Im, AAC(6')-Ip, AAC(6')-Iq, AAC(6')-Ir, AAC(6')-Is, AAC(6')-Isa, AAC(6')-It, AAC(6')-Iu, AAC(6')-Iv, AAC(6')-Iw, AAC(6')-Ix, AAC(6')-Iy, AAC(6')-Iz, ACC-1, ACC-2, ACC-3, ACC-4, ACC-5, ACI-1, ACT-1, ACT-10, ACT-12, ACT-13, ACT-14, ACT-15, ACT-16, ACT-17, ACT-18, ACT-19, ACT-2, ACT-20, ACT-21, ACT-22, ACT-23, ACT-24, ACT-25, ACT-27, ACT-28, ACT-29, ACT-3, ACT-30, ACT-31, ACT-32, ACT-33, ACT-35, ACT-36, ACT-37, ACT-38, ACT-4, ACT-5, ACT-6, ACT-7, ACT-8, ACT-9, ADC-1, ADC-10, ADC-11, ADC-12, ADC-13, ADC-14, ADC-15, ADC-16, ADC-17, ADC-18, ADC-19, ADC-2, ADC-20, ADC-21, ADC-22, ADC-23, ADC-25, ADC-3, ADC-30, ADC-31, ADC-39, ADC-4, ADC-41, ADC-42, ADC-43, ADC-44, ADC-5, ADC-56, ADC-58, ADC-59, ADC-6, ADC-60, ADC-61, ADC-62, ADC-67, ADC-68, ADC-7, ADC-73, ADC-74, ADC-75, ADC-76, ADC-77, ADC-78, ADC-79, ADC-8, ADC-81, ADC-82, AER-1, AIM-1, ANT(2")-Ia, ANT(3")-IIa, ANT(3")-IIb, ANT(3")-IIc, ANT(3")-Ii-AAC(6')-IId 융합 단백질, ANT(4')-IIa, ANT(4')-IIb, ANT(4')-Ia, ANT(4')-Ib, ANT(6)-Ia, ANT(6)-Ib, ANT(9)-Ia, APH(2")-IIIa, APH(2")-IIa, APH(2")-IVa, APH(2")-Ie, APH(2")-If, APH(2")-Ig, APH(3")-Ia, APH(3")-Ib, APH(3")-Ic, APH(3')-IIIa, APH(3')-IIa, APH(3')-IIb, APH(3')-IIc, APH(3')-IVa, APH(3')-IX, APH(3')-Ia, APH(3')-Ib, APH(3')-VI, APH(3')-VIIa, APH(3')-VIIb, APH(3')-VIIa, APH(3')-VIa, APH(3')-Va, APH(3')-Vb, APH(3')-Vc, APH(4)-Ia, APH(4)-Ib, APH(6)-Ia, APH(6)-Ib, APH(6)-Ic, APH(6)-Id, APH(7")-Ia, APH(9)-Ia, APH(9)-Ib, AQU-1, AQU-2, AQU-3, ARL-1, ARL-2, ARL-3, ARL-4, ARL-5, ARL-6, AST-1, AZECL-25, 아시네토박터 바우만니이(Acinetobacter baumannii) AbaF, 아시네토박터 바우만니이 AbaQ, 아시네토박터 바우만니이 AbuO, 아시네토박터 바우만니이 AmvA, 이미페넴에 대한 내성을 부여하는 아시네토박터 바우만니이 OprD, 아시네토박터 바우만니이 ampC 베타-락타마제, 플루오로퀴놀론에 대한 내성을 부여하는 아시네토박터 바우만니이 gyrA, 플루오로퀴놀론에 대한 내성을 부여하는 아시네토박터 바우만니이 parC, AcrE, AcrF, AcrS, 아그로박테리움 파브룸(Agrobacterium fabrum) 클로람페니콜 아세틸트랜스퍼라제, ArmR, AxyX, AxyY, AxyZ, BAT-1, BCL-1, BEL-1,

BEL-2, BEL-3, BES-1, BIC-1, BIL-1, BJP-1, BKC-1, BPU-1, BRO-1, BRO-2, BRP(MBL), BUT-1, 바실루스 클라우시(Bacillus clausii) 클로람페니콜 아세틸트랜스퍼라제, 바실루스 푸밀루스(Bacillus pumilus) cat86, 바실루스 서브틸리스(Bacillus subtilis) mprF, 담토마이신에 대한 내성을 부여하는 돌연변이를 갖는 바실루스 서브틸리스 pgsA, BahA, 플루오로퀴놀론에 대한 내성을 부여하는 바르토넬라 바실리포르미스(Bartonella bacilliformis) gyrA, 아미노카마린에 대한 내성을 부여하는 바르토넬라 바실리포르미스 gyrB, BcI, BcII, 리팜피신에 대한 내성을 부여하는 비피도박테리움 아돌레센티스(Bifidobacterium adolescentis) rpoB 돌연변이체, 뮤피로신에 대한 내성을 부여하는 비피도박테리움 ileS, Bla1, Bla2, 겐타미신에 대한 내성을 부여하는 보렐리아 부르그도르페리(Borrelia burgdorferi) 16S rRNA 돌연변이, 카나마이신에 대한 내성을 부여하는 보렐리아 부르그도르페리 16S rRNA 돌연변이, 스펙티노마이신에 대한 내성을 부여하는 보렐리아 부르그도르페리 16S rRNA 돌연변이, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 보렐리아 부르그도르페리 murA, 티로신에 대한 내성을 부여하는 돌연변이를 갖는 브라키스피라 히오디센테리아에(Brachyspira hyodysenteriae) 23S rRNA, 브루셀라 수이스(Brucella suis) mprF, 부르크홀데리아 슈도말레이(Burkholderia pseudomallei) Omp38, CAM-1, CARB-1, CARB-10, CARB-12, CARB-14, CARB-16, CARB-17, CARB-18, CARB-19, CARB-2, CARB-20, CARB-21, CARB-22, CARB-23, CARB-3, CARB-4, CARB-5, CARB-6, CARB-7, CARB-8, CARB-9, CAU-1, CBP-1, CFE-1, CFE-2, CGA-1, CGB-1, CIA-1, CIA-2, CIA-3, CIA-4, CKO-1, CME-1, CMH-1, CMY-1, CMY-10, CMY-100, CMY-101, CMY-102, CMY-103, CMY-104, CMY-105, CMY-106, CMY-108, CMY-11, CMY-110, CMY-111, CMY-112, CMY-113, CMY-114, CMY-115, CMY-116, CMY-117, CMY-118, CMY-119, CMY-12, CMY-13, CMY-131, CMY-132, CMY-133, CMY-135, CMY-14, CMY-15, CMY-16, CMY-17, CMY-18, CMY-19, CMY-2, CMY-20, CMY-21, CMY-22, CMY-23, CMY-24, CMY-25, CMY-26, CMY-27, CMY-28, CMY-29, CMY-30, CMY-31, CMY-32, CMY-33, CMY-34, CMY-35, CMY-36, CMY-37, CMY-38, CMY-39, CMY-4, CMY-40, CMY-41, CMY-42, CMY-43, CMY-44, CMY-45, CMY-46, CMY-47, CMY-48, CMY-49, CMY-5, CMY-50, CMY-51, CMY-53, CMY-54, CMY-55, CMY-56, CMY-57, CMY-58, CMY-59, CMY-6, CMY-60, CMY-61, CMY-62, CMY-63, CMY-64, CMY-65, CMY-66, CMY-67, CMY-68, CMY-69, CMY-7, CMY-70, CMY-71, CMY-72, CMY-73, CMY-74, CMY-75, CMY-76, CMY-77, CMY-78, CMY-79, CMY-8, CMY-80, CMY-81, CMY-82, CMY-83, CMY-84, CMY-85, CMY-86, CMY-87, CMY-9, CMY-90, CMY-93, CMY-94, CMY-95, CMY-98, CMY-99, CPS-1, CRP, CTX-M-1, CTX-M-10, CTX-M-100, CTX-M-101, CTX-M-102, CTX-M-103, CTX-M-104, CTX-M-105, CTX-M-106, CTX-M-107, CTX-M-108, CTX-M-109, CTX-M-11, CTX-M-110, CTX-M-111, CTX-M-112, CTX-M-113, CTX-M-114, CTX-M-115, CTX-M-116, CTX-M-117, CTX-M-12, CTX-M-121, CTX-M-122, CTX-M-123, CTX-M-124, CTX-M-125, CTX-M-126, CTX-M-129, CTX-M-13, CTX-M-130, CTX-M-131, CTX-M-132, CTX-M-134, CTX-M-136, CTX-M-137, CTX-M-139, CTX-M-14, CTX-M-141, CTX-M-142, CTX-M-144, CTX-M-147, CTX-M-148, CTX-M-15, CTX-M-151, CTX-M-152, CTX-M-155, CTX-M-156, CTX-M-157, CTX-M-158, CTX-M-159, CTX-M-16, CTX-M-160, CTX-M-17, CTX-M-19, CTX-M-2, CTX-M-20, CTX-M-21, CTX-M-22, CTX-M-23, CTX-M-24, CTX-M-25, CTX-M-26, CTX-M-27, CTX-M-28, CTX-M-29, CTX-M-3, CTX-M-30, CTX-M-31, CTX-M-32, CTX-M-33, CTX-M-34, CTX-M-35, CTX-M-36, CTX-M-37, CTX-M-38, CTX-M-39, CTX-M-4, CTX-M-40, CTX-M-41, CTX-M-42, CTX-M-43, CTX-M-44, CTX-M-45, CTX-M-46, CTX-M-47, CTX-M-48, CTX-M-49, CTX-M-5, CTX-M-50, CTX-M-51, CTX-M-52, CTX-M-53, CTX-M-54, CTX-M-55, CTX-M-56, CTX-M-58, CTX-M-59, CTX-M-6, CTX-M-60, CTX-M-61, CTX-M-62, CTX-M-63, CTX-M-64, CTX-M-65, CTX-M-66, CTX-M-67, CTX-M-68, CTX-M-69, CTX-M-7, CTX-M-71, CTX-M-72, CTX-M-74, CTX-M-75, CTX-M-76, CTX-M-77, CTX-M-78, CTX-M-79, CTX-M-8, CTX-M-80, CTX-M-81, CTX-M-82, CTX-M-83, CTX-M-84, CTX-M-85, CTX-M-86, CTX-M-87, CTX-M-88, CTX-M-89, CTX-M-9, CTX-M-90, CTX-M-91, CTX-M-92, CTX-M-93, CTX-M-94, CTX-M-95, CTX-M-96, CTX-M-98, CTX-M-99, 캄필로박터 콜라이(Campylobacter coli) 클로람페니콜 아세틸트랜스퍼라제, 에리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 캄필로박터 제주니(Campylobacter jejuni) 23S rRNA, 플루오로퀴놀론에 대한 내성을 부여하는 캄필로박터 제주니 gyrA, 플루오로퀴놀론에 대한 내성을 부여하는 카프노시토파가 긴기발리스(Capnocytophaga gingivalis) gyrA, CatU, CblA-1, CcrA, CepS, CfxA, CfxA2, CfxA3, CfxA4, CfxA5, CfxA6, 마크롤리드 항생제에 대한 내성을 부여하는 돌연변이를 갖는 클라미디아 트라코마티스(Chlamydia trachomatis) 23S rRNA, 포스포마이신에 대한 내성을 부여하는 클라미디아 트라코마티스 내인성 murA, 스트렙토마이신에 대한 내성을 부여하는 클라미도모나스 레인하르티이(Chlamydomonas reinhardtii) 16S rRNA (rrnS) 돌연변이, 에리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 클라미도모나스 레인하르티이 23S rRNA, 스펙티노마이신에 대한 내성을 부여하는 클라미도필라 프시타시(Chlamydophila psittaci) 16S rRNA 돌연변이, 크리세오박테리움 메닝고셉티쿰(Chryseobacterium meningosepticum) BlaB, 에리트로마이신 및 클린다마이신에 대한 내성을 부여하는 돌연변이를 갖는 클로스트리디오이데스 디피실레(Clostridioides difficile) 23S rRNA, 엘파마이신에 대한 내성을 부여하는 클로스트리디오이데스 디피실레 EF-Tu 돌연변이체, 플루오로퀴놀론에 대한 내성을 부여하는 클로스트리디오이데스 디피실레 gyrA, 플루오로퀴놀론에 대한 내성을 부여하는 클로스트리디오이데스 디피실레 gyrB, 반코마

이신에 대한 내성을 부여하는 돌연변이를 갖는 클로스트리디오테스 디피실레 murG, 리팜피신에 대한 내성을 부여하는 돌연변이를 갖는 클로스트리디오테스 디피실레 rpoB, 반코마이신에 대한 내성을 부여하는 돌연변이를 갖는 클로스트리디오테스 디피실레 rpoC, 클로스트리디움 부티리쿰(*Clostridium butyricum*) catB, 클로스트리디움 페르프링겐스(*Clostridium perfringens*) mprF, 코리네박테리움 스트리아툼(*Corynebacterium striatum*) tetA, CrpP, 테트라시클린에 대한 내성을 부여하는 큐티박테리움 아크네스(*Cutibacterium acnes*) 16S rRNA 돌연변이, 플루오로퀴놀론에 대한 내성을 부여하는 큐티박테리움 아크네스 gyrA, D-Ala-D-Ala 리가제, DES-1, DHA-1, DHA-10, DHA-12, DHA-13, DHA-14, DHA-15, DHA-16, DHA-17, DHA-18, DHA-19, DHA-2, DHA-20, DHA-21, DHA-22, DHA-3, DHA-5, DHA-6, DHA-7, DHA-9, DIM-1, DnaA, EBR-1 베타-락타마제, EBR-2, ERP-1, ESP-1, EXO-1, EdeQ, 엔테로박터 클로아카에(*Enterobacter cloacae*) acrA, 엔테로박터 클로아카에 rob, 답토마이신 내성을 부여하는 돌연변이를 갖는 엔테로코쿠스 파에칼리스(*Enterococcus faecalis*) YvlB, 답토마이신 내성을 부여하는 돌연변이를 갖는 엔테로코쿠스 파에칼리스 YybT, 엔테로코쿠스 파에칼리스 클로람페니콜 아세틸트랜스퍼라제, 답토마이신에 대한 내성을 부여하는 돌연변이를 갖는 엔테로코쿠스 파에칼리스 cls, 답토마이신 내성을 부여하는 돌연변이를 갖는 엔테로코쿠스 파에칼리스 drmA, 답토마이신 내성을 부여하는 돌연변이를 갖는 엔테로코쿠스 파에칼리스 gdpD, 답토마이신 내성을 부여하는 돌연변이를 갖는 엔테로코쿠스 파에칼리스 gshF, 답토마이신 내성을 부여하는 엔테로코쿠스 파에칼리스 liaF 돌연변이체, 답토마이신 내성을 부여하는 엔테로코쿠스 파에칼리스 liaR 돌연변이체, 답토마이신 내성을 부여하는 엔테로코쿠스 파에칼리스 liaS 돌연변이체, GE2270A에 대한 내성을 부여하는 엔테로코쿠스 파에시움(*Enterococcus faecium*) EF-Tu 돌연변이체, 엔테로코쿠스 파에시움 클로람페니콜 아세틸트랜스퍼라제, 답토마이신에 대한 내성을 부여하는 엔테로코쿠스 파에시움 cls, 답토마이신 내성을 부여하는 엔테로코쿠스 파에시움 liaF 돌연변이체, 답토마이신 내성을 부여하는 엔테로코쿠스 파에시움 liaR 돌연변이체, 답토마이신 내성을 부여하는 엔테로코쿠스 파에시움 liaS 돌연변이체, EreA, EreA2, EreB, EreD, Erm(30), Erm(31), Erm(33), Erm(34), Erm(35), Erm(36), Erm(37), Erm(38), Erm(39), Erm(41), Erm(42), Erm(43), Erm(44)v, Erm(47), Erm(48), Erm(49), Erm(K), Erm(O)-lrm, ErmA, ErmB, ErmC, ErmD, ErmE, ErmF, ErmG, ErmH, ErmN, ErmO-srmA, ErmQ, ErmR, ErmS, ErmT, ErmU, ErmV, ErmW, ErmX, ErmY, 스펙티노마이신에 대한 내성을 부여하는 에스케리키아 콜라이(*Escherichia coli*) 16S rRNA (rrnB) 돌연변이, 스트렙토마이신에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrnB) 돌연변이, 테트라시클린에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrnB) 돌연변이, G418에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 젠타미신 C에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 카나마이신 A에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 네오마이신에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 파로모마이신에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 스펙티노마이신에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 스트렙토마이신에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 테트라시클린에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 토브라마이신에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsB) 돌연변이, 카수가마이신에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsC) 돌연변이, 스펙티노마이신에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA (rrsH) 돌연변이, 에데인에 대한 내성을 부여하는 에스케리키아 콜라이 16S rRNA 돌연변이, 클로람페니콜에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 23S rRNA, 클라리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 23S rRNA, 클린다마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 23S rRNA, 에리트로마이신 및 텔리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 23S rRNA, 옥사졸리디논 항생제에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 23S rRNA, 에스케리키아 콜라이 CpxR, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 CyaA, 에나실록신 IIa에 대한 내성을 부여하는 에스케리키아 콜라이 EF-Tu 돌연변이체, 폴보마이신에 대한 내성을 부여하는 에스케리키아 콜라이 EF-Tu 돌연변이체, 키로마이신에 대한 내성을 부여하는 에스케리키아 콜라이 EF-Tu 돌연변이체, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 GlpT, 에스케리키아 콜라이 LamB, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 PtsI, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 UhpA, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 UhpT, 에스케리키아 콜라이 acrA, 다중약물 항생제 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 acrR, 에스케리키아 콜라이 ampC 베타-락타마제, 에스케리키아 콜라이 ampC1 베타-락타마제, 에스케리키아 콜라이 ampH 베타-락타마제, 에스케리키아 콜라이 emrE, 트리클로산에 대한 내성을 부여하는 에스케리키아 콜라이 fabG 돌연변이, 이소니아지드 및 트리클로산에 대한 내성을 부여하는 에스케리키아 콜라이 fabI 돌연변이, 술폰아미드에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 folP, 플루오로퀴놀론에 대한 내성을 부여하는 에스케리키아 콜라이 gyrA, 트리클로산에 대한 내성을

부여하는 돌연변이를 갖는 에스케리키아 콜라이 gyrA, 아미노쿠마린에 대한 내성을 부여하는 에스케리키아 콜라이 gyrB, 항생제 내성을 부여하는 에스케리키아 콜라이 marR 돌연변이체, 에스케리키아 콜라이 mdfA, 에스케리키아 콜라이 mipA, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 murA, 니트로푸란토인에 대한 내성을 부여하는 에스케리키아 콜라이 nfsA 돌연변이, 니트로푸란토인에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 nfsB, 베타-락탐 항생제에 대한 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 ompF, 플루오로퀴놀론에 대한 내성을 부여하는 에스케리키아 콜라이 parC, 플루오로퀴놀론에 대한 내성을 부여하는 에스케리키아 콜라이 parE, 에스케리키아 콜라이 rob, 리팜피신에 대한 내성을 부여하는 에스케리키아 콜라이 rpoB 돌연변이체, 항생제 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 soxR, 항생제 내성을 부여하는 돌연변이를 갖는 에스케리키아 콜라이 soxS, FAR-1, FEZ-1, FIM-1, FONA-1, FONA-2, FONA-3, FONA-4, FONA-5, FONA-6, FOX-1, FOX-10, FOX-2, FOX-3, FOX-4, FOX-5, FOX-7, FOX-8, FOX-9, FPH-1, FRI-1, FRI-2, FRI-3, FTU-1, FomA, FomB, FosA, FosA2, FosA3, FosA4, FosA5, FosA6, FosA7, FosB, FosB1, FosB3, FosB4, FosB5, FosB6, FosC, FosC2, FosD, FosK, FosX, FusF, GES-1, GES-10, GES-11, GES-12, GES-13, GES-14, GES-15, GES-16, GES-17, GES-18, GES-19, GES-2, GES-20, GES-21, GES-22, GES-23, GES-24, GES-26, GES-3, GES-4, GES-5, GES-6, GES-7, GES-8, GES-9, GIM-1, GIM-2, GOB-1, GOB-10, GOB-11, GOB-12, GOB-13, GOB-14, GOB-15, GOB-16, GOB-18, GOB-2, GOB-3, GOB-4, GOB-5, GOB-6, GOB-7, GOB-8, GOB-9, H-NS, HERA-1, HERA-2, HERA-3, HMB-1, 베타-락탐 항생제에 대한 내성을 부여하는 헤모필루스 인플루엔자에 (*Haemophilus influenzae*) PBP3, 플루오로퀴놀론에 대한 내성을 부여하는 헤모필루스 파라인플루엔자에 (*Haemophilus parainfluenzae*) gyrA, 플루오로퀴놀론에 대한 내성을 부여하는 헤모필루스 파라인플루엔자에 parC, 클로람페니콜에 대한 내성을 부여하는 할로박테리움 할로비움(*Halobacterium halobium*) 23S rRNA 돌연변이, 팍타마이신에 대한 내성을 부여하는 할로박테리움 살리나룸(*Halobacterium salinarum*) 16S rRNA 돌연변이, 테트라시클린에 대한 내성을 부여하는 헬리코박터 필로리(*Helicobacter pylori*) 16S rRNA 돌연변이, 클라리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 헬리코박터 필로리 23S rRNA, ICR-Mc, ICR-Mo, IMI-1, IMI-2, IMI-3, IMI-4, IMI-7, IMP-1, IMP-10, IMP-11, IMP-12, IMP-13, IMP-14, IMP-15, IMP-16, IMP-18, IMP-19, IMP-2, IMP-20, IMP-21, IMP-22, IMP-24, IMP-25, IMP-26, IMP-27, IMP-28, IMP-29, IMP-3, IMP-30, IMP-31, IMP-32, IMP-33, IMP-34, IMP-35, IMP-37, IMP-38, IMP-4, IMP-40, IMP-41, IMP-42, IMP-43, IMP-44, IMP-45, IMP-48, IMP-5, IMP-51, IMP-55, IMP-56, IMP-6, IMP-7, IMP-8, IMP-9, IND-1, IND-10, IND-11, IND-12, IND-14, IND-15, IND-2, IND-2a, IND-3, IND-4, IND-5, IND-6, IND-7, IND-8, IND-9, JOHN-1, KHM-1, KPC-10, KPC-11, KPC-12, KPC-13, KPC-14, KPC-15, KPC-16, KPC-17, KPC-19, KPC-2, KPC-22, KPC-24, KPC-3, KPC-4, KPC-5, KPC-6, KPC-7, KPC-8, KPC-9, 클레브시엘라 아에로게네스(*Klebsiella aerogenes*) Omp36, 다중약물 항생제 내성을 부여하는 돌연변이를 갖는 클레브시엘라 아에로게네스 acrR, 콜리스틴에 대한 항생제 내성을 부여하는 클레브시엘라 돌연변이체 PhoP, 클레브시엘라 뉴모니아에(*Klebsiella pneumoniae*) KpnE, 클레브시엘라 뉴모니아에 KpnF, 클레브시엘라 뉴모니아에 KpnG, 클레브시엘라 뉴모니아에 KpnH, 클레브시엘라 뉴모니아에 OmpK35, 클레브시엘라 뉴모니아에 OmpK36, 클레브시엘라 뉴모니아에 OmpK37, 클레브시엘라 뉴모니아에 acrA, 다중약물 항생제 내성을 부여하는 돌연변이를 갖는 클레브시엘라 뉴모니아에 acrR, 클레브시엘라 뉴모니아에 ramR 돌연변이체, L1 베타-락타마제, LAT-1, LCR-1, LEN-1, LEN-10, LEN-11, LEN-12, LEN-13, LEN-14, LEN-15, LEN-16, LEN-18, LEN-19, LEN-2, LEN-20, LEN-21, LEN-22, LEN-23, LEN-24, LEN-26, LEN-3, LEN-4, LEN-5, LEN-6, LEN-7, LEN-8, LEN-9, LRA-1, LRA-10, LRA-12, LRA-13, LRA-17, LRA-18, LRA-19, LRA-2, LRA-3, LRA-5, LRA-7, LRA-8, LRA-9, 락토바실루스 레우테리(*Lactobacillus reuteri*) cat-TC, 라리박터 홍콩겐시스(*Laribacter hongkongensis*) ampC 베타-락타마제, 리스테리아 모노시토키테네스(*Listeria monocytogenes*) mprF, LlmA 23S 리보솜 RNA 메틸트랜스퍼라제, LnuP, LpeA, LpeB, LpxA, LpxC, LpxD, MCR-1.1, MCR-1.10, MCR-1.11, MCR-1.12, MCR-1.13, MCR-1.2, MCR-1.3, MCR-1.4, MCR-1.5, MCR-1.6, MCR-1.7, MCR-1.8, MCR-1.9, MCR-2.1, MCR-2.2, MCR-3.1, MCR-3.10, MCR-3.11, MCR-3.12, MCR-3.2, MCR-3.3, MCR-3.4, MCR-3.5, MCR-3.6, MCR-3.7, MCR-3.8, MCR-3.9, MCR-4.1, MCR-4.2, MCR-4.3, MCR-4.4, MCR-4.5, MCR-5.1, MCR-5.2, MCR-6.1, MCR-7.1, MCR-8.1, MCR-9.1, MIR-1, MIR-10, MIR-11, MIR-12, MIR-13, MIR-14, MIR-15, MIR-16, MIR-17, MIR-2, MIR-3, MIR-4, MIR-5, MIR-6, MIR-8, MIR-9, MOX-1, MOX-2, MOX-3, MOX-4, MOX-5, MOX-6, MOX-7, MOX-8, MOX-9, MSI-1, MSI-OXA, MUS-1, MUS-2, MdtK, Mef(En2), MexA, MexB, MexC, MexD, MexE, MexF, MexG, MexH, MexI, MexJ, MexK, MexL, MexR, MexS, MexT, MexV, MexW, MexZ, 마크롤리드 항생제에 대한 내성을 부여하는 돌연변이를 갖는 모락셀라 카타랄리스(*Moraxella catarrhalis*) 23S rRNA, 모락셀라 카타랄리스 M35, 플루오로퀴놀론에 대한 내성을 부여하는 모르가넬라 모르가니이(*Morganella morganii*) gyrB, MuxA, MuxB, MuxC, MvaT, 클라리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 아비움(*Mycobacterium avium*) 23S rRNA, 아지트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 인트라셀룰라레(*Mycobacterium intracellulare*) 23S

rRNA, 클라리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 인트라셀룰라레 23S rRNA, 클라리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 칸사시이(*Mycobacterium kansasii*) 23S rRNA, 답손에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 레프라에(*Mycobacterium leprae*) folP, 플루오로퀴놀론에 대한 내성을 부여하는 미코박테리움 레프라에 gyrB, 리팜피신에 대한 내성을 부여하는 미코박테리움 레프라에 rpoB 돌연변이, 아미카신에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스(*Mycobacterium tuberculosis*) 16S rRNA 돌연변이, 카나마이신에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 16S rRNA 돌연변이, 스트렙토마이신에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 16S rRNA 돌연변이, 비오마이신에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 16S rRNA 돌연변이, 에탐부톨에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 embA 돌연변이체, 에탐부톨에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 embB, 리팜피신에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 embB, 에탐부톨에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 embR 돌연변이체, 에티온아미드에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 ethA, 파라-아미노살리실산에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 folC, 스트렙토마이신에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 gidB 돌연변이, 플루오로퀴놀론에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 gyrA, 플루오로퀴놀론에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 gyrB 돌연변이체, 이소니아지드에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 inhA 돌연변이, 에탐부톨에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 iniA 돌연변이체, 에탐부톨에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 iniB, 에탐부톨에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 iniC 돌연변이체, 포스포마이신에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 내인성 murA, 이소니아지드에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 kasA 돌연변이체, 이소니아지드에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 katG 돌연변이, 에탐부톨에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 돌연변이체 embC, 이소니아지드에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 ndh, 피라진아미드에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 pncA 돌연변이, 파라-아미노살리실산에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 ribD, 리팜피신에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 rpoB 돌연변이체, 피라진아미드에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 rpsA 돌연변이, 스트렙토마이신에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 rpsL 돌연변이, 파라-아미노살리실산에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 thyA, 아미노글리코시드에 대한 내성을 부여하는 미코박테리움 투베르쿨로시스 tlyA 돌연변이, 에탐부톨에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 변이체 보비스 embB, 이소니아지드에 대한 내성을 부여하는 돌연변이를 갖는 미코박테리움 투베르쿨로시스 변이체 보비스 ndh, 아미카신에 대한 내성을 부여하는 미코박테로이데스 압세수스(*Mycobacteroides abscessus*) 16S rRNA 돌연변이, 겐타미신에 대한 내성을 부여하는 미코박테로이데스 압세수스 16S rRNA 돌연변이, 카나마이신에 대한 내성을 부여하는 미코박테로이데스 압세수스 16S rRNA 돌연변이, 네오마이신에 대한 내성을 부여하는 미코박테로이데스 압세수스 16S rRNA 돌연변이, 토브라마이신에 대한 내성을 부여하는 미코박테로이데스 압세수스 16S rRNA 돌연변이, 클라리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 미코박테로이데스 압세수스 23S rRNA, 아미카신에 대한 내성을 부여하는 미코박테로이데스 켈로나에(*Mycobacteroides chelonae*) 16S rRNA 돌연변이, 겐타미신 C에 대한 내성을 부여하는 미코박테로이데스 켈로나에 16S rRNA 돌연변이, 카나마이신 A에 대한 내성을 부여하는 미코박테로이데스 켈로나에 16S rRNA 돌연변이, 네오마이신에 대한 내성을 부여하는 미코박테로이데스 켈로나에 16S rRNA 돌연변이, 토브라마이신에 대한 내성을 부여하는 미코박테로이데스 켈로나에 16S rRNA 돌연변이, 클라리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 미코박테로이데스 켈로나에 23S rRNA, 플루오로퀴놀론에 대한 내성을 부여하는 미코박테리움 레프라에(*Mycobacterium leprae*) gyrA, 히그로마이신 B에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스(*Mycolicibacterium smegmatis*) 16S rRNA (rrsA) 돌연변이, 카나마이신 A에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스 16S rRNA (rrsA) 돌연변이, 네오마이신에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스 16S rRNA (rrsA) 돌연변이, 비오마이신에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스 16S rRNA (rrsA) 돌연변이, 히그로마이신 B에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스 16S rRNA (rrsB) 돌연변이, 카나마이신 A에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스 16S rRNA (rrsB) 돌연변이, 네오마이신에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스 16S rRNA (rrsB) 돌연변이, 스트렙토마이신에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스 16S rRNA (rrsB) 돌연변이, 비오마이신에 대한 내성을 부여하는 미콜리시박테리움 스메그마티스 16S rRNA (rrsB) 돌연변이, 클라리트로마이신에 대한 내성을 부여하는 돌연변이를 갖는 미콜리시박테리움 스메그마티스 23S rRNA, 이소니아지드에 대한 내성을 부여하는 돌연변이를 갖는 미콜리시박테리움 스메그마티스 ndh, 마크롤리드 항생제에 대한 내성을 부여하는 돌연변이를 갖는 미코플라스마 페르텐

탄스(*Mycoplasma fermentans*) 23S rRNA, 플루로무틸린 항생제에 대한 내성을 부여하는 미코플라스마 갈리셉티쿰(*Mycoplasma gallisepticum*) 23S rRNA 돌연변이, 플루오로퀴놀론 및 마크롤리드 항생제에 대한 내성을 부여하는 미코플라스마 게니탈리움(*Mycoplasma genitalium*) 23S rRNA 돌연변이, 플루오로퀴놀론에 대한 내성을 부여하는 미코플라스마 게니탈리움 *gyrA* 돌연변이, 목시플록사신에 대한 내성을 부여하는 미코플라스마 게니탈리움 *parC* 돌연변이, 마크롤리드 항생제에 대한 내성을 부여하는 돌연변이를 갖는 미코플라스마 호미니스(*Mycoplasma hominis*) 23S rRNA, 플루오로퀴놀론에 대한 내성을 부여하는 미코플라스마 호미니스 *parC*, 에리트로마이신에 대한 내성을 부여하는 미코플라스마 뉴모니아에(*Mycoplasma pneumoniae*) 23S rRNA 돌연변이, NDM-1, NDM-10, NDM-11, NDM-12, NDM-13, NDM-14, NDM-17, NDM-2, NDM-3, NDM-4, NDM-5, NDM-6, NDM-7, NDM-8, NDM-9, NPS-1, 스펙티노마이신에 대한 내성을 부여하는 네이세리아 고노레아에(*Neisseria gonorrhoeae*) 16S rRNA 돌연변이, 플루오로퀴놀론에 대한 내성을 부여하는 네이세리아 고노레아에 *gyrA*, 플루오로퀴놀론에 대한 내성을 부여하는 네이세리아 고노레아에 *parC*, 네이세리아 고노레아에 포린 PIB (*por*), 스펙티노마이신에 대한 내성을 부여하는 네이세리아 메닝기티디스(*Neisseria meningitidis*) 16S rRNA 돌연변이, 베타-락탐에 대한 내성을 부여하는 네이세리아 메닝기티디스 PBP2, NmcA, NmcR, OCH-1, OCH-2, OCH-3, OCH-4, OCH-5, OCH-6, OCH-7, OCH-8, OKP-A-1, OKP-A-10, OKP-A-11, OKP-A-12, OKP-A-13, OKP-A-14, OKP-A-15, OKP-A-16, OKP-A-2, OKP-A-3, OKP-A-4, OKP-A-5, OKP-A-6, OKP-A-7, OKP-A-8, OKP-A-9, OKP-B-1, OKP-B-10, OKP-B-11, OKP-B-12, OKP-B-13, OKP-B-17, OKP-B-18, OKP-B-19, OKP-B-2, OKP-B-20, OKP-B-3, OKP-B-4, OKP-B-5, OKP-B-6, OKP-B-7, OKP-B-8, OKP-B-9, OXA-1, OXA-10, OXA-100, OXA-101, OXA-103, OXA-104, OXA-106, OXA-107, OXA-108, OXA-109, OXA-11, OXA-110, OXA-111, OXA-112, OXA-113, OXA-114a, OXA-115, OXA-116, OXA-117, OXA-118, OXA-119, OXA-12, OXA-120, OXA-121, OXA-122, OXA-123, OXA-124, OXA-125, OXA-126, OXA-127, OXA-128, OXA-129, OXA-13, OXA-130, OXA-131, OXA-132, OXA-133, OXA-134, OXA-135, OXA-136, OXA-137, OXA-138, OXA-139, OXA-14, OXA-140, OXA-141, OXA-142, OXA-143, OXA-144, OXA-145, OXA-146, OXA-147, OXA-148, OXA-149, OXA-15, OXA-150, OXA-151, OXA-152, OXA-153, OXA-154, OXA-155, OXA-156, OXA-157, OXA-158, OXA-16, OXA-160, OXA-161, OXA-162, OXA-163, OXA-164, OXA-165, OXA-166, OXA-167, OXA-168, OXA-169, OXA-17, OXA-170, OXA-171, OXA-172, OXA-173, OXA-174, OXA-175, OXA-176, OXA-177, OXA-178, OXA-179, OXA-18, OXA-180, OXA-181, OXA-182, OXA-183, OXA-184, OXA-185, OXA-19, OXA-192, OXA-194, OXA-195, OXA-196, OXA-197, OXA-198, OXA-199, OXA-2, OXA-20, OXA-200, OXA-201, OXA-202, OXA-203, OXA-204, OXA-205, OXA-206, OXA-207, OXA-208, OXA-209, OXA-21, OXA-210, OXA-211, OXA-212, OXA-213, OXA-214, OXA-215, OXA-216, OXA-217, OXA-219, OXA-22, OXA-223, OXA-224, OXA-225, OXA-226, OXA-228, OXA-229, OXA-23, OXA-230, OXA-231, OXA-232, OXA-233, OXA-234, OXA-235, OXA-236, OXA-237, OXA-239, OXA-24, OXA-240, OXA-241, OXA-242, OXA-243, OXA-244, OXA-245, OXA-246, OXA-247, OXA-248, OXA-249, OXA-25, OXA-250, OXA-251, OXA-252, OXA-253, OXA-254, OXA-255, OXA-256, OXA-257, OXA-258, OXA-259, OXA-26, OXA-260, OXA-261, OXA-262, OXA-263, OXA-264, OXA-265, OXA-266, OXA-267, OXA-268, OXA-269, OXA-27, OXA-270, OXA-271, OXA-272, OXA-273, OXA-274, OXA-275, OXA-276, OXA-277, OXA-278, OXA-279, OXA-28, OXA-280, OXA-281, OXA-282, OXA-283, OXA-284, OXA-285, OXA-286, OXA-287, OXA-288, OXA-29, OXA-291, OXA-292, OXA-293, OXA-294, OXA-295, OXA-296, OXA-297, OXA-298, OXA-299, OXA-3, OXA-300, OXA-301, OXA-302, OXA-303, OXA-304, OXA-305, OXA-306, OXA-308, OXA-309, OXA-31, OXA-312, OXA-313, OXA-314, OXA-315, OXA-316, OXA-317, OXA-32, OXA-320, OXA-322, OXA-323, OXA-324, OXA-325, OXA-326, OXA-327, OXA-328, OXA-329, OXA-33, OXA-330, OXA-331, OXA-332, OXA-333, OXA-334, OXA-335, OXA-336, OXA-337, OXA-338, OXA-339, OXA-34, OXA-340, OXA-341, OXA-342, OXA-343, OXA-344, OXA-345, OXA-346, OXA-347, OXA-348, OXA-349, OXA-35, OXA-350, OXA-351, OXA-352, OXA-353, OXA-354, OXA-355, OXA-356, OXA-357, OXA-358, OXA-359, OXA-36, OXA-360, OXA-361, OXA-362, OXA-363, OXA-364, OXA-365, OXA-366, OXA-368, OXA-37, OXA-370, OXA-371, OXA-372, OXA-373, OXA-374, OXA-375, OXA-376, OXA-377, OXA-378, OXA-379, OXA-380, OXA-381, OXA-382, OXA-383, OXA-384, OXA-385, OXA-386, OXA-387, OXA-388, OXA-389, OXA-390, OXA-391, OXA-392, OXA-397, OXA-398, OXA-4, OXA-400, OXA-401, OXA-402, OXA-403, OXA-404, OXA-405, OXA-406, OXA-407, OXA-408, OXA-409, OXA-411, OXA-412, OXA-413, OXA-414, OXA-415, OXA-416, OXA-417, OXA-418, OXA-42, OXA-420, OXA-421, OXA-422, OXA-423, OXA-424, OXA-425, OXA-426, OXA-427, OXA-429, OXA-43, OXA-430, OXA-431, OXA-432, OXA-433, OXA-435, OXA-436, OXA-437, OXA-438, OXA-439, OXA-440, OXA-441, OXA-442, OXA-443, OXA-444, OXA-446, OXA-447, OXA-448, OXA-449, OXA-45, OXA-450, OXA-451, OXA-452, OXA-453, OXA-454, OXA-455, OXA-457, OXA-458, OXA-459, OXA-46, OXA-460, OXA-461, OXA-464, OXA-465, OXA-466, OXA-47, OXA-470, OXA-471, OXA-472, OXA-473, OXA-474, OXA-475, OXA-476, OXA-477, OXA-478, OXA-479, OXA-48, OXA-480, OXA-482, OXA-483, OXA-484, OXA-485, OXA-486, OXA-488, OXA-49, OXA-5, OXA-50, OXA-51, OXA-53, OXA-535, OXA-54, OXA-55, OXA-56, OXA-57, OXA-58, OXA-59, OXA-60, OXA-61, OXA-62, OXA-63, OXA-64, OXA-65, OXA-66, OXA-663,

OXA-664, OXA-665, OXA-67, OXA-68, OXA-69, OXA-7, OXA-70, OXA-71, OXA-72, OXA-73, OXA-74, OXA-75, OXA-76, OXA-77, OXA-78, OXA-79, OXA-80, OXA-82, OXA-83, OXA-84, OXA-85, OXA-86, OXA-87, OXA-88, OXA-89, OXA-9, OXA-90, OXA-91, OXA-92, OXA-93, OXA-94, OXA-95, OXA-96, OXA-97, OXA-98, OXA-99, OXY-1-1, OXY-1-2, OXY-1-3, OXY-1-4, OXY-1-6, OXY-2-1, OXY-2-10, OXY-2-2, OXY-2-3, OXY-2-4, OXY-2-5, OXY-2-6, OXY-2-7, OXY-2-8, OXY-2-9, OXY-3-1, OXY-4-1, OXY-5-1, OXY-5-2, OXY-6-1, OXY-6-2, OXY-6-3, OXY-6-4, OpmB, OpmD, OpmH, OprA, OprJ, OprM, OprN, OprZ, PC1 베타-락타마제 (blaZ), PDC-1, PDC-10, PDC-2, PDC-3, PDC-4, PDC-5, PDC-6, PDC-7, PDC-73, PDC-74, PDC-75, PDC-76, PDC-77, PDC-78, PDC-79, PDC-8, PDC-80, PDC-81, PDC-82, PDC-83, PDC-84, PDC-85, PDC-86, PDC-87, PDC-88, PDC-89, PDC-9, PDC-90, PDC-91, PDC-92, PDC-93, PEDO-1, PEDO-2, PEDO-3, PER-1, PER-2, PER-3, PER-4, PER-5, PER-6, PER-7, PNGM-1, 스펙티노마이신에 대한 내성을 부여하는 파스테우렐라 물토시다(*Pasteurella multocida*) 16S rRNA 돌연변이, 억제제 GE2270A에 대한 내성을 부여하는 플라노비스포라 로세아(*Planobispora rosea*) EF-Tu 돌연변이체, PmpM, PmrF, 마크롤리드 항생제에 대한 내성을 부여하는 돌연변이를 갖는 프로피오니박테리아(*Propionibacteria*) 23S rRNA, 슈도모나스 아에루기노사(*Pseudomonas aeruginosa*) CpxR, 슈도모나스 아에루기노사 catB6, 슈도모나스 아에루기노사 catB7, 슈도모나스 아에루기노사 emrE, 플루오로퀴놀론에 대한 내성을 부여하는 슈도모나스 아에루기노사 gyrA 및 parC, 플루오로퀴놀론에 대한 내성을 부여하는 슈도모나스 아에루기노사 gyrA, 이미페넴에 대한 내성을 부여하는 돌연변이를 갖는 슈도모나스 아에루기노사 oprD, 플루오로퀴놀론에 대한 내성을 부여하는 슈도모나스 아에루기노사 parE, 슈도모나스 아에루기노사 soxR, 콜리스틴에 대한 내성을 부여하는 슈도모나스 돌연변이체 PhoP, 콜리스틴에 대한 내성을 부여하는 슈도모나스 돌연변이체 PhoQ, PvrR, QepA1, QepA2, QepA3, QepA4, QnrA1, QnrA2, QnrA3, QnrA4, QnrA5, QnrA6, QnrA7, QnrB1, QnrB10, QnrB11, QnrB12, QnrB13, QnrB14, QnrB15, QnrB16, QnrB17, QnrB18, QnrB19, QnrB2, QnrB20, QnrB21, QnrB22, QnrB23, QnrB24, QnrB25, QnrB26, QnrB27, QnrB28, QnrB29, QnrB3, QnrB30, QnrB31, QnrB32, QnrB33, QnrB34, QnrB35, QnrB36, QnrB37, QnrB38, QnrB4, QnrB40, QnrB41, QnrB42, QnrB43, QnrB44, QnrB45, QnrB46, QnrB47, QnrB48, QnrB49, QnrB5, QnrB50, QnrB54, QnrB55, QnrB56, QnrB57, QnrB58, QnrB59, QnrB6, QnrB60, QnrB61, QnrB62, QnrB64, QnrB65, QnrB66, QnrB67, QnrB68, QnrB69, QnrB7, QnrB70, QnrB71, QnrB72, QnrB73, QnrB74, QnrB8, QnrB9, QnrC, QnrD1, QnrD2, QnrS1, QnrS10, QnrS11, QnrS12, QnrS15, QnrS2, QnrS3, QnrS4, QnrS5, QnrS6, QnrS7, QnrS8, QnrS9, QnrVC1, QnrVC3, QnrVC4, QnrVC5, QnrVC6, QnrVC7, R39, RCP-1, ROB-1, RSA-1, RSA-2, RbpA, 로도박터 스파에로이데스(*Rhodobacter sphaeroides*) ampC 베타-락타마제, 로도코쿠스 파시안스(*Rhodococcus fascians*) cmr, RlmA(II), Rm3, SAT-2, SAT-3, SAT-4, SFB-1, SFH-1, SHV-1, SHV-100, SHV-101, SHV-102, SHV-103, SHV-104, SHV-105, SHV-106, SHV-107, SHV-108, SHV-109, SHV-11, SHV-110, SHV-111, SHV-112, SHV-119, SHV-12, SHV-120, SHV-121, SHV-122, SHV-123, SHV-124, SHV-125, SHV-126, SHV-127, SHV-128, SHV-129, SHV-13, SHV-133, SHV-134, SHV-135, SHV-137, SHV-14, SHV-140, SHV-141, SHV-142, SHV-143, SHV-144, SHV-145, SHV-147, SHV-148, SHV-149, SHV-15, SHV-150, SHV-151, SHV-152, SHV-153, SHV-154, SHV-155, SHV-156, SHV-157, SHV-158, SHV-159, SHV-16, SHV-160, SHV-161, SHV-162, SHV-163, SHV-164, SHV-165, SHV-167, SHV-168, SHV-172, SHV-173, SHV-178, SHV-179, SHV-18, SHV-180, SHV-182, SHV-183, SHV-185, SHV-186, SHV-187, SHV-188, SHV-189, SHV-19, SHV-2, SHV-20, SHV-21, SHV-22, SHV-23, SHV-24, SHV-25, SHV-26, SHV-27, SHV-28, SHV-29, SHV-2A, SHV-3, SHV-30, SHV-31, SHV-32, SHV-33, SHV-34, SHV-35, SHV-36, SHV-37, SHV-38, SHV-39, SHV-40, SHV-41, SHV-42, SHV-43, SHV-44, SHV-45, SHV-46, SHV-48, SHV-49, SHV-5, SHV-50, SHV-51, SHV-52, SHV-53, SHV-55, SHV-56, SHV-57, SHV-59, SHV-6, SHV-60, SHV-61, SHV-62, SHV-63, SHV-64, SHV-65, SHV-66, SHV-67, SHV-69, SHV-7, SHV-70, SHV-71, SHV-72, SHV-73, SHV-74, SHV-75, SHV-76, SHV-77, SHV-78, SHV-79, SHV-8, SHV-80, SHV-81, SHV-82, SHV-83, SHV-84, SHV-85, SHV-86, SHV-89, SHV-9, SHV-92, SHV-93, SHV-94, SHV-95, SHV-96, SHV-97, SHV-98, SHV-99, SIM-1, SLB-1, SMB-1, SME-1, SME-2, SME-3, SME-4, SME-5, SPG-1, SPM-1, SRT-1, SRT-2, 스펙티노마이신에 대한 내성을 부여하는 살모넬라 엔테리카(*Salmonella enterica*) 16S rRNA (rrsD) 돌연변이, 살모넬라 엔테리카 cm1A, 플루오로퀴놀론에 대한 내성을 부여하는 살모넬라 엔테리카 gyrA, 트리클로산에 대한 내성을 부여하는 돌연변이를 갖는 살모넬라 엔테리카 gyrA, 플루오로퀴놀론에 대한 내성을 부여하는 살모넬라 엔테리카 parC, 살모넬라 엔테리카 ramR 돌연변이체, 항생제 내성을 부여하는 돌연변이를 갖는 살모넬라 엔테리카 soxR, 플루오로퀴놀론에 대한 내성을 부여하는 살모넬라 세로바르스(*Salmonella serovars*) gyrB, 플루오로퀴놀론에 대한 내성을 부여하는 살모넬라 세로바르스 parE, 항생제 내성을 부여하는 돌연변이를 갖는 살모넬라 세로바르스 soxS, Sed-1, 세라티아 마르세센스(*Serratia marcescens*) Omp1, 시겔라 플렉스네리(*Shigella flexneri*) 클로람페니콜 아세틸트랜스퍼라제, 플루오로퀴놀론에 대한 내성을 부여하는 시겔라 플렉스네리 gyrA, 플루오로퀴놀론에 대한 내성을 부여하는 시겔라 플렉스네리 parC, 리네졸리드에 대한 내성을 부여하는 돌연변이를 갖는 스타필로코쿠스 아우레우스(*Staphylococcus aureus*) 23S rRNA, 스타필로코쿠스 아우레우스 FosB, 포스

포마이신에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 G1pT, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 UhpT, 담토마이신에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 agrA, 담토마이신에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 cls, 푸시드산에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 fusA, 푸시드산에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 fusE, 플루오로퀴놀론에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 gyrA, 아미노쿠마린에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 gyrB, 뮤피로신에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 ileS, 리소신에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 menA, 스태필로코쿠스 아우레우스 mprF, 담토마이신에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 mprF, 포스포마이신에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 murA, 스태필로코쿠스 아우레우스 norA, 플루오로퀴놀론에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 parC, 아미노쿠마린에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 parE, 플루오로퀴놀론에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 parE, 담토마이신에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 pgsA 돌연변이, 담토마이신에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 rpoB 돌연변이체, 리팜피신에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 rpoB 돌연변이체, 담토마이신에 대한 내성을 부여하는 스태필로코쿠스 아우레우스 rpoC, 담토마이신에 대한 내성을 부여하는 돌연변이를 갖는 스태필로코쿠스 아우레우스 walK, 스태필로코쿠스 인테르메디우스(*Staphylococcus intermedius*) 클로람페니콜 아세틸트랜스퍼라제, 뮤피로신에 대한 내성을 부여하는 스태필로코쿠스 mupA, 뮤피로신에 대한 내성을 부여하는 스태필로코쿠스 mupB, 스태필로코쿠스 아우레우스 LmrS, 스트렙토코쿠스 아갈락티아에(*Streptococcus agalactiae*) mprF, 담토마이신 내성을 부여하는 돌연변이를 갖는 스트렙토코쿠스 미티스(*Streptococcus mitis*) CdsA, 마크롤리드 및 스트렙토그라민 항생제에 대한 내성을 부여하는 스트렙토코쿠스 뉴모니아에(*Streptococcus pneumoniae*) 23S rRNA 돌연변이, 마크롤리드 항생제에 대한 내성을 부여하는 돌연변이를 갖는 스트렙토코쿠스 뉴모니아에 23S rRNA, 아목시실린에 대한 내성을 부여하는 스트렙토코쿠스 뉴모니아에 PBP1a, 아목시실린에 대한 내성을 부여하는 스트렙토코쿠스 뉴모니아에 PBP2b, 아목시실린에 대한 내성을 부여하는 스트렙토코쿠스 뉴모니아에 PBP2x, 플루오로퀴놀론에 대한 내성을 부여하는 스트렙토코쿠스 뉴모니아에 parC, 술폰아미드에 대한 내성을 부여하는 돌연변이를 갖는 스트렙토코쿠스 피오게네스(*Streptococcus pyogenes*) folP, 스트렙토코쿠스 수이스(*Streptococcus suis*) 클로람페니콜 아세틸트랜스퍼라제, 마크롤리드 항생제에 대한 내성을 부여하는 돌연변이를 갖는 스트렙토미세스 암보파시엔스(*Streptomyces ambofaciens*) 23S rRNA, 엘파마이신에 대한 내성을 부여하는 스트렙토미세스 신나모네우스(*Streptomyces cinamomeus*) EF-Tu 돌연변이체, 스트렙토미세스 리비단스(*Streptomyces lividans*) cmlR, 아미노쿠마린에 대한 내성을 부여하는 스트렙토미세스 리시리덴시스(*Streptomyces rishiriensis*) parY 돌연변이체, TEM-1, TEM-10, TEM-101, TEM-102, TEM-104, TEM-105, TEM-106, TEM-107, TEM-108, TEM-109, TEM-11, TEM-110, TEM-111, TEM-112, TEM-113, TEM-114, TEM-115, TEM-116, TEM-117, TEM-118, TEM-12, TEM-120, TEM-121, TEM-122, TEM-123, TEM-124, TEM-125, TEM-126, TEM-127, TEM-128, TEM-129, TEM-130, TEM-131, TEM-132, TEM-133, TEM-134, TEM-135, TEM-136, TEM-137, TEM-138, TEM-139, TEM-141, TEM-142, TEM-143, TEM-144, TEM-145, TEM-146, TEM-147, TEM-148, TEM-149, TEM-15, TEM-150, TEM-151, TEM-152, TEM-153, TEM-154, TEM-155, TEM-156, TEM-157, TEM-158, TEM-159, TEM-16, TEM-160, TEM-162, TEM-163, TEM-164, TEM-166, TEM-167, TEM-168, TEM-169, TEM-17, TEM-171, TEM-176, TEM-177, TEM-178, TEM-182, TEM-183, TEM-184, TEM-185, TEM-186, TEM-187, TEM-188, TEM-189, TEM-19, TEM-190, TEM-191, TEM-192, TEM-193, TEM-194, TEM-195, TEM-196, TEM-197, TEM-198, TEM-199, TEM-2, TEM-20, TEM-201, TEM-205, TEM-206, TEM-207, TEM-208, TEM-209, TEM-21, TEM-211, TEM-213, TEM-214, TEM-215, TEM-216, TEM-217, TEM-219, TEM-22, TEM-220, TEM-24, TEM-26, TEM-28, TEM-29, TEM-3, TEM-30, TEM-33, TEM-34, TEM-4, TEM-40, TEM-42, TEM-43, TEM-45, TEM-47, TEM-48, TEM-49, TEM-52, TEM-53, TEM-54, TEM-55, TEM-57, TEM-59, TEM-6, TEM-60, TEM-63, TEM-67, TEM-68, TEM-7, TEM-70, TEM-71, TEM-72, TEM-73, TEM-75, TEM-76, TEM-78, TEM-79, TEM-8, TEM-80, TEM-81, TEM-82, TEM-83, TEM-84, TEM-85, TEM-86, TEM-87, TEM-88, TEM-89, TEM-90, TEM-91, TEM-92, TEM-93, TEM-94, TEM-95, TEM-96, THIN-B, TLA-1, TLA-2, TLA-3, TMB-1, TMB-2, TRU-1, TUS-1, TaeA, Tet(47), Tet(X3), Tet(X4), TolC, TriA, TriB, TriC, Type A NfxB, Type B NfxB, 플루오로퀴놀론에 대한 내성을 부여하는 우레아플라스마 우레아리티쿰(*Ureaplasma urealyticum*) gyrB, 플루오로퀴놀론에 대한 내성을 부여하는 우레아플라스마 우레아리티쿰 parC, VCC-1, VEB-1, VEB-1b, VEB-2, VEB-3, VEB-4, VEB-5, VEB-6, VEB-7, VEB-8, VEB-9, VIM-1, VIM-10, VIM-11, VIM-12, VIM-13, VIM-14, VIM-15, VIM-16, VIM-17, VIM-18, VIM-19, VIM-2, VIM-20, VIM-23, VIM-24, VIM-25, VIM-26, VIM-27, VIM-28, VIM-29, VIM-3, VIM-30, VIM-31, VIM-32, VIM-33, VIM-34, VIM-35, VIM-36, VIM-37, VIM-38, VIM-39, VIM-4, VIM-42, VIM-43, VIM-5, VIM-6, VIM-7, VIM-8, VIM-9, VatI, 비브리오 안구일라룸(*Vibrio anguillarum*) 클로람페니콜 아세틸트랜스퍼라제, 비브리오

콜레라에(*Vibrio cholerae*) OmpT, 비브리오 콜레라에 OmpU, 비브리오 콜레라에 varG, YojI, aacA43, aad(6), aadA, aadA10, aadA11, aadA12, aadA13, aadA14, aadA15, aadA16, aadA17, aadA2, aadA21, aadA22, aadA23, aadA24, aadA25, aadA27, aadA3, aadA4, aadA5, aadA6, aadA6/aadA10, aadA7, aadA8, aadA8b, aadA9, aadK, aadS, abcA, abeM, abeS, acrB, acrD, adeA, adeB, adeC, adeF, adeG, adeH, adeI, adeJ, adeK, adeL, adeN, adeR, adeS, almG, ampS, amrA, amrB, aphA15, apmA, arlR, arlS, armA, arnA, arr-1, arr-2, arr-3, arr-4, arr-5, arr-7, arr-8, bacA, baeR, baeS, basR, basS, bcr-1, bcrA, bcrB, bcrC, blaF, blaI, blaR1, blt, bmr, carA, carO, catA4, catA8, catB10, catB11, catB2, catB3, catB8, catB9, catI, catII, 에스케리키아 콜라이 K-12로부터의 catII, catIII, catP, catQ, catS, catV, cdeA, ceoA, ceoB, cepA, cfr(B), cfrA, cfrC, chrB, cipA, clbA, clbB, clbC, clcD, cmeA, cmeB, cmeC, cmeR, cmlA1, cmlA4, cmlA5, cmlA6, cmlA8, cmlB, cmlB1, cmlv, cmrA, cmx, cpaA, cphA2, cphA3, cphA4, cphA5, cphA6, cphA7, cphA8, cpxA, dfrA1, dfrA10, dfrA12, dfrA13, dfrA14, dfrA15, dfrA15b, dfrA16, dfrA17, dfrA18, dfrA19, dfrA20, dfrA21, dfrA22, dfrA23, dfrA24, dfrA25, dfrA26, dfrA27, dfrA28, dfrA29, dfrA3, dfrA30, dfrA32, dfrA3b, dfrA5, dfrA6, 프로테우스 미라빌리스(*Proteus mirabilis*)로부터의 dfrA6, dfrA7, dfrA8, dfrA9, dfrB1, dfrB2, dfrB3, dfrB4, dfrB5, dfrB6, dfrB7, dfrC, dfrD, dfrE, dfrF, dfrG, dfrI, dfrK, eatAv, efmA, efpA, efrA, efrB, emeA, emrA, emrB, emrD, emrK, emrR, emrY, emtA, eptA, erm(32), erm(40), erm(45), erm(46), ermZ, evgA, evgS, facT, farA, farB, fexA, floR, fusB, fusC, fusD, fusH, gadW, gadX, gimA, golS, hmrM, hpl181, hpl184, imiH, imiS, iri, kamB, kdpE, lfrA, lin, linG, lmrA, lmrB, lmrC, lmrD, lmrP, lnuA, lnuB, lnuC, lnuD, lnuE, lnuF, lnuG, lsaA, lsaB, lsaC, lsaE, macA, macB, marA, mdsA, mdsB, mdsC, mdtA, mdtB, mdtC, mdtE, mdtF, mdtG, mdtH, mdtM, mdtN, mdtO, mdtP, mecA, mecB, mecC, mecD, mecI, mecR1, mef(B), mefC, mefE, mel, mepA, mepR, mexM, mexN, mexP, mexQ, mexX, mexY, mfd, mfpA, mgrA, mgrB, mgtA, mphA, mphB, mphC, mphE, mphF, mphG, mphH, mphI, mphJ, mphK, mphL, mphM, mphN, mphO, msbA, msrA, msrB, msrC, msrE, mtrA, mtrC, mtrD, mtrE, mtrR, myrA, nalC, nalD, norA, norB, novA, npmA, oleB, oleC, oleD, oleI, opcM, opmE, optrA, oqxA, oqxB, otr(A), otr(B), otrC, patA, patB, pexA, pgpB, 플라스미드-코딩된 cat (pp-cat), pmrA, 포린 OmpC, poxtA, pp-flo, qacA, qacB, qacH, qnrE1, qnrE2, ramA, rgt1438, rmtA, rmtB, rmtC, rmtD, rmtD2, rmtE, rmtE2, rmtF, rmtG, rmtH, rosA, rosB, rphA, rphB, rpoB2, rpsJ, salA, sav1866, sdiA, sgm, smeA, smeB, smeC, smeD, smeE, smeF, smeR, smeS, spd, srmB, sta, sul1, sul2, sul3, sul4, tap, tcmA, tcr3, tet(30), tet(31), tet(33), tet(35), tet(38), tet(39), tet(40), tet(41), tet(42), tet(43), tet(44), tet(45), tet(48), tet(49), tet(50), tet(51), tet(52), tet(53), tet(54), tet(55), tet(56), tet(59), tet(A), tet(B), tet(C), tet(D), tet(E), tet(G), tet(H), tet(J), tet(K), tet(L), tet(V), tet(W/N/W), tet(Y), tet(Z), tet32, tet34, tet36, tet37, tetA(46), tetA(58), tetA(60), tetA(P), tetB(46), tetB(58), tetB(60), tetB(P), tetM, tetO, tetQ, tetR, tetR(G), tetS, tetT, tetU, tetW, tetX, 티로신 내성을 부여하는 tlrB, tlrC, tmrB, tsnR, tva(A), ugd, vanA, vanB, vanC, vanD, vanE, vanF, vanG, vanHA, vanHB, vanHD, vanHF, vanHM, vanHO, vanI, vanJ, vanKI, vanL, vanM, vanN, vanO, vanRA, vanRB, vanRC, vanRD, vanRE, vanRF, vanRG, vanRI, vanRL, vanRM, vanRN, vanRO, vanSA, vanSB, vanSC, vanSD, vanSE, vanSF, vanSG, vanSI, vanSL, vanSM, vanSN, vanSO, vanTC, vanTE, vanTG, vanTN, vanTmL, vanTrL, vanUG, vanVB, vanWB, vanWG, vanWI, vanXA, vanXB, vanXD, vanXF, vanXI, vanXM, vanXO, vanXYC, vanXYE, vanXYG, vanXYL, vanXYN, vanYA, vanYB, vanYD, vanYF, vanYG1, vanYM, vanZA, vanZF, vatA, vatB, vatC, vatD, vatE, vatF, vath, vga(E) 스태필로코쿠스 코흐니이(*Staphylococcus cohnii*), vgaA, vgaALC, vgaB, vgaC, vgaD, vgaE, vgbA, vgbB, vgbC, vmlR, vph, y56 베타-락타마제, ykkC, ykkD, 또는 그의 임의의 조합에 결합하는 (또는 그에 의해 결합되는) 풍부화 분자를 포함할 수 있다.

[0242] 일부 실시양태에서, 키트 내의 적어도 1종의 성분은 건조되거나 동결건조된 형태로 제공된다. 다른 실시양태에서, 키트의 적어도 1종의 성분은 가용화된 형태로 제공된다.

[0243] 본원에 제공된 키트는 적합한 포장 내에 있다. 적합한 포장은 바이알, 병, 단지, 가요성 포장 등을 포함하나, 이에 제한되지는 않는다. 특정 장치와 조합하여 사용하기 위한 패키지가 또한 고려된다. "샘플 제조 및 샘플 서열분석을 위한 장치"를 참조한다. 키트는 멸균 접근 포트를 가질 수 있다 (예를 들어, 용기는 피하 주사 바늘로 뚫을 수 있는 마개를 갖는 정맥내 용액 백 또는 바이알일 수 있음). 용기는 또한 멸균 접근 포트를 가질 수 있다.

[0244] 키트는 임의로 추가의 성분, 예컨대 완충제 및 설명 정보를 제공할 수 있다. 일부 실시양태에서, 키트는 적어도 1종의 완충제를 추가로 포함한다. 본원에 기재된 방법에 적합한 완충제는 이전에 기재되었다. 일부 실시양

태에서, 키트는 본원에 기재된 방법 중 임의의 것에서의 사용에 대한 지침서를 추가적으로 포함할 수 있다.

- [0245] 일부 실시양태에서, 본 개시내용은 상기 기재된 키트의 내용물을 포함하는 제조 물품을 제공한다.
- [0246] V. 샘플 제조 및 샘플 서열분석을 위한 장치
- [0247] 일부 측면에서, 본 개시내용은 샘플 제조 및/또는 샘플 서열분석을 위한 장치에 관한 것이다. 일부 실시양태에서, 장치는 샘플 제조 모듈을 포함한다. 일부 실시양태에서, 장치는 샘플 서열분석 모듈을 포함한다. 일부 실시양태에서, 본 장치는 샘플 제조 모듈 및 샘플 서열분석 모듈을 포함한다.
- [0248] A. 샘플 제조를 위한 장치
- [0249] 분석을 위한 샘플을 제조하는 과정에 사용하기 위한 기구, 카트리지 (예를 들어, 채널 (예를 들어, 마이크로유체 채널)을 포함함), 및/또는 펌프 (예를 들어, 연동 펌프)를 포함하는 장치가 일반적으로 제공된다. 본 개시내용에 따른 장치를 사용하여 생물학적 샘플로부터 표적 분자의 풍부화, 농축, 조작, 및/또는 검출을 가능하게 할 수 있다. 일부 실시양태에서, 차세대 서열분석 및/또는 다른 하류 분석 기술을 위한 물질을 생산하기 위한 샘플의 자동화 프로세싱을 위한 장치 및 관련 방법이 제공된다. 장치 및 관련 방법은 본원의 다른 곳에 기재된 샘플 제조 또는 샘플 분석 과정에 따른 핵산 및/또는 폴리펩티드 프로세싱을 위한 반응을 포함하여, 화학적 및/또는 생물학적 반응을 수행하는데 사용될 수 있다.
- [0250] 일부 실시양태에서, 샘플 제조 장치는 표적 분자 또는 복수의 분자 (예를 들어, 표적 핵산 또는 표적 폴리펩티드)를 포함하는 샘플을 서열분석 모듈 또는 장치에 전달하거나 옮기도록 위치된다. 일부 실시양태에서, 샘플 제조 장치는 서열분석 장치에 직접 연결되거나 (예를 들어, 물리적으로 부착되거나) 또는 간접 연결된다.
- [0251] 일부 실시양태에서, 장치는 1개 이상의 카트리지를 수용하도록 구성된 서열 제조 모듈을 포함한다. 일부 실시양태에서, 카트리지는 유체를 수용하고/거나 샘플 제조 과정에 사용되는 1종 이상의 시약을 함유하도록 구성된 1개 이상의 저장소 또는 반응 용기를 포함한다. 일부 실시양태에서, 카트리지는 샘플 제조 과정에 사용되는 유체 (예를 들어, 1종 이상의 시약을 포함하는 유체)를 함유 및/또는 수송하도록 구성된 1개 이상의 채널 (예를 들어, 마이크로유체 채널)을 포함한다. 시약은 완충제, 효소적 시약, 중합체 매트릭스, 풍부화 분자, 포획 시약, 크기-특이적 선택 시약, 서열-특이적 선택 시약, 및/또는 정제 시약을 포함한다. 샘플 제조 과정에 사용하기 위한 추가의 시약은 본원의 다른 곳에 기재되어 있다.
- [0252] 일부 실시양태에서, 카트리지는 1종 이상의 저장된 시약 (예를 들어, 액체 형태로의 재구성에 적합한 액체 또는 동결건조된 형태)을 포함한다. 카트리지의 저장된 시약은 목적하는 과정을 수행하는데 적합한 시약 및/또는 목적하는 샘플 유형을 프로세싱하는데 적합한 시약을 포함한다. 일부 실시양태에서, 카트리지는 단일-사용 카트리지 (예를 들어, 일회용 카트리지) 또는 다중-사용 카트리지 (예를 들어, 재사용가능한 카트리지)이다. 일부 실시양태에서, 카트리지는 사용자-공급 샘플을 수용하도록 구성된다. 사용자-공급 샘플은 카트리지가 장치에 수용되기 전 또는 후에, 예를 들어 사용자에게 의해 수동으로 또는 자동화 과정으로 카트리지에 부가될 수 있다.
- [0253] 일부 실시양태에서, 장치는 본 개시내용에 따른 과정에서 표적 분자의 풍부화를 용이하게 할 수 있다. "폴리펩티드 풍부화 방법"을 참조한다. 이러한 방식으로, 장치는 고도로 멀티플렉스화된 방식으로 관심 폴리펩티드를 풍부화하는 분자의 활용을 가능하게 한다.
- [0254] 일부 실시양태에서, 샘플은 전기영동 방법을 사용하여 표적 분자에 대해 풍부화된다. 일부 실시양태에서, 샘플은 친화도 SCODA를 사용하여 표적 분자에 대해 풍부화된다. 일부 실시양태에서, 샘플은 전계 역전 겔 전기영동 (FIGE)을 사용하여 표적 분자에 대해 풍부화된다. 일부 실시양태에서, 샘플은 펄스 전계 겔 전기영동 (PFGE)을 사용하여 표적 분자에 대해 풍부화된다.
- [0255] 일부 실시양태에서, 장치는 샘플에 존재하는 표적 분자에 (직접 또는 간접) 결합하는 고정된 포획 프로브를 포함하는, 풍부화 동안 사용되는 매트릭스 (예를 들어, 다공성 매질, 전기영동 중합체 겔)를 포함하는 샘플 제조 모듈을 포함한다. 일부 실시양태에서, 풍부화 동안 사용되는 매트릭스는 1, 2, 3, 4, 5종, 또는 그 초과와 고유한 고정된 포획 프로브를 포함하며, 이들 각각은 고유한 표적 분자에 결합하고/거나 상이한 결합 친화도로 동일한 표적 분자에 결합한다.
- [0256] 일부 실시양태에서, 고정된 포획 프로브는 표적 폴리펩티드 또는 폴리펩티드 단편에 결합하는 폴리펩티드 포획 프로브이다. 예를 들어, 일부 실시양태에서, 고정된 포획 프로브는 본원에 기재된 바와 같은 풍부화 분자이다.
- [0257] 일부 실시양태에서, 폴리펩티드 포획 프로브는 표적 폴리펩티드 (또는 폴리펩티드 단편)에 10^{-9} 내지 10^{-8} M, 10^{-}

⁸ 내지 10^{-7} M, 10^{-7} 내지 10^{-6} M, 10^{-6} 내지 10^{-5} M, 10^{-5} 내지 10^{-4} M, 10^{-4} 내지 10^{-3} M, 또는 10^{-3} 내지 10^{-2} M의 결합 친화도로 결합한다. 일부 실시양태에서, 결합 친화도는 피코몰 내지 나노몰 범위 (예를 들어, 약 10^{-12} 내지 약 10^{-9} M)이다. 일부 실시양태에서, 결합 친화도는 나노몰 내지 마이크로몰 범위 (예를 들어, 약 10^{-9} 내지 약 10^{-6} M)이다. 일부 실시양태에서, 결합 친화도는 마이크로몰 내지 밀리몰 범위 (예를 들어, 약 10^{-6} 내지 약 10^{-3} M)이다. 일부 실시양태에서, 결합 친화도는 피코몰 내지 마이크로몰 범위 (예를 들어, 약 10^{-12} 내지 약 10^{-6} M)이다. 일부 실시양태에서, 결합 친화도는 나노몰 내지 밀리몰 범위 (예를 들어, 약 10^{-9} 내지 약 10^{-3} M)이다.

[0258] 일부 실시양태에서, 고정된 포획 프로브는 표적 핵산에 혼성화하는 올리고뉴클레오타이드 포획 프로브이다. 일부 실시양태에서, 올리고뉴클레오타이드 포획 프로브는 표적 핵산에 대해 적어도 50%, 60%, 70%, 80%, 90% 95%, 또는 100% 상보적이다. 일부 실시양태에서, 단일 올리고뉴클레오타이드 포획 프로브는 적어도 50%, 60%, 70%, 80%, 90% 95%, 또는 99% 서열 동일성을 공유하는 복수의 관련 표적 핵산 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50종, 또는 그 초과)의 관련 표적 핵산 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50종, 또는 그 초과)의 관련 표적 핵산 (예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50종, 또는 그 초과)을 풍부화하는데 사용될 수 있다. 복수의 관련 표적 핵산의 풍부화는 메타게놈 라이브러리의 생성을 가능하게 할 수 있다. 일부 실시양태에서, 올리고뉴클레오타이드 포획 프로브는 관련 표적 핵산의 차등 풍부화를 가능하게 할 수 있다. 일부 실시양태에서, 올리고뉴클레오타이드 포획 프로브는 그의 변형 상태 (예를 들어, 메틸화 상태, 아세틸화 상태)가 상이한 동일한 서열의 핵산에 비해 표적 핵산의 풍부화를 가능하게 할 수 있다.

[0259] 일부 실시양태에서, 핵산 표적 분자를 0.5-2 킬로염기의 길이로 풍부화하기 위한 목적으로, 올리고뉴클레오타이드 포획 프로브는 5' 아크리다이트 모이어티를 사용하여 아크릴아미드 매트릭스에 공유 고정될 수 있다. 일부 실시양태에서, 보다 큰 핵산 표적 분자 (예를 들어, >2 킬로염기의 길이를 가짐)를 풍부화하기 위한 목적으로, 올리고뉴클레오타이드 포획 프로브는 아가로스 매트릭스에 고정될 수 있다. 일부 실시양태에서, 올리고뉴클레오타이드 포획 프로브는 티올-에폭시드 화학을 사용하여 (예를 들어, 가교된 아가로스 비드에 공유 부착된 티올-변형된 올리고뉴클레오타이드에 의해) 아가로스 매트릭스에 고정될 수 있다. 아가로스 비드에 연결된 올리고뉴클레오타이드 포획 프로브는 표준 아가로스 매트릭스 내에서 (예를 들어, 동일한 아가로스 백분율로) 조합 및 고체화될 수 있다.

[0260] 일부 실시양태에서, 다중 포획 프로브 (예를 들어, 감염원, 예컨대 아데노바이러스, 스타필로코쿠스, 뉴모니아에, 또는 투베르쿨로시스의 결정적 표적 분자에 결합하는, 예를 들어 다중 포획 프로브 유형의 집단)는 풍부화 매트릭스에 고정될 수 있다. 다중 결정적 포획 프로브가 존재하는 풍부화 매트릭스에의 샘플의 적용으로 질환 또는 상태 (예를 들어, 감염원의 존재)의 진단이 이루어질 수 있다.

[0261] 일부 실시양태에서, 장치는 본 개시내용에 따른 과정에서 비-표적 분자의 제거 후에 풍부화 매트릭스로부터 표적 분자의 방출을 용이하게 할 수 있다. 일부 실시양태에서, 표적 분자는 풍부화 매트릭스의 온도를 증가시킴으로써 풍부화 매트릭스로부터 방출될 수 있다. 매트릭스의 온도를 조정하는 것은 증가된 온도가 보다 높은 포획 프로브 엄격도를 제공하여 표적 분자와 포획 프로브 사이에 보다 큰 결합 친화도를 요구하기 때문에 이동 속도에 추가로 영향을 미친다. 일부 실시양태에서, 관련 표적 분자를 풍부화하는 경우, 계속 증가하는 상동성 단계로 표적 분자를 방출 및 단리하기 위해 매트릭스 온도는 단계적 방식으로 점차적으로 증가될 수 있다. 이는 초기 참조 표적 분자와 관련하여 점점 더 먼 표적 폴리펩티드 또는 표적 핵산의 서열분석을 가능하게 하여, 신규 단백질 (예를 들어, 효소) 또는 기능 (예를 들어, 효소적 기능 또는 유전자 기능)의 발견을 가능하게 할 수 있다. 일부 실시양태에서, 다중 포획 프로브 (예를 들어, 다중 결정적 포획 프로브)를 사용하는 경우에, 매트릭스 온도는 단계적 또는 구배 방식으로 증가될 수 있으며, 이는 상이한 표적 분자의 온도-의존성 방출을 허용하고, 대조군 및 표적 분자의 존재 또는 부재를 나타내는 일련의 바코드부착된 방출 밴드의 생성을 발생시킬 수 있다.

[0262] 본 개시내용에 따른 장치는 일반적으로 본원에 기재된 바와 같은 카트리지를 작동시키는데 사용될 수 있는 기계적 및 전자적 및/또는 광학적 구성요소를 함유한다. 일부 실시양태에서, 장치 구성요소는 카트리지 상에서 또는 카트리지의 특정 영역 상에서 특정 온도를 달성하고 유지하도록 작동한다. 일부 실시양태에서, 장치 구성요소는 카트리지의 전극에 특정 시간 지속기간 동안 특정 전압을 인가하도록 작동한다. 일부 실시양태에서, 장치 구성요소는 액체를 카트리지의 저장소 및/또는 반응 용기로, 그로부터, 또는 그 사이에서 이동시키도록 작동한다. 일부 실시양태에서, 장치 구성요소는 액체를 카트리지의 채널(들)을 통해, 예를 들어 카트리지의 저장소 및/또는 반응 용기로, 그로부터, 또는 그 사이로 이동시키도록 작동한다. 일부 실시양태에서, 장치 구성요소는

카트리지의 엘라스토머, 시약-특이적 저장소 또는 반응 용기와 상호작용하는 연동 펌핑 메카니즘 (예를 들어, 기구)을 통해 액체를 이동시킨다. 일부 실시양태에서, 장치 구성요소는 채널을 통해 유체를 펌핑하기 위해 카트리지의 채널과 회합된 엘라스토머 구성요소 (예를 들어, 엘라스토머를 포함하는 표면 층)와 상호작용하도록 구성된 연동 펌핑 메카니즘 (예를 들어, 기구)을 통해 액체를 이동시킨다. 장치 구성요소는, 예를 들어 샘플 정보가 입력될 수 있는 사용자 인터페이스를 구동하기 위한 컴퓨터 자원을 포함할 수 있고, 특정 과정이 선택될 수 있고, 실행 결과가 보고될 수 있다.

[0263] 하기 비제한적 예는 본원에 기재된 장치, 방법, 및 조성물의 측면을 예시하기 위해 의도된다. 본 개시내용에 따른 샘플 제조 장치의 사용은 하기 기재된 단계 중 1개 이상으로 진행될 수 있다. 사용자는 장치의 뚜껑을 열고, 목적하는 과정을 지지하는 카트리지를 삽입할 수 있다. 이어서, 사용자는 특정 용해 용액과 조합될 수 있는 샘플을 카트리지 상의 샘플 포트에 첨가할 수 있다. 이어서, 사용자는 장치 뚜껑을 닫고, 장치 상의 터치 스크린 인터페이스를 통해 임의의 샘플 특이적 정보를 입력하고, 임의의 과정 특이적 파라미터 (예를 들어, 목적하는 크기 선택의 범위, 표적 분자 포획을 위한 상동성의 목적하는 정도 등)를 선택하고, 샘플 제조 과정 실행을 개시할 수 있다.

[0264] 실행 후에, 사용자는 관련 실행 데이터 (예를 들어, 실행의 성공적인 완료의 확정, 실행 특이적 측정 등)뿐만 아니라, 과정 특이적 정보 (예를 들어, 생성된 샘플의 양, 특이적 표적 서열의 존재 또는 부재 등)를 수신할 수 있다. 실행에 의해 생성된 데이터는 로컬 또는 클라우드 기반일 수 있는 후속 생물정보학 분석에 적용될 수 있다. 과정에 따라, 완료된 샘플은 후속 사용 (예를 들어, 게놈 서열분석, qPCR 정량화, 클로닝 등)을 위해 카트리지에서 추출될 수 있다. 이어서 장치를 열 수 있고, 이어서 카트리지를 빼낼 수 있다.

[0265] 도 8은 풍부화를 수행하기 위한 예시적인 기구를 도시한 예시를 제공한다. 예를 들어, 미국 특허 번호 8608929를 참조하고, 이는 그 전문이 본원에 참조로 포함된다.

[0266] B. 서열분석을 위한 장치

[0267] 폴리펩티드를 포함하는 샘플 (예를 들어, 풍부화된 샘플)을 서열분석하는 과정에 사용하기 위한 기구, 카트리지 (예를 들어, 채널 (예를 들어, 마이크로유체 채널)을 포함함), 및/또는 펌프 (예를 들어, 연동 펌프)를 포함하는 장치가 또한 일반적으로 제공된다. 일부 측면에서, 본 개시내용에 따른 핵산 또는 폴리펩티드의 서열분석은 단일 분자 분석 및/또는 단일 분자의 서열분석을 허용하는 시스템을 사용하여 병행 수행될 수 있다. 시스템은 서열분석 장치 및 서열분석 장치와 인터페이스되도록 구성된 기구를 포함할 수 있다.

[0268] 서열분석 장치는 픽셀의 어레이를 포함하는 서열분석 모듈을 포함할 수 있으며, 여기서 개별 픽셀은 샘플 웰 및 적어도 1개의 광검출기를 포함한다. 서열분석 장치의 샘플 웰은 서열분석 장치의 표면 상에 또는 표면을 통해 형성될 수 있고, 서열분석 장치의 표면 상에 위치된 샘플을 수용하도록 구성될 수 있다. 일부 실시양태에서, 샘플 웰은 장치 내로 삽입될 수 있는 카트리지 (예를 들어, 일회용 또는 단일-사용 카트리지)의 구성요소이다. 집합적으로, 샘플 웰은 샘플 웰의 어레이로서 간주될 수 있다. 복수의 샘플 웰은 샘플 웰의 적어도 일부가 단일 표적 분자 또는 복수의 분자 (예를 들어, 표적 핵산 또는 표적 폴리펩티드)를 포함하는 샘플을 수용하도록 적합한 크기 및 형상을 가질 수 있다. 일부 실시양태에서, 샘플 웰 내의 분자의 수는 일부 샘플 웰이 1종의 분자 (예를 들어, 표적 핵산 또는 표적 폴리펩티드)를 함유하고 다른 것은 0, 2종, 또는 복수의 분자를 함유하도록 서열분석 장치의 샘플 웰 사이에서 분포될 수 있다.

[0269] 일부 실시양태에서, 서열분석 장치는 샘플 제조 장치로부터 복수의 분자 (예를 들어, 1종 이상의 관심 폴리펩티드)를 포함하는 샘플을 수용하도록 위치된다. 일부 실시양태에서, 서열분석 장치는 샘플 제조 장치에 직접 연결되거나 (예를 들어, 물리적으로 부착되거나) 또는 간접 연결된다.

[0270] 서열분석 장치는 픽셀의 어레이를 포함할 수 있으며, 여기서 개별 픽셀은 샘플 웰 및 적어도 1개의 광검출기를 포함한다. 서열분석 장치의 샘플 웰은 서열분석 장치의 표면 상에 또는 표면을 통해 형성될 수 있고, 서열분석 장치의 표면 상에 위치된 샘플을 수용하도록 구성될 수 있다. 집합적으로, 샘플 웰은 샘플 웰의 어레이로서 간주될 수 있다. 복수의 샘플 웰은 샘플 웰의 적어도 일부가 단일 샘플 (예를 들어, 단일 분자, 예컨대 폴리펩티드)을 수용하도록 적합한 크기 및 형상을 가질 수 있다. 일부 실시양태에서, 샘플 웰 내의 샘플의 수는 일부 샘플 웰이 1종의 샘플을 함유하고 다른 것은 0, 2종 또는 그 초과 샘플을 함유하도록 서열분석 장치의 샘플 웰 사이에서 분포될 수 있다.

[0271] 여기 광은 서열분석 장치의 외부 또는 내부일 수 있는 1개 이상의 광원으로부터 서열분석 장치에 제공된다. 서열분석 장치의 광학 구성요소는 광원으로부터 여기 광을 수용하고, 광을 서열분석 장치의 샘플 웰의 어레이를

향하게 하고, 샘플 웰 내의 조명 영역을 조명할 수 있다. 일부 실시양태에서, 샘플 웰은 샘플이 샘플 웰의 표면에 근접하여 보유되도록 하는 구성을 가질 수 있으며, 이는 샘플로의 여기 광의 전달 및 샘플로부터의 방출 광의 검출을 용이하게 할 수 있다. 조명 영역 내에 위치한 샘플은 여기 광에 의해 조명되는 것에 반응하여 방출 광을 방출할 수 있다. 예를 들어, 샘플은 형광 마커로 표지될 수 있으며, 이는 여기 광의 조명을 통해 여기 상태를 달성하는 것에 반응하여 광을 방출한다. 이어서, 샘플에 의해 방출된 방출 광은 분석되는 샘플이 존재하는 샘플 웰에 상응하는 픽셀 내의 1개 이상의 광검출기에 의해 검출될 수 있다. 일부 실시양태에 따라 대략 10,000 픽셀 내지 1,000,000 픽셀의 수의 범위일 수 있는 샘플 웰의 어레이에 걸쳐 수행될 때, 다중 샘플이 병행 분석될 수 있다.

[0272] 서열분석 장치는 여기 광을 수용하고 여기 광을 샘플 웰 어레이 사이로 향하게 하기 위한 광학 시스템을 포함할 수 있다. 광학 시스템은 여기 광을 서열분석 장치에 커플링시키고 여기 광을 다른 광학 구성요소로 향하게 하도록 구성된 1개 이상의 격자 커플러를 포함할 수 있다. 광학 시스템은 여기 광을 격자 커플러로부터 샘플 웰 어레이로 향하게 하는 광학 구성요소를 포함할 수 있다. 이러한 광학 구성요소는 광학 분할기, 광학 조합기, 및 도파관을 포함할 수 있다. 일부 실시양태에서, 1개 이상의 광학 분할기는 격자 커플러로부터의 여기 광을 커플링시키고 여기 광을 도파관 중 적어도 1개로 전달할 수 있다. 일부 실시양태에 따르면, 광학 분할기는 각각의 도파관이 실질적으로 유사한 양의 여기 광을 수용하도록 모든 도파관에 걸친 여기 광의 실질적으로 균일한 전달을 가능하게 하는 구성을 가질 수 있다. 이러한 실시양태는 서열분석 장치의 샘플 웰에 의해 수용되는 여기 광의 균일성을 개선시킴으로써 서열분석 장치의 성능을 개선시킬 수 있다. 서열분석 장치에 포함시키기 위한, 예를 들어 여기 광을 샘플 웰에 커플링시키고/거나 방출 광을 광검출기로 향하게 하는데 적합한 구성요소의 예는 2015년 8월 7일에 출원된 발명의 명칭 "INTEGRATED DEVICE FOR PROBING, DETECTING AND ANALYZING MOLECULES"의 미국 특허 출원 번호 14/821,688 및 2014년 11월 17일에 출원된 발명의 명칭 "INTEGRATED DEVICE WITH EXTERNAL LIGHT SOURCE FOR PROBING, DETECTING, AND ANALYZING MOLECULES"의 미국 특허 출원 번호 14/543,865에 기재되어 있으며, 이들 둘 다는 그 전문이 참조로 포함된다. 서열분석 장치에서 구현될 수 있는 적합한 격자 커플러 및 도파관의 예는 2017년 12월 15일에 출원된 발명의 명칭 "OPTICAL COUPLER AND WAVEGUIDE SYSTEM"의 미국 특허 출원 번호 15/844,403에 기재되어 있으며, 이는 그 전문이 참조로 포함된다.

[0273] 추가의 광자 구조가 샘플 웰과 광검출기 사이에 위치될 수 있고, 여기 광이 광검출기에 도달하는 것을 감소시키거나 막도록 구성될 수 있으며, 이는 그렇지 않으면 방출 광을 검출하는데 있어서 신호 노이즈에 기여할 수 있다. 일부 실시양태에서, 서열분석 장치를 위한 회로로서 작용할 수 있는 금속 층은 또한 공간 필터로서 작용할 수 있다. 적합한 광자 구조의 예는 스펙트럼 필터, 편광 필터, 및 공간 필터를 포함할 수 있고, 2018년 7월 23일에 출원된 발명의 명칭 "OPTICAL REJECTION PHOTONIC STRUCTURES"의 미국 특허 출원 번호 16/042,968에 기재되어 있으며, 이는 그 전문이 참조로 포함된다.

[0274] 서열분석 장치에서 떨어져서 위치하는 구성요소는 여기원을 서열분석 장치에 위치시키고 정렬하는데 사용될 수 있다. 이러한 구성요소는 렌즈, 거울, 프리즘, 윈도우, 개구, 감쇠기, 및/또는 광섬유를 포함한 광학 구성요소를 포함할 수 있다. 1개 이상의 정렬 구성요소의 제어를 가능하게 하기 위해 추가의 기계적 구성요소가 기기에 포함될 수 있다. 이러한 기계적 구성요소는 작동기, 스테퍼 모터, 및/또는 노브를 포함할 수 있다. 적합한 여기원 및 정렬 메카니즘의 예는 2016년 5월 20일에 출원된 발명의 명칭 "PULSED LASER AND SYSTEM"의 미국 특허 출원 번호 15/161,088에 기재되어 있으며, 이는 그 전문이 참조로 포함된다. 빔-스티어링 모듈의 또 다른 예는 2017년 12월 14일에 출원된 발명의 명칭 "COMPACT BEAM SHAPING AND STEERING ASSEMBLY"의 미국 특허 출원 번호 15/842,720에 기재되어 있으며, 이는 본원에 참조로 포함된다. 적합한 여기원의 추가의 예는 2015년 8월 7일에 출원된 발명의 명칭 "INTEGRATED DEVICE FOR PROBING, DETECTING AND ANALYZING MOLECULES"의 미국 특허 출원 번호 14/821,688에 기재되어 있으며, 이는 그 전문이 참조로 포함된다.

[0275] 서열분석 장치의 개별 픽셀과 함께 위치된 광검출기(들)는 픽셀의 상응하는 샘플 웰로부터 방출 광을 검출하도록 구성되고 위치될 수 있다. 적합한 광검출기의 예는 2015년 8월 7일에 출원된 발명의 명칭 "INTEGRATED DEVICE FOR TEMPORAL BINNING OF RECEIVED PHOTONS"의 미국 특허 출원 번호 14/821,656에 기재되어 있으며, 이는 그 전문이 참조로 포함된다. 일부 실시양태에서, 샘플 웰 및 그의 각각의 광검출기(들)는 공통 축을 따라 정렬될 수 있다. 이러한 방식으로, 광검출기(들)는 픽셀 내의 샘플 웰과 중첩될 수 있다.

[0276] 검출된 방출 광의 특징은 방출 광과 연관된 마커를 확인하기 위한 표시를 제공할 수 있다. 이러한 특징은 광검출기에 의해 검출된 광자의 도달 시간, 광검출기에 의해 시간 경과에 따라 축적된 광자의 양, 및/또는 2개 이상의 광검출기에 걸친 광자의 분포를 비롯한 임의의 적합한 유형의 특징을 포함할 수 있다. 일부 실시양태에서, 광검출기는 샘플의 방출 광과 연관된 1종 이상의 타이밍 특징 (예를 들어, 발광 수명)의 검출을 가능하게 하는

구성을 가질 수 있다. 광검출기는 여기 광의 펄스가 서열분석 장치를 통해 전파된 후 광자 도달 시간의 분포를 검출할 수 있고, 도달 시간의 분포는 샘플의 방출 광의 타이밍 특징 (예를 들어, 발광 수명에 대한 프로시)의 표시를 제공할 수 있다. 일부 실시양태에서, 1개 이상의 광검출기는 마커에 의해 방출된 방출 광의 확률 (예를 들어, 발광 강도)의 표시를 제공한다. 일부 실시양태에서, 방출 광의 공간 분포를 포획하도록 복수의 광검출기가 크기조정되고 배열될 수 있다. 이어서, 1개 이상의 광검출기로부터의 출력 신호를 사용하여 마커를 복수의 마커와 구별할 수 있고, 여기서 복수의 마커는 샘플 내의 샘플을 확인하는데 사용될 수 있다. 일부 실시양태에서, 샘플은 다중 여기 에너지에 의해 여기될 수 있고, 다중 여기 에너지에 반응하여 샘플에 의해 방출된 방출 광 및/또는 방출 광의 타이밍 특징은 마커를 복수의 마커와 구별할 수 있다.

[0277] 작동 시, 여기 광을 사용하여 웰 내의 샘플의 일부 또는 전부를 여기시키고 광검출기로 샘플 방출로부터의 신호를 검출하는 것에 의해 샘플 웰 내의 샘플의 병행 분석이 수행된다. 샘플로부터의 방출 광은 상응하는 광검출기에 의해 검출될 수 있고, 적어도 1개의 전기 신호로 전환될 수 있다. 전기 신호는 서열분석 장치와 인터페이스된 기기에 연결될 수 있는, 서열분석 장치의 회로 내의 전도성 배선을 따라 전송될 수 있다. 전기 신호는 후속적으로 프로세싱 및/또는 분석될 수 있다. 전기 신호의 프로세싱 또는 분석은 기기 상에 또는 기기 밖에 위치한 적합한 컴퓨팅 장치 상에서 이루어질 수 있다.

[0278] 기기는 기기 및/또는 서열분석 장치의 작동을 제어하기 위한 사용자 인터페이스를 포함할 수 있다. 사용자 인터페이스는 사용자가 기기에 정보, 예컨대 기기의 기능을 제어하는데 사용되는 명령 및/또는 설정을 입력하게 하도록 구성될 수 있다. 일부 실시양태에서, 사용자 인터페이스는 버튼, 스위치, 다이얼, 및 음성 명령을 위한 마이크로폰을 포함할 수 있다. 사용자 인터페이스는 사용자가 기기 및/또는 서열분석 장치의 성능에 대한 피드백, 예컨대 적절한 정렬 및/또는 서열분석 장치 상의 광검출기로부터의 판독 신호에 의해 수득된 정보를 수신하게 할 수 있다. 일부 실시양태에서, 사용자 인터페이스는 스피커를 사용하여 피드백을 제공하여 가청 피드백을 제공할 수 있다. 일부 실시양태에서, 사용자 인터페이스는 사용자에게 시각적 피드백을 제공하기 위한 표시등 및/또는 디스플레이 스크린을 포함할 수 있다.

[0279] 일부 실시양태에서, 기기는 컴퓨팅 장치와 연결되도록 구성된 컴퓨터 인터페이스를 포함할 수 있다. 컴퓨터 인터페이스는 USB 인터페이스, 파이어와이어 인터페이스, 또는 임의의 다른 적합한 컴퓨터 인터페이스일 수 있다. 컴퓨팅 장치는 임의의 범용 컴퓨터, 예컨대 랩톱 또는 데스크톱 컴퓨터일 수 있다. 일부 실시양태에서, 컴퓨팅 장치는 적합한 컴퓨터 인터페이스를 통해 무선 네트워크를 통해 접근가능한 서버 (예를 들어, 클라우드-기반 서버)일 수 있다. 컴퓨터 인터페이스는 기기와 컴퓨팅 장치 사이의 정보의 통신을 용이하게 할 수 있다. 기기를 제어 및/또는 구성하기 위한 입력 정보가 컴퓨팅 장치에 제공될 수 있고, 컴퓨터 인터페이스를 통해 기기에 전송될 수 있다. 기기에 의해 생성된 출력 정보는 컴퓨팅 장치에 의해 컴퓨터 인터페이스를 통해 수신될 수 있다. 출력 정보는 기기의 성능, 서열분석 장치의 성능, 및/또는 광검출기의 판독 신호로부터 생성된 데이터에 대한 피드백을 포함할 수 있다.

[0280] 일부 실시양태에서, 기기는 서열분석 장치의 1개 이상의 광검출기로부터 수신된 데이터를 분석하고/거나 제어 신호를 여기원(들)에 전송하도록 구성된 프로세싱 장치를 포함할 수 있다. 일부 실시양태에서, 프로세싱 장치는 범용 프로세서, 특수-적용화된 프로세서 (예를 들어, 중앙 프로세싱 유닛 (CPU), 예컨대 1개 이상의 마이크로프로세서 또는 마이크로컨트롤러 코어, 필드-프로그램가능한 게이트 어레이 (FPGA), 주문형 집적 회로 (ASIC), 맞춤형 집적 회로, 디지털 신호 프로세서 (DSP), 또는 그의 조합)를 포함할 수 있다. 일부 실시양태에서, 1개 이상의 광검출기로부터의 데이터의 프로세싱은 기기의 프로세싱 장치 및 외부 컴퓨팅 장치 둘 다에 의해 수행될 수 있다. 다른 실시양태에서, 외부 컴퓨팅 장치는 생략될 수 있고, 1개 이상의 광검출기로부터의 데이터의 프로세싱은 서열분석 장치의 프로세싱 장치에 의해서만 수행될 수 있다.

[0281] 일부 실시양태에 따르면, 발광 방출 특징에 기초하여 샘플을 분석하도록 구성된 기기는 상이한 발광 분자 사이의 발광 수명 및/또는 강도의 차이, 및/또는 상이한 환경에서의 동일한 발광 분자의 수명 및/또는 강도 사이의 차이를 검출할 수 있다. 본 발명자들은 발광 방출 수명의 차이가 상이한 발광 분자의 존재 또는 부재 사이를 식별하고/거나 발광 분자가 적용되는 상이한 환경 또는 조건 사이를 식별하는데 사용될 수 있음을 인식하고 인지하였다. 일부 경우에, (예를 들어, 방출 파장보다는) 수명에 기초하여 발광 분자를 식별하는 것이 시스템의 측면을 단순화할 수 있다. 예로서, 수명에 기초하여 발광 분자를 식별할 때 파장-구별 광학장치 (예컨대 파장 필터, 각각의 파장에 대한 전용 검출기, 상이한 파장에서의 전용 펄스 광학 공급원, 및/또는 회절 광학장치)는 수가 감소되거나 제거될 수 있다. 일부 경우에, 단일의 특징적인 파장에서 작동하는 단일 펄스 광학 공급원을 사용하여, 광학 스펙트럼의 동일한 파장 영역 내에서 방출하지만 측정가능하게 상이한 수명을 갖는 상이한 발광 분자를 여기시킬 수 있다. 동일한 파장 영역에서 방출되는 상이한 발광 분자를 여기시키고 식별하기 위해, 상

이한 파장에서 작동하는 다중 공급원보다는 단일 펄스 광학 공급원을 사용하는 분석 시스템은 작동 및 유지가 덜 복잡할 수 있고, 더 콤팩트할 수 있고, 더 낮은 비용으로 제조될 수 있다.

[0282] 발광 수명 분석에 기초한 분석 시스템이 특정 이점을 가질 수 있지만, 분석 시스템에 의해 수득된 정보의 양 및/또는 검출 정확도는 추가의 검출 기술을 허용함으로써 증가될 수 있다. 예를 들어, 시스템의 일부 실시양태는 발광 파장 및/또는 발광 강도에 기초하여 샘플의 1종 이상의 특성을 식별하도록 추가적으로 구성될 수 있다. 일부 구현에서, 상이한 발광 표지 사이를 구별하기 위해 발광 강도가 추가적으로 또는 대안적으로 사용될 수 있다. 예를 들어, 일부 발광 표지는 그의 붕괴 속도가 유사할 수 있더라도 유의하게 상이한 강도로 방출될 수 있거나 또는 그의 여기 확률에 있어서 유의한 차이 (예를 들어, 적어도 약 35%의 차이)를 가질 수 있다. 비닝된 신호를 측정된 여기 광에 참조함으로써, 강도 수준에 기초하여 상이한 발광 표지를 구별하는 것이 가능할 수 있다.

[0283] 일부 실시양태에 따르면, 상이한 발광 수명은 발광 표지의 여기 후 발광 방출 사건을 시간-비닝하도록 구성된 광검출기에 의해 구별될 수 있다. 시간 비닝은 광검출기에 대한 단일 전하-축적 사이클 동안 일어날 수 있다. 전하-축적 사이클은 광-생성 캐리어가 시간-비닝 광검출기의 빈에 축적되는 동안의 판독 사건 사이의 간격이다. 시간-비닝 광검출기의 예는 2015년 8월 7일에 출원된 발명의 명칭 "INTEGRATED DEVICE FOR TEMPORAL BINNING OF RECEIVED PHOTONS"의 미국 특허 출원 번호 14/821,656에 기재되어 있으며, 이는 본원에 참조로 포함된다. 일부 실시양태에서, 시간-비닝 광검출기는 광자 흡수/캐리어 생성 영역에서 전하 캐리어를 생성하고, 전하 캐리어 저장 영역에서 전하 캐리어 저장 빈으로 전하 캐리어를 직접 전달할 수 있다. 이러한 실시양태에서, 시간-비닝 광검출기는 캐리어 이동/포획 영역을 포함하지 않을 수 있다. 이러한 시간-비닝 광검출기는 "직접 비닝 픽셀"로 지칭될 수 있다. 직접 비닝 픽셀을 포함한 시간-비닝 광검출기의 예는 2017년 12월 22일에 출원된 발명의 명칭 "INTEGRATED PHOTODETECTOR WITH DIRECT BINNING PIXEL"의 미국 특허 출원 번호 15/852,571에 기재되어 있으며, 이는 본원에 참조로 포함된다.

[0284] 일부 실시양태에서, 동일한 유형의 상이한 개수의 형광단이 샘플 내의 상이한 시약에 연결될 수 있고, 이에 각각의 시약이 발광 강도에 기초하여 확인될 수 있다. 예를 들어, 2개의 형광단이 제1 표지된 친화성 시약에 연결될 수 있고, 4개 이상의 형광단이 제2 표지된 친화성 시약에 연결될 수 있다. 상이한 수의 형광단으로 인해, 상이한 친화성 시약과 연관된 상이한 여기 및 형광단 방출 확률이 존재할 수 있다. 예를 들어, 신호 축적 간격 동안 제2 표지된 친화성 시약에 대한 보다 많은 방출 사건이 존재할 수 있어서, 빈의 걸보기 강도가 제1 표지된 친화성 시약에 대한 것보다 유의하게 더 높다.

[0285] 본 발명자들은 형광단 붕괴 속도 및/또는 형광단 강도에 기초하여 뉴클레오타이드 또는 임의의 다른 생물학적 또는 화학적 샘플을 구별하는 것이 광학 여기 및 검출 시스템의 단순화를 가능하게 할 수 있음을 인식하고 인지하였다. 예를 들어, 광학 여기는 단일-파장 공급원 (예를 들어, 다중 공급원이 아닌 1개의 특징적인 파장을 생산하는 공급원 또는 다중의 상이한 특징적인 파장에서 작동하는 공급원)에 의해 수행될 수 있다. 추가적으로, 파장 구별 광학장치 및 필터가 검출 시스템에 필요하지 않을 수 있다. 또한, 단일 광검출기를 각각의 샘플 웰에 사용하여 상이한 형광단으로부터의 방출을 검출할 수 있다. 어구 "특징적인 파장" 또는 "파장"은 방사선의 제한된 대역폭 내의 중심 또는 우세 파장 (예를 들어, 펄스 광학 공급원에 의해 출력된 20 nm 대역폭 내의 중심 또는 피크 파장)을 지칭하는데 사용된다. 일부 경우에, "특징적인 파장" 또는 "파장"은 공급원에 의해 출력된 방사선의 전체 대역폭 내의 피크 파장을 지칭하기 위해 사용될 수 있다.

[0286] 등가물 및 범주

[0287] 청구범위에서, 단수형은 달리 나타내지 않는 한 또는 달리 문맥으로부터 명백하지 않는 한 하나 또는 하나 초과를 의미할 수 있다. 군의 하나 이상의 구성원 사이에 "또는"을 포함하는 청구범위 또는 설명은 달리 나타내지 않는 한 또는 달리 문맥으로부터 명백하지 않는 한, 군 구성원 중 하나, 하나 초과, 또는 모두가 주어진 생성물 또는 과정에 존재하거나, 그에 사용되거나, 또는 달리 그와 관련되는 경우에 충족된 것으로 간주된다. 본 발명은 군의 정확하게 하나의 구성원이 주어진 생성물 또는 과정에 존재하거나, 그에 사용되거나, 또는 달리 관련되는 실시양태를 포함한다. 본 발명은 군 구성원 중 하나 초과 또는 모두가 주어진 생성물 또는 과정에 존재하거나, 그에 사용되거나, 또는 달리 관련되는 실시양태를 포함한다.

[0288] 또한, 본 발명은 열거된 청구항 중 하나 이상으로부터의 하나 이상의 제한, 요소, 조항, 및 서술적 용어가 또 다른 청구항에 도입되는 모든 변형, 조합, 및 순열을 포괄한다. 예를 들어, 또 다른 청구항에 종속항인 임의의 청구항은 동일한 기본 청구항에 종속항인 임의의 다른 청구항에서 발견되는 하나 이상의 제한을 포함하도록 변형될 수 있다. 요소가 목록으로서, 예를 들어 마쿠쉬 군 포맷으로 제시되는 경우에, 요소의 각각의 하위군이

또한 개시되고, 임의의 요소(들)는 군으로부터 제거될 수 있다. 일반적으로, 본 발명 또는 본 발명의 측면이 특정한 요소 및/또는 특색을 포함하는 것으로 언급되는 경우에, 본 발명 또는 본 발명의 측면의 특정 실시양태는 이러한 요소 및/또는 특색으로 이루어지거나 또는 본질적으로 이루어지는 것으로 이해되어야 한다. 단순성의 목적을 위해, 이들 실시양태는 본원에 구체적으로 제시되지 않았다.

[0289] 본원의 명세서 및 청구범위에 사용된 어구 "및/또는"은 이와 같이 연동된 요소 중 "어느 하나 또는 둘 다", 즉 일부 경우에는 함께 존재하고 다른 경우에는 분리되어 존재하는 요소를 의미하는 것으로 이해되어야 한다. "및/또는"으로 열거된 다중 요소는 동일한 방식으로, 즉 이와 같이 연동된 요소 중 "하나 이상"으로 해석되어야 한다. 구체적으로 확인된 요소와 관련되든 관련되지 않든, "및/또는" 절에 의해 구체적으로 확인된 요소 이외의 다른 요소가 임의로 존재할 수 있다. 따라서, 비제한적 예로서, "A 및/또는 B"에 대한 언급은, "포함하는"과 같은 개방형 언어와 함께 사용되는 경우에, 한 실시양태에서, A 단독 (B 이외의 요소를 임의로 포함함); 또 다른 실시양태에서, B 단독 (A 이외의 요소를 임의로 포함함); 또 다른 실시양태에서, A 및 B 둘 다 (다른 요소를 임의로 포함함) 등을 지칭할 수 있다.

[0290] 본원의 명세서 및 청구범위에 사용된 "또는"은 상기 정의된 바와 같은 "및/또는"과 동일한 의미를 갖는 것으로 이해되어야 한다. 예를 들어, 목록에서 항목을 분리할 때, "또는" 또는 "및/또는"은 포괄적인 것으로, 즉 다수의 요소 또는 요소의 목록 중 적어도 하나를 포함할 뿐만 아니라 하나 초과, 및 임의로, 추가의 열거되지 않은 항목을 포함하는 것으로 해석되어야 한다. "중 오직 하나" 또는 "중 정확하게 하나", 또는 청구범위에서 사용 시, "로 이루어진"과 같이 달리 명백하게 나타낸 용어만이 다수의 요소 또는 요소의 목록 중 정확하게 하나의 요소를 포함하는 것을 지칭할 것이다. 일반적으로, 본원에 사용된 용어 "또는"은 "어느 하나", "중 하나", "중 오직 하나" 또는 "중 정확하게 하나"와 같은 배타적 용어가 이어질 때 배타적 대안 (즉, "하나 또는 다른 것이지만 둘 다는 아님")을 나타내는 것으로만 해석될 것이다. 청구범위에서 사용 시 "로 본질적으로 이루어진"은 특허법 분야에서 사용되는 바와 같은 그의 통상의 의미를 가질 것이다.

[0291] 본원의 명세서 및 청구범위에 사용된 바와 같은, 하나 이상의 요소의 목록과 관련하여 어구 "적어도 하나"는 요소의 목록에서 요소 중 어느 하나 이상으로부터 선택된 적어도 하나의 요소를 의미하지만, 요소의 목록 내에 구체적으로 열거된 각각의 및 모든 요소 중 적어도 하나를 반드시 포함하는 것은 아니며, 요소의 목록에서 요소의 임의의 조합을 배제하는 것은 아닌 것으로 이해되어야 한다. 이러한 정의는 또한, 구체적으로 확인된 요소와 관련되든 관련되지 않든, 어구 "적어도 하나"가 지칭하는 요소의 목록 내에서 구체적으로 확인된 요소 이외의 요소가 임의로 존재할 수 있음을 허용한다. 따라서, 비제한적 예로서, "A 및 B 중 적어도 하나" (또는 동등하게, "A 또는 B 중 적어도 하나", 또는 동등하게 "A 및/또는 B 중 적어도 하나")는, 한 실시양태에서, B가 존재하지 않는, 임의로 하나 초과를 포함한 적어도 하나의 A (및 B 이외의 요소를 임의로 포함함); 또 다른 실시양태에서, A가 존재하지 않는, 임의로 하나 초과를 포함한 적어도 하나의 B (및 A 이외의 요소를 임의로 포함함); 또 다른 실시양태에서, 임의로 하나 초과를 포함한 적어도 하나의 A, 및 임의로 하나 초과를 포함한 적어도 하나의 B (및 다른 요소를 임의로 포함함) 등을 지칭할 수 있다.

[0292] 또한, 달리 명백하게 나타내지 않는 한, 하나 초과와 단계 또는 행위를 포함하는 본원에 청구된 임의의 방법에서, 방법의 단계 또는 행위의 순서는 반드시 방법의 단계 또는 행위가 언급된 순서로 제한되지는 않는 것으로 이해되어야 한다.

[0293] 청구범위뿐만 아니라 상기 명세서에서, 모든 연결 어구, 예컨대 "포함하는", "비롯한", "보유하는", "갖는", "함유하는", "수반하는", "보유한", "로 구성된" 등은 개방형인 것으로, 즉 포함하나 이에 제한되지는 않는 것을 의미하는 것으로 이해되어야 한다. 다만 연결 어구 "로 이루어진" 및 "로 본질적으로 이루어진"은 미국 특허청 특허 심사 절차 매뉴얼, 섹션 2111.03에 제시된 바와 같이 각각 폐쇄형 또는 반-폐쇄형 연결 어구일 것이다. 개방형 연결 어구 (예를 들어, "포함하는")를 사용하여 본 문헌에 기재된 실시양태는 또한 대안적 실시양태에서 개방형 연결 어구에 의해 기재된 특색으로 "로 이루어진" 및 "로 본질적으로 이루어진" 것으로 고려된다는 것이 인지되어야 한다. 예를 들어, 본 출원이 "A 및 B를 포함하는 조성물"을 기재하는 경우에, 본 출원은 또한 대안적 실시양태 "A 및 B로 이루어진 조성물" 및 "A 및 B로 본질적으로 이루어진 조성물"을 고려한다.

[0294] 범위가 주어지는 경우에는 중점이 포함된다. 또한, 달리 나타내지 않는 한 또는 달리 문맥 및 관련 기술분야의 통상의 기술자의 이해로부터 명백하지 않는 한, 범위로서 표현된 값은 본 발명의 상이한 실시양태에서 언급된 범위 내의 임의의 구체적 값 또는 하위-범위를, 문맥이 달리 명백하게 지시하지 않는 한 범위의 하한치 단위의 1/10까지 가정할 수 있다.

[0295] 본 출원은 다양한 허여된 특허, 공개 특허 출원, 학술지 논문 및 다른 공개물을 언급하며, 이들 모두는 본원에

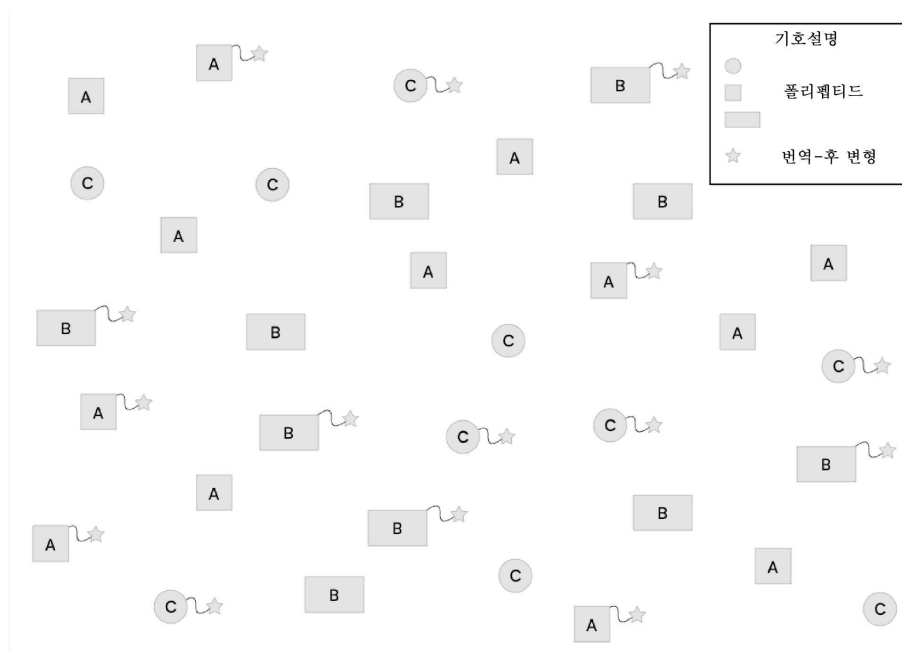
참조로 포함된다. 포함된 참고문헌 중 임의의 것과 본 명세서 사이에 상충이 존재하는 경우, 본 명세서가 우선할 것이다. 또한, 선행 기술 내에 속하는 본 발명의 임의의 특정한 실시양태는 청구항 중 어느 하나 이상으로부터 명백하게 배제될 수 있다. 이러한 실시양태는 관련 기술분야의 통상의 기술자에게 공지된 것으로 간주되기 때문에, 배제가 본원에 명백하게 제시되지 않더라도 이는 배제될 수 있다. 본 발명의 임의의 특정한 실시양태는 선행 기술의 존재와 관련되든 관련되지 않든 임의의 이유로 임의의 청구항으로부터 배제될 수 있다.

[0296] 관련 기술분야의 통상의 기술자는 상용 실험만을 사용하여 본원에 기재된 구체적 실시양태에 대한 많은 등가물을 인식하거나 확인할 수 있을 것이다. 본원에 기재된 본 실시양태의 범주는 상기 설명에 제한되는 것으로 의도되지 않으며, 오히려 첨부된 청구범위에 제시된 바와 같다. 관련 기술분야의 통상의 기술자는 하기 청구범위에 규정된 바와 같은 본 발명의 취지 또는 범주로부터 벗어나지 않으면서 본 설명에 대한 다양한 변화 및 변형이 이루어질 수 있음을 인지할 것이다.

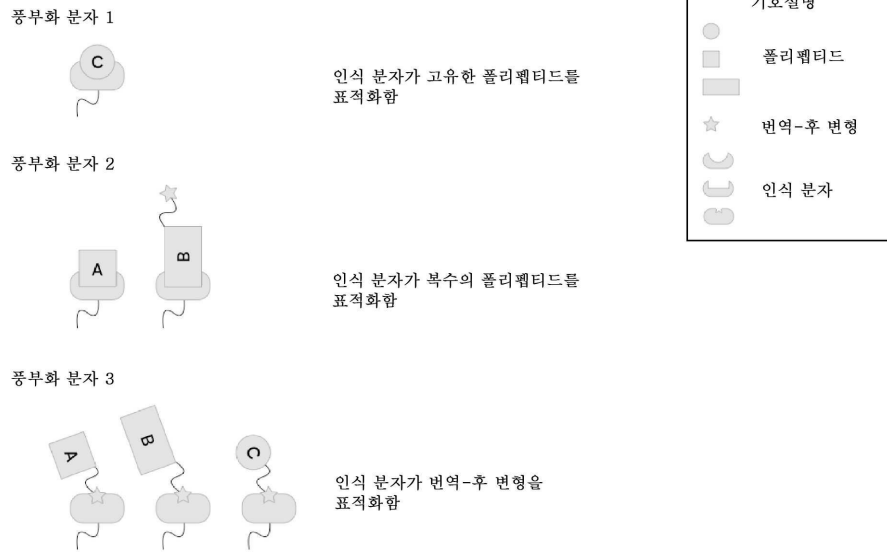
[0297] 본원에서 가변기의 임의의 정의에서 화학적 기의 목록의 언급은 임의의 단일 기 또는 열거된 기의 조합으로서의 그러한 가변기의 정의를 포함한다. 본원에서 가변기에 대한 실시양태의 언급은 임의의 단일 실시양태로서의 또는 임의의 다른 실시양태 또는 그의 부분과 조합된 그러한 실시양태를 포함한다. 본원에서 실시양태의 언급은 임의의 단일 실시양태로서의 또는 임의의 다른 실시양태 또는 그의 부분과 조합된 그러한 실시양태를 포함한다.

도면

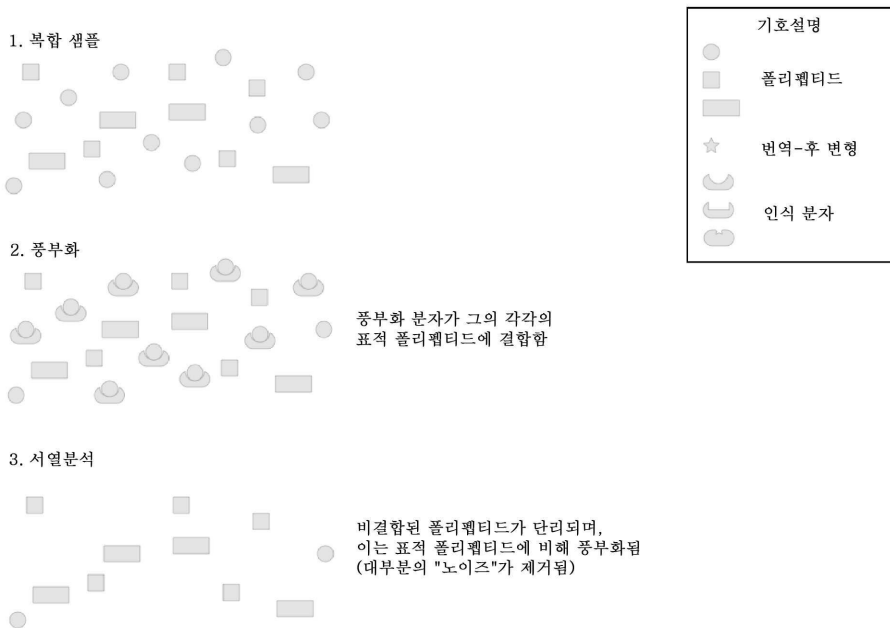
도면1



도면2



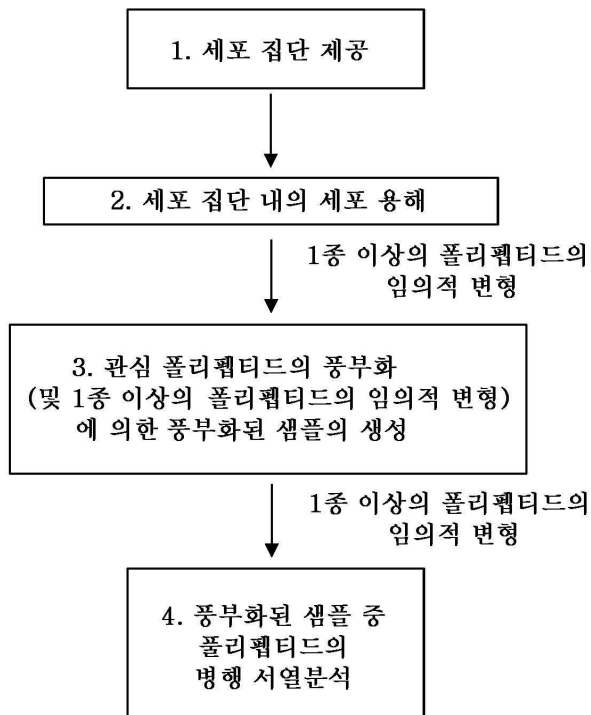
도면3



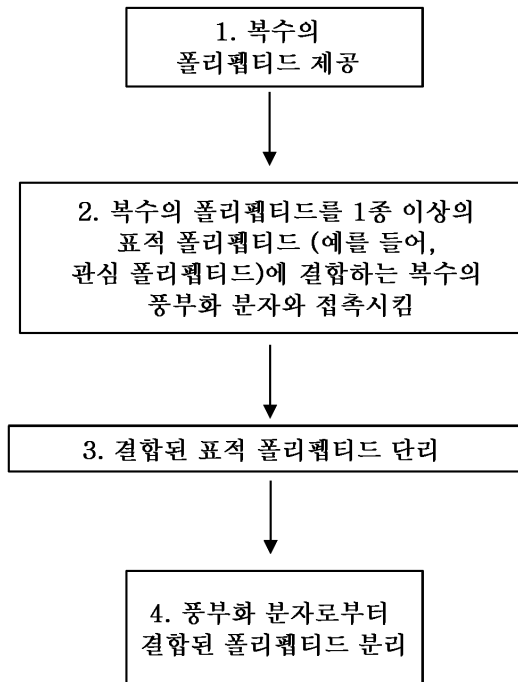
도면4



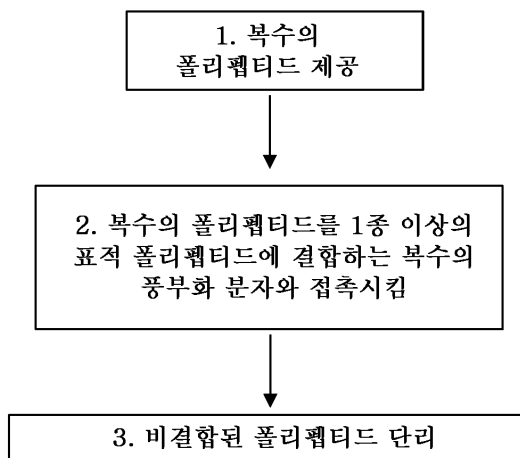
도면5



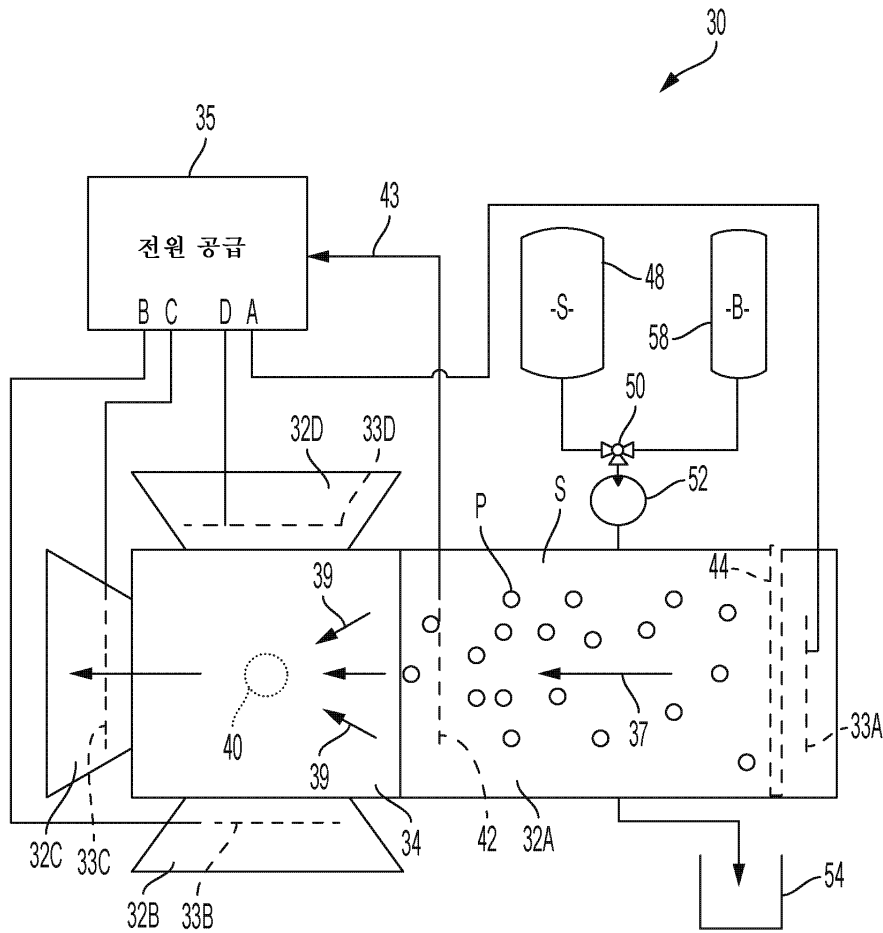
도면6



도면7



도면8



서열목록

SEQUENCE LISTING

<110> Quantum-Si Incorporated

<120> METHODS OF PREPARING AN ENRICHED SAMPLE FOR POLYPEPTIDE

SEQUENCING

<130> R0708.70076W000

<140> Not Yet Assigned

<141> Concurrently Herewith

<150> US 62/926,897

<151> 2019-10-28

<160> 25

<170> PatentIn version 3.5

<210> 1

<211> 921

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 1

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro

1 5 10 15

Arg Gly Ser His Met Met Val Lys Gln Gly Val Phe Met Lys Thr Asp

20 25 30

Gln Ser Lys Val Lys Lys Leu Ser Asp Tyr Lys Ser Leu Asp Tyr Phe

35 40 45

Val Ile His Val Asp Leu Gln Ile Asp Leu Ser Lys Lys Pro Val Glu

50 55 60

Ser Lys Ala Arg Leu Thr Val Val Pro Asn Leu Asn Val Asp Ser His

65 70 75 80

Ser Asn Asp Leu Val Leu Asp Gly Glu Asn Met Thr Leu Val Ser Leu

85 90 95

Gln Met Asn Asp Asn Leu Leu Lys Glu Asn Glu Tyr Glu Leu Thr Lys

100 105 110

Asp Ser Leu Ile Ile Lys Asn Ile Pro Gln Asn Thr Pro Phe Thr Ile

115 120 125

Glu Met Thr Ser Leu Leu Gly Glu Asn Thr Asp Leu Phe Gly Leu Tyr

130 135 140

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

145 150 155 160

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

165 170 175

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

180 185 190

Gly Val Leu Ile Glu Lys Lys Glu Leu Pro Leu Gly Leu His Ser Val

195 200 205

Thr Trp Leu Asp Asp Val Pro Lys Pro Ser Tyr Leu Phe Ala Leu Val

210 215 220
 Ala Gly Asn Leu Gln Arg Ser Val Thr Tyr Tyr Gln Thr Lys Ser Gly
 225 230 235 240
 Arg Glu Leu Pro Ile Glu Phe Tyr Val Pro Pro Ser Ala Thr Ser Lys
 245 250 255
 Cys Asp Phe Ala Lys Glu Val Leu Lys Glu Ala Met Ala Trp Asp Glu
 260 265 270

 Arg Thr Phe Asn Leu Glu Cys Ala Leu Arg Gln His Met Val Ala Gly
 275 280 285
 Val Asp Lys Tyr Ala Ser Gly Ala Ser Glu Pro Thr Gly Leu Asn Leu
 290 295 300
 Phe Asn Thr Glu Asn Leu Phe Ala Ser Pro Glu Thr Lys Thr Asp Leu
 305 310 315 320
 Gly Ile Leu Arg Val Leu Glu Val Val Ala His Glu Phe Phe His Tyr
 325 330 335

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

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

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

 Phe Glu Asn Ile Gln Thr Arg Pro Ile Pro Ser Leu Leu Arg Ser Phe
 530 535 540
 Ser Ala Pro Val His Met Lys Tyr Glu Tyr Ser Tyr Gln Asp Leu Leu
 545 550 555 560
 Leu Leu Met Gln Phe Asp Thr Asn Leu Tyr Asn Arg Cys Glu Ala Ala
 565 570 575
 Lys Gln Leu Ile Ser Ala Leu Ile Asn Asp Phe Cys Ile Gly Lys Lys
 580 585 590

 Ile Glu Leu Ser Pro Gln Phe Phe Ala Val Tyr Lys Ala Leu Leu Ser
 595 600 605
 Asp Asn Ser Leu Asn Glu Trp Met Leu Ala Glu Leu Ile Thr Leu Pro
 610 615 620
 Ser Leu Glu Glu Leu Ile Glu Asn Gln Asp Lys Pro Asp Phe Glu Lys
 625 630 635 640
 Leu Asn Glu Gly Arg Gln Leu Ile Gln Asn Ala Leu Ala Asn Glu Leu
 645 650 655

 Lys Thr Asp Phe Tyr Asn Leu Leu Phe Arg Ile Gln Ile Ser Gly Asp
 660 665 670
 Asp Asp Lys Gln Lys Leu Lys Gly Phe Asp Leu Lys Gln Ala Gly Leu
 675 680 685
 Arg Arg Leu Lys Ser Val Cys Phe Ser Tyr Leu Leu Asn Val Asp Phe
 690 695 700
 Glu Lys Thr Lys Glu Lys Leu Ile Leu Gln Phe Glu Asp Ala Leu Gly

705 710 715 720

Lys Asn Met Thr Glu Thr Ala Leu Ala Leu Ser Met Leu Cys Glu Ile

725 730 735

Asn Cys Glu Glu Ala Asp Val Ala Leu Glu Asp Tyr Tyr His Tyr Trp

740 745 750

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

755 760 765

Ala His Ser Pro Asp Val Ile Glu Arg Val Lys Lys Leu Met Arg His

770 775 780

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

785 790 795 800

Ser Phe Ile Lys Asn Pro Phe Gly Phe His Ser Val Thr Gly Glu Gly

805 810 815

Tyr Gln Leu Val Ala Asp Ala Ile Phe Asp Leu Asp Lys Ile Asn Pro

820 825 830

Thr Leu Ala Ala Asn Leu Thr Glu Lys Phe Thr Tyr Trp Asp Lys Tyr

835 840 845

Asp Val Asn Arg Gln Ala Met Met Ile Ser Thr Leu Lys Ile Ile Tyr

850 855 860

Ser Asn Ala Thr Ser Ser Asp Val Arg Thr Met Ala Lys Lys Gly Leu

865 870 875 880

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

885 890 895

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

900 905 910

Asn Lys Glu Asn Ala Tyr Gln Leu His

915 920

 $\langle 210 \rangle$ 2

<211> 273

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 2

Met Ala His His His His His His Met Gly Thr Ala Ile Ser Ile Lys

1 5 10 15

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

20 25 30

Glu Val Leu Glu Met Ile Glu Pro Tyr Val Lys Pro Gly Val Ser Thr

35 40 45

Gly Glu Leu Asp Arg Ile Cys Asn Asp Tyr Ile Val Asn Glu Gln His

50 55 60

Ala Val Ser Ala Cys Leu Gly Tyr His Gly Tyr Pro Lys Ser Val Cys

65 70 75 80

Ile Ser Ile Asn Glu Val Val Cys His Gly Ile Pro Asp Asp Ala Lys

85 90 95

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

100 105 110

Asp Gly Phe His Gly Asp Thr Ser Lys Met Phe Ile Val Gly Lys Pro

115 120 125

Thr Ile Met Gly Glu Arg Leu Cys Arg Ile Thr Gln Glu Ser Leu Tyr

130 135 140

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

145 150 155 160

Ala Ala Ile Gln Lys Phe Val Glu Ala Glu Gly Phe Ser Val Val Arg

165 170 175

Glu Tyr Cys Gly His Gly Ile Gly Arg Gly Phe His Glu Glu Pro Gln

180 185 190

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

195 200 205

Met Thr Phe Thr Ile Glu Pro Met Val Asn Ala Gly Lys Lys Glu Ile

210 215 220

Arg Thr Met Lys Asp Gly Trp Thr Val Lys Thr Lys Asp Arg Ser Leu

225 230 235 240

 Ser Ala Gln Tyr Glu His Thr Ile Val Val Thr Asp Asn Gly Cys Glu
 245 250 255
 Ile Leu Thr Leu Arg Lys Asp Asp Thr Ile Pro Ala Ile Ile Ser His
 260 265 270
 Asp

 <210> 3
 <211> 330
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 3
 Met Ala His His His His His His Met Gly Thr Leu Glu Ala Asn Thr
 1 5 10 15

 Asn Gly Pro Gly Ser Met Leu Ser Arg Met Pro Val Ser Ser Arg Thr
 20 25 30
 Val Pro Phe Gly Asp His Glu Thr Trp Val Gln Val Thr Thr Pro Glu
 35 40 45
 Asn Ala Gln Pro His Ala Leu Pro Leu Ile Val Leu His Gly Gly Pro
 50 55 60
 Gly Met Ala His Asn Tyr Val Ala Asn Ile Ala Ala Leu Ala Asp Glu
 65 70 75 80

 Thr Gly Arg Thr Val Ile His Tyr Asp Gln Val Gly Cys Gly Asn Ser
 85 90 95
 Thr His Leu Pro Asp Ala Pro Ala Asp Phe Trp Thr Pro Gln Leu Phe
 100 105 110
 Val Asp Glu Phe His Ala Val Cys Thr Ala Leu Gly Ile Glu Arg Tyr
 115 120 125
 His Val Leu Gly Gln Ser Trp Gly Gly Met Leu Gly Ala Glu Ile Ala
 130 135 140

Val Arg Gln Pro Ser Gly Leu Val Ser Leu Ala Ile Cys Asn Ser Pro
 145 150 155 160
 Ala Ser Met Arg Leu Trp Ser Glu Ala Ala Gly Asp Leu Arg Ala Gln
 165 170 175
 Leu Pro Ala Glu Thr Arg Ala Ala Leu Asp Arg His Glu Ala Ala Gly
 180 185 190
 Thr Ile Thr His Pro Asp Tyr Leu Gln Ala Ala Ala Glu Phe Tyr Arg
 195 200 205

Arg His Val Cys Arg Val Val Pro Thr Pro Gln Asp Phe Ala Asp Ser
 210 215 220
 Val Ala Gln Met Glu Ala Glu Pro Thr Val Tyr His Thr Met Asn Gly
 225 230 235 240
 Pro Asn Glu Phe His Val Val Gly Thr Leu Gly Asp Trp Ser Val Ile
 245 250 255
 Asp Arg Leu Pro Asp Val Thr Ala Pro Val Leu Val Ile Ala Gly Glu
 260 265 270

His Asp Glu Ala Thr Pro Lys Thr Trp Gln Pro Phe Val Asp His Ile
 275 280 285
 Pro Asp Val Arg Ser His Val Phe Pro Gly Thr Ser His Cys Thr His
 290 295 300
 Leu Glu Lys Pro Glu Glu Phe Arg Ala Val Val Ala Gln Phe Leu His
 305 310 315 320
 Gln His Asp Leu Ala Ala Asp Ala Arg Val
 325 330

<210> 4

<211> 452

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 4

Met Thr Gln Gln Glu Tyr Gln Asn Arg Arg Gln Ala Leu Leu Ala Lys
 1 5 10 15

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

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

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

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

 435 440 445
 His His His His
 450
 <210> 5
 <211> 303
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 5
 Met Asp Thr Glu Lys Leu Met Lys Ala Gly Glu Ile Ala Lys Lys Val

1 5 10 15
 Arg Glu Lys Ala Ile Lys Leu Ala Arg Pro Gly Met Leu Leu Leu Glu
 20 25 30
 Leu Ala Glu Ser Ile Glu Lys Met Ile Met Glu Leu Gly Gly Lys Pro

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

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

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

 225 230 235 240
 Gln Asn Asp Met Pro Glu Gly Gln Leu Lys Leu Ala Leu Lys Thr Leu
 245 250 255

Glu Lys Ala Gly Ala Ile Tyr Gly Tyr Pro Val Leu Lys Glu Ile Arg
260 265 270
Asn Gly Ile Val Ala Gln Phe Glu His Thr Ile Ile Val Glu Lys Asp
275 280 285
Ser Val Ile Val Thr Gln Asp Met Ile Asn Lys Ser Thr Leu Glu

290 295 300

<210> 6

<211> 428

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 6

His Met Ser Ser Pro Leu His Tyr Val Leu Asp Gly Ile His Cys Glu
1 5 10 15
Pro His Phe Phe Thr Val Pro Leu Asp His Gln Gln Pro Asp Asp Glu
20 25 30
Glu Thr Ile Thr Leu Phe Gly Arg Thr Leu Cys Arg Lys Asp Arg Leu
35 40 45

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

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

Pro Asp Ser Leu His Glu Val Tyr Leu Thr Gly Gly Val Ala Pro Ile
165 170 175

Gly Arg Ser Ala Asp Glu Val Tyr Arg Ala Thr Tyr Gln Arg Val Ala
180 185 190

Asp Lys Asn Arg Ala Phe Phe Ala Arg Phe Pro His Ala Gln Ala Ile
195 200 205

Ala Asn Arg Leu Ala Thr His Leu Gln Arg His Asp Val Arg Leu Pro
210 215 220

Asn Gly Gln Arg Leu Thr Val Glu Gln Leu Gln Gln Gln Gly Leu Asp
225 230 235 240

Leu Gly Ala Ser Gly Ala Phe Glu Glu Leu Tyr Tyr Leu Leu Glu Asp
245 250 255

Ala Phe Ile Gly Glu Lys Leu Asn Pro Ala Phe Leu Tyr Gln Val Gln
260 265 270

Ala Met Gln Pro Phe Asn Thr Asn Pro Val Phe Ala Ile Leu His Glu
275 280 285

Leu Ile Tyr Cys Glu Gly Ala Ala Ser His Trp Ala Ala Glu Arg Val
290 295 300

Arg Gly Glu Phe Pro Ala Leu Ala Trp Ala Gln Gly Lys Asp Phe Ala
305 310 315 320

Phe Thr Gly Glu Met Ile Phe Pro Trp Met Phe Glu Gln Phe Arg Glu
325 330 335

Leu Ile Pro Leu Lys Glu Ala Ala His Leu Leu Ala Glu Lys Ala Asp
340 345 350

Trp Gly Pro Leu Tyr Asp Pro Val Gln Leu Ala Arg Asn Lys Val Pro
355 360 365

Val Ala Cys Ala Val Tyr Ala Glu Asp Met Tyr Val Glu Phe Asp Tyr
370 375 380

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

Asn Glu Tyr Glu His Asn Gly Leu Arg Val Asp Gly Glu Gln Ile Leu

405 410 415
 Asp Arg Leu Ile Arg Leu Asn Arg Asp Cys Leu Glu
 420 425
 <210> 7
 <211
 > 348
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 7
 Met Lys Glu Arg Leu Glu Lys Leu Val Lys Phe Met Asp Glu Asn Ser
 1 5 10 15
 Ile Asp Arg Val Phe Ile Ala Lys Pro Val Asn Val Tyr Tyr Phe Ser
 20 25 30
 Gly Thr Ser Pro Leu Gly Gly Gly Tyr Ile Ile Val Asp Gly Asp Glu
 35 40 45
 Ala Thr Leu Tyr Val Pro Glu Leu Glu Tyr Glu Met Ala Lys Glu Glu
 50 55 60
 Ser Lys Leu Pro Val Val Lys Phe Lys Lys Phe Asp Glu Ile Tyr Glu
 65 70 75 80
 Ile Leu Lys Asn Thr Glu Thr Leu Gly Ile Glu Gly Thr Leu Ser Tyr
 85 90 95
 Ser Met Val Glu Asn Phe Lys Glu Lys Ser Asn Val Lys Glu Phe Lys
 100 105 110
 Lys Ile Asp Asp Val Ile Lys Asp Leu Arg Ile Ile Lys Thr Lys Glu
 115 120 125
 Glu Ile Glu Ile Ile Glu Lys Ala Cys Glu Ile Ala Asp Lys Ala Val
 130 135 140
 Met Ala Ala Ile Glu Glu Ile Thr Glu Gly Lys Arg Glu Arg Glu Val
 145 150 155 160
 Ala Ala Lys Val Glu Tyr Leu Met Lys Met Asn Gly Ala Glu Lys Pro
 165 170 175

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

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

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

305 310 315 320
Leu Gly Gly Val Arg Ile Glu Asp Thr Val Leu Ile Thr Glu Asn Gly
325 330 335
Ala Lys Arg Leu Thr Lys Thr Glu Arg Glu Leu Leu
340 345

<210> 8

<211> 298

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 8

Met Ile Pro Ile Thr Thr Pro Val Gly Asn Phe Lys Val Trp Thr Lys
1 5 10 15

Arg Phe Gly Thr Asn Pro Lys Ile Lys Val Leu Leu Leu His Gly Gly

20	25	30
Pro Ala Met Thr His Glu Tyr Met Glu Cys Phe Glu Thr Phe Phe Gln		
35	40	45
Arg Glu Gly Phe Glu Phe Tyr Glu Tyr Asp Gln Leu Gly Ser Tyr Tyr		
50	55	60
Ser Asp Gln Pro Thr Asp Glu Lys Leu Trp Asn Ile Asp Arg Phe Val		
65	70	75
80		
Asp Glu Val Glu Gln Val Arg Lys Ala Ile His Ala Asp Lys Glu Asn		
85	90	95
Phe Tyr Val Leu Gly Asn Ser Trp Gly Gly Ile Leu Ala Met Glu Tyr		
100	105	110
Ala Leu Lys Tyr Gln Gln Asn Leu Lys Gly Leu Ile Val Ala Asn Met		
115	120	125
Met Ala Ser Ala Pro Glu Tyr Val Lys Tyr Ala Glu Val Leu Ser Lys		
130	135	140
Gln Met Lys Pro Glu Val Leu Ala Glu Val Arg Ala Ile Glu Ala Lys		
145	150	155
160		
Lys Asp Tyr Ala Asn Pro Arg Tyr Thr Glu Leu Leu Phe Pro Asn Tyr		
165	170	175
Tyr Ala Gln His Ile Cys Arg Leu Lys Glu Trp Pro Asp Ala Leu Asn		
180	185	190
Arg Ser Leu Lys His Val Asn Ser Thr Val Tyr Thr Leu Met Gln Gly		
195	200	205
Pro Ser Glu Leu Gly Met Ser Ser Asp Ala Arg Leu Ala Lys Trp Asp		
210	215	220
Ile Lys Asn Arg Leu His Glu Ile Ala Thr Pro Thr Leu Met Ile Gly		
225	230	235
240		
Ala Arg Tyr Asp Thr Met Asp Pro Lys Ala Met Glu Glu Gln Ser Lys		
245	250	255
Leu Val Gln Lys Gly Arg Tyr Leu Tyr Cys Pro Asn Gly Ser His Leu		
260	265	270

Ala Met Trp Asp Asp Gln Lys Val Phe Met Asp Gly Val Ile Lys Phe
275 280 285

Ile Lys Asp Val Asp Thr Lys Ser Phe Asn
290 295

<210> 9
<211> 428
<212> PRT
<213> Artificial Sequence
<220><223> Synthetic
<400> 9

His Met Ser Ser Pro Leu His Tyr Val Leu Asp Gly Ile His Cys Glu
1 5 10 15

Pro His Phe Phe Thr Val Pro Leu Asp His Gln Gln Pro Asp Asp Glu
20 25 30

Glu Thr Ile Thr Leu Phe Gly Arg Thr Leu Cys Arg Lys Asp Arg Leu
35 40 45

Asp Asp Glu Leu Pro Trp Leu Leu Tyr Leu Gln Gly Gly Pro Gly Phe
50 55 60

Gly Ala Pro Arg Pro Ser Ala Asn Gly Gly Trp Ile Lys Arg Ala Leu
65 70 75 80

Gln Glu Phe Arg Val Leu Leu Leu Asp Gln Arg Gly Thr Gly His Ser
85 90 95

Thr Pro Ile His Ala Glu Leu Leu Ala His Leu Asn Pro Arg Gln Gln
100 105 110

Ala Asp Tyr Leu Ser His Phe Arg Ala Asp Ser Ile Val Arg Asp Ala
115 120 125

Glu Leu Ile Arg Glu Gln Leu Ser Pro Asp His Pro Trp Ser Leu Leu
130 135 140

Gly Gln Ser Phe Gly Gly Phe Cys Ser Leu Thr Tyr Leu Ser Leu Phe
145 150 155 160

Pro Asp Ser Leu His Glu Val Tyr Leu Thr Gly Gly Val Ala Pro Ile
165 170 175

Gly Arg Ser Ala Asp Glu Val Tyr Arg Ala Thr Tyr Gln Arg Val Ala
180 185 190

Asp Lys Asn Arg Ala Phe Phe Ala Arg Phe Pro His Ala Gln Ala Ile
195 200 205

Ala Asn Arg Leu Ala Thr His Leu Gln Arg His Asp Val Arg Leu Pro
210 215 220

Asn Gly Gln Arg Leu Thr Val Glu Gln Leu Gln Gln Gln Gly Leu Asp
225 230 235 240

Leu Gly Ala Ser Gly Ala Phe Glu Glu Leu Tyr Tyr Leu Leu Glu Asp
245 250 255

Ala Phe Ile Gly Glu Lys Leu Asn Pro Ala Phe Leu Tyr Gln Val Gln
260 265 270

Ala Met Gln Pro Phe Asn Thr Asn Pro Val Phe Ala Ile Leu His Glu
275 280 285

Leu Ile Tyr Cys Glu Gly Ala Ala Ser His Trp Ala Ala Glu Arg Val
290 295 300

Arg Gly Glu Phe Pro Ala Leu Ala Trp Ala Gln Gly Lys Asp Phe Ala
305 310 315 320

Phe Thr Gly Glu Met Ile Phe Pro Trp Met Phe Glu Gln Phe Arg Glu
325 330 335

Leu Ile Pro Leu Lys Glu Ala Ala His Leu Leu Ala Glu Lys Ala Asp
340 345 350

Trp Gly Pro Leu Tyr Asp Pro Val Gln Leu Ala Arg Asn Lys Val Pro
355 360 365

Val Ala Cys Ala Val Tyr Ala Glu Asp Met Tyr Val Glu Phe Asp Tyr
370 375 380

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

Asn Glu Tyr Glu His Asn Gly Leu Arg Val Asp Gly Glu Gln Ile Leu
405 410 415

Asp Arg Leu Ile Arg Leu Asn Arg Asp Cys Leu Glu

420 425

<210> 10

<211> 310

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 10

Met Tyr Glu Ile Lys Gln Pro Phe His Ser Gly Tyr Leu Gln Val Ser

1 5 10 15

Glu Ile His Gln Ile Tyr Trp Glu Glu Ser Gly Asn Pro Asp Gly Val

20 25 30

Pro Val Ile Phe Leu His Gly Gly Pro Gly Ala Gly Ala Ser Pro Glu

35 40 45

Cys Arg Gly Phe Phe Asn Pro Asp Val Phe Arg Ile Val Ile Ile Asp

50 55 60

Gln Arg Gly Cys Gly Arg Ser His Pro Tyr Ala Cys Ala Glu Asp Asn

65 70 75 80

Thr Thr Trp Asp Leu Val Ala Asp Ile Glu Lys Val Arg Glu Met Leu

85 90 95

Gly Ile Gly Lys Trp Leu Val Phe Gly Gly Ser Trp Gly Ser Thr Leu

100 105 110

Ser Leu Ala Tyr Ala Gln Thr His Pro Glu Arg Val Lys Gly Leu Val

115 120 125

Leu Arg Gly Ile Phe Leu Cys Arg Pro Ser Glu Thr Ala Trp Leu Asn

130 135 140

Glu Ala Gly Gly Val Ser Arg Ile Tyr Pro Glu Gln Trp Gln Lys Phe

145 150 155 160

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

165 170 175

His Gly Leu Leu Phe His Gln Asp Glu Glu Val Cys Leu Ser Ala Ala

180 185 190

Lys Ala Trp Ala Asp Trp Glu Ser Tyr Leu Ile Arg Phe Glu Pro Glu

195 200 205
 Gly Val Asp Glu Asp Ala Tyr Ala Ser Leu Ala Ile Ala Arg Leu Glu
 210 215 220

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

 Ala Phe Asp Pro Pro Leu Ala Asp Ala Leu Val Gln Ala Val Glu Asp
 290 295 300
 Ile Leu Pro Arg Leu Leu
 305 310
 <210> 11
 <211> 891
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 11
 Met Gly Ser Ser His His His His His His Ser Ser Gly Glu Asn Leu
 1 5 10 15
 Tyr Phe Gln Gly His Met Thr Gln Gln Pro Gln Ala Lys Tyr Arg His
 20 25 30

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

85

90

95

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

100

105

110

Leu Lys Ile Ile Asn Glu Ile Ser Pro Ala Ala Asn Thr Ala Leu Glu

115

120

125

Gly Leu Tyr Gln Ser Gly Asp Ala Leu Cys Thr Gln Cys Glu Ala Glu

130

135

140

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

145

150

155

160

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

165

170

175

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

180

185

190

His Trp Val Gln Trp Gln Asp Pro Phe Pro Lys Pro Cys Tyr Leu Phe

195

200

205

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

210

215

220

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

225

230

235

240

Leu Asp Arg Ala Pro Trp Ala Met Thr Ser Leu Lys Asn Ser Met Lys

245

250

255

Trp Asp Glu Glu Arg Phe Gly Leu Glu Tyr Asp Leu Asp Ile Tyr Met

260

265

270

Ile Val Ala Val Asp Phe Phe Asn Met Gly Ala Met Glu Asn Lys Gly

275

280

285

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

290

295

300

Thr Asp Lys Asp Tyr Leu Asp Ile Glu Arg Val Ile Gly His Glu Tyr

305

310

315

320

Phe His Asn Trp Thr Gly Asn Arg Val Thr Cys Arg Asp Trp Phe Gln

325

330

335

Leu Ser Leu Lys Glu Gly Leu Thr Val Phe Arg Asp Gln Glu Phe Ser
 340 345 350

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

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

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

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

580 585 590
 Leu Ala Thr Tyr Ile Lys Leu Asn Val Ala Arg His Gln Gln Gly Gln
 595 600 605

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

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

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

 Asn Arg Ile Arg Ser Leu Ile Gly Ala Phe Ala Gly Ser Asn Pro Ala
 805 810 815
 Ala Phe His Ala Glu Asp Gly Ser Gly Tyr Leu Phe Leu Val Glu Met
 820 825 830

Leu Thr Asp Leu Asn Ser Arg Asn Pro Gln Val Ala Ser Arg Leu Ile
 835 840 845
 Glu Pro Leu Ile Arg Leu Lys Arg Tyr Asp Ala Lys Arg Gln Glu Lys
 850 855 860

 Met Arg Ala Ala Leu Glu Gln Leu Lys Gly Leu Glu Asn Leu Ser Gly
 865 870 875 880
 Asp Leu Tyr Glu Lys Ile Thr Lys Ala Leu Ala
 885 890

 <210> 12
 <211> 889
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 12
 Pro Lys Ile His Tyr Arg Lys Asp Tyr Lys Pro Ser Gly Phe Ile Ile
 1 5 10 15
 Asn Gln Val Thr Leu Asn Ile Asn Ile His Asp Gln Glu Thr Ile Val

 20 25 30
 Arg Ser Val Leu Asp Met Asp Ile Ser Lys His Asn Val Gly Glu Asp
 35 40 45
 Leu Val Phe Asp Gly Val Gly Leu Lys Ile Asn Glu Ile Ser Ile Asn
 50 55 60
 Asn Lys Lys Leu Val Glu Gly Glu Glu Tyr Thr Tyr Asp Asn Glu Phe
 65 70 75 80
 Leu Thr Ile Phe Ser Lys Phe Val Pro Lys Ser Lys Phe Ala Phe Ser

 85 90 95
 Ser Glu Val Ile Ile His Pro Glu Thr Asn Tyr Ala Leu Thr Gly Leu
 100 105 110
 Tyr Lys Ser Lys Asn Ile Ile Val Ser Gln Cys Glu Ala Thr Gly Phe
 115 120 125
 Arg Arg Ile Thr Phe Phe Ile Asp Arg Pro Asp Met Met Ala Lys Tyr
 130 135 140

Asp Val Thr Val Thr Ala Asp Lys Glu Lys Tyr Pro Val Leu Leu Ser

145 150 155 160

Asn Gly Asp Lys Val Asn Glu Phe Glu Ile Pro Gly Gly Arg His Gly

165 170 175

Ala Arg Phe Asn Asp Pro Pro Leu Lys Pro Cys Tyr Leu Phe Ala Val

180 185 190

Val Ala Gly Asp Leu Lys His Leu Ser Ala Thr Tyr Ile Thr Lys Tyr

195 200 205

Thr Lys Lys Lys Val Glu Leu Tyr Val Phe Ser Glu Glu Lys Tyr Val

210 215 220

Ser Lys Leu Gln Trp Ala Leu Glu Cys Leu Lys Lys Ser Met Ala Phe

225 230 235 240

Asp Glu Asp Tyr Phe Gly Leu Glu Tyr Asp Leu Ser Arg Leu Asn Leu

245 250 255

Val Ala Val Ser Asp Phe Asn Val Gly Ala Met Glu Asn Lys Gly Leu

260 265 270

Asn Ile Phe Asn Ala Asn Ser Leu Leu Ala Ser Lys Lys Asn Ser Ile

275 280 285

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

290 295 300

His Gln Tyr Thr Gly Asn Arg Val Thr Leu Arg Asp Trp Phe Gln Leu

305 310 315 320

Thr Leu Lys Glu Gly Leu Thr Val His Arg Glu Asn Leu Phe Ser Glu

325 330 335

Glu Met Thr Lys Thr Val Thr Thr Arg Leu Ser His Val Asp Leu Leu

340 345 350

Arg Ser Val Gln Phe Leu Glu Asp Ser Ser Pro Leu Ser His Pro Ile

355 360 365

Arg Pro Glu Ser Tyr Val Ser Met Glu Asn Phe Tyr Thr Thr Thr Val

370 375 380

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

385 390 395 400
 Glu Glu Tyr Tyr Lys Lys Gly Phe Asp Ile Tyr Ile Lys Lys Asn Asp

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

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

 530 535 540
 Val Tyr Ile Glu Asp Gln Leu Thr Asp Glu Glu Arg Ile Leu Leu Leu
 545 550 555 560
 Lys Tyr Asp Ser Asp Ala Phe Val Arg Tyr Asn Ser Cys Thr Asn Ile
 565 570 575
 Tyr Met Lys Gln Ile Leu Met Asn Tyr Asn Glu Phe Leu Lys Ala Lys
 580 585 590
 Asn Glu Lys Leu Glu Ser Phe Gln Leu Thr Pro Val Asn Ala Gln Phe

 595 600 605
 Ile Asp Ala Ile Lys Tyr Leu Leu Glu Asp Pro His Ala Asp Ala Gly
 610 615 620
 Phe Lys Ser Tyr Ile Val Ser Leu Pro Gln Asp Arg Tyr Ile Ile Asn
 625 630 635 640

Phe Val Ser Asn Leu Asp Thr Asp Val Leu Ala Asp Thr Lys Glu Tyr
 645 650 655
 Ile Tyr Lys Gln Ile Gly Asp Lys Leu Asn Asp Val Tyr Tyr Lys Met
 660 665 670
 Phe Lys Ser Leu Glu Ala Lys Ala Asp Asp Leu Thr Tyr Phe Asn Asp
 675 680 685
 Glu Ser His Val Asp Phe Asp Gln Met Asn Met Arg Thr Leu Arg Asn
 690 695 700
 Thr Leu Leu Ser Leu Leu Ser Lys Ala Gln Tyr Pro Asn Ile Leu Asn
 705 710 715 720
 Glu Ile Ile Glu His Ser Lys Ser Pro Tyr Pro Ser Asn Trp Leu Thr
 725 730 735
 Ser Leu Ser Val Ser Ala Tyr Phe Asp Lys Tyr Phe Glu Leu Tyr Asp
 740 745 750
 Lys Thr Tyr Lys Leu Ser Lys Asp Asp Glu Leu Leu Leu Gln Glu Trp
 755 760 765
 Leu Lys Thr Val Ser Arg Ser Asp Arg Lys Asp Ile Tyr Glu Ile Leu
 770 775 780
 Lys Lys Leu Glu Asn Glu Val Leu Lys Asp Ser Lys Asn Pro Asn Asp
 785 790 795 800
 Ile Arg Ala Val Tyr Leu Pro Phe Thr Asn Asn Leu Arg Arg Phe His
 805 810 815
 Asp Ile Ser Gly Lys Gly Tyr Lys Leu Ile Ala Glu Val Ile Thr Lys
 820 825 830
 Thr Asp Lys Phe Asn Pro Met Val Ala Thr Gln Leu Cys Glu Pro Phe
 835 840 845
 Lys Leu Trp Asn Lys Leu Asp Thr Lys Arg Gln Glu Leu Met Leu Asn
 850 855 860
 Glu Met Asn Thr Met Leu Gln Glu Pro Gln Ile Ser Asn Asn Leu Lys
 865 870 875 880
 Glu Tyr Leu Leu Arg Leu Thr Asn Lys

885

<210> 13

<211> 932

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 13

Met Gly Ser Ser His His His His His His Ser Ser Gly Met Trp Leu

1 5 10 15

Ala Ala Ala Ala Pro Ser Leu Ala Arg Arg Leu Leu Phe Leu Gly Pro

20 25 30

Pro Pro Pro Pro Leu Leu Leu Leu Val Phe Ser Arg Ser Ser Arg Arg

35 40 45

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

50 55 60

Glu Arg Leu Pro Ala Asp Val Ser Pro Ile Asn Tyr Ser Leu Cys Leu

65 70 75 80

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

85 90 95

Ala Gln Val Arg Gln Ala Thr Asn Gln Ile Val Met Asn Cys Ala Asp

100 105 110

Ile Asp Ile Ile Thr Ala Ser Tyr Ala Pro Glu Gly Asp Glu Glu Ile

115 120 125

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

130 135 140

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

145 150 155 160

Phe Val Gly Glu Leu Asn Asp Lys Met Lys Gly Phe Tyr Arg Ser Lys

165 170 175

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

180 185 190

Glu Ala Thr Asp Ala Arg Arg Ala Phe Pro Cys Trp Asp Glu Pro Ala

195 200 205
Ile Lys Ala Thr Phe Asp Ile Ser Leu Val Val Pro Lys Asp Arg Val

210 215 220
Ala Leu Ser Asn Met Asn Val Ile Asp Arg Lys Pro Tyr Pro Asp Asp
225 230 235 240
Glu Asn Leu Val Glu Val Lys Phe Ala Arg Thr Pro Val Met Ser Thr
245 250 255
Tyr Leu Val Ala Phe Val Val Gly Glu Tyr Asp Phe Val Glu Thr Arg
260 265 270
Ser Lys Asp Gly Val Cys Val Arg Val Tyr Thr Pro Val Gly Lys Ala

275 280 285
Glu Gln Gly Lys Phe Ala Leu Glu Val Ala Ala Lys Thr Leu Pro Phe
290 295 300
Tyr Lys Asp Tyr Phe Asn Val Pro Tyr Pro Leu Pro Lys Ile Asp Leu
305 310 315 320
Ile Ala Ile Ala Asp Phe Ala Ala Gly Ala Met Glu Asn Trp Gly Leu
325 330 335
Val Thr Tyr Arg Glu Thr Ala Leu Leu Ile Asp Pro Lys Asn Ser Cys

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

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

Ile Ser Tyr Ser Lys Gly Ala Ser Val Ile Arg Met Leu His Asp Tyr
450 455 460

Ile Gly Asp Lys Asp Phe Lys Lys Gly Met Asn Met Tyr Leu Thr Lys

465 470 475 480

Phe Gln Gln Lys Asn Ala Ala Thr Glu Asp Leu Trp Glu Ser Leu Glu
485 490 495

Asn Ala Ser Gly Lys Pro Ile Ala Ala Val Met Asn Thr Trp Thr Lys
500 505 510

Gln Met Gly Phe Pro Leu Ile Tyr Val Glu Ala Glu Gln Val Glu Asp
515 520 525

Asp Arg Leu Leu Arg Leu Ser Gln Lys Lys Phe Cys Ala Gly Gly Ser

530 535 540

Tyr Val Gly Glu Asp Cys Pro Gln Trp Met Val Pro Ile Thr Ile Ser
545 550 555 560

Thr Ser Glu Asp Pro Asn Gln Ala Lys Leu Lys Ile Leu Met Asp Lys
565 570 575

Pro Glu Met Asn Val Val Leu Lys Asn Val Lys Pro Asp Gln Trp Val
580 585 590

Lys Leu Asn Leu Gly Thr Val Gly Phe Tyr Arg Thr Gln Tyr Ser Ser

595 600 605

Ala Met Leu Glu Ser Leu Leu Pro Gly Ile Arg Asp Leu Ser Leu Pro
610 615 620

Pro Val Asp Arg Leu Gly Leu Gln Asn Asp Leu Phe Ser Leu Ala Arg
625 630 635 640

Ala Gly Ile Ile Ser Thr Val Glu Val Leu Lys Val Met Glu Ala Phe
645 650 655

Val Asn Glu Pro Asn Tyr Thr Val Trp Ser Asp Leu Ser Cys Asn Leu

660 665 670

Gly Ile Leu Ser Thr Leu Leu Ser His Thr Asp Phe Tyr Glu Glu Ile
675 680 685

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

690 695 700
 Trp Asp Pro Lys Pro Gly Glu Gly His Leu Asp Ala Leu Leu Arg Gly
 705 710 715 720
 Leu Val Leu Gly Lys Leu Gly Lys Ala Gly His Lys Ala Thr Leu Glu

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

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

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

 915 920 925
 Pro Pro Thr Val
 930

<210> 14

<211> 932

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 14

Met Gly Ser Ser His His His His His His Ser Ser Gly Met Trp Leu

1 5 10 15

Ala Ala Ala Ala Pro Ser Leu Ala Arg Arg Leu Leu Phe Leu Gly Pro

20 25 30

Pro Pro Pro Pro Leu Leu Leu Leu Val Phe Ser Arg Ser Ser Arg Arg

35 40 45

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

50 55 60

Glu Arg Leu Pro Ala Asp Val Ser Pro Ile Asn Tyr Ser Leu Cys Leu

65 70 75 80

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

85 90 95

Ala Gln Val Arg Gln Ala Thr Asn Gln Ile Val Met Asn Cys Ala Asp

100 105 110

Ile Asp Ile Ile Thr Ala Ser Tyr Ala Pro Glu Gly Asp Glu Glu Ile

115 120 125

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

130 135 140

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

145 150 155 160

Phe Val Gly Glu Leu Asn Asp Lys Met Lys Gly Phe Tyr Arg Ser Lys

165 170 175

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

180 185 190

Glu Ala Thr Asp Ala Arg Arg Ala Phe Pro Cys Trp Asp Glu Pro Ala

195 200 205

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

450 455 460
 Ile Gly Asp Lys Asp Phe Lys Lys Gly Met Asn Met Tyr Leu Thr Lys
 465 470 475 480
 Phe Gln Gln Lys Asn Ala Ala Thr Glu Asp Leu Trp Glu Ser Leu Glu

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

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

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

 675 680 685
 Gln Glu Phe Val Lys Asp Val Phe Ser Pro Ile Gly Glu Arg Leu Gly
 690 695 700

Trp Asp Pro Lys Pro Gly Glu Gly His Leu Asp Ala Leu Leu Arg Gly
705 710 715 720
Leu Val Leu Gly Lys Leu Gly Lys Ala Gly His Lys Ala Thr Leu Glu
725 730 735
Glu Ala Arg Arg Arg Phe Lys Asp His Val Glu Gly Lys Gln Ile Leu
740 745 750
Ser Ala Asp Leu Arg Ser Pro Val Tyr Leu Thr Val Leu Lys His Gly
755 760 765
Asp Gly Thr Thr Leu Asp Ile Met Leu Lys Leu His Lys Gln Ala Asp
770 775 780
Met Gln Glu Glu Lys Asn Arg Ile Glu Arg Val Leu Gly Ala Thr Leu
785 790 795 800
Leu Pro Asp Leu Ile Gln Lys Val Leu Thr Phe Ala Leu Ser Glu Glu
805 810 815
Val Arg Pro Gln Asp Thr Val Ser Val Ile Gly Gly Val Ala Gly Gly
820 825 830
Ser Lys His Gly Arg Lys Ala Ala Trp Lys Phe Ile Lys Asp Asn Trp
835 840 845
Glu Glu Leu Tyr Asn Arg Tyr Gln Gly Gly Phe Leu Ile Ser Arg Leu
850 855 860
Ile Lys Leu Ser Val Glu Gly Phe Ala Val Asp Lys Met Ala Gly Glu
865 870 875 880
Val Lys Ala Phe Phe Glu Ser His Pro Ala Pro Ser Ala Glu Arg Thr
885 890 895
Ile Gln Gln Cys Cys Glu Asn Ile Leu Leu Asn Ala Ala Trp Leu Lys
900 905 910
Arg Asp Ala Glu Ser Ile His Gln Tyr Leu Leu Gln Arg Lys Ala Ser
915 920 925
Pro Pro Thr Val
930
<210> 15

<211> 864

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 15

Met Ile Tyr Glu Phe Val Met Thr Asp Pro Lys Ile Lys Tyr Leu Lys

1 5 10 15

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

20 25 30

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

35 40 45

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

50 55 60

Glu Leu Lys Leu Leu Ser Ile Lys Leu Asn Asn Lys His Leu Ser Pro

65 70 75 80

Ala Glu Phe Ala Val Asn Glu Asn Gln Leu Ile Ile Asn Asn Val Pro

85 90 95

Glu Lys Phe Val Leu Gln Thr Val Val Glu Ile Asn Pro Ser Ala Asn

100 105 110

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

115 120 125

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

130 135 140

Asp Val Met Ala Ala Phe Thr Val Lys Ile Ile Ala Asp Lys Lys Lys

145 150 155 160

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

165 170 175

Ser Asp Asn Gln His Phe Ala Val Trp Lys Asp Pro Phe Lys Lys Pro

180 185 190

Cys Tyr Leu Phe Ala Leu Val Ala Gly Asp Leu Ala Ser Ile Lys Asp

195 200 205

Thr Tyr Ile Thr Lys Ser Gln Arg Lys Val Ser Leu Glu Ile Tyr Ala

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

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

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

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

 435 440 445
 Tyr Ala Gln Ser Gly Thr Pro Asn Ile Lys Val Ser Glu Asn Tyr Asp
 450 455 460

Ala Ser Ser Gln Thr Tyr Ser Leu Thr Leu Glu Gln Thr Thr Leu Pro
465 470 475 480
Thr Ala Asp Gln Lys Glu Lys Gln Ala Leu His Ile Pro Val Lys Met
485 490 495
Gly Leu Ile Asn Pro Glu Gly Lys Asn Ile Ala Glu Gln Val Ile Glu
500 505 510
Leu Lys Glu Gln Lys Gln Thr Tyr Thr Phe Glu Asn Ile Ala Ala Lys
515 520 525
Pro Val Ala Ser Leu Phe Arg Asp Phe Ser Ala Pro Val Lys Val Glu
530 535 540
His Lys Arg Ser Glu Lys Asp Leu Leu His Ile Val Lys Tyr Asp Asn
545 550 555 560
Asn Ala Phe Asn Arg Trp Asp Ser Leu Gln Gln Ile Ala Thr Asn Ile
565 570 575
Ile Leu Asn Asn Ala Asp Leu Asn Asp Glu Phe Leu Asn Ala Phe Lys
580 585 590
Ser Ile Leu His Asp Lys Asp Leu Asp Lys Ala Leu Ile Ser Asn Ala
595 600 605
Leu Leu Ile Pro Ile Glu Ser Thr Ile Ala Glu Ala Met Arg Val Ile
610 615 620
Met Val Asp Asp Ile Val Leu Ser Arg Lys Asn Val Val Asn Gln Leu
625 630 635 640
Ala Asp Lys Leu Lys Asp Asp Trp Leu Ala Val Tyr Gln Gln Cys Asn
645 650 655
Asp Asn Lys Pro Tyr Ser Leu Ser Ala Glu Gln Ile Ala Lys Arg Lys
660 665 670
Leu Lys Gly Val Cys Leu Ser Tyr Leu Met Asn Ala Ser Asp Gln Lys
675 680 685
Val Gly Thr Asp Leu Ala Gln Gln Leu Phe Asp Asn Ala Asp Asn Met
690 695 700
Thr Asp Gln Gln Thr Ala Phe Thr Glu Leu Leu Lys Ser Asn Asp Lys

705 710 715 720
 Gln Val Arg Asp Asn Ala Ile Asn Glu Phe Tyr Asn Arg Trp Arg His
 725 730 735
 Glu Asp Leu Val Val Asn Lys Trp Leu Leu Ser Gln Ala Gln Ile Ser
 740 745 750
 His Glu Ser Ala Leu Asp Ile Val Lys Gly Leu Val Asn His Pro Ala

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

 820 825 830
 Asp Arg Gln Ala Met Met Lys Asn Ala Leu Glu Lys Ile Lys Ala Ser
 835 840 845
 Asn Pro Ser Lys Asn Val Phe Glu Ile Val Ser Lys Ser Leu Glu Ser
 850 855 860
 <210> 16
 <211> 366
 <212> PRT
 <213> Artificial Sequence
 <220><223> Synthetic
 <400> 16
 Met Gly Ser Ser His His His His His His Ser Ser Gly Met Glu Val
 1 5 10 15

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

50	55	60	
Val Ile Ala His Lys Lys Gly Glu Gly Pro Lys Val Met Ile Ala Ala			
65	70	75	80
His Met Asp Gln Ile Gly Leu Met Val Thr His Ile Glu Lys Asn Gly			
	85	90	95
Phe Leu Arg Val Ala Pro Ile Gly Gly Val Asp Pro Lys Thr Leu Ile			
	100	105	110
Ala Gln Arg Phe Lys Val Trp Ile Asp Lys Gly Lys Phe Ile Tyr Gly			
	115	120	125
Val Gly Ala Ser Val Pro Pro His Ile Gln Lys Pro Glu Asp Arg Lys			
130	135	140	
Lys Ala Pro Asp Trp Asp Gln Ile Phe Ile Asp Ile Gly Ala Glu Ser			
145	150	155	160
Lys Glu Glu Ala Glu Asp Met Gly Val Lys Ile Gly Thr Val Ile Thr			
	165	170	175
Trp Asp Gly Arg Leu Glu Arg Leu Gly Lys His Arg Phe Val Ser Ile			
	180	185	190
Ala Phe Asp Asp Arg Ile Ala Val Tyr Thr Ile Leu Glu Val Ala Lys			
195	200	205	
Gln Leu Lys Asp Ala Lys Ala Asp Val Tyr Phe Val Ala Thr Val Gln			
210	215	220	
Glu Glu Val Gly Leu Arg Gly Ala Arg Thr Ser Ala Phe Gly Ile Glu			
225	230	235	240
Pro Asp Tyr Gly Phe Ala Ile Asp Val Thr Ile Ala Ala Asp Ile Pro			
	245	250	255
Gly Thr Pro Glu His Lys Gln Val Thr His Leu Gly Lys Gly Thr Ala			
260	265	270	
Ile Lys Ile Met Asp Arg Ser Val Ile Cys His Pro Thr Ile Val Arg			
275	280	285	
Trp Leu Glu Glu Leu Ala Lys Lys His Glu Ile Pro Tyr Gln Leu Glu			
290	295	300	

Ile Leu Leu Gly Gly Gly Thr Asp Ala Gly Ala Ile His Leu Thr Lys
 305 310 315 320
 Ala Gly Val Pro Thr Gly Ala Leu Ser Val Pro Ala Arg Tyr Ile His
 325 330 335

Ser Asn Thr Glu Val Val Asp Glu Arg Asp Val Asp Ala Thr Val Glu
 340 345 350
 Leu Met Thr Lys Ala Leu Glu Asn Ile His Glu Leu Lys Ile
 355 360 365

<210> 17

<211> 408

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 17

Met Asp Ala Phe Thr Glu Asn Leu Asn Lys Leu Ala Glu Leu Ala Ile
 1 5 10 15
 Arg Val Gly Leu Asn Leu Glu Glu Gly Gln Glu Ile Val Ala Thr Ala

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

Glu Asn Gly Ala Ser Leu Phe Thr Val Leu Tyr Gly Asp Asn Leu Ile
 50 55 60

Ala Arg Lys Arg Leu Ala Leu Val Pro Glu Ala His Leu Asp Arg Ala
 65 70 75 80

Pro Ala Trp Leu Tyr Glu Gly Met Ala Lys Ala Phe His Glu Gly Ala

85 90 95
 Ala Arg Leu Ala Val Ser Gly Asn Asp Pro Lys Ala Leu Glu Gly Leu
 100 105 110

Pro Pro Glu Arg Val Gly Arg Ala Gln Gln Ala Gln Ser Arg Ala Tyr
 115 120 125

Arg Pro Thr Leu Ser Ala Ile Thr Glu Phe Val Thr Asn Trp Thr Ile
 130 135 140

Val Pro Phe Ala His Pro Gly Trp Ala Lys Ala Val Phe Pro Gly Leu

145 150 155 160

Pro Glu Glu Glu Ala Val Gln Arg Leu Trp Gln Ala Ile Phe Gln Ala

165 170 175

Thr Arg Val Asp Gln Glu Asp Pro Val Ala Ala Trp Glu Ala His Asn

180 185 190

Arg Val Leu His Ala Lys Val Ala Phe Leu Asn Glu Lys Arg Phe His

195 200 205

Ala Leu His Phe Gln Gly Pro Gly Thr Asp Leu Thr Val Gly Leu Ala

210 215 220

Glu Gly His Leu Trp Gln Gly Gly Ala Thr Pro Thr Lys Lys Gly Arg

225 230 235 240

Leu Cys Asn Pro Asn Leu Pro Thr Glu Glu Val Phe Thr Ala Pro His

245 250 255

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

260 265 270

Ser Gly Gln Leu Val Glu Gly Leu Trp Ala Arg Phe Glu Gly Gly Val

275 280 285

Ala Val Glu Val Gly Ala Glu Lys Gly Glu Glu Val Leu Lys Lys Leu

290 295 300

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

305 310 315 320

Pro Ala Asp Asn Pro Ile Ala Lys Thr Gly Leu Val Phe Phe Asp Thr

325 330 335

Leu Phe Asp Glu Asn Ala Ala Ser His Ile Ala Phe Gly Gln Ala Tyr

340 345 350

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

355 360 365

Gly Gly Asn Glu Ser Met Val His Val Asp Trp Met Ile Gly Ser Glu

370 375 380

Glu Val Asp Val Asp Gly Leu Leu Glu Asp Gly Thr Arg Val Pro Leu

385					390					395					400
Met	Arg	Arg	Gly	Arg	Trp	Val	Ile								
405															
<210>	18														
<211>	362														
<212>	PRT														
<213>	Artificial Sequence														
<220><223>	Synthetic														
<400>	18														
Met	Ala	Lys	Leu	Asp	Glu	Thr	Leu	Thr	Met	Leu	Lys	Ala	Leu	Thr	Asp
1	5				10				15						
Ala	Lys	Gly	Val	Pro	Gly	Asn	Glu	Arg	Glu	Ala	Arg	Asp	Val	Met	Lys
20				25				30							
Thr	Tyr	Ile	Ala	Pro	Tyr	Ala	Asp	Glu	Val	Thr	Thr	Asp	Gly	Leu	Gly
35				40				45							
Ser	Leu	Ile	Ala	Lys	Lys	Glu	Gly	Lys	Ser	Gly	Gly	Pro	Lys	Val	Met
50				55				60							
Ile	Ala	Gly	His	Leu	Asp	Glu	Val	Gly	Phe	Met	Val	Thr	Gln	Ile	Asp
65	70				75				80						
Asp	Lys	Gly	Phe	Ile	Arg	Phe	Gln	Thr	Leu	Gly	Gly	Trp	Trp	Ser	Gln
85				90				95							
Val	Met	Leu	Ala	Gln	Arg	Val	Thr	Ile	Val	Thr	Lys	Lys	Gly	Asp	Ile
100				105				110							
Thr	Gly	Val	Ile	Gly	Ser	Lys	Pro	Pro	His	Ile	Leu	Pro	Ser	Glu	Ala
115				120				125							
Arg	Lys	Lys	Pro	Val	Glu	Ile	Lys	Asp	Met	Phe	Ile	Asp	Ile	Gly	Ala
130				135				140							
Thr	Ser	Arg	Glu	Glu	Ala	Met	Glu	Trp	Gly	Val	Arg	Pro	Gly	Asp	Met
145	150				155				160						
Ile	Val	Pro	Tyr	Phe	Glu	Phe	Thr	Val	Leu	Asn	Asn	Glu	Lys	Met	Leu
165				170				175							

Leu Ala Lys Ala Trp Asp Asn Arg Ile Gly Cys Ala Val Ala Ile Asp

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

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

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

<210> 19

<211> 490

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400>

19

Met Glu Asp Lys Val Trp Ile Ser Met Gly Ala Asp Ala Val Gly Ser

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

 Val His Thr Ser Ala Gln Gly Ala Met Ala Ala Leu Met Met Pro Glu
 65 70 75 80
 Ser Ile Ala Asn Phe Thr Ile Pro Ala Pro Ser Gln Gln Asp Leu Val
 85 90 95
 Asn Ala Trp Leu Pro Gln Val Ser Ala Asp Gln Ile Thr Asn Thr Ile
 100 105 110
 Arg Ala Leu Ser Ser Phe Asn Asn Arg Phe Tyr Thr Thr Thr Ser Gly
 115 120 125

 Ala Gln Ala Ser Asp Trp Leu Ala Asn Glu Trp Arg Ser Leu Ile Ser
 130 135 140
 Ser Leu Pro Gly Ser Arg Ile Glu Gln Ile Lys His Ser Gly Tyr Asn
 145 150 155 160
 Gln Lys Ser Val Val Leu Thr Ile Gln Gly Ser Glu Lys Pro Asp Glu
 165 170 175
 Trp Val Ile Val Gly Gly His Leu Asp Ser Thr Leu Gly Ser His Thr
 180 185 190

 Asn Glu Gln Ser Ile Ala Pro Gly Ala Asp Asp Asp Ala Ser Gly Ile
 195 200 205
 Ala Ser Leu Ser Glu Ile Ile Arg Val Leu Arg Asp Asn Asn Phe Arg
 210 215 220
 Pro Lys Arg Ser Val Ala Leu Met Ala Tyr Ala Ala Glu Glu Val Gly
 225 230 235 240
 Leu Arg Gly Ser Gln Asp Leu Ala Asn Gln Tyr Lys Ala Gln Gly Lys
 245 250 255

Lys Val Val Ser Val Leu Gln Leu Asp Met Thr Asn Tyr Arg Gly Ser
260 265 270
Ala Glu Asp Ile Val Phe Ile Thr Asp Tyr Thr Asp Ser Asn Leu Thr
275 280 285
Gln Phe Leu Thr Thr Leu Ile Asp Glu Tyr Leu Pro Glu Leu Thr Tyr
290 295 300
Gly Tyr Asp Arg Cys Gly Tyr Ala Cys Ser Asp His Ala Ser Trp His
305 310 315 320

Lys Ala Gly Phe Ser Ala Ala Met Pro Phe Glu Ser Lys Phe Lys Asp
325 330 335
Tyr Asn Pro Lys Ile His Thr Ser Gln Asp Thr Leu Ala Asn Ser Asp
340 345 350
Pro Thr Gly Asn His Ala Val Lys Phe Thr Lys Leu Gly Leu Ala Tyr
355 360 365
Val Ile Glu Met Ala Asn Ala Gly Ser Ser Gln Val Pro Asp Asp Ser
370 375 380

Val Leu Gln Asp Gly Thr Ala Lys Ile Asn Leu Ser Gly Ala Arg Gly
385 390 395 400
Thr Gln Lys Arg Phe Thr Phe Glu Leu Ser Gln Ser Lys Pro Leu Thr
405 410 415
Ile Gln Thr Tyr Gly Gly Ser Gly Asp Val Asp Leu Tyr Val Lys Tyr
420 425 430
Gly Ser Ala Pro Ser Lys Ser Asn Trp Asp Cys Arg Pro Tyr Gln Asn
435 440 445

Gly Asn Arg Glu Thr Cys Ser Phe Asn Asn Ala Gln Pro Gly Ile Tyr
450 455 460
His Val Met Leu Asp Gly Tyr Thr Asn Tyr Asn Asp Val Ala Leu Lys
465 470 475 480
Ala Ser Thr Gln His His His His His
485 490

<210> 20

<211> 494

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 20

Met Glu Asp Lys Val Trp Ile Ser Ile Gly Ser Asp Ala Ser Gln Thr

1 5 10 15

Val Lys Ser Val Met Gln Ser Asn Ala Arg Ser Leu Leu Pro Glu Ser

20 25 30

Leu Ala Ser Asn Gly Pro Val Trp Val Gly Gln Val Asp Tyr Ser Gln

35 40 45

Leu Ala Glu Leu Ser His His Met His Glu Asp His Gln Arg Cys Gly

50 55 60

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

65 70 75 80

Met Pro Gln Ser Leu Val Ala Phe Ser Ile Pro Glu Ile Ser Gln Gln

85 90 95

Asp Thr Val Asn Ala Trp Leu Pro Gln Val Asn Ser Gln Ala Ile Thr

100 105 110

Gly Thr Ile Thr Ser Leu Thr Ser Phe Ile Asn Arg Phe Tyr Thr Thr

115 120 125

Thr Ser Gly Ala Gln Ala Ser Asp Trp Leu Ala Asn Glu Trp Arg Ser

130 135 140

Leu Ser Ala Ser Leu Pro Asn Ala Ser Val Arg Gln Val Ser His Phe

145 150 155 160

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

165 170 175

Pro Asp Glu Trp Ile Val Leu Gly Gly His Leu Asp Ser Thr Ile Gly

180 185 190

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

195 200 205

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

210 215 220
 Asn Phe Gln Pro Lys Arg Ser Ile Ala Phe Met Ala Tyr Ala Ala Glu
 225 230 235 240
 Glu Val Gly Leu Arg Gly Ser Gln Asp Leu Ala Asn Gln Tyr Lys Ala
 245 250 255
 Glu Gly Lys Gln Val Ile Ser Ala Leu Gln Leu Asp Met Thr Asn Tyr

 260 265 270
 Lys Gly Ser Val Glu Asp Ile Val Phe Ile Thr Asp Tyr Thr Asp Ser
 275 280 285
 Asn Leu Thr Thr Phe Leu Ser Gln Leu Val Asp Glu Tyr Leu Pro Ser
 290 295 300
 Leu Thr Tyr Gly Phe Asp Thr Cys Gly Tyr Ala Cys Ser Asp His Ala
 305 310 315 320
 Ser Trp His Lys Ala Gly Phe Ser Ala Ala Met Pro Phe Glu Ala Lys

 325 330 335
 Phe Asn Asp Tyr Asn Pro Met Ile His Thr Pro Asn Asp Thr Leu Gln
 340 345 350
 Asn Ser Asp Pro Thr Ala Ser His Ala Val Lys Phe Thr Lys Leu Gly
 355 360 365
 Leu Ala Tyr Ala Ile Glu Met Ala Ser Thr Thr Gly Gly Thr Pro Pro
 370 375 380
 Pro Thr Gly Asn Val Leu Lys Asp Gly Val Pro Val Asn Gly Leu Ser

 385 390 395 400
 Gly Ala Thr Gly Ser Gln Val His Tyr Ser Phe Glu Leu Pro Ala Gln
 405 410 415
 Lys Asn Leu Gln Ile Ser Thr Ala Gly Gly Ser Gly Asp Val Asp Leu
 420 425 430
 Tyr Val Ser Phe Gly Ser Glu Ala Thr Lys Gln Asn Trp Asp Cys Arg
 435 440 445
 Pro Tyr Arg Asn Gly Asn Asn Glu Val Cys Thr Phe Ala Gly Ala Thr

 450 455 460

Pro Gly Thr Tyr Ser Ile Met Leu Asp Gly Tyr Arg Gln Phe Ser Gly
465 470 475 480

Val Thr Leu Lys Ala Ser Thr Gln His His His His His His
485 490

<210> 21

<211> 877

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 21

Met Thr Gln Gln Pro Gln Ala Lys Tyr Arg His Asp Tyr Arg Ala Pro
1 5 10 15

Asp Tyr Thr Ile Thr Asp Ile Asp Leu Asp Phe Ala Leu Asp Ala Gln
20 25 30

Lys Thr Thr Val Thr Ala Val Ser Lys Val Lys Arg Gln Gly Thr Asp
35 40 45

Val Thr Pro Leu Ile Leu Asn Gly Glu Asp Leu Thr Leu Ile Ser Val
50 55 60

Ser Val Asp Gly Gln Ala Trp Pro His Tyr Arg Gln Gln Asp Asn Thr
65 70 75 80

Leu Val Ile Glu Gln Leu Pro Ala Asp Phe Thr Leu Thr Ile Val Asn
85 90 95

Asp Ile His Pro Ala Thr Asn Ser Ala Leu Glu Gly Leu Tyr Leu Ser
100 105 110

Gly Glu Ala Leu Cys Thr Gln Cys Glu Ala Glu Gly Phe Arg His Ile
115 120 125

Thr Tyr Tyr Leu Asp Arg Pro Asp Val Leu Ala Arg Phe Thr Thr Arg
130 135 140

Ile Val Ala Asp Lys Ser Arg Tyr Pro Tyr Leu Leu Ser Asn Gly Asn
145 150 155 160

Arg Val Gly Gln Gly Glu Leu Asp Asp Gly Arg His Trp Val Lys Trp
165 170 175

Glu Asp Pro Phe Pro Lys Pro Ser Tyr Leu Phe Ala Leu Val Ala Gly
180 185 190

Asp Phe Asp Val Leu Gln Asp Lys Phe Ile Thr Arg Ser Gly Arg Glu
195 200 205

Val Ala Leu Glu Ile Phe Val Asp Arg Gly Asn Leu Asp Arg Ala Asp
210 215 220

Trp Ala Met Thr Ser Leu Lys Asn Ser Met Lys Trp Asp Glu Thr Arg
225 230 235 240

Phe Gly Leu Glu Tyr Asp Leu Asp Ile Tyr Met Ile Val Ala Val Asp
245 250 255

Phe Phe Asn Met Gly Ala Met Glu Asn Lys Gly Leu Asn Val Phe Asn
260 265 270

Ser Lys Tyr Val Leu Ala Lys Ala Glu Thr Ala Thr Asp Lys Asp Tyr
275 280 285

Leu Asn Ile Glu Ala Val Ile Gly His Glu Tyr Phe His Asn Trp Thr
290 295 300

Gly Asn Arg Val Thr Cys Arg Asp Trp Phe Gln Leu Ser Leu Lys Glu
305 310 315 320

Gly Leu Thr Val Phe Arg Asp Gln Glu Phe Ser Ser Asp Leu Gly Ser
325 330 335

Arg Ser Val Asn Arg Ile Glu Asn Val Arg Val Met Arg Ala Ala Gln
340 345 350

Phe Ala Glu Asp Ala Ser Pro Met Ala His Ala Ile Arg Pro Asp Lys
355 360 365

Val Ile Glu Met Asn Asn Phe Tyr Thr Leu Thr Val Tyr Glu Lys Gly
370 375 380

Ser Glu Val Ile Arg Met Met His Thr Leu Leu Gly Glu Gln Gln Phe
385 390 395 400

Gln Ala Gly Met Arg Leu Tyr Phe Glu Arg His Asp Gly Ser Ala Ala
405 410 415

Thr Cys Asp Asp Phe Val Gln Ala Met Glu Asp Val Ser Asn Val Asp

420 425 430
 Leu Ser Leu Phe Arg Arg Trp Tyr Ser Gln Ser Gly Thr Pro Leu Leu
 435 440 445
 Thr Val His Asp Asp Tyr Asp Val Glu Lys Gln Gln Tyr His Leu Phe
 450 455 460

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

 Ile Pro Ser Leu Leu Arg Glu Phe Ser Ala Pro Val Lys Leu Asp Tyr
 530 535 540
 Pro Tyr Ser Asp Gln Gln Leu Thr Phe Leu Met Gln His Ala Arg Asn
 545 550 555 560
 Glu Phe Ser Arg Trp Asp Ala Ala Gln Ser Leu Leu Ala Thr Tyr Ile
 565 570 575
 Lys Leu Asn Val Ala Lys Tyr Gln Gln Gln Gln Pro Leu Ser Leu Pro
 580 585 590

 Ala His Val Ala Asp Ala Phe Arg Ala Ile Leu Leu Asp Glu His Leu
 595 600 605
 Asp Pro Ala Leu Ala Ala Gln Ile Leu Thr Leu Pro Ser Glu Asn Glu
 610 615 620
 Met Ala Glu Leu Phe Thr Thr Ile Asp Pro Gln Ala Ile Ser Thr Val
 625 630 635 640
 His Glu Ala Ile Thr Arg Cys Leu Ala Gln Glu Leu Ser Asp Glu Leu
 645 650 655

 Leu Ala Val Tyr Val Ala Asn Met Thr Pro Val Tyr Arg Ile Glu His
 660 665 670

Gly Asp Ile Ala Lys Arg Ala Leu Arg Asn Thr Cys Leu Asn Tyr Leu
675 680 685
Ala Phe Gly Asp Glu Glu Phe Ala Asn Lys Leu Val Ser Leu Gln Tyr
690 695 700
His Gln Ala Asp Asn Met Thr Asp Ser Leu Ala Ala Leu Ala Ala Ala
705 710 715 720

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

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

Glu Gln Leu Lys Thr Leu Asp Asn Leu Ser Gly Asp Leu Tyr Glu Lys
850 855 860
Ile Thr Lys Ala Leu Ala Ala His His His His His His
865 870 875

<210> 22

<211> 489

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 22

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

245 250 255
 Lys Val Met Ser Val Met Gln Leu Asp Met Thr Asn Tyr Gln Gly Ser
 260 265 270
 Arg Glu Asp Ile Val Phe Ile Thr Asp Tyr Thr Asp Ser Asn Phe Thr

 275 280 285
 Gln Tyr Leu Thr Gln Leu Leu Asp Glu Tyr Leu Pro Ser Leu Thr Tyr
 290 295 300
 Gly Phe Asp Thr Cys Gly Tyr Ala Cys Ser Asp His Ala Ser Trp His
 305 310 315 320
 Ala Val Gly Tyr Pro Ala Ala Met Pro Phe Glu Ser Lys Phe Asn Asp
 325 330 335
 Tyr Asn Pro Asn Ile His Ser Pro Gln Asp Thr Leu Gln Asn Ser Asp

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

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

 465 470 475 480
 Asn Ala Gln His His His His His His
 485

<210> 23

<211> 488

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 23

Met Glu Asp Lys Lys Val Trp Ile Ser Ile Gly Ala Asp Ala Gln Gln

1 5 10 15

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

20 25 30

His Asn Gly Gln Ala Trp Ile Gly Glu Val Ser Glu Ser Glu Leu Ala

35 40 45

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

50 55 60

Ile Val His Ser Ser Ala Gln Ser Ala Met Ala Ala Ser Asn Met Pro

65 70 75 80

Leu Ser Arg Ala Ser Phe Ile Ala Pro Ala Ile Ser Gln Gln Ala Leu

85 90 95

Val Thr Pro Trp Ile Ser Gln Ile Asp Ser Ala Leu Ile Val Asn Thr

100 105 110

Ile Asp Arg Leu Thr Asp Phe Pro Asn Arg Phe Tyr Thr Thr Thr Ser

115 120 125

Gly Ala Gln Ala Ser Asp Trp Ile Lys Gln Arg Trp Gln Ser Leu Ser

130 135 140

Ala Gly Leu Ala Gly Ala Ser Val Thr Gln Ile Ser His Ser Gly Tyr

145 150 155 160

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

165 170 175

Glu Trp Val Val Val Gly Gly His Leu Asp Ser Thr Ile Gly Ser Arg

180 185 190

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

195 200 205

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

450 455 460
Ala Val Met Leu Arg Gly Tyr Arg Thr Tyr Asp Asn Val Ser Leu Ile
465 470 475 480
Val Glu His His His His His His
485

<210> 24

<211> 308

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 24

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

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

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

165 170 175
Leu Gln Leu Asp Met Thr Asn Tyr Lys Gly Ser Ala Gln Asp Val Val

180 185 190
Phe Ile Thr Asp Tyr Thr Asp Ser Asn Phe Thr Gln Tyr Leu Thr Gln

195 200 205
Leu Met Asp Glu Tyr Leu Pro Ser Leu Thr Tyr Gly Phe Asp Thr Cys

210 215 220
Gly Tyr Ala Cys Ser Asp His Ala Ser Trp His Asn Ala Gly Tyr Pro

225 230 235 240
Ala Ala Met Pro Phe Glu Ser Lys Phe Asn Asp Tyr Asn Pro Arg Ile

245 250 255
His Thr Thr Gln Asp Thr Leu Ala Asn Ser Asp Pro Thr Gly Ser His

260 265 270
Ala Lys Lys Phe Thr Gln Leu Gly Leu Ala Tyr Ala Ile Glu Met Gly

275 280 285
Ser Ala Thr Gly Asp Thr Pro Thr Pro Gly Asn Gln Leu Glu His His

290 295 300

His His His His

305

<210> 25

<211> 354

<212> PRT

<213> Artificial Sequence

<220><223> Synthetic

<400> 25

Met Val Asp Trp Glu Leu Met Lys Lys Ile Ile Glu Ser Pro Gly Val

1 5 10 15

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

20 25 30

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

35 40 45

Ala His Phe Lys Gly Ser Ala Pro Lys Val Met Val Ala Ala His Met

50

55

60

Asp Lys Ile Gly Leu Met Val Asn His Ile Asp Lys Asp Gly Tyr Leu

65

70

75

80

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

85

90

95

Lys Ile Arg Phe Phe Thr Glu Lys Gly Glu Arg Tyr Gly Val Val Gly

100

105

110

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

115

120

125

Ile Asp Trp Asp Ser Ile Ile Val Asp Val Gly Ala Ser Ser Arg Glu

130

135

140

Glu Ala Glu Glu Met Gly Phe Arg Ile Gly Thr Ile Gly Glu Phe Ala

145

150

155

160

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

165

170

175

Asp Asp Arg Ile Cys Leu Tyr Ala Met Ile Glu Ala Ala Arg Gln Leu

180

185

190

Gly Glu His Glu Ala Asp Ile Tyr Ile Val Ala Ser Val Gln Glu Glu

195

200

205

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

210

215

220

Val Gly Ile Ala Met Asp Val Thr Phe Ala Lys Gln Pro Asn Asp Lys

225

230

235

240

Gly Lys Ile Val Pro Glu Leu Gly Lys Gly Pro Val Met Asp Val Gly

245

250

255

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

260

265

270

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

275

280

285

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

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