



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0045449
(43) 공개일자 2021년04월26일

(51) 국제특허분류(Int. Cl.)
G01N 33/68 (2006.01)
(52) CPC특허분류
G01N 33/6848 (2013.01)
(21) 출원번호 10-2021-7007865
(22) 출원일자(국제) 2019년08월16일
심사청구일자 없음
(85) 번역문제출일자 2021년03월16일
(86) 국제출원번호 PCT/US2019/046821
(87) 국제공개번호 WO 2020/037205
국제공개일자 2020년02월20일
(30) 우선권주장
62/719,292 2018년08월17일 미국(US)

(71) 출원인
리제너론 파아마슈티컬스, 인크.
미국 뉴욕 10591-6707 테리타운 올드 소우 밀 리버 로드 777
(72) 발명자
마오 위안
미국 뉴욕주 10591 테리타운 올드 소우 밀 리버 로드 777 리제너론 파아마슈티컬스 인크. 내
(74) 대리인
장훈

전체 청구항 수 : 총 20 항

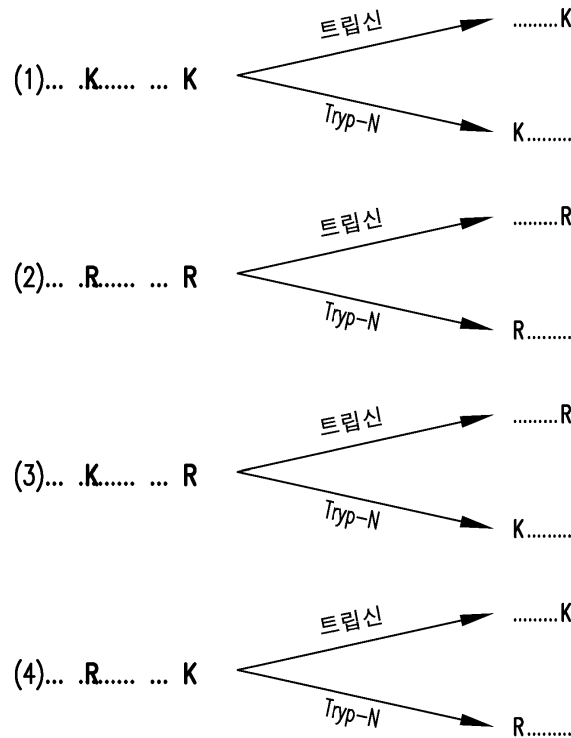
(54) 발명의 명칭 새로운 단백질 서열분석 방법

(57) 요약

폴리펩타이드의 아미노산 서열을 결정하는 방법이 제공되고, 상기 방법은 상기 폴리펩타이드를 함유하는 첫 번째 및 두 번째 샘플을 첫 번째 프로테아제(예를 들어, 트립신) 및 두 번째 프로테아제(예를 들어, Tryp-N)와 각각 접촉시켜 소화된 펩타이드 단편의 첫 번째 및 두 번째 세트를 생성하는 단계; 상기 소화된 펩타이드 단편의 세트

(뒷면에 계속)

대표도 - 도3



들을 단편화하여 단편화 펩타이드 이온의 세트들을 생성하는 단계; 상기 아르기닌 또는 라이신 아미노산 잔기의 질량에 의해 질량이 상이한 펩타이드 이온의 세트들로부터 펩타이드 이온의 쌍을 선택하는 단계; 이온 유형(N-말단 또는 C-말단 펩타이드 이온)을 상기 2개의 세트로부터의 펩타이드 이온의 선택된 쌍에 할당하는 단계; 아미노산 잔기(들)의 질량에 의한 증분 질량을 갖는 펩타이드 이온의 하나의 세트에서 동일한 유형의 펩타이드 이온의 질량 래더를 선택하는 단계, 및 상기 질량 래더로부터의 상기 확인된 아미노산 잔기를 조립하여 관심 폴리펩타이드의 아미노산 서열을 결정하는 단계를 포함한다.

명세서

청구범위

청구항 1

관심 폴리펩타이드의 아미노산 서열을 결정하는 방법으로서,

첫 번째 프로테아제가 관심 폴리펩타이드를 소화하여 소화된 펩타이드 단편의 첫 번째 세트를 생성하도록 허용하는 조건 하에, 상기 관심 폴리펩타이드를 함유하는 첫 번째 샘플을 염기성 아미노산 후에 펩타이드 결합을 절단하는 첫 번째 프로테아제와 접촉시키는 단계;

상기 소화된 펩타이드 단편의 첫 번째 세트를 단편화하여 상기 소화된 펩타이드 단편의 첫 번째 세트의 펩타이드에 상응하는 단편화 펩타이드 이온의 첫 번째 세트를 생성하는 단계;

상기 단편화 펩타이드 이온의 첫 번째 세트의 질량을 결정하는 단계;

상기 두 번째 프로테아제가 관심 폴리펩타이드를 소화하여 소화된 펩타이드 단편의 두 번째 세트를 생성하도록 허용하는 조건 하에, 상기 관심 폴리펩타이드를 함유하는 두 번째 샘플을 염기성 아미노산 전에 펩타이드 결합을 절단하는 두 번째 프로테아제와 접촉시키는 단계;

상기 소화된 펩타이드 단편의 두 번째 세트를 단편화하여 상기 소화된 펩타이드 단편의 두 번째 세트의 펩타이드에 상응하는 단편화 펩타이드 이온의 두 번째 세트를 생성하는 단계;

상기 단편화 펩타이드 이온의 두 번째 세트의 질량을 결정하는 단계;

아르기닌 아미노산 잔기의 질량 또는 라이신 아미노산 잔기의 질량에 의해 질량이 상이한 상기 단편화 펩타이드 이온의 첫 번째 세트 및 상기 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍을 선택하는 단계;

상기 펩타이드 이온의 쌍에 대한 이온 유형을 할당하여 동일한 유형의 펩타이드 이온의 목록을 생성하는 단계;

상기 아미노산 잔기(들)의 질량에 의한 증분 질량을 갖는 펩타이드 이온의 질량 리더를 선택하여 개별 아미노산 잔기를 확인하고 상기 확인된 아미노산 잔기를 조합하여 관심 폴리펩타이드의 아미노산 서열을 결정하는 단계

를 포함하는, 방법.

청구항 2

청구항 1에 있어서, 상기 첫 번째 프로테아제는 트립신 프로테아제인, 방법.

청구항 3

청구항 1 또는 2에 있어서, 상기 두 번째 프로테아제는 Tryp-N 프로테아제인, 방법.

청구항 4

청구항 1 내지 3 중 어느 한 항에 있어서,

상기 소화된 펩타이드 단편의 첫 번째 세트로부터 첫 번째 소화된 펩타이드 단편을 선택하는 단계; 및 상기 첫 번째 소화된 펩타이드 단편과 동일한 질량을 갖는 소화된 펩타이드 단편의 두 번째 세트로부터 두 번째 소화된 펩타이드 단편을 선택하는 단계를 추가로 포함하는, 방법.

청구항 5

청구항 1 내지 4 중 어느 한 항에 있어서,

상기 소화된 펩타이드 단편의 첫 번째 세트로부터 첫 번째 소화된 펩타이드 단편을 선택하는 단계; 및

상기 첫 번째 소화된 펩타이드 단편에 대한 라이신 아미노산 잔기와 아르기닌 아미노산 잔기 사이의 질량 차이와 동일한 질량 차이를 갖는 질량을 갖는 상기 소화된 펩타이드 단편의 두 번째 세트로부터 두 번째 소화된 펩

타이드 단편을 선택하는 단계
를 추가로 포함하는, 방법.

청구항 6

청구항 1 내지 5 중 어느 한 항에 있어서, 상기 소화된 펩타이드 단편의 첫 번째 세트로부터의 첫 번째 소화된 펩타이드 단편은 단편화되어 첫 번째 소화된 펩타이드 단편에 상응하는 단편화 펩타이드 이온의 첫 번째 시리즈를 생성하고, 그리고 상기 첫 번째 소화된 펩타이드 단편에 상응하는 소화된 펩타이드 단편의 두 번째 세트로부터의 두 번째 소화된 펩타이드 단편은 단편화되어 상기 두 번째 소화된 펩타이드 단편에 상응하는 단편화 펩타이드 이온의 두 번째 시리즈를 생성하고; 그리고 아미노산 서열을 유도하기 위해 펩타이드 이온의 쌍을 할당하는 것은,

상기 펩타이드 이온의 쌍에 대한 이온 유형을 할당하여 동일한 유형의 펩타이드 이온의 목록을 생성하는 단계;

아미노산 잔기(들)의 질량에 의한 증분 질량을 갖는 펩타이드 이온의 질량 래터를 선택하여 상기 목록에 대한 개별 아미노산 잔기를 확인하는 단계; 및

상기 확인된 아미노산 잔기를 조립하여 관심 폴리펩타이드의 아미노산 서열을 결정하는 단계

를 포함하는, 방법.

청구항 7

청구항 1 내지 6 중 어느 한 항에 있어서, 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍은, 아르기닌 아미노산 잔기의 질량에 의해 질량이 상이한 것이 선택되는, 방법.

청구항 8

청구항 7에 있어서, 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍 사이의 아르기닌 아미노산 잔기의 음의 질량 차이는, 상기 펩타이드가 N-말단 아르기닌 잔기를 갖는다는 것을 나타내는, 방법.

청구항 9

청구항 7에 있어서, 상기 단편화 펩타이드 이온의 첫 번째 세트 및 상기 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍 사이의 아르기닌 아미노산 잔기의 양의 질량 차이는, 상기 펩타이드가 C-말단 아르기닌 잔기를 갖는다는 것을 나타내는, 방법.

청구항 10

청구항 1 내지 6 중 어느 한 항에 있어서, 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍은, 상기 라이신 아미노산 잔기의 질량에 의해 질량이 상이한 것이 선택되는, 방법.

청구항 11

청구항 10에 있어서, 상기 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍 사이의 라이신 아미노산 잔기의 음의 질량 차이는, 상기 펩타이드가 N-말단 라이신 잔기를 갖는다는 것을 나타내는, 방법.

청구항 12

청구항 10에 있어서, 상기 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍 사이의 라이신 아미노산 잔기의 양의 질량 차이는, 상기 펩타이드가 C-말단 라이신 잔기를 갖는다는 것을 나타내는, 방법.

청구항 13

청구항 1 내지 12 중 어느 한 항에 있어서, 상기 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화

펩타이드 이온은 상기 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화 펩타이드 이온에 상응하는, 방법.

청구항 14

청구항 1 내지 13 중 어느 한 항에 있어서, 상기 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화 펩타이드 이온은 b 이온이고, 그리고 상기 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화 펩타이드 이온은 아르기닌 아미노산 잔기의 질량 또는 라이신 아미노산 잔기의 질량의 차이를 갖는 b 이온인, 방법.

청구항 15

청구항 1 내지 13 중 어느 한 항에 있어서, 상기 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화 펩타이드 이온은 y 이온이고, 그리고 상기 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화 펩타이드 이온은 아르기닌 아미노산 잔기의 질량 또는 라이신 아미노산 잔기의 질량의 차이를 갖는 y 이온인, 방법.

청구항 16

청구항 1 내지 15 중 어느 한 항에 있어서, 상기 질량은 질량 분광분석법을 사용하여 결정되는, 방법.

청구항 17

청구항 1 내지 16 중 어느 한 항에 있어서, 상기 단편 이온은 탠덤 질량 분광분석법을 사용하여 생성되는, 방법.

청구항 18

청구항 1 내지 17 중 어느 한 항에 있어서, 상기 관심 폴리펩타이드는 단백질을 포함하는, 방법.

청구항 19

청구항 1 내지 18 중 어느 한 항에 있어서, 상기 관심 폴리펩타이드는 단클론성 항체를 포함하는, 방법.

청구항 20

청구항 1 내지 19 중 어느 한 항에 있어서, 상기 관심 폴리펩타이드는 단일특이적 항체 또는 이중특이적 항체를 포함하는, 방법.

발명의 설명

기술 분야

[0001] **서열 목록에 대한 참조**

[0002] 본원은 2019년 8월 16일에 창작되고 그 크기가 7,747 바이트인 10478W001-Sequence.txt 파일로 컴퓨터 판독가능한 형태로 제출된 서열 목록을 참조로 통합한다.

[0003] **발명의 분야**

[0004] 본 발명은 생물약품에 속하고, 단백질 또는 폴리펩타이드 서열의 새로운(*do novo*) 결정에 관한 것이다.

배경 기술

[0005] 단백질 서열분석은 전통적으로 에드만 분해 화학(Edman degradation chemistry)을 사용하여 개별적으로 절단된 N-말단 아미노산의 순차적인 검출 및 예를 들어 차등 HPLC 보유 및 UV 흡수와 같은 기술을 사용하여 상이한 아미노산 에드만 유도체의 검출 및 확인에 의존하였다. 보다 최근에는, 질량 분광분석법이 증가된 속도, 정확도 및 민감도로 단백질 또는 폴리펩타이드를 서열화 및/또는 확인하는데 사용되어 왔다. 그러나, 이들 방법은 일반적으로 낮은 처리량이고 여전히 에드만 분해에 의존한다. 많은 수의 상이한 핵산 분자를 동시에 서열분석할 수 있는 고처리량 대규모 병렬 DNA 서열분석 플랫폼에서 상당한 개선이 이루어졌지만, 질량 분광계 성능의 진보가 점진적으로 이루어졌다. 개별 단일 아미노산 잔기 수준에서 글로벌 단백질 서열분석을 위한 "차세대" 플랫폼의 개발을 향한 진행은 비교적 거의 없었다.

발명의 내용

해결하려는 과제

[0006] 따라서, 서열분석 단일 폴리펩타이드의 새로운 방법 및 검정에 대한 요구가 남아 있다.

과제의 해결 수단

[0007] 일 측면에서, 본 발명은 관심 폴리펩타이드의 아미노산 서열의 새로운 결정 위한 방법을 제공하고, 상기 방법은 다음을 포함한다: 첫 번째 프로테아제가 관심 폴리펩타이드를 소화하여 소화된 펩타이드 단편의 첫 번째 세트를 생성하도록 허용하는 조건 하에, 관심 폴리펩타이드를 함유하는 첫 번째 샘플을 염기성 아미노산 후에 펩타이드 결합을 절단하는 첫 번째 프로테아제와 접촉시키는 단계; 상기 소화된 펩타이드 단편의 첫 번째 세트를 단편화하여 상기 소화된 펩타이드 단편의 첫 번째 세트의 펩타이드에 상응하는 단편화 펩타이드 이온의 첫 번째 세트를 생성하는 단계; 상기 단편화 펩타이드 이온의 첫 번째 세트의 질량을 결정하는 단계; 상기 두 번째 프로테아제가 관심 폴리펩타이드를 소화하여 소화된 펩타이드 단편의 두 번째 세트를 생성하도록 허용하는 조건 하에, 관심 폴리펩타이드를 함유하는 두 번째 샘플을 염기성 아미노산 전에 펩타이드 결합을 절단하는 두 번째 프로테아제와 접촉시키는 단계; 상기 소화된 펩타이드 단편의 두 번째 세트를 단편화하여 상기 소화된 펩타이드 단편의 두 번째 세트의 펩타이드에 상응하는 단편화 펩타이드 이온의 두 번째 세트를 생성하는 단계; 상기 단편화 펩타이드 이온의 두 번째 세트의 질량을 결정하는 단계; 아르기닌 아미노산의 질량 또는 라이신 아미노산의 질량에 의해 질량이 상이한 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍을 선택하는 단계; 이온 유형(N-말단 펩타이드 이온 또는 C-말단 펩타이드 이온)을 상기 단편화 펩타이드 이온의 2개의 세트로부터의 펩타이드 이온의 선택된 쌍에 할당하는 단계; 단편화 아미노산 잔기(들)의 질량에 의한 증분 질량을 갖는 펩타이드 이온의 세트 하나에서 동일한 유형의 펩타이드 이온의 질량 래더(mass ladder)를 선택하고 펩타이드 이온의 질량 래더로부터의 확인된 아미노산 잔기를 조립하여 관심 폴리펩타이드의 아미노산 서열을 결정하는 단계.

[0008] 방법의 일부 구현예에서, 첫 번째 프로테아제는 트립신이다.

[0009] 방법의 일부 구현예에서, 두 번째 프로테아제는 Tryp-N이다.

[0010] 일부 구현예에서, 방법은 추가로 소화된 펩타이드 단편의 첫 번째 세트로부터 첫 번째 소화된 펩타이드 단편을 선택하는 단계; 및 첫 번째 소화된 펩타이드 단편과 동일한 질량을 갖는 소화된 펩타이드 단편의 두 번째 세트로부터 두 번째 소화된 펩타이드 단편을 선택하는 단계를 포함한다.

[0011] 일부 구현예에서, 방법은 추가로 소화된 펩타이드 단편의 첫 번째 세트로부터 첫 번째 소화된 펩타이드 단편을 선택하는 단계; 및 첫 번째 소화된 펩타이드 단편에 대한 라이신 아미노산 잔기와 아르기닌 아미노산 잔기 사이의 질량 차이와 동일한 질량 차이를 갖는 질량을 갖는 소화된 펩타이드 단편의 두 번째 세트로부터 두 번째 소화된 펩타이드 단편을 선택하는 단계를 포함한다.

[0012] 방법의 다양한 구현예에서, 아미노산 서열을 유도하기 위해 단편화 펩타이드 이온의 쌍을 할당하는 것은, 다음을 포함한다: 소화된 펩타이드 단편의 첫 번째 세트로부터 첫 번째 소화된 펩타이드 단편을 선택하는 단계; 첫 번째 소화된 펩타이드 단편을 단편화하여 첫 번째 소화된 펩타이드 단편에 상응하는 단편화 펩타이드 이온의 첫 번째 시리즈를 생성하는 단계; 첫 번째 소화된 펩타이드 단편에 상응하는 소화된 펩타이드 단편의 두 번째 세트로부터 두 번째 소화된 펩타이드 단편을 선택하는 단계; 두 번째 소화된 펩타이드 단편을 단편화하여 상기 두 번째 소화된 펩타이드 단편에 상응하는 단편화 펩타이드 이온의 두 번째 시리즈를 생성하는 단계; 이온 유형(N-말단 펩타이드 이온 또는 C-말단 펩타이드 이온)을 단편화 펩타이드 이온의 2개의 세트로부터의 펩타이드 이온의 선택된 쌍에 할당하는 단계; 아미노산 잔기(들)의 질량에 의해 증분 질량을 갖는 단편화 펩타이드 이온의 세트 중 하나에서 동일한 유형의 펩타이드 이온의 질량 래더를 선택하는 단계; 및 선택된 펩타이드 이온의 질량 래더로부터 첫 번째 및 두 번째 소화된 펩타이드 단편의 개별 아미노산 잔기를 결정하여 첫 번째 및/또는 두 번째 단편화 펩타이드의 아미노산 서열을 생성하는 단계.

[0013] 다양한 구현예에서, 아르기닌 아미노산 잔기의 질량에 의해 질량이 상이한 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온의 쌍이 선택된다. 다양한 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온의 쌍에서 아르기닌 아미노산의 질량 잔기의 음의 차이는, 상기 펩타이드가 N-말단 아르기닌 잔기를 갖는다는 것을

나타낸다. 다양한 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온의 쌍에서 아르기닌 잔기의 양의 질량 차이는, 펩타이드가 C-말단 아르기닌 잔기를 갖는다는 것을 나타낸다. 다양한 구현예에서, 라이신 아미노산 잔기의 질량에 의한 질량이 상이한 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온의 쌍이 선택된다. 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온의 쌍에서 라이신 아미노산의 음의 질량 차이 잔기는, 상기 펩타이드가 N-말단 라이신 잔기를 갖는다는 것을 나타낸다. 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍에서 라이신 아미노산 잔기의 양의 질량 차이는, 상기 펩타이드가 C-말단 라이신 잔기를 갖는다는 것을 나타낸다.

- [0014] 방법의 다양한 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화 펩타이드 이온은 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화 펩타이드 이온에 상응한다.
- [0015] 방법의 다양한 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화 펩타이드 이온은 b 이온이고, 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화된 펩타이드 이온은 아르기닌 아미노산 잔기의 질량 또는 라이신 아미노산 잔기의 질량의 차이를 갖는 b 이온이다.
- [0016] 방법의 다양한 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화 펩타이드 이온은 y 이온이고, 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화된 펩타이드 이온은 아르기닌 아미노산 또는 라이신 아미노산의 질량의 차이를 갖는 y 이온이다.
- [0017] 방법의 다양한 구현예에서, 질량은 질량 분광분석법을 사용하여 결정된다.
- [0018] 방법의 다양한 구현예에서, 단편 이온은 탠덤 질량 분광분석법을 사용하여 생성된다.
- [0019] 방법의 다양한 구현예에서, 관심 폴리펩타이드는 단백질을 포함한다.
- [0020] 방법의 다양한 구현예에서, 관심 폴리펩타이드는 항체, 예컨대 단클론성 항체, 단일특이적 항체 또는 이중특이적 항체를 포함한다.

도면의 간단한 설명

- [0021] 도 1a 및 1b는 Tryp-N 프로테아제 소화를 사용하는 소 혈청 알부민(BSA) 서열 적용범위를 도시한다. 서열 적용범위는 91.4%이다. BSA의 Tryp-N 소화물로부터 생성된 다양한 펩타이드 단편은 BSA 단백질 서열(서열번호: 1) 아래에 나타낸다.
 도 2a 및 2b는 트립신 프로테아제 소화를 사용하는 BSA 서열 적용범위를 도시한다. 서열 적용범위는 94.2%이다. BSA의 트립신 소화물(trypsin digest)로부터 생성된 다양한 펩타이드 단편은 BSA 단백질 서열(서열번호: 1) 아래에 나타낸다.
 도 3은 Tryp-N 프로테아제 및 트립신 프로테아제에 의한 모델 폴리펩타이드의 소화로부터 생성된 수득한 펩타이드 단편을 도시한다. 4개의 상이한 일차 단백질 서열 패턴은 고려된다.
 도 4는 도 3의 케이스(1)에 기재된 바와 같이 생성된 BSA로부터의 펩타이드의 분석을 도시한다. 서열 KLVNELTEFAK(서열번호: 2)을 포함하는 폴리펩타이드에 대해 트립신에 의한 소화는 펩타이드 LVNELTEFAK(서열번호: 3)를 산출한다. Tryp-N에 의한 소화는 펩타이드 KLVNELTEFA(서열번호: 4)를 산출한다. 2개의 펩타이드는 동일한 질량을 갖는다. 그러나, 질량 스펙트럼 분석 동안에 단편화될 때, 각 펩타이드로부터의 수득한 b 이온 또는 각 펩타이드로부터의 y 이온은 단일 라이신 아미노산 잔기의 질량에 의해 상이하다.
 도 5a 내지 5c는 트립신 소화물 및 Tryp-N 소화물에 대해 도 4에 도시된 분석으로부터 수득한 질량 스펙트럼, 및 질량 스펙트럼으로부터 생성된 트립신 및 Tryp-N 이온 맵으로부터의 펩타이드의 일차 서열의 생성을 도시한다. 도 5a는 검출된 b 및 y 이온을 보여주는 탠덤 질량 스펙트럼을 거친 펩타이드 LVNELTEFAK(서열번호: 3)의 예시적인 질량 스펙트럼이다. 도 5b는 검출된 b 및 y 이온을 보여주는 탠덤 질량 스펙트럼을 거친 펩타이드 KLVNELTEFA(서열번호: 4)의 예시적인 질량 스펙트럼이다. 도 5c는 트립신 소화물 및 Tryp-N 소화물에 대한 결정된 이온 맵을 사용하여 일차 서열 LVNELTEFA(서열번호: 5)의 생성을 도시한다.
 도 6은 도 3의 케이스(2)에 기재된 바와 같이 생성된 BSA로부터의 펩타이드의 분석을 도시한다. 서열 RHPEYAVSVLLR(서열번호: 6)을 포함하는 폴리펩타이드에 대해 트립신에 의한 소화는 펩타이드 HPEYAVSVLLR(서열번호: 7)를 산출한다. Tryp-N에 의한 소화는 펩타이드 RHPEYAVSVLL(서열번호: 8)를 산출한다. 2개의 펩타이드는

동일한 질량을 갖는다. 그러나, 질량 스펙트럼 분석 동안에 단편화될 때, 각 펩타이드로부터의 수득한 b 이온 또는 각 펩타이드로부터의 y 이온은 단일 아르기닌 아미노산 잔기의 질량에 의해 상이하다.

도 7a 내지 7c는 트립신 소화물 및 Tryp-N 소화물에 대해 도 6에 도시된 분석으로부터 수득한 질량 스펙트럼, 및 질량 스펙트럼으로부터 생성된 트립신 및 Tryp-N 이온 맵으로부터 펩타이드의 일차 서열의 생성을 도시한다. 도 7a는 검출된 b 및 y 이온을 보여주는 탠덤 질량 스펙트럼을 거친 펩타이드 HPEYAVSVLLR(서열번호: 7)의 예시적인 질량 스펙트럼이다. 도 7b는 검출된 b 및 y 이온을 보여주는 탠덤 질량 스펙트럼을 거친 펩타이드 RHPEYAVSVLL(서열번호: 8)의 예시적인 질량 스펙트럼이다. 도 7c는 트립신 소화물 및 Tryp-N 소화물에 대한 결정된 이온 맵을 사용하여 일차 서열 HPEYAVSVLL(서열번호: 9)의 생성을 도시한다.

도 8은 도 3의 케이스(3)에 기재된 바와 같이 생성된 BSA로부터의 펩타이드의 분석을 도시한다. 서열 KCCTESLVNR(서열번호: 10)을 포함하는 폴리펩타이드에 대해 트립신에 의한 소화는 펩타이드 CCTESLVNR(서열번호: 11)를 산출한다. Tryp-N에 의한 소화는 펩타이드 KCCTESLVN(서열번호: 12)를 산출한다. 이 경우에 2개의 펩타이드는 동일한 질량을 갖지 않는다. 그러나, 질량 스펙트럼 분석 동안에 단편화될 때, 각 펩타이드로부터의 수득한 b 이온 또는 각 펩타이드로부터의 y 이온의 단일 라이신 아미노산 잔기(b 이온) 또는 단일 아르기닌 아미노산 잔기(y 이온)의 질량에 의해 상이하다.

도 9a 내지 9c는 트립신 소화물에 대해 도 8에 도시된 분석으로부터 수득한 질량 스펙트럼 및 Tryp-N 소화물에 대해 물에 대해 및 질량 스펙트럼으로부터 생성된 트립신 및 Tryp-N 이온 맵으로부터 펩타이드의 일차 서열의 생성을 도시한다. 도 9a는 검출된 b 및 y 이온을 보여주는 탠덤 질량 스펙트럼을 거친 펩타이드 CCTESLVNR(서열번호: 11)의 예시적인 질량 스펙트럼이다. 도 9b는 검출된 b 및 y 이온을 보여주는 탠덤 질량 스펙트럼을 거친 펩타이드 KCCTESLVN(서열번호: 12)의 예시적인 질량 스펙트럼이다. 도 9c는 트립신 소화물 및 Tryp-N 소화물에 대한 결정된 이온 맵을 사용하여 일차 서열 CCTESLVN(서열번호: 13)의 생성을 도시한다.

도 10은 도 3의 케이스(4)에 기재된 바와 같이 생성된 BSA로부터의 펩타이드의 분석을 도시한다. 서열 RFKDLGEEHFK(서열번호: 14)을 포함하는 폴리펩타이드에 대해, 트립신에 의한 소화는 펩타이드 RFKDLGEEHFK(서열번호: 15)를 산출하였다. Tryp-N에 의한 소화는 펩타이드 RFKDLGEEHF(서열번호: 16)를 산출하였다. 이 경우에 2개의 펩타이드는 동일한 질량을 갖지 않는다. 그러나, 질량 스펙트럼 분석 동안에 단편화될 때, 각 펩타이드로부터의 수득한 b 이온 또는 각 펩타이드로부터의 y 이온은 단일 라이신 아미노산 잔기(y 이온) 또는 단일 아르기닌 아미노산 잔기(b 이온)에 의해 질량이 상이하다.

도 11a 내지 11c는 트립신 소화물 및 Tryp-N 소화물에 대해 도 10에 도시된 분석으로부터의 수득한 질량 스펙트럼, 및 질량 스펙트럼으로부터 생성된 트립신 및 Tryp-N 이온 맵으로부터 펩타이드의 일차 서열의 생성을 도시한다. 도 11a는 검출된 b 및 y 이온을 보여주는 탠덤 질량 스펙트럼을 거친 펩타이드 RFKDLGEEHFK(서열번호: 15)의 예시적인 질량 스펙트럼이다. 도 11b는 검출된 b 및 y 이온을 보여주는 탠덤 질량 스펙트럼을 거친 펩타이드 RFKDLGEEHF(서열번호: 16)의 예시적인 질량 스펙트럼이다. 도 11c는 트립신 소화물 및 Tryp-N 소화물에 대한 결정된 이온 맵을 사용하여 일차 서열 RFKDLGEEHF(서열번호: 17)의 생성을 도시한다.

도 12는 특정 예시적인 구현예에 따라 폴리펩타이드의 서열을 결정하기 위한 방법을 도시하는 블록 선도이다.

도 13은 특정 예시적인 구현예에 따라 새로운 폴리펩타이드 서열분석의 방법의 다양한 단계를 수행하기 위해 사용될 수 있는 예시적인 컴퓨팅 시스템이다.

발명을 실시하기 위한 구체적인 내용

[0022] 본 발명을 기재하기 전에, 본 발명은 기재된 특정 방법 및 실험 조건에 제한되지 않으며, 이러한 방법 및 조건은 변할 수 있음이 이해되어야 한다. 또한, 본 발명의 범위가 첨부된 청구항들에 의해서만 제한될 것이기 때문에, 본 명세서에서 사용되는 용어는 특정 구현예를 설명하기 위한 목적일 뿐, 제한하는 것으로 의도되지 않는 것이 이해되어야 한다. 구현예 중 임의의 구현예 또는 특징은 서로 조합될 수 있고, 이러한 조합들은 본 발명의 범위 내에 명백하게 포괄된다.

[0023] 달리 정의되지 않는 한, 본 명세서에 사용된 모든 기술적 및 과학적 용어는 본 발명이 속하는 분야의 당업자에 의해 통상적으로 이해되는 것과 동일한 의미를 갖는다. 본원에 사용된 용어 "약"은 특정 인용된 수치와 관련하여 사용될 때, 값이 인용된 값으로부터 1% 이하만큼 변할 수 있음을 의미한다. 예를 들어, 본 명세서에 사용된 바와 같이, 표현 "약 100"은 99 및 101 및 이들 사이의 모든 값(예를 들어, 99.1, 99.2, 99.3, 99.4, 등)을 포함한다.

- [0024] 본 명세서에 기재된 것과 유사하거나 동등한 임의의 방법 및 물질이 본 발명의 실시 또는 시험에서 사용될 수 있지만, 바람직한 방법 및 물질은 이제 기재된다. 본 명세서에 언급된 모든 특허, 출원 및 비특허 공보는 그 전 문이 본원에 참조로 포함된다.
- [0025] **본 명세서에 사용된 약어**
- [0026] MS/MS: 탠덤 질량 분광분석법
- [0027] mAb: 단클론성 항체
- [0028] IgG: 면역글로불린 G
- [0029] LC: 경쇄
- [0030] HC: 중쇄
- [0031] MS: 질량 분광분석법
- [0032] **정의**
- [0033] 본 명세서에서 사용된 용어 "항체"는 디설파이드 결합에 의해 상호 연결된 4개의 폴리펩타이드 사슬, 2개의 중 (H) 사슬 및 2개의 경 (L) 사슬을 포함하는 면역글로불린 분자(즉, "완전 항체 분자"), 뿐만 아니라 그의 다량 체(예를 들어 IgM) 또는 이의 항원-결합 단편을 지칭하는 것으로 의도된다. 각 중쇄는 중쇄 가변 영역("HCVR" 또는 "V_H") 및 중쇄 불변 영역(도메인 C_H1, C_H2 및 C_H3을 포함함)을 포함한다. 다양한 구현예에서, 중쇄는 IgG 동형일 수 있다. 일부 경우에, 중쇄는 IgG1, IgG2, IgG3 또는 IgG4로부터 선택된다. 일부 구현예에서, 중쇄는 동형 IgG1/IgG2 또는 IgG4/IgG2의 키메라 힌지 영역을 선택적으로 포함하는 동형 IgG1 또는 IgG4의 것이다. 각 경쇄는 경쇄 가변 영역("LCVR 또는 "V_L") 및 경쇄 불변 영역(C_L)을 포함한다. V_H 및 V_L 영역은 프레임워크 영역 (FR)로 지칭되는 보다 보존된 영역이 산재된 상보성 결정 영역(CDR)으로 지칭되는 초가변성의 영역으로 추가로 세분될 수 있다. 각 V_H 및 V_L은 다음의 순서로 아미노-말단으로부터 카복시-말단으로 배열된 3개의 CDR 및 4개의 FR을 포함한다: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. 용어 "항체"는 임의의 동형 또는 서브클래스의 당화 및 비-당화된 면역글로불린 둘 다에 대한 언급을 포함한다. 용어 "항체"는 항체를 발현하도록 형질감염된 숙주 세포로부터 단리된 항체와 같은 재조합 수단에 의해 제조, 발현, 생성 또는 단리된 항체 분자를 포함한다. 항체 구조에 대한 검토를 위해, Lefranc 등, *IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains*, 27(1) Dev. Comp. Immunol. 55-77 (2003); 및 M. Potter, *Structural correlates of immunoglobulin diversity*, 2(1) Surv. Immunol. Res. 27-42 (1983)를 참조한다.
- [0034] 용어 항체는 또한 1 초과 개의 상이한 에피토프에 결합할 수 있는 이종사량체 면역글로불린을 포함하는 "이중특이적 항체"를 포괄한다. 단일 중쇄 및 단일 경쇄 및 6개의 CDR을 포함하는 이중특이적 항체의 절반은 하나의 항원 또는 에피토프에 결합하고, 그리고 항체의 다른 절반은 상이한 항원 또는 에피토프에 결합한다. 일부 경우에, 이중특이적 항체는 동일한 항원에 결합할 수 있지만, 상이한 에피토프 또는 비-중첩 에피토프에서 결합할 수 있다. 일부 경우에, 이중특이적 항체의 양 절반은 이중 특이성을 유지하면서 동일한 경쇄를 갖는다. 이중 특이적 항체는 일반적으로 U.S. 특허 출원 공개 번호 2010/0331527(2010년 12월 30일)에 기재되어 있다.
- [0035] 항체(또는 "항체 단편")의 용어 "항원-결합부"는, 항원에 특이적으로 결합하는 능력을 보유하는 항체의 하나 이상의 단편을 지칭한다. 항체의 용어 "항원-결합부" 내에 포괄되는 결합 단편의 예는 (i) Fab 단편, VL, VH, CL 및 CH1 도메인으로 이루어진 1가 단편; (ii) F(ab')₂ 단편, 힌지 영역에서 디설파이드 브릿지에 의해 연결된 2개의 Fab 단편을 포함하는 2가 단편, (iii) VH 및 CH1 도메인으로 이루어진 Fd 단편; (iv) 항체의 단일 팔의 VL 및 VH 도메인으로 이루어진 Fv 단편, 및 (v) VH 도메인으로 이루어진 dAb 단편 (Ward 등 (1989) Nature 241:544-546), (vi) 단리된 CDR, 및 (vii) 합성 링커에 의해 연결되어 VL 및 VH 영역이 쌍을 이루어 1가 분자를 형성하는 단일 단백질 사슬을 형성하는, Fv 단편의 2개의 도메인인 VL과 VH로 이루어진 scFv를 포함한다. 다른 형태의 단일 사슬 항체, 예컨대 디아바디도 용어 "항체" 하에 포괄된다(예를 들어, Holliger 등 (1993) 90 PNAS U.S.A. 6444-6448; 및 Poljak 등 (1994) 2 Structure 1121-1123를 참조한다).
- [0036] 또한, 항체 및 이의 항원-결합 단편은 통상적으로 당업계에서 알려진 표준 재조합 DNA 기술을 사용하여 수득될 수 있다(Sambrook 등, 1989을 참조한다).

- [0037] 용어 "인간 항체"는, 인간 생식세포 면역글로불린 서열로부터 유래된 가변 및 불변 영역을 갖는 항체를 포함하는 것으로 의도된다. 본 발명의 인간 mAb는, 예를 들어 CDR 및 특히 CDR3에서 인간 생식세포 면역글로불린 서열에 의해 인코딩되지 않은 아미노산 잔기(예를 들어, 시험관 내에서 무작위 또는 부위-특이적인 돌연변이유발에 의해 또는 생체 내에서 체세포 돌연변이에 의해 도입된 돌연변이)를 포함할 수 있다. 그러나, 본 명세서에서 사용된 용어 "인간 항체"는, 또 다른 포유류 종(예를 들어, 마우스)의 생식세포로부터 유래된 CDR 서열이 인간 FR 서열 상에 그라프팅되었던 mAb를 포함하는 것으로 의도되지 않는다. 용어는 비-인간 포유동물에서, 또는 비-인간 포유동물의 세포에서 재조합으로 생성된 항체를 포함한다. 용어는 인간 대상체로부터 단리되거나 그에서 생성되는 항체를 포함하는 것으로 의도되지 않는다.
- [0038] 본 명세서에 사용된 용어 "샘플"은 예를 들어, 분리, 분석, 추출, 농축 또는 프로파일링을 포함하는 본 발명의 방법에 따라 조작되는 적어도 관심 폴리펩타이드, 예컨대 단클론성 항체를 포함하는 분자의 혼합물을 지칭한다.
- [0039] 본 명세서에 사용된 용어 "분석" 또는 "분석하는"은 상호교환적으로 사용되며, 관심 분자(예를 들어, 폴리펩타이드, 예컨대 단클론성 항체)를 분리, 검출, 단리, 정제, 가용화, 검출 및/또는 특징화하는 임의의 다양한 방법을 지칭한다. 그 예는 비제한적으로, 고상 추출, 고상 마이크로 추출, 전기영동, 질량 분광분석법, 예를 들어, 탠덤 질량 분광분석법, 액체 크로마토그래피, 예를 들어, 고성능, 예를 들어, 역상, 정상, 또는 크기 배제, 이온-쌍 액체 크로마토그래피, 액체-액체 추출, 예를 들어, 가속화된 유체 추출, 초임계 유체 추출, 마이크로파-보조된 추출, 멤브레인 추출, 속슬래 추출, 침전, 정화, 전기화학적 검출, 염색, 원소 분석, 핵자기 공명, 적외선 분석, 유동 주입 분석, 모세관 전기크로마토그래피, 자외선 검출, 및 이들의 조합을 포함한다.
- [0040] 본 명세서에서 사용된 "크로마토그래피"는 혼합물, 예를 들어 펩타이드, 단백질, 폴리펩타이드 및/또는 항체, 예컨대 단클론성 항체를 함유하는 혼합물을 분리하는 공정을 지칭한다. 이는 혼합물을 고정상에 통과시키는 것을 수반하며, 이는 관심 분자를 혼합물 중의 다른 분자로부터 분리하고 하나 이상의 관심 분자가 단리되도록 한다. 크로마토그래피 분리 방법의 예는 기타 중에서 모세관-작용 크로마토그래피, 예컨대 종이 크로마토그래피, 박층 크로마토그래피(TLC), 칼럼 크로마토그래피, 빠른 단백질 액체 크로마토그래피(FPLC), 나노-역상 액체 크로마토그래피, 이온 교환 크로마토그래피, 겔 크로마토그래피, 예컨대 겔 여과 크로마토그래피, 크기 배제 크로마토그래피, 친화성 크로마토그래피, 고성능 액체 크로마토그래피(HPLC), 친수성 상호작용 액체 크로마토그래피(HILIC), 및 역상 고성능 액체 크로마토그래피(RP-HPLC)을 포함한다.
- [0041] 본 명세서에서 사용된 "접촉하는"은 용액 또는 고상에서 2종 이상의 물질을 함께 합치는 것, 예를 들어 샘플을 프로테아제와 접촉시키는 것을 포함한다.
- [0042] 용어 "상응하는"은 위치, 목적 또는 구조의 유사성을 나타내는 상대적 용어이고, N- 또는 C-말단 상의 아르기닌 또는 라이신 아미노산 잔기의 존재 또는 부재를 제외하고는 동일한 구조의 펩타이드를 포함할 수 있다. 일부 구현예에서, N- 또는 C-말단 상의 아르기닌 또는 라이신 아미노산 잔기의 존재 또는 부재를 제외하고는 동일한 구조의 상응하는 펩타이드로 인한 질량 스펙트럼에서의 질량 스펙트럼 신호는 "상응하는" 질량 스펙트럼 신호이다. 특정 펩타이드로 인한 질량 스펙트럼 신호는 또한 펩타이드에 상응하는 신호로 지칭된다. 특정 구현예에서, 특정 펩타이드 서열 또는 아미노산 세트는 상응하는 펩타이드 질량에 할당될 수 있다.
- [0043] 본 명세서에서 사용된 용어들 "단편 펩타이드" 또는 "펩타이드 단편"은, 단편화, 효소적 단백질분해, 또는 화학 가수분해를 포함하는 과정을 통해 전장 폴리펩타이드, 예컨대 단백질 및/또는 단클론성 항체로부터 유래된 펩타이드를 지칭한다. 그와 같은 단백질분해 펩타이드는 펩타이드는 하나 이상의 프로테아제 예컨대 트립신 또는 Tryp-N에 의한 단백질의 처리에 의해 생성된다. 단편 펩타이드, 또는 펩타이드 단편, 소화된 펩타이드일 수 있다.
- [0044] 본 명세서에 사용된 용어 "단리된"은 성분이 자연적으로 발생하거나 형질전환으로 발현되는 유기체의 세포에서 다른 생물학적 성분, 즉 다른 염색체 및 염색체의 DNA 및 RNA, 단백질, 지질 및 대사산물로부터 실질적으로 분리, 생성 또는 정제된 생물학적 성분(예컨대 핵산, 펩타이드, 단백질 또는 대사물)을 지칭한다. 따라서 "단리된" 핵산, 펩타이드, 단백질, 지질 및 대사산물은 표준 또는 비-표준 정제 방법에 의해 정제된 핵산, 펩타이드, 단백질, 지질, 및 대사산물을 포함한다. 용어는 또한 숙주 세포에서 재조합 발현으로 제조된 핵산, 펩타이드, 단백질, 지질, 및 대사산물뿐만 아니라 화학적으로 합성된 펩타이드, 지질, 대사산물, 및 핵산을 포용한다.
- [0045] "질량 분석법"은 샘플로부터 기상 이온을 생성함으로써 샘플을 분석한 다음, 이들을 질량-대-전하 비(m/z)에 따라 분리하고 검출하는 방법이다. 샘플로부터 기상 이온을 생성하는 방법은 전기분무 이온화(ESI), 매트릭스-보

조 레이저 탈착-이온화(MALDI), 표면-향상 레이저 탈착-이온화(SELDI), 화학적 이온화 및 전자-충격 이온화(EI)를 포함한다. m/z 비율에 따른 이온의 분리는 사중극자 질량 분석기(Q), 비과시간(TOF) 질량 분석기, 자기 섹터 질량 분석기, 3D 및 선형 이온 트랩(IT), 오르비트랩(orbitrap) 질량 분석기, 푸리에 변환 이온 사이클로트론 공명(FT-ICR) 분석기, 및 이들의 조합(예를 들어, 사중극자-비과 시간 분석기, 또는 Q-TOF 분석기) 분석기를 포함하는 임의의 유형의 질량 분석기로 달성될 수 있다. 분리 전에, 샘플은 크로마토그래피 분리의 하나 이상의 치수, 예를 들어, 하나 이상의 치수의 액체 또는 크기 배제 크로마토그래피를 거칠 수 있다.

[0046] 탠덤 질량 분광분석법 또는 MS/MS는 선택된 이온(전구체 이온)을 단편(생성물 이온)으로 분해하는 기술이다. 그 후, 단편은 전구체 이온의 화학 구조의 측면을 나타낸다. 탠덤 질량 분광분석법에서, 일단 샘플이 (예를 들어 ESI, MALDI, EI 등에 의해) 이온화되어 이온의 혼합물을 생성하면, 전구체 이온, 예를 들어 특정 질량-대-전하비(m/z)의 소화물로부터 펩타이드가 선택되고 (MS1), 이어서 단편화되어 (MS2) 검출을 위한 생성물 이온을 생성한다. 전형적인 탠덤 MS 기기는 QqQ, QTOF, 및 하이브리드 이온 트랩/FTMS 등을 포함한다. 탠덤 질량 분광분석법 적용의 한 예는 단백질 확인이다. 제1 질량 분석기는 이온 공급원 내로 도입된 다음 이온 공급원으로부터 나오는 많은 것 중에서 펩타이드의 단일 종을 나타내는 특정 m/z 값의 이온을 단리한다. 이어서, 이들 이온은 아르곤과 같은 불활성 기체를 함유하는 충돌 셀로 가속되어 이온 단편화를 유도한다. 이 과정은 충돌 유도 해리(CID) 또는 충돌 활성화 해리(CAD)로 지칭된다. 이어서, 단편 이온의 m/z 값을 제2 질량 분석기에서 측정하여 아미노산 서열 정보를 얻는다. 탠덤 질량 분광분석법은 펩타이드의 서열 및 따라서 본 명세서에 개시된 방법에 따른 전체 또는 부분 길이 단백질을 확인하는데 사용될 수 있다. 탠덤 질량 스펙트럼으로부터 발생하는 펩타이드 단편을 나타내기 위한 표기법이 개발되었다. 본 명세서에 사용된 바와 같이, 펩타이드 단편 이온은, 전하가 N-말단에 보유되는 경우 b 로 표시되고, 전하가 C-말단에서 유지되는 경우 y 로 표시된다. b 또는 y 를 따르는 수는 단편 내의 아미노산의 수를 나타낸다. 전구체 이온은 많은 상이한 방식으로 (내부 에너지가 증가되어) 활성화될 수 있다. 단편화 패턴은 에너지가 전구체 이온으로 전달되는 방법, 전달된 에너지의 양, 및 전달된 에너지가 내부적으로 분배되는 방법에 의존한다. 충돌-유도 해리 및 적외선 다광자 해리는 이온의 볼츠만(Boltzmann) 온도를 증가시키고 따라서 주로 b 및 y 이온을 생성하기 위해 가장 약한 결합을 우선적으로 절단하는 "느린-가열" 기술이다.

[0047] 용어들 "펩타이드", "단백질" 및 "폴리펩타이드"는 상호교환적으로 펩타이드 결합 또는 펩타이드 결합 모방체에 의해 연결된 아미노산 및/또는 아미노산 유사체의 중합체를 지칭한다. 20종의 자연 발생 아미노산 및 그것의 단일 문자 및 3-문자 지정은 아래와 같다: 알라닌 A Ala; 시스테인 C Cys; 아스파르트산 D Asp; 글루탐산 E Glu; 페닐알라닌 F Phe; 글리신 G Gly; 히스티딘 H His; 이소류신 I Ile; 라이신 K Lys; 류신 L Leu; 메티오닌 M Met; 아스파라긴 N Asn; 프롤린 P Pro; 글루타민 Q Gln; 아르기닌 R Arg; 세린 S Ser; 트레오닌 T Thr; 발린 V Val; 트립토판 w Trp; 및 티로신 Y Tyr.

[0048] 아미노산의 질량에 대한 언급은 주어진 동위원소 존재비, 예컨대 자연 존재비에서의 아미노산의 단일동위원소 질량 또는 평균 질량을 의미한다. 일부 예에서, 아미노산의 질량은 예를 들어 아미노산을 동위원소로 표시함으로써 왜곡될 수 있다. 아미노산의 평균 질량 주위의 어느 정도의 가변성은 아미노산의 정확한 동위원소 조성에 기초하여 개별 단일 아미노산에 대해 예상된다. 아미노산에 대한 단일동위원소 및 평균 질량을 포함하는 질량은 당업자에 의해 용이하게 획득가능하다.

[0049] 유사하게, 펩타이드의 질량에 대한 언급은 주어진 동위원소 존재비, 예컨대 자연 존재비에서의 펩타이드의 단일 동위원소 질량 또는 평균 질량을 의미한다. 일부 예에서, 펩타이드의 질량은 예를 들어 펩타이드 내의 하나 이상의 아미노산을 동위원소로 표시함으로써 왜곡될 수 있다. 펩타이드의 평균 질량 주위의 어느 정도의 가변성은 펩타이드의 정확한 동위원소 조성에 기초하여 개별 단일 펩타이드에 대해 예상된다. 특정 펩타이드의 질량은 당업자에 의해 결정될 수 있다.

[0050] **일반적인 설명**

[0051] 본 개시내용의 측면은 관심 폴리펩타이드, 예컨대 단클론성 항체 또는 다른 관심 단백질의 아미노산 서열을 결정하는 방법에 관한 것이다. DNA 서열분석과 매우 유사하게, 개시된 방법은 폴리펩타이드의 서열에 대한 사전 정보를 필요로 하지 않는다. 참조를 위해, 도 12는 개시된 방법에 대한 예시적인 작업-흐름을 도시하지만, 이에 제한되지 않는다. 개시된 방법의 독특한 특징 중 하나는 각각 염기성 아미노산 전 및 후에 펩타이드 결합을 자르거나 절단하는 한 쌍의 프로테아제의 사용이다. 트립신과 같은 첫 번째 프로테아제는 아르기닌 또는 라이신 잔기와 같은 염기성 아미노산 직후에 펩타이드 결합을 절단하는 반면, Tryp-N과 같은 상기 두 번째 프로테아제가 아르기닌 또는 라이신 잔기 직전과 같은 염기성 아미노산의 직전에 펩타이드 결합을 절단한다(각각의 프로테

아제에 의한 소 혈청 알부민 소화물에서 생성된 펩타이드의 맵은 도 1 및 2에 도시됨). 본 발명자들은 이러한 프로테아제 세트가 질량 분광분석 기술과 함께 사용되어 2개의 효소로 별도로 소화된 폴리펩타이드의 서열을 결정할 수 있음을 인식하였다.

[0052] 따라서, 관심 폴리펩타이드의 아미노산 서열을 결정하는 방법이 본 명세서에 개시된다. 상기 방법의 구현예에서, (예를 들어 알려지지 않은 서열의) 관심 폴리펩타이드를 함유하는 샘플, 예컨대 첫 번째 샘플은, 첫 번째 프로테아제가 관심 폴리펩타이드를 소화시키고 소화된 펩타이드 단편, 예를 들어 소화된 펩타이드 단편의 첫 번째 세트를 생성하는 것을 허용하는 조건 하에 첫 번째 프로테아제와 접촉된다. 소화는, 예를 들어, 중첩하는 단편을 생성하기 위한 완전한 소화 또는 불완전한 소화일 수 있다. 구현예에서, 병렬로, 예컨대 동시에, 또는 임의의 순서로 순차적으로, 샘플, 예컨대 관심 폴리펩타이드를 함유하는 첫 번째 샘플로부터 분할된 두 번째 샘플은 두 번째 프로테아제, 예컨대 Tryp-N 프로테아제가 관심 폴리펩타이드를 소화시키고 소화된 펩타이드 단편, 예를 들어 소화된 펩타이드 단편의 두 번째 세트를 생성하는 것을 가능하게 하는 조건 하에 첫 번째 프로테아제와 접촉된다. 소화는, 예를 들어, 중첩 단편을 생성하기 위한 완전한 소화 또는 불완전한 소화일 수 있다.

[0053] 도 3은 표시된 바와 같이 2개의 염기성 아미노산, 예컨대 라이신 잔기 및/또는 아르기닌 잔기에 의해 결합된 폴리펩타이드의 부분에 대해 생성된 소화된 펩타이드 단편을 도시한다. 도 3에서 나타낸 바와 같이, 도시된 바와 같이 라이신 잔기 또는 아르기닌 잔기에 의해 N- 또는 C-말단에서 결합된 8종 (각 효소에 대해 4종)이 생성될 수 있다. 케이스(1)에서, (더 큰 펩타이드 또는 단백질의 맥락에서) 펩타이드 서열은 라이신(K) 아미노산 잔기에 의해 N- 및 C-말단 상에 결합된다. 트립신 프로테아제에 의한 소화가 C-말단 라이신 아미노산 잔기를 갖는 단편 펩타이드를 생성하는 것은, 트립신 프로테아제가 라이신 아미노산 후에 (즉, 그 C-말단에) 펩타이드 사슬을 절단하기 때문이다. Tryp-N 프로테아제에 의한 소화가 N-말단 라이신 아미노산 잔기를 갖는 단편 펩타이드를 생성하는 것은, Tryp-N 프로테아제가 라이신 아미노산 전에 (즉, 그 N-말단에) 펩타이드 사슬을 절단하기 때문이다. 케이스(2)에서, (더 큰 펩타이드 또는 단백질의 맥락에서) 펩타이드 서열은 아르기닌(R) 아미노산 잔기에 의해 N- 및 C-말단 상에 결합된다. 트립신 프로테아제에 의한 소화가 C-말단 아르기닌 아미노산 잔기를 갖는 단편 펩타이드를 생성하는 것은, 아르기닌 아미노산 전에 (즉, 그 C-말단에) 트립신 프로테아제가 펩타이드 사슬을 절단하기 때문이다. Tryp-N 프로테아제에 의한 소화가 N-말단 아르기닌 아미노산 잔기를 갖는 단편 펩타이드를 생성하는 것은, 아르기닌 아미노산 전에 (즉, 그 N-말단에) Tryp-N 프로테아제가 펩타이드 사슬을 절단하기 때문이다. 케이스(3)에서, 펩타이드 서열 (더 큰 펩타이드 또는 단백질의 맥락에서)은 라이신 아미노산 잔기에 의해 N-말단 상에 결합되고, 아르기닌 아미노산 잔기에 의해 C-말단 상에 결합된다. 트립신 프로테아제에 의한 소화는 C-말단 아르기닌 아미노산 잔기를 갖는 단편 펩타이드를 생성한다. Tryp-N 프로테아제에 의한 소화는 N-말단 라이신 아미노산 잔기를 갖는 단편 펩타이드를 생성한다. 케이스(4)에서, 펩타이드 서열 (더 큰 펩타이드 또는 단백질의 맥락에서)은 아르기닌 아미노산 잔기에 의해 N-말단 상에 결합되고, 라이신 아미노산 잔기에 의해 C-말단 상에 결합된다. 트립신 프로테아제에 의한 소화는 C-말단 라이신 아미노산 잔기를 갖는 단편 펩타이드를 생성한다. Tryp-N 프로테아제에 의한 소화는 N-말단 아르기닌 아미노산 잔기를 갖는 단편 펩타이드를 생성한다.

[0054] 본 발명자들은 단편화 동안, 예를 들어 탠덤 질량 분광계에서, 개별 소화된 펩타이드 단편으로부터 생성된 b 및 y 이온(상기 논의된 케이스 1 내지 4를 참조함)이 트립신 또는 Tryp-N에 의해 소화되었는지 여부에 따라 아르기닌 아미노산 잔기 또는 리신 아미노산 잔기의 질량에 의해 질량이 상이할 것임을 인식하였다. 이들은 또한 이러한 차이를 이용하여 단편 이온에 대한 N-말단 또는 C-말단 이온 유형을 결정하고, 아미노산 잔기(들)의 질량에 의한 증분 질량을 갖는 동일한 유형의 단편 이온의 질량 래더로부터 확인된 아미노산 잔기를 펩타이드의 특정 서열로 조립할 수 있고, 이들 펩타이드 서열 자체가 폴리펩타이드, 예컨대 단클론성 항체의 전장 아미노산 서열로 조립될 수 있음을 인식하였다. 표 1은, 명시된 크립신 단편 이온의 질량으로부터 명시된 Tryp-N 단편 이온의 질량을 차감함으로써 트립신 소화물 및 Tryp-N 소화물에 대한 b 및 y 단편 이온에 대한 질량의 차이를 나타낸다.

표 1: b 및 y 단편 이온에 대한 질량 차이.

서열	케이스 1			
K.....K	트립신		Tryp-N	차이
	b(x)	-	b(x+1)	-K
	y(x)	-	y(x-1)	+K
서열	케이스 2			
R.....R	트립신		Tryp-N	
	b(x)	-	b(x+1)	-R
	y(x)	-	y(x-1)	+R
서열	케이스 3			
K.....R	트립신		Tryp-N	
	b(x)	-	b(x+1)	-K
	y(x)	-	y(x-1)	+R
서열	케이스 4			
R.....K	트립신		Tryp-N	
	b(x)	-	b(x+1)	-R
	y(x)	-	y(x-1)	+K

[0055]

[0056]

예로서, 케이스(1)에 대해, 트립신 소화물로부터의 b6 단편 이온의 질량에서, Tryp-N 소화물로부터의 b7 단편 이온의 질량을 뺀 질량은 라이신 잔기의 질량에 대해 음인 질량을 초래할 것이다. 대안적으로, Tryp-N 소화물로부터의 b7 단편 이온의 질량으로부터 트립신 소화물로부터 b6 단편 이온의 질량을 차감하는 것은 라이신 잔기의 질량에 양인 질량을 초래할 것이다. 마찬가지로, 트립신 소화물로부터의 y6 단편 이온의 질량에서, Tryp-N 소화물로부터의 y5 단편 이온의 질량을 뺀 질량은 라이신 잔기의 질량에 대해 양인 질량을 초래할 것이다. 대안적으로, Tryp-N 소화물로부터의 y5 단편 이온의 질량으로부터의 트립신 소화물로부터 y6 단편 이온의 질량을 차감하는 것은 라이신 잔기의 질량에 대해 음인 질량을 초래할 것이다.

[0057]

구현예에서, 첫 번째 프로테아제 소화물로부터의 소화된 펩타이드 단편, 예를 들어, 소화된 펩타이드 단편의 두 번째 세트는, 단편화되어 예를 들어 탠덤 질량 분광계를 사용하여 단편화 펩타이드 이온의 첫 번째 세트를 생성한다. 이어서, 단편화 펩타이드 이온의 첫 번째 세트의 질량은 예를 들어 질량 분광분석법에 의해 결정된다. 구현예에서, 두 번째 프로테아제 소화물로부터의 소화된 펩타이드 단편, 예를 들어, 소화된 펩타이드 단편의 두 번째 세트는, 단편화되어 예를 들어 탠덤 질량 분광계를 사용하여 단편화 펩타이드 이온의 두 번째 세트를 생성한다. 이어서, 단편화 펩타이드 이온의 두 번째 세트의 질량은 예를 들어 질량 분광분석법에 의해 결정된다. 단편화 펩타이드 이온의 첫 번째 및 두 번째 세트의 질량을 사용하여, 아르기닌 아미노산 잔기의 질량 또는 라이신 아미노산 잔기의 질량에 의해 질량이 상이한 단편화 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터의 상응하는 펩타이드 이온의 쌍이 선택된다. 상응하는 펩타이드 이온의 쌍이란, 첫 번째 세트로부터 선택된 하나 및 두 번째 세트로부터 선택된 하나를 의미하기 위한 것이다. 구현예에서, 이온 유형은 단편화 펩타이드 이온으로부터 선택된 펩타이드 이온의 쌍으로부터 결정되고, 단편화 펩타이드 이온의 첫 번째 세트 또는 단편화 펩타이드 이온의 두 번째 세트로부터의 아미노산 잔기(들)의 질량에 의한 증분 질량을 갖는 동일한 유형의 펩타이드 이온의 질량 래더가 생성된다. 개별 20개의 아미노산의 질량에 대한 질량 래더에서의 2개의 인접한 펩타이드 이온의 질량 차이를 조사함으로써, 당업자는 특정 단편화 펩타이드 이온을 구성하는 개별 아미노산을 결정할 수 있다. 특정 구현예에서, 특정 소화된 펩타이드에 해당하는 다중 단편화 펩타이드 이온, 예를 들어 단편화 펩타이드 이온의 b 시리즈 및/또는 단편화 펩타이드 이온의 y 시리즈(예를 들어 b1, b2, b3, b4, b5, 등 또는 y1, y2, y3, y4, 등)을 사용하여, 특정 소화된 펩타이드의 일차 서열은 높은 신뢰로 결정될 수 있다. 상기에 논의된 바와 같이, 케이스 1-4에서 질량 분광분석법에서 이온 맵으로부터 펩타이드의 조합은, 도 4 내지 11c에서 도시된다. 방법의 일부 구현예에서, 아미노산 서열을 유도하기 위해 단편화 펩타이드 이온의 쌍을 할당하는 것은 다음을 포함한다: 소화된 펩타이드 단편의 첫 번째 세트로부터의 첫 번째 소화된 펩타이드

단편을 선택하는 단계; 첫 번째 소화된 펩타이드 단편을 단편화하여 첫 번째 소화된 펩타이드 단편에 상응하는 단편화 펩타이드 이온의 첫 번째 시리즈를 생성하는 단계; 상기 첫 번째 소화된 펩타이드 단편에 상응하는 소화된 펩타이드 단편의 두 번째 세트로부터의 두 번째 소화된 펩타이드 단편을 선택하는 단계; 두 번째 소화된 펩타이드 단편을 단편화하여 상기 두 번째 소화된 펩타이드 단편에 상응하는 단편화 펩타이드 이온의 두 번째 시리즈를 생성하는 단계; 단편화 펩타이드 이온의 2개의 시리즈로부터 선택된 펩타이드 이온의 쌍의 이온 유형을 결정하는 단계; 단편화 펩타이드 이온의 하나의 세트로부터 아미노산 잔기(들)의 질량의 증분 질량을 갖는 동일한 유형의 펩타이드 이온의 질량 래더를 선택하고 펩타이드 이온의 질량 래더로부터 첫 번째 및 두 번째 소화된 펩타이드 단편의 개별 아미노산 잔기를 결정하여 첫 번째 및/또는 두 번째 단편화 펩타이드의 아미노산 서열을 생성하는 단계. 소화물 중 펩타이드의 아미노산 서열이 결정되거나, 또는 이의 분획이 결정되면, 할당된 펩타이드의 아미노산 서열은, 예를 들어 중첩 또는 부분적으로 중첩 서열의 서열 정렬을 사용하여, 관심 폴리펩타이드의 아미노산 서열을 형성할 수 있고, 예를 들어 중첩 또는 부분적으로 중첩 서열의 서열 정렬을 사용하여, BSA에 대해 예를 들어 도 1 및 2를 참조한다.

[0058] 일부 구현예에서, 상기 방법은 소화된 펩타이드 단편의 첫 번째 세트로부터의 첫 번째 소화된 펩타이드 단편을 선택하는 단계, 및 상기 첫 번째 소화된 펩타이드 단편과 동일한 질량을 갖는 소화된 펩타이드 단편의 두 번째 세트로부터 두 번째 소화된 펩타이드 단편을 선택하는 단계를 포함한다 (도 4-7c에서 케이스 (1) 및 (2)를 참조한다). 구현예에서, 상기 방법은 소화된 펩타이드 단편의 첫 번째 세트로부터의 첫 번째 소화된 펩타이드 단편을 선택하고 첫 번째 소화된 펩타이드 단편으로부터의 라이신 아미노산 잔기와 아르기닌 아미노산 잔기 사이의 질량 차이와 동일한 질량 차이를 갖는 질량을 갖는 소화된 펩타이드 단편의 두 번째 세트로부터 두 번째 소화된 펩타이드 단편을 선택하는 단계를 포함한다(도 8 내지 11c에서 케이스(3) 및 (4)를 참조한다). 구현예에서, 상기 방법은 단편 펩타이드의 첫 번째 세트로부터 첫 번째 단편화 펩타이드를 선택하는 단계 및 선택된 단편화 펩타이드를 단편화하여 선택된 단편화 펩타이드에 상응하는 단편화 펩타이드 이온을 생성하는 단계를 포함한다. 구현예에서, 첫 번째 단편화 펩타이드 이온은 단편화 펩타이드 이온의 질량에 기초하여 아미노산 서열이 할당된다. 구현예에서, 아미노산 잔기(들)의 질량의 증분 질량을 갖는 동일한 유형의 펩타이드 이온의 질량 래더는 단편화 펩타이드 이온의 각 세트에서 생성된다. 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트 또는 단편화 펩타이드 이온의 두 번째 세트로부터 확인된 개별 아미노산 잔기는 조립되어 첫 번째 및/또는 두 번째 단편화 펩타이드의 아미노산 서열을 생성한다.

[0059] 특정 구현예에서, 상기 방법은 아르기닌 아미노산 잔기의 질량에 의해 질량이 상이한 단편화된 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍을 선택하는 단계를 포함한다. 특정 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 펩타이드 이온과 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온 사이의 아르기닌 아미노산 잔기의 음의 질량 차이는, 상기 펩타이드가 N-말단 아르기닌 잔기를 갖는다는 것을 나타낸다. 특정 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 펩타이드 이온과 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온 사이의 아르기닌 아미노산 잔기의 양의 질량 차이는, 펩타이드가 C-말단 아르기닌 잔기를 갖는다는 것을 나타낸다. 특정 구현예에서, 상기 방법은 라이신 아미노산 잔기의 질량에 의해 질량이 상이한 단편화된 펩타이드 이온의 첫 번째 세트 및 단편화 펩타이드 이온의 두 번째 세트로부터 펩타이드 이온의 쌍을 선택하는 단계를 포함한다. 특정 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 펩타이드 이온과 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온 사이의 라이신 아미노산 잔기의 음의 질량 차이는, 상기 펩타이드가 N-말단 라이신 잔기를 갖는다는 것을 나타낸다. 특정 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 펩타이드 이온과 단편화 펩타이드 이온의 두 번째 세트로부터의 펩타이드 이온 사이의 라이신 아미노산 잔기의 양의 질량 차이는, 상기 펩타이드가 C-말단 라이신 잔기를 갖는다는 것을 나타낸다. 특정 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화 펩타이드 이온은 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화 펩타이드 이온과 상응한다. 특정 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화 펩타이드 이온은 b 이온이고, 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화 펩타이드 이온은 아르기닌 아미노산 잔기의 질량 또는 라이신 아미노산 잔기의 질량의 차이를 갖는 b 이온이다. 특정 구현예에서, 단편화 펩타이드 이온의 첫 번째 세트로부터의 선택된 단편화 펩타이드 이온은 y 이온이고, 단편화 펩타이드 이온의 두 번째 세트로부터의 선택된 단편화 펩타이드 이온은 아르기닌 아미노산 잔기의 질량 또는 라이신 아미노산 잔기의 질량의 차이를 갖는 y 이온이다.

[0060] 일부 예에서, 샘플은 예를 들어 질량 스펙트럼 분석을 위해 관심 생체분자를 정제하기 위해 샘플 전처리를 거친다. 일부 예에서, 샘플 전처리는 겔 전기영동, 액체 크로마토그래피, 기체 크로마토그래피, 모세관 전기영동, 모세관 겔 전기영동, 등전점 초점조절 크로마토그래피, 종이 크로마토그래피, 박-층 크로마토그래피; 나노-유동

크로마토그래피, 미세-유동 크로마토그래피, 고-유속 크로마토그래피, 역상 크로마토그래피, 정상 크로마토그래피, 친수성-상호작용 크로마토그래피, 이온 교환 크로마토그래피, 다공성 흡연질 크로마토그래피, 크기-배제 크로마토그래피, 친화도-기반, 크로마토그래피, 칩-기반 미세유체공학, 고-성능 액체 크로마토그래피, 초-고압 액체 크로마토그래피 또는 유동-압력 액체 크로마토그래피 중 하나 이상을 포함한다. 일부 구현예에서, 샘플은 펩타이드 및/또는 펩타이드 이온의 질량 측정을 복잡하게 하는 당화와 같은 임의의 번역후 변형을 제거하기 위해 샘플 전처리를 거친다.

[0061] 본 명세서에 개시된 방법의 특정 측면 및/또는 단계는 질량 분광계의 일부일 수 있거나 질량 분광계로부터 분리될 수 있는 하나 이상의 컴퓨팅 머신 상에서 수행될 수 있는 것으로 생각된다.

[0062] 도 13은 폴리펩타이드의 아미노산 서열, 예컨대 단클론성 항체 또는 다른 단백질의 결정을 위해, 특정 예시적인 구현예에 따라 컴퓨팅 머신(2000) 및 모듈(2050)을 도시한다. 컴퓨팅 머신(2000)은 임의의 다양한 컴퓨터, 서버, 모바일 디바이스, 임베디드 시스템, 또는 컴퓨팅 시스템에 해당할 수 있다. 모듈(2050)은 본 명세서에 기재된 다양한 방법 및 프로세싱 기능을 수행함에 있어 컴퓨팅 머신(2000)을 용이하게 하도록 구성된 하나 이상의 하드웨어 또는 소프트웨어 요소를 포함할 수 있다. 컴퓨팅 머신(2000)은 다양한 내부 또는 부착된 구성요소 예컨대 프로세서(2010), 시스템 버스(2020), 시스템 메모리(2030), 저장 매체(2040), 입력/출력 인터페이스(2060), 및 네트워크(2080)와 통신하기 위한 네트워크 인터페이스(2070)를 포함할 수 있다. 일부 예에서, 컴퓨팅 머신은 질량 분광계의 일부일 수 있고, 질량 분광계에 연결되고/거나 질량 분광계, 예컨대 네트워크를 통해 질량 분광계로부터 데이터를 수신할 수 있고, 예를 들어 텐덤 질량 분광계에서 생성된 b 및 y 단편 이온에 상응하는 질량 스펙트럼 데이터를 수신할 수 있다.

[0063] 컴퓨팅 머신(2000)은 종래의 컴퓨터 시스템, 임베디드 컨트롤러, 랩탑, 서버, 모바일 디바이스, 스마트폰, 텔레비전과 관련된 하나 이상의 프로세서, 맞춤형 기계, 임의의 다른 하드웨어 플랫폼, 또는 이들의 임의의 조합 또는 다중도로서 구현될 수 있다. 컴퓨팅 머신(2000)은 데이터 네트워크 또는 버스 시스템을 통해 상호연결된 다중 컴퓨팅 머신을 사용하여 기능하도록 구성된 분산 시스템일 수 있다.

[0064] 프로세서(2010)는 본 명세서에서 설명된 기능으로서 동작들을 수행하고, 요청 흐름 및 어드레스 매핑들을 관리하고, 계산들을 수행하고 명령을 생성하기 위해 코드 또는 명령을 실행하도록 구성될 수 있다. 프로세서(2010)는 컴퓨팅 머신(2000) 내의 구성요소의 동작을 모니터링하고 제어하도록 구성될 수 있다. 프로세서(2010)는 범용 프로세서, 프로세서 코어, 멀티프로세서, 재구성가능 프로세서, 마이크로컨트롤러, 디지털 신호 프로세서("DSP"), 주문형 집적회로("ASIC"), 그래픽 처리 장치("GPU"), 필드 프로그래머블 게이트 어레이("FPGA"), 프로그래머블 로직 디바이스("PLD"), 컨트롤러, 상태 기계(state machine), 게이팅된 로직, 별개의 하드웨어 구성요소, 임의의 다른 처리 장치, 또는 이들의 임의의 조합 또는 다중도일 수 있다. 프로세서(2010)는 단일 처리 장치, 다중 처리 장치, 단일 프로세싱 코어, 다중 프로세싱 코어, 특별한 목적 프로세싱 코어, 코-프로세서, 또는 이들의 임의의 조합일 수 있다. 특정 예시적인 구현예에 따르면, 컴퓨팅 머신(2000)의 다른 구성요소와 함께 프로세서(2010)는 하나 이상의 다른 컴퓨팅 머신 내에서 실행되는 가상화된 컴퓨팅 머신일 수 있다.

[0065] 시스템 메모리(2030)는 비-휘발성 메모리 예컨대 읽기 전용 메모리("ROM"), 프로그래밍가능한 읽기 전용 메모리("PROM"), 소거가능 프로그래밍가능한 읽기 전용 메모리("EPROM"), 플래시 메모리, 또는 인가 전력 여부에 관계없이 프로그램 명령 또는 데이터를 저장할 수 있는 임의의 다른 장치를 포함할 수 있다. 시스템 메모리(2030)는 또한 휘발성 메모리 예컨대 랜덤 액세스 메모리("RAM"), 정적 랜덤 액세스 메모리("SRAM"), 동적 랜덤 액세스 메모리("DRAM"), 및 동시 동적 랜덤 액세스 메모리("SDRAM")를 포함할 수 있다. 다른 유형의 RAM은 또한 시스템 메모리(2030)를 구현하도록 사용될 수 있다. 시스템 메모리(2030)는 단일 메모리 모듈 또는 다중 메모리 모듈을 사용하여 구현될 수 있다. 시스템 메모리(2030)가 컴퓨팅 머신(2000)의 일부로서 도시되지만, 당업자, 시스템 메모리(2030)가 본 기술의 범위를 벗어나지 않으면서 컴퓨팅 머신(2000)으로부터 분리될 수 있음을 인식할 것이다. 또한 시스템 메모리(2030)이 비-휘발성 저장 장치 예컨대 저장 매체(2040)를 포함하거나, 그것과 함께 동작할 수 있음을 이해해야 한다.

[0066] 저장 매체(2040)는 하드 디스크, 플로피 디스크, 콤팩트 디스크 읽기 전용 메모리("CD-ROM"), 디지털 다용도 디스크("DVD"), 블루레이 디스크, 자기 테이프, 플래시 메모리, 다른 비-휘발성 메모리 소자, 솔리드 스테이트 드라이브("SSD"), 임의의 자기 저장 장치, 임의의 광학 저장 장치, 임의의 전기 저장 장치, 임의의 반도체 저장 장치, 임의의 물리적-기반 저장 장치, 임의의 다른 데이터 저장 장치, 또는 이들의 임의의 조합 또는 다중도를 포함할 수 있다. 저장 매체(2040)는 하나 이상의 운영 체제, 애플리케이션 프로그램 및 프로그램 모듈 예컨대 모듈(2050), 데이터, 또는 임의의 다른 정보를 저장할 수 있다. 저장 매체(2040)는 컴퓨팅 머신(2000)의 일부이

거나 그에 연결될 수 있다. 저장 매체(2040)는 또한 컴퓨팅 머신(2000) 예컨대 서버, 데이터베이스 서버, 클라우드 저장, 네트워크 부착 스토리지, 등과 통신하는 하나 이상의 다른 컴퓨팅 머신의 일부일 수 있다.

[0067] 모듈(2050)은 본 명세서에 제시된 다양한 방법 및 처리 기능을 수행하면서 컴퓨팅 머신(2000)을 용이하게 하도록 구성된 하나 이상의 하드웨어 또는 소프트웨어 요소를 포함할 수 있다. 모듈(2050)은 시스템 메모리(2030), 저장 매체(2040), 또는 둘 모두와 관련하여 소프트웨어 또는 펌웨어로서 저장된 하나 이상의 명령 시퀀스를 포함할 수 있다. 따라서 저장 매체(2040)는 프로세서(2010)에 의한 실행을 위해 명령 또는 코드가 저장될 수 있는 머신 또는 컴퓨터 판독가능 매체의 예를 나타낼 수 있다. 머신 또는 컴퓨터 판독가능 매체는 일반적으로 명령을 프로세서(2010)에 제공하기 위해 사용된 임의의 매체 또는 매체들을 지칭할 수 있다. 모듈(2050)과 관련된 그와 같은 머신 또는 컴퓨터 판독가능 매체는 컴퓨터 소프트웨어 제품을 포함할 수 있다. 모듈(2050)을 포함하는 컴퓨터 소프트웨어 제품은 네트워크(2080), 임의의 신호-보유 매체, 또는 임의의 다른 통신 또는 전달 기술을 통해 컴퓨팅 머신(2000)에 모듈(2050)을 전달하기 위한 하나 이상의 프로세스 또는 방법과 또한 연관될 수 있다는 것을 이해해야 한다. 모듈(2050)은 또한 FPGA 또는 다른 PLD에 대한 마이크로코드 또는 구성 정보와 같은 하드웨어 회로를 구성하기 위한 하드웨어 회로 또는 정보를 포함할 수 있다.

[0068] 입력/출력("I/O") 인터페이스(2060)는 하나 이상의 외부 장치에 연결하고, 하나 이상의 외부 장치로부터 데이터를 수신하고, 그리고 데이터를 하나 이상의 외부 장치에 전송하도록 구성될 수 있다. 다양한 내부 장치와 함께 그와 같은 외부 장치는 주변 장치로도 알려져 있을 수 있다. I/O 인터페이스(2060)는 다양한 주변 장치를 컴퓨팅 머신(2000) 또는 프로세서(2010)에 작동가능하게 연결하기 위한 전기적 및 물리적 연결 모두를 포함할 수 있다. I/O 인터페이스(2060)는 주변 장치, 컴퓨팅 머신(2000), 또는 프로세서(2010) 사이에서 데이터를, 어드레스 및 제어 신호를 통신하도록 구성될 수 있다. I/O 인터페이스(2060)는 임의의 표준 인터페이스, 예컨대 소형 컴퓨터 시스템 인터페이스("SCSI"), 직렬 부착 SCSI("SAS"), 파이버 채널, 주변 컴포넌트 인터랙트("PCI"), PCI 익스프레스(PCIe), 직렬 버스, 병렬 버스, 고급 기술 부착("ATA"), 직렬 ATA("SAT A"), 범용 직렬 버스("USB"), 썬더볼트, 파이어 와이어, 다양한 비디오 버스, 등을 구현하도록 구성될 수 있다. I/O 인터페이스(2060)는 단지 하나의 인터페이스 또는 버스 기술을 구현하도록 구성될 수 있다. 대안적으로, I/O 인터페이스(2060)는 다수의 인터페이스 또는 버스 기술을 구현하도록 구성될 수 있다. I/O 인터페이스(2060)는 시스템 버스(2020)의 일부, 전부, 또는 그와 함께 동작하도록 구성될 수 있다. I/O 인터페이스(2060)는 하나 이상의 외부 장치, 내부 장치, 컴퓨팅 머신(2000) 또는 프로세서(2010) 사이의 전송을 버퍼링하기 위한 하나 이상의 버퍼를 포함할 수 있다.

[0069] I/O 인터페이스(2060)는 컴퓨팅 머신(2000)을 마우스, 터치-스크린, 스캐너, 전자적 디지털라이저, 센서, 수신기, 터치패드, 트랙볼, 카메라, 마이크로폰, 키보드, 임의의 다른 포인팅 장치, 또는 이들의 임의의 조합을 포함하는 다양한 입력 디바이스에 연결할 수 있다. I/O 인터페이스(2060)는 컴퓨팅 머신(2000)을 비디오 디스플레이, 스피커, 프린터, 프로젝터, 촉각의 피드백 디바이스, 자동화 제어, 로봇 구성요소, 액추에이터, 모터, 팬, 슬레노이드, 밸브, 펌프, 송신기, 신호 방출기, 조명, 등을 포함하는 다양한 출력 디바이스에 연결할 수 있다.

[0070] 컴퓨팅 머신(2000)은 네트워크(2080)를 통해 하나 이상의 다른 시스템 또는 컴퓨팅 머신으로의 네트워크 인터페이스(2070)를 통한 논리적 접속을 사용하여 네트워크 환경에서 동작할 수 있다. 네트워크(2080)는 광역 네트워크(WAN), 근거리 통신망(LAN), 인트라넷, 인터넷, 무선 액세스 네트워크, 유선 네트워크, 모바일 네트워크, 전화기 네트워크, 광학 네트워크, 또는 이들의 조합을 포함할 수 있다. 네트워크(2080)는 임의의 토폴로지의 패킷 스위칭될 수 있고, 회로 스위칭될 수 있고, 그리고 임의의 통신 프로토콜을 사용할 수 있다. 네트워크(2080) 내의 통신 회선은 다양한 디지털 또는 아날로그 통신 매체 예컨대 광섬유 케이블, 자유 공간 광학, 도파관, 전기 전도체, 무선 링크, 안테나, 무선 주파수 통신, 등을 수반할 수 있다.

[0071] 프로세서(2010)는 시스템 버스(2020)를 통해 컴퓨팅 머신(2000)의 다른 요소 또는 다양한 주변장치에 연결될 수 있다. 시스템 버스(2020)가 프로세서(2010) 내에 있을 수 있고, 프로세서(2010) 외부에 있을 수 있거나, 또는 둘 모두일 수 있는 것으로 인정되어야 한다. 일부 구현예에 따르면, 본 명세서에 논의된 임의의 프로세서(2010), 컴퓨팅 머신(2000)의 다른 요소, 또는 다양한 주변장치는 단일 디바이스 예컨대 시스템 온 칩("SOC"), 시스템 온 패키지("SOP"), 또는 ASIC 디바이스에 통합될 수 있다.

[0072] 구현예는 본 명세서에서 설명되고 예시된 기능들을 구현하는 컴퓨터 프로그램을 포함할 수 있고, 컴퓨터 프로그램은 기계-판독가능 매체에 저장된 명령 및 명령을 실행하는 프로세서를 포함하는 컴퓨터 시스템에서 구현된다. 그러나, 컴퓨터 프로그래밍에서 구현예를 구현하는 많은 상이한 방식들이 존재할 수 있고, 구현예는 임의의 한 세트의 컴퓨터 프로그램 명령으로 제한되는 것으로 해석되어서는 안된다는 것이 명백해야 한다. 또한, 숙련된

프로그래머는 첨부된 흐름도 및/또는 애플리케이션 텍스트에서의 연관된 설명에 기초하여 개시된 구현예들 중 구현예를 구현하기 위해 이러한 컴퓨터 프로그램을 기록할 수 있을 것이다. 따라서, 특정 세트의 프로그램 코드 명령들의 개시는 구현예를 만들고 사용하는 방법에 대한 적절한 이해를 위해 필수적인 것으로 고려되지 않는다. 또한, 당업자는 본 명세서에 기술된 구현예의 하나 이상의 측면이 하나 이상의 컴퓨팅 시스템에서 구현될 수 있는 바와 같이, 하드웨어, 소프트웨어, 또는 이들의 조합에 의해 수행될 수 있다는 것을 인식할 것이다. 또한, 컴퓨터에 의해 수행되는 동작에 대한 임의의 참조는 하나보다 많은 컴퓨터가 동작을 수행할 수 있기 때문에 단일 컴퓨터에 의해 실행되는 것으로 해석되어서는 안된다.

[0073] 본 명세서에 기재된 예시적인 구현예는 앞서 기재된 방법 및 프로세싱 기능을 수행하는 컴퓨터 하드웨어 및 소프트웨어와 함께 사용될 수 있다. 본 명세서에 기재된 시스템, 방법, 및 절차는 프로그래밍가능한 컴퓨터, 컴퓨터-실행가능 소프트웨어, 또는 디지털 회로에서 구현될 수 있다. 소프트웨어는 컴퓨터-판독가능 매체에 저장될 수 있다. 예를 들어, 컴퓨터-판독가능 매체는 플로피 디스크, RAM, ROM, 하드 디스크, 착탈식 매체, 플래시 메모리, 메모리 스틱, 광학 매체, 광자기 매체, CD-ROM, 등을 포함할 수 있다. 디지털 회로는 집적회로, 게이트 어레이, 빌딩 블록 로직, 필드 프로그래머블 게이트 어레이(FPGA), 등을 포함할 수 있다.

[0074] 이전에 제시된 구현예에서 기재된 예시적인 시스템, 방법, 및 동작은 예시적이고, 대안적인 구현예에서, 특정 동작들은 다양한 구현예의 범위 및 사상으로부터 벗어나지 않고, 상이한 순서로, 서로 병렬로, 완전히 생략되어, 및/또는 상이한 예시적인 구현예 사이에서 조합될 수 있고/거나 특정 추가적인 동작들이 수행될 수 있다. 따라서, 이러한 대안적인 구현예는 본 명세서에 기재된 실시예에 포함된다.

[0075] 특정 구현예가 위에서 상세히 설명되었지만, 설명은 단지 예시를 위한 것이다. 따라서, 위에서 설명된 많은 측면이 명시적으로 달리 언급되지 않는 한 요구되거나 필수적인 요소로서 의도되지 않는다는 것이 인정되어야 한다. 전술한 것들에 부가하여, 예시적인 구현예의 개시된 측면의 수정, 및 대응하는 등가의 요소 또는 동작은, 다음의 청구항들에 정의된 구현예의 사상 및 범위로부터 벗어나지 않고, 본 개시내용의 이익을 갖는, 당업자에 의해 이루어질 수 있고, 다음의 청구항들의 범위는 그러한 수정 및 등가의 구조들을 포함하도록 가장 넓은 해석에 따라야 한다.

[0076] 하기 실시예는 특정 구현예의 특정 특징을 예시하기 위해 제공된다. 그러나, 아래에서 설명되는 특정 특징은 본 발명의 범위에 대한 제한들로서 간주되어서는 안되며, 오히려 당업자들에 의해 등가물이 인식될 실시예로서 간주되어야 한다.

[0077] **실시예**

[0078] 하기 실시예는 당업자에게 본 발명의 방법을 제조 및 사용하는 방법에 대한 완전한 개시 및 설명을 제공하기 위해 제시되며, 본 발명자가 본 발명으로 간주하는 것의 범위를 제한하고자 하는 것은 아니다. 사용된 수(예를 들어, 양, 온도 등)에 대한 정확성을 보장하기 위한 노력이 행해졌으나, 일부 실험 오차 및 편차가 고려되어야 한다. 달리 나타내지 않는 한, 부는 중량부이고, 분자량은 지시되지 않는 한 평균 분자량이고, 온도는 섭씨 온도이고, 실온은 약 25°C이고, 압력은 대기압이거나 대기압에 가깝다.

[0079] 소과 혈청 알부민(BSA)를 함유하는 2개의 샘플은 트립신 및 Tryp-N(Cold Spring Harbor 실험실에서 개발되고 Protifi, LLC에서 상업적으로 입수가능한 아르기닌 및 라이신에 대한 N-말단 특이성을 갖는 호열성 메탈로프로테아제)에 의한 소화를 조심스럽게 거친다. 수득한 펩타이드 소화물은 개별적으로 탠덤 질량 분광분석법을 거쳐 BSA의 소화된 펩타이드 단편의 아미노산 서열을 결정하였다. 도 1은 Tryp-N 프로테아제 소화를 사용하는 소혈청 알부민(BSA) 서열 적용범위를 도시한다. 서열 적용범위는 91.4%이다. Tryp-N 소화로부터 생성된 다양한 펩타이드 단편은 BSA 단백질 서열(서열번호: 1) 아래에 나타난다. 도 2는 트립신 프로테아제 소화를 사용한 BSA 서열 적용범위를 도시한다. 서열 적용범위는 94.2%이다. 트립신 소화물로부터 생성된 다양한 펩타이드 단편은 BSA 단백질 서열(서열번호: 1) 아래에 나타난다.

[0080] 각각의 소화된 샘플에서 펩타이드의 서열을 결정하기 위해, 개별 펩타이드를 퀴드루폴로부터 선택하고 충돌 유도된 단편화를 수행하여 b 및 y 펩타이드 단편 이온을 생성하였다(예를 들어, 도, 5a, 5b, 7a, 7b, 9a, 9b, 11a 및 11b를 참조한다).

[0081] 그 다음, 개별 이온 맵은 도 4, 6, 8, 및 10 각각에 도시된 절차에 따라 도 5c, 7c, 9c, 및 11c에서 도시된 바와 같은 개별 펩타이드의 일차 서열을 결정하기 위해 사용되었다. 간단히, 서열 KLVNELTEFAK(서열번호: 2)을 포함하는 폴리펩타이드에 대해 도 4에서 도시된 바와 같이, 트립신에 의한 소화로 펩타이드 LVNELTEFAK(서열번호: 3)를 산출하였다. Tryp-N에 의한 소화로 펩타이드 KLVNELTEFA(서열번호: 4)를 산출하였다. 2개의 펩타이드는 동

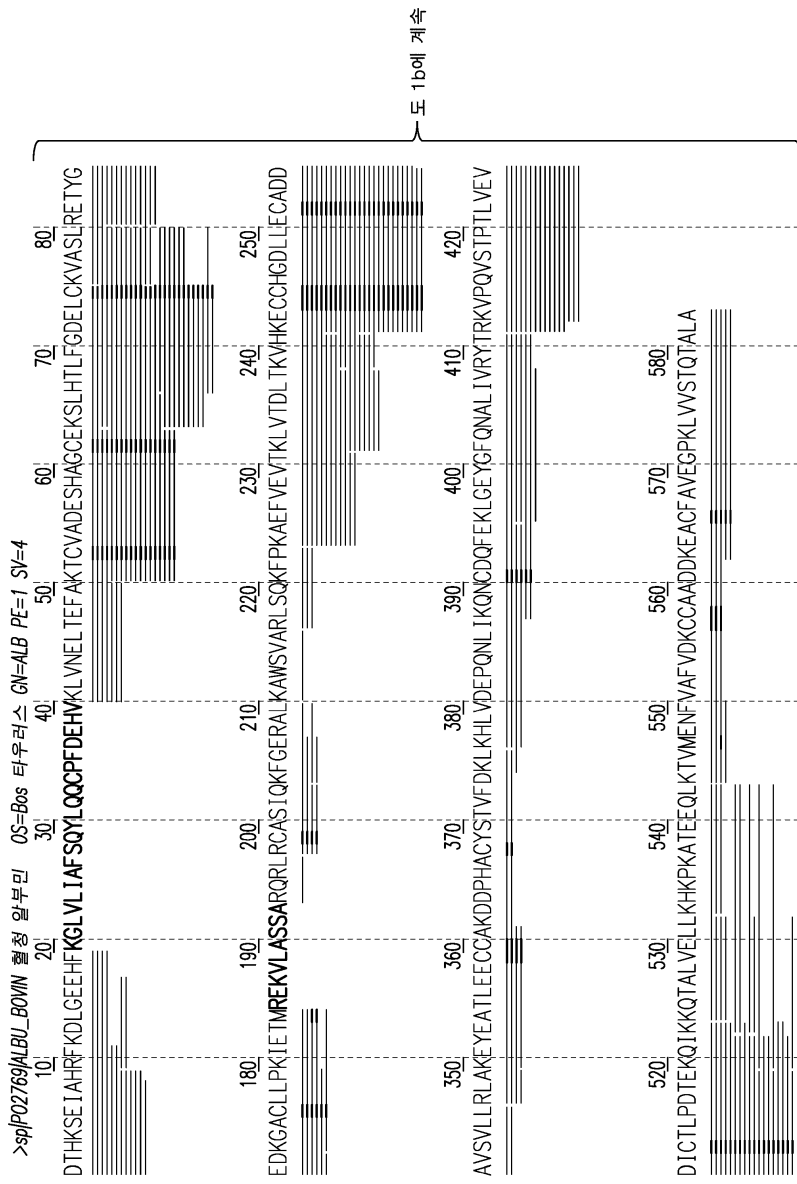
일한 질량을 갖는다. 그러나, 질량 스펙트럼 분석 동안에 단편화될 때, 트립신 소화물로부터의 b 이온 및 y 이온은 Tryp-N 소화물로부터의 b 이온 및 y 이온으로부터의 단일 라이신 잔기에 의해 질량이 상이하다. 도 6의 도시와 유사하게, 트립신으로 소화될 때 서열 RHPEYAVSVLLR(서열번호: 6)을 포함하는 폴리펩타이드는 펩타이드 HPEYAVSVLLR(서열번호: 7)를 산출하였다. Tryp-N에 의한 소화로 펩타이드 RHPEYAVSVLL(서열번호: 8)를 산출하였다. 2개의 펩타이드는 동일한 질량을 갖는다. 그러나, 질량 스펙트럼 분석 동안에 단편화될 때, 트립신 소화물로부터의 b 이온 및 y 이온은 Tryp-N 소화물로부터의 b 이온 및 y 이온으로부터의 단일 아르기닌 잔기에 의해 질량이 상이하다. 약간 상이한 상황은 아르기닌 및 라이신 잔기의 혼합물에 의해 한정된 펩타이드에 대해 관찰되었다. 도 8에서 나타난 바와 같이, 서열 KCCTESLVNR(서열번호: 10)를 포함하는 폴리펩타이드에 대해, 트립신에 의한 소화로 펩타이드 CCTESLVNR(서열번호: 11)를 산출하였다. Tryp-N에 의한 소화로 펩타이드 KCCTESLVN(서열번호: 12)을 산출하였다. 이 경우에 2개의 펩타이드는 동일한 질량을 갖지 않는다. 그러나, 질량 스펙트럼 분석 동안에 단편화될 때, 각 펩타이드로부터의 수득한 b 이온 또는 각 펩타이드로부터의 y 이온은 단일 라이신 아미노산 잔기(b 이온) 또는 단일 아르기닌 아미노산 잔기(y 이온)의 질량에 의해 상이하다. 도 10은, 서열 RFKDLGEEHFK(서열번호: 14)을 포함하는 폴리펩타이드에 대해, 트립신에 의한 소화로 펩타이드 FKDLGEEHFK(서열번호: 15)를 산출하였음을 도시한다. Tryp-N에 의한 소화로 펩타이드 RFKDLGEEHF(서열번호: 16)를 산출하였다. 이 경우에 2개의 펩타이드는 동일한 질량을 갖지 않는다. 그러나, 질량 스펙트럼 분석 동안에 단편화될 때, 각 펩타이드로부터의 수득한 b 이온 또는 각 펩타이드로부터의 y 이온은 단일 라이신 아미노산 잔기(y 이온) 또는 단일 아르기닌 아미노산 잔기(b 이온)에 의해 질량이 상이하다.

[0082] 일단 b 및 y 이온 유형이 결정되면, 동일한 유형의 펩타이드 이온의 목록이 단편 펩타이드 이온의 각 세트에서 생성되고, 아미노산 잔기(들)의 질량에 의한 증분 질량을 갖는 펩타이드 이온의 질량 래더가 목록으로부터 생성된다. 질량 래더에서 2개의 인접한 펩타이드 이온 사이의 질량 차이는 개별 20개 아미노산의 질량을 기준으로 특정 아미노산 잔기(들)에 할당되었다. 트립신 소화물 및 Tryp-N 소화물 둘 다로부터의, 서열번호: 2, 6, 10, 및 14에 제시된 바와 같은 펩타이드에 대한 b 및 y 이온의 세트를 사용하여, 이들로부터 확인된 개별 아미노산 잔기를 사용하여, 이들이 유래된 펩타이드의 일차 서열을 조립하였다. 이어서, 개별 펩타이드를 사용하여 BSA의 일차 서열을 조립하였다.

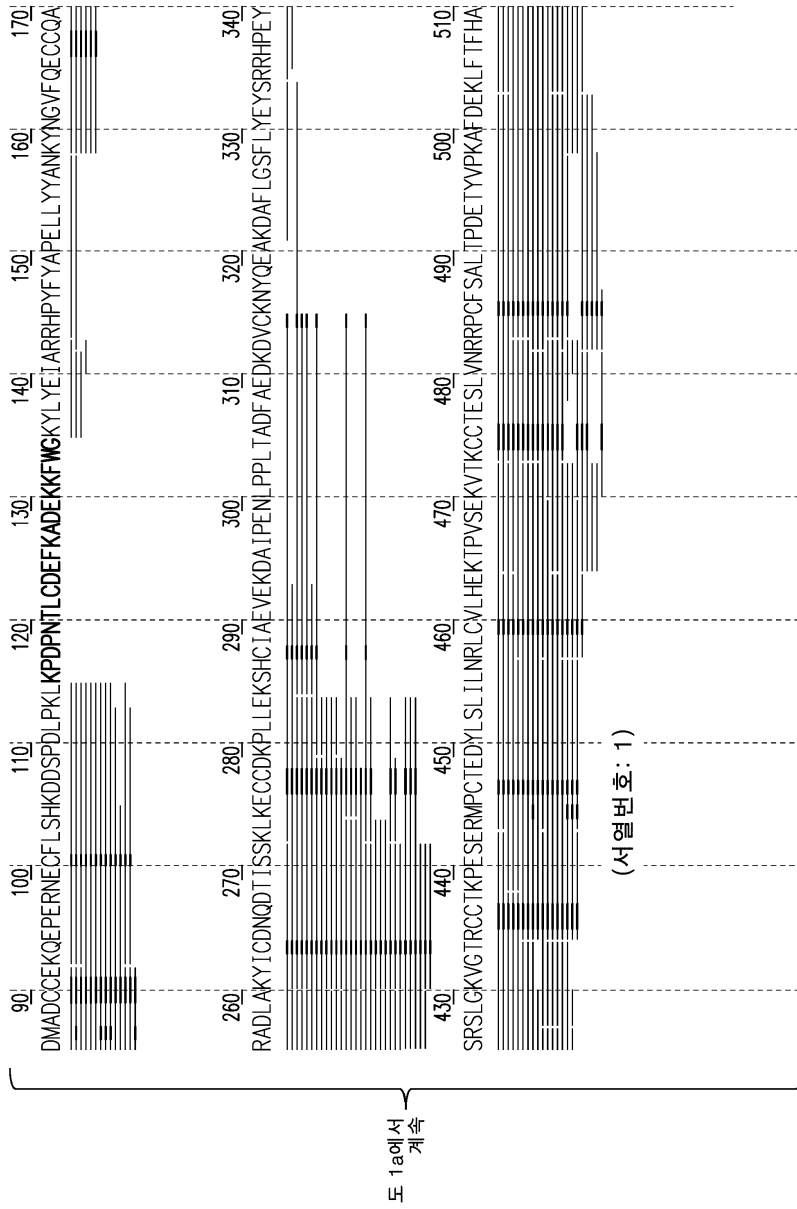
[0083] 본 발명은 본 명세서에 기재된 특정 구현예에 의해 범위가 제한되지 않는다. 실제로, 본 명세서에 설명된 것들에 더하여 본 발명의 다양한 수정들은 전술한 설명 및 수반되는 도면으로부터 당업자에게 명백해질 것이다. 이러한 변형은 첨부된 청구범위의 범위 내에 속하는 것으로 의도된다.

도면

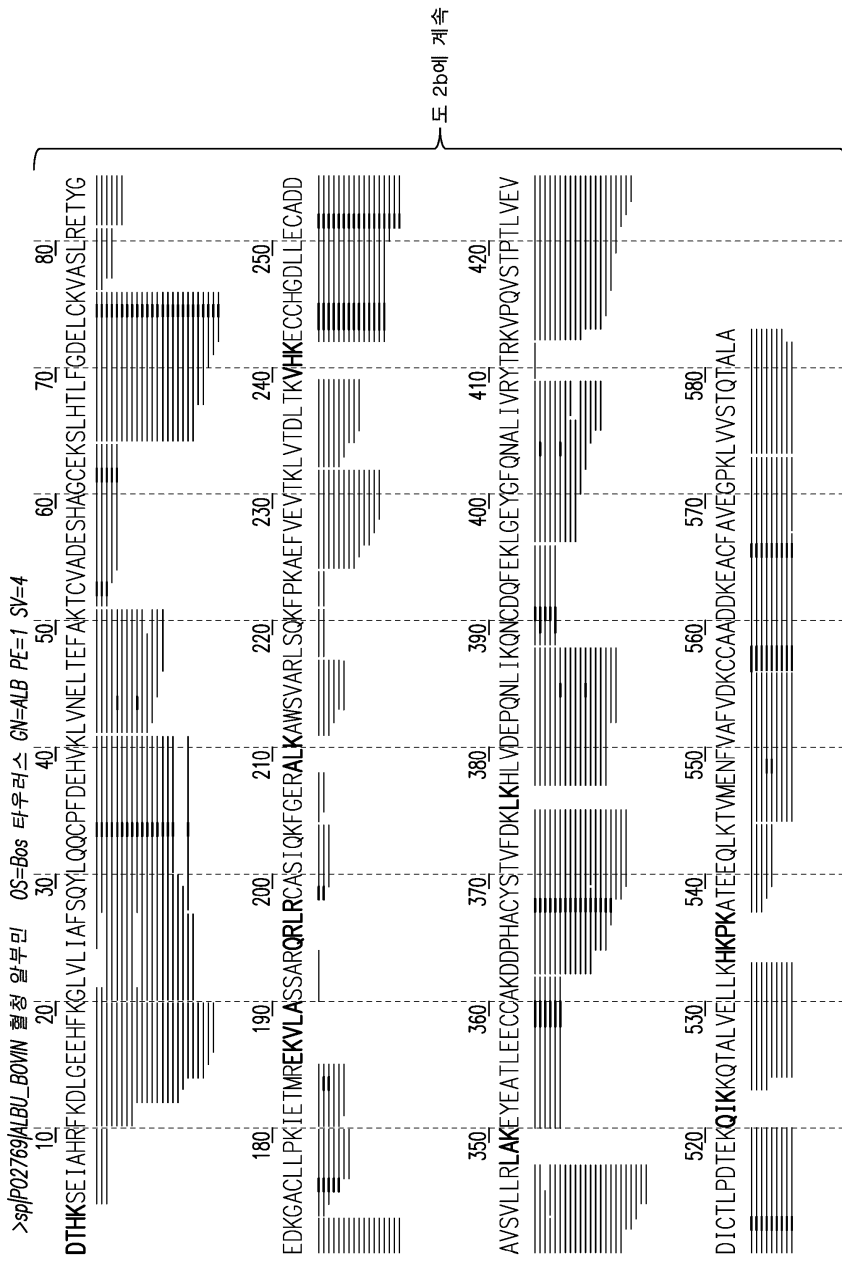
도면1a



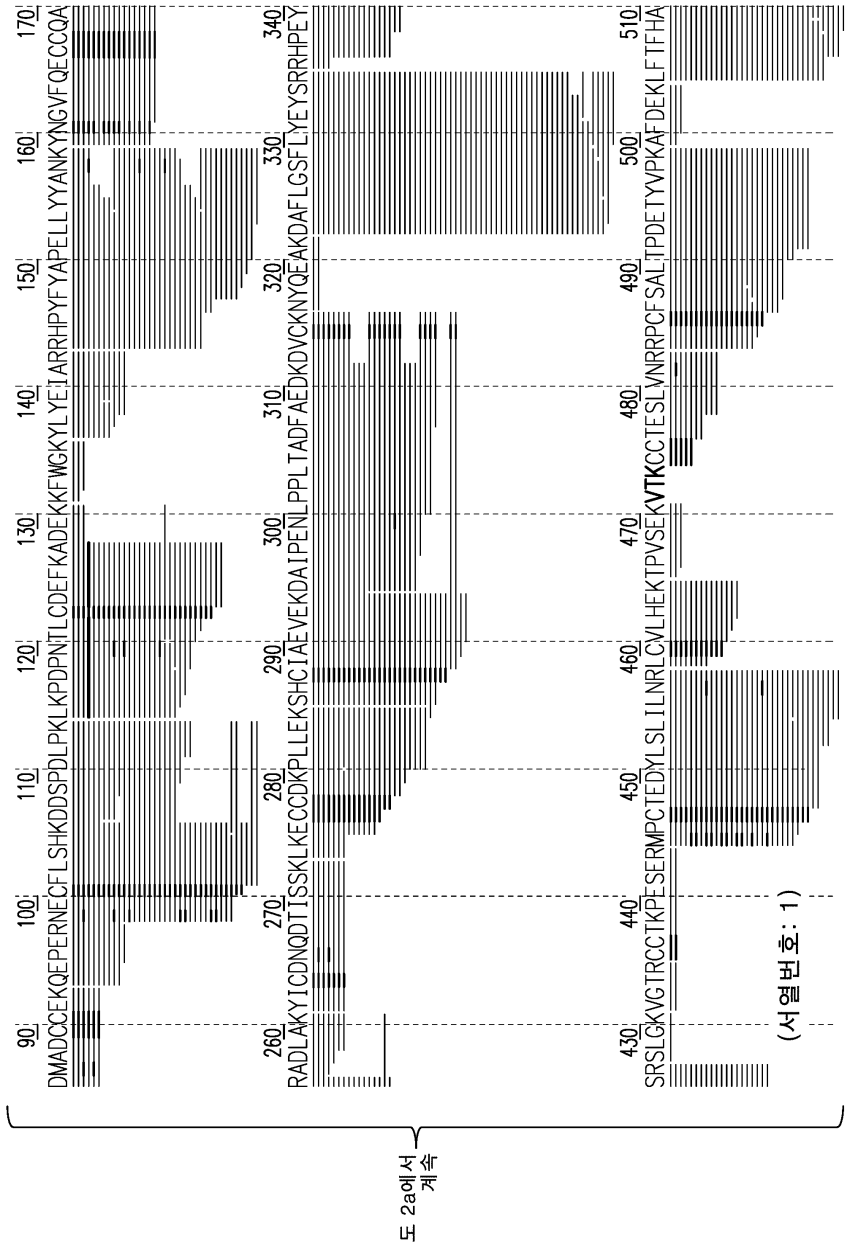
도면1b



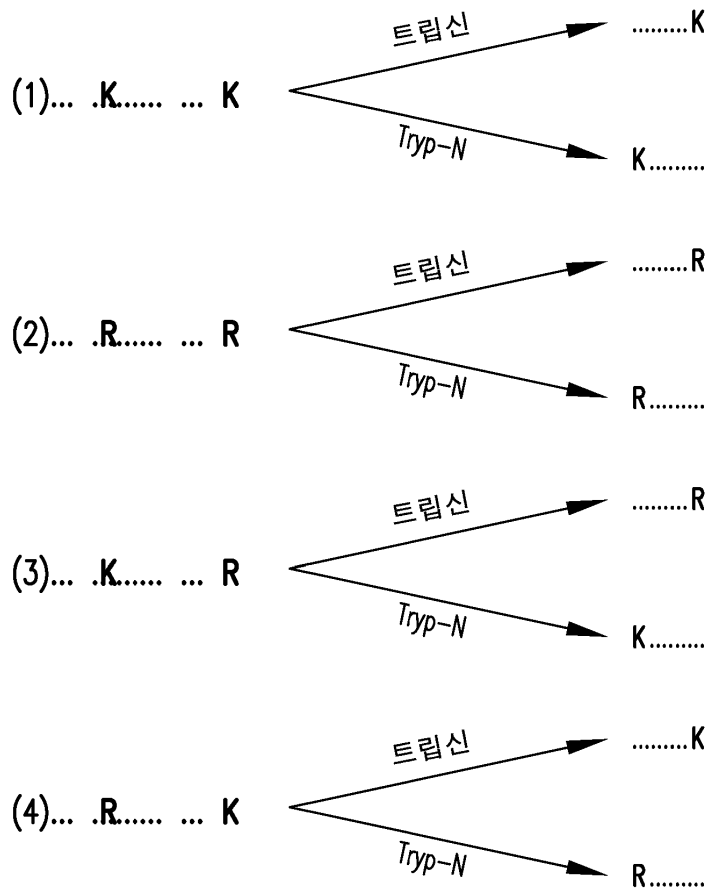
도면2a



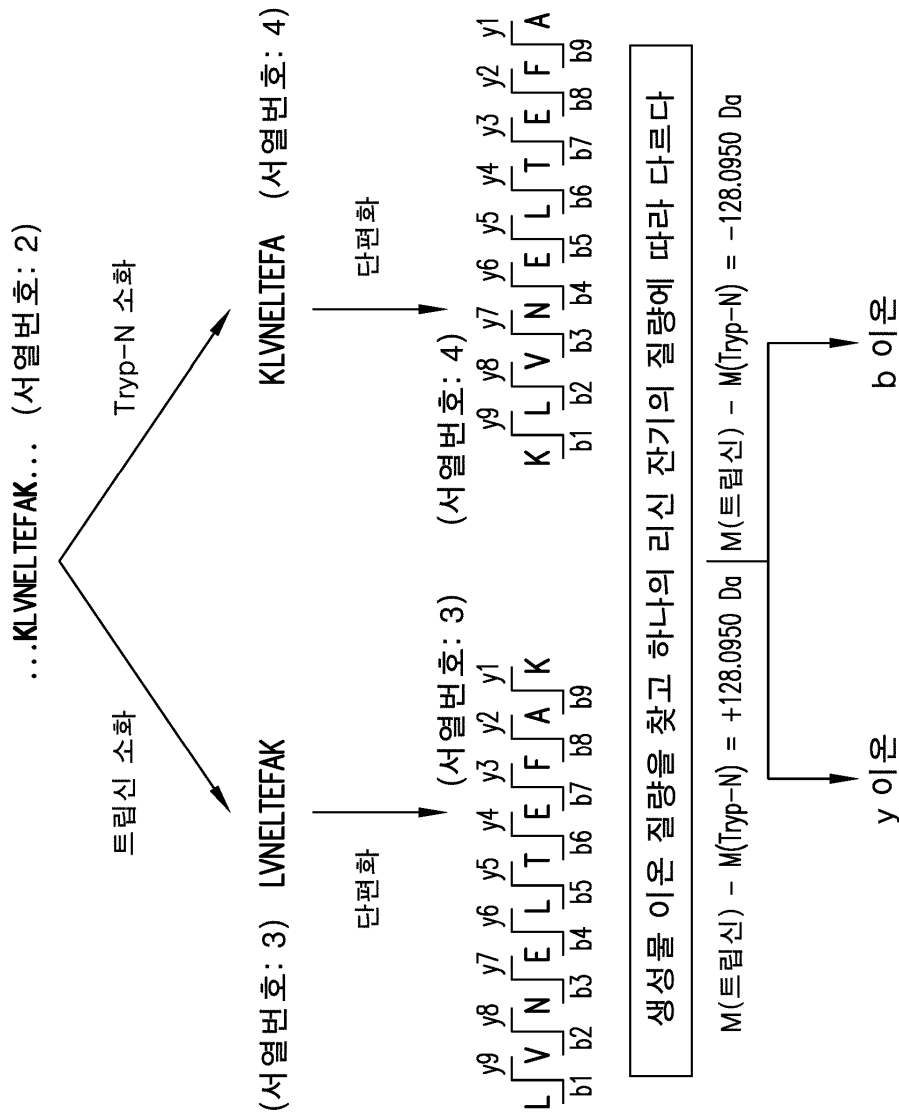
도면2b



도면3



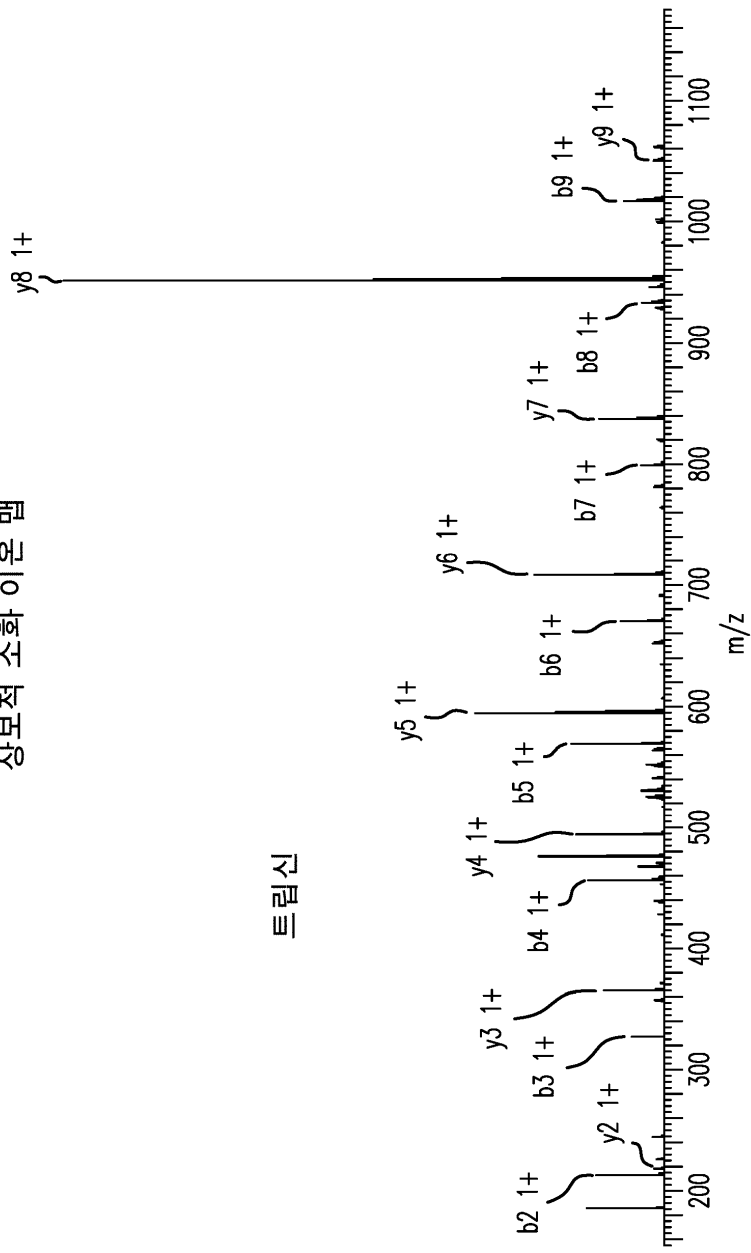
도면4



도면5a

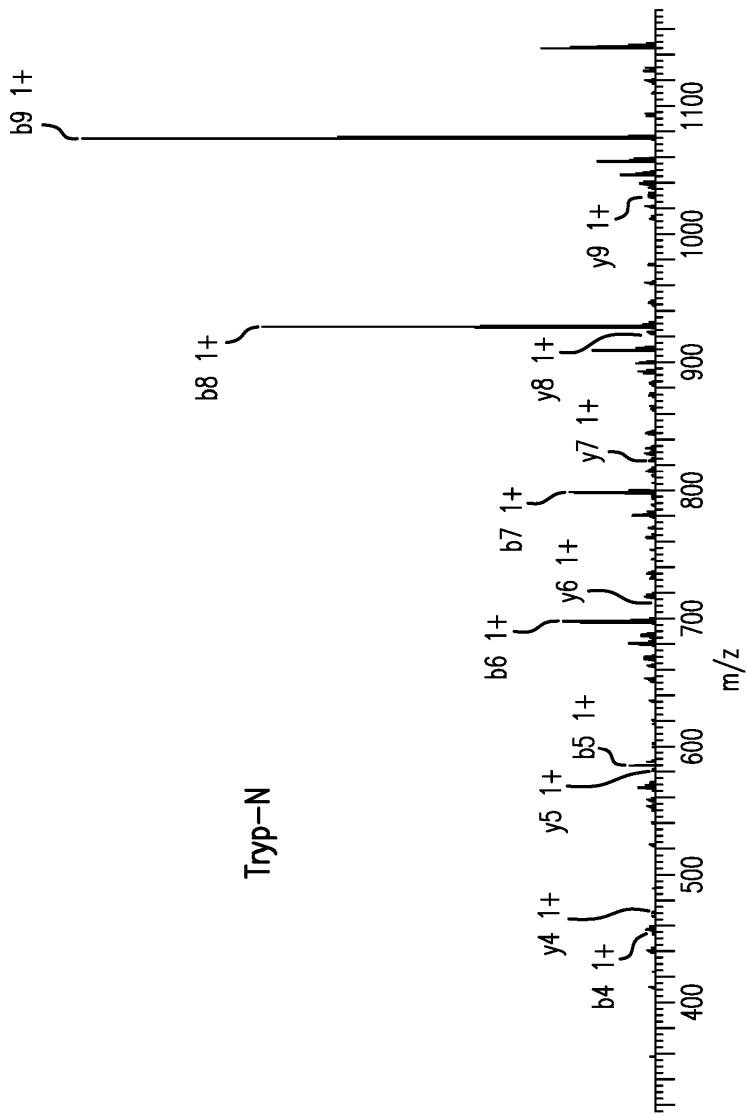
상보적 소화 이온 맵

트립신

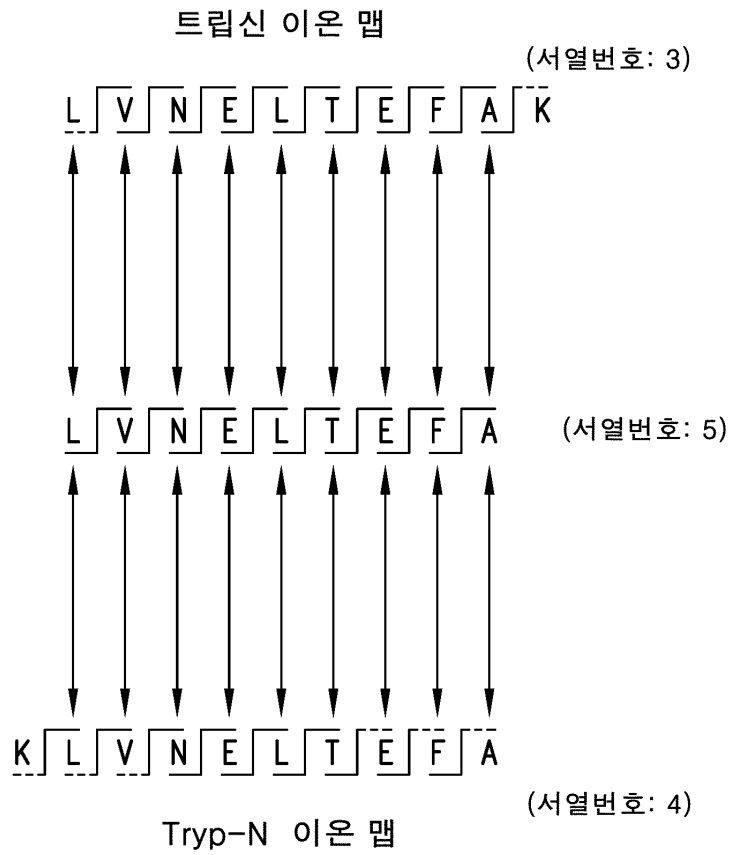


도면5b

상보적 소화 이온 맵

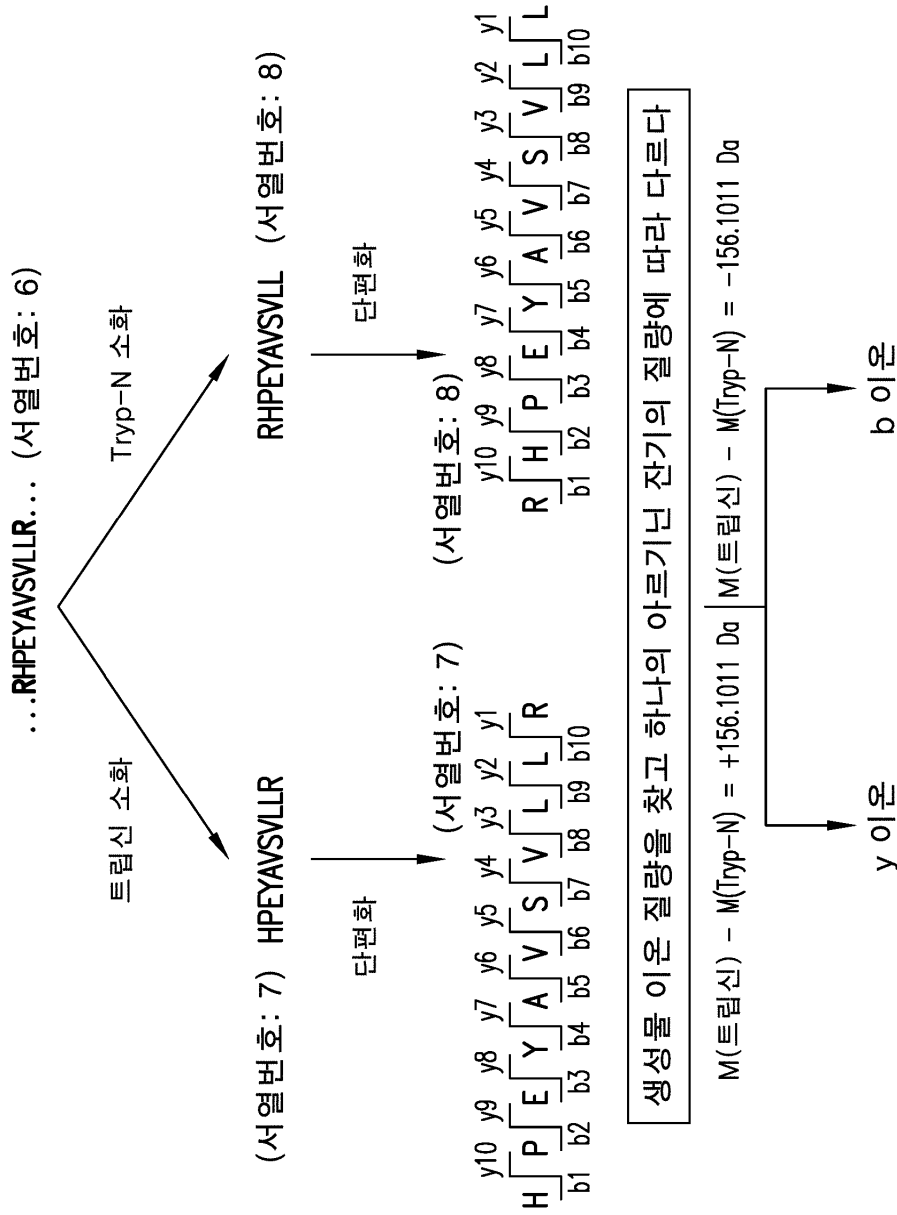


도면5c

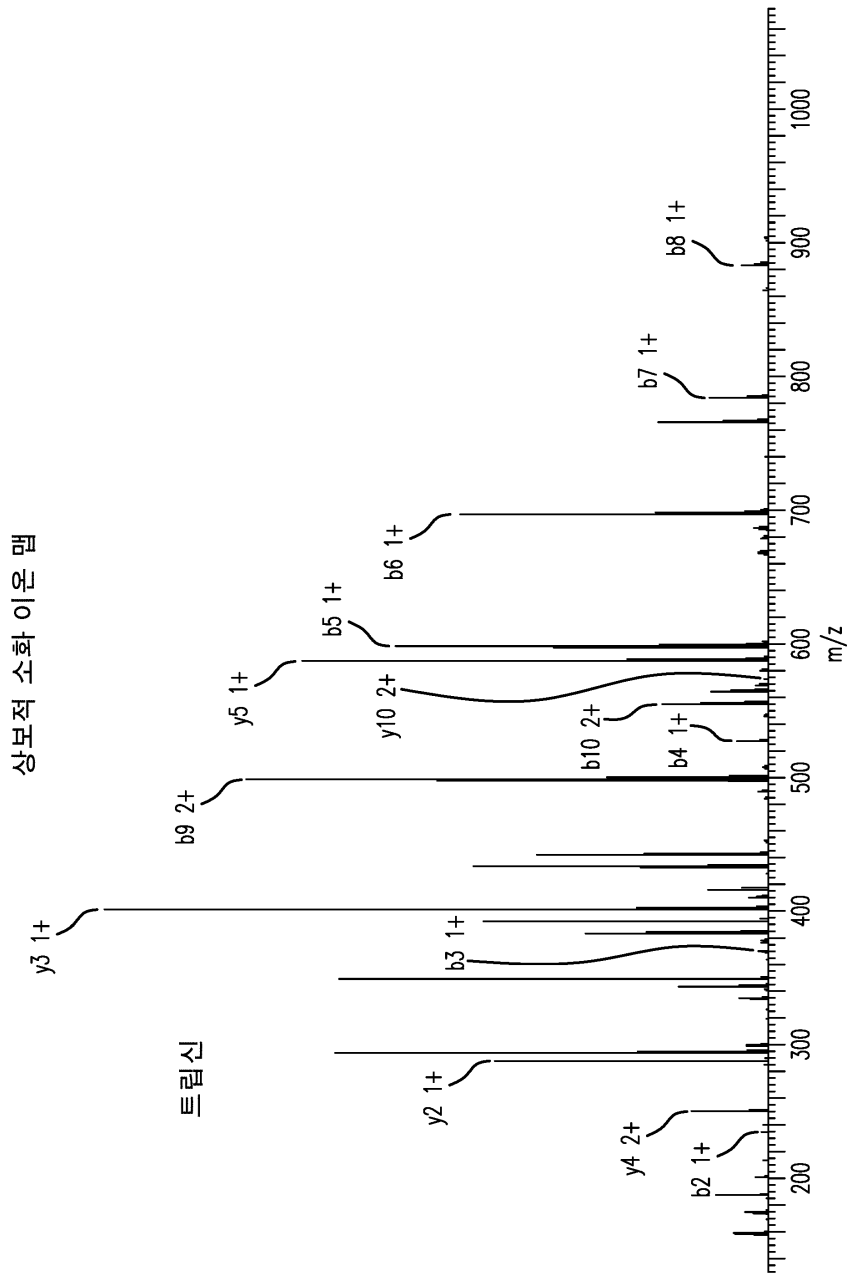


*점선은 단편 이온이 MS2에서 관찰되지 않았지만
y 이온의 질량을 기준으로 계산될 수 있음을 나타낸다

도면6

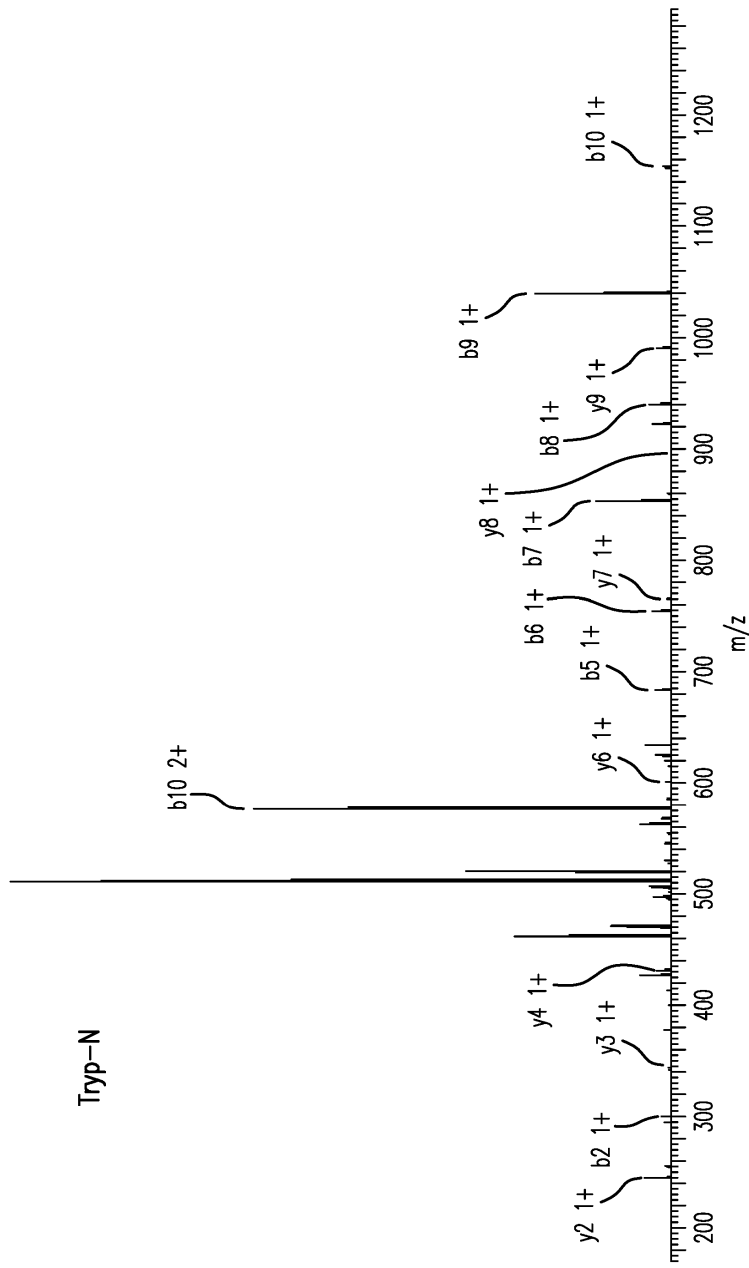


도면7a

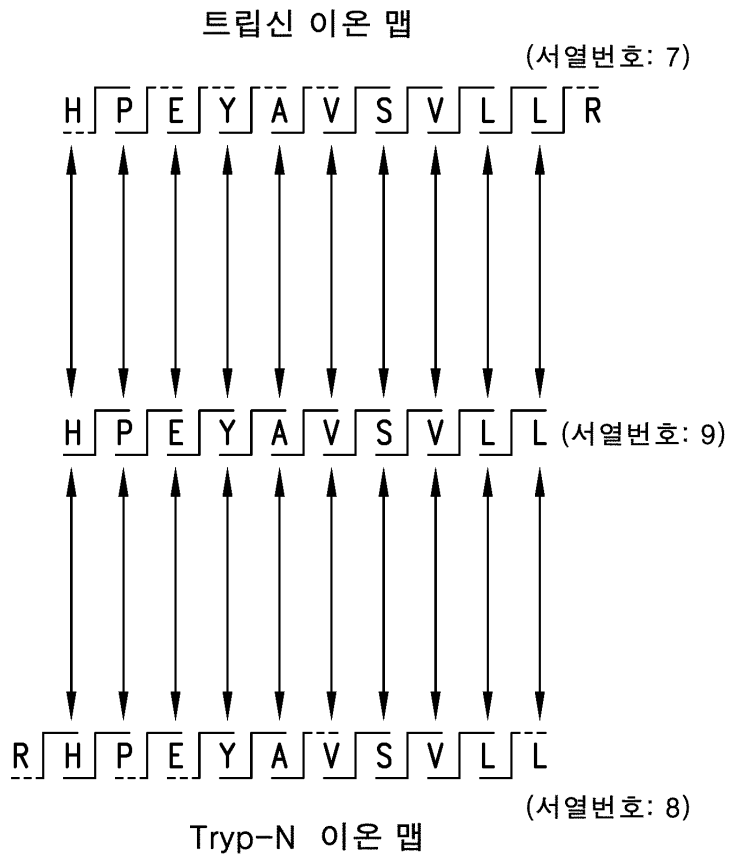


도면7b

상보적 소화 이온 맵

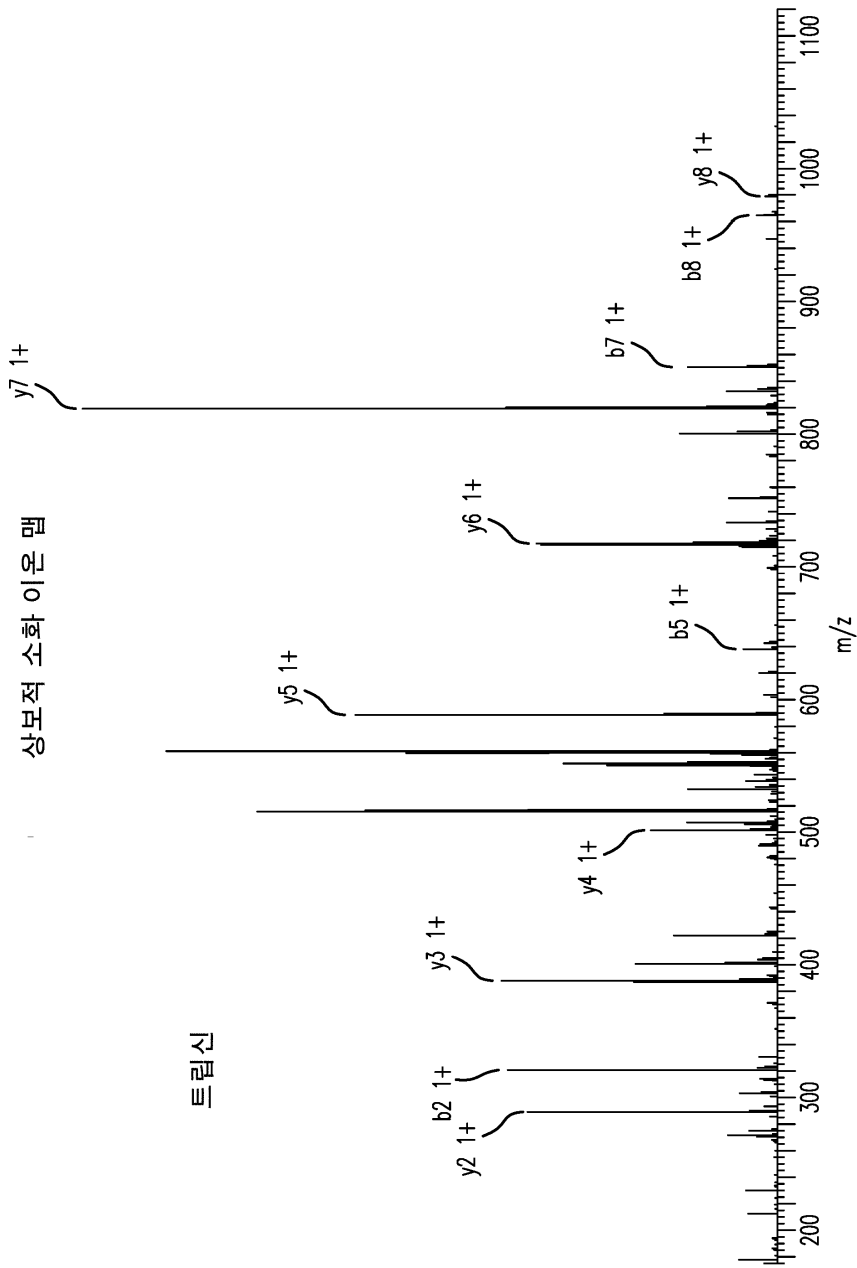


도면7c



*점선은 단편 이온이 MS2에서 관찰되지 않았지만 y 이온의 질량을 기준으로 계산될 수 있음을 나타낸다

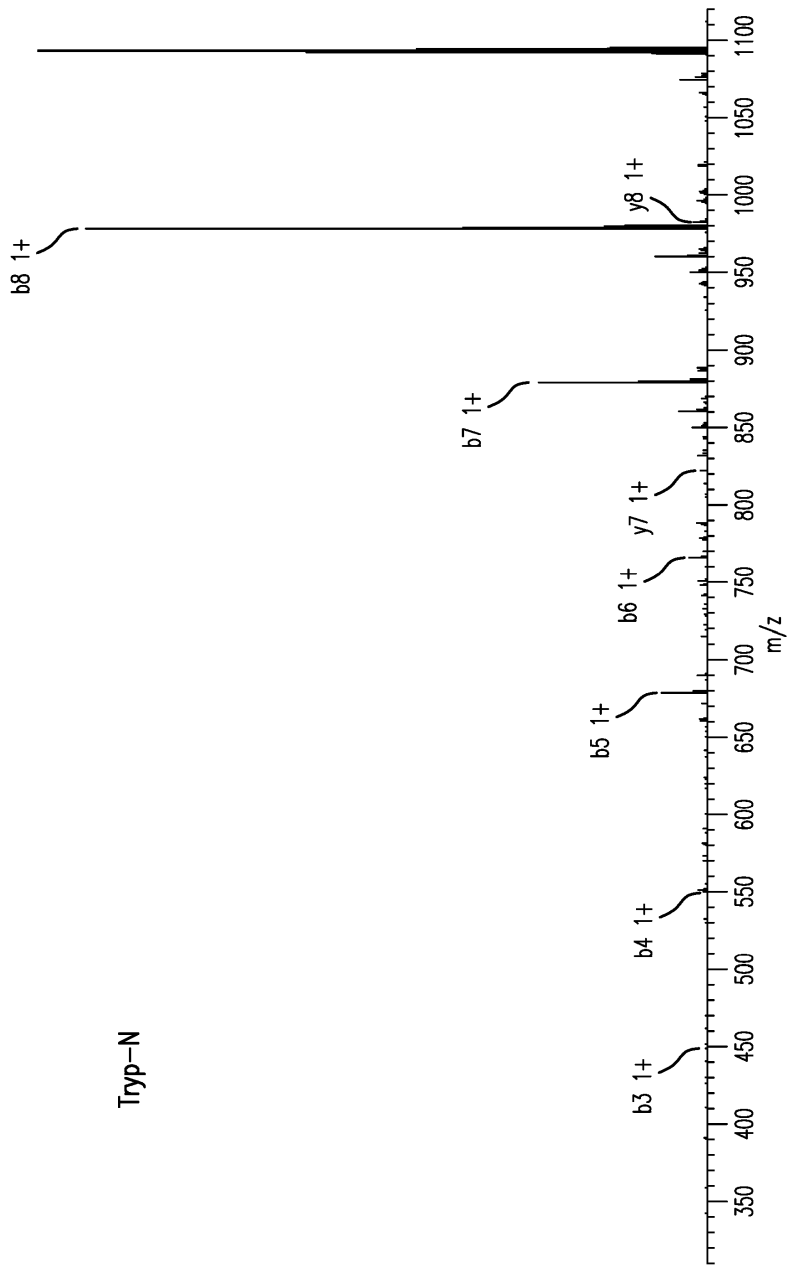
도면9a



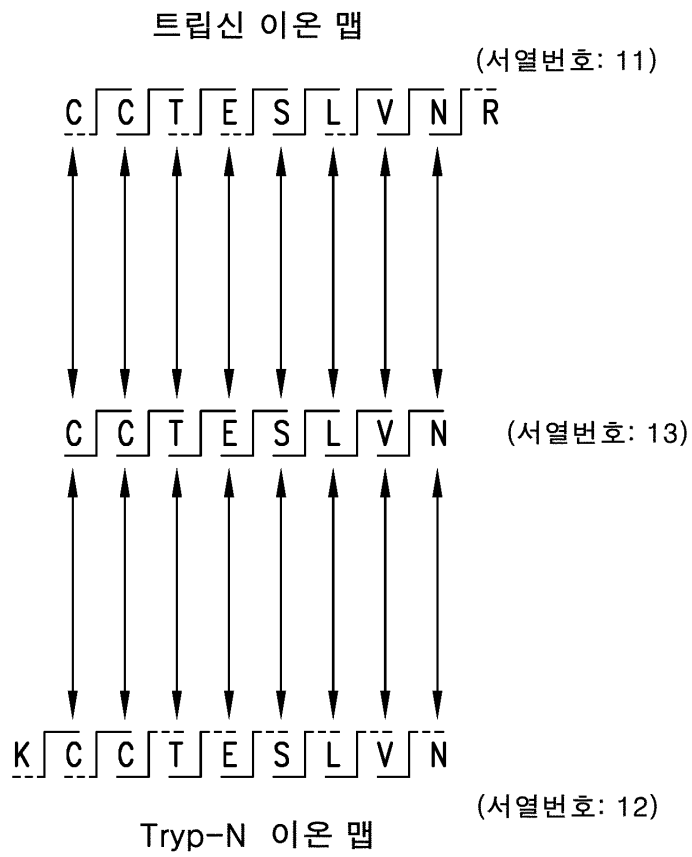
도면9b

상보적 소화 이온 맵

Tryp-N

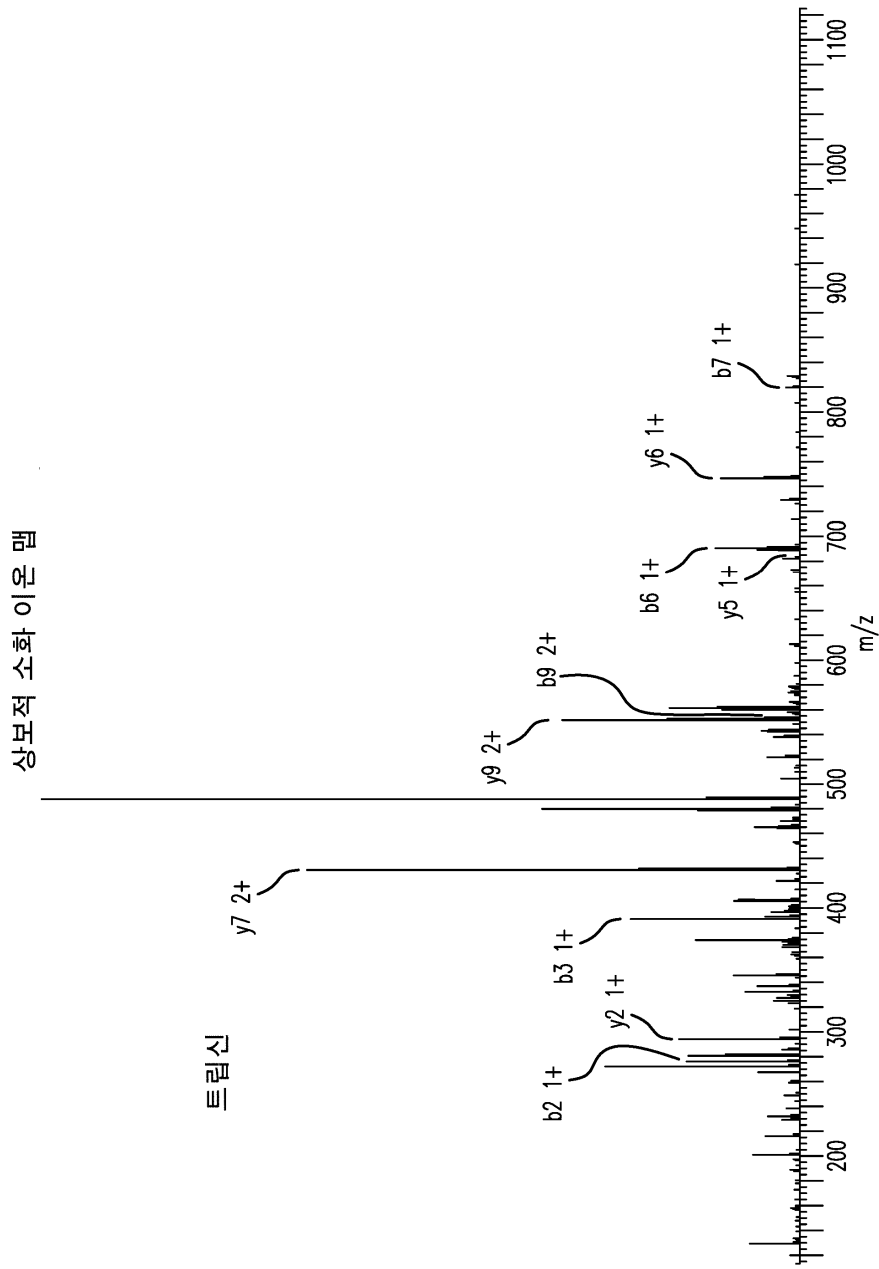


도면9c

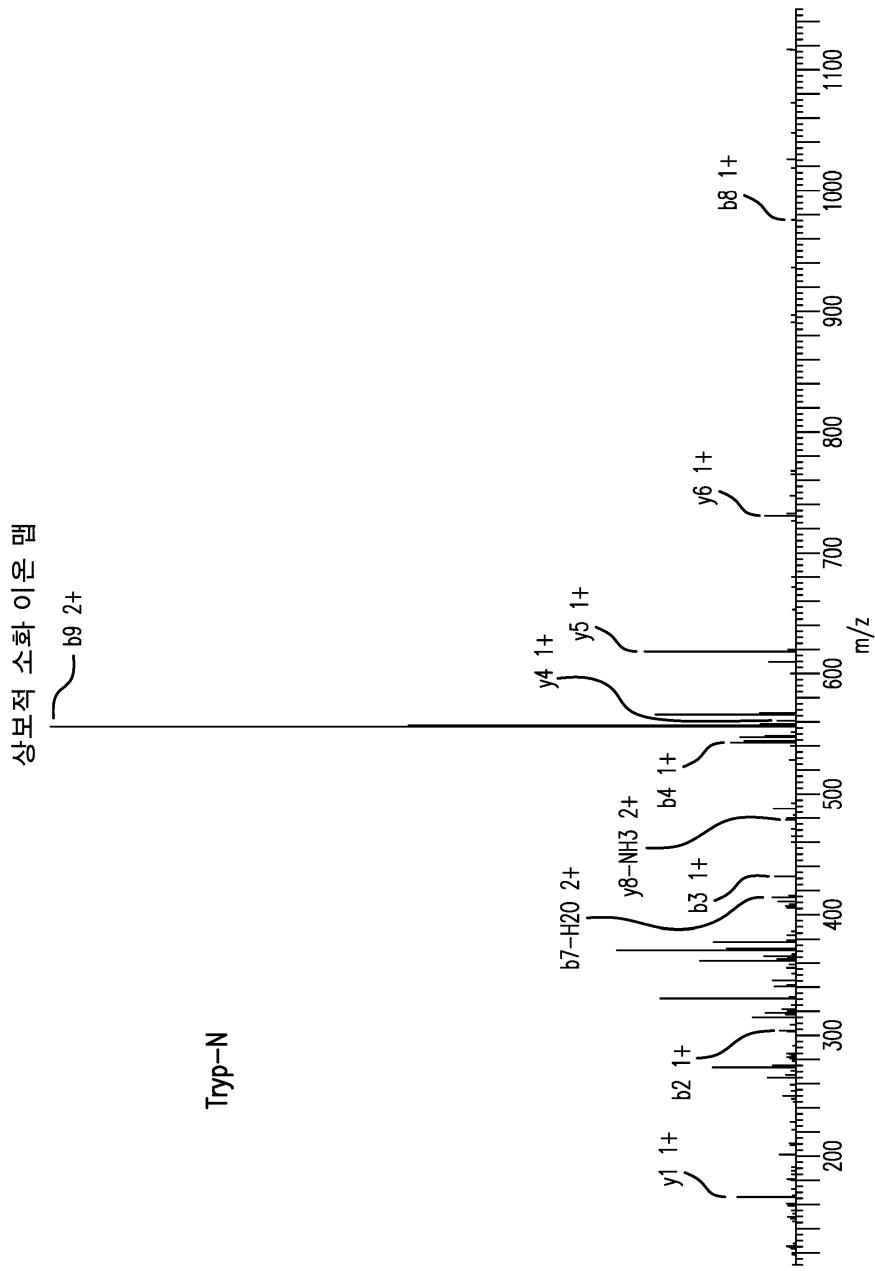


*점선은 단편 이온이 MS2에서 관찰되지 않았지만
y 이온의 질량을 기준으로 계산될 수 있음을 나타낸다

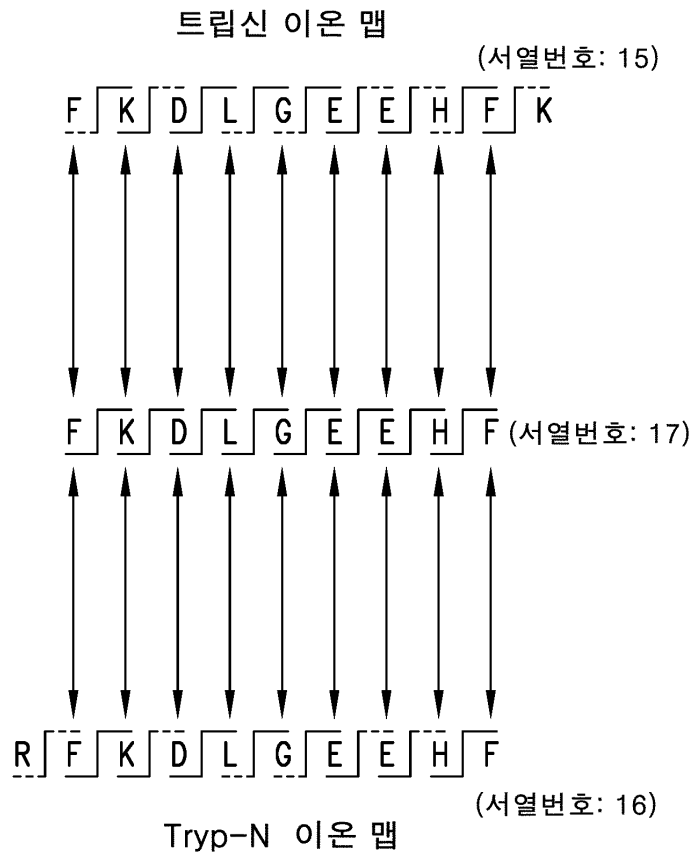
도면11a



도면11b

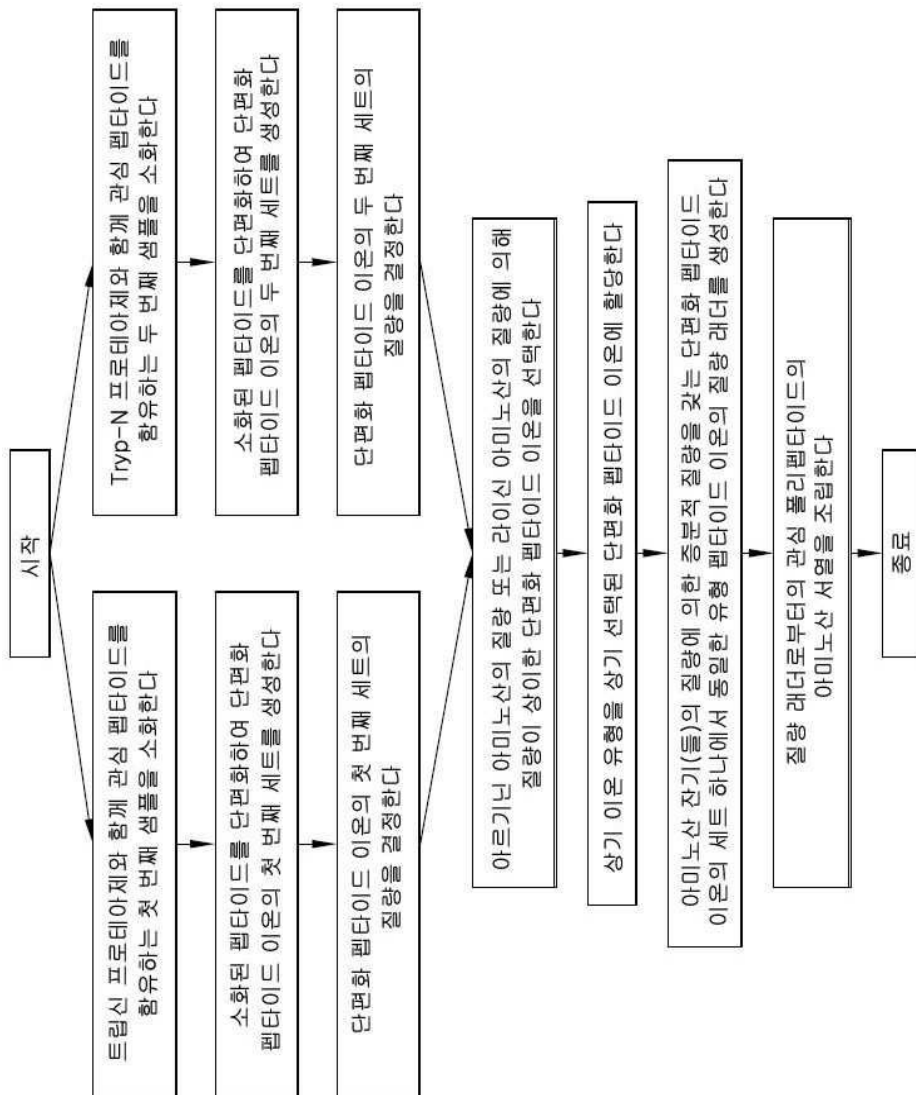


도면11c

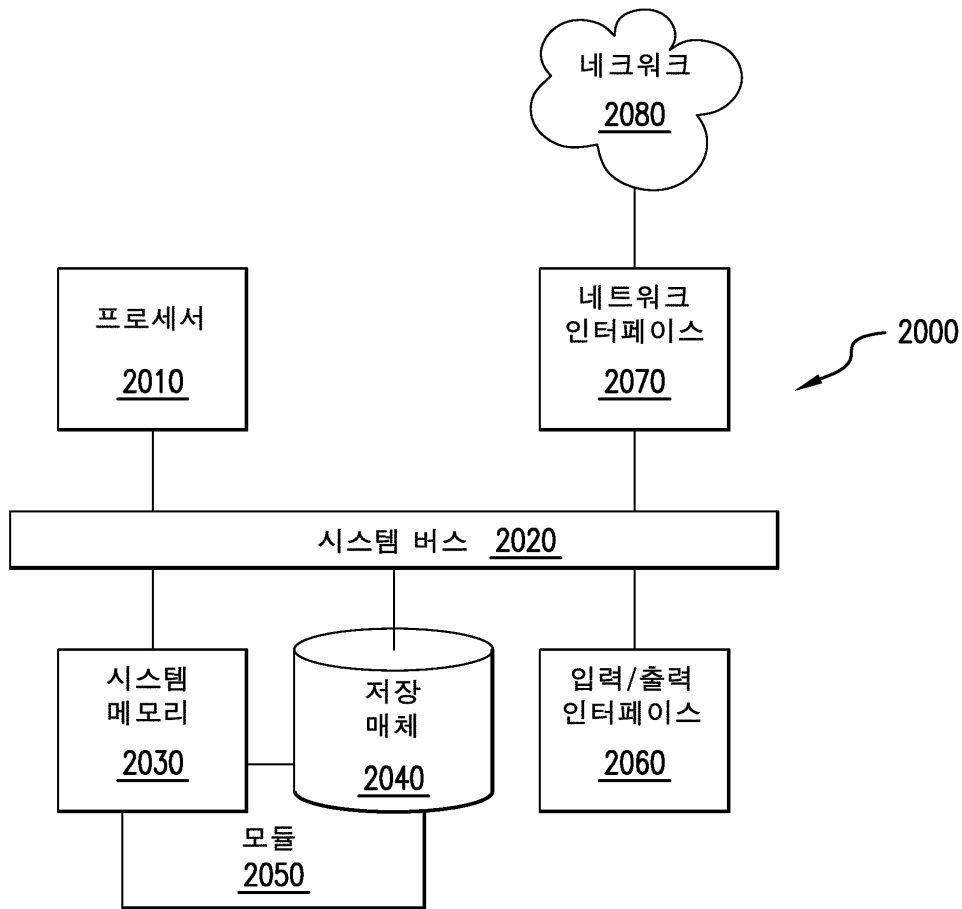


*점선은 단편 이온이 MS2에서 관찰되지 않았지만 y 이온의 질량을 기준으로 계산될 수 있음을 나타낸다

도면12



도면13



서열 목록

<110> Regeneron Pharmaceuticals, Inc.

<120> METHODS FOR DE NOVO PROTEIN SEQUENCING

<130> 10478W001

<150> 62/719,292

<151> 2018-08-17

<160> 17

<170> PatentIn version 3.5

<210> 1

<211> 583

<212> PRT

<213> Bos taurus

<400> 1

Asp Thr His Lys Ser Glu Ile Ala His Arg Phe Lys Asp Leu Gly Glu

1

5

10

15

Glu His Phe Lys Gly Leu Val Leu Ile Ala Phe Ser Gln Tyr Leu Gln
 20 25 30

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

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

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

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

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

Cys Thr Leu Pro Asp Thr Glu Lys Gln Ile Lys Lys Gln Thr Ala Leu
 515 520 525

Val Glu Leu Leu Lys His Lys Pro Lys Ala Thr Glu Glu Gln Leu Lys
 530 535 540

Thr Val Met Glu Asn Phe Val Ala Phe Val Asp Lys Cys Cys Ala Ala
 545 550 555 560

Asp Asp Lys Glu Ala Cys Phe Ala Val Glu Gly Pro Lys Leu Val Val
 565 570 575

Ser Thr Gln Thr Ala Leu Ala
 580

<210> 2

<211> 11

<212> PRT

<213> Bos taurus

<400> 2

Lys Leu Val Asn Glu Leu Thr Glu Phe Ala Lys

1 5 10

<210> 3

<211> 10

<212>

> PRT

<213> Bos taurus

<400> 3

Leu Val Asn Glu Leu Thr Glu Phe Ala Lys

1 5 10

<210> 4

<211> 10

<212> PRT

<213> Bos taurus

<400> 4

Lys Leu Val Asn Glu Leu Thr Glu Phe Ala

1 5 10

<210> 5

<211> 9

<212> PRT

<213> Bos taurus

<400> 5

Leu Val Asn Glu Leu Thr Glu Phe Ala

1 5

<210> 6

<211> 12

<212> PRT

<213> Bos taurus

<400> 6

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

1 5 10

<210> 7

<211> 11

<212> PRT

<213> Bos taurus

<400> 7

His Pro Glu Tyr Ala Val Ser Val Leu Leu Arg

1 5 10

<210> 8

<211> 11

<212> PRT

<213> Bos taurus

<400> 8

Arg His Pro Glu Tyr Ala Val Ser Val Leu Leu

1 5 10

<210> 9

<211> 10

<212> PRT

<213> Bos taurus

<400> 9

His Pro Glu Tyr Ala Val Ser Val Leu Leu

1 5 10

<210> 10
<211> 10
<212> PRT
<213> Bos taurus
<400> 10
Lys Cys Cys Thr Glu Ser Leu Val Asn Arg
1 5 10

<210> 11
<211> 9
<212> PRT
<213> Bos taurus
<400> 11
Cys Cys Thr Glu Ser Leu Val Asn Arg
1 5

<210> 12
<211> 9
<212> PRT
<213> Bos taurus
<400> 12
Lys Cys Cys Thr Glu Ser Leu Val Asn
1 5

<210> 13
<211> 8
<212> PRT
<213> Bos taurus
<400> 13
Cys Cys Thr Glu Ser Leu Val Asn
1 5

<210> 14
<211> 11
<212> PRT
<213> Bos taurus
<400> 14
Arg Phe Lys Asp Leu Gly Glu Glu His Phe Lys

1 5 10

<210> 15

<211> 10

<212> PRT

<213> Bos taurus

<400> 15

Phe Lys Asp Leu Gly Glu Glu His Phe Lys

1 5 10

<210> 16

<211> 10

<212> PRT

<213> Bos taurus

<400> 16

Arg Phe Lys Asp Leu Gly Glu Glu His Phe

1 5 10

<210> 17

<211> 9

<212> PRT

<213> Bos taurus

<400> 17

Phe Lys Asp Leu Gly Glu Glu His Phe

1 5