(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0107990
(43) 공개일자 2024년07월09일

(51) 국제특허분류(Int. Cl.)
G16B 20/00 (2019.01) *G16B 30/10* (2019.01)
G16B 5/00 (2019.01)
(52) CPC특허분류
G16B 20/00 (2019.02)
G16B 30/10 (2019.02)
(21) 출원번호 10-2022-0191149
(22) 출원일자 2022년12월30일
심사청구일자 2022년12월30일

(71) 출원인
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
(72) 발명자
김현수
대전광역시 유성구 대학로 99, 충남대학교 생명시스템과학대학 113호
박종함
대전시 서구 월평북로 13, 210동 1104호(월평동, 은뜰마을)
손아름
10550 North Torrey Pines Rd, SR 302BScripps Research, LaJolla CA USA
(74) 대리인
이은철, 김재문

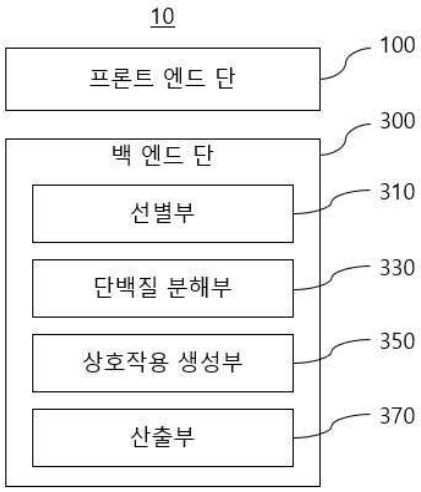
전체 청구항 수 : 총 7 항

(54) 발명의 명칭 **타겟 질량분석법 기반의 단백질 상호작용체 분석 시스템 및 방법**

(57) 요 약

본 발명은, 단백질 상호작용에 연관된 복수개의 데이터를 입력받는 프론트 엔드 단; 및 상기 프론트 엔드 단에서 입력받은 데이터를 이용하여 가능한 이론적 질량 값을 모두 산출하는 백 엔드 단;을 포함하고, 상기 프론트 엔드 단은, 상기 백 엔드 단에서 산출된 이론적 질량 값을 출력하는 것을 일 특징으로 한다.

대 표 도 - 도1



(52) CPC특허분류
G16B 5/00 (2019.02)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711172806
과제번호 2022R1F1A1074235
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 개인기초연구(과기정통부)
연구과제명 구조단백체 데이터 기반의 간암 치료표적 신규 발굴
기 여 율 1/3
과제수행기관명 충남대학교
연구기간 2022.06.01 ~ 2023.05.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711140348
과제번호 2021R1G1A1092933
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 개인기초연구(과기정통부)(R&D)
연구과제명 알츠하이머병 조기예측을 위한 질량분석법 기반 치매취약유전자 (ApoE ε4) 단백질 분석법 개발
기 여 율 1/3
과제수행기관명 충남대학
연구기간 2021.09.01 ~ 2022.08.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711174038
과제번호 00155857
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 인공지능융합혁신인재양성
연구과제명 인공지능융합혁신인재양성(충남대학교)
기 여 율 1/3
과제수행기관명 충남대학교산학협력단
연구기간 2022.07.01 ~ 2025.12.31

명 세 서

청구범위

청구항 1

단백질 상호작용에 연관된 복수개의 데이터를 입력받는 프론트 엔드 단; 및

상기 프론트 엔드 단에서 입력받은 데이터를 이용하여 가능한 이론적 질량 값을 모두 산출하는 백 엔드 단;을 포함하고,

상기 프론트 엔드 단은,

상기 백 엔드 단에서 산출된 이론적 질량 값을 출력하는 것을 특징으로 하는 단백질 상호작용체 분석 시스템.

청구항 2

제 1 항에 있어서,

상기 프론트 엔드 단은,

단백질 상호작용에 연관된 복수개의 데이터로 표적 단백질에 관한 데이터, 상기 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터, 효소에 대한 데이터, 크로스링커에 대한 데이터, 변이에 대한 데이터, 펩타이드 길이에 대한 데이터, 또는 전하에 대한 데이터를 입력받는 것을 특징으로 하는 단백질 상호작용체 분석 시스템.

청구항 3

제 1 항에 있어서,

상기 백 엔드 단은,

표적 단백질에 관한 데이터와 상기 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터로 상기 표적 단백질과 상호작용하는 특정 단백질을 확률에 따라 선별하는 선별부;를 포함하는 것을 특징으로 하는 단백질 상호작용체 분석 시스템.

청구항 4

제 3 항에 있어서,

상기 백 엔드 단은,

표적 단백질에 대한 데이터, 효소에 대한 데이터, 펩타이드 길이에 대한 데이터, 상기 선별부에서 선별된 상호작용하는 특정 단백질의 확률, 크로스링커에 대한 데이터, 및 변이에 대한 데이터를 이용하여,

표적 단백질과 상호작용 가능한 특정 단백질의 아미노산 서열을 효소 처리하였을 때 펩타이드 단편으로 분해되는 경우의 수를 생성하는 단백질 분해부;를 더 포함하는 것을 특징으로 하는 단백질 상호작용체 분석 시스템.

청구항 5

제 4 항에 있어서,

상기 백 엔드 단은,

상기 단백질 분해부가 생성한 단편화된 펩타이드의 길이, 전하에 대한 데이터, 크로스링커에 대한 데이터, 변이

에 대한 데이터를 이용하여,

단백질-단백질 상호작용에 가능한 모든 경우의 수를 생성하는 상호작용 생성부;를 더 포함하는 것을 특징으로 하는 단백질 상호작용체 분석 시스템.

청구항 6

제 5 항에 있어서,

상기 백 엔드 단은,

상기 상호작용 생성부가 생성한 단백질-단백질 상호작용에 가능한 모든 경우의 수에 따른 표적 단백질의 질량값을 산출하는 산출부;를 더 포함하는 것을 특징으로 하는 단백질 상호작용체 분석 시스템.

청구항 7

표적 단백질에 관한 데이터, 상기 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터, 효소에 대한 데이터, 크로스링커에 대한 데이터, 변이에 대한 데이터, 펩타이드 길이에 대한 데이터, 또는 전하에 대한 데이터를 입력받는 입력단계; 및

상기 입력단계에서 입력받은 데이터를 이용하여 가능한 이론적 질량 값을 모두 산출하는 연산단계; 및

상기 연산단계에서 산출된 이론적 질량 값을 출력하는 출력단계;를

포함하는 단백질 상호작용체 분석 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 타겟 질량분석법 기반의 단백질 상호작용체 분석 시스템 및 방법에 관한 것으로서, 특히 XL-MS 방법을 질량분석에 적용하기 위해 단백질의 이론적인 질량 값의 경우의 수를 산출할 수 있는 타겟 질량분석법 기반의 단백질 상호작용체 분석 시스템 및 방법에 관한 것이다.

배 경 기 술

[0003] 단백질-단백질 상호작용은 복잡한 상호작용 네트워크의 형성과 세포의 기능적 일꾼을 나타내는 다중 소단위 단백질 복합체의 조립에 기본적이다. 따라서 이러한 다량체 및 기능적 실체의 구조 및 역학에 대한 자세한 이해는 생물학적 기능을 이해하는 데 중요하다. 현재까지 대부분의 구조 정보는 X선 결정학 및 핵자기 공명(NMR) 분광법에 의해 제공되었으며 누적적으로 모든 PDB(Protein Data Bank) 항목의 98% 이상(각각 89% 및 9%)을 차지한다. 그러나 이러한 전통적인 기술은 샘플 준비 요구 사항에 의해 제한되어 크고 역동적인 단백질 복합체에 적용할 수 없다. cryoelectron microscopy(cryo-EM) 검출기 기술 및 디지털 이미지 처리 알고리즘의 발전으로 이제 단백질 복합 구조 규명을 위한 near-atomic resolution 맵이 허용되지만, cryo-EM 밀도 맵은 여전히 모든 PDB 항목의 약 1%만 나타낸다.

[0004] 질량분석법(MS) 기반 구조 기술은 구조적 및 구성적으로 이질적인 단백질 복합체의 구조와 동역학을 해결하기 위한 전통적인 생물물리학적 구조 방법의 무능력으로 인해 지난 10년 동안 큰 발전을 이루었다. 질량분석법 기반 구조 기술은 개별적으로 완전한 구조를 제공할 수는 없지만 이러한 접근 방식의 장점은 다용도성, 감도 및 처리량으로, 기존 구조 분석을 보완하는 유용한 정보를 제공할 수 있다. 질량분석법 기반 구조 기술은 X선 결정학 및 전자 현미경과 같은 구조 도구와 함께 하이브리드 기술을 사용하면 전통적인 방법론만으로는 다루기 힘든 다합체 단백질 복합체의 구조적 설명이 가능하다.

[0005] XL-MS(Cross-Linking Mass Spectrometry) 분석 기법은 질량 값의 차이를 통해 표적 단백질(Target Protein)과 물리적으로 상호작용하는 단백질을 실험적으로 확인하고, 어느 부위에서 실질적으로 상호작용이 일어나는지 확인하는 분석 기법이다. XL-MS는 기본 환경에서 단백질-단백질 상호 작용(PPI)을 동시에 캡처하고 물리적 상호작

용 접촉을 밝혀내어 세포에서 단백질-단백질 상호작용의 정체성과 연결성을 모두 결정할 수 있는 기능으로 인해 MS 기반 기술 중에서 고유하다.

[0006] 인체 내에서 기능적인 역할을 담당하는 단백질은 특정 자극(예로, 돌연변이 또는 바이러스 침입)으로 인하여 단백질의 구조적, 형태적 변화와 함께 상호작용 단백질 변화가 발생하는데 이러한 상호작용체를 검증하는데 사용되는 XL-MS 방법을 타겟 질량분석에 적용하기 위해서는 현재 질량 값이 요구된다. 그러나, 단백질은 구조적, 형태적 변화와 상호작용 단백질에 의한 변화 등에 의해 질량이 변할 수 있어 단백질의 현재 질량 값을 산출하는데 있어서 많은 어려움이 있다.

[0007] 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(2022R1F1A1074235). This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (2022R1F1A1074235).

[0008] 이 성과는 정부(과학기술정보통신부)의 재원으로 과학기술일자리진흥원의 지원을 받아 수행된 연구임(2022OPENLAB(RND)_02). The work was support the Commercializations Promotion Agency for R&D Outcoms grant funded by the Korea (2022OPENLAB(RND)_02).

선행기술문헌

특허문헌

[0010] (특허문헌 0001) 한국등록특허 제10-0698466호

발명의 내용

해결하려는 과제

[0011] 본 발명은 XL-MS 방법을 질량분석에 적용하기 위해 단백질의 구조적, 형태적 변화와 상호작용 단백질의 변화 등에 의해 발생할 수 있는 이론적인 질량 값의 경우의 수를 모두 자동으로 산출할 수 있는 타겟 질량분석법 기반의 단백질 상호작용체 분석 시스템 및 방법을 제공하는 것을 일 목적으로 한다.

과제의 해결 수단

[0013] 상기 목적을 달성하기 위하여 본 발명은, 단백질 상호작용에 연관된 복수개의 데이터를 입력받는 프론트 엔드 단; 및 상기 프론트 엔드 단에서 입력받은 데이터를 이용하여 가능한 이론적 질량 값을 모두 산출하는 백 엔드 단;을 포함하고, 상기 프론트 엔드 단은, 상기 백 엔드 단에서 산출된 이론적 질량 값을 출력하는 것을 일 특징으로 한다.

[0014] 바람직하게는, 상기 프론트 엔드 단은, 단백질 상호작용에 연관된 복수개의 데이터로 표적 단백질에 관한 데이터, 상기 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터, 효소에 대한 데이터, 크로스링커에 대한 데이터, 변이에 대한 데이터, 펩타이드 길이에 대한 데이터, 또는 전하에 대한 데이터를 입력받을 수 있다.

[0015] 바람직하게는, 상기 백 엔드 단은, 표적 단백질에 관한 데이터와 상기 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터로 상기 표적 단백질과 상호작용하는 특정 단백질의 확률에 따라 선별하는 선별부;를 포함할 수 있다.

[0016] 바람직하게는, 상기 백 엔드 단은, 표적 단백질에 대한 데이터, 효소에 대한 데이터, 펩타이드 길이에 대한 데이터, 상기 선별부에서 선별된 상호작용하는 특정 단백질을 확률, 크로스링커에 대한 데이터, 및 변이에 대한 데이터를 이용하여, 표적 단백질과 상호작용 가능한 특정 단백질의 아미노산 서열을 효소 처리하였을 때 펩타이드 단편으로 분해되는 경우의 수를 생성하는 단백질 분해부;를 더 포함할 수 있다.

[0017] 바람직하게는, 상기 백 엔드 단은, 상기 단백질 분해부가 생성한 단편화된 펩타이드의 길이, 전하에 대한 데이터, 크로스링커에 대한 데이터, 변이에 대한 데이터를 이용하여, 단백질-단백질 상호작용에 가능한 모든 경우의

수를 생성하는 상호작용 생성부;를 더 포함할 수 있다.

[0018] 바람직하게는, 상기 백 엔드 단은, 상기 상호작용 생성부가 생성한 단백질-단백질 상호작용에 가능한 모든 경우의 수에 따른 표적 단백질의 질량값을 산출하는 산출부;를 더 포함할 수 있다.

[0019] 또한 본 발명은, 표적 단백질에 관한 데이터, 상기 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터, 효소에 대한 데이터, 크로스링커에 대한 데이터, 변이에 대한 데이터, 펩타이드 길이에 대한 데이터, 또는 전하에 대한 데이터를 입력받는 입력단계; 및 상기 입력단계에서 입력받은 데이터를 이용하여 가능한 이론적 질량 값을 모두 산출하는 연산단계; 및 상기 연산단계에서 산출된 이론적 질량 값을 출력하는 출력단계;를 포함하는 것을 다른 특징으로 한다.

발명의 효과

[0021] 본 발명은 XL-MS 방법을 질량분석에 적용하기 위해 단백질의 구조적, 형태적 변화와 상호작용 단백질의 변화 등에 의해 발생할 수 있는 이론적인 질량 값의 경우의 수를 모두 자동으로 산출할 수 있다는 이점이 있다.

[0022] 본 발명은 이론적인 질량 값을 자동으로 산출함으로써 연구 효율을 높일 수 있다는 이점이 있다.

도면의 간단한 설명

[0024] 도 1은 본 발명의 실시예에 따른 단백질 상호작용체 분석 시스템의 구성도를 나타낸다.

도 2는 본 발명의 실시예에 따른 단백질 상호작용체 분석 시스템의 흐름도를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0025] 이하, 첨부된 도면들에 기재된 내용들을 참조하여 본 발명을 상세히 설명한다. 다만, 본 발명이 예시적 실시 예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일 참조부호는 실질적으로 동일한 기능을 수행하는 부재를 나타낸다.

[0026] 본 발명의 목적 및 효과는 하기의 설명에 의해서 자연스럽게 이해되거나 보다 분명해 질 수 있으며, 하기의 기재만으로 본 발명의 목적 및 효과가 제한되는 것은 아니다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이, 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다.

[0027] 도 1은 본 발명의 실시예에 따른 단백질 상호작용체 분석 시스템(10)의 구성도를 나타낸다. 도 1을 참조하면, 단백질 상호작용체 분석 시스템(10)은 프론드 엔드 단(100)과 백 엔드 단(300)을 포함할 수 있다.

[0028] XL-MS을 분석 시 타겟 질량 분석법에 적용하기 위해서 필요한 질량 값을 산출해는데 많은 어려움이 있다. 단백질 상호작용체 분석 시스템(10)은 XL-MS 방법을 질량분석에 적용하기 위해 단백질의 구조적, 형태적 변화와 상호작용 단백질의 변화 등에 의해 발생할 수 있는 이론적인 질량 값의 경우의 수를 모두 자동으로 산출할 수 있다. 단백질 상호작용체 분석 시스템(10)은 XL-MS 분석 시 필요한 표적 단백질 및 표적 단백질과 상호작용하는 특정 단백질의 상호작용체를 예측해 주며, 실험적으로 상호작용을 검증하는 데 사용되는 크로스링커와 상호작용체의 펩타이드 및 단편화된 펩타이드의 전하, 변이 등을 고려하여 이론적으로 상호작용 가능한 모든 경우의 수에 대한 질량 값을 자동으로 산출할 수 있다. 단백질 상호작용체 분석 시스템(10)은 이론적인 질량 값을 자동으로 산출함으로써 연구 효율을 높일 수 있다.

[0029] 도 2는 본 발명의 실시예에 따른 단백질 상호작용체 분석 시스템(10)의 흐름도를 나타낸다. 도 2를 참조하면, 단백질 상호작용체 분석 시스템(10)은 프론트 엔드 단(100)을 통해 사용자로부터 입력 데이터를 입력받을 수 있고, 사용자에게 질량 값 산출 결과를 출력할 수 있다. 프론트 엔드 단(100)은 사용자가 화면에 보이는 웹 화면에 나타나는 기능을 의미할 수 있다.

[0030] 단백질 상호작용체 분석 시스템(10)은 백 엔드 단(300)에서 프론트 엔드 단(100)을 통해 입력받은 데이터를 조회하여 선택적으로 데이터를 불러오고, 불러온 데이터를 이용하여 이론적으로 상호작용 가능한 모든 경우의 수에 대한 질량 값을 산출할 수 있다. 백 엔드 단(300)은 사용자가 볼 수 없는 영역, 즉 프론트 엔드 단(100)에서

전달된 데이터를 처리하고 관리하는 기능을 의미할 수 있다.

[0031] 프론드 엔드 단(100)은 단백질 상호작용에 연관된 복수개의 데이터를 입력받을 수 있다. 단백질 상호작용 연관된 데이터는 특정 단백질 상호작용체에 관련된 단백질, 단백질-단백질 상호작용, 효소, 크로스링커, 변이, 전하 등에 대한 것이 있다.

[0032] 프론드 엔드 단(100)은 단백질 상호작용에 연관된 복수개의 데이터로 표적 단백질에 관한 데이터, 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터, 효소(Enzyme)에 대한 데이터, 크로스링커(Cross-Linker)에 대한 데이터, 변이(Modification)에 대한 데이터, 펩타이드(Peptide) 길이에 대한 데이터, 또는 전하(Charge)에 대한 데이터를 입력받을 수 있다.

[0033] 프론트 엔드 단(100)은 표적 단백질에 관한 데이터를 Uniprot(https://www.uniprot.org/) Public DB에서 불러올 수 있다. 여기에서, 표적 단백질이란 질량 분석법을 위해 이론적 질량을 구할 대상 단백질을 의미한다.

[0034] 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터는 STRING(https://string-db.org/) Public DB에서 불러올 수 있다. 표적 단백질과 상호작용하는 특정 단백질은 상호작용을 통해 특정 단백질의 변화를 발생시킬 수 있는 단백질을 의미한다. 표적 단백질은 특정 자극으로 인해 단백질의 구조적, 형태적 변화와 함께 상호작용 단백질에 의해 변화가 발생할 수 있으므로 이를 모두 고려해야 한다.

[0035] 프론트 엔드 단(100)은 기존에 알려진 효소(Enzyme)에 대한 데이터, 크로스링커(Cross-Linker)에 대한 데이터, 변이(Modification)에 대한 데이터, 펩타이드(Peptide) 길이에 대한 데이터, 또는 전하(Charge)에 대한 데이터에 관한 whiter paper가 취합된 기구축된 데이터베이스에서 데이터를 불러올 수 있다.

[0036] 프론드 엔드 단(100)은 백 엔드 단(300)에서 산출된 이론적 질량 값을 출력할 수 있다. 프론드 엔드 단(100)은 가능한 이론적 질량 값을 출력하므로 복수개의 질량 값을 출력할 수 있다. 프론트 엔드 단(100)은 표적 단백질과 특정 단백질 사이에 얼마나 상호작용하는지 확률을 입력받을 수 있다. 즉, 프론트 엔드 단(100)으로 확률이 입력되면, 표적 단백질과 상호작용할 확률이 입력된 확률 이상인 특정 단백질만이 선별될 수 있다. 프론트 엔드 단(100)은 입력된 확률에 맞는 구체적인 결과를 출력할 수 있다. 프론트 엔드 단(100)으로 입력된 확률은 선별부(310)에 전달될 수 있다.

[0037] 백 엔드 단(300)은 선별부(310), 단백질 분해부(330), 상호작용 생성부(350), 및 산출부(370)를 포함할 수 있다.

[0038] 백 엔드 단(300)은 프론트 엔드 단(100)에서 입력받은 데이터를 이용하여 가능한 이론적 질량 값을 모두 산출할 수 있다. 백 엔드 단(300)은 단백질 상호작용에 연관된 복수개의 데이터를 이용하여 표적 단백질의 변화 가능한 이론적 구조, 형태, 상호작용을 모두 고려하여 이론적인 질량 값을 모두 산출할 수 있다. 백 엔드 단(300)은 발생할 수 있는 이론적인 질량 값의 경우의 수를 모두 자동으로 산출할 수 있다.

[0039] 선별부(310)는 표적 단백질에 관한 데이터와 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터로 표적 단백질과 상호작용하는 특정 단백질을 확률에 따라 선별할 수 있다. 표적 단백질은 상호작용하는 특정 단백질에 의해 변화가 발생할 수 있다. 표적 단백질은 근처에 상호작용하는 특정 단백질이 다수 존재하는 경우 어떠한 특정 단백질과 상호작용하는지는 확률적으로 통계 되어 있다.

[0040] 선별부(310)는 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터로 표적 단백질과 상호작용하는 특정 단백질의 랭킹(Ranking)을 선별할 수 있다. 이후, 표적 단백질과 상호작용하는 특정 단백질의 랭킹을 이용하여 이론적인 질량 값의 경우의 수를 확률에 따라 나타낼 수 있다.

[0041] 선별부(310)는 선별과정에서 단백질-단백질 간 상호작용할 확률에 따라 highest confidence (>0.9), high confidence (>0.7), medium confidence(>0.4), low confidence (>0.15)로 특정하여 선별할 수 있다. 선별부(310)는 프론트 엔드 단(100)에 입력된 확률을 전달받아, 표적 단백질과 상호작용할 확률이 전달된 확률 이상인 특정 단백질만을 선별할 수 있다.

[0042] 단백질 분해부(330)는 표적 단백질에 대한 데이터, 효소에 대한 데이터, 펩타이드 길이에 대한 데이터, 선별부(310)에서 선별된 상호작용하는 특정 단백질의 확률, 크로스링커에 대한 데이터, 및 변이에 대한 데이터를 이용하여, 표적 단백질과 상호작용 가능한 특정 단백질의 아미노산 서열을 효소 처리하였을 때 펩타이드 단편으로 분해되는 경우의 수를 생성할 수 있다.

[0043] 단백질 분해부(330)는 표적 단백질이 변화되는 경우의 수를 생성하기 위해 표적 단백질이 펩타이드 단편으로 분

해될 수 있는 가능성을 모두 고려할 수 있다. 단백질 분해부(330)는 생성된 펩타이드 단편으로 분해되는 경우의 수를 상호작용 생성부(350)로 전달할 수 있다.

[0044] 상호작용 생성부(350)는 단백질 분해부(330)가 생성한 단편화된 펩타이드의 길이, 전하에 대한 데이터, 크로스링커에 대한 데이터, 변이에 대한 데이터를 이용하여, 단백질-단백질 상호작용에 가능한 모든 경우의 수를 생성할 수 있다.

[0045] 상호작용 생성부(350)는 단편화된 펩타이드가 재결합하거나 특정 단백질과 결합하는 등 단백질-단백질 상호작용할 수 있는 모든 경우의 수를 생성할 수 있다. 상호작용 생성부(350)는 단백질-단백질 상호작용을 통해 변화된 표적 단백질의 경우의 수를 모두 생성할 수 있다. 상호작용 생성부(350)는 단백질-단백질 상호작용을 통해 변화된 표적 단백질의 경우의 수를 생성하여 산출부(370)로 전달할 수 있다.

[0046] 산출부(370)는 상호작용 생성부(350)가 생성한 단백질-단백질 상호작용에 가능한 모든 경우의 수에 따른 표적 단백질의 질량값을 산출할 수 있다.

[0047] 본 발명의 다른 실시예인 단백질 상호작용체 분석 방법은 입력단계, 연산단계, 및 출력단계를 포함할 수 있다.

[0048] 입력단계는 표적 단백질에 관한 데이터, 상기 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터, 효소에 대한 데이터, 크로스링커에 대한 데이터, 변이에 대한 데이터, 펩타이드 길이에 대한 데이터, 또는 전하에 대한 데이터를 입력받을 수 있다. 입력단계는 전술한 프론트 엔드 단(100)에서 수행되는 동작을 의미한다.

[0049] 연산단계는 입력단계에서 입력받은 데이터를 이용하여 가능한 이론적 질량 값을 모두 산출할 수 있다. 연산단계는 전술한 백 엔드 단(300)에서 수행되는 동작을 의미한다.

[0050] 연산단계는 표적 단백질에 관한 데이터와 상기 표적 단백질과 상호작용하는 특정 단백질에 대한 확률 데이터로 상기 표적 단백질과 상호작용하는 특정 단백질을 확률에 따라 선별하는 선별단계를 포함할 수 있다. 선별단계는 전술한 선별부(310)에서 수행되는 동작을 의미한다.

[0051] 연산단계는 표적 단백질에 대한 데이터, 효소에 대한 데이터, 펩타이드 길이에 대한 데이터, 선별단계에서 선별된 상호작용하는 특정 단백질을 확률, 크로스링커에 대한 데이터, 및 변이에 대한 데이터를 이용하여, 표적 단백질과 상호작용 가능한 특정 단백질의 아미노산 서열을 효소 처리하였을 때 펩타이드 단편으로 분해되는 경우의 수를 생성하는 단백질 분해단계를 포함할 수 있다. 단백질 분해단계는 전술한 단백질 분해부(330)에서 수행되는 동작을 의미한다.

[0052] 연산단계는 단백질 분해단계에서 생성된 단편화된 펩타이드의 길이, 전하에 대한 데이터, 크로스링커에 대한 데이터, 변이에 대한 데이터를 이용하여, 단백질-단백질 상호작용에 가능한 모든 경우의 수를 생성하는 상호작용 생성단계를 포함할 수 있다. 상호작용 생성단계는 전술한 상호작용 생성부(350)에서 수행되는 동작을 의미한다.

[0053] 연산단계는 상호작용 생성단계에서 생성된 단백질-단백질 상호작용에 가능한 모든 경우의 수에 따른 표적 단백질의 질량값을 산출하는 산출단계를 포함할 수 있다. 산출단계는 전술한 산출부(370)에서 수행되는 동작을 의미한다.

[0054] 이상에서 대표적인 실시예를 통하여 본 발명을 상세하게 설명하였으나, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 상술한 실시예에 대하여 본 발명의 범주에서 벗어나지 않는 한도 내에서 다양한 변형이 가능함을 이해할 것이다. 그러므로 본 발명의 권리 범위는 설명한 실시예에 국한되어 정해져서는 안 되며, 후술하는 특허청구범위뿐만 아니라 특허청구범위와 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태에 의하여 정해져야 한다.

부호의 설명

[0056] 10 : 단백질 상호작용체 분석 시스템

100 : 프로트 엔드 단

300 : 백 엔드 단

310 : 선별부

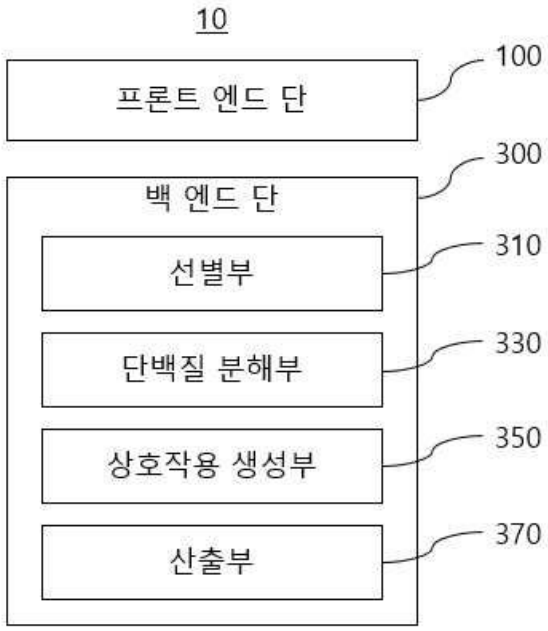
330 : 단백질 분해부

350 : 상호작용 생성부

370 : 산출부

도면

도면1



도면2

