



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년01월10일
(11) 등록번호 10-2622760
(24) 등록일자 2024년01월04일

- (51) 국제특허분류(Int. Cl.)
G16B 15/00 (2019.01) G16B 20/30 (2019.01)
G16B 30/00 (2019.01) G16B 5/00 (2019.01)
G16C 20/20 (2019.01) G16C 20/40 (2019.01)
G16C 20/80 (2019.01)
- (52) CPC특허분류
G16B 15/00 (2019.02)
G16B 20/30 (2019.02)
- (21) 출원번호 10-2023-0060736
(22) 출원일자 2023년05월10일
심사청구일자 2023년05월10일
- (56) 선행기술조사문헌
K-E Choi 외, "TWN-RENCOD: A novel method for protein binding site comparison", Computational and Structural Biotechnology Journal 21, pp.425-431, 2022.12.19.*
J. Lee 외, "Water polygons in high-resolution protein crystal structures", Protein Science, 18권, pp.1370-1376, 2009.
K-E Choi 외, "Topological Water Network Analysis Around Amino Acids", Molecules, 24, 2653, 2019.
*는 심사관에 의하여 인용된 문헌
- (73) 특허권자
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
- (72) 발명자
강남숙
대전광역시 유성구 어은로 57 한빛아파트 101-201
최광은
대전광역시 유성구 대학로 159번길 13(궁동) 시티빌 204호
바루푸리 아난드
대전광역시 유성구 어은로51번길 42, 202호
- (74) 대리인
특허법인공룡

전체 청구항 수 : 총 7 항

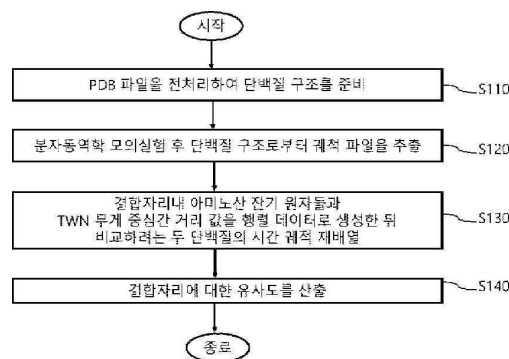
심사관 : 성경아

(54) 발명의 명칭 위상학적 물분자 네트워크 기반의 단백질 결합 자리 유사도 분석 방법

(57) 요약

본 발명은 위상학적 물분자 네트워크 기반의 단백질 결합 자리 유사도 분석 방법 에 대한 것이다. 본 발명에 따른 단백질 결합 자리 유사도 분석 방법은 RCSB 웹사이트로부터 단백질 정보 파일(Protein Data Bank, PDB) 파일을 수집하고, 수집된 PDB 파일을 전처리하여 단백질 구조를 준비하는 단계, 준비된 단백질 구조로부터 분자동역학 모의실험 후 궤적파일을 추출하고, 추출된 궤적 파일을 이용하여 결합 자리내 물 분자를 추출하는 단계, 상기 추출된 물분자를 이용하여 TWN 분석을 수행하여, 결합자리에 아미노산 잔기 원자들과 TWN 무게 중심간 거리 값을 행렬 데이터로 생성한 뒤 비교하려는 두 단백질의 시간 궤적을 재배열하는 단계, 그리고 상기 TWN 무게 중심 간 거리값을 이용하여 결합자리에 대한 유사도를 산출하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G16B 30/00 (2019.02)

G16B 5/00 (2019.02)

G16C 20/20 (2019.02)

G16C 20/40 (2019.02)

G16C 20/80 (2019.02)

공지예외적용 : 있음

명세서

청구범위

청구항 1

RCSB 웹사이트로부터 단백질 정보 파일(Protein Data Bank, PDB) 파일을 수집하고, 수집된 PDB 파일을 전처리하여 단백질 구조를 준비하는 단계,

준비된 단백질 구조로부터 분자동역학 모의실험 후 궤적파일을 추출하고, 추출된 궤적 파일을 이용하여 결합 자리내 물 분자를 추출하는 단계,

상기 추출된 물분자를 이용하여 TWN 분석을 수행하여, 결합자리에 아미노산 잔기 원자들과 TWN 무게 중심간 거리 값을 행렬 데이터로 생성한 다음, 비교하려는 두 단백질의 시간 궤적을 재배열하는 단계, 그리고

상기 TWN 무게 중심간 거리값을 이용하여 결합자리에 대한 유사도를 산출하는 단계를 포함하는 단백질 결합 자리 유사도 분석 방법.

청구항 2

제1항에 있어서,

상기 시간 궤적을 재배열하는 단계는,

각각의 시간 궤적 파일에서 하기의 수학적식을 이용하여 물 분자들 사이의 상호작용 포텐셜 에너지를 산출하는 단백질 결합 자리 유사도 분석 방법:

$$v(a, b) = \sum_i^{on} \sum_j^{on} \frac{q_i q_j e^2}{r_{ij}} + \frac{A}{r_{oo}^{12}} - \frac{C}{r_{oo}^6}$$

여기서, $v(a, b)$ 는 물 분자 a, b 사이의 상호작용 에너지를 나타내고, r_{oo} 는 산소 원자 사이의 거리를 나타내고, q_i 는 i 위치에서의 부분 전하를 나타내고, q_j 는 j 위치에서 부분 전하를 나타내고, r_{ij} 는 q_i 와 q_j 사이의 거리를 나타내고, A 는 i 와 j 사이의 척력을 나타내며, C 는 i 와 j 사이의 인력을 나타낸다.

청구항 3

제2항에 있어서,

상기 시간 궤적을 재배열하는 단계는,

상기 포텐셜 에너지를 이용하여 TWN를 추출하고, 추출된 TWN을 구성하는 물 분자들의 산소와 수소를 포함하는 좌표 x, y, z 축 값들을 각 원자 개수로 나눈 평균 값으로 TWN 무게중심 좌표를 구하는 단백질 결합 자리 유사도 분석 방법.

청구항 4

제3항에 있어서,

상기 시간 궤적을 재배열하는 단계는,

단백질 결합 자리 내부를 공결정화된 리간드의 좌표로부터 5 내지 6 Å 반경으로 지정해서 탐색하여 TWN 무게중심 좌표에서 결합 자리 내부에 위치한 각 아미노산 잔기의 골격 원자들 사이의 3차원 유클리드 거리를 산출하고,

모든 시간 궤적 파일에서 계산된 거리 값들을 기반으로 행렬 데이터를 생성하는 단백질 결합 자리 유사도 분석

방법.

청구항 5

제4항에 있어서,

상기 행렬데이터는,

TWN 시간 궤적을 행으로 나타내고, 결합 자리내 아미노산 잔기 번호를 열로 나타내는 단백질 결합 자리 유사도 분석 방법.

청구항 6

제4항에 있어서,

상기 시간 궤적을 재배열하는 단계는,

비교하려는 두 단백질을 기준으로 생성된 행렬 데이터에서 첫 번째 단백질의 행렬 데이터에 있는 시간 궤적을 기준으로 두 번째 단백질의 모든 시간 궤적에 있는 TWN 무게중심 좌표와 아미노산 잔기 골격 원자들 사이의 거리 값의 차를 산출한 다음,

최소값을 나타내는 두 번째 단백질의 시간 궤적을 선정하여 첫번째 단백질의 시간 궤적과 행렬 데이터 값이 최소인 두번째 단백질의 시간 궤적을 재배열하는 단백질 결합 자리 유사도 분석 방법.

청구항 7

제1항에 있어서,

상기 결합자리에 대한 유사도를 산출하는 단계는,

TWN과 아미노산 잔기 골격 원자들 사이의 거리 값을 이용하여 복수의 클래스로 분류하고, 분류된 클래스를 이용하여 결합 자리 유사도 값을 계산하는 단백질 결합 자리 유사도 분석 방법.

발명의 설명

기술 분야

[0001] 본 발명은 위상학적 물분자 네트워크 기반의 단백질 결합 자리 유사도 분석 방법에 관한 것으로서, 더욱 상세하게는 위상학적 물분자 네트워크(Topological Water Network, TWN)를 기반으로 산출된 단백질 결합 자리 내 유사도(binding site similarity)를 분석하는 단백질 결합 자리 유사도 분석 방법에 관한 것이다.

배경 기술

[0002] 단백질의 3차원 구조는 생물학적 기능을 이해하는 데 가장 중요하며, 단백질 구조의 분석, 특히 결합 자리의 연구는 신약 설계에서 핵심적인 역할을 한다.

[0003] 결합 자리는 단백질 기능과 밀접한 관련이 있으며 결합 자리의 특성을 이해하는 것은 구조 기반 신약 설계에서 중요한 과정이다. 많은 단백질들이 유사한 결합 자리를 가지고 있으며, 이러한 결합 자리 유사도 계산은 알려진 약물들의 신약 재창출(Repurposing), 부작용 (Off-target) 가능성 식별, 다중약물학(Polypharmacology) 연구에 유용한 정보를 제공한다.

[0004] 현재까지 단백질 구조에서 리간드 결합 자리를 비교하기 위해서 여러 가지 계산 방법이 개발되었다. 단백질 결합자리 유사도 계산 계산방법은 아미노산 잔기 기반(residue-based), 형태 기반(shape-based) 및 상호작용 기반(interaction-based)으로 분류할 수 있다.

[0005] 아미노산 잔기 기반은 구조 표현자를 이용하고, 형태 기반은 그래프를 이용하며, 상호 작용 기반은 상호 작용 필드를 이용한다. 결합 자리 비교 방법 프로그램들은 계속해서 개발되고 있지만 특정 연구 분야에 가장 적합한 방법이 무엇인지 결정하는 것이 점점 더 어려워지고 있다.

[0006] 예를 들어 단백질 구조 및 서열은 유사한 반면, 리간드 (혹은 약물)의 활성 값에서 차이를 발생시키는 문제점이 있었다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 대한민국 등록특허공보 제10-0546779호(2006.01.25. 공고)

발명의 내용

해결하려는 과제

[0008] 본 발명이 이루고자 하는 기술적 과제는 위상학적 물분자 네트워크(Topological Water Network, TWN)를 기반으로 산출된 단백질 결합 자리 내 유사도(binding site similarity)를 분석하는 단백질 결합 자리 유사도 분석 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0009] 이러한 기술적 과제를 이루기 위한 본 발명의 실시예에 따른 단백질 결합 자리 유사도 분석 방법은 RCSB 웹사이트로부터 단백질 정보 파일(Protein Data Bank, PDB) 파일을 수집하고, 수집된 PDB 파일을 전처리하여 단백질 구조를 준비하는 단계, 준비된 단백질 구조로부터 분자동역학 모의실험 후 궤적파일을 추출하고, 추출된 궤적파일을 이용하여 결합 자리내 물 분자를 추출하는 단계, 상기 추출된 물분자를 이용하여 TWN 분석을 수행하여, 결합자리에 아미노산 잔기 원자들과 TWN 무게 중심간 거리 값을 행렬 데이터로 생성한 다음, 비교하려는 두 단백질의 시간 궤적을 재배열하는 단계, 그리고 상기 TWN 무게 중심간 거리값을 이용하여 결합자리에 대한 유사도를 산출하는 단계를 포함한다.

[0010] 상기 시간 궤적을 재배열하는 단계는, 각각의 시간 궤적 파일에서 하기의 수학식을 이용하여 물 분자들 사이의 상호작용 포텐셜 에너지를 산출할 수 있다.

$$v(a, b) = \sum_i^{on} \sum_j^{on} \frac{q_i q_j e^2}{r_{ij}} + \frac{A}{r_{oo}^{12}} - \frac{C}{r_{oo}^6}$$

[0011]

[0012] 여기서, $v(a, b)$ 는 물 분자 a, b 사이의 상호작용 에너지를 나타내고, r_{oo} 는 산소 원자 사이의 거리를 나타내고, q_i 는 i 위치에서의 부분 전하를 나타내고, q_j 는 j 위치에서 부분 전하를 나타내고, r_{ij} 는 q_i 와 q_j 사이의 거리를 나타내고, A 는 i 와 j 사이의 척력을 나타내며, C 는 i 와 j 사이의 인력을 나타낸다.

[0013] 상기 시간 궤적을 재배열하는 단계는, 상기 포텐셜 에너지를 이용하여 TWN을 추출하고, 추출된 TWN을 구성하는 물 분자들의 산소와 수소를 포함하는 좌표 x, y, z 축 값들을 각 원자 개수로 나눈 평균 값으로 TWN 무게중심 좌표를 구할 수 있다.

[0014] 상기 시간 궤적을 재배열하는 단계는, 단백질 결합 자리 내부를 공결정화된 리간드의 좌표로부터 5 내지 6 Å 반경으로 지정해서 탐색하여 TWN 무게중심 좌표에서 결합 자리 내부에 위치한 각 아미노산 잔기의 골격 원자들 사이의 3차원 유클리드 거리를 산출하고, 모든 시간 궤적 파일에서 계산된 거리 값들을 기반으로 행렬 데이터를 생성할 수 있다.

[0015] 상기 행렬데이터는, TWN 시간 궤적을 행으로 나타내고, 결합 자리내 아미노산 잔기 번호를 열로 나타낼 수 있다.

[0016] 상기 시간 궤적을 재배열하는 단계는, 비교하려는 두 단백질을 기준으로 생성된 행렬 데이터에서 첫 번째 단백질의 행렬 데이터에 있는 시간 궤적을 기준으로 두 번째 단백질의 모든 시간 궤적에 있는 TWN 무게중심 좌표와 아미노산 잔기 골격 원자들 사이의 거리 값의 차를 산출한 다음, 최소값을 나타내는 두 번째 단백질의 시간 궤적을 선정하여 첫번째 단백질의 시간 궤적과 행렬 데이터 값이 최소인 두번째 단백질의 시간 궤적을 재배열할 수 있다.

[0017] 상기 결합자리에 대한 유사도를 산출하는 단계는, TWN과 아미노산 잔기 골격 원자들 사이의 거리 값을 이용하여 복수의 클래스로 분류하고, 분류된 클래스를 이용하여 결합 자리 유사도 값을 계산할 수 있다.

도면의 간단한 설명

[0018] 도 1은 본 발명의 실시예에 따른 위상학적 물분자 네트워크 기반의 단백질 결합 자리 유사도 분석 방법을 설명하기 위한 순서도이다.

도 2는 도 1에 도시된 단백질 결합 자리 유사도 분석 방법을 설명하기 위한 예시도이다.

도 3은 도 1에 도시된 S130단계에서 분자동역학 시간계적들의 재배열 과정을 설명하기 위한 도면이다.

도 4는 본 발명의 실시예에 따른 단백질 결합 분석 방법을 통해 활성값 차이에 따른 경향성을 나타내는 히트맵이다.

도 5는 도4에 도시된 결과를 표로 나타낸 것이다.

도 6은 도 4 및 도 5에 도시된 데이터의 상관관계를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0019] 이하 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하기로 한다. 이 과정에서 도면에 도시된 선들의 두께나 구성요소의 크기 등은 설명의 명료성과 편의상 과장되게 도시되어 있을 수 있다.

[0020] 또한 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서, 이는 측정 대상자, 운용자의 의도 또는 관례에 따라 달라질 수 있다. 그러므로 이러한 용어들에 대한 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0021] 이하에서는 도 1을 이용하여 본 발명의 실시예에 따른 네트워크 기반의 단백질 결합 분석 방법에 대해 더욱 상세하게 설명한다.

[0022] 도 1은 본 발명의 실시예에 따른 위상학적 물분자 네트워크 기반의 단백질 결합 자리 유사도 분석 방법을 설명하기 위한 순서도이다.

[0023] 도 1에 도시된 바와 같이, 단백질 결합 자리 유사도 분석 장치는 RCSB 웹사이트로부터 단백질 정보 파일(Protein Data Bank, PDB) 파일을 수집하고, 수집된 PDB 파일을 전처리하여 단백질 구조를 준비한다(S110).

[0024] PDB 파일은 수니티넵이 결합된 X선 결정 구조들에 관한 것으로서, KIT (PDB: 3GOE), PDGFRA (PDB: 6JOK), VEGFR2 (PDB: 4AGD), PHKG2 (PDB: 2Y7J), ITK (PDB: 3MIY), HPK1 (PDB: 6NFY), MST3 (PDB: 4QMZ), PAK6 (PDB: 4KS8), CDK2 (PDB: 3TI1)을 포함한다.

[0025] 그 다음, 단백질 결합 자리 유사도 분석 장치는 수집된 PDB 파일로부터 리간드를 제거하는 전처리 과정을 수행하여 단백질 구조를 준비한다. 이때, 단백질 결합 자리 유사도 분석 장치는 아미노산 잔기들과 수소 원자들을 추가하되, 원자결합 방향 및 형식 전하들을 부여하여 pH 7.4 기준으로 이온화 계산을 수행한다.

[0026] S110단계가 완료되면, 단백질 결합 자리 유사도 분석 장치는 준비된 단백질 구조로부터 분자동역학 모의실험 후 궤적파일을 추출하고, 추출된 궤적 파일을 이용하여 결합 자리내 물 분자를 추출한다(S120).

[0027] 단백질 결합 자리 유사도 분석 장치는 GROMACS를 실행한 상태에서 준비된 단백질 구조를 입력데이터로 사용하여 분자동역학 모의 실험을 수행하여 궤적파일을 추출한다.

[0028] 부연하자면, 단백질 결합 자리 유사도 분석 장치는 CHARMM27 역장을 부여하고, TIP3P 모델을 이용하여 용매화를 수행한다.

[0029] 그리고, 단백질 결합 자리 유사도 분석 장치는 에너지 최소화 및 평형화를 수행한다. 에너지 최소화는 경사하강(steepest descent) 알고리즘을 이용하여 에너지를 최소화한다. 그리고, 단백질 결합 자리 유사도 분석 장치는 NVT(입자 수, 부피, 온도가 일정) 조건에서 v-rescale thermostat으로 100 피코초(ps) 동안 평형화를 거치고, Parrinello-Rahman barostat 알고리즘으로 200 피코초(ps) NPT(입자 수, 압력, 온도가 일정) 조건에서 평형화를 거친다.

[0030] 다만, 단백질 결합 자리 유사도 분석 장치는 분자동역학 최종 모의실험을 298K 온도에서 10 나노초(ns) 동안 수

행한다. 단백질의 거동은 제한하고, 장범위 정전기적상호작용(Long-range electrostatics)는 PME로 수행하고, 단범위 정전기적 상호작용(Short-range electrostatics)은 12Å 로 설정한다. 또한, 모든 결합 길이들은 LINC 알고리즘으로 설정하여, 분자동역학 궤적들을 1 피코초(ps) 마다 추출한다.

[0031] 그 다음, 단백질 결합 자리 유사도 분석 장치는 추출된 물분자를 이용하여 TWN 분석을 수행하여, 결합자리에 아미노산 잔기 원자들과 TWN 무게 중심간 거리 값을 행렬 데이터로 생성한 다음, 비교하려는 두 단백질의 시간 궤적을 재배열한다(S130)

[0032] 도 2는 도 1에 도시된 단백질 결합 자리 유사도 분석 방법을 설명하기 위한 예시도이고, 도 3은 도 1에 도시된 S130단계에서 분자동역학 시간궤적들의 재배열 과정을 설명하기 위한 도면이다.

[0033] 먼저, TWN는 물 분자들의 고리형 네트워크를 나타낸다. 단백질 결합 자리 유사도 분석 장치는 삼환 (3-ring) TWN을 추출한다. 이때, 포텐셜 함수는 TIP3P 물 분자 모델을 이용하며 레너드존스 힘과 쿨롱힘으로 물 분자들 사이의 상호작용 에너지를 계산하며, 상호작용 포텐셜 에너지는 하기의 수학적 식 1을 통해 산출된다.

수학적 식 1

$$v(a,b) = \sum_i^{on} \sum_j^{on} \frac{q_i q_j e^2}{r_{ij}} + \frac{A}{r_{oo}^{12}} - \frac{C}{r_{oo}^6}$$

[0034]

[0035] 여기서, v(a, b)는 물분자 a,b 사이의 상호작용 에너지를 나타내고, r_{oo} 는 산소 원자 사이의 거리를 나타내고, q_i 는 i 위치에서의 부분전하를 나타내고, q_j 는 j 위치에서 부분전하를 나타낸다. 또한, r_{ij} 는 q_i 와 q_j 사이의 거리를 나타내고, A는 i 와 j 사이의 척력을 나타내며, C는 i 와 j 사이의 인력을 나타낸다.

[0036] 단백질 결합 자리 유사도 분석 장치는 추출된 TWN을 구성하는 물 분자들의 산소와 수소를 포함하는 좌표 x,y,z 축 값들을 각 원자 개수로 나눈 평균 값으로 TWN 무게중심 좌표를 구한다.

[0037] 그 다음, 단백질 결합 자리 유사도 분석 장치는 단백질 결합 자리 내부를 공결정화된 리간드의 좌표로부터 5 내지 6 Å 반경으로 지정해서 탐색하여 TWN 무게중심 좌표에서 결합 자리 내부에 위치한 각 아미노산 잔기의 골격 원자들(N, CA, C, O) 사이의 3차원 유클리드 거리를 산출한다.

[0038] 그리고, 단백질 결합 자리 유사도 분석 장치는 모든 시간 궤적 파일에서 상기 산출된 3차원 유클리드 거리 값들을 기반으로 행렬 데이터를 생성한다.

[0039] 여기서, 행렬 데이터는 TWN 시간 궤적을 행으로 나타내고, 결합 자리내 아미노산 잔기 번호를 열로 나타낸다.

[0040] 그 다음 도 3에 도시된 바와 같이, 단백질 결합 자리 유사도 분석 장치는 단백질의 서로 다른 궤적에 따라 단백질의 결합 자리 내 특정위치에 있는 TWN과 비슷하게 재배열하고, 두 번째 단백질은 첫 번째 단백질의 분자동역학 궤적에 따라 재배열한다. 이때, 재배열과정은 각 MD 궤적의 거리의 최소값에 따라 이루어진다.

[0041] 부연하자면, 단백질 결합 자리 유사도 분석 장치는 비교하려는 두 단백질을 기준으로 생성된 행렬 데이터에서 첫 번째 단백질의 행렬 데이터에 있는 시간 궤적을 기준으로 두 번째 단백질의 모든 시간 궤적에 있는 TWN 무게 중심 좌표와 아미노산 잔기 골격 원자들 사이의 거리 값의 차를 산출한다.

[0042] 또한, 단백질 결합 자리 유사도 분석 장치는 최소값을 나타내는 두 번째 단백질의 시간 궤적을 선정하여 첫번째 단백질의 시간 궤적과 행렬 데이터 값이 최소인 두번째 단백질의 시간 궤적을 재배열한다.

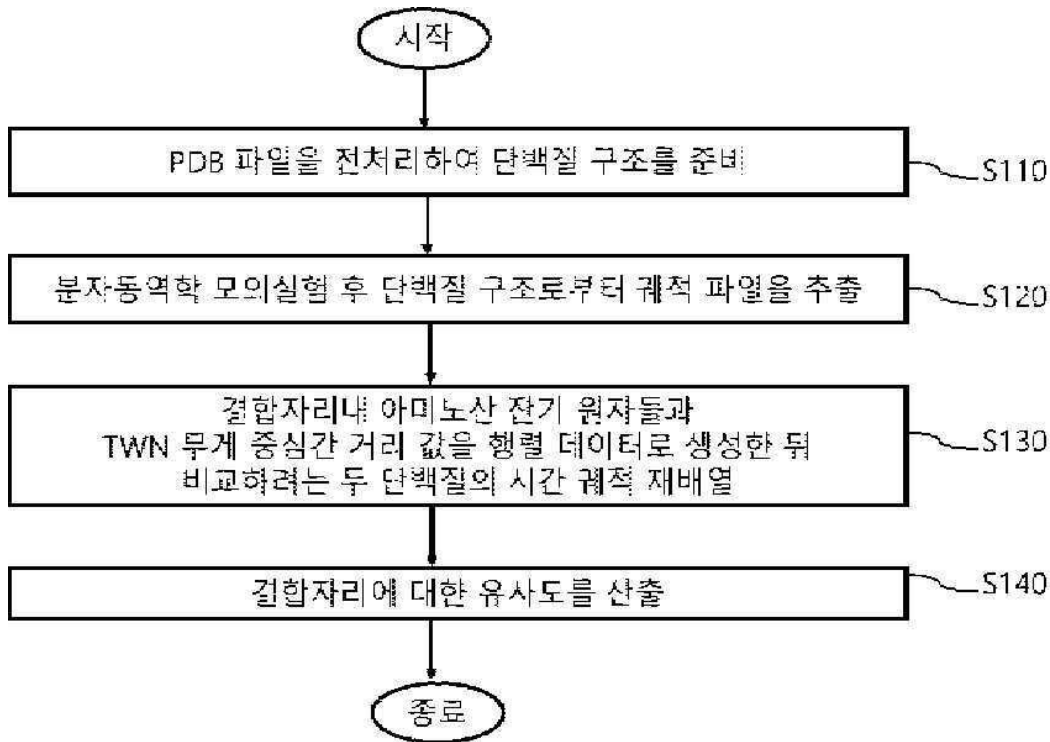
[0043] S130단계가 완료되면, 단백질 결합 자리 유사도 분석 장치는 TWN 무게 중심간 거리값을 이용하여 결합자리에 대한 유사도를 산출한다(S140). 단백질 결합 자리 유사도 분석 장치는 TWN과 아미노산 잔기 골격 원자들 사이의 거리 값을 이용하여 복수의 클래스로 분류한다. 예를 들면, TWN과 아미노산 잔기 골격 원자들 사이의 거리 값이 5 Å보다 작으면, 단백질 결합 자리 유사도 분석 장치는 해당하는 단백질의 행렬 데이터를 클래스 A로 분류하고, TWN과 아미노산 잔기 골격 원자들 사이의 거리 값이 5Å보다 크고 7.5 Å보다 작으면, 단백질 결합 자리 유사도

분석 장치는 해당하는 단백질의 행렬 데이터를 클래스 B로 분류한다. 또한, TWN과 아미노산 잔기 골격 원자들 사이의 거리 값이 $7.5 \text{ \AA}$ 보다 크고 $10 \text{ \AA}$ 보다 작으면, 단백질 결합 자리 유사도 분석 장치는 해당하는 단백질의 행렬 데이터를 클래스 C로 분류한다. 또한, TWN과 아미노산 잔기 골격 원자들 사이의 거리 값이 $10 \text{ \AA}$ 보다 크면, 단백질 결합 자리 유사도 분석 장치는 해당하는 단백질의 행렬 데이터를 클래스 D로 분류한다.

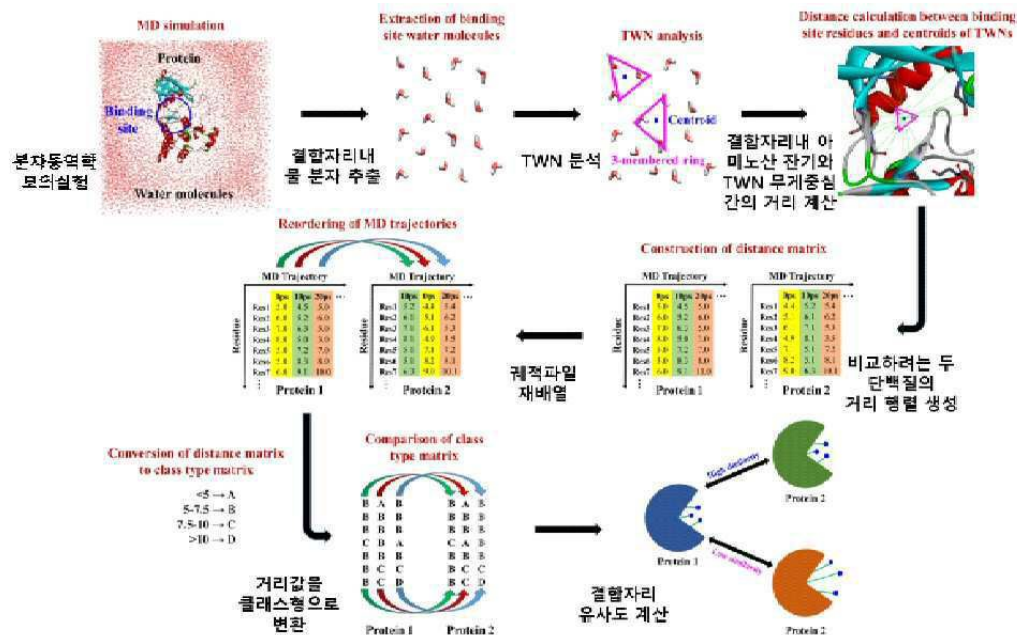
- [0044] 그 다음, 단백질 결합 자리 유사도 분석 장치는 분류된 클래스를 이용하여 결합 자리 유사도 값을 계산한다.
- [0045] 여기서, 유사도는 전체 수의 클래스형에 대한 매칭된 수의 클래스형에 따라 정의될 수 있다.
- [0046] 이하에서는 도 4 내지 6을 이용하여 본 발명의 실시예에 따른 단백질 결합 분석 방법에 대한 실험 결과에 대해 더욱 상세하게 설명한다.
- [0047] 도 4는 본 발명의 실시예에 따른 단백질 결합 분석 방법을 통해 활성값 차이에 따른 경향성을 나타내는 히트맵이고, 도 5는 도4에 도시된 결과를 표로 나타낸 것이다.
- [0048] 도 4에 도시된 바와 같이, 히트맵의 가로축은 두 단백질들의 수니티넵 활성값 차이를 로그값으로 나타낸 것이고, 세로축은 각각의 단백질 결합자리내 유사도 방법을 나타낸다.
- [0049] 세로축에 기재된 TWN-RENCOD는 본 발명의 실시예에 따른 위상학적 물분자 네트워크 기반의 단백질 결합 자리 유사도 분석 방법을 나타내고, IsoMIF는 상호작용 기반의 방법을 나타내고, ProBis는 형태 기반의 방법을 나타내고, PoocketMatch P-max, P-min는 아미노산 잔기 기반의 방법을 나타내고, Sequence similarity는 아미노산 서열 유사성을 나타내고, Sequence Identity는 아미노산 서열 동일성을 나타낸다.
- [0050] 그리고 색깔로 표시된 각각의 칸은 유사도를 나타내며, TWN-RENCOD을 기반으로 하는 방법이 활성 값 차이에 따른 경향성이 가장 크게 나타나는 것을 알 수 있다.
- [0051] 도 6은 도 4 및 도 5에 도시된 데이터의 상관관계를 나타낸다.
- [0052] 도 6에 도시된 바와 같이, 각 인산화효소 단백질들의 수니티넵 활성값 차이와 결합자리내 유사도를 계산한 결과 본 발명의 실시예에 따른 TWN-RENCOD을 기반으로 하는 방법이 가장 높은 값을 나타내는 것을 알 수 있다.
- [0053] 이와 같이 본 발명에 따른 단백질 결합 분석 방법은 자유 에너지 계산을 필요로 하지 않고, 인산화효소 저해제를 설계할 수 있고, 인산화효소들의 선택성을 규명할 수 있어 약물재창출에 적용 가능한 장점이 있다. 또한, 본 발명에 따르면, 결합 자리 유사도 계산 방법을 이용하여 동일한 리간드가 결합되는 여러 동종의 단백질도 분석 가능하다.
- [0054] 본 발명은 도면에 도시된 실시예를 참고로 하여 설명되었으나 이는 예시적인 것에 불과하며, 당해 기술이 속하는 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서 본 발명의 진정한 기술적 보호범위는 아래의 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

도면

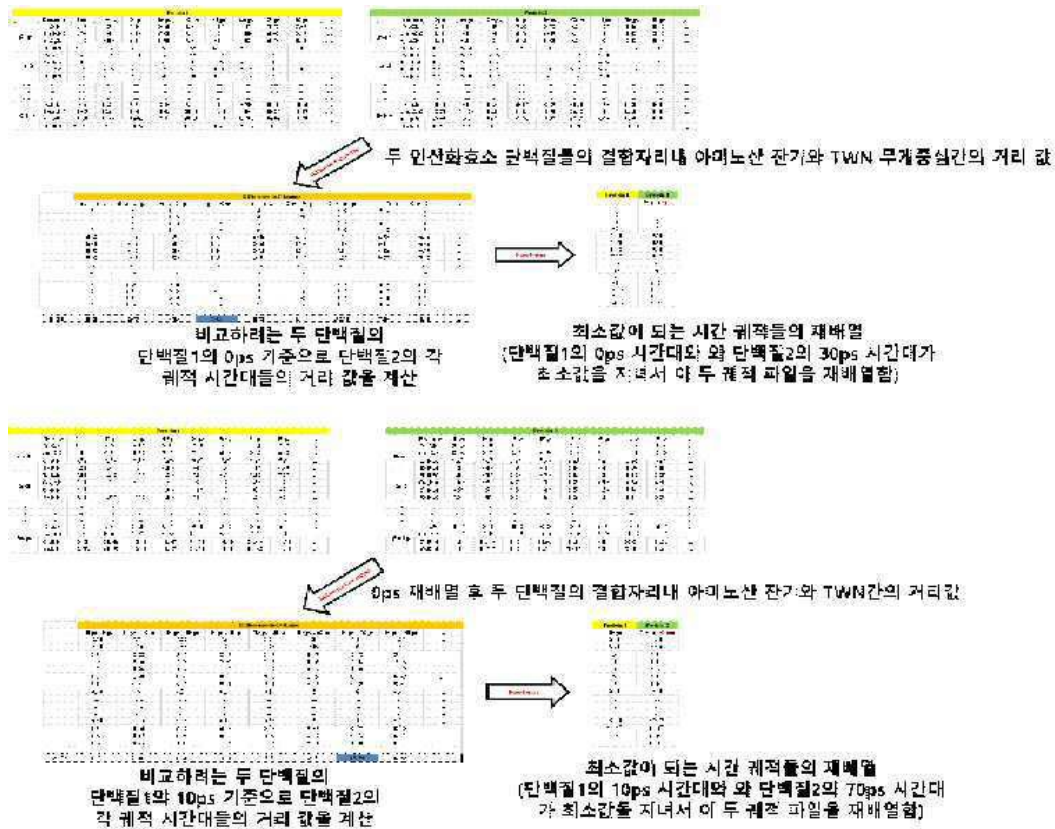
도면1



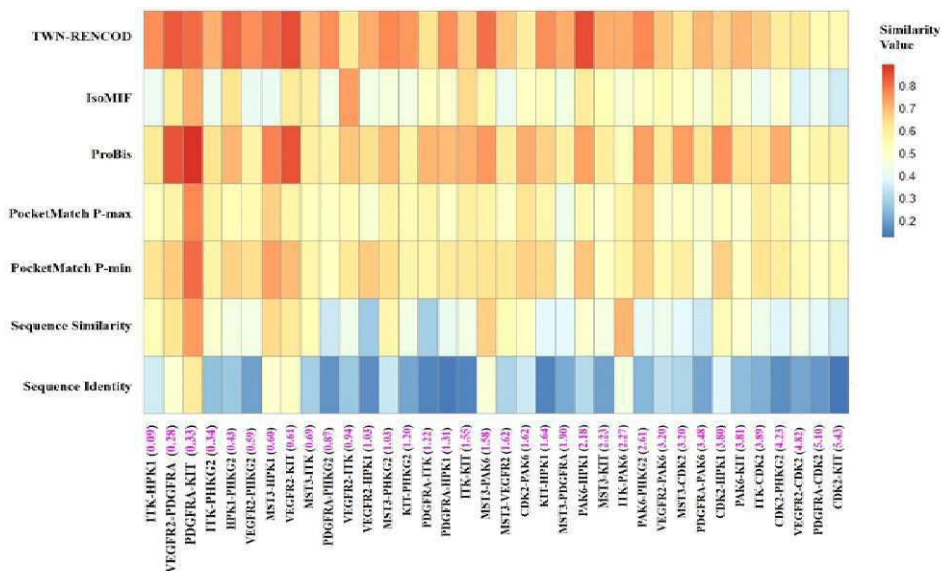
도면2



도면3



도면4



도면5

Protein-Pair	Activity-Difference	Sequence		PocketMatch		ProBis	IsoMIF	TWN-RENCOD
	(log)	Identity	Similarity	P-min	P-max			
ITK-HPK1	0.09	0.35	0.53	0.63	0.51	0.62	0.41	0.76
VEGFR2-PDGFR	0.28	0.47	0.63	0.67	0.58	0.84	0.61	0.83
PDGFR-KIT	0.33	0.61	0.74	0.81	0.77	0.89	0.71	0.81
ITK-PHK2	0.34	0.25	0.48	0.58	0.58	0.63	0.42	0.71
HPK1-PHK2	0.43	0.26	0.44	0.66	0.53	0.70	0.63	0.82
VEGFR2-PHK2	0.59	0.19	0.43	0.64	0.55	0.58	0.40	0.76
MST3-HPK1	0.60	0.48	0.65	0.73	0.66	0.78	0.40	0.80
VEGFR2-KIT	0.61	0.50	0.61	0.69	0.57	0.84	0.60	0.85
MST3-ITK	0.69	0.28	0.53	0.57	0.52	0.59	0.59	0.70
PDGFR-PHK2	0.87	0.17	0.34	0.51	0.51	0.56	0.43	0.76
VEGFR2-ITK	0.94	0.26	0.42	0.62	0.54	0.68	0.74	0.62
VEGFR2-HPK1	1.03	0.16	0.27	0.67	0.47	0.64	0.43	0.72
MST3-PHK2	1.03	0.33	0.56	0.64	0.58	0.69	0.45	0.77
KIT-PHK2	1.20	0.20	0.43	0.56	0.54	0.58	0.44	0.75
PDGFR-ITK	1.22	0.15	0.28	0.56	0.56	0.70	0.50	0.63
PDGFR-HPK1	1.31	0.13	0.40	0.61	0.49	0.69	0.48	0.76
ITK-KIT	1.55	0.15	0.43	0.62	0.59	0.71	0.65	0.65
MST3-PAK6	1.58	0.46	0.66	0.63	0.57	0.75	0.55	0.80
MST3-VEGFR2	1.62	0.29	0.54	0.62	0.49	0.59	0.41	0.68
CDK2-PAK6	1.62	0.34	0.50	0.66	0.63	0.72	0.50	0.60
KIT-HPK1	1.64	0.15	0.38	0.65	0.56	0.66	0.52	0.76
MST3-PDGFR	1.90	0.21	0.38	0.47	0.42	0.58	0.43	0.72
PAK6-HPK1	2.18	0.30	0.48	0.68	0.55	0.73	0.60	0.85
MST3-KIT	2.23	0.19	0.44	0.51	0.48	0.60	0.53	0.72
ITK-PAK6	2.27	0.43	0.70	0.57	0.57	0.52	0.46	0.72
PAK6-PHK2	2.61	0.23	0.39	0.66	0.66	0.74	0.51	0.77
VEGFR2-PAK6	3.20	0.32	0.42	0.55	0.48	0.61	0.54	0.68
MST3-CDK2	3.20	0.29	0.38	0.59	0.51	0.74	0.51	0.63
PDGFR-PAK6	3.48	0.20	0.34	0.47	0.47	0.62	0.46	0.70
CDK2-HPK1	3.60	0.37	0.54	0.66	0.51	0.76	0.56	0.66
PAK6-KIT	3.61	0.24	0.48	0.51	0.48	0.63	0.46	0.70
ITK-CDK2	3.89	0.22	0.42	0.63	0.60	0.63	0.41	0.67
CDK2-PHK2	4.23	0.16	0.37	0.60	0.57	0.72	0.48	0.60
VEGFR2-CDK2	4.82	0.20	0.43	0.55	0.50	0.53	0.37	0.62
PDGFR-CDK2	5.10	0.18	0.38	0.51	0.48	0.57	0.41	0.54
CDK2-KIT	5.43	0.12	0.35	0.55	0.49	0.58	0.35	0.55

도면6

단백질 결합자리 유사도	상관 관계		
	Pearson	Kendall	Spearman
서열 동일성	0.37	0.25	0.34
서열 유사성	0.35	0.29	0.40
<u>PocketMatch</u> (P-min)	0.46	0.32	0.47
<u>PocketMatch</u> (P-max)	0.34	0.24	0.35
<u>Pro-Bis</u>	0.32	0.15	0.26
<u>Iso-MIF</u>	0.36	0.15	0.22
TWN-RENCOD	0.67	0.47	0.62