(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년11월01일
(11) 등록번호 10-2725935
(24) 등록일자 2024년10월30일

(51) 국제특허분류(Int. Cl.)
G16B 5/00 (2019.01) *G16B 15/30* (2019.01)
G16C 10/00 (2019.01) *G16C 20/20* (2019.01)
G16C 20/40 (2019.01) *G16C 20/64* (2019.01)
(52) CPC특허분류
G16B 5/00 (2019.02)
G16B 15/30 (2019.02)
(21) 출원번호 10-2024-0037803
(22) 출원일자 2024년03월19일
심사청구일자 2024년03월19일
(56) 선행기술조사문헌
Myeong Hwi Lee 외, Design of a Novel and Selective IRAK4 Inhibitor Using Topological Water Network Analysis and Molecular Modeling Approaches, Molecules, 2018.11.29., Vol.23, No.3136, pp.1-15.*
(뒷면에 계속)
전체 청구항 수 : 총 5 항

(73) 특허권자
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
(72) 발명자
강남숙
대전광역시 유성구 어은로 57, 101동 201호 (어은동, 한빛아파트)
박경진
대전광역시 유성구 대학로 159 13 백운빌라 307호
(뒷면에 계속)
(74) 대리인
특허법인공룡

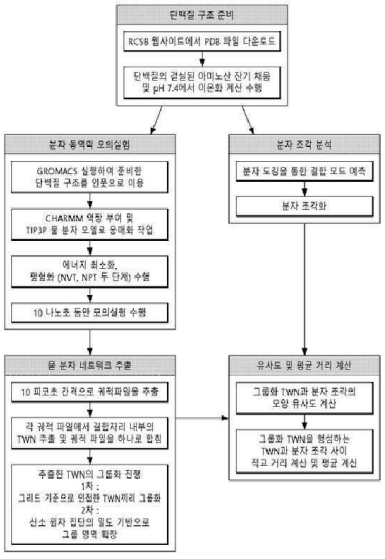
심사관 : 정태수

(54) 발명의 명칭 **위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치 및 방법**

(57) 요 약

본 발명은 위상학적 물분자 네트워크 기반 분자 조각 적합 스크리닝 장치 및 방법에 관한 것으로, 보다 상세하게는 각각의 키나아제에 특이적으로 상호작용하는 화합물의 분자 조각을 효과적으로 도출하는 위상학적 물분자 네트워크 기반 분자 조각 적합 스크리닝 장치 및 방법에 관한 것이다.

대 표 도 - 도2



(52) CPC특허분류
G16C 10/00 (2019.02)
G16C 20/20 (2019.02)
G16C 20/40 (2019.02)
G16C 20/64 (2019.02)

(72) 발명자
윤혜리
세종특별자치시 남세종로 302 새샘마을6단지 603동 1602호
정래진
세종특별자치시 달빛1로 158 범지기마을7단지 707동 402호
손동현
대전광역시 유성구 문화원로37번길 20 202호

(56) 선행기술조사문헌
Kwang-Eun Choi 외, TWN-RENCOD: A novel method for protein binding site comparison, Computational and Structural Biotechnology Journal, 2022.12.19., Vol.21, pp.425-431.
Woo Dae Jang 외, The analysis of water network for kinase selectivity based on the MD simulations, Journal of Molecular Liquids, 2013.12.01., Vol.191, pp.37-41.
Axel Rudling 외, JCIM, 2018.01.08., Vol.58, Issue.2, pp.350-361.
KR102622760 B1
*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업
과제고유번호 1711181708
과제번호 2020R1A2C1006915
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 중견연구
연구과제명 웹기반 TWN 분석을 이용한 생체고분자 핫스팟 특성에 대한 체계적 연구
과제수행기관명 충남대학교
연구기간 2023.03.01 ~ 2024.02.29
공지예외적용 : 있음

명 세 서

청구범위

청구항 1

제어부가 단백질 구조를 준비하는 단계;

단백질 구조의 분자 동역학 모의 실험하는 단계;

단백질 구조의 위상학적 물 분자 네트워크를 추출하는 단계;

위상학적 물 분자 네트워크를 그룹화하는 단계;

단백질과 결합하는 화합물의 분자 조각을 분석하는 단계; 및

그룹화 위상학적 물 분자 네트워크와 분자 조각 사이의 모양 유사도 및 평균 거리를 계산하는 단계;를 수행하고,

상기 단백질 구조의 위상학적 물 분자 네트워크를 추출하는 단계는,

10피코초 간격으로 궤적 파일을 추출하는 단계;

각 궤적 파일에서 결합 자리 내부의 물 분자 네트워크를 추출 및 궤적 파일을 하나로 합치는 단계; 및

추출한 물 분자 네트워크의 그룹화를 진행하고, 1차: 그리드 기준으로 인접 한 물 분자 네트워크끼리 그룹화하고, 2차: 산소 원자 집단의 밀도 기반으로 그룹 영역을 확장하는 단계;를 수행하는 것을 특징으로 하는, 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 방법.

청구항 2

제1항에 있어서,

상기 단백질 구조를 준비하는 단계는,

PDB 파일을 다운로드하는 단계; 및

단백질의 결실된 아미노산 잔기 채움 및 pH 7.4에서 이온화 계산을 수행 하는 단계;를 수행하는 것을 특징으로 하는, 위상학적 물 분자 네트 워크 기반 분자 조각 적합 스크리닝 방법.

청구항 3

제1항에 있어서,

상기 단백질 구조의 분자 동역학 모의 실험하는 단계는,

준비한 단백질 구조를 입력하는 단계;

역장 부여 및 물 분자 모델로 용매화 작업하는 단계;

에너지 최소화, 평형화를 수행하는 단계; 및

10나노초 동안 모의 실험을 수행하는 단계;를 수행하는 것을 특징으로 하는, 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 방법.

청구항 4

제1항에 있어서,

상기 단백질 구조와 결합하는 화합물의 분자 조각을 분석하는 단계 는,

분자 도킹을 통한 결합 모드를 예측하는 단계; 및

분자 조각화를 수행하는 단계;를 수행하는 것 을 특징으로 하는, 위상학적 물 분자 네트워크 기반 분자 조각 적

합 스크리닝 방법.

청구항 5

제1항에 있어서,

상기 분석된 그룹화 물 분자 네트워크와 분자 조각 사이의 모양 유사도 및 평균 거리를 계산하는 단계는,

그룹화 물 분자 네트워크와 분자 조각의 모양 유사도를 계산하는 단계; 및

그룹화 물 분자 네트워크를 형성하는 물 분자 네트워크와 분자 조각 사이 직교 거리를 계산 및 평균 거리를 계산하는 단계;를 수행하는 것을 특징으로 하는, 위상학적 물분자 네트워크 기반 분자 조각 적합 스크리닝 방법.

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

발명의 설명

기 술 분 야

[0001] 본 발명은 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치 및 방법에 관한 것으로, 보다 상세하게는 각각의 키나아제에 특이적으로 상호작용하는 화합물의 분자 조각을 효과적으로 도출하는 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 본 발명에 관련된 종래기술에는 수계 약리작용단, 물 분자 네트워크 분석이 있다. 특허문헌 1 수계 약리작용단을 이용하는 방법은 표적 단백질의 결합 부위의 주형을 제조하기 위한 방법을 제공한다. 또한, 특허문헌 2 물 분자 네트워크 분석 방법은 직관적인 UI(User Interface)로 구성된 웹 기반의 플랫폼을 통해 분자동역학 모의실험과 물 분자 네트워크 분석을 용이하게 수행할 수 있고, 물 분자 네트워크 분석에 대한 실험 데이터와 결과 데이터의 이력관리를 용이하게 할 수 있다.

선행기술문헌

특허문헌

[0003] (특허문헌 0001) 공개특허공보 제10-2016-0128288호(2016.11.07)

(특허문헌 0002) 등록특허공보 제10-1902150호(2018.10.12)

발명의 내용

해결하려는 과제

[0004] 본 발명은 단백질 구조의 특정 부분을 효과적으로 추출하고, 해당 부분의 모양과 거리적 특성을 고려하여 적합한 분자 조각을 선별하는 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치 및 방법을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0005] 본 발명의 바람직한 일 실시예에 따른 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치는, 제어

부가 단백질 구조의 분자 동역학을 모의 실험하여 단백질 구조의 위상학적 물 분자 네트워크를 추출하는 모의 실험 단계; 추출한 위상학적 물 분자 네트워크에서 빈번하게 나타나는 위상학적 물 분자 네트워크를 그룹화하는 단계; 단백질 구조와 결합하는 화합물의 분자 조각을 분석하는 분자 조각 단계; 및 추출된 위상학적 물 분자 네트워크의 그룹화 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도 및 평균 거리를 계산하여 임계값 이상인 분자 조각을 스크리닝하는 계산 단계;를 수행하는 것을 특징으로 한다.

[0006] 또한, 상기 계산 단계는, 그룹화 위상학적 물 분자 네트워크를 구성하는 각 첫 번째 위상학적 물 분자 네트워크 그룹화의 4개의 산소 원자 중심을 꼭지점으로 하여 사각형 면을 만들고, 분자 조각의 중심과 직교 거리를 계산하고 평균화하는 단계; 및 그룹화 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산하여 각 계산 항목의 임계값 이상인 분자 조각을 스크리닝하는 단계;를 수행하는 것을 특징으로 한다.

[0007] 또한, 제어부가 단백질 구조를 준비하는 단계; 단백질 구조의 분자 동역학 모의 실험하는 단계; 단백질 구조의 결합 공간 내의 위상학적 물 분자 네트워크를 추출하는 단계; 추출한 위상학적 물 분자 네트워크에서 빈번하게 나타나는 위상학적 물 분자 네트워크를 그룹화하는 단계; 단백질과 결합하는 화합물의 분자 조각을 분석하는 단계; 및 분석된 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도 및 평균 거리를 계산하는 단계;를 수행하는 것을 특징으로 한다.

[0008] 또한, 상기 단백질 구조를 준비하는 단계는, PDB 파일을 다운로드하는 단계; 및 단백질의 결실된 아미노산 잔기 채움 및 pH 7.4에서 이온화 계산을 수행하는 단계;를 수행하는 것을 특징으로 한다.

[0009] 또한, 상기 단백질 구조의 분자 동역학 모의 실험하는 단계는, 준비한 단백질 구조를 입력하는 단계; 역장 부여 및 물 분자 모델로 용매화 작업하는 단계; 에너지 최소화, 평형화를 수행하는 단계; 및 10나노초 동안 모의 실험을 수행하는 단계;를 수행하는 것을 특징으로 한다.

[0010] 또한, 상기 단백질 구조의 위상학적 물 분자 네트워크를 추출하는 단계는, 10피코초 간격으로 궤적 파일을 추출하는 단계; 각 궤적 파일에서 결합자리 내부의 위상학적 물 분자 네트워크를 추출 및 궤적 파일을 하나로 합치는 단계; 및 추출한 위상학적 물 분자 네트워크의 그룹화를 진행하고, 1차: 그리드 기준으로 인접한 위상학적 물 분자 네트워크끼리 그룹화하고, 2차: 1차 그룹화에서 나타나는 산소 원자 집단의 밀도 기반으로 그룹 영역을 확장하는 단계;를 수행하는 것을 특징으로 한다.

[0011] 또한, 상기 단백질과 결합하는 화합물의 분자 조각을 분석하는 단계는, 분자 도킹을 통한 결합 모드를 예측하는 단계; 및 분자 조각화를 수행하는 단계;를 수행하는 것을 특징으로 한다.

[0012] 또한, 상기 분석된 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도 및 평균 거리를 계산하는 단계는, 그룹화 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도를 계산하는 단계; 및 그룹화 위상학적 물 분자 네트워크를 형성하는 위상학적 물 분자 네트워크와 분자 조각 사이 직교 거리를 계산 및 평균 거리를 계산하는 단계;를 수행하는 것을 특징으로 한다.

발명의 효과

[0013] 본 발명은 단백질 구조의 특정 부분을 효과적으로 추출하고, 해당 부분의 모양과 거리적 특성을 고려하여 적합한 분자 조각을 선별함으로써 각각의 키나아제에 특이적으로 상호작용하는 화합물의 분자 조각을 효과적으로 도출하는 효과를 가진다.

도면의 간단한 설명

[0014] 도 1은 본 발명 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치의 구성을 보인 블록도이다.

도 2는 본 발명 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 방법의 동작을 보인 흐름도이다.

도 3은 본 발명 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 분석 방법의 동작을 시각화한 흐름도이다.

도 4는 본 발명 사환 위상학적 물 분자 네트워크를 보인 예시도이다.

도 5는 본 발명 그룹화 위상학적 물 분자 네트워크를 그룹화하는 과정을 보인 예시도이다.

도 6은 본 발명 임계값 설정을 위한 27종의 키나아제 정보를 나타낸다.

도 7은 본 발명 모양 유사도 임계값 설정을 위한 모양 유사도 및 평균 거리 분석을 보인 표이다.

도 8은 본 발명 평균 거리 임계값 설정을 위한 각 키나아제의 X선 결정 분자 조각이 스크리닝된 수 및 비율을 보인 표이다.

도 9는 본 발명 CDK2 키나아제의 아데닌 포켓 위치에 있는 그룹화 위상학적 물 분자 네트워크 위치 구역을 보인 예시도이다.

도 10은 본 발명 CDK2 키나아제 아데닌 포켓의 특정 위치 구역 1(site Ⅰ)에서 TWN-FS을 통해 스크리닝된 화합물의 분자 조각 스크리닝 결과를 보인 예시도이다.

도 11은 본 발명 CHK1 키나아제 아데닌 포켓의 특정 위치 구역 1(site Ⅰ)의 그룹화 위상학적 물 분자 네트워크(도 11a)와 해당 위치에서 TWN-FS를 통해 스크리닝된 화합물의 분자 조각 스크리닝 결과(도 11b 및 11c)를 보인 예시도이다.

도 12는 본 발명 IGF1R 키나아제 아데닌 포켓의 특정 위치 구역 1 (site Ⅰ)의 그룹화 위상학적 물 분자 네트워크 (도 12a)와 해당 위치에서 TWN-FS를 통해 스크리닝된 화합물의 분자 조각 스크리닝 결과(도 12b 및 12c)를 보인 예시도이다.

도 13은 본 발명 ERBB4 키나아제 아데닌 포켓의 특정 위치 구역 1(site Ⅰ)의 그룹화 위상학적 물 분자 네트워크(도 13a)와 해당 위치에서 TWN-FS를 통해 스크리닝된 화합물의 분자 조각 스크리닝 결과(도 13b 및 13c)를 보인 예시도이다.

도 14는 본 발명 키나아제별 활성 및 비활성 화합물 스크리닝 결과를 보인 표이다.

발명을 실시하기 위한 구체적인 내용

[0015] 이하, 도면을 참조하여 본 발명의 바람직한 일 실시 예에 따른 위상학적 물 분자 네트워크(Topological Water Network, TWN) 기반 분자 조각 적합 스크리닝 장치 및 방법에 대하여 상세히 설명하기로 한다. 이하에서 종래 주지된 사항에 대한 설명은 본 발명의 요지를 명확히 하기 위해 생략하거나 간단히 한다. 본 발명의 설명에 포함된 구성은 개별 또는 복합 결합 구성되어 동작한다.

[0016] 도 1은 본 발명 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치의 구성을 보인 블록도이다.

[0017] 도 1을 참조하면, 제어부가 단백질 구조의 분자 동역학을 모의 실험하여 단백질 구조의 위상학적 물 분자 네트워크를 추출하는 모의 실험 단계; 단백질과 결합하는 화합물의 분자 조각을 분석하는 분자 조각 단계; 및 추출된 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도 및 평균 거리를 계산하여 임계값 이상인 분자 조각을 스크리닝하는 계산 단계;를 수행한다.

[0018] 분자 동역학 모의 실험 모듈은 단백질 구조의 분자 동역학을 시뮬레이션하는 역할을 하고, 초기 단백질 구조를 설정하고, 시뮬레이션을 실행하는데 필요한 다양한 매개 변수를 조절한다.

[0019] 위상학적 물 분자 네트워크 추출 모듈은 분자 동역학 시뮬레이션의 결과에서 물 분자의 상호 작용을 분석하여 네트워크를 추출하는 모듈이고, 단백질 결합 공간 내에 빈번하게 나타나는 위상학적 물 분자 네트워크를 그룹화한다.

[0020] 분자 조각 분석 모듈은 단백질과 결합하는 화합물의 결합 모드를 예측하고 조각화하는 단계이다.

[0021] 모양 유사도 및 평균 거리 계산 모듈은 추출된 그룹화 위상학적 물 분자 네트워크와 분자 조각 간의 유사도와 평균 거리를 계산하는 모듈이고, 그룹화 위상학적 물 분자 네트워크의 특성과 각 분자 조각의 특성을 비교하여 유사성을 측정하고 거리를 계산한다.

[0022] 스크리닝 및 임계값 적용 모듈은 계산된 모양 유사도와 평균 거리를 기반으로, 미리 정의된 임계값을 넘는 분자 조각을 식별하는 모듈이고, 적합한 분자 조각을 선별하여 다음 단계로 전달한다.

[0023] 분자 동역학 모의 실험 단계에서 제어부는 초기에 단백질 구조를 설정하고, 분자 동역학 시뮬레이션을 수행하여 시간에 따른 단백질의 다양한 구조를 캡처한다.

[0024] 위상학적 물 분자 네트워크 추출 단계에서 제어부는 모의 실험 결과에서 물 분자와의 상호 작용을 분석하여 위상학적 물 분자 네트워크를 추출한다.

[0025] 분자 조각 분석 단계에서 제어부는 분자 도킹을 통해 예측된 화합물의 결합 모드를 기반으로 단백질과 결합하는 화합물을 조각화하고 식별한다.

[0026] 모양 유사도 및 평균 거리 계산 단계에서 제어부는 화합물의 분자 조각과 그룹화 위상학적 물 분자 네트워크 간의 유사도와 평균 거리를 계산하여 각 분자 조각의 특성을 정량화한다.

[0027] 스크리닝 단계에서 제어부는 계산된 특성을 기반으로 미리 설정된 임계값을 넘는 분자 조각을 식별하여 선택한다.

[0028] 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치는 단백질 구조에 대한 유용한 정보를 추출하고, 특정 조건을 충족하는 분자 조각을 자동으로 식별한다.

[0029] 상기 계산 단계는, 산소 원자 중심을 꼭지점으로 하여 사각형 면을 만들고, 분자 조각의 중심과 직교 거리를 계산 및 평균화하는 단계; 그룹화 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도를 계산하는 단계; 각 단계의 임계값 이상인 분자 조각을 스크리닝하는 단계;를 수행한다.

[0030] 산소 원자 중심 계산 모듈은 분자 동역학 시뮬레이션 결과에서 그룹화 위상학적 물 분자 네트워크를 형성하는 과정에서 나타나는 산소 원자 집단 중심을 꼭지점으로 하는 사각형 면을 생성하고, 해당 면의 특성을 추출한다.

[0031] 직교 거리 계산 모듈은 산소 원자 집단 중심을 기준으로 생성한 면과 분자 조각의 중심과의 직교 거리를 계산하는 모듈이고, 거리 정보를 이용하여 분자 조각의 위치와 특성을 정확히 파악한다.

[0032] 모양 유사도 및 평균 거리 계산 모듈은 그룹화된 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산하는 모듈이고, 분자 조각과 그룹화 위상학적 물 분자 네트워크 간의 형태적 유사성과 거리적 유사성을 평가하여 각 분자 조각의 특성을 측정한다.

[0033] 스크리닝 및 임계값 적용 모듈은 계산된 모양 유사도와 평균 거리를 기반으로, 미리 정의된 임계값을 넘는 분자 조각을 식별하는 모듈이고, 적합한 분자 조각을 선별하여 다음 단계로 전달한다.

[0034] 산소 원자 중심 계산 단계에서 제어부는 분자 동역학 모의 실험 결과에서 산소 원자 중심을 꼭지점으로 하는 사각형 면을 계산한다.

[0035] 직교 거리 계산 단계에서 제어부는 산소 원자 중심을 기준으로 계산한 사각형 면과 분자 조각의 중심과의 직교 거리를 계산하여 각 분자 조각의 위치 정보를 얻는다.

[0036] 모양 유사도 및 평균 거리 계산 단계에서 제어부는 그룹화된 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산하여 분자 조각의 특성을 평가한다.

[0037] 스크리닝 단계에서 제어부는 계산된 특성을 기반으로 미리 설정된 임계값을 넘는 분자 조각을 식별하여 선택한다.

[0038] 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 장치는 단백질 구조의 특정 부분을 효과적으로 추출하고, 해당 부분의 모양과 거리적 특성을 고려하여 적합한 분자 조각을 선별한다.

[0039] 도 2는 본 발명 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 방법의 동작을 보인 흐름도이다.

[0040] 도 2를 참조하면, 제어부가 단백질 구조를 준비하는 단계; 단백질 구조의 분자 동역학 모의 실험하는 단계; 단백질 구조의 위상학적 물 분자 네트워크를 추출하는 단계; 위상학적 물 분자 네트워크를 그룹화하는 단계; 단백질 구조의 분자 조각을 분석하는 단계; 및 그룹화된 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도 및 평균 거리를 계산하는 단계;를 수행한다.

[0041] 상기 단백질 구조를 준비하는 단계는, PDB(Protein Data Bank) 파일을 다운로드하는 단계; 단백질의 결실된 아미노산 잔기 채움 및 pH 7.4에서 이온화 계산을 수행하는 단계;를 수행한다.

[0042] 단백질 구조 준비는 X선 결정 구조들의 (AURKA (PDB: 5DT0), BRAF (PDB: 3Q4C), BTK (PDB: 6X3N), CDK2 (PDB: 4ERW), CHK1 (PDB: 1NVR), CHK2 (PDB: 2CN5), CK2A1 (PDB: 5T1H), DYRK1A (PDB: 4NCT), EGFR (PDB; 4ZAU), EPHA2 (PDB: 5IA2), ERBB4 (PDB: 2R4B), ERK2 (PDB; 4ZZN), FGFR1 (PDB: 5EW8), GSK3B (PDB: 1Q3D), HCK (PDB: 2HK5), IGF1R (PDB: 3NW6), JAK1 (PDB: 6N7A), JAK2 (PDB: 4IVA), JAK3 (PDB: 5LWM), JNK3 (PDB: 4Z9L), LCK (PDB: 3AC1), MELK (PDB: 5M5A), p38a (PDB: 4F9Y), PDK1 (PDB: 1OKY), SYK (PDB: 1XBC), TGFBR1 (PDB: 5QIM), WEE1 (PDB: 1X8B) PDB 파일들을 RCSB 웹사이트에서 다운받는다. PDB 파일에서 결합된 리간드를

제거하고 Discovery studio 2022 (DS program)를 이용하여 'Prepare protein' 모듈을 실행한다. 전처리한 단백질 구조가 준비하는 동안 결실된 아미노산 잔기들과 수소 원자들이 추가되며 원자결합 방향 및 형식 전하들이 부여되고 pH 7.4 기준으로 단백질 전처리가 된다.

[0043] PDB 파일 다운로드 단계에서 제어부는 RCSB 웹사이트로부터 단백질 구조에 해당하는 PDB 파일들을 다운로드한다. 각 PDB 파일은 특정 단백질에 대한 X선 결정 구조를 포함하고 있다.

[0044] 단백질 구조 전처리 단계에서 제어부는 Discovery Studio 2022의 'Prepare protein' 모듈을 실행하여 단백질 구조를 전처리한다. 전처리 단계에서 제어부는 결합된 리간드를 제거하고, 결실된 아미노산 잔기와 수소 원자를 추가하며, 원자결합 방향과 형식 전하를 부여한다.

[0045] 이온화 계산 단계에서 제어부는 pH 7.4에서 이온화 상태를 계산하여 단백질의 전하를 정의하고, 단백질의 화학적 환경을 모의 실험에 더 적합하게 만들어준다.

[0046] 분자 동역학 모의 실험 단계에서 제어부는 전처리된 단백질 구조에 대해 분자 동역학 모의 실험을 수행하고, 단백질의 다양한 구조를 시뮬레이션하여 시간에 따른 움직임을 캡처한다.

[0047] 위상학적 물 분자 네트워크 추출 단계에서 제어부는 분자 동역학 모의 실험 결과에서 물 분자와의 상호 작용을 분석하여 위상학적 물 분자 네트워크를 추출하고, 단백질 결합 부위에서 물 분자 간의 상호 작용을 나타내는 위상학적 물 분자 네트워크를 생성한다. 추출한 위상학적 물 분자 네트워크 그룹화 단계를 진행한다.

[0048] 분자 조각 분석 단계에서 제어부는 단백질과 결합하는 화합물의 분자 조각을 식별하고 분석하고, 각 분자 조각의 특성과 위치를 정밀하게 파악한다.

[0049] 모양 유사도 및 평균 거리 계산 단계에서 제어부는 분자 조각과 그룹화된 위상학적 물 분자 네트워크 간의 모양 유사도와 평균 거리를 계산하여 각 분자 조각의 특성을 정량화하고, 정보는 위상학적 물 분자 네트워크의 형태와 분자 조각 간의 공간적 배치를 고려하여 계산된다.

[0050] 산소 원자 중심 계산 모듈은 그룹화 위상학적 물 분자 네트워크를 형성하는 과정에서 나타나는 산소 원자 집단 중심을 꼭지점으로 하는 사각형 면을 계산하여 추가적인 공간적 특성을 고려한다.

[0051] 직교 거리 계산 모듈은 분자 조각의 중심과의 직교 거리를 계산하여 더 정확한 분자 조각의 위치 정보를 얻는다.

[0052] 모양 유사도 및 평균 거리 계산 모듈은 그룹화된 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산하여 분자 조각의 특성을 보다 정교하게 평가한다.

[0053] 스크리닝 및 임계값 적용 모듈은 계산된 특성을 기반으로 미리 설정된 임계값을 넘는 분자 조각을 식별하여 선별한다.

[0054] 위상학적 물분자 네트워크 기반 분자 조각 적합 스크리닝은 단백질 구조의 특정 부분을 분석하고, 해당 부분에 대한 적절한 분자 조각을 식별하는 데 효과적으로 활용될 수 있다.

[0055] 상기 단백질 구조의 분자 동역학 모의 실험하는 단계는, 준비한 단백질 구조를 입력하는 단계; 역장 부여 및 물 분자 모델로 용매화 작업하는 단계; 에너지 최소화, 평형화를 수행하는 단계; 10나노초 동안 모의 실험을 수행하는 단계;를 수행한다.

[0056] 모든 분자 동역학 모의실험들은 GROMACS(버전 5.1)를 이용하여 수행한다. CHARMM27 전체 원자 파라미터의 역장으로 했으며 정육면체 박스를 씌워 TIP3P 물 분자 모델로 용매화시킨다. 시스템은 카운터 이온을 넣어 중성화시키고, 경사하강(steepest descent) 알고리즘을 5만 단계 거쳐서 에너지 최소화 과정을 거친다. NVT(입자 수, 부피, 온도가 일정) 조건에서 v-rescale thermostat 으로 100 ps 동안 평형화를 거치고, Parrinello-Rahman barostat 알고리즘으로 200 ps NPT(입자 수, 압력, 온도가 일정) 조건에서 평형화를 수행한다. 분자 동역학 최종 모의실험을 298K 온도에서 10 나노초 동안 수행한다. 단백질의 거동은 제한하고. 장범위 정전기적 상호작용(Long-range electrostatics)는 PME로 수행하고, 단범위 정전기적 상호작용(Short-range electrostatics)은 12 Å 로 설정한다. 모든 결합 길이들은 LINC 알고리즘으로 설정한다. 분자 동역학 궤적들은 1ps 마다 추출한다. TWN-FS 분석에는 분자 동역학 궤적 파일을 모두 합쳐서 사용한다.

[0057] 산소 원자 중심 계산 모듈은 그룹화 위상학적 물 분자 네트워크를 형성하는 과정에서 나타나는 산소 원자 집단 중심을 꼭지점으로 하는 사각형 면을 계산하여 추가적인 공간적 특성을 고려한다.

[0058] 직교 거리 계산 모듈은 분자 조각의 중심과의 직교 거리를 계산하여 더 정확한 분자 조각의 위치 정보를 얻는다.

[0059] 모양 유사도 및 평균 거리 계산 모듈은 그룹화된 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산하여 분자 조각의 특성을 보다 정교하게 평가한다.

[0060] 스크리닝 및 임계값 적용 모듈은 계산된 특성을 기반으로 미리 설정된 임계값을 넘는 분자 조각을 식별하여 선별한다.

[0061] 분자 동역학 모의 실험 단계에서 제어부는 단백질 구조 입력 단계를 실행하고, 전처리된 단백질 구조를 GROMACS를 이용하여 분자 동역학 모의 실험을 수행한다.

[0062] 역장 부여 및 물 분자 모델로 용매화 단계에서 제어부는 CHARMM27 전체 원자 파라미터를 사용하여 역장을 부여하고, TIP3P 물 분자 모델로 용매화 작업을 수행한다.

[0063] 에너지 최소화 및 평형화 단계에서 제어부는 시스템에 카운터 이온을 추가하여 중성화하고, 경사하강 알고리즘을 통해 5만 단계 동안 에너지 최소화를 수행한다. 이후, NVT 조건에서 v-rescale thermostat을 사용하여 100피코초 동안 평형화를 거치고, Parrinello-Rahman barostat 알고리즘으로 200피코초 동안 NPT 조건에서 평형화를 수행한다.

[0064] 분자 동역학 최종 모의 실험 단계에서 제어부는 298K 온도에서 10 나노초 동안 분자 동역학 최종 모의 실험을 수행하고, 단백질의 거동을 제한하며, 장범위 정전기적 상호작용은 PME로 수행하고, 단범위 정전기적 상호작용은 12Å로 설정한다.

[0065] 분자 동역학 궤적 추출 단계에서 제어부는 1ps 간격으로 분자 동역학 궤적을 추출하여 나중의 분석에 사용한다.

[0066] TWN-FS 분석 단계에서 제어부는 분자 동역학 궤적 파일을 모두 합쳐 서 TWN-FS 분석에 사용하고, 분자 동역학 시뮬레이션 결과를 상세하게 분석하여 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도 및 평균 거리를 계산하는데 활용한다.

[0067] 단백질의 특정 부분에 대한 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝은 실험적인 조건에서 분자 동역학 시뮬레이션을 수행하고 이를 통해 분자 조각을 스크리닝하며, 이를 상세하게 분석하여 유용한 정보를 도출할 수 있다.

[0068] 상기 단백질 구조의 위상학적 물 분자 네트워크를 추출하는 단계는, 10피코초 간격으로 궤적 파일을 추출하는 단계; 각 궤적 파일에서 결합자리 내부의 위상학적 물 분자 네트워크를 추출 및 궤적 파일을 하나로 합치는 단계; 추출한 위상학적 물 분자 네트워크의 그룹화를 진행하고, 1차: 그리드 기준으로 인접한 위상학적 물 분자 네트워크끼리 그룹화하고, 2차: 1차 그룹화에서 형성된 산소 원자 집단의 밀도의 중심 기반으로 그룹 영역을 확장하는 단계;를 수행한다.

[0069] 물 분자들은 고리형 네트워크를 형성하며, 이를 TWN이라고 정 의한다. 사환 (4-membered ring) TWN을 추출한다. 포텐셜 함수는 TIP3P 물 분자 모델을 이용하며 레너드존스 힘과 쿨롱힘으로 물 분자들 사이의 상호작용 에너지를 계산한다. 상호작용 포텐셜 에너지는 수학식 1과 같다.

수학식 1

[0070]
$$v(a,b) = \sum_{i}^{on} \sum_{j}^{on} \frac{q_i q_j e^2}{r_{ij}} + \frac{A}{r_{oo}^{12}} - \frac{C}{r_{oo}^{6}}$$

[0071] v(a, b) = 물분자 a,b 사이의 상호작용 에너지이고, roo = 산소 원자 사이의 거리이고, qi = i 위치에서 부분전하 (0.834e)이고, qj = j 위치에서 부분전하 (0.417e)이고, rij = qi 와 qj 사이의 거리이고, A = i 와 j 사이의 척력 (582,000 kcal$Å^{12}mol^{-1}$) 이고, C = i 와 j 사이의 인력 (595 kcal$Å^{6}mol^{-1}$)이다.

[0072] 물 분자들 사이의 수소결합을 결정하기 위해, 에너지 값 기준을 -2.25 kcal/mol로 지정한다. 따라서 TWN 분석을 위해, 이전의 분자동역학 모의실험 수행이 요구되며, 4 ring TWN은 결합자리내 수용액 환경을 대변하여 분자 조

각으로 치환을 위한 스크리닝 사용에 이용한다.

[0073] 그룹화 위상학적 물 분자 네트워크는 분자 동역학 궤적 파일을 합친 상태에서 분석한다. 추출된 TWN을 합치면 가시적으로 그룹이 형성됨을 볼 수 있다. 이는 분자 조각으로 대체될 수 있는 위치에 TWN이 존재한다는 가설로 TWN을 그룹화 진행한다. TWN 그룹화는 2단계로 진행한다. 그룹화 진행을 위해 0.5Å 단위로 그리드를 형성하고 3차원 그리드 좌표 시스템 (3D grid coordinate system)에 TWN의 좌표를 저장한다. 첫 번째 그룹화는 TWN을 형성하는 각 산소 원자에서 0.5Å 단위로 xyz 축으로 확장한 후 비교하는 TWN의 4개의 산소 원자가 확장된 그리드 공간 내에 존재할 때 그룹화를 진행한다. 첫 번째 그룹화에서 모든 TWN을 비교하여 진행한다. 두 번째 그룹화는 첫 번째 그룹을 형성하는 4개의 산소 원자 집단을 밀도 기반 중심 알고리즘 (DBSCAN algorithm)을 이용하여 중심을 계산한다. 첫 번째 그룹 TWN들 사이에서 4개의 중심이 각각 반경 1Å 내에 4개의 중심이 위치할 때 두 번째 그룹으로 지정한다. 모든 첫 번째 그룹을 비교 계산하며 최종 형성된 그룹을 그룹화 TWN (grouped TWN)으로 정의하며, 분자 조각을 비교 계산하는데 사용한다.

[0074] 궤적 파일 추출 단계에서 제어부는 10피코초 간격으로 분자 동역학 궤적 파일을 추출하여 분석에 사용한다.

[0075] 위상학적 물 분자 네트워크 추출 및 통합 단계에서 제어부는 각 궤적 파일에서 결합자리 내부의 위상학적 물 분자 네트워크를 추출하고, 이를 하나의 궤적 파일로 합친다.

[0076] 위상학적 물 분자 네트워크 그룹화 단계에서 제어부는 1차 그룹화: 그리드 기준으로 인접한 위상학적 물 분자 네트워크끼리 그룹화하고, 2차 그룹화: 1차 그룹화에서 형성된 산소 원자의 밀도를 기반으로 그룹 영역을 확장하여 위상학적 물 분자 네트워크를 더 큰 그룹으로 형성한다.

[0077] 상호작용 에너지 계산 단계에서 제어부는 TIP3P 물 분자 모델을 사용하여 상호작용 에너지를 계산한다.

[0078] 상호작용 포텐셜 에너지 = 레너드존스 힘 + 쿨롱힘

[0079] 수소결합 결정 단계에서 제어부는 상호작용 포텐셜 에너지 기준으로 수소결합을 판별하며, 에너지 값이 -2.25 kcal/mol 이상인 경우를 수소결합으로 간주한다.

[0080] TWN 그룹화 단계에서 제어부는 0.5Å 단위로 그리드를 형성하고 3D 그리드 좌표 시스템에 TWN 좌표를 저장하고, 1차 그룹화: TWN의 산소 원자를 각각 0.5 단 위로 xyz 축으로 확장한 후, 비교하여 인접한 TWN의 산소 원자가 포함되는 경우 그룹화하고, 2차 그룹화: 첫 번째 그룹 TWN을 형성하는 4개의 산소 원자 집단의 중심을 계산하여 밀도 기반 중심 알고리즘(DBSCAN algorithm)을 사용하여 두 번째 그룹을 형성한다.

[0081] 분자 조각 스크리닝에 활용 단계에서 제어부는 그룹화된 위상학적 물 분자 네트워크를 분자 조각으로 활용하여 분자 조각 적합 스크리닝에 활용한다.

[0082] 그룹화된 위상학적 물 분자 네트워크는 각 그룹이 분자 조각으로 치환될 수 있는 위치를 나타내며, 스크리닝에 활용된다.

[0083] 산소 원자 중심 계산 모듈은 그룹화 위상학적 물 분자 네트워크를 형성하는 과정에서 나타나는 산소 원자 집단 중심을 꼭지점으로 하는 사각형 면을 계산하여 추가적인 공간적 특성을 고려한다.

[0084] 직교 거리 계산 모듈은 분자 조각의 중심과의 직교 거리를 계산하여 더 정확한 분자 조각의 위치 정보를 얻는다.

[0085] 모양 유사도 및 평균 거리 계산 모듈은 그룹화된 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산하여 분자 조각의 특성을 보다 정교하게 평가한다.

[0086] 스크리닝 및 임계값 적용 모듈은 계산된 특성을 기반으로 미리 설정된 임계값을 넘는 분자 조각을 식별하여 선별한다.

[0087] 그룹화된 위상학적 물 분자 네트워크를 기반으로 한 분자 조각 적합 스크리닝은 단백질 구조의 특정 부분에 대한 분석 및 단백질과 결합하는 화합물의 분자 조각 스크리닝을 효과적으로 수행할 수 있다.

[0088] 상기 단백질과 결합하는 구조의 분자 조각을 분석하는 단계는, 분자 도킹을 통한 결합 모드를 예측하는 단계; 분자 조각화를 수행하는 단계;를 수행한다.

[0089] 분자 도킹은 단백질 구조 준비에서 전처리 과정을 거친 단백질을 이용하며 , Discovery studio 2022 내의 LigandFit 모듈을 실행하여 도킹 연구를 진행한다. 도킹 포즈는 도킹 점수가 가장 높은 포즈를 선택하고, 예측 결합 모드에서 화합물의 스캐폴드 부분을 분자 조각 라이브러리 (KinFragLib)에서 조각화한 방법과 유사하게

분자 조각화한다.

[0090] 분자 도킹을 통한 결합 모드 예측 단계에서 제어부는 도킹 프로세스로 전처리된 단백질 구조를 이용하여 Discovery Studio 2022의 LigandFit 모듈을 실행하여 도킹 연구를 수행한다.

[0091] 도킹 포즈 선택에서 제어부는 도킹 점수가 가장 높은 포즈를 선택하여 화합물이 단백질과 결합할 때의 가장 가능성 있는 결합 모드를 예측한다.

[0092] 분자 조각화 수행 단계에서 제어부는 도킹 포즈에서의 분자 조각화로 도킹된 화합물의 예측된 결합 모드에서 분자 조각을 도출한다.

[0093] 제어부는 X선 결정학 구조의 분자 조각 라이브러리인 KinFragLib의 분자 조각과 유사하게 화합물의 스캐폴드 부분을 분자 조각화한다.

[0094] 결합 모드 예측 단계에서 제어부는 단백질과 화합물 간의 상호 작용을 고려 하여 도킹을 통해 결합 모드를 예측하고, 분자 조각 도출 단계에서 도킹된 화합물의 결합 모드에서 분자 조각을 도출한다.

[0095] 스크리닝과 임계값 적용 단계에서 제어부는 분자 조각을 스크 리닝하고, 설정된 임계값을 기반으로 가장 유망한 조각을 선별한다.

[0096] 단백질과 결합하는 구조의 분자 조각을 분석하는 과정에서 도킹 및 분자 조각화를 통한 실험적인 결과를 활용하여 분자 조각의 적합성을 스크리닝하는 효과적인 방법을 제공할 수 있다.

[0097] 상기 분석된 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도 및 평균 거리를 계산하는 단계는, 그룹화 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도를 계산하는 단계; 그룹화 위상학적 물 분자 네트워크를 형성하는 위상학적 물 분자 네트워크와 분자 조각 사이 직교 거리를 계산 및 평균 거리를 계산하는 단계;를 수행한다.

[0098] 그룹화 TWN과 분자 조각 사이의 모양 유사도 계산은 ShaEP (reminiscent of Shape and Electrostatic Potential)방법을 사용한다. 이 방법에서 좌표 이동 없이 계산을 진행하며, 모양 유사도 결과값을 선택하여 사용하고, 계산은 모양 밀도로 구형 가우시안으로 표현되며 수학식 2와 같다.

수학식 2

$$\rho_i(\boldsymbol{r}) = p_i \, exp(-\alpha_i(\boldsymbol{r} - \boldsymbol{R}_i)^2)$$

[0099]

[0100] ρi(r): 원자 i에 대한 특정 위치 r에서의 모양 밀도 값이고, pi: 가우스 가중치이고, αi: 감쇠 계수이고, r: 로컬 좌표이고, Ri: 원자 i의 원자 좌표이다.

[0101] 평균 거리 계산은 첫 번째 그룹화에서 DBSCAN algorithm을 통해 계산된 산소 원자 집단 중심을 꼭지점으로 하여 사각형의 면을 만들고, 분자 조각의 중심과 직교 거리를 계산한다. 그룹화 위상학적 물 분자 네트워크를 형성한 위상학적 물 분자 네트워크의 수가 그룹화 위상학적 물 분자 네트워크간의 차이가 있는 점을 고려하여 평균 거리로 선정한다.

[0102] 스크리닝은 그룹화 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산하여 임계값 이상인 분자 조각을 스크리닝한다.

[0103] 그룹화 위상학적 물 분자 네트워크와 분자 조각의 모양 유사도 계산 단계에서 제어부는 ShaEP 방법 활용 단계에서 그룹화된 위상학적 물 분자 네트워크와 분자 조각 사이의 모양 유사도를 계산하는데 ShaEP 방법을 활용한다.

[0104] 제어부는 좌표 이동 없이 계산을 진행하며, 분자 조각의 위치와 모양을 고려하여 유사성을 평가한다.

[0105] 제어부는 수학식 2 사용하여 모양 밀도를 구형 가우시안으로 표현하며, ShaEP 방법에서는 수학식 2를 사용하여 모양 유사성을 측정한다.

[0106] 그룹화 위상학적 물 분자 네트워크를 형성하는 위상학적 물 분자 네트워크와 분자 조각 사이 직교 거리 계산 및 평균 거리 계산 단계에서 제어부는 평균 거리 계산 단계에서 DBSCAN 알고리즘을 통해 계산된 산소 원자 집단 중

심을 꼭지점으로 하는 사각형 면을 만들고, 분자 조각의 중심과 그룹화 위상학적 물 분자 네트워크를 구성하는 위상학적 물 분자 네트워크까지의 직교 거리를 계산한다.

[0107] 그룹화 위상학적 물 분자 네트워크간의 차이 고려 단계에서 제어부는 그룹화 위상학적 물 분자 네트워크를 형성한 위상학적 물 분자 네트워크의 수가 다를 경우, 그룹 간의 차이를 고려하여 평균 거리를 선정한다.

[0108] 스크리닝 단계에서 제어부는 모양 유사도 및 평균 거리 기반 스크리닝 단계에서 그룹화된 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산한다.

[0109] 임계값 설정 단계에서 제어부는 미리 정의된 임계값을 기반으로 계산된 모양 유사도와 평균 거리가 이 임계값을 초과하는 분자 조각을 선별하여 스크리닝한다.

[0110] 모양 유사도 계산 모듈은 ShaEP 방법의 매개 변수를 조절하여 유사성 계산의 정확성과 효율성을 극대화한다.

[0111] 스크리닝 및 임계값 설정 모듈은 임계값의 조절 가능성을 포함하여 실험적 결과와의 비교를 통해 가장 효과적인 스크리닝을 수행할 수 있도 록 최적화한다.

[0112] 결과 시각화 모듈은 최종 결과를 사용자에게 직관적으로 전달하기 위한 시각화 도구를 추가하여 분석을 용이하게 한다.

[0113] 제어부는 위상학적 물 분자 네트워크와 분자 조각 간의 모양 유사성 및 평균 거리를 계산하여 스크리닝하는 효과적인 방법을 제공한다.

[0114] 도 3은 본 발명 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 분석 방법의 동작을 보인 시각화한 흐름도이다.

[0115] 도 3을 참조하면, 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 분석 방법은 물 분자 네트워크 분석 단계; 프래그먼트(분자 조각) 분석 단계; 그룹화 위상학적 물 분자 네트워크와 프래그먼트(분자 조각) 분석 단계;를 포함한다.

[0116] 위상학적 물 분자 네트워크 분석 단계에서 제어부는 물 분자 동역학 궤적 추출 단계에서 앞서 수행한 단백질 분자 동역학 모의 실험 에서 얻은 궤적 파일을 10 피코초 간격으로 물 분자를 추출하고 상호작용 에너지 계산을 통해 위상학적 물 분자 네트워크를 추출한다.

[0117] 위상학적 물 분자 네트워크 형성단계에서 제어부는 추출된 물 분자 좌표를 기반으로 물 분자들 간의 연결을 확인하여 상호작용 에너지 계산을 통해 추출된 위상학적 물 분자 네트워크를 형성한다.

[0118] TWN 그룹화 단계에서 제어부는 0.5Å 단위로 그리드를 형성하고 3D 그리드 좌표 시스템에 TWN 좌표를 저장하고, 1차 그룹화: TWN의 산소 원자를 각각 0.5Å 단위로 xyz 축으로 확장한 후, 비교하여 인접한 TWN의 산소 원자가 포함되는 경우 그룹화하고, 2차 그룹화: 첫 번째 그룹 TWN을 형성하는 산소 원자 집단의 중심을 계산하여 밀도 기반 중심 알고리즘(DBSCAN algorithm)을 사용하여 두 번째 그룹을 형성한다. 프래그먼트(분자 조각) 분석 단계에서 제어부는 도킹 포즈에서의 분자 조각화로 도킹된 화합물의 예측된 결합 모드에서 분자 조각을 도출한다.

[0119] 제어부는 BRICS algorithm을 활 용한 분자 조각화로 분자 조각 라이브러리인 KinFragLib의 분자 조각 과 유사하게 화합물의 스캐폴드 부분을 분자 조각화한다.

[0120] 그룹화 위상학적 물 분자 네트워크 프래그먼트(분자조각) 분석 단계에서 제어부는 그룹화 위상학적 물 분자 네트워크와 도킹 포즈에서 조각화한 분자 조각 사이의 모양 유사도와 평균 거리를 계산한다.

[0121] 활성 프래그먼트(분자 조각) 추적 단계에서 제어부는 그룹화 위상학적 물 분자 네트워크와 분자 조각 사이 모양 유사도 및 평균 거리가 임계값을 초과하는 분자 조각을 선별한다.

[0122] 분자 조각 선별 제어부는 예측된 도킹 포즈 화합물을 추적하여 특정 프래그 먼트의 활성값을 추적하고, 이를 통해 단백질과의 상호작용 또는 활성 부위를 파악한다.

[0123] 물 분자 네트워크 분석 방법은 보다 효과적이고 심층적인 결과를 제공할 수 있다.

[0124] 도 4는 본 발명 사환 물 분자 네트워크를 보인 예시도이다.

[0125] 도 4를 참조하면, 사환 물 분자 네트워크는 물 분자 (W1-W4)는 막대 모형으로 표현하고, 수소 결합은 점선으로

나타낸다.

[0126] 사환 물 분자 네트워크는 물 분자들이 사환 구조를 형성하는 모습을 시각적으로 나타낸다. 각 물 분자는 막대 모형으로 표현되고, 수소 결합은 점선으로 나타낸다. 이는 분자 간의 상호작용 및 결합을 이해하고자 하는 목적으로 사용된다.

[0127] 물 분자의 막대 모형 표현에서 빨간색은 산소 원자를 나타내고 흰색은 수소 원자들을 나타낸다.

[0128] 각 수소 원자는 산소 원자와의 결합을 나타내는 막대로 연결되어 있다. 막대는 실제 수소 원자의 방향과 결합을 표현한다.

[0129] 수소 결합의 점선 표현에서 물 분자 간의 수소 결합은 점선으로 나타낸다. 수소 결합은 두 물 분자 간에 수소 원자와 산소 원자 사이의 강한 전기적 상호작용을 나타낸다.

[0130] 점선의 방향은 수소 결합의 방향을 나타내며, 이것은 물 분자 간의 특정 방향성 있는 상호작용을 나타낸다.

[0131] 사환 구조의 형성에서 여러 물 분자가 수소 결합을 통해 사환 구조를 형성한 다. 사환은 한 물 분자가 주변의 두 물 분자 또는 세 물 분자와 수소 결합을 형성 하여 구성된 구조를 의미한다.

[0132] 사환 구조는 물 분자 간의 상호작용 패턴을 시각적으로 나타낸다. 시각적 해석은 사환 위상학적 물 분자 네트워크를 통해 물 분자들 간의 패턴 , 구조, 그리고 수소 결합의 방향성을 시각적으로 해석할 수 있다.

[0133] 사환 위상학적 물 분자 네트워크는 분자 동역학적인 상황에서 물 분자들의 상호작용을 이해하는 데 유용하며, 이러한 시각적인 표현은 연구나 시뮬레이션 결 과를 전달하는데 효과적이다.

[0134] 도 5는 본 발명 그룹화 위상학적 물 분자 네트워크를 그룹화하는 과정을 보인 예시도이다.

[0135] 도 5를 참조하면, 그룹화 위상학적 물 분자 네 트워크는 사환 위상학적 물 분자 네트워크의 산소 원자는 빨간색 구 로 표시하고, 사환 위상학적 물 분자 네트워크는 녹색, 빨간색, 파란 색 선으로 표시한다. 인접한 위상학적 물 분자 네트워크로 구성된 첫번째 그룹에서 각 산소 원자의 집단은 검은색 점선으로 표시하고, 검은색 점선의 밀도 중심은 보라색과 노란색 삼각형으로 표시하고, 노란색과 보라색 사이의 거리는 d로 나타낸다. 각 밀도 중심이 거리 d의 조건인 1Å 이내를 만족하는 두 번째 그룹이 최종 그룹화 위상학적 물 분자 네트워크로 정의된다.

[0136] 그룹화 위상학적 물 분자 네트워크는 사환 위상학적 물 분자 네트워크에서 특정 그룹으로 묶인 구조를 시각적으로 나타낸다. 다양한 색상과 기호를 사용하여 산소 원자, 수소 결합, 그룹화된 구조 등을 표현한다.

[0137] 산소 원자의 표시에서 빨간색 구로 표시된 산소 원자는 그룹화된 위상학적 물 분자 네트워크에서 해당 그룹에 속하는 산소 원자들은 빨간색 구 형태로 표시되고, 특정 그룹을 시각적으로 강조하는데 사용된다.

[0138] 수소 결합 및 사환 구조의 표시에서 녹색, 빨간색, 파란색 선은 사환 위상학적 물 분자 네트워크에서의 수소 결합을 나타내고, 각 색상은 다른 사환 구조를 구분한다.

[0139] 검은색 점선은 첫 번째 그룹에서 각 산소 원자의 집단을 나타내고, 해당 그룹 내에서의 산소 원자의 배치를 나타낸다.

[0140] 점선의 밀도 중심 표시에서 검은색 점선의 밀도 중심은 보라색과 노란색 삼각형으로 표시되고, 그룹 내에서의 밀도 중심을 나타낸다.

[0141] 노란색과 보라색 사이의 거리는 d로 나타내며, 밀도 중심 간의 거리를 의미한다.

[0142] 이와 같은 표시는 그룹화된 물 분자 네트워크에서의 구조와 상호작용을 명확하게 시각화하여 사용자에게 전달하고, 특히, 그룹 내의 밀도 중심과 거리 정보는 해당 그룹의 특성을 뚜렷하게 표현한다.

[0143] 도 6은 본 발명 임계값 설정을 위한 27종의 kinase 정보를 나타 낸다.

[0144] 도 6을 참조하면, 키나아제 이름(Kinase, Protein) 및 단백질 데이터 은행 (PDB) 코드 (PDB code), X선 결정 구조 해상도 (Resolution), 분자 조각 라이브러리(KinFragLib)의 해당 키나아제의 아데닌 포켓 위치에 존재하는 분자 조각 수 (AP site fragments), 중복 제거 후 분자 조각 수 (Unique fragments), 각 키나아제에서 추출한 위상학적 물 분자 네트워크 그룹화에 이용된 그리드 수 (Units), 궤적 파일내에 존재하는 위상학적 물 분자 네트워크를 모두 합친 수 (TWNs), 그룹화 위상학적 물 분자 네트워크의 수 (Grouped TWNs)를 나타낸다.

[0145] 임계값 설정을 위해 사용된 27종의 키나아제 정보에 대한 데이터이며, 데이터에는 키나아제 이름, PDB 코드, X선 결정 구조 해상도, 아데닌 포켓 위치에 존재하는 분자 조각 수, 중복 제거 후 분자 조각 수, 그룹화에 이용된 그리드 수, 궤적 파일을 모두 합친 위상학적 물 분자 네트워크의 수, 그룹화 위상학적 물 분자 네트워크의 수를 제공한다. 이러한 정보는 각 키나아제에 대한 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝에 사용될 수 있다.

[0146] 도 7은 본 발명 모양 유사도 임계값 설정을 위한 모양 유사도 및 평균 거리 분석을 보인 표이다.

[0147] 도 7을 참조하면, 각 키나아제에서 조건에 따라 X선 결정 분자 조각 라이브러리 (KinFragLib)의 1,315 분자 조각에서 스크리닝된 분자 조각의 수 및 비율을 나타내고, 모양 유사도는 S로 나타냈으며, 평균 거리는 D로 나타낸다.

[0148] 모양 유사도 (S) 분석은 키나아제에 대해 알려진 X선 결정 분자 조각 라이브러리 (KinFragLib)를 사용하여 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝을 수행하고, 다양한 모양 유사도 (S) 임계값을 설정하고자 S≥0.5, S≥0.6, S≥0.7 을 선택했다.

[0149] 각각의 모양 유사도 임계값에 대해, 분자 조각 라이브러리 (KinFragLib)에서 스크리닝된 분자 조각의 수와 해당 비율을 계산하고, 정보를 토대로 어떤 모양 유사도 임계값이 적절한지 결정할 수 있다.

[0150] 평균 거리 (D) 분석은 키나아제에 대해 알려진 X선 결정 분자 조각 라이브러리 (KinFragLib)를 사용하여 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝을 수행하고, 다양한 평균 거리 (D) , D≤0.3 Å, D≤0.5 Å, D≤1.0 Å 를 선택하여 분석했다.

[0151] 결과에서 모양 유사도 0.6 이상에서 각 평균 거리의 조건에 따라 최소 7.4~11.6%, 최대 39.5~58.7%를 나타내어 스크리닝에 적합성을 보인다.

[0152] 위의 과정을 통해 모양 유사도와 평균 거리 중 모양 유사도에 대한 적절한 임계값은 0.6 이상이다.

[0153] 도 8은 본 발명 평균 거리 임계값 설정을 위한 각 카나아제의 X선 결정 분자 조각이 스크리닝된 수 및 비율을 보인 표이다.

[0154] 도 8을 참조하면, 모양 유사도 0.6 이상에서 평균 거리에 따라 해당 키나아제의 X선 결정 분자 조각이 스크리닝된 수 및 비율이고, 해당 키나아제의 분자 조각을 중복 제거한 수 (Unique fragments), 모양 유사도는 S로 표현하고, 평균 거리는 D로 표현한다.

[0155] 임계값으로 설정한 S=0.6 이상의 모양 유사도에서의 평균 거리(D) D=0.3 Å, D=0.5 Å D=1.0 Å 이하를 선택하고, 각각의 평균 거리 임계값에 대해, X선 결정 분자 조각이 스크리닝된 수 와 해당 비율을 계산한다.

[0156] 중복 제거된 분자 조각 (Unique fragments)은 중복 제거된 분자 조각의 수를 표현한다. 중복 제거는 해당 키나아제에서 동일한 분자 조각이 있는 경우로 중복을 제거하는 과정이다. 이후 27종의 키나아제에서 해당 키나아제의 분자 조각을 50% 이상 스크리닝하는 평균 거리를 분석한다.

[0157] 결과 요약은 27종의 키나아제에서 D=0.3Å 이하에서 12개, 0.5Å 이하에서 20개, 1.0Å이하에서 23개로 임계값은 S=0.6 이상의 모양 유사도와 평균거리는 0.5Å 이하로 선정한다.

[0158] 도 9는 본 발명 CDK2 키나아제의 아데닌 포켓 위치에 있는 그룹화 위상학적 물 분자 네트워크 및 위치 구역을 보인 예시도이다.

[0159] 도 9를 참조하면, CDK2 키나아제 아데닌 포켓에 존재하는 그룹화 위상학적 물 분자 네트워크 및 위치 구역이다.

[0160] 위상학적 물 분자 네트워크 구성은 아데닌 포켓 내에서 분자 동역학 모의실험을 통해 얻은 궤적 파일로부터 물 분자 간의상호작용 에너지를 계산하여 수소 결합 및 다른 상호작용을 고려한 위상학적 물 분자 네트워크를 구성한다.

[0161] 위상학적 물 분자 네트워크 형성은 궤적 파일을 일정 간격으로 분할하고, 각 분할된 궤적 파일에서 물 분자의 상호작용을 고려하여 위상학적 물 분자 네트워크를 형성한다.

[0162] 그룹화 구역 정의에서 그룹화는 위상학적 물 분자 네트워크가 단백질 결합 공간 내에서 밀집하게 위치한 일부 위상학적 물 분자 네트워크들의 군집을 형성하는 과정이고, 군집은 물 분자들이 특정 위치에서 결합 또는 교류하는 영역을 나타낸다.

[0163] 색상 및 표시는 CDK2 키나아제 아데닌 포켓 내에서 특정 구역에 속한 위상학적 물 분자 네트워크들을 시각적으로 다르게 표시하고, 도 9의 5개 구역 중에서 특정 부위 구역 1(site Ⅰ)이 분자 조각을 스크리닝에 적합한 지역이다.

[0164] CDK2 키나아제 아데닌 포켓의 그룹화된 물 분자 네트워크 위치 구역을 시각적으로 이해하고, 해당 영역에서의 물 분자들의 역할 및 특성을 분석할 수 있다.

[0165] 도 10은 본 발명 CDK2 키나아제 아데닌 포켓의 특정 부위 구역 1 (site Ⅰ)에서 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 방법을 기반으로 화합물의 분자 조각 스크리닝 결과를 보인 예시도이다.

[0166] 도 10을 참조하면, CDK2 키나아제 아데닌 포켓 (AP)의 특정 부위 구역 1 (site Ⅰ의 그룹화 위상학적 물 분자 네트워크와 ChEMBL DB내에 단일 CDK2 IC_{50} (활성값, 활성: ≤1,000nM, 비활성: ≥10,000nM) 결과를 가지는 화합물을 도킹 진행 및 조각화 후 분자 조각 스크리닝 결과이다.

[0167] 그룹화된 위상학적 물 분자 네트워크형성은 분자 동력학 모의실험을 통해 얻은 궤적 파일을 기반으로 그룹화된 위상학적 물 분자 네트워크를 형성하고, CDK2 키나아제 아데닌 포켓의 특정 부위인 site Ⅰ에 대해 분석한다.

[0168] ChEMBL 데이터베이스 검색은 ChEMBL 데이터베이스에서 CDK2 IC_{50} 값이 있는 화합물들을 검색하고, 활성 값 (IC_{50} ≤ 1,000 nM) 및 비활성 값 (IC_{50} ≥ 10,000 nM)을 설정하여 필터링한다.

[0169] 화합물의 분자 조각 스크리닝은 검색된 화합물들의 결합 모드를 예측하고 화합물을 분자 조각으로 조각화하고, 각 화합물의 분자 조각들을 CDK2 아데닌 포켓의 site Ⅰ에 대한 그룹화 위상 학적 물 분자 네트워크와 비교한다.

[0170] 스크리닝 결과 해석은 모양 유사도 및 평균 거리를 기반으로 스크리 닝 결과를 해석하고, 모양 유사도가 0.6 이상이고 평균 거리가 0.5 이하의 일정 범위 내에 있는 분자 조각들은 CDK2 아데닌 포켓 site Ⅰ에서 안정적으로 상호작용할 가능성이 높다.

[0172] CDK2 아데닌 포켓의 특정 영역인 site Ⅰ에서 상호작용하는 화합물의 분자 조각을 스크리닝 결과, ChEMBL 데이터베이스활성 화합물을 나타내는 유의미한 분자 조각을 도출할 수 있다.

[0173] 도 11은 본 발명 CHK1 키나아제 아데닌 포켓의 특정 부위 구역 1 (site Ⅰ)에서 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 방법을 기반으로 화합물의 분자 조각 스크리닝 결과를 보인 예시도이다.

[0174] 도 11을 참조하면, CHK1 키나아제 아데닌 포켓 (AP)의 특정 부위 구역 1 (site Ⅰ의 그룹화 위상학적 물 분자 네트워크와 ChEMBL DB내에 단일 CHK1 IC_{50} (활성값, 활성: ≤1,000nM, 비활성: ≥10,000nM) 결과를 가지는 화합물의 분자 조각 스크리닝 결과이다.

[0175] 도 11a는 CHK1 키나아제 아데닌 포켓의 특정 부위 구역 1(site Ⅰ의 그룹화 TWN을 나타낸다.

[0176] 도 11b는 화합물의 분자 조각 스크리닝 결과에서 활성값을 나타내는 분자 조각을 나타내며, 도 11c는 화합물의 분자 조각 스크리닝 결과에서 비활성값을 나타내는 분자 조각을 나타낸다. 빨간색으로 표시된 공통 분자 조각은 활성값 및 비활성값을 나타내는 분자 조각을 나타낸다.

[0177] 그룹화된 위상학적 물 분자 네트워크 형성은 분자 동역학 시뮬레이션을 통해 얻은 궤적 파일을 기반으로 그룹화된 위상학적 물 분자 네트워크를 형성하고, CHK1 아데닌 포켓의 특정 부위인 site Ⅰ에 대해 분 석한다.

[0178] ChEMBL 데이터베이스 검색은 ChEMBL 데이터베이스에서 CHK1 IC_{50} 값이 있는 화합물들을 검색하고, 활성 값 (IC_{50} ≤ 1,000 nM) 및 비활성 값 (IC_{50} ≥ 10,0 00 nM)을 설정하여 필터링한다.

[0179] 화합물의 분자 조각 스크리닝은 검색된 화합물들의 결합 모드를 예측하고 화합물을 분자 조각으로 조각화하고, 각 화합물의 분자 조각 들을 CHK1 아데닌 포켓의 site Ⅰ에 대한 그룹화 위상학적 물 분자 네트워크와 비교한다.

[0180] 스크리닝 결과 해석은 모양 유사도 및 평균 거리를 기반으로 스크리닝 결과를 해석하고, 모양 유사도가 0.6 이상이고 평균 거리가 0.5 이하의 일정 범위 내에 있는 분자 조각들은 CHK1 아데닌 포켓 site Ⅰ에서 안정적으로 상호작용할 가능성이 높다.

[0182] CHK1 아데닌 포켓의 특정 영역인 site I에서 상호작용하는 화합물의 분자 조각을 스크리닝 결과, ChEMBL 데이

터베이스에서 활성 화합물을 나타내는 유의미한 분자 조각을 도출할 수 있다.

[0183] 도 12는 본 발명 IGF1R 키나아제 아데닌 포켓의 특정 부위 구역 1 (site Ⅰ)에서 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 방법을 기반으로 화합물의 분자 조각 스크리닝 결과를 보인 예시도이다.

[0184] 도 12를 참조하면, IGF1R 키나아제 아데닌 포켓 (AP)의 특정 부위 구역 1 (site Ⅰ)의 그룹화 위상학적 물 분자 네트워크와 ChEMBL DB내에 단일 IGF1R IC_{50} (활성값, 활성: ≤ 1,000nM, 비활성: ≥10,000nM) 결과를 가지는 화합물의 분자 조각 스크리닝 결과이다.

[0185] 도 12a는 IGF1R 키나아제 아데닌 포켓의 특정 부위 구역 1 (site Ⅰ)의 그룹화 위상학적 물 분자 네트워크를 나타낸다.

[0186] 도 12b는 화합물의 분자 조각 스크리닝 결과에서 활성값을 나타내는 분자 조각을 나타내며, 도 12c는 화합물의 분자 조각 스크리닝 결과에서 비활성값을 나타내는 분자 조각을 나타낸다. 빨간색으로 표시된 공통 분자 조각은 활성값 및 비활성값을 나타내는 분자 조각을 나타낸다.

[0187] 그룹화된 위상학적 물 분자 네트워크 형성은 분자 동력학 모의실험을 통해 얻은 궤적 파일을 기반으로 그룹화 위상학적 물 분자 네트워크를 형성하고, IGF1R 키나아제 아데닌 포켓의 특정 부위인 site I에 대해 분석한다.

[0188] ChEMBL 데이터베이스 검색은 ChEMBL 데이터베이스에서 IGF1R IC50 값이 있는 화합물들을 검색하고, 활성 값 (IC_{50}≤ 1,000 nM) 및 비활성 값 (IC_{50}≥ 10,000 nM)을 설정하여 필터링한다.

[0189] 화합물의 분자 조각 스크리닝은 검색된 화합물들의 결합모드를 예측하고 화합물을 분자 조각으로 조각화하고, 각 화합물의 분자 조각들 을 IGF1R 아데닌 포켓의 site I에 대한 그룹화 위상학적 물 분자 네트워크와 비교한다.

[0190] 스크리닝 결과 해석은 모양 유사도 및 평균 거리를 기반으로 스크리닝 결과 를 해석하고, 모양 유사도가 0.6이상이고 평균 거리가 0.5 이하의 일정 범위 내에 있는 분자 조각들은 IGF1R 아데닌 포켓 site I에서 안정적으로 상호작용할 가능성이 높다.

[0192] IGF1R 아데닌 포켓 특정 영역인 site I에서 상호작용하는 화합물의 분자 조각을 스크리닝 결과, ChEMBL 데이터베이스에서 활성 화합물을 나타내는 유의미한 분자 조각을 도출할 수 있다.

[0193] 도 13은 본 발명 ERBB4 키나아제 아데닌 포켓의 특정 부위 구역 1(site Ⅰ)에서 위상학적 물 분자 네트워크 기반 분자 조각 적합 스크리닝 방법을 기반으로 화합물의 분자 조각 스크리닝 결과 를 보인 예시도이다.

[0194] 도 13을 참조하면, ERBB4 키나아제 아데닌 포켓(AP)의 특정 부위 구역 1(site Ⅰ의 그룹화 위상학적 물 분자 네트워크와 ChEMBL DB내에 단일 ERBB4 IC_{50} (활성값, 활성: ≤ 1,000nM, 비활성: ≥10,000nM) 결과를 가지는 화합물의 분자 조각 스크리닝 결과이다.

[0195] 도 13a는 ERBB4 키나아제 아데닌 포켓의 특정 부위 구역 1(site Ⅰ의 그룹화 위상학적 물 분자 네트워크를 나타낸다.

[0196] 도 13b는 화합물의 분자 조각 스크리닝 결과에서 활성값을 나타내는 분자 조각을 나타내며, 도 13c는 화합물의 분자 조각 스크리닝 결과에서 비활성값을 나타내는 분자 조각을 나타낸다.

[0197] 그룹화된 위상학적 물 분자 네트워크 형성은 분자 동력학 모의실험을 통해 얻은 궤적 파일을 기반으로 그룹화 위상학적 물 분자 네트워크를 형성하고, ERBB4 키나아제 아데닌 포켓의 특정 부위인 site I에 대해 분석한다.

[0198] ChEMBL 데이터베이스 검색은 ChEMBL 데이터베이스에서 ERBB4 IC_{50} 값이 있는 화합물들을 검색하고, 활성 값 (IC_{50}≤ 1,000 nM) 및 비활성 값 (IC_{50}≥ 10,000 nM) 을 설정하여 필터링한다.

[0199] 화합물의 분자 조각 스크리닝은 검색된 화합물들의 결합 모드를 예측하고 화합물을 분자 조각으로 조각화하고, 각 화합물의 분자 조각들을 ERBB4 아데닌 포켓의 site Ⅰ에 대한 그룹화 위상학 적 물 분자 네트워크와 비교한다.

[0200] 스크리닝 결과 해석은 모양 유사도 및 평균 거리를 기반으로 스크리닝 결과를 해석하고, 모양 유사도가 0.6이상이고 평균 거리가 0.5 이하의 일정 범위 내에 있는 분자 조각들은 ERBB4 아데닌 포켓 site I에서 안정적으로 상호 작용할 가능성이 높다.

[0202] ERBB4 아데닌 포켓 특정 영역인 site I에서 상호작용하는 화합물의 분자 조각을 스크리닝 결과, ChEMBL 데이터베이스에서 활성 화합물을 나타내는 유의미한 분자 조각을 도출할 수 있다.

[0203] 도 14는 본 발명 키나아제별 활성 및 비활성 화합물 스크리닝 결과를 보인 표이다.

[0204] 도 14를 참조하면, 키나아제별 활성, 비활성 화합물의 수 및 그룹화 위상학적 물 분자 기반 분자 조각 적합 스크리닝 결과를 설명한다.

[0205] 각 키나아제에 대해 ChEMBL 데이터베이스에서 얻은 화합물에서 활성 화합물 ($IC_{50} \leq 1{,}000$ nM) 및 비활성 화합물 ($IC_{50} \geq 10{,}000$ nM)로 분류하고 화합물의 결합 모드를 예측하고 분자 조각으로 조각화한다.

[0206] 그룹화 위상학적 물 분자 네트워크 분석은 각 키나아제 단백질에 대해 분자 동력학 모의실험을 통해 얻은 그룹화 위상학적 물 분자 네트워크를 키나아제 아데닌 포켓 특정 영역인 site I에 대해 분석한다.

[0207] 각 화합물을 조각화한 분자 조각을 위상학적 물 분자 기반 분자 조각 적합 스크리닝방법 기반으로 스크리닝하고, 키나아제 아데닌 포켓 site I에 위치하는 분자 조각을 설명한다. 임계값을 만족하는 안정적으로 상호작용하는 화합물의 분자 조각들을 도출한다.

[0208] 도출된 분자 조각은 활성 화합물을 나타내는 특성을 높게 가진다.

[0209] 각 키나아제에 특이적으로 상호작용하는 화합물의 분자 조각을 효과 적으로 도출할 수 있다.

[0210] 본 발명은 상술한 특정의 바람직한 실시 예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 해당 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 누구든지 다양한 변형실시가 가능한 것은 물론이고, 그와 같은 변경은 청구범위 기재의 범위 내에 있게 된다.

부호의 설명

[0211] 1: 프로세서

2: 메모리

3: 입출력장치

4: 운영체제

5: 제어부

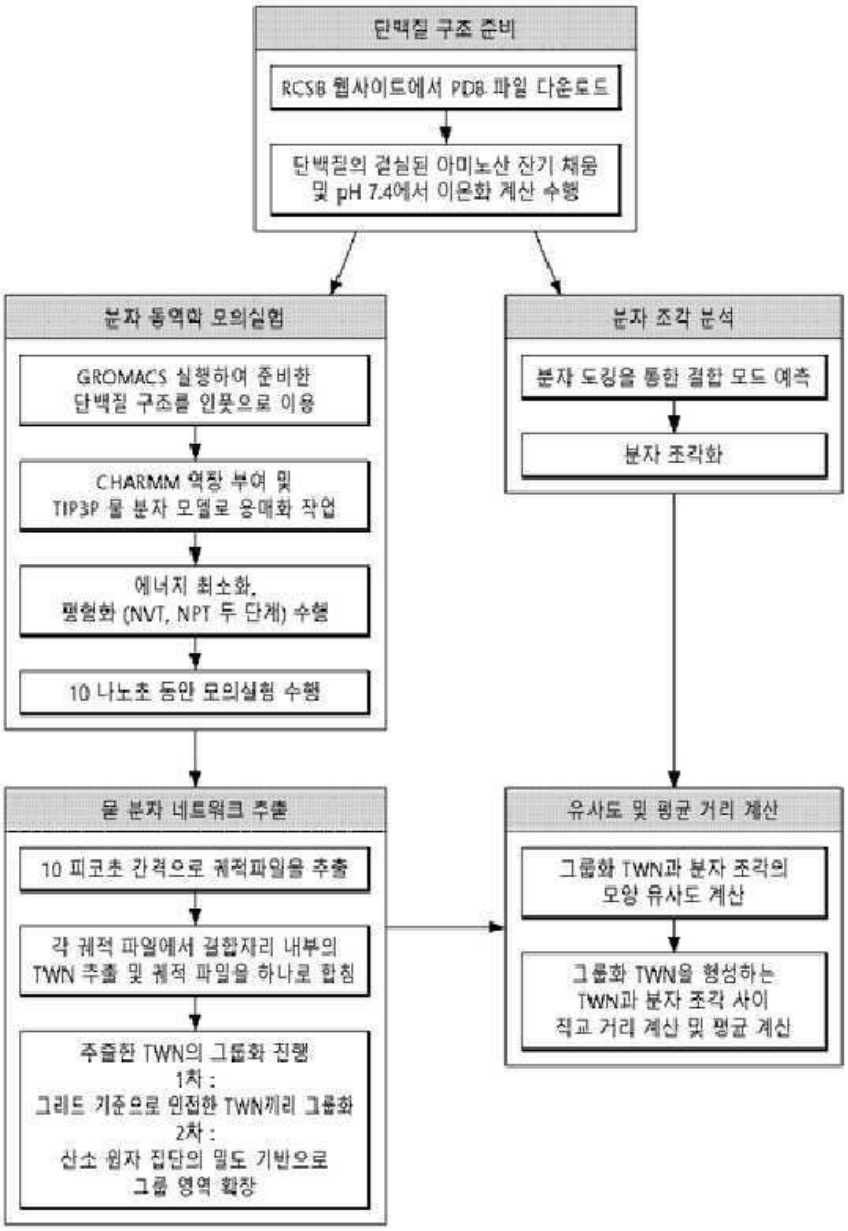
6: 단말기

도면

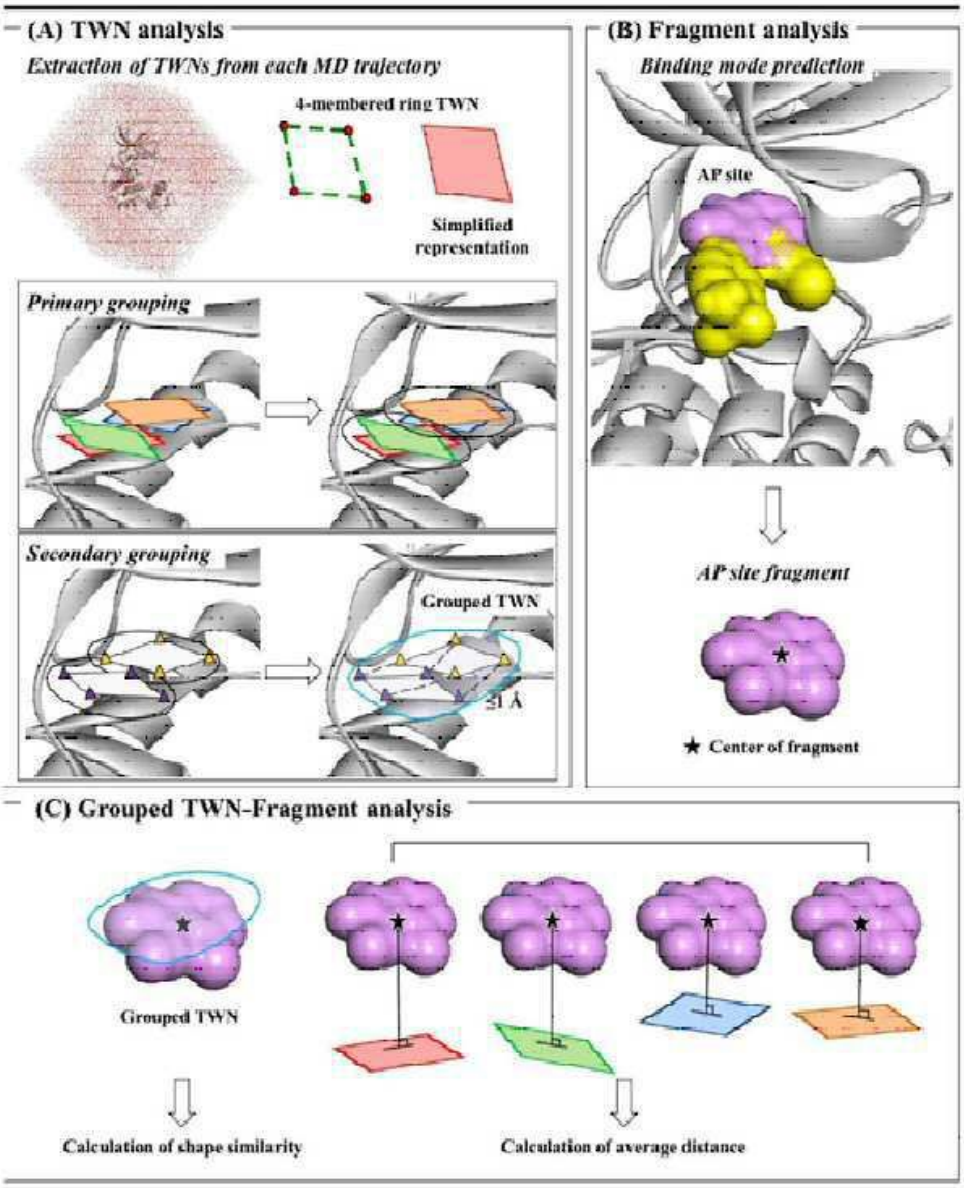
도면1



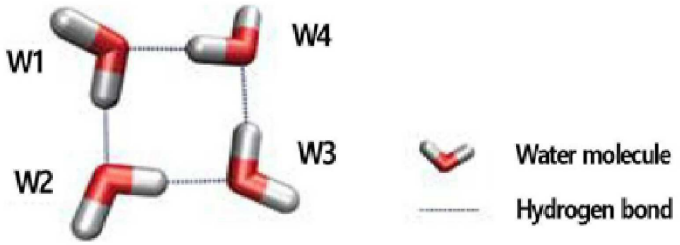
도면2


단백질 구조 준비
RCSB 웹사이트에서 PDB 파일 다운로드
단백질의 결실된 아미노산 잔기 채움 및 pH 7.4에서 이온화 계산 수행
분자 동역학 모의실험
GROMACS 실행하여 준비한 단백질 구조를 인풋으로 이용
CHARMM 력장 부여 및 TIP3P 물 분자 모델로 용매화 작업
에너지 최소화, 평형화 (NVT, NPT 두 단계) 수행
10 나노초 동안 모의실험 수행
분자 조각 분석
분자 도킹을 통한 결합 모드 예측
분자 조각화
물 분자 네트워크 추출
10 피코초 간격으로 궤적파일을 추출
각 궤적 파일에서 결합재리 내부의 TWN 추출 및 궤적 파일을 하나로 합침
추출한 TWN의 그룹화 진행
1차 :
그리드 기준으로 인접한 TWN끼리 그룹화
2차 :
산소 원자 집단의 밀도 기반으로 그룹 영역 확장
유사도 및 평균 거리 계산
그룹화 TWN과 분자 조각의 모양 유사도 계산
그룹화 TWN을 형성하는 TWN과 분자 조각 사이 직교 거리 계산 및 평균 계산

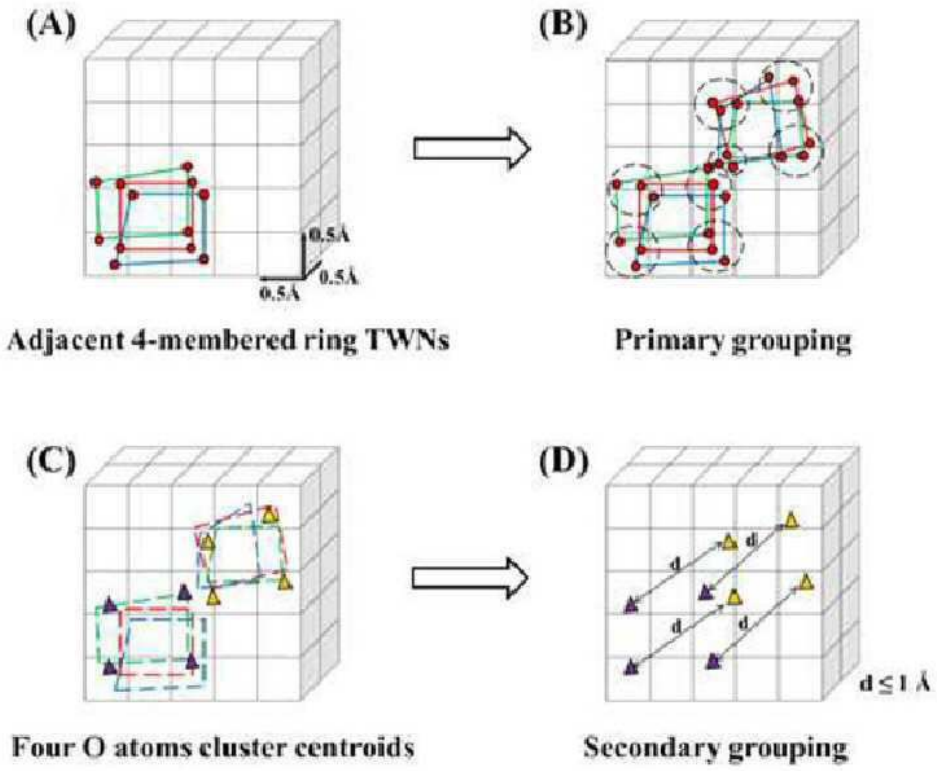
도면3


(A) TWN analysis
Extraction of TWNs from each MD trajectory
4-membered ring TWN
Simplified representation
Primary grouping
Secondary grouping
Grouped TWN
≤1 Å
(B) Fragment analysis
Binding mode prediction
AP site
AP site fragment
★ Center of fragment
(C) Grouped TWN-Fragment analysis
Grouped TWN
Calculation of shape similarity
Calculation of average distance

도면4


W1
W2
W3
W4
Water molecule
Hydrogen bond

도면5


(A)
(B)
0.5Å
0.5Å
0.5Å
Adjacent 4-membered ring TWNs
Primary grouping
(C)
(D)
d
d
d
d
d ≤ 1 Å
Four O atoms cluster centroids
Secondary grouping

도면6

Kinase	Protein	PDB code	Resolution (Å)	AP site fragments	Unique fragments	Units	TWNs	Grouped TWNs
AURKA	Aurora Kinase A	5DTO	2.15	66	35	23,492	9,230	468
BRAF	B-Raf Proto-Oncogene, Serine/Threonine Kinase	3Q4C	3.20	43	28	24,598	9,982	469
BTK	Bruton Tyrosine Kinase	6X3N	1.95	52	33	27,255	10,669	529
CDK2	Cyclin Dependent Kinase 2	4ERW	2.00	330	168	22,244	9,039	898
CHK1	Checkpoint Kinase 1	1NVR	1.80	121	87	32,281	12,760	726
CHK2	Checkpoint Kinase 2	2CN5	2.25	29	23	34,991	13,740	1,165
CK2A1	Casein Kinase 2 Alpha 1	5T1H	2.11	58	40	26,995	10,916	1,586
DYRK1A	Dual Specificity Tyrosine Phosphorylation Regulated Kinase 1A	4NCT	2.60	31	28	29,811	12,179	915
EGFR	Epidermal Growth Factor Receptor	4ZAU	2.80	66	37	26,771	10,656	653
EPHA2	EPH Receptor A2	5IAZ	1.62	26	8	26,221	10,525	598
ERBB4	Erb-B2 Receptor Tyrosine Kinase 4	2R4B	2.40	1	1	24,237	9,769	841
ERK2	Mitogen-Activated Protein Kinase 1	4ZZN	1.33	77	40	23,857	9,855	1,467
FGFR1	Fibroblast Growth Factor Receptor 1	5EW8	1.63	37	28	26,550	10,748	674
GSK3B	Glycogen Synthase Kinase 3 Beta	1Q3D	2.20	45	36	31,051	12,390	696
HCK	HCK Proto-Oncogene, Src Family Tyrosine Kinase	2HK5	2.00	23	5	32,365	13,349	692
IGF1R	Insulin Like Growth Factor 1 Receptor	3NW6	2.20	13	9	29,039	11,937	1,615
JAK1	Janus Kinase 1	6N7A	1.33	31	21	23,550	9,550	642
JAK2	Janus Kinase 2	4IVA	1.50	60	40	23,615	9,428	594
JAK3	Janus Kinase 3	5LWM	1.55	22	12	25,273	10,327	1,048
JNK3	Mitogen-Activated Protein Kinase 10	4Z9L	2.10	40	27	20,168	8,233	1,557
LCK	LCK Proto-Oncogene, Src Family Tyrosine Kinase	3ACI	1.99	25	24	30,975	12,774	946
MELK	Maternal Embryonic Leucine Zipper Kinase	5M5A	1.90	21	17	22,298	8,870	815
p38a	Mitogen-Activated Protein Kinase 14	4F9Y	1.85	101	70	22,330	8,886	643
PDK1	Pyruvate Dehydrogenase Kinase 1	1OKY	2.30	30	20	29,083	11,331	417
SYK	Spleen Associated Tyrosine Kinase	1XBC	2.00	59	39	27,181	10,503	520
TGFBR1	Transforming Growth Factor Beta Receptor 1	5QIM	1.75	24	19	25,084	9,964	717
WEE1	WEE1 G2 Checkpoint Kinase	1X8B	1.81	20	11	27,984	11,341	599

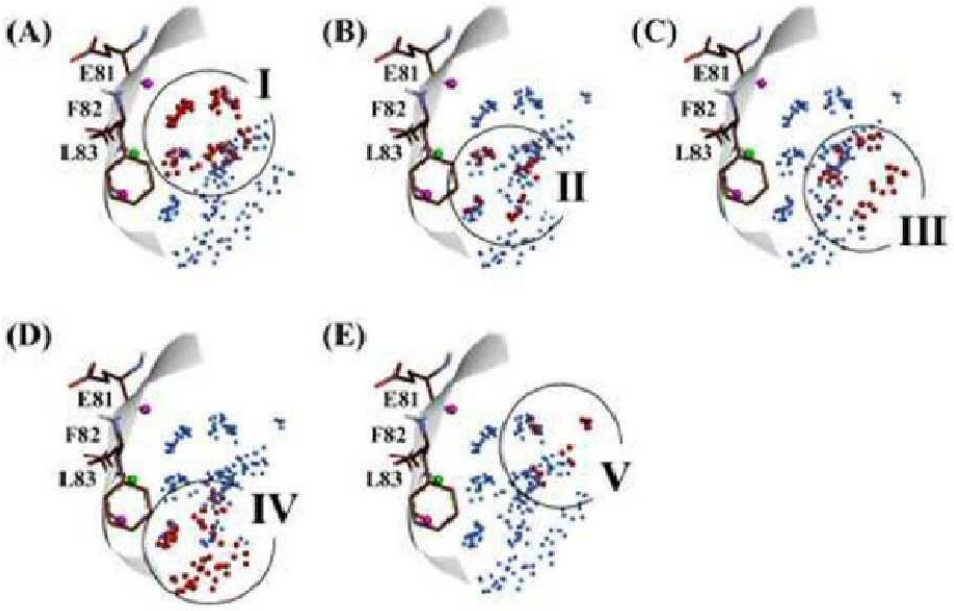
도면7

Kinase	S ≥0.5			S ≥0.6			S ≥0.7		
	D ≤0.3	D ≤0.5	D ≤1.0	D ≤0.3	D ≤0.5	D ≤1.0	D ≤0.3	D ≤0.5	D ≤1.0
AURKA	617 (46.9%)	842 (64.0%)	1035 (78.7%)	350 (26.6%)	475 (36.1%)	553 (42.1%)	43 (3.3%)	67 (5.1%)	81 (6.2%)
BRAF	664 (50.5%)	873 (66.4%)	1,042 (79.2%)	468 (35.6%)	613 (46.6%)	697 (53.0%)	180 (13.7%)	237 (18.0%)	271 (20.6%)
BTK	620 (47.1%)	811 (61.7%)	932 (70.9%)	346 (26.3%)	418 (31.8%)	478 (36.3%)	55 (4.2%)	76 (5.8%)	85 (6.5%)
CDK2	280 (21.3%)	504 (38.3%)	918 (69.8%)	99 (7.5%)	217 (16.5%)	496 (37.7%)	2 (0.2%)	25 (1.9%)	113 (8.6%)
CHK1	471 (35.8%)	740 (56.3%)	1,089 (82.8%)	319 (24.3%)	523 (39.8%)	772 (58.7%)	49 (3.7%)	115 (8.7%)	185 (14.1%)
CHK2	642 (48.8%)	897 (68.2%)	1,087 (82.7%)	348 (26.5%)	486 (37.0%)	570 (43.3%)	64 (4.9%)	96 (7.3%)	107 (8.1%)
CK2A1	303 (23.0%)	495 (37.6%)	923 (70.2%)	112 (8.5%)	227 (17.3%)	502 (38.2%)	10 (0.8%)	57 (4.3%)	140 (10.6%)
DYRK1A	660 (50.2%)	911 (69.3%)	1,085 (82.5%)	467 (35.5%)	612 (46.5%)	689 (52.4%)	175 (13.3%)	213 (16.2%)	224 (17%)
EGFR	504 (38.3%)	698 (53.1%)	891 (67.8%)	229 (17.4%)	339 (25.8%)	424 (32.2%)	50 (3.8%)	78 (5.9%)	110 (8.4%)
EPHA2	721 (54.8%)	941 (71.6%)	1,108 (84.3%)	520 (39.5%)	665 (50.6%)	754 (57.3%)	195 (14.8%)	252 (19.2%)	287 (21.8%)
ERBB4	424 (32.2%)	687 (52.2%)	1,070 (81.4%)	330 (25.1%)	529 (40.2%)	741 (56.3%)	163 (12.4%)	262 (19.9%)	290 (22.1%)
ERK2	626 (47.6%)	802 (61.0%)	916 (69.7%)	363 (27.6%)	443 (33.7%)	478 (36.3%)	44 (3.3%)	68 (5.2%)	77 (5.9%)
FGFR1	775 (58.9%)	965 (73.4%)	1,105 (84.0%)	501 (38.1%)	617 (46.9%)	691 (52.5%)	81 (6.2%)	115 (8.7%)	135 (10.3%)
GSK3B	430 (32.7%)	653 (49.7%)	1,107 (84.2%)	261 (19.8%)	401 (30.5%)	759 (57.7%)	61 (4.6%)	127 (9.7%)	327 (24.9%)
HCK	620 (47.1%)	776 (59.0%)	897 (68.2%)	344 (26.2%)	428 (32.5%)	504 (38.3%)	66 (5.0%)	99 (7.5%)	132 (10.0%)
IGF1R	231 (17.6%)	334 (25.4%)	499 (37.9%)	97 (7.4%)	134 (10.2%)	152 (11.6%)	18 (1.4%)	24 (1.8%)	26 (2.0%)
JAK1	606 (46.1%)	813 (61.8%)	1,046 (79.5%)	444 (33.8%)	571 (43.4%)	689 (52.4%)	116 (8.8%)	156 (11.9%)	174 (13.2%)
JAK2	436 (33.2%)	691 (52.5%)	1,084 (82.4%)	340 (25.9%)	529 (40.2%)	768 (58.4%)	155 (11.8%)	235 (17.9%)	289 (22.0%)
JAK3	264 (20.1%)	499 (37.9%)	1,022 (77.7%)	193 (14.7%)	368 (28.0%)	695 (52.9%)	64 (4.9%)	130 (9.9%)	226 (17.2%)
JNK3	366 (27.8%)	492 (37.4%)	635 (48.3%)	156 (11.9%)	198 (15.1%)	223 (17.0%)	29 (2.2%)	32 (2.4%)	33 (2.5%)
LCK	636 (48.4%)	890 (67.7%)	1,124 (85.5%)	437 (33.2%)	603 (45.9%)	764 (58.1%)	110 (8.4%)	183 (13.9%)	245 (18.6%)
MELK	473 (36.0%)	680 (51.7%)	1,008 (76.7%)	185 (14.1%)	327 (24.9%)	553 (42.1%)	23 (1.7%)	63 (4.8%)	153 (11.6%)
p38a	451 (34.3%)	635 (48.3%)	865 (65.8%)	294 (22.4%)	409 (31.1%)	507 (38.6%)	78 (5.9%)	97 (7.4%)	108 (8.2%)
PDK1	716 (54.4%)	931 (70.8%)	1,057 (80.4%)	426 (32.4%)	529 (40.2%)	589 (44.8%)	106 (8.1%)	136 (10.3%)	151 (11.5%)
SYK	644 (49.0%)	871 (66.2%)	1,078 (82.0%)	361 (27.5%)	477 (36.3%)	579 (44.0%)	17 (1.3%)	25 (1.9%)	26 (2.0%)
TGFBR1	644 (49.0%)	882 (67.1%)	1,062 (80.8%)	389 (29.6%)	493 (37.5%)	555 (42.2%)	79 (6.0%)	102 (7.8%)	113 (8.6%)
WEE1	316 (24.0%)	553 (42.1%)	1,039 (79.0%)	192 (14.6%)	367 (27.9%)	681 (51.8%)	43 (3.3%)	83 (6.3%)	115 (8.7%)

도면8

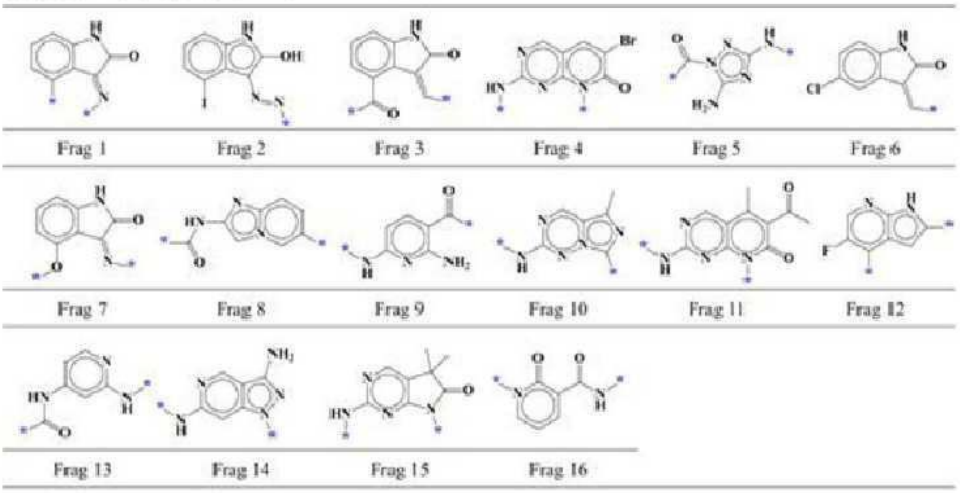
Kinase	Unique fragments	S ≥0.6		
		D ≤0.3	D ≤0.5	D ≤1.0
AURKA	35	17 (48.6%)	22 (62.9%)	24 (68.6%)
BRAF	28	13 (46.4%)	20 (71.4%)	22 (78.6%)
BTK	33	19 (57.6%)	23 (69.7%)	26 (78.8%)
CDK2	168	21 (12.5%)	32 (19.0%)	49 (29.2%)
CHK1	87	32 (36.8%)	54 (62.1%)	67 (77.0%)
CHK2	23	13 (56.5%)	14 (60.9%)	14 (60.9%)
CK2A1	40	4 (10.0%)	6 (15.0%)	18 (45.0%)
DYRK1A	28	18 (64.3%)	20 (71.4%)	23 (82.1%)
EGFR	37	14 (37.8%)	21 (56.8%)	23 (62.2%)
EPHA2	8	4 (50.0%)	6 (75.0%)	7 (87.5%)
ERBB4	1	0 (0.0%)	1 (100.0%)	1 (100.0%)
ERK2	40	22 (55.0%)	25 (62.5%)	26 (65.0%)
FGFR1	28	12 (42.9%)	18 (64.3%)	24 (85.7%)
GSK3B	36	10 (27.8%)	16 (44.4%)	23 (63.9%)
HCK	5	3 (60.0%)	3 (60.0%)	4 (80.0%)
IGF1R	9	5 (55.6%)	8 (88.9%)	8 (88.9%)
JAK1	21	18 (85.7%)	18 (85.7%)	18 (85.7%)
JAK2	40	28 (70.0%)	34 (85.0%)	38 (95.0%)
JAK3	12	3 (25.0%)	10 (83.3%)	11 (91.7%)
JNK3	27	13 (48.1%)	14 (51.9%)	14 (51.9%)
LCK	24	2 (8.3%)	5 (20.8%)	14 (58.3%)
MELK	17	1 (5.9%)	1 (5.9%)	2 (11.8%)
p38a	70	14 (20.0%)	19 (27.1%)	25 (35.7%)
PDK1	20	11 (55.0%)	13 (65.0%)	13 (65.0%)
SYK	39	25 (64.1%)	28 (71.8%)	29 (74.4%)
TGFBR1	19	15 (78.9%)	15 (78.9%)	16 (84.2%)
WEE1	11	2 (18.2%)	4 (36.4%)	7 (63.6%)

도면9

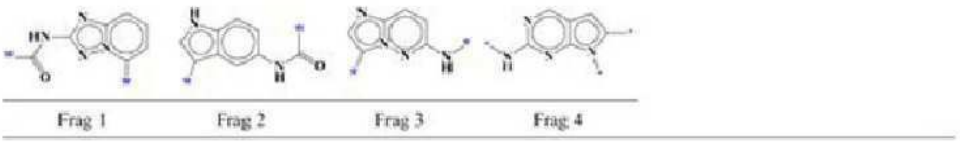


도면10

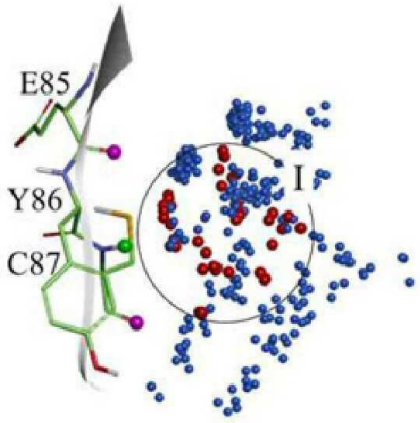
활성값을 나타내는 분자 조각



비활성값을 나타내는 분자 조각

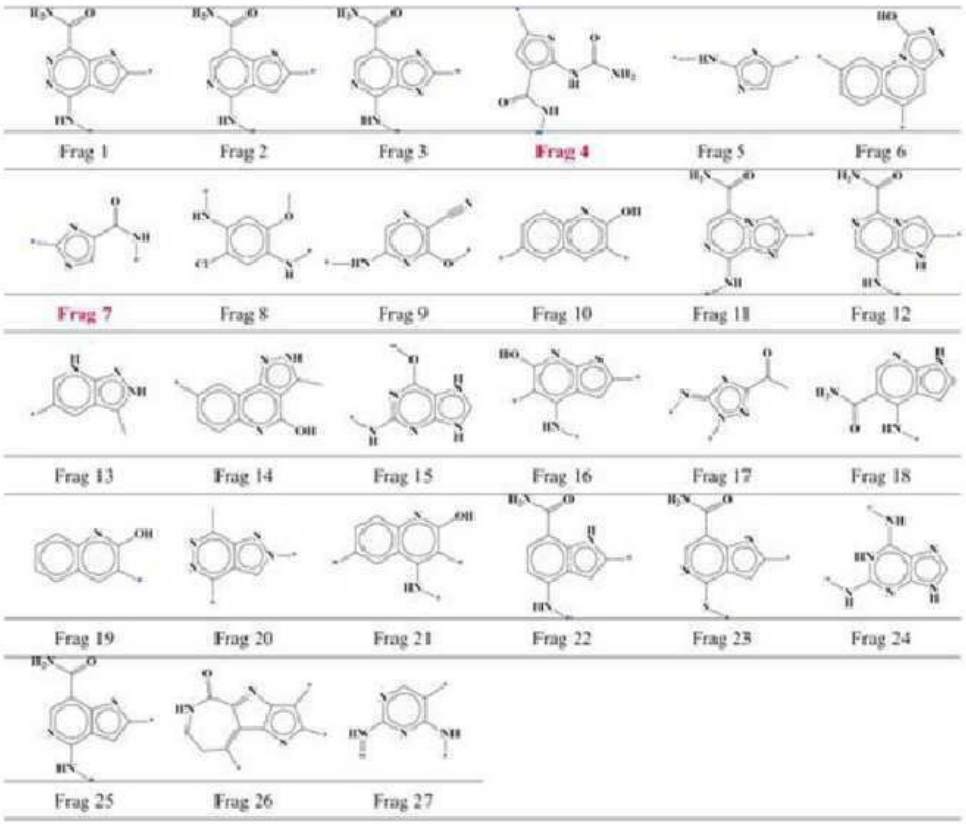


도면11a



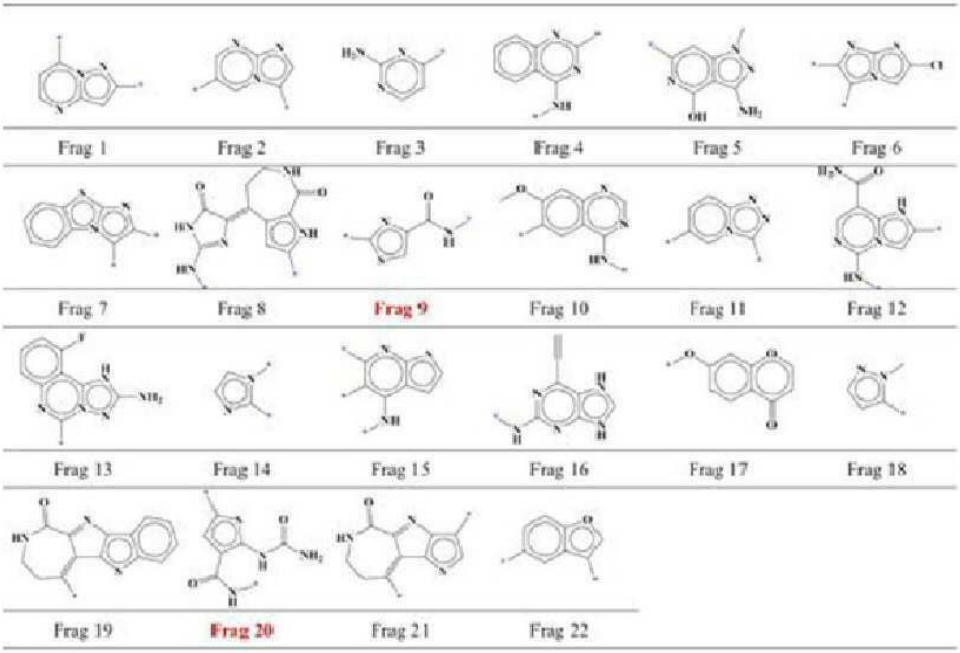
도면11b

활성값을 나타내는 분자 조각

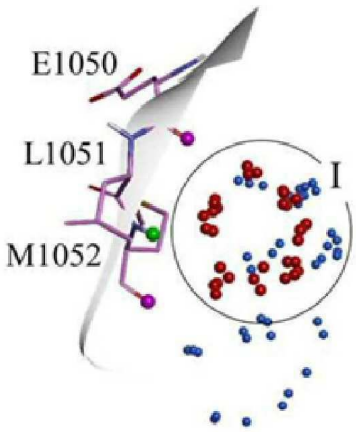


도면11c

비활성값을 나타내는 분자 조각

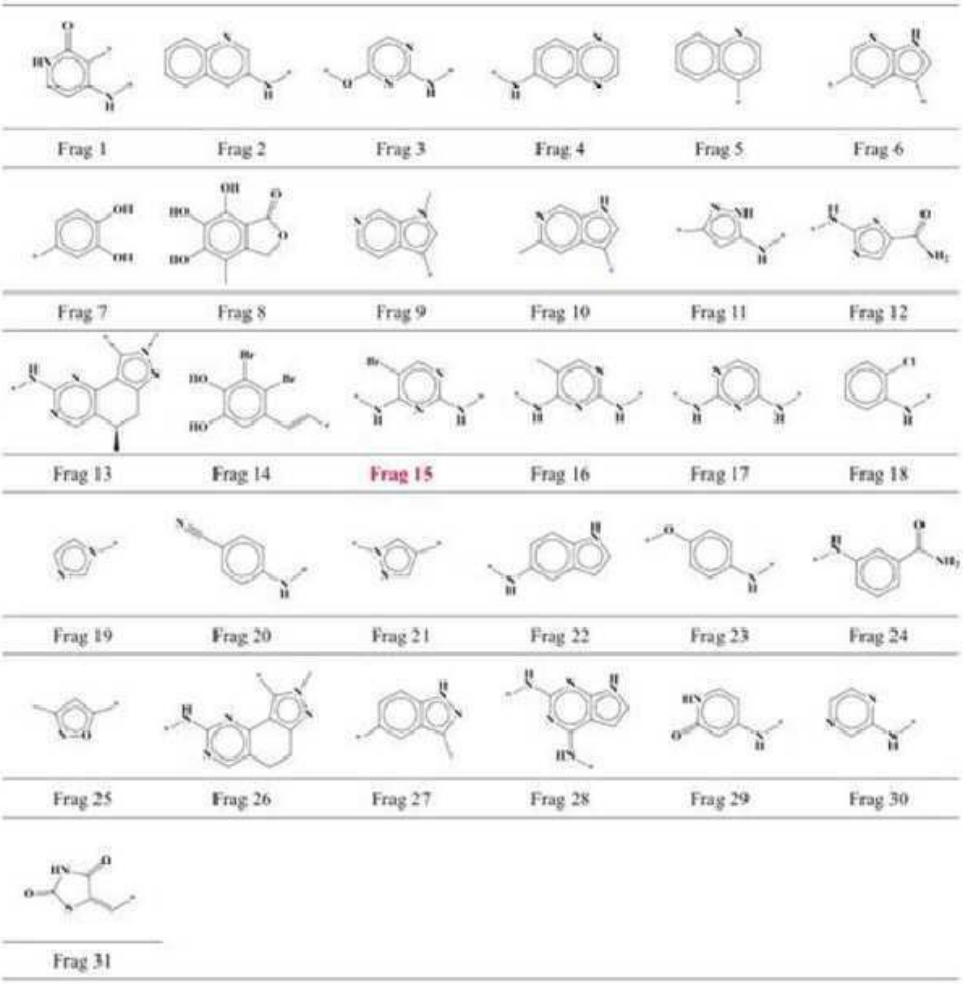


도면12a



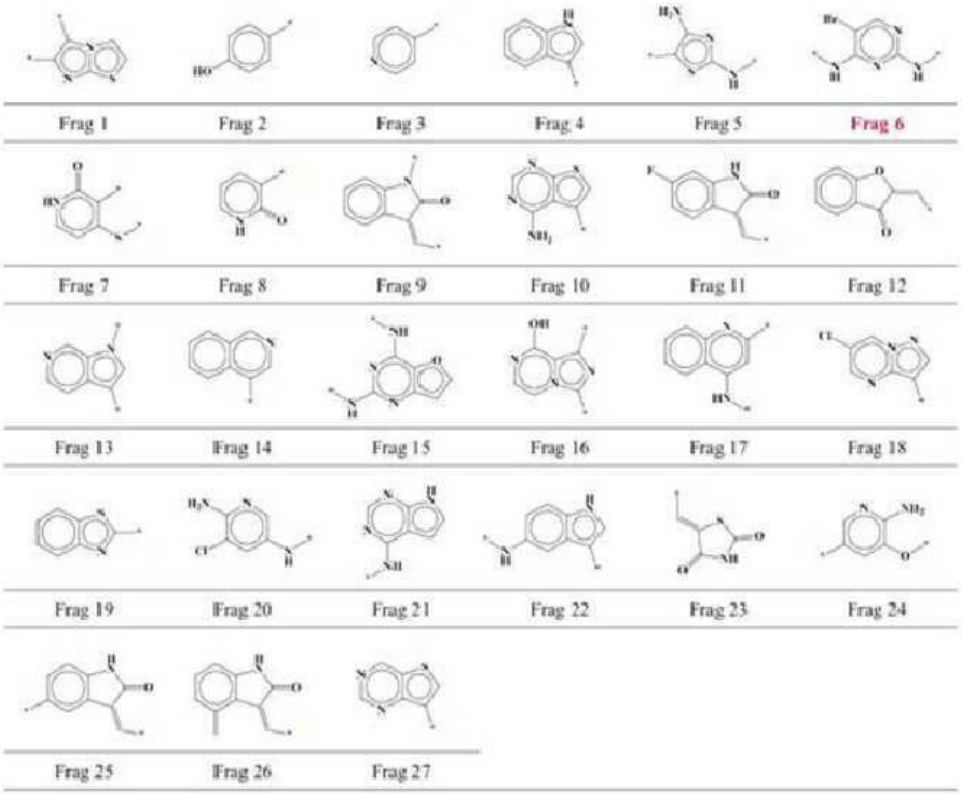
도면12b

활성값을 나타내는 분자 조각

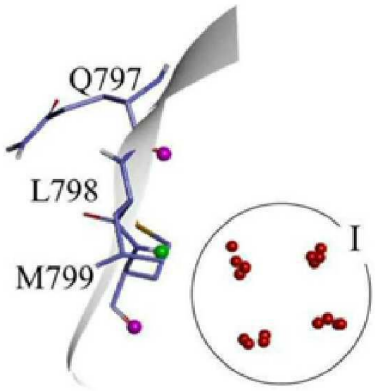


도면12c

비활성값을 나타내는 분자 조각

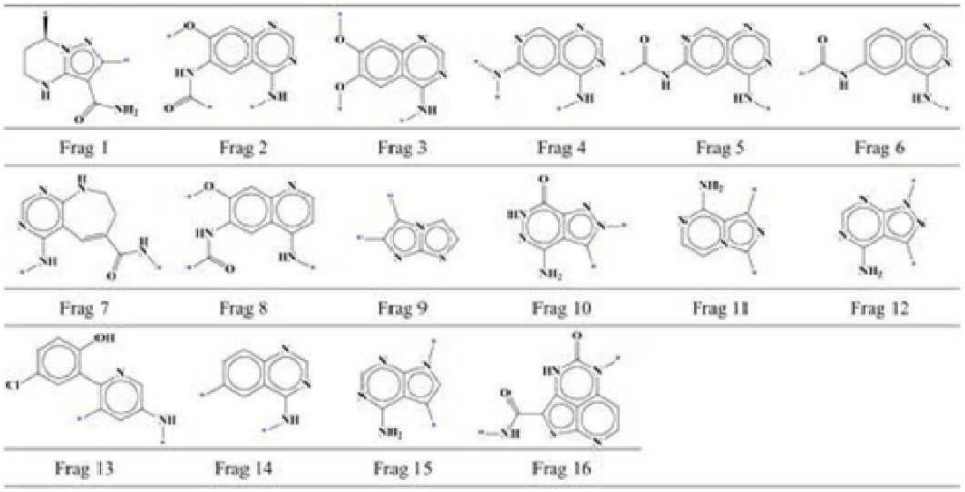


도면13a



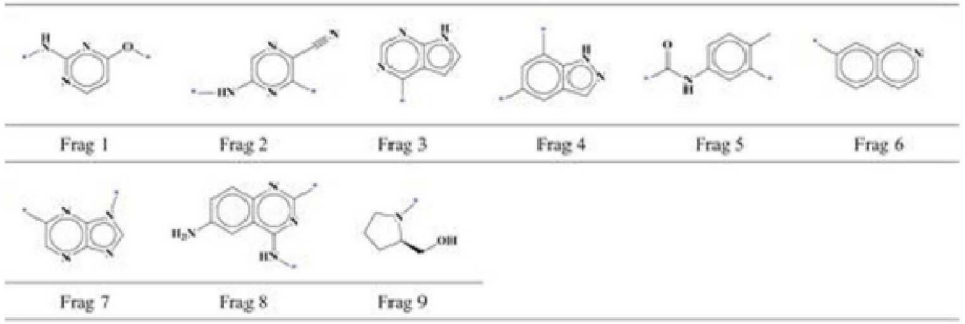
도면13b

활성값을 나타내는 분자 조각



도면13c

비활성값을 나타내는 분자 조각



도면14

ChEMBL 데이터베이스 화합물	CDK2	CHK1	IGF1R	ERBB4
전체 화합물 수	1016	986	1126	157
활성 화합물 수 ($IC_{50} \leq 1,000nM$)	571	743	795	119
비활성 화합물 수 ($IC_{50} \geq 10,000nM$)	445	243	331	38
Site I에 분자 조각이 위치하는 화합물 수	228	356	280	81
TWN-FS으로 스크리닝된 화합물 수	37	175	175	77
스크리닝된 화합물의 분자 조각 유형	20	49	58	25
활성 화합물의 분자 조각 유형	16	27	31	16
비활성 화합물의 분자 조각 유형	4	22	27	9
활성과 비활성 화합물의 중복되는 분자 조각 유형	0	2	1	0