



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년10월12일
(11) 등록번호 10-1902150
(24) 등록일자 2018년09월18일

(51) 국제특허분류(Int. Cl.)
G06F 19/12 (2011.01) G06F 19/00 (2018.01)
G06F 19/16 (2011.01) G06F 19/28 (2011.01)
(52) CPC특허분류
G06F 19/12 (2013.01)
G06F 19/16 (2013.01)
(21) 출원번호 10-2018-0029647
(22) 출원일자 2018년03월14일
심사청구일자 2018년03월14일
(56) 선행기술조사문헌
J. Lee 외 1인, "Water polygons in
high-resolution protein crystal structures",
Protein Science, 18권, pp.1370-1376,
2009.05.13.
KR1020160128288 A

(73) 특허권자
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
주식회사 유클리트소프트
대전광역시 서구 대덕대로317번길 20 (월평동선
사엔조이빌딩458호)
(72) 발명자
강남숙
대전광역시 유성구 대학로 99 충남대학교 산학연
별관 419
채은경
대전광역시 서구 청사로 70, 115-205
(뒷면에 계속)
(74) 대리인
김상욱

전체 청구항 수 : 총 3 항

심사관 : 성경아

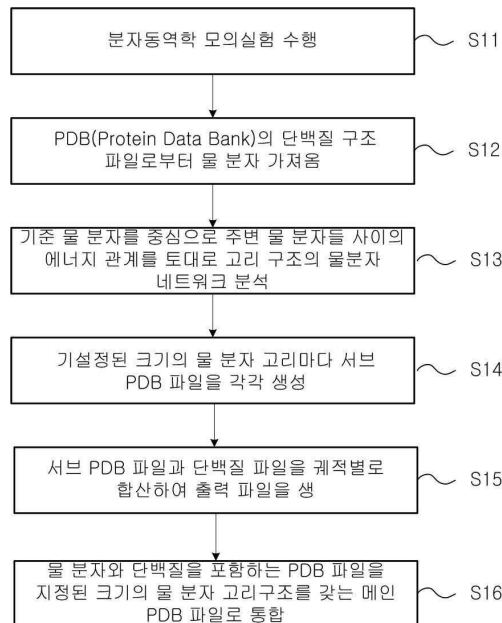
(54) 발명의 명칭 물 분자 네트워크 분석 방법

(57) 요약

본 발명은 물 분자 네트워크 분석 방법에 관한 것으로서, 분자동역학(Molecular Dynamics, MD) 모의실험 소프트웨어 패키지를 실행하여 분석 결과를 제공하는 분석 서버 및 분자동역학 모의 실험에 필요한 데이터를 관리하는 데이터베이스를 포함하는 물 분자 네트워크 분석 시스템에 의해 수행되는 물 분자 네트워크 분석 방법에 있어서,

(뒷면에 계속)

대표도 - 도2



웹 기반의 플랫폼 UI(User Interface)를 통해 명령어를 입력하여 상기 분자동역학 모의실험을 실행하는 제1 단계; 분자동역학 모의실험이 수행된 PDB(Protein Data Bank)의 단백질 구조 파일(trajecory file)로부터 물 분자를 가져오는 제2 단계; 상기 물 분자들 중 기준 물 분자를 중심으로 주변 물 분자들 사이의 에너지 관계를 토대로 고리 구조(Water ring)의 물 분자 네트워크를 분석하는 제3 단계; 기설정된 크기의 물 분자 고리마다 서버 PDB 파일을 각각 생성하는 제4 단계; 상기 서버 PDB 파일과 단백질 파일을 궤적별로 합산하여 출력 파일을 생성하는 제5 단계; 및 상기 궤적별로 고리 구조를 형성하는 물 분자와 단백질을 포함하는 PDB 파일을 지정된 크기의 물 분자 고리구조를 갖는 메인 PDB 파일로 통합하는 제 6단계를 포함한다. 따라서, 본 발명은 에너지 기반으로 물 분자 고리를 계산하거나 거리 기반으로 물 분자 고리를 계산하여 단백질 결합 자리 내의 물 분자 움직임을 분석할 수 있기 때문에 단백질-리간드 결합시 물 분자가 두 물질의 결합력과 결합 양식에 중요한 역할을 하고 있으므로 유사한 결합 자리를 가지는 서로 다른 단백질에 대한 선택적인 리간드 설계에 중용한 정보를 제공할 수 있고, 웹 기반의 플랫폼 UI를 통해 프로그래밍 경험이 부족한 신약개발 연구자들이 직관적이 UI를 통해 분석에 대한 접근성을 높이고, 실험실 외부나 원격지에서 분석 서버의 분석 진행 상태를 웹 또는 모바일에서 제어 또는 확인할 수 있으며, 분석 결과와 단계별로 생성되는 결과 데이터(OUTPUT DATA)를 웹 리포팅 형식으로 제공하여 연구의 효율성을 높일 수 있다.

(52) CPC특허분류

G06F 19/28 (2013.01)

G06F 19/701 (2013.01)

(72) 발명자

이명휘

대전광역시 서구 정림서로 162-18 강변들보람아파트 103동 704호

최광은

대전광역시 유성구 궁동 494-2 시티빌 204호

명세서

청구범위

청구항 1

분자동역학(Molecular Dynamics, MD) 모의실험 소프트웨어 패키지를 실행하여 분석 결과를 제공하는 분석 서버 및 분자동역학 모의 실험에 필요한 데이터를 관리하는 데이터베이스를 포함하는 물 분자 네트워크 분석 시스템에 의해 수행되는 물 분자 네트워크 분석 방법에 있어서,

웹 기반의 플랫폼 UI(User Interface)를 통해 명령어를 입력하여 PDB(Protein Data Bank)의 단백질 구조 파일(trajjectory file)로부터 물 분자와 헤테로 원자(hetero atom)를 선택하여 새로운 PDB 파일을 생성하는 제1 단계; 상기 물 분자들 중 기준 물 분자를 중심으로 산소 원자끼리의 거리 기준으로 주변 물 분자들 사이의 고리 구조(Water ring)의 물 분자 네트워크를 분석하기 위하여 하기한 포텐셜 함수를 이용하는 제2 단계; 단백질과 물이 결합된 상태에서 물 분자만 분리하기 위해 기설정된 크기의 물 분자 고리마다 서브 PDB 파일을 각각 생성하는 제3 단계; 상기 서브 PDB 파일과 단백질 파일을 궤적별로 합산하여 서브 pdb 파일별로 PDB 파일 형태의 출력 파일을 생성하는 제4 단계; 및 상기 궤적별로 고리 구조를 형성하는 물 분자와 단백질을 포함하는 PDB 파일 형태의 각 출력 파일들을 하나의 메인 PDB 파일로 통합하고, 상기 물 분자 고리에서 최단거리의 단백질 원자들을 식별하여 단백질을 구성하는 각 아미노산 또는 원자들의 물 분자 고리 패턴을 분석하는 제5 단계를 포함하되,

$$\epsilon_{mn} = \sum_i^{on\ m} \sum_j^{on\ n} \frac{q_i q_j e^2}{r_{ij}} + \frac{A}{r_{oo}^{12}} - \frac{C}{r_{oo}^6}$$

상기 포텐셜 함수는

이고, 상기 포텐셜 함수에서 q i

는 전하 값, A는 4εσ¹², C는 4εσ⁶인 것을 특징으로 하는 물 분자 네트워크 분석 방법.

청구항 2

제1항에 있어서,

사용자가 상기 웹 기반의 플랫폼 UI(User Interface)를 통해 상기 분석 서버에서 실행되는 상기 물 분자 네트워크의 분석 진행 상태를 웹 또는 모바일 상에서 실시간으로 제어 및 확인하고, 상기 물 분자 네트워크의 분석 결과를 웹 리포팅 형식으로 제공하며, 사용자별로 입력 데이터(INPUT DATA)를 관리하여, 기 분석된 데이터(INPUT DATA)를 재활용하는 것을 특징으로 하는 물 분자 네트워크 분석 방법.

청구항 3

제1항에 있어서,

상기 분석 서버는 데이터베이스와 연계하여 상기 물 분자 네트워크의 분석 데이터와 설정 값(OPTIONS)의 이력 관리를 수행하고, 3차원 단백질 구조를 가시화 도구 라이브러리를 이용하여 웹 리포팅 형식으로 제공하는 것을 특징으로 하는 물 분자 네트워크 분석 방법.

발명의 설명

기술 분야

본 발명은 물 분자 네트워크 분석 방법에 관한 것으로, 보다 상세하게는 단계별 명령어를 웹 기반의 플랫폼 UI를 통해 입력하는 방식으로 에너지 기반으로 물 분자 고리를 계산하거나 거리 기반으로 물 분자 고리를 계산하여 단백질 결합 자리 내의 물 분자 움직임을 분석할 수 있는 물 분자 네트워크 분석 방법에 관한 것이다.

[0001]

배경 기술

- [0002] 신약 개발은 초기의 질병표적물질 결정 및 후보물질 발굴을 위해 화학 기술, 생명과학 기술, IT 기반의 컴퓨팅 기술을 동시 복합적으로 사용하거나 상호 교차하여 사용함으로써 연구 개발에 소요되는 시간과 경비를 최소화하고 있으며, 점차 더 효율적인 방법을 모색하며 발전해가고 있다.
- [0003] 신약 개발을 위해 1980년대 후반부터 계산화학이 발달하면서 컴퓨터를 도구로 적극 활용하고 있다. 신약개발의 가장 첫 단계인 신약의 초기 유효 물질(Hit compounds)과 선도물질(Lead compounds)을 탐색하고 최적화하는 과정은 컴퓨터 활용으로 인해 이전보다 빠르고 효율적으로 진행되고 있으며, 이와 관련한 다양한 자료를 정보화, 지식화, 시각화하는 과정이 신약 연구에 있어 반드시 필요한 과정이 되고 있다.
- [0004] 컴퓨터를 이용한 신약설계(Computer-Aided Drug Design, CADD)는 신약개발에 드는 비용을 절감하고 연구개발 시간을 단축시킬 수 있도록 한다. 이러한 인실리코(in silico) 접근법은 리간드 기반 신약설계(Ligand-based drug design)와 구조기반 신약설계(Structure-based drug design)로 구분된다.
- [0005] 초기에는 2차원, 3차원 QSAR(Quantitative Structure Activity Relationships)이 연구 분야의 주종을 이루고 있지만, 최근에는 더욱 강력하고 다양한 소프트웨어와 대용량을 처리할 수 있는 데이터베이스의 개발로 인해 신약개발에 있어서 컴퓨터를 이용한 신약설계(CADD)는 없어서는 안 되는 중요한 연구 분야로 자리 잡게 되었다.
- [0006] 신약개발에 있어서 비용을 절감하고, 임상시험에서 실패율을 낮추면서도 획기적인 치료 약물을 개발하기 위해서는 다양한 방법을 발굴해내야 한다. 그중, 컴퓨터 기반의 분자설계기술은 신약개발을 위한 표적 물질들과 표적 물질에 작용하는 약물 후보군 화합물을 찾아내는데 매우 효율적인 방법으로서, 시간과 비용을 절감할 수 있으며, 임상 시험에서의 위험을 최소화하기 위해 독성이나 부작용을 예상하는데 도움이 된다. 또한, 컴퓨터 기반의 분자설계기술은 연구 단계 초기에 가상 탐색 방법을 통해 효율적이고, 신속하게 초기 유효물질을 도출할 수가 있다.
- [0007] 그러나, 기존의 CADD(Computer-Aided Drug Discovery)를 이용한 단백질-리간드와의 상호작용 분석은 SAR(구조-활성 상관관계)가 정확히 적용되지 않는다는 문제점이 있다. 이는 표적 단백질을 포함한 분자들의 작용은 공기 중이 아닌 용매(물)가 있는 곳에서 이루어지며 열역학적 에너지 변화가 일어난다. 실제 실험으로는 ITC, 등온적정열량법으로 연구되고 있다. 그래서, 신약개발 과정 중, 질병에 대한 작용점을 선정하는 단계에서 용매(solvent)인 물에서 단백질(질병표적물질)과 리간드(단백질에 작용하는 저분자 화합물)의 상호작용 특성을 정확하게 분석하는 것은 매우 중요하다. 리간드 기반 분자 설계 기술(LBDD, Ligand-Based Drug Design)은 약물 후보 화합물인 리간드가 물이 존재하는 조건에서 표적 단백질(target protein)에 결합하여 적응증(질병 치료), 즉 약물의 활성을 나타내기 때문이다.
- [0008] 최근 용매(solvent)인 물이 단백질과 리간드의 상호작용에 중요하게 관여한다는 사실이 밝혀졌고, 이를 바탕으로 물 분자 분석 정보를 활용해 합리적 신약설계를 할 수 있는 다양한 방법이 제시되고 있다.
- [0009] 물은 모든 생체의 구조와 기능을 결정하는데 매우 중요한 역할을 한다. 세포 대부분 물 분자로 구성되어 있으며, 물 분자는 촉매 작용(catalysis), 단백질 접힘(protein folding) 및 생체 분자 결합의 핵심 요소로 관여한다. 물을 통한 분자동역학 모의실험은 물의 존재에서 생체 분자의 특성을 연구하기 위해 오래 동안 사용되어 왔다.
- [0010] 물은 단백질(생체 분자)과 리간드(저분자 화합물, 약물) 상호작용에서 소수성 효과에 영향을 주며 중요한 역할을 한다. 최근 결합 부위 주변 물 분자의 열역학적 특성을 효율적으로 계산할 수 있는 방법이 고안되었다.
- [0011] 기존에 알려진 물 관련 연구 사례는 Fernandez 등에 의한 'de-hydron' 이론을 통해 단백질 주위의 물 분자 구조를 분석하는 방법과 미국 Schrodinger사에서 개발한 소프트웨어 WaterMap이 있다. WaterMap은 물 분자 분석연구를 통해 키나아제의 선택성을 분석하는 또 다른 방법이다.
- [0012] Fernandez 등에 의한 dehydron 이론은 단백질 주위의 물 분자 구조를 분석함으로써 단백질들 간의 3차원적 구조 차이를 분석한다. 단백질 내의 근접한 골격(backbone) 사이의 수소결합이 단백질 구조 안정화에 중요한 역할을 하는데, 이러한 수소결합이 용매에 노출될 경우 단백질 구조의 불안정화를 야기하게 된다. 이러한 이유로 단백질은 골격 수소 결합 주위로 소수성 아미노산 곁사슬이 위치하여 수소결합을 물의 공격으로부터 보호하도록 진화되어 있다. 하지만 어떤 이유에 의해서 소수성 아미노산 곁사슬에 의해서 물의 공격을 제대로 막지 못 하는 불완전한 접힘의 구조적 모티프가 존재할 수 있는데, 이를 'de-hydron'이라고 명명하였다.

- [0013] 분자동력학적 모의 실험방법으로 서로 다른 키나아제 사이에서 'de-hydron' 위치에 차이가 있음을 발견하였고, 이러한 차이를 이용하여 단백질의 선택성을 분석하였다.
- [0014] WaterMap은 분자동역학 모의실험 방법과 통계 열역학적 분석을 통해 단백질 결합자리(binding site) 내에 존재하는 물 분자의 자유에너지를 계산하여 여러 단백질들에 대한 리간드의 활성을 예측 할 수 있게 해주는 소프트웨어이다.
- [0015] 위의 사례들을 통해 분자동력학적 모의실험 방법을 기반으로 한 물 분자의 분석이 신약개발 과정 중 단백질의 구조 분석에 중요한 지표가 될 수 있음을 알 수 있다.
- [0016] 분자동역학 모의실험(Molecular Dynamics Simulation)은 물리화학 법칙을 사용한 계산을 통해 원자나 단백질과 같은 분자의 실시간 움직임을 자세하게 관찰하거나 예측하는 기법으로서, 실험과 유사한 온도, 압력 등의 조건을 반영하며, 확률에 의존하여 이론적으로 컴퓨터 모의실험을 하는 연구 방법이다.
- [0017] 일반적으로 분자동역학 모의실험을 하기 위해서는 빠르게 연산처리가 가능한 고성능 초고속 컴퓨터가 사용된다. 연구자들은 서버에 설치된 분자동역학(MD) 소프트웨어를 구동시키는 과정을 수행하는데, 익숙하지 않은 UNIX 명령어를 접해야 하고, 서버 환경 설정시 돌발적으로 발생하는 오류를 처리하기 위한 불필요한 시간을 소비해야 한다. 게다가, 연구자들이 분자동역학 모의실험의 분석 결과와 이력을 데이터베이스를 통해 효과적으로 관리하기 힘들 뿐만 아니라, 다양한 분석 방법을 시도해보기 어렵다는 문제점이 있다.
- [0018] 게다가, java와 R script를 활용해 구현된 물 네트워크 분석 알고리즘은 일반 연구자들이 사용하기에는 매우 까다롭고, 다양한 변수 선언, 파라미터 설정과 build 및 실행 과정은 만만치 않다는 문제점이 있다.

선행기술문헌

특허문헌

- [0019] (특허문헌 0001) 한국등록특허 제10-0757040호 " 단백질 도메인을 기반으로 한 단백질 간 또는 단백질과화합물 간의 상호작용면 분석 시스템 및 방법과 이를 위한기록매체 "

발명의 내용

해결하려는 과제

- [0020] 본 발명은 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 있는 물 분자 네트워크 분석 방법을 제공한다.

과제의 해결 수단

- [0021] 실시예들 중에서, 물 분자 네트워크 분석 방법은, 분자동역학(Molecular Dynamics, MD) 모의실험 소프트웨어 패키지를 실행하여 분석 결과를 제공하는 분석 서버 및 분자동역학 모의 실험에 필요한 데이터를 관리하는 데이터 베이스를 포함하는 물 분자 네트워크 분석 시스템에 의해 수행되는 물 분자 네트워크 분석 방법에 있어서, 웹 기반의 플랫폼 UI(User Interface)를 통해 명령어를 입력하여 상기 분자동역학 모의실험을 실행하는 제1 단계; 분자동역학 모의실험이 수행된 PDB(Protein Data Bank)의 단백질 구조 파일(trajecory file)로부터 물 분자를 가져오는 제2 단계; 상기 물 분자들 중 기준 물 분자를 중심으로 주변 물 분자들 사이의 에너지 관계를 토대로 고리 구조(Water ring)의 물 분자 네트워크를 분석하는 제3 단계; 기설정된 크기의 물 분자 고리마다 서브 PDB 파일을 각각 생성하는 제4 단계; 상기 서브 PDB 파일과 단백질 파일을 캐적별로 합산하여 출력 파일을 생성하는 제5 단계; 및 상기 캐적별로 고리 구조를 형성하는 물 분자와 단백질을 포함하는 PDB 파일을 지정된 크기의 물 분자 고리구조를 갖는 메인 PDB 파일로 통합하는 제 6단계를 포함하는 것을 특징으로 한다.
- [0022] 상기 제3 단계는, 상기 물 분자 사이의 에너지 값이 -2.25 kcal/mol 를 기준으로 비공유결합-반데르발스 힘과 정전기적 상호작용을 이용하여 물 분자 고리를 계산하는 것을 특징으로 한다.
- [0023] 상기 물 분자들 사이의 상호 작용을 분석하기 위해 하기한 포텐셜 함수를 이용하는 것을 특징으로 한다.
- [0024] 상기 제2 단계는, 상기 단백질 구조 파일로부터 결합 부위(binding site)를 중심으로 주변 물 분자만 추출하고,

상기 결합 부위 중심의 아미노산을 기준 아미노산으로 선정한 후 상기 선정된 결합 부위 중심으로부터 추출된 좌표값(coordinates)과 기준 아미노산 중심으로 일정 거리 이내의 물 분자들을 추출하는 것을 특징으로 한다.

[0025] 상기 제3 단계는, 상기 물 분자 네트워크를 고리 구조의 위치와 수로 분석하고, 상기 물 분자 고리를 구성하는 물 분자의 개수에 따라 삼각수, 사각수, 오각수로 분류하여 계산하는 것을 특징으로 한다.

[0026] 실시예들 중에서, 물 분자 네트워크 분석 방법은, 분자동역학(Molecular Dynamics, MD) 모의실험 소프트웨어 패키지를 실행하여 분석 결과를 제공하는 분석 서버 및 분자동역학 모의 실험에 필요한 데이터를 관리하는 데이터 베이스를 포함하는 물 분자 네트워크 분석 시스템에 의해 수행되는 물 분자 네트워크 분석 방법에 있어서, 웹 기반의 플랫폼 UI(User Interface)를 통해 명령어를 입력하여 PDB(Protein Data Bank)의 단백질 구조 파일(trajjectory file)로부터 물 분자와 헤테로 원자(hetero atom)를 선택하여 새로운 PDB 파일을 생성하는 제1 단계; 상기 물 분자들 중 기준 물 분자를 중심으로 산소 원자끼리의 거리 기준으로 주변 물 분자들 사이의 고리 구조(Water ring)의 물 분자 네트워크를 분석하는 제2 단계; 기설정된 크기의 물 분자 고리마다 서브 PDB 파일을 각각 생성하는 제3 단계; 상기 서브 PDB 파일과 단백질 파일을 궤적별로 합산하여 출력 파일을 생성하는 제4 단계; 및 상기 물 분자 고리에서 최단거리의 단백질 원자들을 식별하여 단백질을 구성하는 각 아미노산 또는 원자들의 물 분자 고리 패턴을 분석하는 제5 단계를 포함하는 것을 특징으로 한다.

[0027] 한편, 물 분자 네트워크 분석 방법은 사용자가 상기 웹 기반의 플랫폼 UI(User Interface)를 통해 상기 분석 서버에서 실행되는 상기 물 분자 네트워크의 분석 진행 상태를 웹 또는 모바일 상에서 실시간으로 제어 및 확인하고, 상기 물 분자 네트워크의 분석 결과를 웹 리포팅 형식으로 제공하며, 사용자별로 입력 데이터(INPUT DATA)를 관리하여, 기 분석된 데이터(INPUT DATA)를 재활용하는 것을 특징으로 한다.

[0028] 물 분자 네트워크 분석 방법은, 상기 분석 서버는 데이터베이스와 연계하여 상기 물 분자 네트워크의 분석 데이터와 설정 값(OPTIONS)의 이력 관리를 수행하고, 3차원 단백질 구조를 가시화 도구 라이브러리를 이용하여 웹 리포팅 형식으로 제공하는 것을 특징으로 한다.

발명의 효과

[0029] 본 발명은 에너지 기반으로 물 분자 고리를 계산하거나 거리 기반으로 물 분자 고리를 계산하여 단백질 결합 자리 내의 물 분자 움직임을 분석할 수 있기 때문에 단백질-리간드 결합시 물 분자가 두 물질의 결합력과 결합 양식에 중요한 역할을 하고 있으므로 유사한 결합 자리를 가지는 서로 다른 단백질에 대한 선택적인 리간드 설계에 중용한 정보를 제공할 수 있는 효과가 있다.

[0030] 또한, 본 발명은 웹 기반의 플랫폼 UI를 통해 프로그래밍 경험이 부족한 신약개발 연구자들이 직관적이 UI를 통해 분석에 대한 접근성을 높이고, 실험실 외부나 원격지에서 분석 서버의 분석 진행 상태를 웹 또는 모바일에서 제어 또는 확인할 수 있으며, 분석 결과와 단계별로 생성되는 결과 데이터(OUTPUT DATA)를 웹 리포팅 형식으로 제공하여 연구의 효율성을 높일 수 있는 효과가 있다.

[0031] 본 발명은 신약개발 과정 중, 가상 탐색 단계에서 매뉴얼화되지 않은 과정을 조금 더 효율적이고 신속하게 처리하게 하기 위해 일련의 공정을 웹을 통해서 쉽게 수행할 수 있으며, 분석 결과와 이력을 RDBMS로 연동하여 관리할 수 있으므로 대량으로 누적된 연구결과의 활용도를 높일 수 있는 효과가 있다.

도면의 간단한 설명

[0032] 도 1은 본 발명의 일 실시예에 따른 물 분자 네트워크 분석 시스템을 설명하는 도면이다.

도 2는 본 발명의 제1 실시예에 따른 물 분자 네트워크 분석 방법을 설명하는 순서도이다.

도 3은 도 2의 물 분자 고리에 대한 다양한 구조를 설명하는 도면이다.

도 4는 본 발명의 제2 실시예에 따른 물 분자 네트워크 분석 방법을 설명하는 순서도이다.

도 5는 본 발명의 일 실시예에 따른 물 분자 네트워크 분석 방법의 인터페이스 화면을 설명하는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0033] 본 발명에 기재된 실시예 및 도면에 도시된 구성은 본 발명의 바람직한 실시예에 불과할 뿐이고, 본 발명의 기술적 사상을 모두 표현하는 것은 아니므로, 본 발명의 권리범위는 본문에 설명된 실시예 및 도면에 의하여 제한되는 것으로 해석되어서는 아니 된다. 즉, 실시예는 다양한 변경이 가능하고 여러 가지 형태를 가질 수 있으며

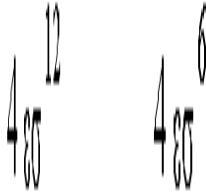
로 본 발명의 권리범위는 기술적 사상을 실현할 수 있는 균등물들을 포함하는 것으로 이해되어야 한다. 또한, 본 발명에서 제시된 목적 또는 효과는 특정 실시예가 이를 전부 포함하여야 한다거나 그러한 효과만을 포함하여야 한다는 의미는 아니므로, 본 발명의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.

- [0034] 여기서 사용되는 모든 용어들은 다르게 정의되지 않는 한, 본 발명이 속하는 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 것으로 해석되어야 하며, 본 발명에서 명백하게 정의하지 않는 이상 적이거나 과도하게 형식적인 의미를 지니는 것으로 해석될 수 없다.
- [0035] 도 1은 본 발명의 일 실시예에 따른 물 분자 네트워크 분석 시스템을 설명하는 도면이다.
- [0036] 도 1을 참고하면, 물 분자 네트워크 분석 시스템은 분석 서버 () 및 데이터베이스(110)를 포함한다.
- [0037] 분석 서버(100)는 분자동역학(Molecular Dynamics, MD) 모의실험 소프트웨어 패키지를 실행하여 분석 결과를 제공하고, 데이터베이스(110)는 분자동역학 모의 실험에 필요한 데이터를 관리한다.
- [0038] 사용자는 로컬 PC나 모바일 단말 등을 이용하여 웹 기반의 플랫폼 UI(User Interface)를 통해 분석 서버(100)에 설치된 분자동역학 모의실험 소프트웨어 패키지에 접근하여 제어할 수 있고, 물 분자 네트워크 분석과 관련된 데이터들을 관계형 데이터베이스(110)와 연동하여 관리할 수 있도록 한다.
- [0039] 이때, 분자동역학 모의실험 소프트웨어 패키지는 CHARMM(Chemistry at HARvard Macromolecular Mechanics), AMBER(Assisted Model Buildingwith Energy Refinement), NAMD, Discover, GROMACS(GRONingen MACHINE for Chemical Simulations) 등이 있다.
- [0040] 이 중에서 GROMACS 패키지는 GNU 정신(정보공유정신, 오픈소스 운동)에 입각하여 누구나 다운받을 수 있고, 무료로 사용할 수 있으며, 소스코드가 공개되어 있는 프리웨어이기 때문에 자기에게 맞는 기능을 추가해 넣을 수 있고, 조작이 비교적 쉽고, 계산 속도가 빠르다는 장점이 있다.
- [0041] 분석 서버(100)는 분자동역학 모의실험 소프트웨어 패키지로 계산 속도도 빠르고, 다양한 분석이 가능한 GROMACS 패키지를 선택하여 웹 기반의 플랫폼을 구현한다.
- [0042] 사용자는 자신이 실행할 새로운 프로젝트를 생성하고, INPUT DATA(*.pdb)를 업로드 한 후, 설정 값을 선택하고 job을 실행시키면 분석된 결과로 OUTPUT DATA를 얻게 되고, 플랫폼에서 OUTPUT DATA를 웹 리포팅 형식으로 제공할 수 있다.
- [0043] 또한, 사용자는 'MY DATA' 를 통해 INPUT DATA를 관리할 수 있기 때문에 동일한 데이터에 여러 설정 값을 적용해 job을 실행할 수 있고, 분석된 실험 데이터(INPUT DATA)를 다시 재활용할 수 있으며, 모의실험에 사용되는 UNIX 명령어는 python script로 연계하여 반복적인 데이터 분석 과정을 자동화할 수 있다.
- [0044] 또한, 분석 서버(100)는 python script를 활용해서 관계형 데이터베이스(110) 연동 처리하여 job status의 실시간 업데이트가 가능하여, 사용자가 서버의 상태를 모니터링 할 수 있도록 하고, 분석 데이터(INPUT DATA) 정보와 분석 시, 설정 값(OPTIONS)의 이력 관리가 용이하도록 한다.
- [0045] 도 2는 본 발명의 제1 실시예에 따른 물 분자 네트워크 분석 방법을 설명하는 순서도이고, 도 3은 도 2의 물 분자 고리에 대한 다양한 구조를 설명하는 도면이다.
- [0046] 도 2 및 도 3을 참고하면, 에너지 기반의 물 분자 네트워크 분석 방법은, 에너지 값이 -2.25 kcal/mol을 기준으로 비공유결합-반데르발스 힘과 정전기적 상호작용을 활용해 물 분자 고리를 계산하는 방법이고, 모든 단계의 알고리즘은 java로 구현되어 있다. 또한, 물 분자 네트워크 분석 방법은 웹 기반의 플랫폼으로 변환하는 과정에서 각 단계에 대한 연계는 python script를 활용하고, 설계 프레임워크는 검증된 MVC 모델을 적용하며, 비즈니스 로직은 java로 구현하고, 관계형 데이터베이스는 MySQL을 활용한다.
- [0047] 에너지 기반의 물 분자 네트워크 분석 방법은 단백질 내의 물 분자 특성을 정량적으로 분석하기 위해 분자동역학 모의실험을 실행한다.(S11) 시스템관리자(admin)는 분자동역학 모의실험 소프트웨어 패키지를 실행할 경우, publicdata(reference data)가 필요하면 참조할 데이터를 미리 업로드 할 수 있고, 새로운 설정 값을 구성할 수 있다.
- [0048] 에너지 기반의 물 분자 네트워크 분석 방법은, 분자동역학 모의실험이 수행된 PDB(Protein Data Bank)의 단백질 구조 파일(trajjectory file)로부터 결합 부위(binding site)를 중심으로 주변 물 분자만 추출한다.(S12) 결합

부위 중심의 아미노산을 기준 아미노산으로 선정하고, 결합 부위 중심으로부터 추출된 좌표값(coordinates)과 기준 아미노산을 토대로 일정 거리내의 물 분자들을 추출한다.

[0049] 에너지 기반의 물 분자 네트워크 분석 방법은 상기에서 추출된 물 분자들을 이용하여 기준 물 분자를 지정하고, 기준 물 분자를 중심으로 주변의 물 분자들과의 거리를 계산한다.(S13) 이때, 물 분자 네트워크 분석 방법은 단백질 구조 파일 내의 물 분자들을 모두 돌아가면서 계산하고, 계산된 값으로 결합된 관계, 즉 고리 구조인지를 확인하여 물 분자들 사이의 물 분자 네트워크를 분석한다.

[0050] 에너지 기반의 물 분자 네트워크 분석 방법은 선택된 물 분자들 사이의 상호 작용을 분석하기 위해 하기한 수학적 식 1의 포텐셜 함수를 이용하고, 포텐셜 함수에서 q_i 는 전하 값, A와 C는 Lennard-Jones 위치 에너지 계산을 위



한 상수 값으로써 A는 , C는 이다.

수학적 식 1

$$\epsilon_{mn} = \sum_i^{on} \sum_j^{m} \frac{q_i q_j e^2}{r_{ij}} + \frac{A}{r_{oo}^{12}} - \frac{C}{r_{oo}^6}$$

[0051]

[0052] 에너지 기반의 물 분자 네트워크 분석 방법은 물 분자 네트워크를 고리 구조의 위치와 수로 분석하고, 물 분자 고리를 구성하는 물 분자의 개수에 따라 삼각수, 사각수, 오각수로 분류하여 계산한다.

[0053] 도 3에 도시된 바와 같이, 물 분자들이 일정 거리에 있다는 것은 서로 상호적 반응하여 연결되어 있다는 뜻이므로 고리 구조를 형성하고 있을 가능성이 매우 높다. 기준 물 분자를 시작으로 꼬리에 꼬리를 무는 형식으로 찾아가다가 원점으로 돌아오는 것을 찾는 방식으로 고리 구조를 정의한다.

[0054] 에너지 기반의 물 분자 네트워크 분석 방법은 CHARMM의 TIP3P parameter를 이용하여 계산된 물 분자들 사이의 에너지 관계를 기반으로 물 분자 고리 구조를 찾아간다. R3, R4, R5, R6에 해당하는 물 분자 일련번호를 정의하고, 그에 해당하는 평균 값을 생성한다.

[0055] 에너지 기반의 물 분자 네트워크 분석 방법은 고리구조를 형성하는 물분자들 중에서 특정 크기의 고리구조를 형성하는 물 분자를 각각 추출하고, 특정 크기의 물 분자 고리 구조만 포함하는 PDB 포맷(*.pdb)의 출력 파일을 서버 PDB 파일로 각각 생성한다. (S14) S14 단계는 단백질과 물이 결합된 상태에서 물 분자만 분리하는 과정이다.

[0056] 에너지 기반의 물 분자 네트워크 분석 방법은 각각 나눠져 있는 물 분자 고리 구조의 서버 PDB 파일과 단백질 파일을 제적별로 합산하여 출력 파일을 생성한다.(S15)

[0057] 에너지 기반의 물 분자 네트워크 분석 방법은 각 제적별로 고리 구조를 형성하는 물분자와 단백질이 같이 있는 PDB 파일을 특정 크기의 물 분자 고리구조를 갖는 하나의 메인 PDB 파일로 통합한다.(S16) 물 분자 고리 구조의 크기에 따라 각 1개의 PDB 파일이 형성된다.

[0058] 도 4는 본 발명의 제2 실시예에 따른 물 분자 네트워크 분석 방법을 설명하는 순서도이다.

[0059] 도 4를 참고하면, 거리 기반의 물 분자 네트워크 분석 방법은, 물 분자의 고리를 거리를 기반기준으로 계산해 나가는 방법으로서, S21, S23, S24의 알고리즘은 java로 구현되며, S22와 S25는 R script를 사용한다. 물 분자 네트워크 분석 방법은 웹 기반의 플랫폼으로 변환하는 과정에서 각 단계에 대한 연계를 python script를 활용하고, 설계 프레임워크는 검증된 MVC 모델을 적용하며, 비즈니스 로직은 java로 구현하고, 관계형 데이터베이스는 MySQL을 활용한다.

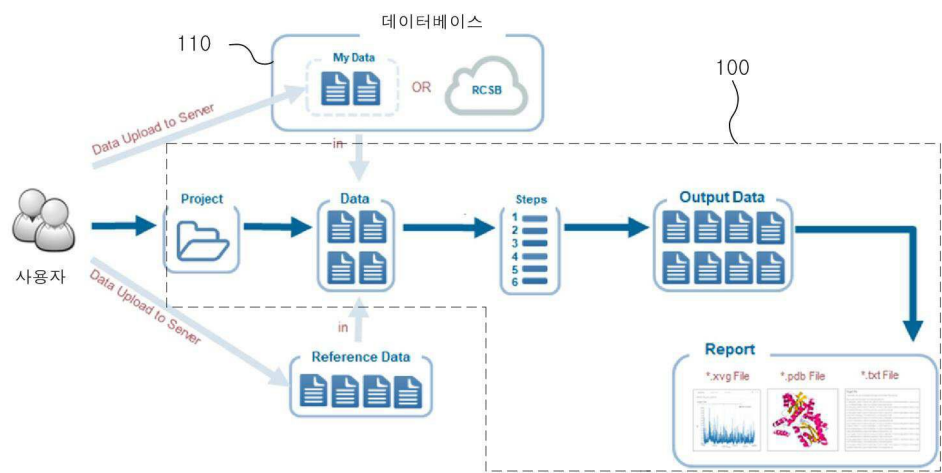
- [0060] 이러한 거리 기반의 물 분자 네트워크 분석 방법은 분자동역학 모의실험(GROMACS)의 실행 과정이 필요 없고, 물 분자를 포함하고 있는 PDB 포맷의 단백질 구조 파일을 데이터베이스(110)에서 제공받아 입력 파일로 사용한다.
- [0061] 거리 기반의 물 분자 네트워크 분석 방법은 PDB(Protein Data Bank)의 단백질 구조 파일(trajecory file)로부터 물 분자와 헤테로 원자(hetero atom)를 선택하여 새로운 PDB 파일을 생성한다.(S21)
- [0062] 거리 기반의 물 분자 네트워크 분석 방법은 물 분자들 사이의 물 분자 네트워크를 분석하는데, 산소 원자끼리의 거리 기준(대부분 수소 원자가 없는 크리스탈 구조)으로 한 계산식은 매트릭스 구조의 알고리즘을 구현하기 적합한 R script를 이용한다.(S22)
- [0063] 거리 기반의 물 분자 네트워크 분석 방법은 특정 크기의 물 분자 고리(water-ring) 마다 각각의 서브 PDB 파일 (*.pdb) 생성한다.(S23)
- [0064] 거리 기반의 물 분자 네트워크 분석 방법은 S23 단계에서 추출된 각 크기별 물 분자 고리 파일 (*.pdb)과 S21 단계에서 출력되는 단백질 파일을 하나의 메인 PDB 파일로 통합한다.(S24)
- [0065] 거리 기반의 물 분자 네트워크 분석 방법은 단백질을 구성하는 각 아미노산 또는 원자들의 물리화학적 성질에 따른 물 분자 고리 패턴을 분석하기 위해 물 분자 고리에서 최단 거리에 있는 가장 가까이 있는 protein의 원자(atom)들을 식별한다. (S25)
- [0066] 도 5는 본 발명의 일 실시예에 따른 물 분자 네트워크 분석 방법의 인터페이스 화면을 설명하는 도면이다.
- [0067] 본 발명의 실시예에 따른 물 분자 네트워크 분석 방법은 웹 기반으로 개발된 플랫폼 UI를 통해 분석 서버에 설치되어 있는 분자동역학 모의실험 소프트웨어 패키지(GROMACS)를 구동 및 제어할 수 있고, 세부 분석 단위마다 사용자들이 설정 값을 선택할 수 있으며, 해당 분석 결과를 웹으로 리포팅하여 사용자들간의 데이터 비교 분석을 용이하게 할 수 있다.
- [0068] 따라서, 사용자들은 외부에서 웹/모바일로 분석 서버의 분석 진행 상태(완료, 실패, 진행중, 일시정지, 대기중 표시)를 실시간 제어 또는 확인할 수 있고, 소프트웨어 환경 설정시 자주 발생하는 오류로 인해 소요되는 불필요한 시간과 UNIX 명령어 사용의 어려움 등 연구에 방해되는 여러 가지 요소들을 절감할 수 있다.
- [0069] 본 발명의 실시예에 따른 물 분자 네트워크 분석 방법은 모의실험 단계별로 생성된 결과 파일 중 단백질 구조 파일 (*.pdb)의 경우는 다운로드 받아 로컬 PC의 어플리케이션에서 확인하는 것이 아니라, 결과를 직관적으로 웹에서 바로 확인할 수 있고, 3차원 단백질 구조를 가시화 도구 라이브러리(jmol, jsmol)을 활용해 웹 리포팅 형식으로 확인할 수 있다.
- [0070] 도 5에 도시된 바와 같이, 사용자는 웹 기반으로 개발된 플랫폼 UI를 통해 프로젝트 명을 입력하고 우측의 [Search] 버튼을 누르면 입력한 값이 포함된 프로젝트명을 검색할 수 있고, 순번, 상태, 프로젝트명, 실행될 파이프라인, 프로젝트를 생성한 계정, 프로젝트 생성 일시, 프로젝트 관련 활성화 버튼 등을 이용하여 프로젝트 관리, 프로젝트 목록 조회 및 프로젝트 등록을 수행할 수 있으며, [Report] 버튼을 이용하여 레포트 화면으로 이동하여 단백질 구조 파일 (*.pbd), 에너지 파일 (*.xvg), 텍스트 파일 (*.txt)을 확인할 수 있다.
- [0071] 또한, 본 발명의 실시예에 따른 물 분자 네트워크 분석 방법은 분석시에 생성된 에너지 파일의 경우 (*.xvg)는 jquery로 구현하여 웹 차트 형식으로 리포팅되어 사용자들이 자료로 바로 활용할 수 있고, 분자동역학 모의실험 결과로 얻어진 trajectory data를 tree 형태로 하나하나 웹에서 확인할 수 있으며, 필요시에 사용자가 직접 결과 파일을 웹에서 로컬로 다운로드 받을 수 있다.
- [0072] 상기에서는 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

부호의 설명

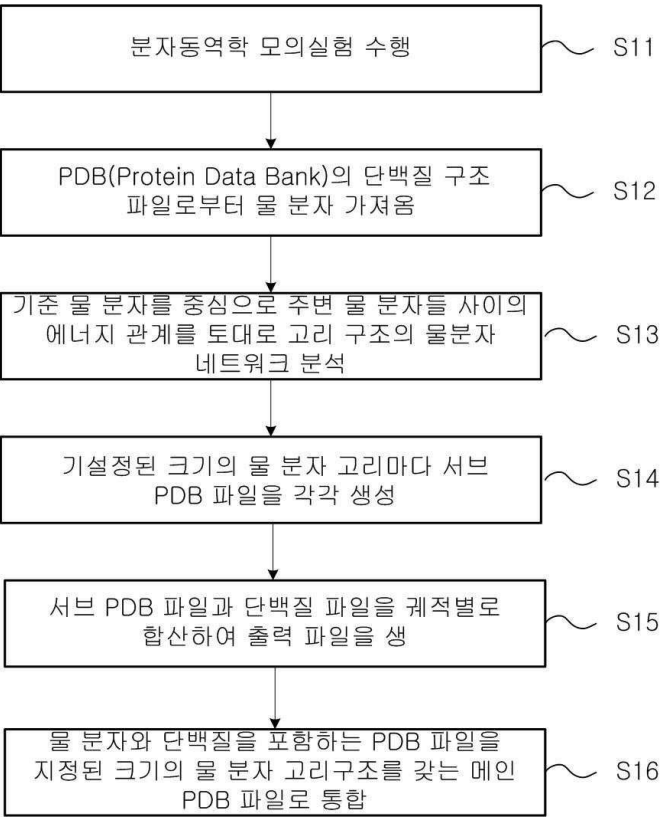
- [0073] 100 : 분석 서버 110 : 데이터베이스

도면

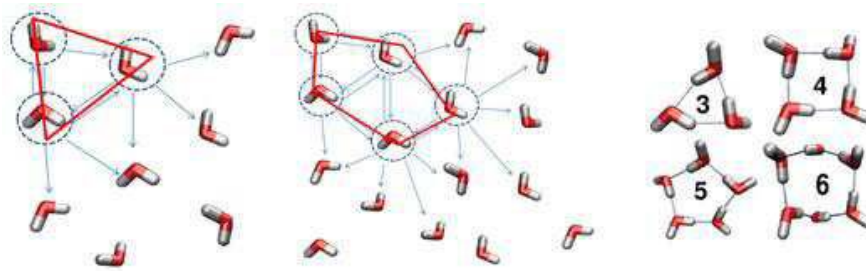
도면1



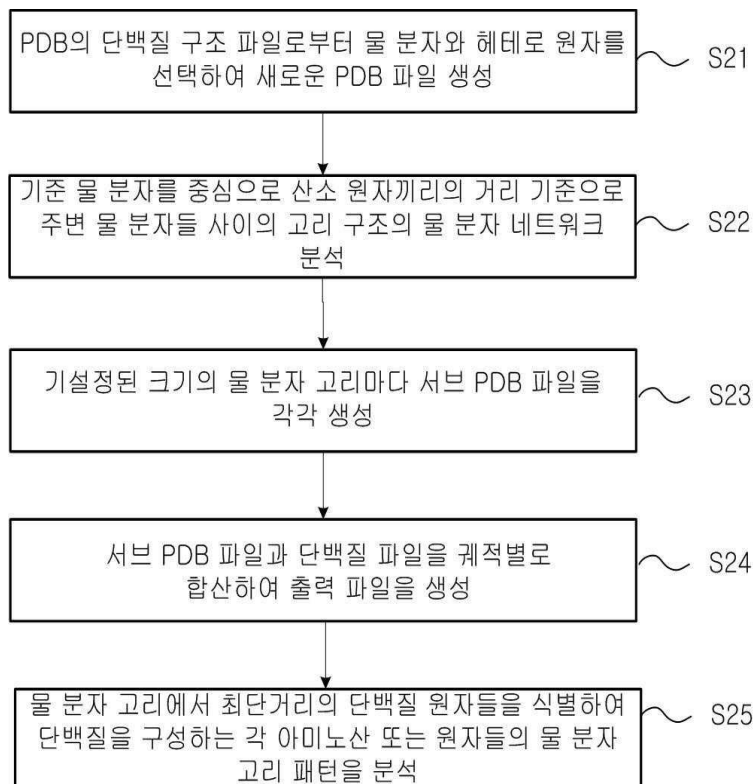
도면2



도면3



도면4



도면5

Protein Water Network Analysis

test Log out

Dashboard Projects Data Report new

Home / Projects

프로젝트명

Search Refresh + 신규

#	Status	Project	Pipeline	Created by	Created	Action
5	RUNNING	running project	OneClick	test	Aug 11, 2016 16:54:59	상세
4	COMPLETED	completed project	STEP	test	Aug 11, 2016 15:55:45	Report 상세
3	FAIL	fail project	OneClick	test	Aug 11, 2016 15:54:01	상세
2	WAITING	wating project		test	Aug 11, 2016 15:52:13	상세
1	STOP	stop project	OneClick	test	Aug 11, 2016 14:50:15	상세

1