



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2013년07월26일  
(11) 등록번호 10-1289948  
(24) 등록일자 2013년07월19일

(51) 국제특허분류(Int. Cl.)  
*G01N 33/50* (2006.01) *G06F 19/10* (2011.01)

(73) 특허권자  
인제대학교 산학협력단

(21) 출원번호 10-2012-0062277

경상남도 김해시 어방동 607 인제대학교 내

(22) 출원일자 2012년06월11일

(72) 발명자

심사청구일자 2012년06월11일

오석준

(56) 선행기술조사문헌

부산광역시 기장군 정관면 모전리 755 혼진에버빌  
103동 1103호

JP2005242493 A

(74) 대리인

JP2003083864 A

특허법인태백

JP2004170364 A

JP05322817 A

전체 청구항 수 : 총 10 항

심사관 : 박환수

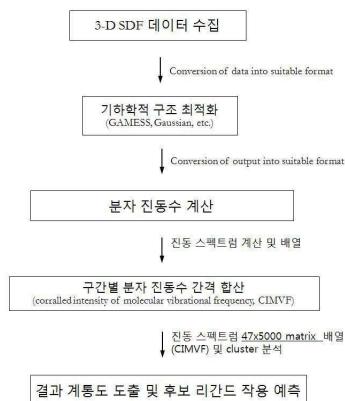
(54) 발명의 명칭 분자 진동수 패턴을 이용한 리간드 분류 방법 및 장치

### (57) 요 약

본 발명은 후보 리간드의 3차원 구조 데이터 정보를 획득하는 단계; 상기 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하는 단계; 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하는 단계; 및 상기 진동수를 이용하여 후보 리간드의 작용을 예측하는 단계를 포함하는 리간드 분류 방법에 관한 것이다.

본 발명의 리간드 분류 방법에 따르면, 리간드 분자 고유의 분자 진동수 패턴을 이용하여 리간드가 실제 수용체에 작용할 때 작용제(agonist)로 작용할지 길항제(antagonist)로 작용할지 여부를 간편하게 예측, 분석할 수 있어, 다양한 후보 리간드의 작용 양상을 실험을 거치지 않고 예측이 가능하다.

### 대 표 도 - 도1



## 특허청구의 범위

### 청구항 1

후보 리간드의 3차원 구조 데이터 정보를 획득하는 단계;

상기 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하는 단계;

변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하는 단계; 및

상기 진동수를 이용하여 후보 리간드의 작용을 예측하는 단계를 포함하는 리간드 분류 방법.

### 청구항 2

제1항에 있어서,

상기 후보 리간드의 작용을 예측하는 단계는, 후보 리간드 및 본 리간드 분자의 진동수의 유사도를 비교하여 계통도를 구하는 것인 리간드 분류 방법.

### 청구항 3

제1항에 있어서,

상기 후보 리간드의 작용을 예측하는 단계는,

후보 리간드의 문자 진동수를 일정한 간격으로 나누는 단계; 동일 간격 구간에 존재하는 문자 진동수의 강도를 합산하여 그 구간의 대표값으로 하는 단계; 및 상기 대표값의 패턴을 바탕으로 본 리간드와의 유사도를 비교하여 계통도를 구하는 단계를 포함하는 것인 리간드 분류 방법.

### 청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 리간드는 로돕신(rhodopsin) 유사 G-단백질-연결 수용체(G protein-coupled receptor, GPCR) 결합 리간드인 것인 리간드 분류 방법.

### 청구항 5

제4항에 있어서,

상기 로돕신(rhodopsin) 유사 G-단백질-연결 수용체(G protein-coupled receptor, GPCR)는 히스타민 수용체 또는 아데노신 수용체인 것인 리간드 분류 방법.

### 청구항 6

후보 리간드의 3차원 구조 데이터 정보를 입력받는 입력부;

상기 입력부로 입력된 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하는 제1연산부;

변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하는 제2연산부;

상기 진동수를 이용하여 후보 리간드의 작용을 예측하는 제3연산부; 및

상기 제3연산부를 통해 예측된 후보 리간드의 작용이 출력되는 출력부를 포함하는 리간드 분류 장치.

### 청구항 7

제6항에 있어서,

상기 제3연산부는 후보 리간드의 문자 진동수를 일정한 간격으로 나누는 단계; 동일 간격 구간에 존재하는 문자 진동수의 강도를 합산하여 그 구간의 대표값으로 하는 단계; 및 상기 대표값의 패턴을 바탕으로 본 리간드와의

유사도를 비교하여 계통도를 구하는 단계를 포함하여 후보 리간드의 작용을 예측하는 것인 리간드 분류 장치.

### 청구항 8

제6항 또는 제7항에 있어서,

상기 리간드는 로돕신(rhodopsin) 유사 G-단백질-연결 수용체(G protein-coupled receptor, GPCR) 결합 리간드인 것인 리간드 분류 장치.

### 청구항 9

제8항에 있어서,

상기 로돕신(rhodopsin) 유사 G-단백질-연결 수용체(G protein-coupled receptor, GPCR)는 히스타민 수용체 또는 아데노신 수용체인 것인 리간드 분류 장치.

### 청구항 10

컴퓨터 입력장치를 통해 후보 리간드의 3차원 구조 데이터 정보를 입력받는 단계;

컴퓨터 연산장치가 상기 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하도록 하는 단계;

컴퓨터 연산장치가 상기 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하도록 하는 단계;

컴퓨터의 연산장치가 상기 진동수를 이용하여 후보 리간드의 작용을 예측하게 하는 단계; 및

컴퓨터의 출력장치가 상기 연산장치에 의해 예측된 후보 리간드의 작용을 출력하게 하는 단계를 실행시키기 위한 리간드 분류 프로그램을 기록한 컴퓨터로 읽을 수 있는 매체.

## 명세서

### 기술 분야

[0001] 본 발명은 문자 진동수 패턴을 이용한 리간드 분류 방법, 분류 장치 및 리간드 분류 프로그램을 기록한 컴퓨터로 읽을 수 있는 매체에 관한 것이다.

### 배경 기술

[0002] G-단백질-연결 수용체(G protein-coupled receptor, GPCR)는 세포막 수용체를 구성하는 큰 그룹으로, 작동자의 자극에 의해 반응하여 G-단백질을 활성화시키고, 그 결과 고리형 AMP/단백질 키나아제 A 케스케이드와 같은 적어도 하나의 신호전달 케스케이드를 활성화시킨다. 이 커다란 수용체 집단은 효모로부터 사람에 이르기까지 유기체에서 널리 발견되며, 예컨대, 호르몬, 신경 및 후각 신호의 변환에 관여한다.

[0003] G-단백질-연결 수용체(G protein-coupled receptor, GPCR)는 시장 유통중인 인간 대상 의약품의 25% 이상의 수용체로 작용함이 밝혀져 있다. GPCR을 타겟으로 하는 의약품은 전체 의약업계 시장의 40%의 시장 점유율을 보이는 주요 제품군에 해당한다.

[0004] GPCR은 막관통 단백질로, 다양한 길이의 3개의 세포간 및 3개의 세포외 루프로 연결된 7개의 막 관통 부분(segment)으로 구성되어 있다. GPCR의 결합 리간드의 결정 구조는 리간드와 결합하는 포켓 및 세포외 루프와 관련이 있음이 밝혀져 있다(Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science*. 2007;318(5854):1258-65.; Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, et al. Structure of the human histamine H1 receptor complex with doxepin. *Nature*. 2011;475(7354):65-70). GPCR의 리간드는 수용체 구조를 활성형으로 변화시키는 기작을 통해 수용체를 활성화시킨다.

[0005] GPCR의 리간드-수용체 상호작용의 분자적 기전은 아직 알려져 있지 않으며, 현재 많은 연구가 수행되고 있다. 상기 리간드-수용체 상호작용 기작을 설명하기 위하여, 모양 이론, 결합 이론 및 진동 이론과 같은 다양한 이론

들이 제시되고 있다(Keller A, Vosshall LB. A psychophysical test of the vibration theory of olfaction. Nat Neurosci. 2004;7(4):337-8.; Takane SY, Mitchell JB. A structure-odour relationship study using EVA descriptors and hierarchical clustering. Org Biomol Chem. 2004;2(22):3250-5; 한국공개특허 10-2005-0029109호).

[0006] 그러나, 히스타민 수용체 또는 아데노신 수용체와 같은 GPCR에 작용하는 리간드는 구조가 매우 다양하여, 구조적 분석으로 리간드의 작용 효과를 예측하는데는 한계가 있다. 또한, 실제 *in vitro* 실험 등이 수행되는 경우, 분리, 여과 및 세척과 같은 다단계의 준비과정을 거쳐야 한다는 번거로움이 있다. 따라서, GPCR 리간드의 작용 효과를 예측할 수 있는 보다 신속하고 간편한 방법의 개발이 요구되고 있다.

[0007] 이에 따라 본 발명의 발명자는 보다 신속하고 간편하게 G-단백질-연결 수용체 리간드의 작용 효과를 예측하고 분류하고자 연구한 결과, 리간드 분자 고유의 분자 진동수 패턴의 분석만으로 리간드의 작용 여부를 간단히 예측할 수 있음을 확인하고 본 발명을 완성하게 되었다.

## 발명의 내용

### 해결하려는 과제

[0008] 따라서 본 발명은 분자 진동수 패턴을 이용한 리간드 분류 방법, 분류 장치 및 리간드 분류 프로그램을 기록한 컴퓨터로 읽을 수 있는 매체를 제공하는 것이다.

### 과제의 해결 수단

[0009] 상기 과제의 해결을 위해, 본 발명은

[0010] 후보 리간드의 3차원 구조 데이터 정보를 획득하는 단계; 상기 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하는 단계; 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하는 단계; 및 상기 진동수를 이용하여 후보 리간드의 작용을 예측하는 단계를 포함하는 리간드 분류 방법을 제공한다.

[0011] 또한 본 발명은 후보 리간드의 3차원 구조 데이터 정보를 입력받는 입력부; 상기 입력부로 입력된 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하는 제1연산부; 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하는 제2연산부; 상기 진동수를 이용하여 후보 리간드의 작용을 예측하는 제3연산부; 및 상기 제3연산부를 통해 예측된 후보 리간드의 작용이 출력되는 출력부를 포함하는 리간드 분류 장치를 제공한다.

[0012] 또한 본 발명은 컴퓨터 입력장치를 통해 후보 리간드의 3차원 구조 데이터 정보를 입력받는 단계; 컴퓨터 연산장치가 상기 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하도록 하는 단계; 컴퓨터 연산장치가 상기 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하도록 하는 단계; 컴퓨터의 연산장치가 상기 진동수를 이용하여 후보 리간드의 작용을 예측하게 하는 단계; 및 컴퓨터의 출력장치가 상기 연산장치에 의해 예측된 후보 리간드의 작용을 출력하게 하는 단계를 실행시키기 위한 리간드 분류 프로그램을 기록한 컴퓨터로 읽을 수 있는 매체를 제공한다.

## 발명의 효과

[0013] 본 발명의 리간드 분류 방법 및 분류 장치에 따르면, 리간드 분자 고유의 분자 진동수 패턴을 이용하여 리간드가 실제 수용체에 작용할 때 작용제로 작용할지 길항제로 작용할지 여부를 간편하게 예측, 분석할 수 있어, 다양한 후보 리간드의 작용 여부를 실험을 거치지 않고 예측이 가능하다.

### 도면의 간단한 설명

[0014] 도 1은 본 발명에 따른 리간드 분류 방법이 도시된 순서도이다.

도 2는 본 발명에 따른 리간드 분류 장치의 구성이 도시된 블록도이다.

도 3은 본 발명의 일실시예에 따라 얻어진 분자 진동수의 일 예이다.

도 4은 본 발명의 일실시예에 따른 히스타민 수용체 결합 리간드의 계통도를 나타낸다.

도 5는 본 발명의 일실시예에 따른 아데노신 수용체 결합 리간드의 계통도를 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

- [0015] 본 발명은 후보 리간드의 3차원 구조 데이터 정보를 획득하는 단계; 상기 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하는 단계; 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하는 단계; 및 상기 진동수를 이용하여 후보 리간드의 작용을 예측하는 단계를 포함하는 리간드 분류 방법을 제공한다.
- [0016] 본 발명의 한 구체예에서, 상기 후보 리간드의 작용을 예측하는 단계는, 후보 리간드 및 본 리간드 분자의 진동수의 유사도를 비교하여 계통도를 구하는 단계일 수 있다.
- [0017] 본 발명의 다른 구체예에서, 상기 후보 리간드의 작용을 예측하는 단계는, 후보 리간드의 분자 진동수를 일정한 간격으로 나누는 단계; 동일 간격 구간에 존재하는 분자 진동수의 강도를 합산하여 그 구간의 대표값으로 하는 단계; 및 상기 대표값의 패턴을 바탕으로 본 리간드와의 유사도를 비교하여 계통도를 구하는 단계를 포함할 수 있다.
- [0018] 본 발명의 또 다른 구체예에서, 상기 리간드는 로돕신(rhodopsin) 유사 G-단백질-연결 수용체(G protein-coupled receptor, GPCR) 결합 리간드일 수 있으며, 보다 구체적으로 상기 로돕신(rhodopsin) 유사 G-단백질-연결 수용체(G protein-coupled receptor, GPCR)는 히스타민 수용체 또는 아데노신 수용체일 수 있으나 이에 제한되는 것은 아니다.
- [0019] 본 발명에서, '후보 리간드'는 작용 효과를 알고자 하는 분석 대상 리간드이고, '본 리간드'는 본래 수용체에 결합하는 것으로 알려진 리간드를 의미한다. 본 발명의 일실시예에서, 수용체가 히스타민 수용체인 경우, 히스타민이 본 리간드가 되고, 수용체가 아데노신 수용체인 경우 본 리간드는 아데노신이 된다. 본 발명은 후보 리간드와 본 리간드의 진동수를 비교하여, 이의 유사성을 통해, 후보 리간드의 작용 효과를 예측하는 것이다. 본 리간드와의 진동수 패턴이 유사하게 나타나 계통도 상에서 인접하게 위치하는 경우, 후보 리간드는 작용제로 작용한다고 예측할 수 있으며, 반대의 경우 길항제로 작용한다고 예측할 수 있다.
- [0020] 도 1은 본 발명의 일실시예에 따른 리간드 분류 방법이 도시된 순서도이다.
- [0021] 이하 각 단계를 상세히 설명한다.
- [0022] 후보 리간드의 3차원 구조 데이터 정보를 획득하는 단계는, 후보 리간드에 대해 일반적인 정보를 제공하는 웹사이트 등에서 다운로드 하는 방식으로 수행될 수 있으며, 보다 구체적으로 NCBI의 PubChem Compound Database로부터 다운로드 받을 수 있으나, 이에 제한되는 것은 아니다.
- [0023] 상기 방법으로 얻어진 3차원 구조 데이터 정보는 최소 에너지 상태의 구조를 제공하는 것이 아니기 때문에, 이후 최소 에너지 상태로 후보 리간드의 구조를 변환하는 후보 리간드의 기하학적 최적화 단계가 수행된다. 이는 GAMESS 프로그램 패키지를 사용하여 제한된 하트리-포크 계산법(Restricted Hartree-Fock calculations)을 통해 수행될 수 있으나 이에 제한되는 것은 아니다.
- [0024] 이후 변환된 최소 에너지 준위 상태 구조의 후보 리간드 분자의 진동수가 측정된다. 이후 진동수를 보다 간편히 비교하기 위해, 후보 리간드의 분자 진동수를 일정한 간격으로 나누는 단계; 동일 간격 구간에 존재하는 분자 진동수의 강도를 합산하여 그 구간의 대표값으로 하는 단계가 수행된다(corralled intensity of molecular vibrational frequency, CIMVF).
- [0025] 상기 구간(corrall 의 interval)은 계산상의 편의를 위하여 임의값으로 설정될 수 있으며, 보다 구체적으로, 3 내지 50의 범위 내의 자연수로 설정될 수 있고, 바람직하게는 결과의 정확도를 위해 5 내지 10으로 설정될 수 있으나 이에 제한되는 것은 아니다.
- [0026] 이후 상기 대표값의 패턴을 바탕으로 본 리간드와의 유사도를 비교하여 계통도를 구하는 단계가 수행된다.
- [0027] 리간드의 유사도를 비교하기 위하여 J-Express 등 군집화 (agglomerative hierarchical clustering; complete linkage) 응용 프로그램 등을 사용할 수 있으며, 계통도를 그리기 위해서는 TreeView 등의 프로그램을 사용하여 통상의 방법을 통해 계통도를 구할 수 있다.
- [0028] 계통도에서 후보 물질이 본 리간드와 가까이 위치하는 경우, 후보 리간드는 작용제로 작용한다고 예측할 수 있으며, 반대의 경우 길항제로 작용한다고 예측할 수 있다.

- [0029] 본 발명은 또한 후보 리간드의 3차원 구조 데이터 정보를 입력받는 입력부; 상기 입력부로 입력된 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하는 제1연산부; 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하는 제2연산부; 상기 진동수를 이용하여 후보 리간드의 작용을 예측하는 제3연산부; 및 상기 제3연산부를 통해 예측된 후보 리간드의 작용이 출력되는 출력부를 포함하는 리간드 분류 장치를 제공한다.
- [0030] 도 2는 본 발명에 따른 리간드 분류 장치의 구성이 도시된 블록도이다.
- [0031] 본 발명에 따른 리간드 분류 장치는 입력부(10)와, 연산부(20)와, 출력부(30)를 포함하여 구성된다. 또한 상기 연산부(20)은 제1연산부(21), 제2연산부(22) 및 제3연산부(23)를 포함하여 구성된다.
- [0032] 상기 입력부(10)는 키보드 또는 마우스 등의 입력장치를 사용할 수 있다. 또한 입력부(10)는 후보 리간드의 3차원 구조를 측정하는 장치를 포함할 수 있으며, 이들로부터 후보 리간드의 3차원 구조 데이터가 입력될 수 있다.
- [0033] 상기 연산부(20)는 입력부(10)로부터 입력된 데이터를 이용하여 후보 리간드의 작용체/길항체로서의 작용 가능성을 계산, 예측한다. 연산부(20)은 마이컴 또는 컴퓨터 등으로 구현될 수 있다.
- [0034] 본 발명의 한 구체예에서, 상기 연산부(20)는 입력부로 입력된 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하는 제1연산부(21); 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하는 제2연산부(22); 및 상기 진동수를 이용하여 후보 리간드의 작용을 예측하는 제3연산부(23)를 포함할 수 있다.
- [0035] 본 발명의 다른 구체예에서, 상기 제3연산부(23)는 후보 리간드의 분자 진동수를 일정한 간격으로 나누는 단계; 동일 간격 구간에 존재하는 분자 진동수의 강도를 합산하여 그 구간의 대표값으로 하는 단계; 및 상기 대표값의 패턴을 바탕으로 본 리간드와의 유사도를 비교하여 계통도를 구하는 단계를 포함하여 후보 리간드의 작용을 예측하는 것일 수 있다.
- [0036] 본 발명의 또 다른 구체예에서, 상기 리간드는 로돕신(rhodopsin) 유사 G-단백질-연결 수용체(G protein-coupled receptor, GPCR) 결합 리간드일 수 있으며, 보다 구체적으로 상기 로돕신(rhodopsin) 유사 G-단백질-연결 수용체(G protein-coupled receptor, GPCR)는 히스타민 수용체 또는 아데노신 수용체일 수 있으나 이에 제한되는 것은 아니다.
- [0037] 상기 출력부(30)는 연산부(20)에 의해 예측된 후보 리간드의 작용을 출력한다. 출력부(30)는 구체적으로 이에 제한되는 것은 아니나, 모니터, LCD 패널 등의 출력장치로 구현될 수 있다.
- [0038] 본 발명은 또한 상기 리간드의 분류 방법이 컴퓨터 상에서 수행될 수 있도록, 컴퓨터 프로그램으로 제작되어 이를 저장한 기록 매체로 구현될 수 있다.
- [0039] 즉, 본 발명에 따른 리간드 분류 프로그램을 기록한 컴퓨터로 읽을 수 있는 매체에는,
- [0040] 컴퓨터 입력장치를 통해 후보 리간드의 3차원 구조 데이터 정보를 입력받는 단계;
- [0041] 컴퓨터 연산장치가 상기 데이터를 기하학적으로 최적화하여 후보 리간드의 구조를 최소 에너지준위 상태의 구조로 변환하도록 하는 단계;
- [0042] 컴퓨터 연산장치가 상기 변환된 최소 에너지준위 상태 구조의 후보 리간드 분자의 진동수를 계산하도록 하는 단계;
- [0043] 컴퓨터의 연산장치가 상기 진동수를 이용하여 후보 리간드의 작용을 예측하게 하는 단계; 및
- [0044] 컴퓨터의 출력장치가 상기 연산장치에 의해 예측된 후보 리간드의 작용을 출력하게 하는 단계를 실행시키기 위한 리간드 분류 프로그램이 기록될 수 있다.
- [0045] 이하, 본 발명의 이해를 돋기 위하여 실시예를 들어 상세하게 설명하기로 한다. 다만 하기의 실시예는 본 발명의 내용을 예시하는 것일 뿐 본 발명의 범위가 하기 실시예에 한정되는 것은 아니다. 본 발명의 실시예는 당업

계에서 평균적인 지식을 가진 자에게 본 발명을 보다 완전하게 설명하기 위해 제공되는 것이다.

[0046] <실시예 1> 히스타민 수용체 결합 리간드의 분류

### 1. 데이터 세트

[0048] 데이터 세트의 SMILES 및 3차원 구조 데이터 포맷(SDF, structure data format) 파일은 NCBI (National Center for Biotechnology Information)의 PubChem Compound Database로부터 다운로드 받았고, 이후의 분석에 사용되었다. 데이터 세트의 모든 47 리간드 분자는 9 히스타민 작용제(agonist)와 38 히스타민 길항제(antagonist)를 포함하였으며, 이는 이하의 표 1과 같다.

**표 1**

PubChem ID	Compound name	Receptor	Function type	PubChem ID	Compound name	Receptor	Function type
87653	2-thiazoleethanamine	HRH1	agonist	3957	loratadine	HRH1	antagonist
2366	betahistine	HRH1	agonist	4034	meclozine	HRH1	antagonist
75919	demethylbetahistine	HRH1	agonist	4761	pheniramine	HRH1	antagonist
7741	betazole	HRH2	agonist	4927	promethazine	HRH1	antagonist
3077	dimaprit	HRH2	agonist	5002	quetiapine	HRH1	antagonist
3692	imetit	HRH3	agonist	3032915	burimamide	HRH2	antagonist
126688	amthamine	HRH3, H4	agonist	2756	cimetidine	HRH2	antagonist
41376	impromidine	HRH4	agonist	5282136	lafutidine	HRH2	antagonist
5227	SKF91488	HR	agonist	3033637	nizatidine	HRH2	antagonist
2267	azelastine	HRH1	antagonist	5282450	pibutidine	HRH2	antagonist
2678	cetirizine	HRH1	antagonist	3001055	ranitidine	HRH2	antagonist
2725	chlorpheniramine	HRH1	antagonist	5105	roxatidine	HRH2	antagonist
26987	clemastine	HRH1	antagonist	50287	tiotidine	HRH2	antagonist
6726	cyclizine	HRH1	antagonist	9954017	A-349821	HRH3	antagonist
124087	desloratadine	HRH1	antagonist	9818903	ABT-239	HRH3	antagonist
33036	dexchlorpheniramine	HRH1	antagonist	2366	betahistine	HRH3	antagonist
21855	dimetindene	HRH1	antagonist	6422124	ciproxifan	HRH3	antagonist
3100	diphenhydramine	HRH1	antagonist	3035746	iodophenpropit	HRH3	antagonist
667477	doxepin	HRH1	antagonist	9948102	pitolisant	HRH3	antagonist
3162	doxylamine	HRH1	antagonist	3035905	thioperamide	HRH3, H4	antagonist
3191	ebastine	HRH1	antagonist	2790	clobenpropit	HRH3	antagonist
19105	embramine	HRH1	antagonist	4908365	JNJ-7777120	HRH4	antagonist

3348	fexofenadine	HRH1	antagonist	10446295	VUF6002	HRH4	antagonist
1549 000	levocetirizine	HRH1	antagonist				

[0050] 2. 기하학적 최적화(Geometry optimization) 및 분자 진동수 계산

[0051] 분자 진동수를 계산하기 위해, 화합물의 구조가 반드시 기하학적으로 최적화되어야 한다. 이론적 3차원 형상 SDF가 최소 에너지가 아니어서 진공, 용매, 또는 결합 주머니 내에서 가장 낮은 에너지를 갖는 형태를 대변하지 않으므로, 리간드 분자 각각의 SDF 파일은 GAMESS 프로그램 패키지를 사용하여 단일 저-에너지 형태로의 전환 과정을 거쳤다. 제한된 하트리-포크 계산법(Restricted Hartree-Fock calculations)이 6-31G 기초 세트와의 BLYP DFT 방법을 사용하여 분자를 기하학적으로 최적화시키기 위해 수행되었다. 분자 진동수의 계산은 분자 형태와 연관이 있음에도 불구하고, 각각의 결과는 분자의 형태를 대변하였다. 각각의 기하학적 최적화 결과는 GAMESS 프로그램의 HESSIAN 의 RUNTYP 으로 분자 진동수를 계산하는 단계에 사용되었다. 얻어진 분자 진동수의 일 예를 도 3에 나타내었다.

[0052] 3. 분자 진동수의 군집 강도(Corralled Intensity of Molecular Vibrational Frequency, CIMVF)의 계보적 군집화(Hierarchical clustering)

[0053] 분자 비교를 간편화하기 위해, 계산된 분자 진동수는 오름차순으로 정렬되었고, 5 간격으로 나누어 군집화하였다. 각 간격 내의 각각의 진동수 강도를 합산하여 0 내지  $5000 \text{ cm}^{-1}$  의 진동수 범위에서 각 간격의 대푯값으로 하였다. 최종 결과물을 위해, 상기 각각의 분자에 대한 잠재적 분자 지표는 1000 요소를 포함하는 1차원 벡터로 표시되었다. 마지막으로, 히스타민 리셉터의 47 리간드의 지표를 포함하는 유사도 매트릭스가 계층적으로 군집화되었다. 본 실시예에서, 유사도 매트릭스는 최종적으로 47 꼭지점을 가진 나무로 군집을 형성하였다. CIMVF 계산은 Python에 의해 작성된 in-house scripts에 의해 수행되었다. CIMVF를 통한 유사도 매트릭스의 계층적 군집화를 통해 나타낸 계통도를 도 4에 나타내었다.

[0054] 도 4에 나타낸 바에 따르면, impropidine 을 제외하고는 8개의 작용제가 히스타민 근처에 위치하였으며(점선 원으로 표시), 모든 길항제는 히스타민 근처의 방사상의 트리에서 서로에 가깝게 군집화되었다. 따라서, 트리에서 작용제 및 저해제의 지역적 차이가 존재하며, 분자 진동수를 통해서 히스타민 수용체에 대한 작용제/저해제 분류에 대한 정보를 제공할 수 있다는 점을 알 수 있었다.

[0055] <비교예 1> 히스타민 수용체 리간드 구조 비교를 통한 리간드의 분류

[0056] 히스타민 및 다른 46 화합물간의 분자 구조의 유사성이 화합물의 SMILES로부터 분석되었다. 구조적 유사성이 계산되었으며, 이는 히스타민으로부터 각각의 분자의 Tanimoto distance로 표현되었다. 분자의 쌍별 비교(pairwise comparison)를 위한 Tanimoto 계수는 분자 구조 유사성의 측정에 가장 광범위하게 사용되는 측정값이고,  $T_c = N_{ab} / (N_a + N_b - N_{ab})$  로 규정되었다. 상기 식에서  $N_{ab}$  는 분자인식기술(molecular fingerprint)을 사용한 공통 비트(common bit) 수,  $N_a$  는 분자 a의 고유 비트(unique bit)이고,  $N_b$  는 분자 b의 고유 비트(unique bit)이다. 본 실시예에서, 분자 유사성은 38-비트 세트를 사용한 Tanimoto 계수로 계산되었다.

[0057] 히스타민 및 다른 리간드 분자들의 Tanimoto 계수를 아래 표 2에 나타내었다.

표 2

Compound	Tanimoto coefficient	Function type	Compound	Tanimoto coefficient	Function type
histamine	1	agonist			
imetit	0.45	agonist	loratadine	0.13	antagonist
burimamide	0.43	antagonist	JNJ-7777120	0.13	antagonist

thioperamide	0.41	antagonist	pitolisant	0.12	antagonist
iodophenpropit	0.3	antagonist	embramine	0.12	antagonist
demethylbetahistine	0.3	agonist	cyclizine	0.11	antagonist
ciproxifan	0.3	antagonist	dimetindene	0.11	antagonist
improxidine	0.29	agonist	meclozine	0.11	antagonist
clobenpropit	0.29	antagonist	nizatidine	0.1	antagonist
betahistine	0.27	antagonist	promethazine	0.1	antagonist
amthamine	0.25	agonist	azelaistine	0.1	antagonist
2-thiazoleethanamine	0.25	agonist	cetirizine	0.1	antagonist
cimetidine	0.24	antagonist	levocetirizine	0.1	antagonist
betazole	0.22	agonist	tiotidine	0.09	antagonist
VUF6002	0.19	antagonist	ABT-239	0.09	antagonist
doxylamine	0.18	antagonist	doxepin	0.09	antagonist
pheniramine	0.18	antagonist	dimaprit	0.08	agonist
chlorpheniramine	0.18	antagonist	A-349821	0.08	antagonist
dexchlorpheniramine	0.18	antagonist	pibutidine	0.08	antagonist
fexofenadine	0.16	antagonist	SKF91488	0.08	agonist
desloratadine	0.14	antagonist	roxatidine	0.08	antagonist
diphenhydramine	0.14	antagonist	ranitidine	0.07	antagonist
clemastine	0.13	antagonist	lafutidine	0.07	antagonist
ebastine	0.13	antagonist	quetiapine	0.07	antagonist

[0059] 일반적으로 Tanimoto 계수가 > 0.85인 경우, 매우 높은 유사성이 있음을 의미하며, > 0.75인 경우 군집 분자와 유사한 목적을 갖는 유사한 생물학적 활성 특성을 갖는 물질임을 의미한다. 작용제(agonist) 중 가장 높은 Tanimoto 계수는 0.45 (imetit)로 나타났으며, 잠재적인 작용제로 볼 만큼 유사도가 높게 나타나지 못했다. 또한, 작용제(agonist) 중 가장 낮은 Tanimoto 계수 값은 0.08 (dimaprit 및 SKF91488)로 나타났으며, 저해제 (antagonist)로서도 낮은 값을 보였다.

[0060] 따라서, Tanimoto 계수 및 분자의 기능간에 연관성이 미미하게 나타난 바, 리간드의 구조를 통해 리간드의 생물학적 활성을 예측하기는 어려운 것을 알 수 있었다.

#### [0061] <실시예 2> 아데노신 수용체 결합 리간드의 분류

[0062] 아데노신 수용체 리간드 데이터 세트의 SMILES 및 3차원 구조 데이터 포맷(SDF, structure data format) 파일은 NCBI (National Center for Biotechnology Information)의 PubChem Compound Database로부터 다운로드 받았고, 이후의 분석에 사용되었다. 데이터 세트는 이하의 표 3과 같다.

**표 3**

[0063]	PubChem ID	Compound name	Receptor	Function type	PubChem ID	Compound name	Receptor	Function type
	60961	adenosine	AdoR	agonist	1329	DPCPX	AdoRA1	antagonist
	10475082	2-MeCCPA	AdoRA1	agonist	6439091	FK453	AdoRA1	antagonist
	16218845	ADAC	AdoRA1	agonist	9821511	FR194921	AdoRA1	antagonist
	123807	CCPA	AdoRA1	agonist	64627	KW-3902	AdoRA1	antagonist
	4402	CPA	AdoRA1	agonist	9953065	SLV320	AdoRA1	antagonist
	158795	CVT-510	AdoRA1	agonist	9902054	WRC-0571	AdoRA1	antagonist
	11957545	ENBA	AdoRA1	agonist	2690	CGS15943	AdoRA2A	antagonist
	130214	GR79236	AdoRA1	agonist	6603761	CSC	AdoRA2A	antagonist

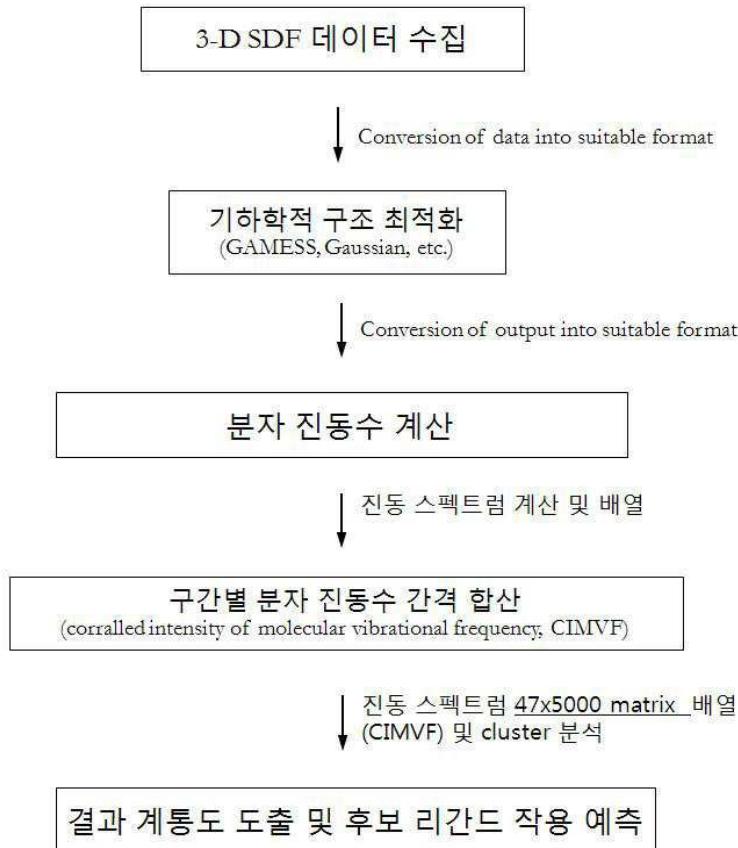
104795	NECA	AdoRA1	agonist	5311037	KW-6002	AdoRA2A	antagonist
993608 4	NNC21-0136	AdoRA1	agonist	11223144	MSX-2	AdoRA2A	antagonist
93205	R-PIA (L-PIA)	AdoRA1	agonist	10668061	SCH442416	AdoRA2A	antagonist
531143 1	S-EBNA	AdoRA1	agonist	176408	SCH58261	AdoRA2A	antagonist
164305	SDZWAG994	AdoRA1	agonist	2153	theophylline	AdoRA2A	antagonist
308659 9	CGS21680	AdoRA2A	agonist	9949541	VER-6947	AdoRA2A	antagonist
984128 4	CHA	AdoRA2A	agonist	9923269	VER-7835	AdoRA2A	antagonist
1585	CV1808	AdoRA2A	agonist	5697	XAC	AdoRA2A	antagonist
219024	CVT-3146	AdoRA2A	agonist	176407	ZM241385	AdoRA2A	antagonist
164437	HENECA	AdoRA2A	agonist	5372720	alloxazine	AdoRA2B	antagonist
131432	SHA40	AdoRA2A	agonist	1882	compd19a	AdoRA2B	antagonist
983351 9	UK-432097	AdoRA2A	agonist	10196114	CVT-5440	AdoRA2B	antagonist
957691 2	WRC-0470(bino)	AdoRA2A	agonist	1676	enprofylline	AdoRA2B	antagonist
8974	2-CADO	AdoRA2B	agonist	11716665	LAS-38096	AdoRA2B	antagonist
117178 31	BAY60-6583	AdoRA2B	agonist	CSID_19920712	MRE2029-F20	AdoRA2B	antagonist
997503 6	LUF5835	AdoRA2B	agonist	6603931	MRS1754	AdoRA2B	antagonist
2519	caffeine	AdoR	antagonist	5311041	OSIP-339391	AdoRA2B	antagonist
158540	BG9719	AdoRA1	antagonist	11689583	PSB-36	AdoRA2B	antagonist
216466	BG9928	AdoRA1	antagonist				

[0064] 이후, 상기 실시예 1의 2, 3단계를 동일하게 수행하였으며, 그 결과 얻어진 계통도를 도 5에 나타내었다. 도 5에 의하면, ZM241385를 제외하고 앞에 t로 표시된 길항제(antagonist)들은 모두 아데노신과 다른 군집을 형성하는 것을 알 수 있는 바, 본 발명의 방법을 통해 높은 확률로 길항제/작용제를 구별할 수 있음을 알 수 있었다.

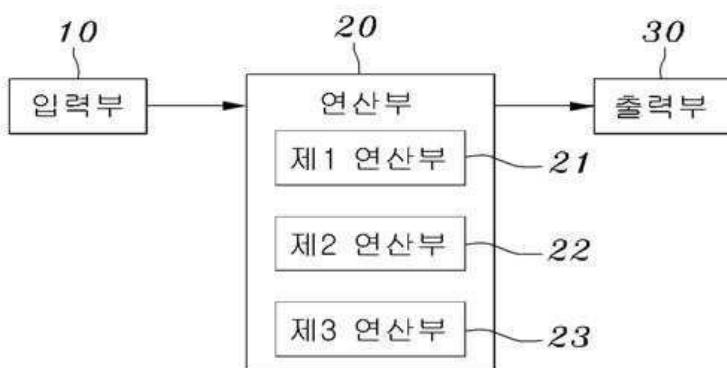
[0065] 이상으로 본 발명의 특정한 부분을 상세히 기술하였는 바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적 기술은 단지 바람직한 실시예일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서, 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의하여 정의된다고 할 것이다.

## 도면

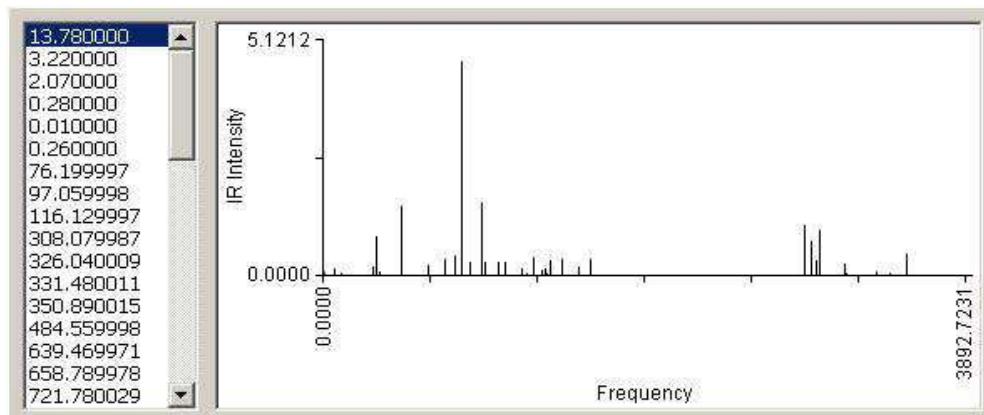
## 도면1



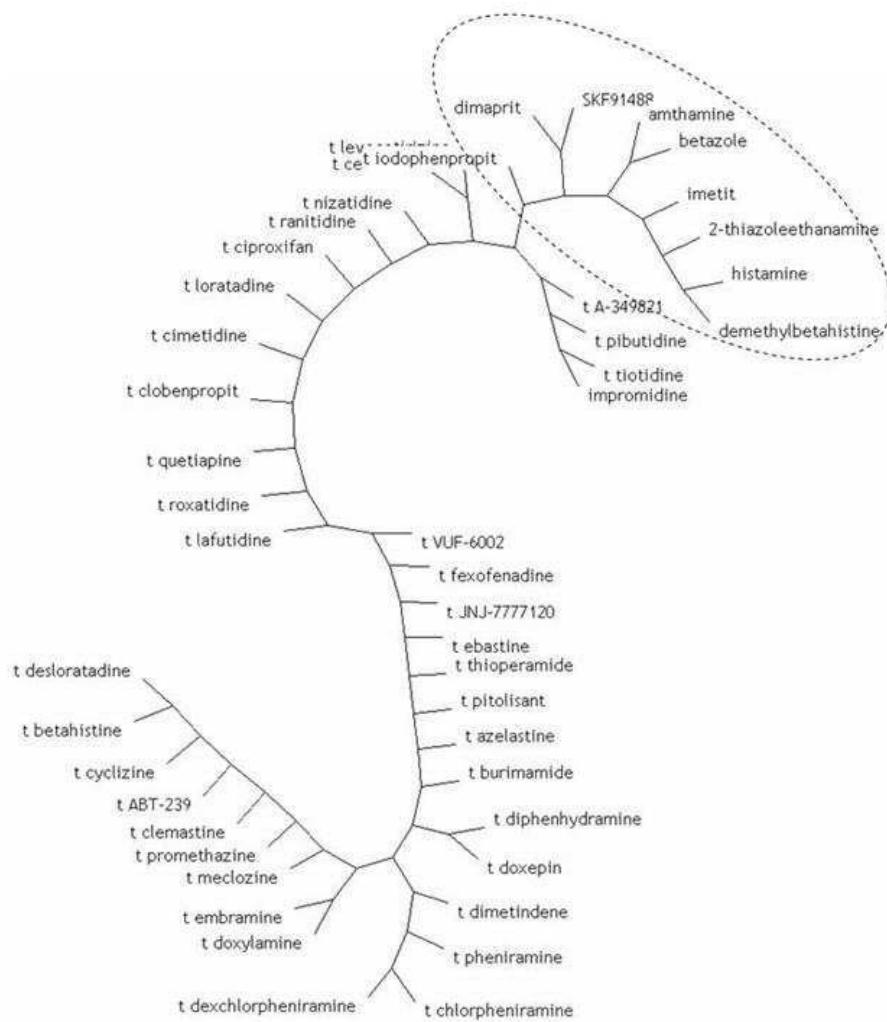
## 도면2



## 도면3



## 도면4



## 도면5

