(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2011년07월06일
(11) 등록번호 10-1047192
(24) 등록일자 2011년06월30일

(51) Int. Cl.
C07K 1/113 (2006.01) *G06F 19/00* (2011.01)
C12N 15/63 (2006.01) *A61K 38/00* (2006.01)
(21) 출원번호 10-2010-0034505
(22) 출원일자 2010년04월14일
심사청구일자 2010년04월14일
(65) 공개번호 10-2010-0114478
(43) 공개일자 2010년10월25일
(30) 우선권주장
61/169,636 2009년04월15일 미국(US)
(56) 선행기술조사문헌
KR1020080099278 A
Curr Opin Biotechnol. 2007 Aug;18(4):295-304.
KR1020100019964 A
전체 청구항 수 : 총 13 항

(73) 특허권자
포항공과대학교 산학협력단
경상북도 포항시 남구 효자동 산31 포항공과대학교내
(72) 발명자
최윤섭
부산광역시 수영구 남천동 558 뉴비치아파트 503동 806호
정준호
서울시 종로구 연건동 29 서울대학교의과대학 삼성암연구빌딩 510호
(뒷면에 계속)
(74) 대리인
손민

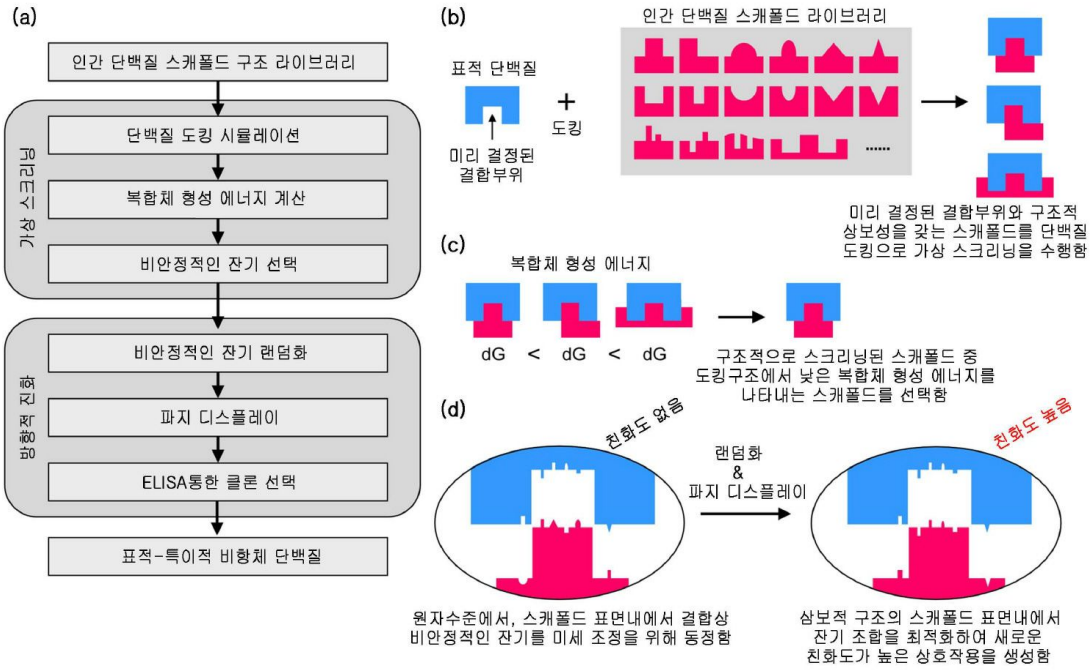
심사관 : 손영희

(54) 표적 특이적 비항체 단백질 및 이의 제조 방법

(57) 요 약

본 발명은 표적 특이적 비항체 단백질의 제조 방법에 관한 것으로, 보다 구체적으로, 비항체 단백질의 라이브러리에서 표적 단백질의 표적 부위와 구조적 상보성을 가지는 비항체 단백질을 선별하고, 선별된 비항체 단백질과 표적 단백질의 결합에너지를 계산하여, 선별된 비항체 단백질 중에서 안정적인 결합에너지를 갖는 비항체 단백질을 선별하고, 선정된 비항체 단백질과 표적 단백질이 직접 접촉하는 아미노산 잔기 중 결합에너지가 높은 아미노산 잔기를 선택하여, 선택된 아미노산 잔기를 결합에너지가 낮은 아미노산 잔기로 치환하는 단계를 포함하는 표적 특이적 비항체 단백질의 제조 방법에 관한 것이다. 또한 본 발명은 상기 방법에 의해 제조된 EGFR (Epidermal Growth Factor Receptor) 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 및 이를 포함하는 암 치료용 조성물에 관한 것이다.

대 표 도



(72) 발명자

유지호

서울시 동작구 신대방1동 565번지 우성아파트 10동 1502호

조현수

경기도 고양시 덕양구 행신동 SK아파트 201동 801호

윤수민

서울시 종로구 연건동 29 서울대학교의과대학 삼성암연구빌딩 510호

김경록

부산광역시 남구 대연4동 현대아파트 102동 306호

류성호

경상북도 포항시 남구 지곡동 교수아파트 5동 204호

김상욱

경상북도 포항시 남구 지곡동 교수아파트 9동 302호

이 발명을 지원한 국가연구개발사업

과제고유번호 2009K000859

부처명 교육과학기술부

연구관리전문기관

연구사업명 21C프론티어연구

연구과제명 인슐린 신호 프로테옴 규명 및 조절 물질 개발

기여율

주관기관 포항공과대학교 산학협력단

연구기간 2009년 04월 01일~ 2010년 03월 31일

특허청구의 범위

청구항 1

(a) 비항체 단백질의 라이브러리에서 표적 단백질의 표적 부위와 구조적 상보성을 가지는 비항체 단백질을 도킹 시뮬레이션을 통해 선별하는 단계;

(b) 상기 선별된 비항체 단백질과 표적 단백질의 결합에너지를 계산하는 단계;

(c) 상기 선별된 비항체 단백질 중에서 가장 낮은 결합에너지를 갖는 비항체 단백질을 선별하는 단계;

(d) 선정된 비항체 단백질과 표적 단백질이 직접 접촉하는 아미노산 잔기 중 결합에너지가 가장 높은 순서대로 1개 내지 6개의 아미노산 잔기를 선택하는 단계; 및

(e) 상기 (d) 단계에서 선택된 아미노산 잔기를 결합에너지가 (d) 단계에서 선택된 아미노산 잔기의 결합에너지보다 낮은 아미노산 잔기로 치환하는 단계를 포함하는, 표적 특이적 비항체 단백질의 제조 방법.

청구항 2

제1항에 있어서, 상기 (a) 단계 및 (b) 단계는 순차적으로, 또는 동시에 수행될 수 있는 것인 표적 특이적 비항체 단백질의 제조 방법.

청구항 3

제1항에 있어서, 상기 (a) 단계의 비항체 단백질 라이브러리는 인간 단백질 중 10 내지 40kDa의 분자량의 가지고, 단량체 (monomer), 동종이량체 (homodimer), 및 동종삼량체 (homotrimer) 중 어느 하나 이상을 형성할 수 있는 비항체 단백질을 포함하는 것인 표적 특이적 비항체 단백질의 제조 방법.

청구항 4

제1항에 있어서, 상기 (a) 단계의 라이브러리는

(i) 인간 단백질 중 10 내지 40kDa의 분자량을 갖는 단백질을 선택하는 단계;

(ii) 상기 (i) 단계에서 선택된 단백질에서 막 단백질 및 항체 단백질을 제거하는 단계;

(iii) 상기 (ii) 단계에서 제거하고 남은 단백질 중에서 공지의 데이터 베이스에서 알려진 단백질의 기능과 관련한 상호작용 파트너가 10개 이하인 단백질을 선택하는 단계; 및

(iv) 상기 (iii) 단계에서 선택된 단백질에 단량체 (monomer), 동종이량체 (homodimer), 및 동종삼량체 (homotrimer) 중 어느 하나 이상을 형성할 수 있는 단백질을 선택하는 단계를 포함하는 방법으로 제조하는 것인, 표적 특이적 비항체 단백질의 제조 방법.

청구항 5

삭제

청구항 6

제1항에 있어서, 상기 (a) 단계의 도킹 시뮬레이션은 상기 표적 단백질의 표적 부위 아미노산 잔기들 중 30% 이상에 결합하는 비항체 단백질을 선택하는 단계인 표적 특이적 비항체 단백질의 제조 방법.

청구항 7

제1항에 있어서, 상기 (e) 단계는 파지 디스플레이 및 바이오 패닝 (biopanning)을 수행하여 이루어지는 것인 표적 특이적 비항체 단백질의 제조 방법.

청구항 8

제1항에 있어서, 상기 (a) 단계의 표적 단백질은 EGFR (Epidermal Growth Factor Receptor) 도메인 2인 것인 표적 특이적 비항체 단백질의 제조 방법.

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

서열번호 3, 4 또는 5의 아미노산 서열을 갖는 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질.

청구항 13

제12항의 표적 특이적 비항체 단백질을 코딩하는 핵산.

청구항 14

제13항에 있어서, 상기 핵산은 서열번호 7, 8 또는 9의 핵산 서열을 갖는 것인 핵산.

청구항 15

제13항의 핵산을 포함하는 벡터.

청구항 16

제15항의 벡터로 형질전환된 형질전환체.

청구항 17

제12항의 표적 특이적 비항체 단백질을 포함하는 암치료용 조성물.

명 세 서

기 술 분 야

[0001] 본 발명은 표적 특이적 비항체 단백질의 제조 방법에 관한 것으로, 보다 구체적으로 비항체 단백질의 라이브러리에서 표적 단백질의 표적 부위와 구조적 상보성을 가지는 비항체 단백질을 선별하고, 선별된 비항체 단백질과 표적 단백질의 결합에너지를 계산하여, 선별된 비항체 단백질 중에서 안정적인 결합에너지를 갖는 비항체 단백질을 선별하고, 선정된 비항체 단백질과 표적 단백질이 직접 접촉하는 아미노산 잔기 중 결합에너지가 높은 아미노산 잔기를 선택하여, 선택된 아미노산 잔기를 결합에너지가 낮은 아미노산 잔기로 치환하는 단계를 포함하는 표적 특이적 비항체 단백질의 제조 방법에 관한 것이다. 또한 본 발명은 상기 방법에 의해 제조된 EGFR (Epidermal Growth Factor Receptor) 도메인 2에 특이적으로 결합하는 표적 특이적 비항체, 및 이를 포함하는 암치료용 조성물에 관한 것이다.

배 경 기 술

[0002] 질병을 치료하기 위한 단백질 신약으로 종래에는 항체 (antibody)를 사용하여 왔다. 항체는 면역계 내의 백혈구의 B세포에서 항원의 자극에 의하여 만들어지는 단백질로써, 항원을 만나게 되면 세포에 있는 수용체를 통하여 항원을 인식하고 수용체를 통해 결합한다. 특히 항체는 그 속성상, 특정 단백질을 특이적으로 인식하여 강하게 결합하는 특징이 있다. 이러한 성질을 이용하여 암 등의 질병을 일으키는 원인이 되는 단백질에 결합하여, 악의적인 단백질 상호작용을 막는 목적으로 사용되어 왔던 것이다.

[0003] 그러나, 치료용 항체의 경우 분자의 크기가 크고, 대량 생산이 어려우며, 생산 공정이 길고 가격이 높다는 문제점에 의해, 최근에는 환자 맞춤형 표적 치료를 위한 대안으로 비항체 단백질 (non-antibody protein scaffold)을 사용하고자 하는 연구가 활발히 이루어지고 있다. 비항체 단백질은 환자 맞춤형 표적 치료를 위한 신약의 한 부류로써 현재 신약 시장에 있어서 각광받고 있는 치료용 항체 (therapeutic antibody)가 가지는 상기 한계점을 극복하기 위한 대안으로 최근 새롭게 주목받기 시작하였다. 이러한 비항체 단백질은 항체 치료제와 같이 특정 표적 분자를 선택적으로 인식하여 강하게 결합함으로써 원하는 표적 분자의 활성을 저해하여 암, 자가 면역 질환과 같은 여러 질병을 치료하거나 진행을 막는다는 목적을 가지고 있다.

[0004] 하지만 기존의 비항체 단백질은 기존의 항체와 마찬가지로 표적 분자의 구조에 상관없이 오직 하나의 스캐폴드 (scaffold)에 기반하여 제작되고 있다. 단백질 들은 3차원적 구조에 기반하여, 레고 블럭이 엇물리듯 서로 결합하게 되므로, 특정 표적에 맞는 특이적인 맞춤형 단백질 스캐폴드를 사용하는 것이 이상적이다. 그러나, 표적과 특이적으로 결합할 수 있는, 상보적인 구조를 가지는 비항체 단백질을 찾아낼 수 있는 기술이 없었기 때문에, 표적 특이적인 비항체 단백질을 찾고, 이를 질병 치료용으로 이용하는데에 어려움이 있다는 문제점이 있었다. 특히, 단백질의 구조와 종류는 매우 다양하므로 수많은 단백질 중 어떤 것이 특정 표적과 잘 결합하는지 여부를 실험을 통해 확인하는 것은 불가능하다는 문제점이 있었다.

[0005] 또한 하나의 스캐폴드에 기반하는 기존의 항체와 비항체 단백질의 경우, 표적 분자에의 결합 부위가 우연히 결정되며, 이 결합 부위를 미리 디자인할 수 없다. 또한 기존의 비항체 단백질 및 항체가 결합할 수 있는 표적 단백질의 부위는 구조적으로 매우 제한되어 있으며, 특히 표적 부위가 요면 (concave)일 경우에는 접근이 원천적으로 불가능하였다. 따라서 표적 분자에 결합하는 단백질들 중 원하는 부위에 결합하는 단백질을 찾기 위해, 표적 결합 부위 확인 실험 (epitope mapping) 등 이후 오랜 기간 복잡한 후보 선택과정을 거쳐야 한다는 문제점이 있었다.

[0006] 따라서, 상기와 같은 기존 항체나 비항체 단백질의 한계점을 극복하는 표적에 최적화된 스캐폴드를 선정하여, 표적 분자의 원하는 표적 부위에 결합하는 비항체 단백질을 논리적으로 디자인할 수 있는 기술의 개발이 절실히 요구되고 있다. 원하는 표적 단백질의 원하는 부위에 강한 결합력과 특이성을 가지고 붙는 비항체 단백질을 디자인 및 제조할 수 있는 기술은 다양한 표적 단백질에 폭넓게 적용 가능한 기반 기술 (platform technology)이며, 이는 항체 신약이나 압타머 등에서도 아직까지 구현되지 못한 기술이다. 이에 따라 혁신적인 환자 맞춤형 표적 치료제 개발을 위해서도 절실히 요구되는 기술이다.

[0007] 이에 본 발명자들은 표적 특이적 비항체 단백질을 제조하는 방법을 찾기 위해 예의 노력한 결과, 3차 구조가 결정되어 있는 단백질의 구조 라이브러리를 기반으로, 주어진 표적 단백질의 원하는 표적 부위에 대한 가상 스크리닝을 통해 구조 상보성이 가장 높은 최적의 단백질 스캐폴드를 선정하고, 표적 분자와의 결합을 방해하는 아

미노산을 파지 디스플레이 및 바이오 패닝을 이용한 랜덤화를 수행하여 미리 결정된 표적 단백질의 특정 부위와 최적의 결합을 할 수 있는 표적 특이적 비항체 단백질을 제조할 수 있음을 확인하였다. 나아가 실제로 항암 표적인 EGFR (Epdermal Growth Factor Receptor)의 도메인 2를 표적으로 하는 비항체 단백질의 제조 및 상기 표적과의 결합력을 확인하여 본 발명을 완성하였다.

발명의 내용

해결하려는 과제

[0008] 본 발명의 하나의 목적은 (a) 비항체 단백질의 라이브러리에서 표적 단백질의 표적 부위와 구조적 상보성을 가지는 비항체 단백질을 선별하는 단계; (b) 상기 선별된 비항체 단백질과 표적 단백질의 결합에너지를 계산하는 단계; (c) 상기 선별된 비항체 단백질 중에서 안정적인 결합에너지를 갖는 비항체 단백질을 선별하는 단계: (d) 선정된 비항체 단백질과 표적 단백질이 직접 접촉하는 아미노산 잔기 중 결합에너지가 높은 아미노산 잔기를 선택하는 단계; 및 (e) 상기 (d) 단계에서 선택된 아미노산 잔기를 결합에너지가 낮은 아미노산 잔기로 치환하는 단계를 포함하는, 표적 특이적 비항체 단백질의 제조 방법을 제공하는 것이다.

[0009] 본 발명의 다른 목적은 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질을 제공하는 것이다.

[0010] 본 발명의 또 다른 목적은 상기 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질을 코딩하는 핵산, 상기 핵산을 포함하는 벡터, 또는 상기 벡터로 형질전환된 형질전환체를 제공하는 것이다.

[0011] 본 발명의 또 다른 목적은 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질을 포함하는 암 치료용 조성물을 제공하는 것이다.

과제의 해결 수단

[0012] 상기의 목적을 달성하기 위한 하나의 양태로서, 본 발명은 (a) 비항체 단백질의 라이브러리에서 표적 단백질의 표적 부위와 구조적 상보성을 가지는 비항체 단백질을 선별하는 단계; (b) 상기 선별된 비항체 단백질과 표적 단백질의 결합에너지를 계산하는 단계; (c) 상기 선별된 비항체 단백질 중에서 안정적인 결합에너지를 갖는 비항체 단백질을 선별하는 단계: (d) 선정된 비항체 단백질과 표적 단백질이 직접 접촉하는 아미노산 잔기 중 결합에너지가 높은 아미노산 잔기를 선택하는 단계; 및 (e) 상기 (d) 단계에서 선택된 아미노산 잔기를 결합에너지가 낮은 아미노산 잔기로 치환하는 단계를 포함하는, 표적 특이적 비항체 단백질의 제조 방법에 관한 것이다.

[0013] 본 발명의 표적 특이적 비항체 단백질의 제조 방법의 전체적인 모식도는 도 1a에 나타낸 바와 같다. 본 발명자들은 본 발명자가 개발한 도킹 시뮬레이션 및 결합 에너지 계산에 의한 가상 스크리닝 단계, 및 파지 디스플레이 및 바이오패닝을 통한 잔기의 랜덤화를 이용한 방향적 진화 단계를 수행하는 본 표적 특이적 비항체 단백질의 제조 방법을 PELEX (Protein Engineering with Laboratory Evolution & Extensive Computation)이라고 명명하였다.

[0014] 본 발명에서 사용된 용어 "비항체 단백질"은 분자량 10 내지 40 kDa의 크기로 기존의 치료용 항체에 비해 분자량이 1/10 이상으로 줄어든 단백질로 목표 조직으로의 침투가 용이하며, 박테리아 등에서 대량생산이 가능한 항체가 아닌 단백질로서, 분자 표적 치료를 위해 사용될 수 있는 특정 표적 분자를 선택적으로 인식하여 강하게 결합함으로써 원하는 표적분자의 활성을 저해하여 암, 자가 면역 질환과 같은 여러 질병을 치료하거나 진행을 막을 수 있는 단백질을 의미한다. 상기 "비항체 단백질"은 본원에서 "스캐폴드", "스캐폴드 단백질", "결합 단백질"과 혼용될 수 있다.

[0015] 본 발명에서 사용된 용어 "비항체 단백질의 라이브러리"는 인간이 자연계에서 그 3차 구조가 밝혀진 모든 단백질의 구조체로 이루어진 단백질 골격의 총 집합체를 의미하며, 이러한 집합체는 데이터 베이스 상에서 존재하도록 제조할 수 있다. 비항체 단백질의 라이브러리의 제조는 PDB (Protein Data Bank), SCOP (Structural Classification of Protein)과 같은 인간 단백질의 3차 구조를 모두 포함하는 데이터 베이스를 사용할 수 있으며, 이에 제한되는 것은 아니나 공지의 다양한 인간 단백질의 3차원 구조가 알려진 다양한 데이터 베이스를 통하여 비항체 단백질의 라이브러리의 제조를 위한 단백질을 선택할 수 있다. 이와 같은 데이터 베이스에는 인간 단백질 각각이 가지는 3차원적 구조에 따라 유형별로 복수의 그룹으로 분류되어 있다. 비항체 단백질 라이브러

리의 구체적인 제조 방법은 다음과 같다. 우선, 인간 단백질 중에서 조직 침투를 용이하게 하기 위해 분자량을 한정하여 10 내지 40kDa의 분자량을 갖는 단백질만을 선택할 수 있다. 이때, 각각의 인간 단백질 구조 유형별 그룹에서 10 내지 40kDa의 분자량을 갖는 대표 단백질을 5개 이하로 선택하여 모든 유형의 인간 단백질을 포함함과 동시에 전자 라이브러리의 규모를 축소함으로써 검색 속도를 향상시킬 수 있다. 이후에는 인간의 단백질 가운데 비항체 단백질이 아닌 단백질을 제거하기 위한 단계로서, 막 단백질 및 항체 단백질을 제거하는 단계를 포함할 수 있다. 막단백질 및 항체 단백질은 공지의 기능이 알려진 단백질의 데이터 베이스를 제한 없이 사용할 수 있으며, 그 예로 PDBTM (Protein Data Bank of Transmembrane Proteins)와 같은 데이터 베이스를 이용하여 키워드 검색 등을 이용하여 수행될 수 있다. 이 후, 무차별적인 결합을 피하기 위해 10개 이하의 상호작용이 알려진 단백질만을 선택하는 단계를 포함할 수 있다. 이와 같은 상호 작용과 관련된 정보도 공지의 단백질-단백질 상호 작용 정보를 제공하는 다양한 데이터 베이스를 사용할 수 있으나, 그 예로 HPRD (Human Protein Reference Database)를 이용하여 단백질의 상호 작용의 수를 계산할 수 있다. 그 후 동종사량체 (homotetramer), 동종육량체 (homohexamer) 등과 같이 동시에 많은 상호작용을 가지는 단백질을 제외하기 위해 단량체 (monomer), 동종이량체 (homodimer), 및 동종삼량체 (homotrimer) 중 어느 하나 이상을 형성할 수 있는 단백질을 선택하는 단계를 포함할 수 있으며, 이는 단백질간의 결합을 계산할 수 있는 공지의 데이터 베이스를 제한 없이 이용하여 수행될 수 있으나, 그 예로 SWISS-PROT 데이터 베이스를 이용할 수 있다. 상기와 같은 단계를 모두 거쳐 제조된 비항체 단백질 라이브러리는 1000개 내지 2000개의 구조가 알려진 비항체 단백질이 포함될 수 있다. 본 발명의 일 구현예에 따르면, 작은 분자량을 가지면서 무작위 결합의 위험이 적은 1261 개의 비항체 수용성 단백질로 구성된 비항체 단백질 라이브러리를 제작하였으며 이를 기반으로 도킹 시뮬레이션 등 가상 스크리닝을 수행하였다.

[0016] 바람직하게 상기 (a) 비항체 단백질의 라이브러리에서 표적 단백질의 표적 d부위와 구조적 상보성을 가지는 비항체 단백질을 선별하는 단계; 및 (b) 상기 선별된 비항체 단백질과 표적 단백질의 결합에너지를 계산하는 단계는 순차적으로 또는 동시에 수행될 수 있다.

[0017] 본 발명에서 용어 "표적 단백질"이란 비항체 단백질이 결합하는 목적 단백질을 의미하는 것으로서, 비항체 단백질의 결합을 통해 목적하고자 하는 용도에 따라 다양하게 선택될 수 있으며, 암 유발 단백질, 면역 질환 유발 단백질, 외래 감염 단백질 등이 제한 없이 선택될 수 있으나, 그 예로 EGFR (Epidermal Growth Factor Receptor) 일 수 있다.

[0018] 본 발명에서 용어 "표적 부위"란 표적 단백질의 3차 구조에서 병리적인 단백질 상호작용 및 결합에 직간접적으로 참여하는 부위로서, 해당 부위에 비항체 단백질이 강하게 결합하여 악의적인 단백질 상호작용을 막을 수 있으며, 그 예로 EGFR의 도메인 2일 수 있다. 또한, 본 발명의 비항체 단백질이 결합할 수 있는 표적 단백질의 결합 면은 요면 (concave) 등 다양한 표면에 결합할 수 있어서 본 발명의 비항체 단백질의 제조는 그 결합하는 표적 단백질의 3차 구조에 제한이 없다.

[0019] 본 발명에서 용어 "EGFR (Epidermal Growth Factor Receptor)"이란 대표적인 성장 인자 수용체로서, 유방암, 폐암, 대장암, 신장암, 담낭암, 두경부암, 난소암, 전립선암, 자궁경부암 및 위암 등 여러 다양한 암세포의 표면에 과발현되어 있으며, 성장인자 EGF의 결합에 의해 활성화되어 암세포의 증식, 침윤, 혈관생성이나 전이에 중심적인 역할을 하는 EGF의 수용체를 의미한다. 현재까지 다양하게 EGFR을 선택적으로 저해하여 암세포만을 선택적으로 공격하는 표적치료제의 개발에 대한 연구가 활발히 이루어지고 있었으나, 본 발명에서 표적으로 사용한 EGFR 도메인 2에 대한 연구는 거의 이루어지고 있지 않았다.

[0020] EGFR은 도메인 4개로 이루어져 있으며, 리간드인 EGF와 결합 후에 구조적 변화를 거쳐 노출된 도메인 2를 통해 동종이량체 (homodimer)를 형성함으로써 활성화된다. 따라서 본 발명이 표적으로 하는 EGFR 도메인 2에 강하게 결합하는 비항체 단백질은 리간드인 EGF가 결합되어 노출된 도메인 2에 특이적으로 결합하여 동종이량체를 형성할 수 없도록 하여서 EGFR의 활성을 특이적으로 저해할 수 있다 (도 3). 따라서 활성화된 EGFR만을 선택적으로 저해하기 때문에 정상세포가 아닌 암세포만을 선택적으로 공격하여 암치료제로서의 기능을 할 수 있을 것이다.

[0021] 상기 (a) 비항체 단백질의 라이브러리에서 표적 단백질과 구조적 상보성을 가지는 비항체 단백질을 선별하는 단계는 제조된 비항체 단백질의 라이브러리에 있는 다수의 비항체 단백질 또는 스캐폴드를 슈퍼 컴퓨터를 이용하여 표적 단백질과 구조적으로 상보성이 있는지 여부를 확인하는 단계로, 이는 도킹 시뮬레이션을 수행하여 가상으로 결합하여 이루어질 수 있다. 이는 스캐폴드 라이브러리에 있는 다수의 스캐폴드를 모두 표적과 가상으로 결합하여 다양한 결합 구조들을 만들어보는 과정으로, 도킹 시뮬레이션은 공지의 프로그램을 제한 없이 사용할 수 있으나, 그 예로 ZDOCK, PIPER, ClusPro, ICM-DISCO, RosettaDock, PatchDock, MolFit과 같은 프로그램일

수 있으며, 본 발명의 일 구현예에 따르면, PatchDock을 사용하여 구조적으로 상보성을 가진 비항체 단백질을 선별하였다. 하나의 표적과 하나의 비항체 단백질이 구조적 상보성에 의해 결합 가능한 방법은 무수히 많으며, PatchDock 프로그램은 상보성 등을 기준으로 하여 도킹 결합 구조의 순위를 매겨 준다. 이 때 하나의 비항체 단백질에 대한 다양한 결합구조들 중 1위 내지 10위에 있는 결합 구조를 선별하여 이후 단계의 분석에 사용하게 된다. 바람직하게 도킹 시뮬레이션 단계는 상기 표적 단백질의 표적 부위 아미노산 잔기들 중 기 설정된 개수 이상 결합되는 비항체 단백질을 선택할 수 있다. 더욱 바람직하게는 표적 단백질의 표적 부위인 EGFR 도메인 2인 경우, 비항체 단백질과 결합할 표적 잔기는 28개일 수 있으며, 기 설정된 개수는 10개로 28개의 잔기 중 10개 이상의 잔기가 결합에 참여하면 도킹 구조 중 스캐폴드가 안정된 구조로 EGFR 도메인 2에 결합한다고 판단하여, 이러한 도킹 구조를 형성하는 비항체 단백질을 선별할 수 있다. 이와 같은 잔기의 개수는 표적 단백질 혹은 표적 부위의 종류 등에 따라 개수를 다양하게 설정할 수 있다.

[0022] 상기 (b) 단계는 상기 (a) 단계에서 선별된 비항체 단백질과 표적 단백질의 결합에너지를 계산하는 단계로서, 이는 도킹 시뮬레이션 결과 표적 단백질의 원하는 부위에 일정 기준 이상으로 결합이 잘 이루어지는 비항체 단백질을 선택하기 위한 단계이다. 이때 표적 부위의 잔기들 중 결합되는 잔기의 수가 일정 개수 이상일 때 결합이 안정적으로 이루어질 수 있다.

[0023] 하기의 수학식 1에 따라 비항체 단백질과 표적의 결합 에너지를 계산한다. 이러한 결합 에너지의 계산은 다양한 공지의 프로그램으로 계산할 수 있으며, 이에 제한되는 것은 아니나, 그 예로 EGAD, RossettaDesign 등을 사용할 수 있다. 본 발명의 일 구현예에 따르면 EGAD 프로그램을 사용하여 결합 에너지를 계산하였다.

[0024] [수학식 1]

[0025] $\Delta G_{binding} = \Delta G_{complex} - \Delta G_{free}$

[0026] 결합에너지($G_{binding}$)는 표적 단백질과 비항체 단백질의 결합에 필요한 에너지로서, 표적 단백질과 비항체 단백질이 결합된 상태에서의 에너지($G_{complex}$)와 결합 전의 표적 단백질과 비항체 단백질 각각의 에너지(G_{free})의 차로 나타낼 수 있다. 결합에너지가 가장 낮은 비항체 단백질부터 순차적으로 정렬한다.

[0027] 상기 (c) 단계는 선별된 비항체 단백질 중에서 안정적인 결합에너지를 갖는 비항체 단백질을 선별하는 단계로서, 결합에너지가 낮을수록 표적과 비항체 단백질의 결합이 안정적으로 이루어질 수 있어, 표적에 특이적으로 강하게 결합할 확률이 높아진다. 이때, 원하는 개수를 설정하여, 설정된 개수의 비항체 단백질만을 정렬함이 바람직하다.

[0028] 상기 (d) 단계는 선정된 비항체 단백질과 표적 단백질이 직접 접촉하는 아미노산 잔기 중 결합에너지가 높은 잔기를 선택하는 단계로서, 직접 접촉하는 각 아미노산 잔기별 결합 에너지를 계산하여 각 잔기가 결합 에너지에 기여하는 정도를 산출한다. 비안정적인 결합 에너지를 가지는 잔기를 골라서, 추후 랜덤화하는 단계에 사용할 수 있다. 이 경우 비안정적인 결합 에너지를 갖는 표면 잔기의 개수가 너무 많으면 추후 랜덤화 과정에서 무리가 올 수 있으므로, 상위 스캐폴드 사이에서 상대적인 비교를 통해 제외를 할 수 있다. 본 발명의 일 구현예에 따르면, EGFR 도메인 2와 도킹 시뮬레이션에 의해 결합된 비항체 단백질 중 구조적 상보성이 높은 결합 에너지가 안정적인 스캐폴드로 선별된 10ZJ의 결합에 참여하는 각 잔기별 결합 에너지를 계산하여 결합에너지가 높게 나타난 잔기인 44Lys, 107 Glu, 110Glu, 114Asn 및 118Asp인 5개의 잔기를 비안정적인 잔기로 선택하였다 (도 4c, 표 2).

[0029] 상기 (e) 단계는 상기 (d) 단계에서 선택된 비안정적인 아미노산 잔기를 결합에너지가 낮은 아미노산 잔기로 치환하는 단계로서, 표적 단백질과 비항체 단백질간의 결합하는 표면의 아미노산 잔기 중 결합 에너지가 높아서 결합을 방해하는 아미노산 잔기를 결합에너지가 낮은 아미노산으로 치환하여 안정적인 결합을 할 수 있게 하는 방향적 진화 과정이다. 아미노산 잔기의 치환은 공지의 분자생물학 방법을 제한 없이 사용할 수 있으나, 바람직하게는 파지 디스플레이 및 바이오 패닝과정을 수행하여 이루어질 수 있다. 본 발명의 일 구현예에 따르면, 선택된 스캐폴드인 10ZJ의 결합에 참여하는 각 잔기별 결합 에너지를 계산하여 결합에너지가 높게 나타난 잔기인 44Lys, 107 Glu, 110Glu, 114Asn 및 118Asp인 5개의 아미노산의 치환을 위해 파지 디스플레이 랜덤 라이브러리를 제작한 후 5회의 바이오 패닝을 수행한 후 파지 ELISA를 통해 야생형의 10ZJ 스캐폴드 보다 결합력이 강화된 클론을 확인하였다 (도 5).

[0030] 본 발명자는 이미 알려져 있는 인간 단백질을 거의 모두 포함하는 단백질 중 10 내지 40 kDa의 분자량을 갖으면서 무작위 결합의 위험이 적은 1,261개의 비항체 수용성 단백질로 이루어진 인간 스캐폴드 라이브러리를 제작한 후, EGFR 도메인 2에 상보적인 모양을 가지는 스캐폴드를 스크리닝하기 위해 PatchDock 프로그램을 통해 대규모 도킹 시뮬레이션을 수행하여 상위 10개의 결합 모델을 선택하고, EGFR의 도메인 2의 28개의 잔기 중 10개의 잔기가 결합에 참여하면 안정된 결합 구조로 판별하여 EGFR 도메인 2의 결합 영역과 가장 상보적인 모양을 갖는 단백질 스캐폴드를 찾아내었다. 선정된 스캐폴드의 각 잔기의 결합에 관여하는 에너지 기여도를 EGAD를 통해 계산하여, 가장 안정적인 복합체 형성 에너지를 갖는 스캐폴드 1OZJ를 선정 및 5개의 비안정적인 결합 에너지를 갖는 표면 잔기를 선정하였다 (표 1 및 도 4). 그 후 파지 디스플레이 라이브러리 제작 및 바이오 패닝을 통해 상기 5개의 잔기를 안정적인 결합 에너지를 갖는 잔기로 치환하는 실험을 수행한 후, EGFR과의 결합을 파지 ELISA를 통해 확인하여 결합력이 증가된 15개의 클론을 선별하였으며 (표 2 및 도 5), 그 중 가장 결합력이 강한 클론 6, 7 및 9를 선정하여 서열 분석을 통해 원하는 서열이 변경되어서 결합력이 강화된 표적-특이적 비항체 단백질을 제조할 수 있음을 확인하였다 (도 7 내지 9).

[0031] 또 하나의 양태로서, 본 발명은 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질에 관한 것이다.

[0032] EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질은 상기에서 설명한 바와 같으며, 바람직하게 상기의 표적 특이적 비항체 단백질의 제조 방법에 의해 제조될 수 있다. 바람직하게는 상기 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질은 서열번호 3, 4 또는 5의 아미노산 서열일 수 있다. 본 발명의 일 구현예에 따르면 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질로 선정된 1OZJ 중 파지 ELISA 결과, EGFR과 결합력이 증가한 클론 6, 7 및 9의 아미노산 서열을 분석한 결과 각각 서열번호 3, 4 및 5의 서열을 갖는 것을 확인하였다 (표 2, 도 5 및 7 내지 9).

[0033] 또 하나의 양태로서, 본 발명은 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질을 코딩하는 핵산에 관한 것이다. 바람직하게는 서열번호 7, 8 또는 9의 핵산 서열일 수 있다. 본 발명의 일 구현예에 따르면 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질로 선정된 1OZJ 중 파지 ELISA 결과, EGFR과 결합력이 증가한 클론 6, 7 및 9를 코딩하는 핵산 서열을 분석한 결과 각각 서열번호 7, 8 및 9의 서열을 갖는 것을 확인하였다 (표 2, 도 5 및 7 내지 9)

[0034] 또 하나의 양태로서, 본 발명은 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질을 코딩하는 핵산을 포함하는 벡터에 관한 것이다. 본 발명의 벡터는 플라스미드 벡터, 코즈미드 벡터, 박테리오파아지 벡터 및 바이러스 벡터 등을 포함하나 이에 제한되지 않는다. 적합한 발현벡터는 프로모터, 오퍼레이터, 개시코돈, 종결코돈, 폴리아데닐화 시그넬 및 인핸서 같은 발현 조절 엘리먼트 외에도 막 표적화 또는 분비를 위한 시그널 서열 또는 리더 서열을 포함할 수 있다. 벡터의 프로모터는 구성적 또는 유도성일 수 있다. 또한 발현벡터는 벡터를 함유하는 숙주 세포를 선택하기 위한 선택 마커를 포함한다.

[0035] 또 하나의 양태로서, 본 발명은 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질을 코딩하는 핵산을 포함하는 벡터로 형질전환된 형질전환체에 관한 것이다.

[0036] 형질전환은 핵산을 유기체, 세포, 조직 또는 기관에 도입하는 어떤 방법도 포함되며, 당 분야에서 공지된 바와 같이 숙주 세포에 따라 적합한 표준 기술을 선택하여 수행할 수 있다. 이런 방법에는 전기충격유전자전달법 (electroporation), 원형질 융합, 인산 칼슘 (CaPO4), 침전, 염화 칼슘 (CaCl2) 침전, 실리콘 카바이드 섬유 이용한 교반, 아그로 박테리아 매개된 형질전환, PEG, 덱스트란 셀페이트, 리포펙타민 등이 포함되나 이로 제한되지 않는다. 숙주세포에 따라서 단백질의 발현량과 수식 등이 다르게 나타나므로, 목적에 가장 적합한 숙주세포를 선택하여 사용하면 된다. 숙주세포로는 에스케리치아 콜라이 (Escherichia coli), 바실러스 서브틸리스 (Bacillus subtilis), 스트렙토마이세스 (Streptomyces), 슈도모나스 (Pseudomonas), 프로테우스 미라빌리스 (Proteus mirabilis) 또는 스타필로코쿠스 (Staphylococcus)와 같은 원핵 숙주 세포가 있으나 이로 제한되는

것은 아니다. 또한, 진균 (예를 들어, 아스페르길러스(Aspergillus)), 효모 (예를 들어, 피치아 파스토리스(Pichia pastoris), 사카로마이세스 세르비시애 (Saccharomyces cerevisiae), 쉬조사카로마세스(Schizosaccharomyces), 뉴로스포라 크라사 (Neurospora crassa))등과 같은 하등 진핵 세포, 곤충 세포, 식물 세포, 포유동물 등을 포함하는 고등 진핵생물 유래의 세포를 숙주세포로 사용할 수 있다.

[0037] 또 하나의 양태로서, 본 발명은 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질을 포함하는 암 치료 또는 예방용 조성물에 관한 것이다.

[0038] 본 발명의 암 치료 또는 예방용 조성물은 EGFR의 활성화에 의해 발생되는 모든 종류의 암에 적용이 가능하며, 상기 암이 발생할 수 있는 임의의 동물에 적용이 가능하다. 동물은 인간 및 영장류뿐만 아니라, 소, 돼지, 양, 말, 개 및 고양이 등의 가축을 제한 없이 포함한다.

[0039] 본 발명에서 용어, "예방"이란 본 발명의 비항체 단백질을 포함하는 조성물 투여에 의해 암 유발이 억제되거나 지연되는 모든 행위를 의미하고, "치료"란 본 발명의 비항체 단백질을 포함하는 조성물 투여에 의해 암 증세가 호전되거나 이롭게 변경하는 모든 행위를 의미한다. 치료용 조성물로 사용될 경우, EGFR 도메인 2 특이적 비항체 단백질은 기존의 치료제와 직접 또는 링커 등을 통하여 간접적으로 커플링(예를 들어, 공유결합)되어 비항체 단백질-치료제 결합체 형태로 생체내로 투입하여 암 예방 또는 치료에 이용할 수 있다. 사용될 수 있는 치료제는 화학치료제, 면역치료제, 사이토킨, 케모킨, 항바이러스제, 생물작용제, 효소 저해물질 등을 포함한다. 본 발명의 EGFR 도메인 2 특이적 비항체 단백질을 포함하는 암 치료용 조성물은 약학적으로 허용가능한 담체를 추가로 포함할 수 있으며, 담체와 함께 제제화될 수 있다. 본 발명에서 용어, "약학적으로 허용가능한 담체"란 생물체를 자극하지 않고 투여 화합물의 생물학적 활성 및 특성을 저해하지 않는 담체 또는 희석제를 말한다. 액상 용액으로 제제화되는 조성물에 있어서 허용되는 약제학적 담체로는, 멸균 및 생체에 적합한 것으로서, 식염수, 멸균수, 링거액, 완충 식염수, 알부민 주사용액, 덱스트로즈 용액, 말토 덱스트린 용액, 글리세롤, 에탄올 및 이들 성분 중 1 성분 이상을 혼합하여 사용할 수 있으며, 필요에 따라 항산화제, 완충액, 정균제 등 다른 통상의 첨가제를 첨가할 수 있다. 또한 희석제, 분산제, 계면활성제, 결합제 및 윤활제를 부가적으로 첨가하여 수용액, 현탁액, 유탁액 등과 같은 주사용 제형, 환약, 캡슐, 과립 또는 정제로 제제화할 수 있다.

[0040] 본 발명의 EGFR 도메인 2에 특이적으로 결합하는 표적 특이적 비항체 단백질 및 약학적으로 허용가능한 담체를 포함하는 항암 조성물은 이를 유효성분으로 포함하는 어떠한 제형으로도 적용가능하며, 경구용 또는 비경구용 제형으로 제조할 수 있다. 본 발명의 약학적 제형은 구강(oral), 직장(rectal), 비강(nasal), 국소(topical; 볼 및 혀 밑을 포함), 피하, 질(vaginal) 또는 비경구(parenteral; 근육내, 피하 및 정맥내를 포함) 투여에 적당한 것 또는 흡입(inhalation) 또는 주입(insufflation)에 의한 투여에 적당한 형태를 포함한다.

[0041] 본 발명의 조성물을 유효성분으로 포함하는 경구 투여용 제형으로는, 예를 들어 정제, 트로키제, 로렌지, 수용성 또는 유성현탁액, 조제분말 또는 과립, 에멀젼, 하드 또는 소프트 캡슐, 시럽 또는 엘릭시르제로 제제화할 수 있다. 정제 및 캡슐 등의 제형으로 제제화하기 위해, 락토오스, 사카로오스, 솔비톨, 만니톨, 전분, 아밀로펙틴, 셀룰로오스 또는 젤라틴과 같은 결합제, 디칼슘 포스페이트와 같은 부형제, 옥수수 전분 또는 고구마 전분과 같은 붕괴제, 스테아르산 마스네슘, 스테아르산 칼슘, 스테아릴푸마르산 나트륨 또는 폴리에틸렌글리콜 왁스와 같은 윤활유를 포함할 수 있으며, 캡슐제형의 경우 상기 언급한 물질 외에도 지방유와 같은 액체 담체를 더 함유할 수 있다.

[0042] 본 발명의 조성물을 유효성분으로 포함하는 비경구 투여용 제형으로는, 피하주사, 정맥주사 또는 근육내 주사 등의 주사용 형태, 좌제 주입방식 또는 호흡기를 통하여 흡입이 가능하도록 하는 에어로졸제 등 스프레이용으로 제제화할 수 있다. 주사용 제형으로 제제화하기 위해서는 본 발명의 조성물을 안정제 또는 완충제와 함께 물에서 혼합하여 용액 또는 현탁액으로 제조하고, 이를 앰플 또는 바이알의 단위 투여용으로 제제화할 수 있다. 좌제로 주입하기 위해서는, 코코아버터 또는 다른 글리세라이드 등 통상의 좌약 베이스를 포함하는 좌약 또는 체료 관장제와 같은 직장투여용 조성물로 제제화할 수 있다. 에어로졸제 등의 스프레이용으로 제형화하는 경우, 수분산된 농축물 또는 습윤 분말이 분산되도록 추진제 등이 첨가제와 함께 배합될 수 있다.

발명의 효과

[0043] 본 발명의 표적 특이적 비항체 단백질의 제조 방법을 이용할 경우, 기존의 항체 신약이나 단백질 신약, 압타머

등 기존의 물질로 접근이 불가능하였던 다양한 표적에 적용 가능한 바이오 신약의 개발을 할 수 있는 기반 기술(platform technology)를 제공할 수 있다. 따라서, 본 발명의 방법은 혁신적인 환자 맞춤형 표적 치료제 개발을 가능케 하여, 사회적, 경제적 파급 효과가 매우 클 것으로 예상된다. 또한 본 발명의 EGFR 도메인 2를 저해하는 표적 특이적 단백질은 암 치료의 표적인 EGFR를 새로운 개념으로 EGFR의 활성을 특이적으로 저해하여 정상 세포가 아닌 암세포만을 선택적으로 공격하여 효율적인 암 치료 효과를 제공할 수 있을 것이다.

도면의 간단한 설명

[0044] 도 1은 결합 부위에 특이적인 스캐폴드를 디자인하기 위한 본 발명의 방법인 PELEX (Protein Engineering with Laboratory Evolution & Extensive Computaion)의 대략적인 방법을 도시화한 것이다.

(a)는 PELEX를 2 단계로 나타낸 순서도이다. 인간 단백질 스캐폴드 구조 라이브러리를 근거로 하여, 도킹 시뮬레이션과 복합체 형성 에너지 계산을 포함하는 가상 스크리닝을 수행하는 가상 스크리닝 단계와 스크리닝된 스캐폴드를 미리 결정된 결합 부위와 결합 친화력을 최적화되도록 방향적 진화하는 단계로 구성된다.

(b)는 표적 단백질의 미리 결정된 결합 부위와 상보적인 모양을 갖는 단백질 스캐폴드를 인간 단백질 스캐폴드 구조 라이브러리와 도킹 시뮬레이션을 바탕으로 선별하는 것을 나타낸 도면이다. 대규모의 단백질 스캐폴드 라이브러리는 알려진 모든 인간 단백질 구조를 포함한다.

(c)는 스크리닝된 도킹 구조를 사용한 복합체 형성 에너지 계산에 의해 방향적 진화를 통한 추가적인 최적화를 위한 안정적인 결합 에너지를 나타내는 스캐폴드를 선택하는 것을 나타낸 도면이다.

(d)는 결합 부위에 특이적인 스캐폴드의 미세 조정을 위해서, 서열 랜덤화와 파지 디스플레이를 이용한 방향적 진화를 수행하는 것을 나타낸 도면이다. 이러한 과정에서, 결합에 비안정적인 잔기를 표적과 스캐폴드 사이의 친화도를 최적화하기 위해 랜덤화하였다.

도 2는 EGFR의 아미노산 서열을 나타낸 도면이다. 파란색으로 표시한 아미노산이 EGFR 도메인 2 부위의 아미노산으로 166번째부터 309번째 아미노산부분이다. 빨간색으로 표시한 28개의 아미노산 잔기를 표면 잔기로 결합에 참여하는 아미노산의 후보로 선정하였다.

도 3은 EGFR의 활성 메커니즘과 EGFR의 활성을 억제하기 위해 디자인된 스캐폴드의 활성 기전을 나타낸 도면이다.

(a)는 자동억제 상태에서, EGFR은 도메인 2 (초록색)와 도메인 4 (빨간색) 사이의 분자 내 상호작용에 의해 비활성화되는 것을 나타낸다. 이 상태에서 도메인 2의 "이량체화 팔 (dimerization arm)"이 차단되어 있다.

(b)는 EGF가 결합하면 도메인 2와 도메인 4 사이의 분자 내 상호작용을 제거하고, 영역 재배치를 촉진하여, 도메인 2를 노출시켜 EGFR을 활성화시키는 것을 나타낸 것이다. 마지막으로 도메인 2의 "이량체화 팔" 매개의 상호 작용을 통해 동종이량체 (homodimer)를 형성한다. 동종이량체에서 오른쪽 EGFR 분자의 도메인 2는 회색으로 표시하였다.

(c)는 디자인된 스캐폴드가 EGFR 도메인 2에 결합하여 동종이량체가 되는 것을 막아 EGFR의 활성을 억제하는 기작을 도시화한 것이다. 이러한 결합 단백질은 활성화된 EGFR만 인식하고, 자동-억제된 EGFR은 인식하지 않는다.

도 4는 1OZJ의 구조 및 에너지를 도식화한 것이다.

(a)는 1OZJ (사슬 B)의 구조를 도식화한 것이다.

(b)는 1OZJ-EGFR의 도킹 구조를 도식화한 것이다. 마젠타색 사슬은 스캐폴드를 나타낸 것이다. 파랑, 초록, 노랑, 빨간 사슬은 각각 EGFR의 도메인 1, 2, 3 및 4를 나타낸 것이다. 원 안의 잔기는 EGFR과 스캐폴드의 접촉 잔기를 나타낸 것이다. (a)의 스캐폴드의 구조는 (b)에서의 스캐폴드 구조와 같은 방향에서 표시하였다.

(c)는 EGFR과의 복합체를 형성하기 위한 1OZJ의 표면 잔기의 에너지 기여를 나타낸 것이다. 방향적 진화를 통한 추가적인 개선을 위해 선택된 불안정한 결합에너지를 나타내는 잔기는 빨간색 막대로 표시하였다. 이는 추후 랜덤화를 통해 돌연변이시킬 아미노산 잔기를 나타낸 것이다.

도 5는 1OZJ의 비안정적인 결합에너지를 갖는 잔기를 돌연변이 시켰을 때의 결합 친화도를 나타낸 ELISA 결과를

나타낸 것이다. 가로 축은 클론 번호를 나타낸다. 세로 축은 ELISA 분석으로 얻은 광학 밀도를 나타낸다. 1OZJ 클론의 서열 랜덤화 결과는 표 2에 나타내었다.

도 6은 1OZJ 스캐폴드의 랜덤화를 통한 돌연변이 이전의 야생형 스캐폴드의 아미노산 서열 (윗 부분) 및 DNA 서열 (아래 부분)을 나타낸 도면이다. 빨간색으로 표시한 서열이 비안정적인 결합에너지를 나타내는 잔기로 돌연변이시킬 서열이다.

도 7은 1OZJ 스캐폴드의 6번 클론의 아미노산 서열 (윗 부분) 및 DNA 서열 (아래 부분)을 나타낸 도면이다. 빨간색으로 표시한 서열이 비안정적인 결합에너지를 나타내는 잔기로 돌연변이된 서열이다.

도 8은 1OZJ 스캐폴드의 7번 클론의 아미노산 서열 (윗 부분) 및 DNA 서열 (아래 부분)을 나타낸 도면이다. 빨간색으로 표시한 서열이 비안정적인 결합에너지를 나타내는 잔기로 돌연변이된 서열이다.

도 9는 1OZJ 스캐폴드의 9번 클론의 아미노산 서열 (윗 부분) 및 DNA 서열 (아래 부분)을 나타낸 도면이다. 빨간색으로 표시한 서열이 비안정적인 결합에너지를 나타내는 잔기로 돌연변이된 서열이다.

발명을 실시하기 위한 구체적인 내용

[0045] 이하, 하기 실시예를 통하여 본 발명을 더욱 상세하게 설명하기로 한다. 이들 실시예는 단지 본 발명을 예시하기 위한 것으로, 본 발명의 범위가 이들 실시예에 의해 제한되는 것으로 해석되지는 않는다.

[0046] **실시예 1: 구조적으로 결정된 인간 스캐폴드 라이브러리의 제작**

[0047] 이미 알려져 있는 인간 단백질을 거의 모두 포함하는 인간 단백질 스캐폴드의 라이브러리 제작을 위하여, SCOP (Structural Classification of Proteins) 데이터베이스에 있는 인간 단백질 각각이 가지는 3차원적 구조에 따라 유형별로 복수의 그룹으로 분류되어 있는 그룹에서 5개의 대표적인 구조를 선택하였다. 조직 침투를 용이하게 하기 위해 분자량을 한정하여 10-40 kDa의 분자량을 갖는 대표 단백질을 5개 이하로 선택하여 모든 유형의 인간 단백질을 포함함과 동시에 라이브러리의 규모를 축소하였다. 그리고 본 발명자들은 PDBTM (Protein Data Bank of Transmembrane Proteins) 데이터베이스의 주석을 이용하여 막 단백질 구조 및 키워드 검색으로 항체를 제외시켰다. 무차별적인 결합을 피하기 위하여 10개 이하의 상호작용이 알려진 단백질만을 선택하였다. 단백질의 상호작용의 수는 HPRD (Human Protein Reference Database)로 계산하였다. 또한, SWISS-PROT 데이터베이스의 'SUBNIT' 섹션을 바탕으로 동종사량체 (homotetramer), 동종육량체 (homohexamer) 등과 같은 동시에 많은 상호작용을 갖는 동종다량체 (homomultimer)를 제외하고, 단량체 (monomer), 동종이량체 (homodimer) 또는 동종삼량체 (homotrimer)를 형성하는 만드는 단백질을 선택하였다.

[0048] 그 결과, 작은 분자량을 가지면서 무작위 결합의 위험이 적은 1,261개의 비항체 수용성 단백질로 구성된 인간 단백질 구조의 스캐폴드 라이브러리를 만들었다.

[0049] **실시예 2: 구조 기반 가상 스크리닝**

[0050] 상기 실시예 1에서 제작한 인간 단백질 스캐폴드 라이브러리를 기반으로, EGFR 도메인 2에 상보적인 모양을 가지는 스캐폴드를 스크리닝하기 위해서 EGFR을 가지고 대규모 도킹 시뮬레이션에 의한 가상 스크리닝을 실시하였다 (도 1b). 그 방법은 하기와 같다. 단백질 도킹 시뮬레이션의 본래 목적은 현존하는 단백질-단백질 상호작용을 예측하는 것이지만, 본 발명에서는 두 단백질 사이의 상보적인 모양을 찾고, 새로운 단백질-단백질 상호작용을 만들기 위한 목적으로 사용하였다.

[0051] EGFR 도메인 2와 상보적인 구조를 갖는 스캐폴드를 스크리닝하기 위하여, 인간 EGFR 세포 외 영역(PDB id: 1IVO, Ogiso H et.al, Cell. Sep 20;110(6):775-87, 2002)의 A 사슬과 라이브러리 내의 각 스캐폴드 간의 도킹 시뮬레이션을 PatchDock 프로그램을 이용하여 수행하였다 (Schneidman-Duhovny, D. et al. *Proteins* 52, 107-112 (2003), Schneidman-Duhovny, D. et al. *Nucleic Acids Res* 33, W363-367 (2005)). 그 다음, 각 스캐폴드-EGFR 도킹 결과 중에서 상위 10개의 도킹 모델을 선택하였고, 상기 모델들의 결합 양상을 분석하여, EGFR의 무슨 잔기가 복합체 형성에 관여하는지 확인 및 EGFR 도메인 2의 도킹 구조를 찾았다. EGFR의 표적 부위인 도메인 2의 28개의 잔기를 기설정하고, 이중 복합체 형성에 10개 이상의 잔기가 관여하면 도킹 복합체 중 스캐폴드가 안정된 구조로 EGFR 도메인 2 (서열번호 1의 EGFR의 아미노산 중 166번째 부터 309번째 서열, 도 2) 에 결합한다고 판단하였다: 229, 230, 239, 242, 244, 245, 246, 248, 249, 250, 251, 252, 253, 262, 263, 264,

265, 275, 278, 279, 280, 282, 283, 284, 285, 286, 303, 304 (서열번호 1의 EGFR의 아미노산 서열 중 기 설정된 아미노산 서열인 28개 서열의 위치, 도 2).

[0052] 본 발명에서는 두 사슬 사이의 도킹 복합체 형성시 SASA(solvent accessible surface area) 내에서 $1Å^2$ 보다 큰 변화를 겪는 잔기를 결합에 관여하는 잔기라고 판단하였다. 각 잔기의 SASA는 Naccess (http://www.bioinf.manchester.ac.kr)를 이용하여 계산하였다. 이를 통하여 특정 잔기가 결합에 참여하는 지를 판단할 수 있다.

[0053] 그 결과, EGFR 도메인 2의 표적 결합 영역과 가장 상보적인 모양을 가지는 단백질 스캐폴드를 찾아냈다.

[0054] **실시예 3: 복합체 형성 에너지 계산 및 안정적인 결합 에너지를 갖는 스캐폴드의 선택**

[0055] EGFR 도메인 2의 결합에 관여하는 스캐폴드-EGFR 도킹 복합체에서 선택된 것들 중에서, 가장 안정적인 복합체 형성 에너지를 가지는 소수의 주도적인 스캐폴드를 추가적인 최적화를 위해 스크리닝하였다 (도 1c).

[0056] 스캐폴드-EGFR 도킹 구조의 복합체 형성 에너지 및 스캐폴드 내의 각 결합에 관여하는 잔기의 기여도를 EGAD (http://egad.berkeley.edu/EGAD_manual/ index.html)를 이용하여 계산하였다. 템플릿 (template)으로서 PatchDock에 의해 생성된 도킹 구조를 사용하였고, 명령어 중 JOBTYPE 항목에 complex_formation_energy를 사용하여 결합 에너지를 산출하였다. 다른 모든 옵션을 디폴트 (default) 값으로 사용하였다. 계산 결과, pseudo_DELTA_G 복합체 형성 값을 복합체 형성 에너지로 얻었고, level_1으로 표시된 표면 잔기의 dG 값을 비안정적인 잔기를 선택하기 위해 수집하였다. 이때 결합에 간접적으로 관여하는 level_2의 잔기를 제외하였다. 이 과정에서, 안정적인 에너지 이외에, 2 가지의 다른 기준을 고려하였다. 첫 번째, 적합하지 않은 구조를 가지는 스캐폴드를 매뉴얼 검사에 의해 제거하였다. 두 번째, 방향적 진화를 위한 무작위 서열 라이브러리를 제작하기 편리하도록 비안정적인 잔기가 일차 서열에서 서로 가깝게 위치한 스캐폴드를 선호하였다.

[0057] 결합 에너지의 계산은 하기의 수학식 1에 의해 계산하였다.

[0058] [수학식 1]

[0059] $$\Delta G_{binding} = \Delta G_{complex} - \Delta G_{free}$$

[0060] 결합에너지($G_{binding}$)는 표적과 비항체 바인더 단백질의 결합에 필요한 에너지로서, 표적과 비항체 바인더 단백질이 결합된 상태에서의 에너지($G_{complex}$)와 결합 전의 표적과 비항체 바인더 단백질 각각의 에너지(G_{free})의 차이를 계산하여 복합체일 경우에 더 안정한가, 그렇지 않은가를 계산하였다.

[0061] 그 결과, 가장 안정적인 복합체 형성 에너지를 갖는 한 개의 스캐폴드, 1OZJ (Crystal structure of Smad3-MH1)의 B 사슬 (도 4a)를 선택하였다 (표 1). EGFR에 대한 1OZJ의 복합체 형성 에너지를 도킹 모델로 산출하였다 (도 4b). 상기 스캐폴드를 마젠타색으로 표시하였고, 스캐폴드의 구조는 같은 방법으로 도 4a에 나타내었다. EGFR에 대한 스캐폴드의 접촉 잔기는 원 (sphere)으로 나타내었다.

표 1

PDB id	chain	PatchDock rank	ddG (kcal/mol)
2A5E	-	9	36.86
1KZY	D	4	62.75
1ZOP	B	8	76.26
1OZJ	**B**	**9**	**84.69**
4FAP	B	1	86.45
1AUT	L	9	88.62
1RK9	A	10	94.61
1XMJ	A	4	110.12
1SI3	A	9	111.94
1SI3	A	1	114.67

[0062]

[0063] 그 후, 선택된 스캐폴드에 있는 각 표면 잔기의 에너지 기여 실험을 통해 복합체를 형성하는 데 비안정적인 아미노산 잔기를 서열 랜덤화를 위해 선택하였다. 도 4c는 1OZJ의 각 접촉 잔기에 대한 에너지 기여를 각각 나타낸 결과이다.

[0064] 그 결과, 빨간색 막대로 나타낸 1OZJ의 5개 잔기를 랜덤화를 통한 방향적 진화를 위해 선택하였다 (44Lys, 107Glu, 110Glu, 114Asn,118Asp, 도 4c).

[0065] **실시예 4: 랜덤 라이브러리 제작**

[0066] 상기 실시예 3에서 선택한 EGFR 결합 스캐폴드 (1OZJ)의 DNA (서열번호 6, 도 6)를 Genscript(Piscataway, NJ)에서 합성하였고, 결합 영역의 특정 위치의 랜덤화를 위해 디자인한 NNK 프라이머를 (Genotech, Daejon, Korea) 이용하여 랜덤화하였다. 랜덤화한 스캐폴드를 *Sfi*I (Roche, Indianapolis, IN)로 절단하여 파지미드 벡터 pComb3X에 라이게이션하고, electrocompetent ER2738 (New England Biolabs, Beverly, MA)에 형질전환시켜서 랜덤 라이브러리를 제작하였다.

[0067] 그 결과, 선택된 1OZJ의 5개의 비안정적인 표면 잔기를 랜덤화 라이브러리의 구축에 의해 충분히 랜덤화하였다. 선택된 표면 잔기로 총 $4X10^8$ 조합을 생성하였다. 그 후 하기의 파지 디스플레이를 바탕으로 돌연변이 중에서 강한 결합 단백질을 스크리닝하였다.

[0068] **실시예 5: 바이오패닝 (biopanning), 파지 디스플레이 및 파지 ELISA**

[0069] <5-1> 바이오패닝 및 파지 디스플레이

[0070] EGFR에 특이적인 결합 단백질을 질적으로 향상시키기 위하여, 친화도 선택 기술인 5회의 바이오패닝을 아래와 같이 수행하였다.

[0071] 본 발명자들은 제조사의 지시에 따라 Dynabead M-270 Epoxy (Dynal, Invitrogen, Carlsbad, CA)를 코팅하고 블로킹하였다. 1.5 ㎍ EGFR (Sigma, St. Louis, MO)를 1 ㎖/mg이 되게 PBS (137 mM 염화 나트륨, 10 mM 인산염, 2.7 mM 염화 칼륨, ph 7.4)을 스톡 용액 (stock solution)으로 사용하여 5 X 10^6 개 Dynabead를 패닝의 각 회마다 코팅하였다. EGFR로 코팅된 Dynabead를 랜덤화한 파지 라이브러리 500 ㎕와 함께 2시간 동안 실온 로테이터 (rotator)에서 배양하였다. 결합하지 않은 파지를 제거하기 위해 Dynabead를 0.05% Tween 20를 포함하는 PBS(v/v)로 첫번째 바이오패닝 후에는 한 번, 2 및 3번째 바이오패닝 후에는 3번, 남은 2회에는 5번 세척하였다. EGFR-코팅된 Dynabead에 결합한 파지를 0.1M 글리신-HCl (pH 2.2) 30 ㎕를 2번 추가해서 용리하고, 2M Tris-HCl (pH9.1)를 추가해서 중성화시켰다. 용리된 파지를 신선하게 성장시킨 ER2738에 감염시키고,

다음 바이오패닝을 위해 하룻밤 배양했다.

[0072] <5-2> 파지 ELISA

[0073] 각 콜로니를 패닝 마지막 주기의 산출량 적정 플레이트에서 무작위로 선택하였고, 1 ㎖의 super broth (3% 트립톤, 2% 효모 추출, 1% 3-[N-Morpholino] propanesulfonic acid [MOPS], pH 7.0)에 접종하였다. 37 ℃에서 하룻밤 배양 후에, 배양 상등액을 파지 ELISA (enzyme-linked immunosorbant assay)를 수행하는 데 사용하였다. 미세적정 플레이트 웰을 4 ℃에서 하룻밤 동안 4 ㎍/㎖ EGFR을 포함하는 PBS로 코팅하였고, 37 ℃에서 1시간 동안 3% BSA를 포함하는 PBS로 블로킹하였다. 상기 플레이트를 파지를 포함하는 배양 상등액과 함께 37 ℃에서 2시간 동안 배양하였고, 0.05% Tween-20를 포함하는 PBS(PBST)로 3회 세척하였다. 그 후, 블로킹 완충용액(1:2000)에 희석된 HRP (horseradish peroxidase)와 결합한 항-M13 항체 (Sigma)를 더하고 37 ℃에서 1 시간 동안 배양하였다. 50 ㎕ TMB 기질 용액 (Pierce, Rockford, IL)을 각 웰에 더하고, 650 nm에서 OD를 측정하였다.

[0074] 그 결과, 패닝 과정 후에, 1OZJ 스캐폴드에서 나온 클론들의 EGFR과 결합력이 비약적으로 증가된 것을 확인하였다 (도 5). 1OZJ 야생형이 EGFR과의 친화도를 보여주진 않았지만, 상기 실시예 3에서 선택한 5개 아미노산들의 랜덤화를 거친 후에는 15 개 이상의 1OZJ 돌연변이 클론에서 EGFR에 대한 강한 결합력이 관찰되었다 (도 5 및 표 2). 가장 결합력이 높은 6, 7 및 9번 클론의 아미노산 서열 및 DNA 서열을 공지의 분자 생물학적 방법으로 분석하였다. 그 결과, 6번 클론은 서열번호 3의 아미노산 서열 및 서열번호 7의 DNA 서열을 갖고 있었다 (도 7). 7번 클론은 서열번호 4의 아미노산 서열 및 서열번호 9의 DNA 서열을 갖고 있었다 (도 8). 9번 클론은 서열번호 5의 아미노산 서열 및 서열번호 8의 DNA 서열을 갖고 있었다 (도 9).

표 2

Clone	residue					OD at 650nm
	44	107	110	114	118	
WT	K	E	E	N	D	0.086
6	S	G	S	N	L	1.826
9	A	Q	I	L	P	1.748
7	L	A	A	M	C	1.741
8	G	P	N	T	P	1.18
54	C	I	D	P	R	1.156
24	S	A	D	V	E	1.144
21	T	L	A	A	M	1.089
35	T	L	A	A	M	1.041
20	S	T	T	Q	G	1.028
53	A	S	S	S	I	1.02
34	S	T	T	Q	G	0.987
55	I	K	R	P	D	0.964
31	G	H	P	K	W	0.959
18	S	A	D	V	E	0.927
43	R	H	N	G	A	0.919

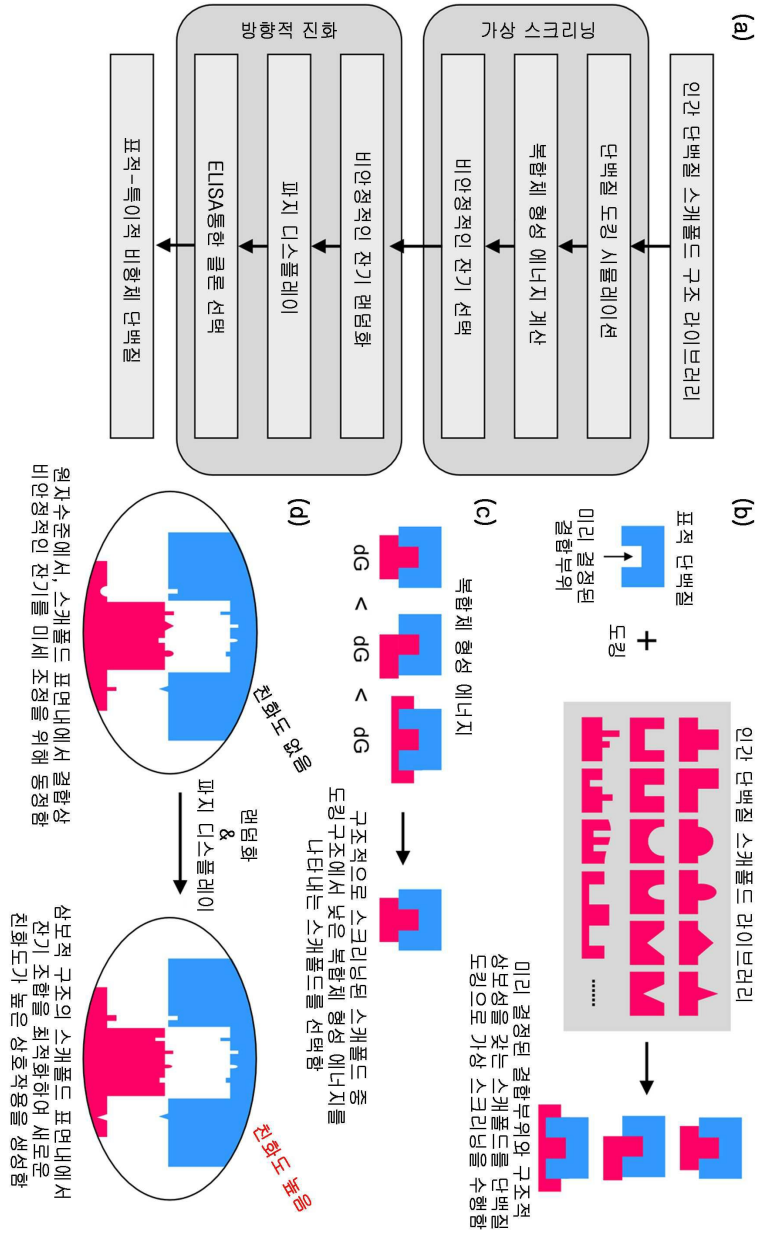
[0075]

[0076] 표 2 및 도 5와 같은 결과는 상보적인 구조와 안정적인 결합 에너지 외에는 서로 전혀 관련이 없는 스캐폴드 단백질과 EGFR 사이에 새로운 결합이 실제로 생성되었다는 것을 시사하는 결과이다. 아울러, 이와 같은 결과는 본 발명의 방법인 가상 스크리닝 및 바이오 패닝을 통한 랜덤화 과정을 거치는 표적 특이적 비항체 단백질의 제

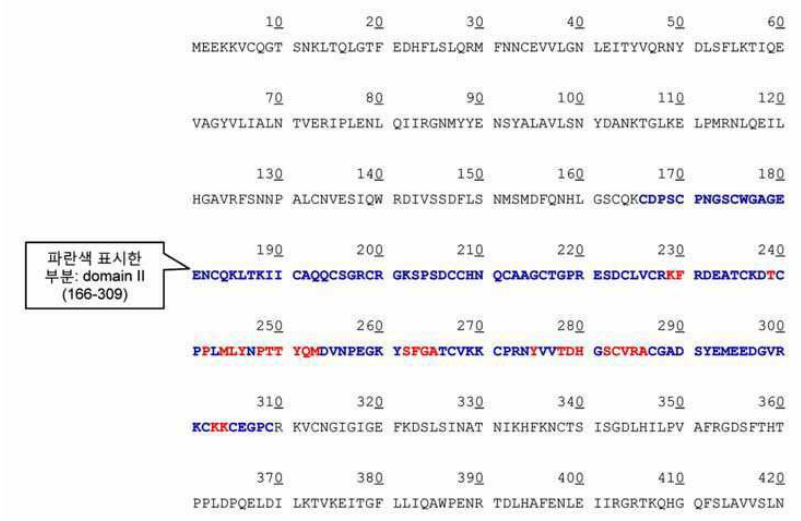
조방법을 이용할 경우, 원하는 표적을 미리 특정하고 이에 결합력과 특이성을 최적화한 결합 단백질을 제조할 수 있음을 시사하는 결과이다.

도면

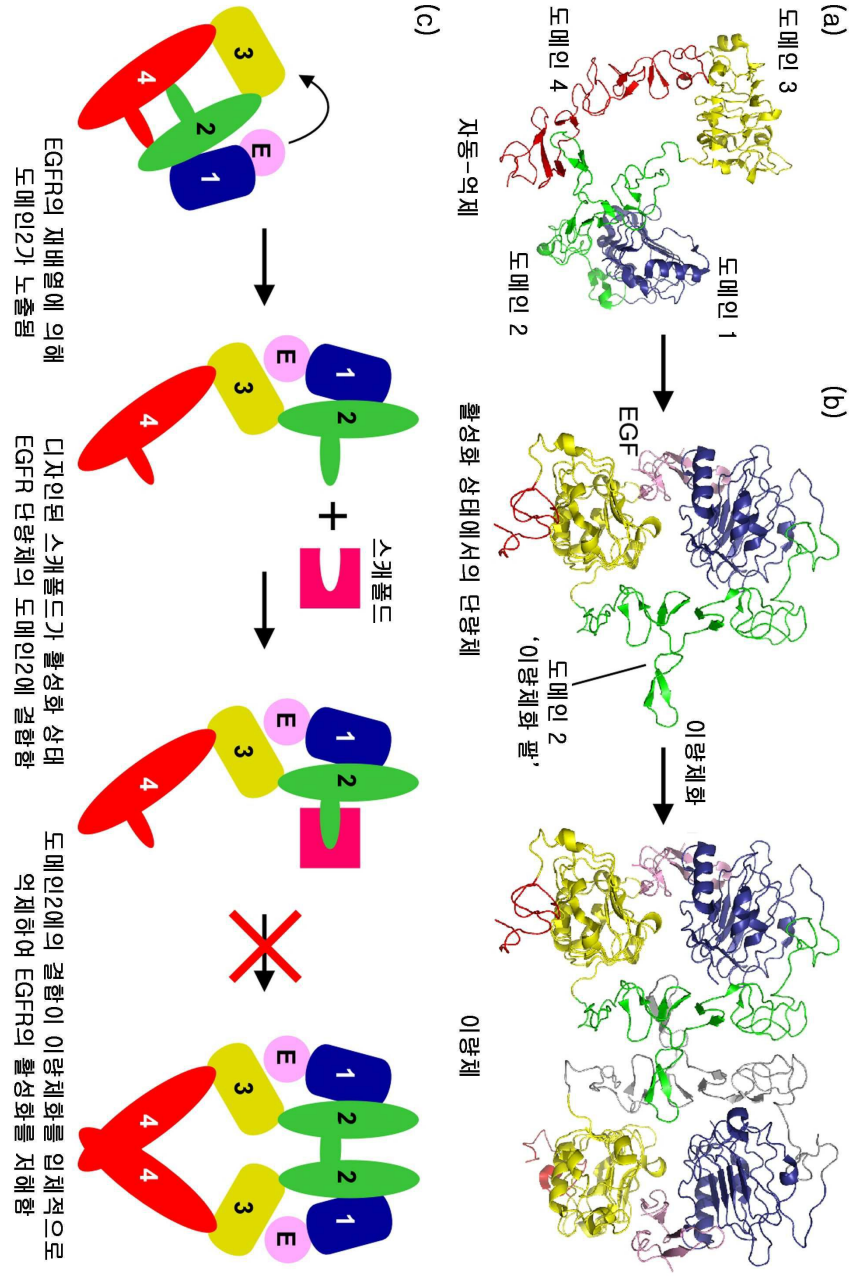
도면1



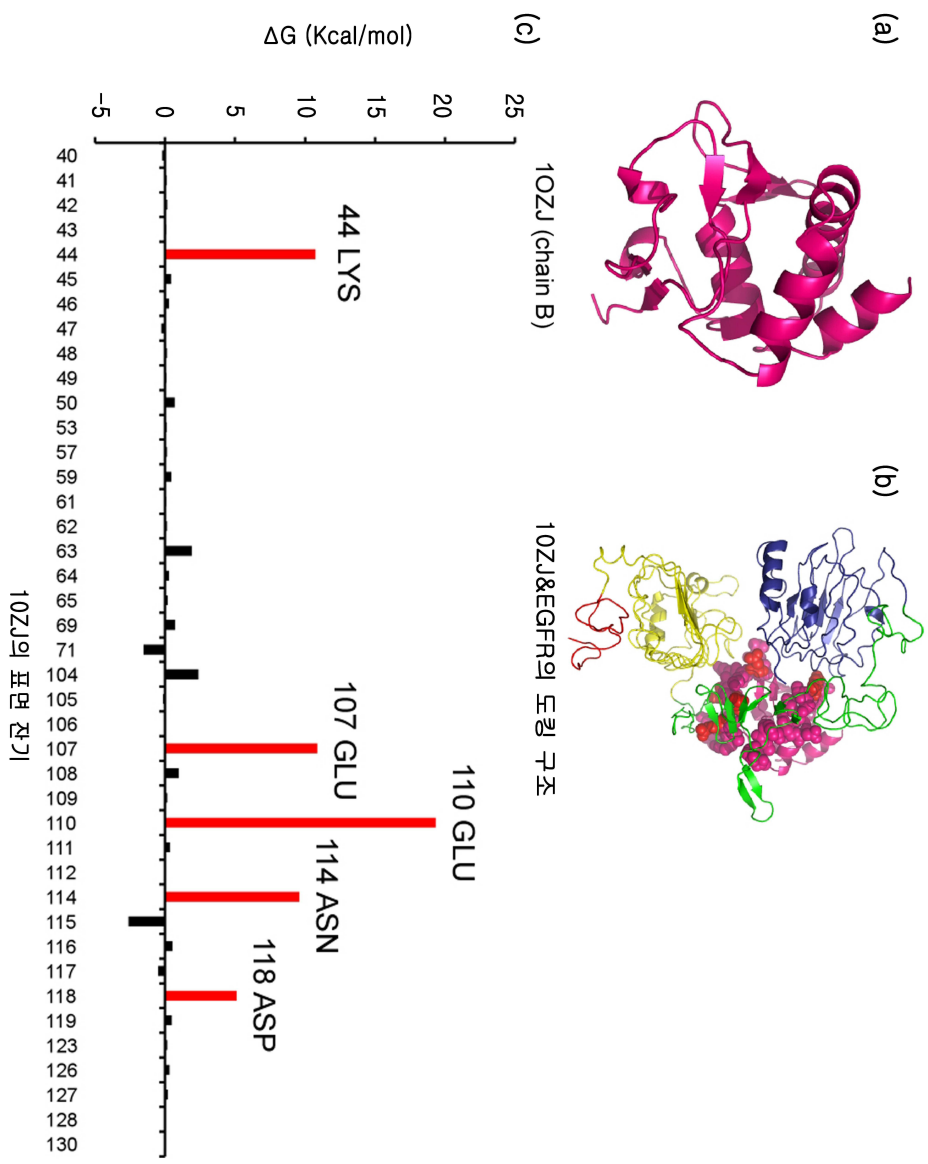
도면2


10 20 30 40 50 60
MEEKKVCQGT SNKLTQLGTF EDHFLSLQRM FNNCEVVLGN LEITYVQRNY DLSFLKTIQE
70 80 90 100 110 120
VAGYVLIALN TVERIPLENL QIIRGNMYYE NSYALAVLSN YDANKTGLKE LPMRNLQEIL
130 140 150 160 170 180
HGAVRFSNNP ALCNVESIQW RDIVSSDFLS NMSMDFQNHL GSCQKCDPSC PNGSCWGAGE
190 200 210 220 230 240
ENCQKLTKII CAQQCSGRCR GKSPSDCCHN QCAAGCTGPR ESDCLVCRKF RDEATCKDTC
250 260 270 280 290 300
PPLMLYNPTT YQMDVNPEGK YSFGATCVKK CPRNYVVTDH GSCVRACGAD SYEMEEDGVR
310 320 330 340 350 360
KCKKCEGPCR KVCNGIGIGE FKDSLSINAT NIKHFKNCTS ISGDLHILPV AFRGDSFTHT
370 380 390 400 410 420
PPLDPQELDI LKTVKEITGF LLIQAWPENR TDLHAFENLE IIRGRTKQHG QFSLAVVSLN
파란색 표시한 부분: domain II (166-309)

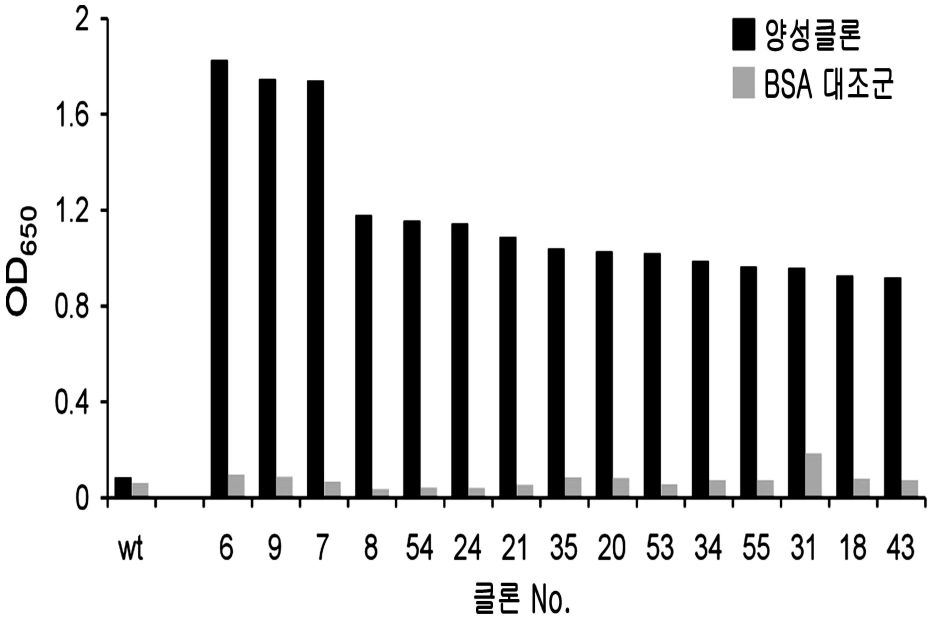
도면3


(a)
도메인 3
도메인 4
도메인 2
도메인 1
자동-억제
(b)
EGF
도메인 2
'이량체화 팔'
이량체화
활성화 상태에서의 단량체
이량체
(c)
스캐폴드
EGFR의 재배열에 의해
도메인2가 노출됨
디자인된 스캐폴드가 활성화 상태
EGFR 단량체의 도메인2에 결합함
도메인2에의 결합이 이량체화를 입체적으로
억제하여 EGFR의 활성화를 저해함

도면4


(a)
1OZJ (chain B)
(b)
1OZJ&EGFR의 도킹 구조
(c)
ΔG (Kcal/mol)
−5
0
5
10
15
20
25
10ZJ의 표면 잔기
40
41
42
43
44
45
46
47
48
49
50
53
57
59
61
62
63
64
65
69
71
104
105
106
107
108
109
110
111
112
114
115
116
117
118
119
123
126
127
128
130
44 LYS
107 GLU
110 GLU
114 ASN
118 ASP

도면5



도면6

```
MSSILPFTPP IVKRLLGWKK GEQNGQEEKW CEKAVKSLVK KLKKTGQLDE LEKAITTQNV
NTKCITIPRS LDGRLQVSHR KGLPHVIYCR LWRWPDLHSH HELRAMELCE FAFNMKKDEV
CVNPYHYQRV ETPVLPPVLV PRHT

ATGTCGTCCA TCCTGCCTTT CACTCCCCCG ATCGTGAAGC GCCTGCTGGG CTGGAAGAAG
GGCGAGCAGA ACGGGCAGGA GGAGAAATGG TGCGAGAAGG CGGTCAAGAG CCTGGTCAAG
AAACTCAAGA AGACGGGGCA GCTGGACGAG CTGGAGAAGG CCATCACCAC GCAGAACGTC
AACACCAAGT GCATCACCAT CCCCAGGTCC CTGGATGGCC GGTTGCAGGT GTCCCATCGG
AAGGGGCTCC CTCATGTCAT CTACTGCCGC CTGTGGCGAT GGCCAGACCT GCACAGCCAC
CACGAGCTAC GGGCCATGGA GCTGTGTGAG TTCGCCTTCA ATATGAAGAA GGACGAGGTC
TGCGTGAATC CCTACCACTA CCAGAGAGTA GAGACACCAG TTCTACCTCC TGTGTTGGTG
CCACGCCACA CA
```

도면7

```
MSSILPFTPP IVKRLLGWKK GEQNGQEEKW CEKAVKSLVK KLKSTGQLDE LEKAITTQNV
NTKCITIPRS LDGRLQVSHR KGLPHVIYCR LWRWPDLHSH HELRAMGLCS FAFNMKKLEV
CVNPYHYQRV ETPVLPPVLV PRHT

ATGTCGTCCA TCCTGCCTTT CACTCCCCCG ATCGTGAAGC GCCTGCTGGG CTGGAAGAAG
GGCGAGCAGA ACGGGCAGGA GGAGAAATGG TGCGAGAAGG CGGTCAAGAG CCTGGTCAAG
AAACTCAAGA GCACGGGGCA GCTGGACGAG CTGGAGAAGG CCATCACCAC GCAGAACGTC
AACACCAAGT GCATCACCAT CCCCAGGTCC CTGGATGGCC GGTTGCAGGT GTCCCATCGG
AAGGGGCTCC CTCATGTCAT CTACTGCCGC CTGTGGCGAT GGCCAGACCT GCACAGCCAC
CACGAGCTAC GGGCCATGGG TCTGTGTTCC TTCGCCTTCA ATATGAAGAA GTTGGAGGTC
TGCGTGAATC CCTACCACTA CCAGAGAGTA GAGACACCAG TTCTACCTCC TGTGTTGGTG
CCACGCCACA CA
```

도면8

MSSILPFTPP IVKRLLGWKK GEQNGQEEKW CEKAVKSLVK KLK**L**TGQLDE LEKAITTQNV
NTKCITIPRS LDGRLQVSHR KGLPHVIYCR LWRWPDLHSH HELRAM**A**LC**A** FAF**M**MKK**C**EV
CVNPYHYQRV ETPVLPPVLV PRHT

ATGTCGTCCA TCCTGCCTTT CACTCCCCCG ATCGTGAAGC GCCTGCTGGG CTGGAAGAAG
GGCGAGCAGA ACGGGCAGGA GGAGAAATGG TGCGAGAAGG CGGTCAAGAG CCTGGTCAAG
AAACTCAAG**C** **TT**ACGGGGCA GCTGGACGAG CTGGAGAAGG CCATCACCAC GCAGAACGTC
AACACCAAGT GCATCACCAT CCCCAGGTCC CTGGATGGCC GGTTGCAGGT GTCCCATCGG
AAGGGGCTCC CTCATGTCAT CTACTGCCGC CTGTGGCGAT GGCCAGACCT GCACAGCCAC
CACGAGCTAC GGGCCATG**GG** **T**CTGTGT**GCG** TTCGCCTTC**A** **TG**ATGAAGAA G**TGT**GAGGTC
TGCGTGAATC CCTACCACTA CCAGAGAGTA GAGACACCAG TTCTACCTCC TGTGTTGGTG
CCACGCCACA CA

도면9

MSSILPFTPP IVKRLLGWKK GEQNGQEEKW CEKAVKSLVK KLK**A**TGQLDE LEKAITTQNV
NTKCITIPRS LDGRLQVSHR KGLPHVIYCR LWRWPDLHSH HELRAM**Q**LC**I** FAF**L**MKK**P**EV
CVNPYHYQRV ETPVLPPVLV PRHT

ATGTCGTCCA TCCTGCCTTT CACTCCCCCG ATCGTGAAGC GCCTGCTGGG CTGGAAGAAG
GGCGAGCAGA ACGGGCAGGA GGAGAAATGG TGCGAGAAGG CGGTCAAGAG CCTGGTCAAG
AAACTCAAG**G** **CG**ACGGGGCA GCTGGACGAG CTGGAGAAGG CCATCACCAC GCAGAACGTC
AACACCAAGT GCATCACCAT CCCCAGGTCC CTGGATGGCC GGTTGCAGGT GTCCCATCGG
AAGGGGCTCC CTCATGTCAT CTACTGCCGC CTGTGGCGAT GGCCAGACCT GCACAGCCAC
CACGAGCTAC GGGCCATG**CA** **G**CTGTGT**ATT** TTCGCCTTC**C** **TG**ATGAAGAA G**CCG**GAGGTC
TGCGTGAATC CCTACCACTA CCAGAGAGTA GAGACACCAG TTCTACCTCC TGTGTTGGTG
CCACGCCACA CA

서 열 목 록

서열목록 전자파일 첨부