(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0087766
(43) 공개일자 2016년07월22일

(51) 국제특허분류(Int. Cl.)
C40B 50/02 (2006.01) *C07K 16/00* (2006.01)
C40B 40/10 (2006.01)
(52) CPC특허분류
C40B 50/02 (2013.01)
C07K 16/00 (2013.01)
(21) 출원번호 10-2016-0004232
(22) 출원일자 2016년01월13일
심사청구일자 2016년01월13일
(30) 우선권주장
1020150006358 2015년01월13일 대한민국(KR)

(71) 출원인
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
심현보
서울시 강남구 압구정로 29길 23, 210동 702호 (압구정동 현대아파트)
김지혜
인천광역시 서구 신현동 278 ~295
백설련
서울특별시 마포구 고산14길 24-13 202호 (대흥동)
(74) 대리인
손민

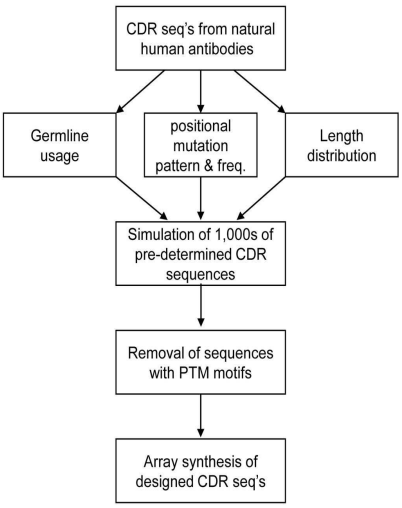
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **신규 항체 라이브러리 제조방법 및 이로부터 제조된 라이브러리**

(57) 요 약

본 발명은 신규 항체 라이브러리 제조방법 및 이로부터 제조된 라이브러리에 관한 것이다. 본 발명에 의해 제조된 항체 라이브러리는 다수의 항원에 대해 우수한 물성을 가지는 항체들을 포함하여, 기능적 다양성을 가질 뿐 아니라 고유서열을 다수 포함하는 등 항체 라이브러리로써 유용하게 사용될 수 있다.

대 표 도 - 도4



(52) CPC특허분류
C40B 40/10 (2013.01)
C07K 2317/54 (2013.01)
C07K 2317/55 (2013.01)
C07K 2317/565 (2013.01)
C07K 2317/622 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 NRF-2010-0029762
부처명 미래창조과학부
연구관리전문기관 한국연구재단
연구사업명 글로벌프론티어 사업
연구과제명 도전적 표적분자에 대한 고효율 항체 도출 기술
기 여 율 1/1
주관기관 이화여자대학교
연구기간 2010.10.22 ~ 2019.08.31

명 세 서

청구범위

청구항 1

항체의 상보성 결정부위(complementarity Determining Regions, CDRs) 서열을 개별적으로 설계하는 단계; 및 상기 설계한 서열을 가지는 상보성 결정부위들을 포함하는 항체를 합성하여 라이브러리를 제조하는 단계를 포함하는, 항체 라이브러리를 제작하는 방법.

청구항 2

제1항에 있어서, 상기 항체 라이브러리를 구성하는 항체의 상보성 결정부위를 이루는 중쇄 상보성 결정부위 1(CDR-H1), 중쇄 상보성 결정부위 2(CDR-H2), 중쇄 상보성 결정부위 3(CDR-H3), 경쇄 상보성 결정부위 1(CDR-L1), 경쇄 상보성 결정부위 2(CDR-L2) 및 경쇄 상보성 결정부위 3(CDR-L3)은 다양성(polymorphism)을 가지는 것인, 항체 라이브러리를 제작하는 방법.

청구항 3

제1항에 있어서, 상보성 결정부위 서열의 개별적인 설계시, CDR-H1, CDR-H2, CDR-L1 또는 CDR-L2에 대하여 실제 인간유래 숙성 항체의 상보성 결정부위의 i) 생식계열(germline) 면역글로불린 유전자 사용빈도, ii) 각 아미노산 위치별로 체세포 초돌연변이에 의한 각 20종의 아미노산으로의 변이 빈도, iii) 상보성 결정부위 서열의 길이별 빈도 또는 iv) 이들의 조합을 분석하여 계산된 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하는 것인, 항체 라이브러리를 제작하는 방법.

청구항 4

제1항에 있어서, 상보성 결정부위 서열의 개별적인 설계시, CDR-L3에 대하여

a) 상기 상보성 결정부위의 N 말단으로부터 7개 내지 8개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 i) 생식계열(germline) 면역글로불린 유전자 사용빈도, ii) 각 아미노산 위치별로 체세포 초돌연변이에 의한 각 20종의 아미노산으로의 변이 빈도, iii) 상보성 결정부위 CDR-L3 서열의 길이별 빈도 또는 iv) 이들의 조합을 분석하여 계산된 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하고;

b) 상기 상보성 결정부위의 C 말단으로부터 2개 내지 3개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 각 아미노산 위치별 빈도를 분석하여 계산된 빈도를 반영하여 모사한 서열을 설계하는 것이며;

상기 CDR-L3는 9개 내지 11개의 아미노산으로 이루어지고, 상기 빈도 분석은 각 길이별로 나누어 분석하며, 상기 CDR-L3 서열의 설계는 설계하고자 하는 CDR-L3의 길이와 동일한 아미노산 길이를 가지는 인간유래 숙성 항체의 상보성 결정부위 CDR-L3를 분석한 결과를 기반으로 설계하는 것인, 항체 라이브러리를 제작하는 방법.

청구항 5

제1항에 있어서, 상기 상보성 결정부위 서열 중 경쇄를 설계하는 경우에 해당 경쇄는 카파(kappa) 경쇄 또는 람다(lambda) 경쇄인 것인, 항체 라이브러리를 제작하는 방법.

청구항 6

제1항에 있어서, 상보성 결정부위 서열의 개별적인 설계시, CDR-H3에 대하여

a) 상기 상보성 결정부위의 C 말단으로부터 3개의 아미노산을 제외한 서열들은 실제 인간유래 숙성 항체의 상보성 결정부위의 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하고;

b) 상기 상보성 결정부위의 C 말단으로부터 3개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 해당 3개 아미노산 서열의 위치별 빈도를 분석하여 계산된 빈도를 반영하여 모사한 서열을 설계하는 것이며;

상기 CDR-H3는 9개 내지 20개의 아미노산으로 이루어지고, 상기 빈도 분석은 각 길이별로 나누어 분석하며, 상기 CDR-H3 서열의 설계는 설계하고자 하는 CDR-H3의 길이와 동일한 아미노산 길이를 가지는 인간유래 숙성 항체의 상보성 결정부위 CDR-H3를 분석한 결과를 기반으로 설계하는 것인, 항체 라이브러리를 제작하는 방법.

청구항 7

제1항에 있어서, 상기 상보성 결정부위 아미노산 서열을 설계한 후 설계된 서열에서 N-당화, 이성질체화, 탈아미드화, 절단, 산화가 일어날 수 있는 서열을 배제하는 단계를 추가로 포함하는 것인, 항체 라이브러리를 제작하는 방법.

청구항 8

제1항에 있어서, 상기 상보성 결정부위 아미노산 서열을 설계한 후 설계한 서열을 폴리뉴클레오티드로 역번역한 후, 번역된 폴리뉴클레오티드 5' 및 3' 말단에 각각 해당하는 인간항체 생식계열(germline) 유전자의 가변영역 골격부위 서열을 연결한 올리고뉴클레오티드 서열을 설계하는 단계를 추가로 포함하는 것인, 항체 라이브러리를 제작하는 방법.

청구항 9

제1항에 있어서, 상기 항체는 VH3-23 (Genebank accession No. Z12347), VK3-A27 (Genebank accession No. X93639), VL1g(GenBank accession No. Z73663) 또는 이들의 단편으로부터 코딩되는 아미노산 서열을 포함하는 것인, 항체 라이브러리를 제작하는 방법.

청구항 10

제1항에 있어서, 상기 항체는 IgA, IgD, IgE, IgM, IgG, Fc 단편, Fab, Fab', $F(ab')_2$ 및 Fv로 이루어진 군에서 선택된 것인, 항체 라이브러리를 제작하는 방법.

청구항 11

제3항에 있어서, 하기 1) 내지 6) 중 어느 하나 이상에 해당하는 것인, 항체 라이브러리를 제작하는 방법:

1) 상기 중쇄의 CDR-H1에 대한 생식계열 CDR 서열은 서열번호 54 내지 84의 아미노산 서열을 사용;

2) 상기 중쇄의 CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열을 사용;

3) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열을 사용;

4) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 146 내지 165의 아미노산 서열을 사용;

5) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열로써 서열번호 166 내지 189의 아미노산 서열을 사용; 및

6) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L2에 대한 생

식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열을 사용.

청구항 12

제4항에 있어서, 하기 1) 내지 2) 중 어느 하나 이상에 해당하는 것인, 항체 라이브러리를 제작하는 방법:

1) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열을 사용; 및

2) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열을 사용.

청구항 13

제11항에 있어서, 상기 각 생식계열 CDR 서열의 사용빈도는 항체서열 데이터베이스의 분석을 통해 알아낸 자연 인간항체에서의 각 생식계열 CDR 서열의 사용빈도를 모사한 것인, 항체 라이브러리를 제작하는 방법.

청구항 14

제12항에 있어서, 상기 각 생식계열 CDR 서열의 사용빈도는 항체서열 데이터베이스의 분석을 통해 알아낸 자연 인간항체에서의 각 생식계열 CDR 서열의 사용빈도를 모사한 것인, 항체 라이브러리를 제작하는 방법.

청구항 15

제1항 내지 제14항 중 어느 한 항에 따른 방법에 의해 제조된 항체 라이브러리.

발명의 설명

기 술 분 야

[0001] 본 발명은 신규 항체 라이브러리 제조방법 및 이로부터 제조된 라이브러리에 관한 것이다.

배 경 기 술

[0002] 파지 표면제시 (phage display) 기법은 박테리아를 숙주로 하는 박테리오파지(bacteriophage)를 유전적으로 조작하여 유전형(유전자)와 표현형(단백질)이 하나의 파지 입자를 통해 연결되도록 만드는 기법이다. 이 경우 유전형으로서의 유전자는 파지 유전자의 일부로 삽입되며, 표현형으로서의 단백질은 그 단백질의 유전자가 들어있는 파지 입자의 표면에 제시된다. 이러한 유전형과 표현형의 물리적 결합은 단백질 공학에서 매우 중요한 개념이며, 표현형으로 나타나는 성질에 의해 선택된 단백질 클론의 유전자를 쉽게 획득하여 복제, 증폭, 분석 및 조작할 수 있도록 해 준다.

[0003] 항체는 특별히 파지 표면제시 기법이 매우 유용하게 적용된 예이다. 매우 높은 다양성을 가지는 항체 라이브러리를 파지의 표면에 제시한 후 표면흡착된 항원과 결합시키면, 항원에 선택적으로 결합하는 항체 클론의 유전자를 획득할 수 있다. 이는 실험동물을 거치지 않고 항체를 얻는 매우 효과적인 방법이며, 특히 임의의 항원에 대한 인간항체를 얻을 수 있으므로 인체에서의 면역반응이 적은 치료용 항체 신약 등의 개발에 매우 높은 활용성을 가진다. 높은 결합친화도를 가지는 좋은 항체를 획득하기 위해서는 라이브러리의 품질이 중요하며, 특히 라이브러리의 크기와 기능적 다양성 및 클론의 품질이 중요하다.

[0004] 라이브러리의 크기는 그로부터 선택되는 항체의 품질을 결정하는 가장 중요한 요소 중 하나이다. 항체 라이브러리의 항원 결합부위는 이론적으로 어떤 특정 항원에도 경도되지 않은 무작위적 다양성을 가지며, 순전히 우연에 의해 특정 항원에 선택적으로 결합하는 항체가 무작위적 다양성 중에서 나타나게 된다. 따라서 라이브러리의 크기가 커질수록, 즉 라이브러리 내의 서로 다른 항체의 숫자가 증가할수록, 우연에 의해 높은 선택성 및 친화도

를 가지는 항체가 발견될 확률이 높아지는 것이다. 일반적으로 최소 10^8 이상의 크기가 필요한 것으로 인식되고 있으며, 많은 수의 항체 라이브러리가 10^9 - 10^{11} 정도의 크기를 가지고 있다.

[0005] 라이브러리의 기능적 다양성은, 라이브러리를 이루는 클론들 중 실제로 항체를 발현할 수 있는 클론의 비율이다. 라이브러리의 크기가 크더라도 기능적 다양성이 낮으면 실질적인 라이브러리 크기는 줄어들게 된다. 기능적 다양성이 낮아지는 이유는 대부분 라이브러리 구축 과정에서 DNA 합성 및 증폭의 오류로 인한 것이다. 항체 라이브러리는 여러 단계의 중합효소 연쇄반응(PCR)을 통해 구축되는데, 이 때 효소 및 반응의 특성상 필연적으로 낮은 빈도의 오류가 발생하게 되며 이러한 오류가 축적이 되면 최종 라이브러리의 기능적 다양성이 떨어지게 된다. 또한 특히 합성 라이브러리의 경우, 염기 합성반응의 효율성 문제 때문에 오류가 도입될 가능성이 높아지기도 한다. 이렇듯 기능적 다양성의 문제는 특히 합성 라이브러리에서 두드러지는 경향이 있으며 대부분의 합성 라이브러리는 이러한 문제점을 회피하기 위한 설계상의 고려를 해야만 한다.

[0006] 라이브러리를 이루는 개별 클론들의 품질, 즉 발현성과 안정성 및 면역유발성(immunogenicity) 등도 항체 라이브러리의 성능을 결정하는 요소 중 하나이다. 항체공학적으로 우수한 클론을 선별하기 위해서는 합성 항체라이브러리의 설계단계에서부터 이러한 요소들이 고려되어야 한다. 특히 인공적 다양성을 기존의 항체유전자에 도입하는 단계에서, 생성된 다양성이 항체 골격과의 적합성(compatibility)을 가지도록 설계하여야 하며, 자연 항체(natural antibody)의 아미노산 서열로부터의 급격한 변화는 인공합성 항체의 적합성과 안정성 등을 저해할 위험이 있다. 따라서 인공적 다양성을 설계할 때 자연적 다양성을 효율적으로 모방하는 것은 합성 항체 라이브러리의 설계와 구축에 있어 매우 중요하다. 또한 다양한 서열의 항체들로 이루어진 항체 라이브러리에는, 당화(glycosylation)이나 산화, 이성질체화, 탈아미드화 등의 원치 않는 단백질의 변형을 일으킬 수 있는 부위가 포함될 수 있으며, 이는 항체의 물성과 산업적 개발과정에 나쁜 영향을 줄 수 있다.

[0007] 동물의 체내에서 항체가 만들어질 때, 유전체 내에 존재하는 수십~수백개의 생식계열(germline) 면역글로불린 유전자의 재조합을 통해 매우 큰 다양성을 가지는 항체 서열들이 생성되며, 이들 중에서 특정 항원에 반응하는 항체가 선택된다. 선택된 항체는 초돌연변이 (hypermutation) 과정을 거쳐 항원에 대한 결합력이 향상되며, 최종적으로 숙성 (mature) 항체가 도출된다. 따라서 각 숙성 항체의 서열, 특히 항원과 결합하는 상보성 결정부위(CDR)의 서열은 생식계열 유전자 서열로부터 유래하되, 재조합 및 돌연변이 과정을 통해 생식계열 CDR 서열과는 다른 다양한 서열들이 만들어지게 된다. 합성 항체 라이브러리의 설계에서는 이러한 과정을 통해 만들어진 CDR 서열을 모사하여 자연유래 항체서열과 유사하게 제작하는 전략의 수립이 필요하다.

[0008] 자연유래 항체 제작방법은 하나의 항체를 만드는데 과한 노력과 시간이 요구됨에 따라 합성 항체 라이브러리를 이용하고자 하는 시도가 최근에 주목받고 있다. 그러나, 기존 합성 항체 라이브러리는 CDR에 해당하는 변이 구역을 무작위로 합성하여 제작하였는바 실질적으로 항체로서 기능하는 비율이 낮거나 그 효율이 보장되지 않는 문제점이 있었다.

[0009] 이러한 배경하에 본 발명자들은 효율적이고 빠른 과정을 통해 자연항체의 서열과 유사하고 우수한 물성을 가지는 항체들로 이루어진, 기능적 다양성이 높은 고품질의 합성 인간항체 라이브러리를 구축하고자 노력한 결과, 합성 항체 라이브러리를 자연유래 항체의 CDR 서열을 분석하여 서열을 구성한 다음 합성하여 자연항체의 서열과 유사하고 우수한 물성을 가지는 항체들로 이루어진 라이브러리 구축하였음을 확인함에 따라, 본 발명을 완성하였다.

발명의 내용

해결하려는 과제

[0010] 본 발명의 목적은 항체의 상보성 결정부위(complementarity Determining Regions, CDRs) 서열을 개별적으로 설계하는 단계; 및 상기 설계한 서열을 가지는 상보성 결정부위들을 포함하는 항체를 합성하여 라이브러리를 제조하는 단계를 포함하는, 항체 라이브러리를 제작하는 방법을 제공하는 것이다.

[0011] 본 발명의 다른 목적은 상기 항체 라이브러리를 제작하는 방법에 의해 제조된 항체 라이브러리를 제공하는 것이다.

과제의 해결 수단

[0012] 상기 목적을 달성하기 위하여, 본 발명의 일 양태는 항체의 상보성 결정부위(complementarity Determining

Regions, CDRs) 서열을 개별적으로 설계하는 단계; 및 상기 설계한 서열을 가지는 상보성 결정부위들을 포함하는 항체를 합성하여 라이브러리를 제조하는 단계를 포함하는, 항체 라이브러리를 제작하는 방법을 제공한다.

[0013] 하나의 구체예로서, 본 발명은 상기 항체 라이브러리를 구성하는 항체의 상보성 결정부위를 이루는 중쇄 상보성 결정부위 1(CDR-H1), 중쇄 상보성 결정부위 2(CDR-H2), 중쇄 상보성 결정부위 3(CDR-H3), 경쇄 상보성 결정부위 1(CDR-L1), 경쇄 상보성 결정부위 2(CDR-L2) 및 경쇄 상보성 결정부위 3(CDR-L3)은 다양성(polymorphism)을 가지는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0014] 한편, 상기 상보성 결정부위 서열 중 경쇄를 설계하는 경우에 해당 경쇄는 카파(kappa) 경쇄 및/또는 람다(lambda) 경쇄일 수 있다.

[0015] 다른 하나의 구체예로서, 본 발명은 상보성 결정부위 서열의 개별적인 설계시, CDR-H1, CDR-H2, CDR-L1 또는 CDR-L2에 대하여 실제 인간유래 숙성 항체의 상보성 결정부위의 i) 생식계열(germline) 면역글로불린 유전자 사용빈도, ii) 각 아미노산 위치별로 체세포 초돌연변이에 의한 각 20종의 아미노산으로의 변이 빈도, iii) 상보성 결정부위 서열의 길이별 빈도 또는 iv) 이들의 조합을 분석하여 계산된 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0016] 또 다른 하나의 구체예로서, 본 발명은 상보성 결정부위 서열의 개별적인 설계시, CDR-L3에 대하여

[0017] a) 상기 상보성 결정부위의 N 말단으로부터 7개 내지 8개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 i) 생식계열(germline) 면역글로불린 유전자 사용빈도, ii) 각 아미노산 위치별로 체세포 초돌연변이에 의한 각 20종의 아미노산으로의 변이 빈도, iii) 상보성 결정부위 CDR-L3 서열의 길이별 빈도 또는 iv) 이들의 조합을 분석하여 계산된 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하고;

[0018] b) 상기 상보성 결정부위의 C 말단으로부터 2개 내지 3개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 각 아미노산 위치별 빈도를 분석하여 계산된 빈도를 반영하여 모사한 서열을 설계하는 것이며;

[0019] 상기 CDR-L3는 9개 내지 11개의 아미노산으로 이루어지고, 상기 빈도 분석은 각 길이별로 나누어 분석하며, 상기 CDR-L3 서열의 설계는 설계하고자 하는 CDR-L3의 길이와 동일한 아미노산 길이를 가지는 인간유래 숙성 항체의 상보성 결정부위 CDR-L3를 분석한 결과를 기반으로 설계하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0020] 또 다른 하나의 구체예로서, 본 발명은 상기 상보성 결정부위 서열 중 경쇄를 설계하는 경우에 해당 경쇄는 카파(kappa) 경쇄 또는 람다(lambda) 경쇄인 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0021] 또 다른 하나의 구체예로서, 본 발명은 상보성 결정부위 서열의 개별적인 설계시, CDR-H3에 대하여

[0022] a) 상기 상보성 결정부위의 C 말단으로부터 3개의 아미노산을 제외한 서열들은 실제 인간유래 숙성 항체의 상보성 결정부위의 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하고;

[0023] b) 상기 상보성 결정부위의 C 말단으로부터 3개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 해당 3개의 아미노산 서열의 위치별 빈도를 분석하여 계산된 빈도를 반영하여 모사한 서열을 설계하는 것이며;

[0024] 상기 CDR-H3는 9개 내지 20개의 아미노산으로 이루어지고, 상기 빈도 분석은 각 길이별로 나누어 분석하며, 상기 CDR-H3 서열의 설계는 설계하고자 하는 CDR-H3의 길이와 동일한 아미노산 길이를 가지는 인간유래 숙성 항체의 상보성 결정부위 CDR-H3를 분석한 결과를 기반으로 설계하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0025] 또 다른 하나의 구체예로서, 본 발명은 상기 상보성 결정부위 아미노산 서열을 설계한 후 설계된 서열에서 N-당화, 이성질체화, 탈아미드화, 절단, 산화가 일어날 수 있는 서열을 배제하는 단계를 추가로 포함하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0026] 또 다른 하나의 구체예로서, 본 발명은 상기 상보성 결정부위 아미노산 서열을 설계한 후 설계한 서열을 폴리뉴클레오티드로 역번역한 후, 번역된 폴리뉴클레오티드 5' 및 3' 말단에 각각 해당하는 인간항체 생식계열(germline) 유전자의 가변영역 골격부위 서열을 연결한 올리고뉴클레오티드 서열을 설계하는 단계를 추가로 포함하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0027] 또 다른 하나의 구체예로서, 본 발명은 상기 항체는 VH3-23 (Genebank accession No. Z12347), VK3-A27

(Genebank accession No. X93639), VL1g(GenBank accession No. Z73663) 또는 이들의 단편으로부터 코딩되는 아미노산 서열을 포함하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0028] 또 다른 하나의 구체예로서, 본 발명은 상보성 결정부위 서열의 개별적인 설계시, 경쇄 상보성 결정부위를 설계함에 있어 카파 경쇄와 람다 경쇄에 대하여 각각 CDR을 설계하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0029] 또 다른 하나의 구체예로서, 본 발명은 카파 경쇄와 람다 경쇄에 대하여 각 경쇄 가변영역을 설계시, 카파 경쇄의 CDR은 VK3-A27의 카파 경쇄 골격부위와 연결하여 조립하고, 람다 경쇄의 CDR은 VL1g의 람다 경쇄 골격부위와 연결하여 조립하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0030] 또 다른 하나의 구체예로서, 본 발명은 상기 항체는 IgA, IgD, IgE, IgM, IgG, Fc 단편, Fab, Fab', $F(ab')_2$, scFv, 단일 가변도메인 항체 및 Fv로 이루어진 군에서 선택된 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0031] 또 다른 하나의 구체예로서, 본 발명은 하기 1) 내지 6) 중 어느 하나 이상에 해당하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다:

[0032] 1) 상기 중쇄의 CDR-H1에 대한 생식계열 CDR 서열은 서열번호 54 내지 84의 아미노산 서열을 사용;

[0033] 2) 상기 중쇄의 CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열을 사용;

[0034] 3) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열을 사용;

[0035] 4) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 146 내지 165의 아미노산 서열을 사용;

[0036] 5) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열로써 서열번호 166 내지 189의 아미노산 서열을 사용; 및

[0037] 6) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열을 사용.

[0038] 또 다른 하나의 구체예로서, 본 발명은 하기 1) 내지 2) 중 어느 하나 이상에 해당하는 것인, 항체 라이브러리를 제작하는 방법을 제공한다:

[0039] 1) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열을 사용; 및

[0040] 2) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열을 사용.

[0041] 또 다른 하나의 구체예로서, 본 발명은 상기 각 생식계열 CDR 서열의 사용빈도는 항체서열 데이터베이스의 분석을 통해 알아낸 자연인간항체에서의 각 생식계열 CDR 서열의 사용빈도를 모사한 것인, 항체 라이브러리를 제작하는 방법을 제공한다.

[0042] 본 발명의 다른 양태는, 상기 항체 라이브러리를 제작하는 방법에 의해 제조된 항체 라이브러리를 제공한다.

[0043] 하나의 구체예로서, 본 발명은 상기 항체 라이브러리는 하기 1) 내지 9) 중 하나 이상에 해당하는 것인, 항체 라이브러리를 제공한다:

[0044] 1) 중쇄의 CDR-H1에 대한 생식계열 CDR 서열은 서열번호 54 내지 84의 아미노산 서열을 사용;

[0045] 2) 중쇄의 CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열을 사용;

[0046] 3) 중쇄의 CDR-H3는 길이가 아미노산 9개 내지 20개이며, 길이별 빈도 및 각 위치에서의 아미노산 사용빈도가 자연인간항체 서열들의 CDR-H3의 길이별 빈도 및 각 위치에서의 아미노산 사용빈도를 모사한 형태;

[0047] 4) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열을 사용;

[0048] 5) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L2에 대한 생식계열

CDR 서열은 서열번호 146 내지 165의 아미노산 서열을 사용;

[0049] 6) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열을 사용;

[0050] 7) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 166 내지 189의 아미노산 서열을 사용;

[0051] 8) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열을 사용; 및

[0052] 9) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열을 사용.

[0053] 또 다른 하나의 구체예로서, 본 발명은 상기 1) 내지 2) 및 4) 내지 9)의 경우 각 생식계열 CDR 서열의 사용빈도는 항체서열 데이터베이스의 분석을 통해 알아낸 자연인간항체에서의 각 생식계열 CDR 서열의 사용빈도를 모사한, 항체 라이브러리를 제공한다.

[0054] 상기 본 발명을 설명함에 있어, 사용되는 용어를 하기 정의하였다.

[0055] **A. 정의**

[0056] "파지표면제시 (phage display)"는 M13 박테리오파지의 유전자를 조작하여 그 표면단백질 중 하나의 유전자에 외부단백질의 유전자를 융합하고, 생산된 파지의 표면단백질에 외부단백질이 융합되어 파지의 표면에 제시되는 기법이다. 단백질을 표면제시하는 경우 흔히 gIII 유전자의 5' 쪽에 외부유전자를 융합한다.

[0057] "항체"는 목표항원에 특이적으로 결합하는 단백질을 의미하며, 다클론 항체 및 단일클론항체를 모두 포함한다. 또한, 상기 용어는 키메라성 항체(예를 들면, 인간화 뮤린 항체) 및 이종결합항체(예를 들면, 양특이성 항체)와 같은 유전공학에 의해 생산된 형태를 포함한다. 특히, 항체는 경쇄(light chain)과 중쇄(heavy chain)이 각각 2개씩 모여 이루어지는 헤테로테트라머이며 각각의 사슬은 아미노산 서열이 가변적인 가변도메인(variable domain)과 일정한 서열을 가지는 고정도메인(constant domain)으로 이루어질 수 있으나, 상기 형태에 제한되지 않는다.

[0058] 본 발명에서 항체는 IgA, IgD, IgE, IgM 및 IgG를 포함하며, IgG는 아형(subtype)으로, IgG1, IgG2, IgG3 및 IgG4를 포함하며, 항체 단편을 포함할 수 있다. 항체 단편은 항원 결합 기능을 보유하고 있는 단편을 의미하며, Fc 단편, Fab, Fab', $F(ab')_2$, scFv, 단일 가변도메인 항체 및 Fv 등을 포함하며 항체의 항원 결합 형태를 포함한다. 상기 Fc 단편은 Fc 수용체와 같은 세포 표면 수용체와 결합할 수 있는 항체의 말단 부위를 의미하며, 항체의 두 개의 중쇄의 2번 또는 3번째 보존 도메인(constant domain)으로 구성된다. 상기 Fab는 경쇄 및 중쇄의 가변영역과 경쇄의 불변 영역 및 중쇄의 첫 번째 불변 영역(CH1)을 가지는 구조로 1개의 항원 결합 부위를 가진다. Fab'는 중쇄 CH1 도메인의 C 말단에 하나 이상의 시스테인 잔기를 포함하는 힌지 영역(hinge region)을 가진다는 점에서 Fab와 차이가 있다. $F(ab')_2$ 항체는 Fab'의 힌지 영역의 시스테인 잔기가 디설파이드 결합을 이루면서 생성된다. Fv(variable fragment)는 중쇄 가변부위 및 경쇄 가변부위만을 가지고 있는 최소의 항체조각을 의미한다. 이중쇄 Fv(dsFv, disulfide-stabilized variable frament)는 디설파이드 결합으로 중쇄 가변부위와 경쇄 가변부위가 연결되어 있고 단쇄 Fv(scFv, single chain variable frament)는 일반적으로 펩타이드 링커를 통하여 중쇄의 가변 영역(VH)과 경쇄의 가변 영역(VL)이 공유 결합으로 연결되어 있다. 단일 가변도메인 항체는 중쇄 또는 경쇄 가변도메인 하나만으로 이루어진 항체조각을 의미한다.

[0059] 본 발명의 항체는 재조합 단일 사슬 Fv 단편 (scFv)을 포함하며, 이가(bivalent) 또는 양특이성 분자, 디아바디(Diabody), 트리아바디(Triabody) 및 테트라바디 (Tetrabody)를 모두 제한없이 포함한다.

[0060] "상보성 결정부위" (Complementarity determining region, CDR)는 가변도메인의 경쇄와 중쇄에 각각 3개씩 존재하는 부위이며 가변도메인 중에서도 아미노산 서열의 가변성이 특히 높은 부분으로 이러한 높은 가변성에 의해 다양한 항원에 대해 특이적 항체가 찾아질 수 있다. 중쇄의 상보성 결정부위 3개를 아미노 말단부터 카르복실 말단방향으로 차례로 CDR-H1, CDR-H2, CDR-H3라 하며, 경쇄의 상보성 결정부위 3개를 아미노 말단부터 카르복실 말단방향으로 차례로 CDR-L1, CDR-L2, CDR-L3라 한다. 하나의 항체에서 이 여섯 개의 상보성 결정부위가 모여 항원 결합부위를 형성한다.

[0061] "골격 부위"는 가변도메인 서열 중 상보성 결정부위를 제외한 나머지 부위를 말하며, 상보성 결정부위에 비해

서열의 가변성 및 다양성이 낮고 일반적으로 항원-항체 반응에 참여하지 않는 부위를 의미한다.

[0062] "면역글로불린"은 항체와 구조적 특성이 동일하며, 항원특이성이 없는 항체유사분자와 항체를 포함하는 개념이다.

[0063] "생식계열 면역글로불린 유전자"는 동물의 생식세포에 존재하는, B 세포로 분화되어 면역글로불린 유전자의 재조합이나 체세포 초돌연변이(somatic hypermutation) 과정을 거치지 않은 항체 유전자이며, 그 수는 동물의 종에 따라 다르나 보통 수십 내지 수백개이다.

[0064] "숙성 항체"는 B 세포 분화를 통해 생식계열 면역글로불린 유전자들의 재조합, 또는 재조합과 체세포 초돌연변이 과정을 거쳐 만들어진 항체 유전자로부터 발현된 항체 단백질이다.

[0065] "단일사슬단편항체(single-chain Fv, scFv)"는 항체의 경쇄 및 중쇄의 가변도메인을 15개 내외의 아미노산이 연결된 펩타이드 사슬로 이루어진 연결부위(linker)로 연결한 단백질이다. 경쇄가변도메인 - 연결부위 - 중쇄가변도메인, 또는 중쇄가변도메인 - 연결부위 - 경쇄가변도메인의 순서 모두 가능하며, 원 항체와 동일 혹은 유사한 항원특이성을 가진다. 연결부위는 주로 글리신(glycine)과 세린(serine)으로 이루어진 친수성의 유연한 펩타이드 사슬로서 "$(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$"의 15개의 아미노산 서열 혹은 유사한 서열이 주로 사용될 수 있다.

[0066] "항체 라이브러리"는 서로 다른 서열을 가지는 다양한 항체 유전자들의 집합이다. 항체라이브러리로부터 임의의 항원에 대해 특이적 항체를 분리하기 위해서는 매우 높은 다양성이 요구되며, 일반적으로 10^9 내지 10^{11} 개의 서로 다른 항체 클론들로 이루어진 라이브러리가 구축되어 사용된다. 이러한 항체라이브러리를 이루는 항체 유전자는 파지미드(phagemid) 벡터에 클로닝되어 대장균에 트랜스폼 된다.

[0067] "파지미드(phagemid)" 벡터는 파지복제시작점(phage origin of replication)을 가지는 플라스미드 DNA이다. 통상적으로 항생제내성유전자를 선택 마커(selection marker)로 가지고 있다. 파지표면제시에 사용되는 파지미드 벡터의 경우 M13 파지의 gIII 유전자 혹은 그 일부가 포함되어 있으며, 라이브러리 유전자는 gIII 유전자의 5' 말단에 라이게이션(ligation)되어 대장균에서 융합단백질로서 발현된다.

[0068] "보조파지(helper phage)"는 파지미드가 파지 입자로 조립되도록 필요한 유전정보를 제공하는 파지이다. 파지미드에는 파지 유전자 중 gIII 혹은 그 일부만이 존재하므로 파지미드로 형질전환된 대장균을 보조파지로 감염시켜 나머지 파지 유전자를 공급하게 된다. M13K07 혹은 VCSM13 등의 종류가 있으며 대부분 카나마이신(kanamycin) 등 항생제 내성 유전자를 포함하여 보조파지에 감염된 대장균을 선택할 수 있도록 하고 있다. 또한 조립신호(packaging signal)에 결함이 있으므로 보조파지 유전자보다 파지미드 유전자가 선별적으로 파지입자 속으로 조립된다.

[0069] "패닝"은 파지 표면에 디스플레이된 항체 등 단백질의 라이브러리로부터 특정 분자에 결합하는 클론만을 선택적으로 증폭시키는 과정을 뜻한다. 표면에 고정된 표적분자에 파지 라이브러리를 가하여 결합을 유도하고, 결합하지 않은 파지 클론을 세척하여 제거한 후, 결합된 파지 클론들만을 용출하여 다시 대장균 숙주에 감염시키고 보조 파지를 이용하여 표적 결합 파지 클론들을 증폭하는 절차를 거친다. 대부분 이러한 과정을 3 내지 4회 혹은 그 이상 반복하여 결합 클론의 비율을 최대한 높이게 된다.

[0070] **B. 라이브러리의 설계, 구축 및 검증**

[0071] 상기 목적을 달성하기 위한 하나의 양태로서, 본 발명은 항체의 상보성 결정부위(complementarity Determining Regions, CDRs) 서열을 개별적으로 설계하는 단계; 및 상기 설계한 서열을 가지는 상보성 결정부위들을 포함하는 항체를 합성하여 라이브러리를 제조하는 단계를 포함하는, 항체 라이브러리를 제작하는 방법을 제공한다.

[0072] 본 발명에 기술된 방법을 이용해 인간항체 라이브러리를 구축하고 그로부터 임의의 항원에 대한 인간항체를 얻을 수 있다. 항체 라이브러리는 골수, 비장, 혈액 등에 포함된 B 세포로부터 그 다양성을 얻는 방법과 인공적인 설계와 합성을 통해 다양성을 얻는 방법으로 구축될 수 있으며 본 발명은 합성 인간항체 라이브러리의 구축 및 검증을 제공한다.

[0073] 파지 표면제시 항체라이브러리는 면역 글로불린 분자의 일부인 Fab 혹은 scFv 단편의 형태로 구축되어 진다. 이들 단편은 150 kDa 크기의 면역 글로불린보다 작기 때문에 단백질 공학적인 조작의 효율성을 기할 수 있으며, 면역 글로불린 분자와 동일한 항원선택성을 가진다. 본 발명에서는 25 kDa 크기를 가지는 scFv 단편을 이용한 라이브러리를 구축하였으며, 구체적으로 면역 글로불린의 VH 도메인과 VL 도메인을 15개의 아미노산으로 이루어진 사슬 즉 $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$ 로 연결한 하나의 폴리펩티드 사슬이다.

[0074] 합성 라이브러리의 구축을 위해서는 라이브러리 서열의 설계가 선행되어야 한다. 비교적 높은 골격 다양성을 가지는 B세포 유래 항체 라이브러리 혹은 자연항체 라이브러리(natural antibody library)와 달리, 합성 항체라이브러리는 한 개 혹은 제한된 개수의 골격 서열을 기반으로 구축된다.

[0075] 본 발명의 구체적인 일 실시예에서 제작된 항체 라이브러리는 두 개의 골격을 가진다. 구체적으로 라이브러리를 이루는 모든 클론들이 인간 면역글로불린의 VH3-23 유전자와 VK3-A27 유전자가 연결부위로 이어진 scFv, 혹은 VH3-23 유전자와 VL1g 유전자가 연결부위로 이어진 scFv를 골격으로 가지도록 구축하였으며, 이 중 상보성 결정부위에 인공적 다양성을 도입하여 라이브러리를 구축하였다. 즉 상기 라이브러리의 골격에 다양한 상보성 결정부위(CDR) 서열을 삽입하여 scFv 항체 라이브러리를 제작하였다.

[0076] 구체적으로 본 발명의 항체 라이브러리는 항체 라이브러리를 구성하는 항체의 상보성 결정부위를 이루는 중쇄 상보성 결정부위 1(CDR-H1), 중쇄 상보성 결정부위 2(CDR-H2), 중쇄 상보성 결정부위 3(CDR-H3), 경쇄 상보성 결정부위 1(CDR-L1), 경쇄 상보성 결정부위 2(CDR-L2) 및 경쇄 상보성 결정부위 3(CDR-L3)은 다양성(polymorphism)을 가지는 것일 수 있다.

[0077] 특히, 본 발명에서 상보성 결정부위 서열의 개별적인 설계시, CDR-H1, CDR-H2, CDR-L1 또는 CDR-L2에 대하여 실제 인간유래 숙성 항체의 상보성 결정부위의 i) 생식계열(germline) 면역글로불린 유전자 사용빈도, ii) 각 아미노산 위치별로 체세포 초돌연변이에 의한 각 20종의 아미노산으로의 변이 빈도, iii) 상보성 결정부위 서열의 길이별 빈도 또는 iv) 이들의 조합을 분석하여 계산된 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하는 것일 수 있다.

[0078] 한편, 본 발명은 상보성 결정부위 서열의 개별적인 설계시, CDR-L3에 대하여 a) 상기 CDR-L3 상보성 결정부위의 N 말단으로부터 7개 내지 8개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 i) 생식계열(germline) 면역글로불린 유전자 사용빈도, ii) 각 아미노산 위치별로 체세포 초돌연변이에 의한 각 20종의 아미노산으로의 변이 빈도, iii) 상보성 결정부위 CDR-L3 서열의 길이별 빈도 또는 iv) 이들의 조합을 분석하여 계산된 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하고;

[0079] b) 상기 CDR-L3 상보성 결정부위의 C 말단으로부터 2개 내지 3개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 각 아미노산 위치별 빈도를 분석하여 계산된 빈도를 반영하여 모사한 서열을 설계하는 것이며;

[0080] 상기 CDR-L3는 9개 내지 11개의 아미노산으로 이루어지고, 상기 빈도 분석은 각 길이별로 나누어 분석하며, 상기 CDR-L3 서열의 설계는 설계하고자 하는 CDR-L3의 길이와 동일한 아미노산 길이를 가지는 인간유래 숙성 항체의 상보성 결정부위 CDR-L3를 분석한 결과를 기반으로 설계하는 것일 수 있다.

[0081] 아울러, 본 발명은 상보성 결정부위 서열의 개별적인 설계시, CDR-H3에 대하여 a) 상기 상보성 결정부위의 C 말단으로부터 3개의 아미노산을 제외한 서열들은 실제 인간유래 숙성 항체의 상보성 결정부위의 각 아미노산 위치별 빈도를 반영하여 모사한 서열을 설계하고;

[0082] b) 상기 상보성 결정부위의 C 말단으로부터 3개의 아미노산 서열은 실제 인간유래 숙성 항체의 상보성 결정부위의 해당 3개 아미노산 서열의 위치별 빈도를 분석하여 계산된 빈도를 반영하여 모사한 서열을 설계하는 것이며;

[0083] 상기 CDR-H3는 9개 내지 20개의 아미노산으로 이루어지고, 상기 빈도 분석은 각 길이별로 나누어 분석하며, 상기 CDR-H3 서열의 설계는 설계하고자 하는 CDR-H3의 길이와 동일한 아미노산 길이를 가지는 인간유래 숙성 항체의 상보성 결정부위 CDR-H3를 분석한 결과를 기반으로 설계하는 것일 수 있다.

[0084] 본 발명에서 본 발명의 라이브러리 중쇄 CDR 설계에 사용하는 생식계열 CDR 서열로써, CDR-H1에 대한 생식계열 CDR 서열은 서열번호 54 내지 84의 아미노산 서열이며, CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열일 수 있다. 한편, CDR-H3의 설계에 있어서는 생식계열 CDR 서열을 사용하지 않고, 그 길이를 아미노산 9개 내지 20개로 하여, 길이별 빈도 및 각 위치에서의 아미노산 사용빈도가 자연인간항체 서열들의 CDR-H3의 길이별 빈도 및 각 위치에서의 아미노산 사용빈도를 모사하도록 할 수 있다.

[0085] 또한, 본 발명의 라이브러리 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로, 람다 경쇄 또는 카파 경쇄를 가질 수 있으며, 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열이고, 카파 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 146 내지 165의 아미노산 서열일 수 있다. 한편, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열일 수 있다.

[0086] 한편, 본 발명의 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 166 내지 189의 아미노산 서열이고,

람다 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열일 수 있다. 한편, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열일 수 있다.

[0087] 상기 항체 라이브러리를 제작하는 방법, 특히 본 발명의 항체 라이브러리에 포함되는 CDR-H1, CDR-H2, CDR-L1 및 CDR-L2에 대한 설계를 하는 방법은 하기 1) 내지 6) 중 어느 하나 이상에 해당할 수 있다:

[0088] 1) 상기 중쇄의 CDR-H1에 대한 생식계열 CDR 서열은 서열번호 54 내지 84의 아미노산 서열을 사용;

[0089] 2) 상기 중쇄의 CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열을 사용;

[0090] 3) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열을 사용;

[0091] 4) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 146 내지 165의 아미노산 서열을 사용;

[0092] 5) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열로써 서열번호 166 내지 189의 아미노산 서열을 사용; 및

[0093] 6) 상기 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열을 사용.

[0094] 또한, 상기 항체 라이브러리를 제작하는 방법, 특히 본 발명의 항체 라이브러리에 포함되는 CDR-L3에 대한 설계를 하는 방법은 하기 1) 내지 2) 중 어느 하나 이상에 해당할 수 있다:

[0095] 1) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열을 사용; 및

[0096] 2) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열을 사용.

[0097] 상기 각 생식계열 CDR 서열의 사용빈도는 항체서열 데이터베이스의 분석을 통해 알아낸 자연인간항체에서의 각 생식계열 CDR 서열의 사용빈도를 모사한 것일 수 있다.

[0098] 본 발명의 구체적인 실시예에서 본 발명의 라이브러리 중쇄 CDR 설계에 사용된 생식계열 CDR 서열은, CDR-H1에 대한 생식계열 CDR 서열은 서열번호 54 내지 84의 아미노산 서열이며, CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열이었다. 중쇄의 CDR-H3의 설계에 있어 그 길이를 아미노산 9개 내지 20개로 하여, 길이별 빈도 및 각 위치에서의 아미노산 사용빈도가 자연인간항체 서열들의 CDR-H3의 길이별 빈도 및 각 위치에서의 아미노산 사용빈도를 모사하도록 설계하였다.

[0099] 또한, 본 발명의 라이브러리 경쇄 CDR 설계에 사용된 생식계열 CDR 서열로, 람다 경쇄 또는 카파 경쇄 중에서 먼저 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열이고, 카파 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 146 내지 165의 아미노산 서열이었다. 한편, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열이었다.

[0100] 한편, 본 발명의 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 166 내지 189의 아미노산 서열이고, 람다 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열이었다. 한편, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열이었다.

[0101] 본 발명의 항체 라이브러리 제조방법에서 상보성 결정부위 서열의 개별적인 설계시, 경쇄 상보성 결정부위를 설계함에 있어 카파 경쇄와 람다 경쇄에 대하여 각각 CDR을 설계할 수 있으며, 카파 경쇄와 람다 경쇄에 대하여 각 경쇄 가변영역을 설계시, 카파 경쇄의 CDR은 VK3-A27의 카파 경쇄 골격부위와 연결하여 조립하고, 람다 경쇄의 CDR은 VL1g의 람다 경쇄 골격부위와 연결하여 조립하는 것일 수 있다.

[0102] 본 발명에서 "모사"란 아미노산 서열 등의 발현 빈도 또는 변이 빈도 등을 반영하여 랜덤으로 시뮬레이션하여 서열을 설계하는 것을 의미하며, 자연항체, 특히 인간유래 숙성 항체의 아미노산의 발현 빈도 또는 변이 빈도를 모사한다는 의미를 내포하고 있다.

[0103] 본 발명의 구체적인 일 실시예에서 상보성 결정부위의 다양성은 알려진 인간항체들의 CDR 서열의 특성을 분석하고 시뮬레이션하여 인간항체 CDR과 유사한 서열적 다양성을 가지도록 설계하였다. 구체적으로, 먼저 IMGT 데이터베이스 (http://imgt.org)로부터 8,846개의 인간 면역글로불린 중쇄 가변부위 (VH), 3,110개의 카파 경쇄 가

변부위 (Vκ), 2,440개의 람다 경쇄 가변부위 (Vλ) 서열을 다운로드하여 CDR 서열을 추출하였다. 다음으로, V-base (http://www2.mrc-lmb.cam.ac.uk/vbase/alignments2.php)에 있는 인간항체 생식계열(germline) 면역글로불린 유전자의 CDR 서열들과 상기 IMGT 데이터베이스에서 추출된 숙성(mature) CDR 서열들을 비교 분석하여, i) 각 숙성 CDR서열에 가장 가까운 생식계열 CDR 서열을 찾고, 이로부터 각 숙성 CDR에 일어난 돌연변이의 위치, 종류, 빈도를 확인하고 ii) 각 생식계열 CDR이 숙성 인간항체에서 사용되는 빈도를 계산하여 이용하였다.

[0104] 먼저, 재조합이 일어나지 않고 체세포 초돌연변이만 일어나는 중쇄 및 경쇄의 CDR1과 CDR2에 대해서는 컴퓨터를 이용한 시뮬레이션을 통하여, 인간 생식계열 CDR 서열에 가상의 돌연변이를 도입하여 인간유래 숙성 항체의 CDR과 유사한 서열 및 생식계열 CDR 사용빈도를 가지도록 CDR 서열들을 설계하였다. 각 CDR 별로 1,500개의 시뮬레이션된 서열들이 설계되었다.

[0105] 다음으로, 재조합 및 체세포 초돌연변이가 일어나는 경쇄의 CDR3에 대하여, 카파 경쇄의 경우 첫 7개의 아미노산에 대해서 CDR1 및 CDR2에서와 동일한 작업을 수행하고, 끝의 2개 혹은 3개의 아미노산 서열은 숙성 인간항체의 카파 경쇄 CDR3에서의 서열빈도를 모사하여 총 9개 혹은 10개 아미노산으로 이루어진 1,500개의 서열들을 설계하였다. 람다 경쇄의 경우 첫 7개 혹은 8개의 아미노산에 대하여 CDR1 및 CDR2에서와 동일한 작업을 수행하고, 끝의 2개 혹은 3개의 아미노산 서열은 숙성 인간항체의 람다 경쇄 CDR3에서의 서열빈도를 모사하여 총 9개, 10개, 혹은 11개 아미노산으로 이루어진 1,500개의 서열들을 설계하였다.

[0106] 마지막으로, 재조합 및 체세포 초돌연변이가 일어나는 중쇄의 CDR3에서는 VDJ 재조합 및 접합부위 유연성(junctional flexibility), P-첨가 (P-addition), N-첨가 (N-addition) 등의 기전에 의하여 생식계열 서열의 확인이 어려운 경우가 많으므로 다른 CDR 같은 분석이 곤란함에 따라, 우선 길이별로 9개에서 20개의 아미노산 길이의 숙성 인간항체 CDR-H3 서열을 나누어 분석하여, 각 길이의 각 위치에서의 아미노산 사용빈도를 분석하여 시뮬레이션 하였다. 다만 마지막 3개의 아미노산 서열은 J 유전자로부터 유래하는 경우가 대부분이므로 각 CDR-H3 길이별로 자주 사용되는 최대 8개의 삼(3)아미노산 서열을 숙성항체에서의 사용빈도를 고려하여 설계하였다 (도 1).

[0107] 특히, 중쇄의 CDR-H3를 설계하는데 있어 자연인간항체의 CDR-H3의 길이별 빈도 및 각 위치에서의 아미노산 사용빈도는 항체서열 데이터베이스의 분석을 통해 알아낼 수 있으며, 대표적으로 IMGT database (http://www.imgt.org)에 취합된 항체서열을 분석할 수 있다.

[0108] 한편, 본 발명의 항체 라이브러리를 제조하는 방법은 상기 상보성 결정부위 아미노산 서열을 설계한 후 설계된 서열에서 N-당화, 이성질체화, 탈아미드화, 절단, 산화가 일어날 수 있는 서열을 배제하는 단계를 추가로 포함하는 것일 수 있다.

[0109] 본 발명의 구체적인 일 실시예에서는, CDR 서열의 설계시, 다음의 번역후 변형(post translational modification) 서열을 포함하는 것들을 배제하였다. 단백질의 번역후 변형은 단백질의 기능 및 물성에 영향을 줄 수 있으므로 항체 라이브러리의 설계시 가능한 한 배제하는 것이 유리하다. 상기 번역 후 변형 서열에는, N-당화 (N-glycosylation) 서열, 즉 Asp-Xaa-Ser/Thr (Asp: 아스파라긴산, Xaa: 시스테인을 제외한 임의의 19종의 아미노산, Ser/Thr: 세린 혹은 트레오닌) 서열, 이성질체화 (isomerization) 서열, 탈아미드화 (deamidation) 서열, 절단(cleavage) 서열 및 산화(oxidation) 서열을 제외하였다.

[0110] 한편, 본 발명의 항체 라이브러리를 제조하는 방법은 상기 상보성 결정부위 아미노산 서열을 설계한 후 설계한 서열을 폴리뉴클레오티드로 역번역한 후, 번역된 폴리뉴클레오티드 5' 및 3' 말단에 각각 해당하는 인간항체 생식계열(germline) 유전자의 가변영역 골격부위 서열을 연결한 올리고뉴클레오티드 서열을 설계하는 단계를 추가로 포함하는 것일 수 있다.

[0111] 본 발명의 구체적인 일 실시예에서는 상기 설계된 CDR의 아미노산 서열은 뉴클레오티드 서열로 역번역하였으며 CDR 서열의 양 옆에 항체 가변부위의 골격부위 서열을 더하여 100개의 뉴클레오티드로 이루어진 올리고뉴클레오티드 서열들을 최종설계하였다. 총 19,836개의 서열을 설계하고, 이를 배열 합성(array synthesis)법을 통하여 올리고뉴클레오티드 혼합물의 형태로 합성하였다 (미국 텍사스주 휴스턴 소재 LC Sciences사).

[0112] 본 발명의 항체 라이브러리를 제조하는 방법에서 항체는 항체 라이브러리 제작의 골격으로 VH3-23, VK3-A27, VL1g 또는 이들의 단편으로부터 코딩되는 아미노산 서열을 포함하는 것일 수 있다. 상기 VH3-23는 Genebank accession No. Z12347에 해당하는 것일 수 있고, VK3-A27는 Genebank accession No. X93639에 해당하는 것일 수 있으며, VL1g는 GenBank accession No. Z73663에 해당하는 것일 수 있다.

[0113] 구체적인 일 실시예로 본 발명에서는 항체 라이브러리 제작의 골격으로 사용하기 위하여 코돈 최적화된 scFv 유

전자들을 합성하였으며 (미국 뉴저지 주 피스카타웨이 시 소재 Genscript 사), 이 유전자들은 생식계열 유전자인 VH3-23와 VK3-A27, 또는 VH3-23과 VL1g가 $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$의 연결부위로 연결된 scFv 유전자들로 pUC57 벡터에 클로닝되어있으며 파지미드 벡터로의 클로닝을 위해 pComb3X 벡터와 호환되는 두 개의 SfiI 제한효소 부위가 포함된다. 라이브러리 제작의 편의를 위하여, pUC57 벡터에 클로닝된 이 유전자들을 pComb3X 벡터로도 클로닝하여 PCR의 주형으로 사용하였다. 골격으로 사용하기 위한 용도이므로 이 유전자들은 서열적 다양성이 없는 단일서열들이며, 포유동물세포에서 발현이 향상되도록 코돈이 최적화되었으나 PCR 과정에서의 비특이적 어닐링(annealing)을 방지하기 위하여 아미노산 서열에 변화가 없는 범위에서 일부 최적화 DNA서열에 변화를 가하였다(서열번호 19 및 서열번호 20).

[0114] 본 발명의 항체 라이브러리를 제조하는 방법에서 항체는 IgA, IgD, IgE, IgM, IgG, Fc 단편, Fab, Fab', $F(ab')_2$, scFv, 단일 가변도메인 항체 및 Fv로 이루어진 군에서 선택된 것일 수 있다.

[0115] 특히, 본 발명의 구체적인 일 실시예에서는 scFv 항체 라이브러리를 제작하였으며, 제작한 라이브러리로부터 scFv의 발현을 확인하였다 (도 2). 항체 라이브러리의 구축을 검증하기 위하여, 또한 scFv의 서열을 분석하였다. 라이브러리로부터 임의의 클론들을 선택하고 scFv 유전자를 PCR 기법으로 증폭한 후, DNA를 정제하여 서열분석을 의뢰하였다. 서열분석 결과, 설계한 대로 다양한 서열과 길이를 가지는 CDR들이 라이브러리를 이루는 scFv 클론들에 도입되어 있음을 확인할 수 있다(도 3). 구체적인 서열분석 결과, 18개의 scFv 서열들 중 절단(아스파라긴산-프롤린)서열이 PCR 오류로 도입된 1건을 제외하면 배제하고자 하였던 번역후 변형 서열들이 존재하지 않는 것을 확인하였다. 이에 비하여 인간 유래 숙성항체의 각 CDR에서의 이러한 번역후 변형을 유발할 수 있는 서열들의 존재비율을 분석한 결과 중쇄의 CDR-H1에서 5.7%, CDR-H2에서 39.7%, CDR-H3에서 34.5%, 카파 경쇄의 CDR-L1에서 12.6%, CDR-L2에서 0.6%, CDR-L3에서 15.9%, 람다 경쇄의 CDR-L1에서 8.3%, CDR-L2에서 6.4%, CDR-L3에서 24.0% 발견되었다. 이로부터 설계의도대로 번역후 변형이 최소화될 수 있는 항체 라이브러리가 구축되었음을 확인하였다.

[0116] 또 하나의 양태로서, 본 발명은 본 발명의 항체 라이브러리를 제작하는 방법에 의해 제조된 항체 라이브러리를 제공한다.

[0117] 본 발명의 항체 라이브러리는 CDR-H1에 대한 생식계열 CDR 서열로써 서열번호 54 내지 84의 아미노산 서열을 사용할 수 있으며, CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열을 사용할 수 있다. 한편, CDR-H3의 설계에 있어 그 길이를 아미노산 9개 내지 20개로 하여, 길이별 빈도 및 각 위치에서의 아미노산 사용빈도가 자연인간항체 서열들의 CDR-H3의 길이별 빈도 및 각 위치에서의 아미노산 사용빈도를 모사하여 설계된 것일 수 있다.

[0118] 본 발명의 항체 라이브러리는 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄 또는 카파 경쇄를 사용할 수 있으며, 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열을 사용할 수 있고, 카파 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 146 내지 165의 아미노산 서열을 사용하여 제작된 것일 수 있다. 아울러, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열을 사용하여 제작된 것일 수 있다.

[0119] 한편, 본 발명의 항체 라이브러리는 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열로써 서열번호 166 내지 189의 아미노산 서열을 사용할 수 있고, 람다 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열을 사용하여 제작된 것일 수 있다. 아울러, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열을 사용하여 제작된 것일 수 있다.

[0120] 구체적으로, 본 발명의 항체 라이브러리는 하기 1) 내지 9) 중 하나 이상에 해당하는 것인, 항체 라이브러리일 수 있다:

[0121] 1) 중쇄의 CDR-H1에 대한 생식계열 CDR 서열은 서열번호 54 내지 84의 아미노산 서열을 사용;

[0122] 2) 중쇄의 CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열을 사용;

[0123] 3) 중쇄의 CDR-H3는 길이가 아미노산 9개 내지 20개이며, 길이별 빈도 및 각 위치에서의 아미노산 사용빈도가 자연인간항체 서열들의 CDR-H3의 길이별 빈도 및 각 위치에서의 아미노산 사용빈도를 모사한 형태;

[0124] 4) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열을 사용;

[0125] 5) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 146 내지 165의 아미노산 서열을 사용;

[0126] 6) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 카파 경쇄를 사용하고, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열을 사용;

[0127] 7) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 166 내지 189의 아미노산 서열을 사용;

[0128] 8) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열을 사용; 및

[0129] 9) 경쇄 CDR 설계에 사용하는 생식계열 CDR 서열로 람다 경쇄를 사용하고, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열을 사용.

[0130] 특히, 본 발명은 상기 1) 내지 2) 및 4) 내지 9)의 경우 각 생식계열 CDR 서열의 사용빈도는 항체서열 데이터베이스의 분석을 통해 알아낸 자연인간항체에서의 각 생식계열 CDR 서열의 사용빈도를 모사한, 항체 라이브러리일 수 있다.

[0131] 본 발명의 구체적인 일 실시예에서는 형질전환된 대장균 라이브러리를 배양하고 VCSM13 보조 파지(helper phage)로 감염시켜 scFv 클론이 표면제시된 항체파지 라이브러리를 얻었다. 이 라이브러리를 통하여 라이소자임(Hen egg white lysozyme, HEWL) 및 AIMP1(101-192) 항원에 대하여 항원특이적 항체를 선별하는 패닝 실험을 통해 항체파지 라이브러리의 기능성을 검증한 결과, HEWL 항원에 대하여 188개의 클론들을 ELISA로 스크리닝하여 배경신호 대비 3배 이상의 결합신호를 보이는 ELISA 양성 클론들을 150개 확인하였으며 이 중 16개를 서열분석하여 5개의 고유 서열을 발견하였다. 또한, AIMP1(101-192)에 대하여 94개의 클론글을 ELISA로 스크리닝하여 배경신호 대비 3배 이상의 결합신호를 보이는 ELISA양성 클론들을 18개 확인하였으며 이 중 5개를 서열분석하여 2개의 고유 서열을 발견하였다.

[0132] 본 항체 라이브러리 설계 및 구축기술을 이용하여, 복수의 인간 생식계열 면역글로불린 가변부위 유전자를 골격으로 하는 개선된 라이브러리를 제작할 수 있다.

[0133] 즉, 본 발명에 기술된 항체 라이브러리가 중쇄, 카파 경쇄, 람다 경쇄당 각 1개씩의 골격 (각각 VH3-23, VK3-A27, VL1g)을 사용하면서 여기에 서로 다른 다양한 생식계열 면역글로불린 유전자로부터 유래하는 CDR에 돌연변이를 도입한 합성 CDR을 삽입한 설계방식을 따라 제조될 수 있을 뿐만 아니라, 개선된 설계를 통하여 복수의 다양한 생식계열 면역글로불린 가변부위의 골격 유전자를 기반으로 하여 여기에 각 해당 골격과 동일한 유전자군의 생식계열 면역글로불린 유전자로부터 유래하는 CDR에 돌연변이를 도입한 합성 CDR을 각각 도입함으로써 골격과 CDR의 연결 조합이 자연 인간유래 항체와 유사하게 제작된 라이브러리를 얻을 수 있다.

[0134] 또한 항체 라이브러리 제작과정에서 CDR의 길이에 따른 합성 및 증폭효율의 차이로 인하여 짧은 CDR이 선호되는 편향성을 해결하기 위하여, CDR 길이별로 서브 라이브러리를 각각 제작한 후 자연 인간항체에서 발견되는 CDR 길이분포와 유사한 비율로 혼합하여 최종 라이브러리를 제작하는 것이 가능하다.

[0135] 또 하나의 양태로서, 본 발명은 본 발명의 항체 라이브러리를 제작하는 방법에 의해 제조된 항체 라이브러리로부터 제작된 항체를 제공한다.

[0136] 본 발명에서 항체, 항체 라이브러리 등은 상기 설명한 바와 같다.

발명의 효과

[0137] 본 발명에 의해 제조된 항체 라이브러리는 다수의 항원에 대해 우수한 물성을 가지는 항체들을 포함하여, 기능적 다양성을 가질 뿐 아니라 고유서열을 다수 포함하는 등 항체 라이브러리로써 유용하게 사용될 수 있다.

도면의 간단한 설명

[0138] 도 1은 CDR 부위의 설계에 대한 개념을 나타낸 모식도이다.

도 2는 6개의 비조합 다양화 CDR(non-combinatorially diversified CDRs)을 이용한 scFv 라이브러리 제조방법을 나타낸 모식도이다. 설계된 CDR 서열이 있는 올리고뉴클레오티드의 풀(pool)이 배열 생성되고 PCR을 통해 증폭된다. 각 CDR 위치에 대한 단일 CDR 라이브러리(6개의 CDR 중 하나만 다양화된 scFv 라이브러리)를 제조하고,

CDR 레퍼토리가 in-frame 서열인 것을 확인하기 위한 scFv의 C 말단 HA-태그에 결합하는 항-HA 항체를 이용하여 패닝하였다.

도 3은 본 발명의 항체 라이브러리에 scFv가 발현되는지 여부를 확인하기 위하여, 라이브러리로부터 셀렉션 전 scFv 클론을 무작위로 선택하여 HA 태그를 이용하여 발현을 확인한 도면이다. 구체적으로 92개의 클론을 무작위로 선택하여 배양 및 IPTG 처리하고 용출물에서 전장 scFV 발현 여부를 항-HA 항체로 확인한 결과이며, 총 92개 중 58개에서 발현이 확인되었다 (63%).

도 4는 본 발명의 항체 라이브러리 제조 단계를 도식화한 모식도이다.

도 5a는 본 발명의 항체 라이브러리 중 특히 OL (odd-lambda) 및 EL (even-lambda) 라이브러리로부터 임의의 클론들을 선택하고 scFv 유전자를 PCR 기법으로 증폭한 후, DNA를 정제하여 서열분석한 결과를 개시하고 있는 도이다.

도 5b는 본 발명의 항체 라이브러리 중 특히 OK (odd-kappa) 및 EK (even-kappa) 라이브러리로부터 임의의 클론들을 선택하고 scFv 유전자를 PCR 기법으로 증폭한 후, DNA를 정제하여 서열분석한 결과를 개시하고 있는 도이다.

도 6은 라이브러리의 설계 및 실제 CDR 레퍼토리의 고유 CDR 서열의 빈도에 관한 그래프이다. 실제 제조된 scFv 라이브러리의 NGS 분석으로부터 발견한 CDR 서열과 설계된 CDR 레파토리에서 각 고유 CDR 서열이 발생한 빈도를 XY 분포 그래프로 표시하였다. 그래프 내의 각 점은 고유 CDR 서열을 의미한다. 한편, CDR-H3 서열은 대부분의 서열이 설계된 레퍼토리에서 오직 한번만 발생하기 때문에 분석하지 않았다.

도 7은 제조된 라이브러리의 다양한 도메인 서열 중복도를 나타낸 도이다. Illumina MiSeq 플랫폼 상의 300 bp paired-end 서열분석을 통해 셀렉션되지 않은 라이브러리의 다양한 도메인 서열을 수득하고, 다양한 도메인 서열에 대한 중복회수(n)를 분석하였다. V_H 및 V_λ의 약 98% 및 V_κ의 약 88.5% 서열이 단 한번 발견되었다 (n = 1).

도 8은 설계된 CDR 및 실제 CDR의 길이 분포를 나타낸 도이다. (a) CDR-H2, L1 (kappa 및 lambda), 및 L3 (kappa 및 lambda)은 다양한 길이의 서열을 포함한다. 설계된 레파토리에서의 길이 분포와 차세대 서열분석 (Next generation sequencing, NGS)에서 분석된 길이 분포의 비교를 통하여, 짧은 CDR이 실제 라이브러리에서 선호됨을 시사하는 결과를 얻었다. (b) 짧은 CDR에 대한 선호성은 다른 CDR에 비해 넓은 길이 다양성 범위를 가지는 CDR-H3에서 특히 분명하게 보여졌다. (a) 및 (b) 둘 모두에서, 청색 막대("Designed")는 설계 레파토리에서 각 CDR 길이의 빈도수를 보여주며, 주황색 막대("Found")는 제조된 라이브러리의 차세대 서열분석(Next generation sequencing, NGS)을 통해 확인된 각 CDR 길이의 빈도수를 보여준다.

도 9는 CDR-H3의 아미노산 분포를 나타낸 도이다. 자연 인간 항체(N), 설계된 레파토리 (D), 및 실제 제조된 라이브러리 (L)의 CDR-H3의 각 위치별 아미노산 분포를 나타내고 있다. 각 중첩막대는 각기 다른 길이의 CDR-H3의 각 Kabat 위치에서의 아미노산 빈도의 총합을 나타내고 있다. 각기 다른 길이를 가지는 모든 CDR-H3에 대해서, 마지막 세 개의 아미노산 잔기를 각각 100j, 101 및 102로 표시하였다.

발명을 실시하기 위한 구체적인 내용

[0139] 이하 본 발명을 실시예를 통하여 보다 상세하게 설명한다. 그러나 이들 실시예는 본 발명을 예시적으로 설명하기 위한 것으로, 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다.

[0140] **실시예 1: 항체 라이브러리 구축**

[0141] 본 발명은 합성 인간항체 라이브러리를 구축하고자 하였으며, 특히, 본 발명에서는 25 kDa 크기를 가지는 scFv 단편을 이용한 항체 라이브러리를 구축하였으며, 구체적으로 본 발명의 항체는 면역글로불린의 VH 도메인과 VL 도메인을 15개의 아미노산으로 이루어진 사슬 즉 $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$ 로 연결하여 제작하였다.

[0142] 합성 라이브러리의 구축을 위해서는 라이브러리 서열의 설계가 선행되어야 함에 따라 본 발명자들은 하기와 같이 서열을 설계하였다.

[0143] 1-1. 항체 라이브러리 골격 서열 설계

[0144] 비교적 높은 골격 다양성을 가지는 B세포 유래 항체 라이브러리 혹은 자연항체 라이브러리(natural antibody library)와 달리, 합성 항체라이브러리는 한 개 혹은 제한된 개수의 골격 서열을 기반으로 구축된다. 본 발명에

서는 두 개의 항체 골격을 이용하여 항체 라이브러리를 제조하였다.

[0145] 구체적으로 본 발명의 항체 라이브러리를 이루는 모든 클론들이 인간 면역글로불린의 VH3-23 유전자와 VK3-A27 유전자가 연결부위로 이어진 scFv, 혹은 VH3-23 유전자와 VL1g 유전자가 연결부위로 이어진 scFv를 골격으로 가지며, 이 중 상보성 결정부위에 인공적 다양성을 도입하여 라이브러리를 구축하였다. 이에 사용한 골격 서열은 하기 표 1과 같다.

표 1

[0146]

골격형태	서열
(VH3-23)-linker-(Vk3-A27)	5'-gaagtgcagctgctggaaagtggaggtggactggtgcagcctggcggcagcctgcgcctgagctgtgccgccagcggattcaccttcagcNNNNNNNNNNNNNNNtgggttcgccaagcacctggcaaaggcctggaatgggtgNNNNNNNNNNNNNNNcgctttaccatcagccgcgataacagcaaaaacaccctgtatctgcagatgaacagcctgcgcgccgaggacaccgcagtctactactgtNNNNNNNNNNNNNNNtggggacaaggtactctggtgaccgtgagcagc**ggtggaggaggtagcggaggtggtggatctggaggtggaggtagt**gaaatcgtgctgacccagagccctggcaccctgagcctgagccctggcgaacgcgcaacactgtcatgcNNNNNNNNNNNNNNNtggtatcagcagaaaccaggtcaggctccacgtctgctgatctatNNNNNNNNNNNNNNNggcatccctgatcgcttctcaggatctggaagcggtaccgattttaccctgaccatcagccgcctggaacctgaggactttgccgtgtattattgtNNNNNNNNNNNNNNNttcggtcaggcactaaagtggaaatcaaa-3' (서열번호 50)
	N'-EVQLLESGGGLVQPGGSLRLSCAASGFTFSXXXXXWVRQAPGKGLEWVXXXXXRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKXXXXXWGQGTLVTVSS**GGGGSGGGGSGGGGS**EIVLTQSPGTLSLSPGERATLSCXXXXXWYQQKPGQAPRLLIYXXXXXGIPDRFSGSGSGTDFTLTISRLEPEDFAVYYCXXXXXFGQGTKVEIK-C' (서열번호 51)
(VH3-23)-linker-(Vl1g)	5'-gaagtgcagctgctggaaagtggaggtggactggtgcagcctggcggcagcctgcgcctgagctgtgccgccagcggattcaccttcagcNNNNNNNNNNNNNNNtgggttcgccaagcacctggcaaaggcctggaatgggtgNNNNNNNNNNNNNNNcgctttaccatcagccgcgataacagcaaaaacaccctgtatctgcagatgaacagcctgcgcgccgaggacaccgcagtctactactgtNNNNNNNNNNNNNNNtggggacaaggtactctggtgaccgtgagcagc**ggtggaggaggtagcggaggtggtggatctggaggtggaggtagt**cagagcgtgctgacccagcctcctagcgcctccggtacaccaggacagcgcgtgactattagctgtNNNNNNNNNNNNNNNtggtaccagcaactgcctggaactgcacctaagctgctgatctatNNNNNNNNNNNNNNNggcgtgcctgatcgctttagcggtagcaaatcaggcaccagcgccagcctggccatcagcggccttcgctccgaagatgaagccgattattattgtNNNNNNNNNNNNNNNtttggtggcggtaccaagctgaccgtgctg-3' (서열번호 52)
	N'-EVQLLESGGGLVQPGGSLRLSCAASGFTFSXXXXXWVRQAPGKGLEWVXXXXXRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAKXXXXXWGQGTLVTVSS**GGGGSGGGGSGGGGS**QSVLTQPPSASGTPGQRVTISCXXXXXWYQQLPGTAPKLLIYXXXXXGVPDRFSGSKSGTSASLAISGLRSEDEADYYCXXXXXFGGGTKLTVL-C' (서열번호 53)

[0147] 상기 표의 핵산서열에서 N는 임의의 핵산, 아미노산 서열에서 X는 임의의 아미노산을 의미하고 밑줄친 부위가 CDR이고, 볼드체로 표시된 부위가 링커(연결부위)이다.

[0148] 1-2. 항체 라이브러리 상보성 결정부위(CDR) 서열 설계

[0149] 상기의 항체 라이브러리의 골격에 다양한 상보성 결정부위(CDR) 서열을 삽입하여 scFv 항체 라이브러리를 제작하였다. 상보성 결정부위의 다양성은 알려진 인간항체들의 CDR 서열의 특성을 분석하고 시뮬레이션하여 인간항체 CDR과 유사한 서열적 다양성을 가지도록 설계되었다. 그 구체적인 방법은 다음과 같다.

[0150] - IMGT 데이터베이스 (http://imgt.org)로부터 8,846개의 인간 면역글로불린 중쇄 가변부위 (V_H), 3,110개의 카파 경쇄 가변부위 (V_K), 2,440개의 람다 경쇄 가변부위 (V_λ) 서열을 다운로드하여 CDR 서열을 추출하였다.

[0151] 다운로드한 가변부위 서열들은 각 가변부위의 3개의 CDR이 모두 존재하는, 서로 동일하지 않은 서열들이다. CDR을 추출하기 위하여 http://www.bioinf.org.uk/abs/index.html#cdrid 에 있는 규칙을 사용하였다.

[0152] - V-base(http://www2.mrc-lmb.cam.ac.uk/vbase/alignments2.php)에 있는 인간항체 생식계열(germline) 면역글로불린 유전자의 CDR 서열들과 상기 IMGT 데이터베이스에서 추출된 숙성(mature) CDR 서열들을 비교 분석하여,

[0153] (i) 각 숙성 CDR서열에 가장 가까운 생식계열 CDR 서열을 찾고, 이로부터 각 숙성 CDR에 일어난 돌연변이의 위치, 종류, 빈도를 확인하고 (ii) 각 생식계열 CDR이 숙성 인간항체에서 사용되는 빈도를 확인하였다.

[0154] 본 발명의 라이브러리 중쇄 CDR 설계에 사용된 생식계열 CDR 서열로써, CDR-H1에 대한 생식계열 CDR 서열은 서열번호 54 내지 84의 아미노산 서열이며, CDR-H2에 대한 생식계열 CDR 서열은 서열번호 85 내지 121의 아미노산 서열을 사용하였다. 한편, CDR-H3의 설계에 있어서는 생식계열 CDR 서열을 사용하지 않고, 그 길이를 아미노산 9개 내지 20개로 하여, 길이별 빈도 및 각 위치에서의 아미노산 사용빈도가 자연인간항체 서열들의 CDR-H3의 길이별 빈도 및 각 위치에서의 아미노산 사용빈도를 모사하도록 하였다.

[0155] 한편, 본 발명의 라이브러리 경쇄 CDR 설계에 사용된 생식계열 CDR 서열로, 람다 경쇄 또는 카파 경쇄를 가질 수 있으며, 카파 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 122 내지 145의 아미노산 서열이고, 카파 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 146 내지 165의 아미노산 서열이었다. 한편, 카파 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 210 내지 236의 아미노산 서열이다.

[0156] 한편, 본 발명의 람다 경쇄의 CDR-L1에 대한 생식계열 CDR 서열은 서열번호 166 내지 189의 아미노산 서열이고, 람다 경쇄의 CDR-L2에 대한 생식계열 CDR 서열은 서열번호 190 내지 209의 아미노산 서열이었다. 한편, 람다 경쇄의 CDR-L3에 대한 생식계열 CDR 서열은 서열번호 237 내지 252의 아미노산 서열이다.

[0157] - 재조합이 일어나지 않고 체세포 초돌연변이만 일어나는 중쇄 및 경쇄의 CDR1과 CDR2에 대해서 다음과 같은 작업을 수행하였다. 즉 컴퓨터를 이용한 시뮬레이션을 통하여, 인간 생식계열 CDR 서열에 가상의 돌연변이를 도입하여 인간유래 숙성 항체의 CDR과 유사한 서열 및 생식계열 CDR 사용빈도를 가지도록 CDR 서열들을 설계하였다. 각 CDR 별로 1,500개의 시뮬레이션된 서열들이 설계되었다.

[0158] - 재조합 및 체세포 초돌연변이가 일어나는 경쇄의 CDR3에 대하여, 카파 경쇄의 경우 첫 7개의 아미노산에 대해서 CDR1 및 CDR2에서와 동일한 작업을 수행하고, 끝의 2개 혹은 3개의 아미노산 서열은 숙성 인간항체의 카파 경쇄 CDR3에서의 아미노산 빈도를 모사하여 총 9개 혹은 10개 아미노산으로 이루어진 1,500개의 서열들을 설계하였다. 람다 경쇄의 경우 첫 7개 혹은 8개의 아미노산에 대하여 CDR1 및 CDR2에서와 동일한 작업을 수행하고, 끝의 2개 혹은 3개의 아미노산 서열은 숙성 인간항체의 람다 경쇄 CDR3에서의 아미노산 빈도를 모사하여 총 9개, 10개, 혹은 11개 아미노산으로 이루어진 1,500개의 서열들을 설계하였다.

[0159] -구체적으로, 대표적으로 수행했던 상기 CDR-L3 상보성 결정부위의 C 말단으로부터 2개 내지 3개의 아미노산 서열의 빈도수는 하기 표 2와 같았다.

표 2

[0160]

아미노산	CDR-L3(%)												
	카파 CDR-L3 길이 9		카파 CDR-L3 길이 10			람다 CDR-L3 길이 9		람다 CDR-L3 길이 10			람다 CDR-L3 길이 11		
	8번째	9번째	8번째	9번째	10번째	8번째	9번째	8번째	9번째	10번째	9번째	10번째	11번째
A	0.4	2.5	1.2	0.0	0.4	12.8	0.6	3.7	2.5	1.0	13.2	1.2	0.6
C	0.7	0.0	0.2	1.8	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.3	0.0
D	0.4	0.1	0.2	0.8	0.0	0.6	0.0	2.9	0.2	0.0	0.8	0.5	0.0
E	0.2	0.0	0.2	0.4	0.0	1.1	0.0	0.2	1.0	0.0	0.5	1.9	0.0
F	7.2	0.0	0.4	10.1	0.0	0.0	0.0	1.2	3.4	0.0	2.9	0.8	0.0
G	1.5	0.3	2.8	1.8	0.2	5.6	0.0	2.9	4.9	0.0	32.9	5.7	0.0
H	2.1	0.1	0.4	1.4	0.2	0.6	0.0	4.7	0.7	0.0	26.5	0.2	0.0
I	6.2	0.2	0.6	9.3	0.4	1.7	6.7	2.7	0.2	6.9	0.2	0.0	2.8
K	0.2	0.0	0.2	0.4	0.0	1.7	0.0	0.7	0.0	0.0	0.1	0.3	0.0
L	22.7	0.3	7.5	14.1	0.2	2.8	3.9	3.2	16.4	3.9	6.8	6.5	5.2
M	0.1	0.1	2.8	0.0	0.2	0.6	0.0	0.0	0.5	0.0	0.0	0.4	0.0
N	1.7	0.1	0.0	1.2	0.2	0.6	0.0	14.7	0.0	0.0	0.3	0.5	0.0
P	7.6	0.3	59.4	0.0	0.0	1.1	0.0	2.2	0.5	0.0	2.5	8.9	0.0
Q	4.9	0.0	2.0	0.0	0.0	6.7	0.0	0.7	1.2	0.0	1.3	1.7	0.0
R	16.2	0.1	14.7	1.6	0.0	4.5	0.0	1.5	7.8	0.0	1.4	7.0	0.0
S	0.8	2.8	4.4	1.2	4.0	0.6	0.0	8.1	0.7	0.0	2.9	2.7	0.0
T	0.4	92.8	2.0	0.6	94.1	1.7	0.0	47.5	0.2	0.0	1.3	0.0	0.0
V	1.1	0.3	0.6	3.8	0.0	42.5	88.8	0.5	25.2	88.2	2.3	29.7	91.4

W	10.9	0.0	0.2	27.5	0.0	9.5	0.0	0.0	22.8	0.0	0.1	18.4	0.0
Y	14.9	0.0	0.0	23.8	0.0	5.6	0.0	2.5	10.3	0.0	4.2	13.1	0.0

[0161] 이에 따라, CDR-L3의 경우에는 생식계열 CDR 서열을 기반으로 앞쪽 7개(길이 11을 가지는 람다 경쇄 CDR-L3의 경우는 8개) 아미노산 서열을 고른 후, 그 후에 2 또는 3개의 아미노산을 상기 표 2의 빈도수를 기반으로 덧붙였다. 이는 V-J 재조합에 따라 CDR-L3의 C-말단 부위의 서열이 생식계열 서열에서 유래하지 않는 경우가 많고 불확정성이 높기 때문이다.

[0162] - 재조합 및 체세포 초돌연변이가 일어나는 중쇄의 CDR3에서는 VDJ 재조합 및 접합부위 유연성(junctional flexibility), P-첨가 (P-addition), N-첨가 (N-addition) 등의 기전에 의하여 생식계열 서열의 확인이 어려운 경우가 많으므로 다른 CDR 같은 분석이 불가하다.

[0163] 이에 따라 길이별로 9개에서 20개의 아미노산 길이의 CDR-H3 서열을 분석하여, 각 길이의 각 위치에서의 아미노산 사용빈도를 분석하여 시뮬레이션하였다. 다만 마지막 3개의 아미노산 서열은 J 유전자로부터 유래하는 경우가 대부분이므로 각 CDR-H3 길이별로 자주 사용되는 최대 8개의 삼(3)아미노산 서열을 숙성항체에서의 사용빈도를 고려하여 설계하였다.

[0164] 항원-항체반응에서 CDR-H3의 기여도는 다른 CDR들보다 더 크므로, 다른 CDR보다 많은 수의 CDR-H3 서열을 설계하였다. 한 번의 배열합성에 3,918개의 올리고뉴클레오티드 서열을 합성할 수 있어서, CDR-H3의 길이에 따라 짝수와 홀수로 나누어 각각 3,918개씩 총 7,836개의 서열을 설계하였다. 즉 CDR-H3의 길이가 9,11,13,15,17,19개인 것들을 함께 합성하고, 10,12,14,16,18,20개인 것들을 함께 합성하도록 설계하였다.

[0165] 상기 서술한, 설계된 CDR들의 가짓수는 중복서열을 고려하지 않은 것으로, 실제 동일한 서열들이 중복되는 것을 제외하면 실제 고유 CDR은 하기 표 3와 같이 측정되었다.

표 3

[0166] scFv 라이브러리

CDR	H1	H2	H3	K1	K2	K3	L1	L2	L3
전체	1,500	1,500	1,500	1,500	1,500	1,500	1,500	1,500	1,500
고유	502	1,300	7,526	657	202	920	678	279	1,015

[0167] 상기 CDR 부위의 설계에 대한 개념설명이 도 1에 제시되어 있다.

[0168] - 중쇄, 카파 경쇄, 람다 경쇄에 각각 3개씩의 CDR이 있으며, 총합 9개의 CDR이 있다. 이 중에 중쇄의 첫 번째 CDR (CDR-H1)과 카파 및 람다 경쇄의 두 번째 CDR (CDR-L2)를 제외한 나머지 CDR들은 인간유래 항체에서의 길이의 다양성이 크기 때문에, 다양한 길이를 가지도록 설계되었다. 구체적으로 중쇄의 CDR-H2는 16개 혹은 17개의 아미노산, 카파 경쇄의 CDR-L1은 11,12,16개의 아미노산, 람다 경쇄의 CDR-L1은 11,13,14개의 아미노산, 카파 경쇄의 CDR-L3은 9,10개의 아미노산, 람다 경쇄의 CDR-L3은 9,10,11개의 아미노산, 중쇄의 CDR-H3은 9개에서 20개의 아미노산을 가질 수 있도록 설계하였으며, 각 길이를 가지는 CDR 수의 비율은 인간유래항체에서의 각 길이의 비율을 따랐다.

[0169] 1-3. 항체 라이브러리 상보성 결정부위(CDR) 서열 배제 서열

[0170] CDR 서열의 설계시, 다음의 번역후 변형(post translational modification) 서열을 포함하는 것들을 배제하였다. 단백질의 번역후 변형은 단백질의 기능 및 물성에 영향을 줄 수 있으므로 항체 라이브러리의 설계시 가능한 한 배제하는 것이 유리하다.

[0171] - N-당화 (N-glycosylation) 서열, 즉 Asp-Xaa-Ser/Thr (Asp: 아스파라긴산, Xaa: 시스테인을 제외한 임의의 19종의 아미노산, Ser/Thr: 세린 혹은 트레오닌) 서열을 배제하였다.

[0172] - 이성질체화 (isomerization) 서열을 제외하였다. 단백질 내의 아스파라긴산은 수용액 상태에서 자발적으로 느리게 이소아스파라긴산(isoaspartic acid)으로 이성질체화할 수 있음이 알려져 있으며, 이 반응의 속도는 특히 아스파라긴산의 C-말단에 위치하는 아미노산 잔기에 의해 영향을 받는다. 이 잔기가 글리신일 경우 특히 이성질체화 속도가 빨라짐이 알려져 있으므로, 아스파라긴산-글리신 서열을 가지는 CDR 서열들을 배제하였다.

[0173] - 탈아미드화 (deamidation) 서열을 제외하였다. 단백질 내의 아스파라긴은 수용액 상태에서 자발적으로 느리게 아스파라긴산으로 탈아미드화할 수 있음이 알려져 있으며, 이 반응의 속도는 특히 아스파라긴의 C-말단에 위치하는 아미노산 잔기에 의해 영향을 받는다. 이 잔기가 글리신일 경우 특히 탈아미드화 속도가 빨라짐이 알려져 있으므로, 아스파라긴-글리신 서열을 가지는 CDR 서열들을 배제하였다.

[0174] - 절단(cleavage) 서열을 제외하였다. 단백질을 이루는 펩타이드 결합은 수용액 상태에서 서서히 절단되며, 특히 아스파라긴산-프롤린 서열에서 이러한 절단이 더 빠르게 일어나므로 아스파라긴산-프롤린 서열을 가지는 CDR 서열들을 배제하였다.

[0175] - 산화(oxidation) 서열을 제외하였다. 단백질을 이루는 아미노산 20종 중 여러 개가 산화적 변형을 일으킬 수 있음이 알려져 있으며, 특히 이 중에서 황을 포함하는 시스테인과 메티오닌이 상대적으로 쉽게 산화된다. 따라서 기본적으로 시스테인과 메티오닌이 포함된 CDR 서열들을 배제하되, 생식계열 유전자의 CDR 내에 이미 우세하게 존재하는 메티오닌의 경우는 배제하지 않았다. 구체적으로 중쇄의 CDR-H1의 34번 잔기나 CDR-H3의 100번 잔기가 이에 해당한다.

[0176] - 설계된 CDR의 아미노산 서열은 뉴클레오티드 서열로 역번역하였으며 CDR 서열의 양 옆에 항체 가변부위의 프레임워크 서열을 더하여 100개의 뉴클레오티드로 이루어진 올리고뉴클레오티드 서열들을 최종설계하였다. 총 19,836개의 서열을 설계하고, 이를 배열 합성(array synthesis)법을 통하여 올리고뉴클레오티드 혼합물의 형태로 합성하였다 (미국 텍사스주 휴스턴 소재 LC Sciences사).

[0177] 합성된 올리고뉴클레오티드 혼합물에서 각 CDR들을 중합효소 연쇄반응(PCR)을 통해 증폭하였으며, 이 때 사용된 프라이머 조합을 하기 표 4에 개시하였다. PCR 조건은 다음과 같다: 섭씨 94도, 2분; (섭씨 94도, 30초/섭씨 56도, 30초/섭씨 72도, 30초) 25회 반복; 섭씨 72도, 7분. PCR 반응의 주형으로 1.6 나노그램의 올리고뉴클레오티드 혼합물이 사용되었으며, 프라이머는 각각 600 nM 농도이고 dNTP는 0.8 mM, 중합효소는 Taq (2.5 유닛), 반응부피는 100마이크로리터이다. 반응 완료후 2% 아가로즈 젤에서 전기영동으로 분리한 후 약 100 염기쌍 (base pair, bp) 크기에 해당하는 DNA 밴드로부터 DNA를 추출, 정제한다.

표 4

[0178]

CDR	프라이머	서열 (5' -> 3')
CDR-H1	OPALS-h1-f (서열번호 1)	CAGCGGATTCACCTTCAGC
	OPALS-h1-b (서열번호 2)	AGGTGCTTGGCGAACCCA
CDR-H2	OPALS-h2-f (서열번호 3)	GGCCTGGAATGGGTGAGC
	OPALS-h2-b (서열번호 4)	GCGGCTGATGGTAAAGCG
CDR-H3-odd/even	OPALS-h3-f (서열번호 5)	GGACACCGCAGTCTACTACT
	OPALS-h3-b (서열번호 6)	CACCAGAGTACCTTGTCCC
CDR-L1 카파	OPALS-k1-f (서열번호 7)	CGCGCAACACTGTCATGC
	OPALS-k1-b (서열번호 8)	TGGAGCCTGACCTGGTTTC
CDR-L2 카파	OPALS-k2-f (서열번호 9)	CCAGGTCAGGCTCCACGT
	OPALS-k2-b (서열번호 10)	ACCGCTTCCAGATCCTGAG
CDR-L3 카파	OPALS-k3-f (서열번호 11)	CTGGAACCTGAGGACTTTG
	OPALS-k3-b (서열번호 12)	ACTTTAGTGCCCTGACCG

CDR-L1 람다	OPALS-l1-f (서열번호 13)	GCGCGTGACTATTAGCTGT
	OPALS-l1-b (서열번호 14)	AGGTGCAGTTCCAGGCAGT
CDR-L2 람다	OPALS-l2-f (서열번호 15)	GCCTGGAACTGCACCTAAG
	OPALS-l2-b (서열번호 16)	GCCTGATTTGCTACCGCTA
CDR-L3 람다	OPALS-l3-f (서열번호 17)	CTTCGCTCCGAAGATGAAG
	OPALS-l3-b (서열번호 18)	GTCAGCTTGGTACCGCCA

[0179] 항체 라이브러리 제작의 골격으로 사용하기 위하여 코돈 최적화된 scFv 유전자들을 합성하였으며 (미국 뉴저지주 피스카타웨이 시 소재 Genscript 사), 이 유전자들은 생식계열 유전자인 VH3-23와 VK3-A27, 또는 VH3-23과 VL1g가 $(Gly-Gly-Gly-Gly-Ser)_3$의 연결부위로 연결된 scFv 유전자들로 pUC57 벡터에 클로닝되어있으며 파지미드 벡터로의 클로닝을 위해 pComb3X 벡터와 호환되는 두 개의 SfiI 제한효소 부위가 포함된다. 라이브러리 제작의 편의를 위하여, pUC57 벡터에 클로닝된 이 유전자들을 pComb3X 벡터로도 클로닝하여 PCR의 주형으로 사용하였다. 골격으로 사용하기 위한 용도이므로 이 유전자들은 서열적 다양성이 없는 단일서열들이며, 포유동물세포에서 발현이 향상되도록 코돈이 최적화되었으나 PCR 과정에서의 비특이적 어닐링(annealing)을 방지하기 위하여 아미노산 서열에 변화가 없는 범위에서 일부 최적화 DNA서열에 변화를 가하였다. 이들(cFv VH3-23_linker_VK3-A27, scFv VH3-23_linker_VL1g)의 서열은 각각 서열번호 19 및 서열번호 20에 개시하였다.

[0180] 1-4. CDR 라이브러리 제작

[0181] 코돈 최적화된 scFv 골격의 다양한 절편들을 PCR을 통하여 증폭하고, 이를 역시 PCR 증폭된 CDR 부위 DNA와 중첩연장(overlap extension) PCR을 통하여 연결함으로써 하나의 CDR에만 다양성을 가지는 단일 CDR 라이브러리들을 제작하였다.

[0182] PCR 조건은 위와 동일하며, 다만 섭씨 72 ℃에서의 연장 시간은 예측되는 DNA 절편 결과물의 길이에 따라 30초, 1분, 혹은 1분 30초로 조정하였다. 구체적으로 예측되는 길이가 500 염기쌍 이하인 경우는 30초, 500에서 1,000 염기쌍 사이인 경우는 1분, 1,000에서 1,500 염기쌍 사이인 경우에는 1분 30초의 연장 시간을 사용하였다. 이 과정들 및 이에 사용된 프라이머 서열들을 하기 표 5 및 표 6에 정리하였다.

표 5

[0183]

산출물명	DNA 주형	프라이머	
단일 CDR 라이브러리 제작을 위한 골격부위 PCR			
1	VH3-23/VL1g-pUC57	pUC57-b	OPALS-H1-rc-b
2	VH3-23/VL1g-pUC57	pUC57-b	OPALS-H2-rc-b
3	VH3-23/VL1g-pUC57	pUC57-b	OPALS-H3-rc-b
4	VH3-23/VL1g-pUC57	pUC57-b	OPALS-L1-rc-b
5	VH3-23/VL1g-pUC57	pUC57-b	OPALS-L2-rc-b
6	VH3-23/VL1g-pUC57	pUC57-b	OPALS-L3-rc-b
7	VH3-23/VL1g-pComb3X	OPALS-H1-rc-f	pC3X-b
8	VH3-23/VL1g-pComb3X	OPALS-H2-rc-f	pC3X-b
9	VH3-23/VL1g-pComb3X	OPALS-H3-rc-f	pC3X-b
10	VH3-23/VL1g-pComb3X	OPALS-L1-rc-f	pC3X-b
11	VH3-23/VL1g-pComb3X	OPALS-L2-rc-f	pC3X-b
12	VH3-23/VL1g-pComb3X	OPALS-L3-rc-f	pC3X-b
13	VH3-23/VK3-A27-pComb3X	OPALS-K1-rc-f	pC3X-b
14	VH3-23/VK3-A27-pComb3X	OPALS-K2-rc-f	pC3X-b
15	VH3-23/VK3-A27-pComb3X	OPALS-K3-rc-f	pC3X-b
26	VH3-23/VK3-A27-pUC57	pUC57-b	OPALS-K1-rc-b
27	VH3-23/VK3-A27-pUC57	pUC57-b	OPALS-K2-rc-b
28	VH3-23/VK3-A27-pUC57	pUC57-b	OPALS-K3-rc-b

단일 CDR 라이브러리 제작			
H1	1 + CDR-H1 + 7	pC3x-b	pUC57-b
H2	2 + CDR-H2 + 8		
Odd	3 + CDR-H3-odd + 9		
Even	3 + CDR-H3-even + 9		
λ1	4 + CDR-L1-람다 + 10		
λ2	5 + CDR-L2-람다 + 11		
λ3	6 + CDR-L3-람다 + 12		
κ1	26 + CDR-L1-카파 + 13		
κ2	27 + CDR-L2-카파 + 14		
κ3	28 + CDR-L3-카파 + 15		

표 6

[0184]

프라이머	서열(5'->3')
OPALS-H1-rc-f (서열번호 21)	TGGGTTCGCCAAGCACCT
OPALS-H2-rc-f (서열번호 22)	CGCTTTACCATCAGCCGC
OPALS-H3-rc-f (서열번호 23)	GGGACAAGGTACTCTGGTG
OPALS-L1-rc-f (서열번호 24)	ACTGCCTGGAACTGCACCT
OPALS-L2-rc-f (서열번호 25)	TAGCGGTAGCAAATCAGGC
OPALS-L3-rc-f (서열번호 26)	TGGCGGTACCAAGCTGAC
OPALS-K1-rc-f (서열번호 27)	GAAACCAGGTCAGGCTCCA
OPALS-K2-rc-f (서열번호 28)	CTCAGGATCTGGAAGCGGT
OPALS-K3-rc-f (서열번호 29)	CGGTCAGGGCACTAAAGT
OPALS-H1-rc-b (서열번호 30)	GCTGAAGGTGAATCCGCTG
OPALS-H2-rc-b (서열번호 31)	GCTCACCCATTCCAGGCC
OPALS-H3-rc-b (서열번호 32)	AGTAGTAGACTGCGGTGTCC
OPALS-L1-rc-b (서열번호 33)	ACAGCTAATAGTCACGCGC
OPALS-L2-rc-b (서열번호 34)	CTTAGGTGCAGTTCCAGGC
OPALS-L3-rc-b (서열번호 35)	CTTCATCTTCGGAGCGAAG
OPALS-K1-rc-b (서열번호 36)	GCATGACAGTGTTGCGCG
OPALS-K2-rc-b (서열번호 37)	ACGTGGAGCCTGACCTGG
OPALS-K3-rc-b (서열번호 38)	CAAAGTCCTCAGGTTCCAG
pC3x-b (서열번호 39)	AACCATCGATAGCAGCACCG
pUC57-b (서열번호 40)	TTCGCCATTCAGGCTGCG

[0185] 제작된 단일 CDR 라이브러리는 1% 아가로즈 젤 전기영동을 통해 분리 정제하였고, 이를 SfiI 제한효소로 섭씨 50 ℃에서 12시간 이상 처리하여 절단한 후, 역시 SfiI 제한효소로 같은 방식으로 절단한 pComb3X 파지미드 벡터와 라이게이션 하였다. 1 ㎍의 벡터와 1 ㎍의 scFv DNA를 T4 라이게이즈 1,000 유닛이 첨가된 50 ㎕의 T4 라이게이즈 완충용액에서 16시간 상온 반응하였다. 반응물에는 5 ㎕의 3M 아세트산 소듐 (pH 5.2)과 110 ㎕의 에탄올을 첨가하여 섭씨 영하 20 ℃에서 1시간 이상 DNA를 침전시키고, 침전된 DNA를 원심분리하여 차가운 70% 에탄올로 세척하고 20 ㎕의 멸균순수에 용해하였다. 이것을 ER2537 대장균에 전기천공법으로 트랜스폼하고 암피실린 항생제를 함유하는 배지에서 12 내지 16시간 배양하여 단일 CDR 라이브러리를 획득하였다.

[0186] 단일 CDR 라이브러리를 제작한 목적은 합성된 CDR 올리고뉴클레오티드 내에 존재하는 뉴클레오티드의 삽입/결손/치환 등으로 일어나는 부정확하고 번역 조기종결(premature translational termination)을 일으키는 CDR 서열들을 제거하기 위해서이다. 단일 CDR 라이브러리가 형질전환된 대장균을 배양하고, 대수기(log phase)일 때 보조 파지와 카나마이신을 차례로 첨가 후 밤새 배양하여 파지 라이브러리를 얻었다. 이를 면역시험관(immunotube) 표면에 흡착된 항-HA 항체에 대하여 1회 패닝하여 HA-태그를 발현하는 파지 클론만을 선별하였다. HA-태그는 9개의 아미노산(YPYDVPDYA)으로 이루어진 짧은 펩타이드 서열로 pComb3X 파지미드 벡터의 SfiI 클로닝 위치의 3' 말단에 위치하며, 따라서 발현된 scFv 단백질의 C-말단에 존재하게 된다. 올리고뉴클레오티드의 화학적 합성법의 단계적 효율성은 완벽하지 않으므로, 합성된 CDR 올리고뉴클레오티드 중에는 번역 조기종결을 일으키는 삽입/결손/치환이 생긴 서열들이 다수 존재하게 되고, 이들을 포함하는 파지 클론들은 HA-태그를 가질 수 없으므로 패닝에 의해 제거된다. 결과적으로 이 과정을 통해 결함이 없는 CDR 서열들을 선별할 수 있다.

[0187] 1-5. scFv 항체 라이브러리 제작

[0188] 선별된 CDR 서열들은 전술한 조건에 따라 PCR을 통해 증폭되었고, 증폭된 DNA 단편들은 다시 중첩 연장 PCR을 통해 경쇄와 중쇄 가변부위로 조합되었다. 경쇄 가변부위에 연결부위를 중첩연장 PCR로 연결하고, 이를 다시 중쇄 가변부위와 중첩연장 PCR로 조합하여 최종적으로 scFv 유전자 라이브러리를 획득하였다. 본 최종 CDR 라이브러리를 제작하기 위한 PCR 과정을 도 2에 도식화하였다(도 2). 이 과정들 및 이에 사용된 프라이머 서열들을 하기 표 7 및 표 8에 정리하였다.

표 7

[0189]

산출물명	주형	프라이머	
선택된 CDR과 연결부위의 증폭			
VH-CDR1	H1	pC3x-f	OPALS-H2-rc-b
VH-CDR2	H2	OPALS-H1-rc-f	OPALS-H3-rc-b
VH-CDR3-odd	Odd	OPALS-H2-rc-f	OPALS-FR4-b
VH-CDR3-even	Even		
VL-CDRλ1	λ1	OPALS-L1-f	OPALS-L2-rc-b
VL-CDRλ2	λ2	OPALS-L1-rc-f	lFR3-b
VL-CDRλ3	λ3	lFR3-f	pC3x-b
Linker-λ	OPALS-lambda	OPALS-FR4-f	OPALS-L1-rc-b
VL-CDRκ1	κ1	kFR1-f	OPALS-K2-rc-b
VL-CDRκ2	κ2	OPALS-K1-rc-f	OPALS-K3-rc-b
VL-CDRκ3	κ3	OPALS-K2-rc-f	pC3x-b
Linker-κ	OPALS-kappa	OPALS-FR4-f	kFR1-b
경쇄 및 중쇄 가변부위의 증폭			
VH-Odd	VH-(CDR1+CDR2+CDR3-odd)	pC3x-f	OPALS-FR4-b
VH-Even	VH-(CDR1+CDR2+CDR3-even)		
VL-Lambda	VL-(CDRλ1+λ2+λ3)	OPALS-L1-f	pC3x-b
VL-Kappa	VL-(CDRκ1+κ2+κ3)	OPALS-K1-f	
경쇄 가변부위와 연결부위의 연결			
λ	VL-Lambda+linker-λ	OPALS-FR4-f	pC3x-b
κ	VL-Kappa+linker-κ		
ScFv 라이브러리 제작을 위한 최종 PCR			
Oλ	VH-Odd+λ	pC3-seq	dp-seq
Eλ	VH-Even+λ		
Oκ	VH-Odd+κ		
Eκ	VH-Even+κ		

표 8

[0190]

프라이머	서열(5'->3')
pC3x-f (서열번호 41)	GCACGACAGGTTTCCCGAC
lFR3-f (서열번호 42)	CTGGCCATCAGCGGCCTTC
lFR3-b (서열번호 43)	CTGGGTCAGCACGATTTC
kFR1-f (서열번호 44)	GAAATCGTGCTGACCCAG
kFR1-b (서열번호 45)	CTGGGTCAGCACGATTTC
OPALS-FR4-f (서열번호 46)	CTGGTGACCGTGAGCAGC
OPALS-FR4-b (서열번호 47)	GCTGCTCACGGTCACCAG
pC3-seq (서열번호 48)	GTGAGCGGATAACAATTGA
dp-seq (서열번호 49)	AGAAGCGTAGTCCGGAACG

[0191] scFv 유전자 라이브러리는 전술한 방법대로 SfiI 제한효소로 절단하고 pComb3X 벡터에 라이게이션한 후 ER2537 대장균에 트랜스폼하여 scFv 라이브러리를 완성하였다. 구체적으로 CDR-H3의 아미노산 개수(길이)가 짝수인 것과 홀수인 것을 따로 제작하고, 카파 경쇄와 람다 경쇄를 따로 제작하여 이들을 PCR 조합하여 라이게이션-트랜스폼하였으므로 총 4개의 서브 라이브러리가 만들어졌다. 이를 각각 OL (odd-lambda), EL (even-lambda), OK (odd-kappa), EK (even-kappa) 라이브러리로 칭하였으며, 각각의 라이브러리로 형질전환된 박테리아 역가

(titer), 즉 라이브러리의 크기가 $1.3x10^8$, $1.4x10^8$, $1.9x10^8$, $3.7x10^8$ 로 총합 $8.3x10^8$ 크기인 1차 라이브러리, 총합 10^{10}에 이르른 2차 라이브러리를 제작하였다. 제작된 OL 라이브러리, EL 라이브러리의 서열비교를 도 5a에 표시하였으며, OK 라이브러리, EK 라이브러리의 서열비교를 도 5b에 표시하였다.

[0192] 각 서브 라이브러리로부터 24개씩의 클론을 임의로 선택하여 그 박테리아를 배양하고, IPTG로 scFv 항체의 발현을 유도한 후 섭씨 30 ℃에서 밤새 배양하였다. 이로부터 수크로즈 완충용액을 이용한 삼투 충격법(osmotic shock)으로 페리플라즘 추출물(periplasmic extract)를 얻은 후 이것을 1 ㎕씩 니트로셀룰로즈 박막에 가하였다. 박막을 건조시킨 후 3% 탈지분유 완충용액에 담가 블로킹하고, 항 HA 항체-HRP(horseradish peroxidase)를 1시간 동안 결합시킨 후 세척하였다. 루미놀을 이용하여 HA-태그의 존재여부를 검출함으로써 scFv의 발현을 확인하였다 (도 3).

[0193] 본 실시예 1의 항체 라이브러리 제조 단계를 도식화한 결과는 도 4에 정리하였다.

[0194] **실시예 2: 서열 분석**

[0195] 제조된 라이브러리에 반영된 CDR 설계의 정확도를 측정하기 위해, 상기 라이브러리의 차세대 서열분석(Next generation sequencing, NGS)을 수행하였다. 수백만 개의 CDR 서열을 분석하고 설계 서열과 비교하였다(표 9).

표 9

[0196] CDR 라이브러리의 차세대 서열분석

CDR	H1	H2	H3	K1	K2	K3	L1	L2	L3
제조된 서열 수 ($x\ 10^6$)	4.75	4.70	4.53	2.58	2.31	2.52	2.14	2.12	2.09
인프레임 %	91.3	89.7	90.3	92.2	91.6	93.2	91.5	78.7	90.7
설계 길이 %	89.4	88.4	85.6	87.1	82.3	90.7	84.6	56.2	89.4
설계 서열 %	80.9	52.0	51.8	62.4	70.3	74.3	56.2	47.6	67.2
proofreading 전 인프레임%	85.7	50.0	67.9	53.3	44.4	66.7	38.5	31.3	66.7
설계 커버리지 %	100	100	97.3	99.8	100	99.9	100	100	100

[0197] 비록 긴 CDR에 대해서는 고유 서열의 대부분이 전체 설계서열 중 오직 한번 또는 두번만 반복되도록 설계되어 정량 계수(r^2)가 상대적으로 낮으나, 제조된 라이브러리에는 설계된 서열의 거의 전부가 커버된 것을 확인하였으며, 각 설계된 CDR 서열의 발생 빈도 또한 실제 라이브러리에 나타났다(도 6). 당연하게도, 해당 라이브러리는 제조 오류에 따라 어느 설계 서열과도 매칭되지 않은 저-빈도 CDR 서열을 많이 포함하고 있는 것으로 확인되었다. 이 중에는 핵산 첨가 또는 결실 등에 의해서 프레임 시프트가 일어난 비-기능(non-functional) 서열이 포함되어 있었다. 항-HA 항체를 이용한 단일-CDR 라이브러리의 proofreading 패닝을 통해서 비-기능 CDR 서열을 제거하였으며, proofreading 패닝 전에 활성 인프레임 CDR 서열의 비율은 31 내지 86 %인데 비하여, proofreading 패닝 이후에는 대략 90에서 93% 정도임을 확인하였다. 다만, CDR-L2만 약 79%에 불과했으나 이는 중첩 PCR의 부정확한 어닐링 때문이었던 것으로 사료된다. 설계된 길이에 비교한 CDR 서열의 비율은 약 82 내지 91 %(CDR-L2는 약 56%)였다.

[0198] 종합적으로, V_H, V_κ, 및 V_λ 서열의 약 75%는 종결 코돈없는 다양한 기능 도메인들이었고, 라이브러리에서 기능성 scFv 클론은 약 55%에 이르는 것으로 추정되었다. 이러한 추정은 무작위적으로 선택한 라이브러리 클론에 대한 닷-블랏 어세이 결과와 일치하였다. 닷-블랏을 위하여 셀렉션되지 않은 라이브러리로부터 무작위적으로 선택된 scFv 클론의 페리플라즘 추출물을 니트로셀룰로즈 막에 블랏팅을 하였고, 및 추출물 상의 수용성 발현 scFv의 존재를 C 말단의 HA 태그를 측정하여 확인하였다. 그 결과 수용성 발현 scFv 클론은 대략 60% 가량되었다(도 3).

[0199] NGS 데이터로부터 다양한 도메인 중에서 고유한 도메인이 분석되었다. 종결 코돈을 포함하지 않는 각 heavy, 카파 및 람다의 다양한 도메인 서열 약 1.3×10^6개를 분석하였으며, V_H의 98%, V_κ의 89% 및 V_λ의 98%이 반복되지 않았다(도 7).

[0200] 전체 서열 중에서 상이한 다양한 도메인 서열의 수의 %는 V_H, V_κ, 및 V_λ 각각이 99%, 92% 및 99%였다. 한편, 이러한 수치를 기 보고된 다른 항체 라이브러리들의 CDR-H3에서의 고유도(uniqueness)인 97 내지 98%에 비교하였을 때, 셀렉션되지 않은 라이브러리 내의 scFv 클론 중에서 중복도는 그다지 심각하지 않음을 시사한다.

[0201] 제조된 라이브러리에서 CDR 길이의 분포, 특히 CDR-H3 길이의 분포는 설계와 상이하였고, 긴 CDR에 비교해서 짧은 길이의 CDR이 현저히 높은 빈도로 나타났다(도 8). 이는 일부분 올리고뉴클레오티드 어레이 제조 과정에서 프레임 시프트가 일어나거나 조기 스탑 코돈이 도입되는 등의 부정확함 때문인 것으로 사료된다. 이러한 오류는 긴 CDR을 제조하는 동안에 더 자주 발생하고 그들 중 대부분은 항-HA-태그 항체를 이용한 단일 CDR 라이브러리 proofreading 패닝 동안에 제거되기 때문에, 라이브러리에서 긴 CDR이 더 많이 제거되는 원인일 것으로 사료된다. 또한, 짧은 CDR을 가진 scFv가 항-HA-태그 항체에 대한 패닝에 의해서 우선적으로 선별 및 증폭되었을 가능성도 있다.

[0202] 라이브러리 CDR 서열과 자연 CDR 서열간의 유사도를 가장 가까운 생식 CDR 서열로부터 각 CDR 서열의 아미노산 변이 수를 분석하여 측정하였다. CDR을 자연 체세포 초돌연변이(SHM) 패턴을 시뮬레이션하여 설계되었기 때문에, 설계 서열은 자연 유사도가 매우 높을 것이 예상되었다. 이에 따라, 설계 CDR 서열과 자연 CDR 서열 간에 CDR 서열 당 평균 변이 횟수를 비교하였다(표 10).

표 10

[0203] CDR 서열의 유사 인간 생식 CDR 서열로부터 평균 아미노산 변이수

CDR	H1	H2		K1			K2	K3		L1			L2	L3		
길이	5	16	17	11	12	16	7	9	10	11	13	14	7	9	10	11
설계	0.75	2.14	2.55	1.0	1.18	1.85	0.47	0.73	0.65	1.03	1.10	0.98	0.61	0.83	1.18	0.70
비-설계	1.84	4.30	3.67	2.77	3.44	3.82	2.01	1.55	1.45	2.82	3.22	2.75	1.85	1.67	2.11	1.45
자연	0.82	2.09	2.37	1.08	1.29	0.93	0.49	0.72	0.54	0.95	1.10	1.00	0.66	0.88	1.14	0.75

[0204] 제조 오류에 의하여 설계 서열 중 어느 것과도 매치되지 않는 라이브러리 CDR(비-설계 CDR) 서열을 분석할 때, 가장 유사한 생식계열 CDR 서열로부터의 평균 아미노산 변이 수는 설계 CDR 서열의 평균 아미노산 변이 수에 비해서 평균 1 내지 2 아미노산만 차이났다. 이러한 결과는 라이브러리의 CDR 서열이 인간 생식 CDR 서열로부터 오직 적은 수의 변이만을 포함하고 있으며, 자연 인간 항체의 CDR 서열과 매우 유사함을 시사한다. CDR-H3에 대해, 각 위치의 아미노산 분포를 분석하였다 (도 9). 자연 인간 항체, 시뮬레이션된 레파토리 및 제조 라이브러리의 CDR-H3 사이에서 아미노산 분포 패턴이 유사함이 확인되었으며, 이는 라이브러리 CDR의 높은 자연-유사도를 보여주고 있다.

[0205] 예측한 대로, 상기 CDR에서 바람직하지 않은 번역-후 수식(Post-translation modification, PTM) 모티프의 발생 빈도가 자연 인간 항체 CDR보다 현저히 낮았다. 다만, 각각 5개 및 7개의 아미노산으로 이루어져 짧고 자연 인간 항체에서 상대적으로 적은 PTM 모티프를 가지고 있는 CDR-H1 및 CDR-L2는 예외였다 (표 11).

표 11

[0206] 라이브러리 및 자연 인간 항체의 CDR에서 번역-후 수식 비율

모티프	CDR	H1	H2	H3	K1	K2	K3	L1	L2	L3
Asp-Gly	라이브러리	0.05	0.31	0.66	0.03	0.34	0.10	0.32	3.41	0.34
	자연	0.08	10.79	6.45	6.03	0.32	0.09	0.28	0.30	1.98
Asn-Gly	라이브러리	0.04	0.62	0.32	0.05	0.63	0.12	0.60	4.51	0.32
	자연	0.02	5.96	3.03	4.54	0.53	0.50	0.06	0.12	7.61
Asp-Pro	라이브러리	0.01	0.16	0.25	0.04	0.08	0.10	0.04	0.20	0.05
	자연	0.00	2.40	10.13	0.00	0.04	0.14	0.11	0.00	0.66
N-glyc	라이브러리	0.66	1.91	2.77	1.35	1.46	1.00	1.97	0.94	1.30
	자연	0.59	0.78	7.73	0.61	0.07	1.69	6.74	0.30	8.20

Met	라이브러리	0.35	1.03	1.80	0.90	0.55	0.29	1.78	0.70	1.70
	자연	0.06	2.72	9.04	0.88	0.11	12.36	0.39	0.18	0.86
Cys	라이브러리	0.30	1.96	0.49	1.30	0.64	0.77	2.10	0.87	1.29
	자연	0.35	17.50	1.46	0.92	0.18	1.19	0.61	0.24	1.98
총 PTM	라이브러리	1.36	5.44	5.87	3.60	3.58	2.32	6.16	10.48	4.81
	자연	1.09	36.70	32.92	12.55	1.23	15.56	8.08	1.13	20.70

[0207] 상이한 CDR에서 PTM 모티프는 독립적으로 발생하는 것으로 가정했을 때, 라이브러리의 scFv 서열에서 최소한 하나의 PTM 모티프 발생 확률이 대략 20 내지 30 %이고 70 내지 80%의 클론은 PTM 모티프 없는 것으로 추정된다. 이에 반해, 자연 유래의 경우에는 PTM 모티프를 가지지 않은 scFv의 비율이 24 내지 27 %에 불과하다.

[0208] **실시예 3: 항원을 이용한 라이브러리 패닝 및 스크리닝**

[0209] 제조된 라이브러리는 이의 기능성을 확인하기 위하여, 4가지 항원에 대해서 패닝을 수행하였다. 플라스틱 표면에 고정된 항원에 대한 패닝을 4번 반복한 후 다수의 타겟-결합 scFv 클론들을 분리하였다. 3번째 또는 4번째 패닝으로부터 분리된 콜로니들을 ELISA를 통해 스크리닝하고, 양성 신호를 보이는 클론 중 일부를 서열분석하였다(표 12).

표 12

[0210] 항원에 대한 라이브러리의 패닝 및 스크리닝

항원	패닝 횟수	ELISA 양성 수/스크리닝 수	고유 서열/ 전체 분석 서열	Kappa/lambda
AIMP1	3	10/94	6/6	1/5
SerRS	4	16/94	3/14	0/14
hEpCAM-ECD	4	18/94	3/15	3/12
HER3-ECD	4	46/188	6/16	8/8
HEW Lysozyme	4	151/188	6/13	1/12

[0211] 테스트된 항원의 숫자와 서열분석된 클론의 수가 경쇄 클래스 선호성을 일반화하기에는 충분하지 않았으나, 분리된 클론의 대부분은 람다 경쇄 클래스(lambda light chain class)였다. 파지 항체 라이브러리로부터 특정 경쇄 패밀리/클래스를 가지는 클론에 대한 우선 셀렉션(preferential selection)이 기존에 알려졌으며, 특정 V_H-V_L 도메인의 선호 우선 쌍(preferential pairing)과 서브-라이브러리의 기능 길이상 차이에 기여하였다. 커다란 자연 scFv 라이브러리의 항원-유래 파지 표시 셀렉션 이후에 람다 경쇄 강한 선호성도 알려졌으며, 이는 람다 쇄 선호성이 특정 항체 라이브러리의 특징이라기 보다 scFv 라이브러리의 파지 표시 셀렉션의 보편적인 현상임을 시사한다.

[0212] 라이브러리로부터 분리된 ELISA-양성 scFv 클론 중 일부의 결합 에너지를 SPR을 이용하여 분석하였다 (표 13).

표 13

[0213] SPR에 의해 측정된 셀렉션된 목적-특이적 scFv 클론의 결합 동력학 상수

항원-클론	$K_{on}(M^{-1}s^{-1})$	$K_{off}(s^{-1})$	$K_D(M)$
HER3-G3	6.8×10^{5}	2.8×10^{-3}	4.2×10^{-9}
HER3-C5	1.2×10^{5}	1.0×10^{-3}	8.1×10^{-9}
HER3-D11	7.8×10^{4}	5.3×10^{-3}	6.8×10^{-8}
HEWL-B6	1.1×10^{5}	5.3×10^{-3}	4.8×10^{-8}
HEWL-F8	5.8×10^{4}	1.9×10^{-3}	3.3×10^{-8}

HEWL-B2	1.1×10^{6}	3.0×10^{-3}	2.8×10^{-9}
HEWL-G7	1.4×10^{5}	3.2×10^{-3}	2.4×10^{-8}
SRS-A3	2.7×10^{4}	1.8×10^{-3}	6.7×10^{-8}
SRS-C4	9.5×10^{4}	1.3×10^{-3}	1.4×10^{-8}
SRS-D8	6.5×10^{4}	1.4×10^{-3}	2.2×10^{-8}

[0214] 3개의 상이한 항원에 대한 10개의 scFv들에 대한 해리 상수 (*Kd*)가 10^{-9} 내지 10^{-7} M 범위로 측정되었다.

[0215] 셀렉션되지 않은 라이브러리의 NGS 결과를 패닝 이후에 라이브러리로부터 선택된 고유 scFv의 서열과 비교하였다. 패닝 과정이 설계된 CDR 서열의 비율 또는 CDR 잔기당 평균 변이 횟수를 현저히 바꿔주는 것은 아닌 것으로 나타났다 (표 14).

표 14

[0216] 패닝 셀렉션 전후의 CDR 및 FR(framework region, 골격부위)의 변이

CDR 또는 FR	H1	H2	H3	K1	K2	K3	L1	L2	L3	V_HFR	V_κFR	V_λFR
셀렉션 전 서열 중 설계된 서열의 비율 (%)	80.9	52.0	51.8	62.4	70.3	74.3	56.2	47.6	67.2			
셀렉션 후 서열 중 설계된 서열의 비율 (%)	92.0	48.3	64.5	57.1	60.0	83.3	60.0	60.0	60.0			
잔기당 평균변이수 셀렉션 전	0.21	0.18	N/A	0.13	0.11	0.06	0.11	0.12	0.13	0.0067	0.0144	0.0187
잔기당 평균변이수,셀렉션 후	0.17	0.17	N/A	0.14	0.10	0.09	0.14	0.11	0.10	0.0018	0.0019	0.0046

[0217] 올리고뉴클레오티드 합성 또는 PCR 과정에서의 오류에 의한 추가적인 CDR 다양성은 라이브러리의 다양성에 큰 영향을 주지 않음을 확인할 수 있다. 한편, 골격부위(framework region, FR)에서 변이 빈도가 패닝 셀렉션 이후에 감소하는 것을 확인하였다.

[0218] **실시예 4: 라이브러리 구체적 구축 확인**

[0219] 상기 실시예 1-5을 통한 항체 라이브러리의 구축을 검증하기 위하여, scFv의 서열을 분석하였다. 각 서브 라이브러리로부터 임의의 클론들을 선택하고 scFv 유전자를 PCR 기법으로 증폭한 후, DNA를 정제하여 서열분석을 의뢰하였다. 서열분석 결과가 도 5에 제시되어 있으며, 설계한 대로 다양한 서열과 길이를 가지는 CDR들이 라이브러리를 이루는 scFv 클론들에 도입되어 있음을 확인할 수 있다.

[0220] 구체적인 서열분석 결과, 18개의 scFv 서열들 중 절단(아스파라긴산-프롤린)서열이 PCR 오류로 도입된 1건을 제외하면 배제하고자 하였던 번역후 변형 서열들이 존재하지 않는 것을 확인하였다. 이에 비하여 인간유래 숙성항체의 각 CDR에서의 이러한 번역 후 변형을 유발할 수 있는 서열들의 존재비율을 분석한 결과 중쇄의 CDR-H1에서 5.7%, CDR-H2에서 39.7%, CDR-H3에서 34.5%, 카파 경쇄의 CDR-L1에서 12.6%, CDR-L2에서 0.6%, CDR-L3에서 15.9%, 람다 경쇄의 CDR-L1에서 8.3%, CDR-L2에서 6.4%, CDR-L3에서 24.0% 발견되었다. 이로부터 설계의도대로 번역후 변형이 최소화될 수 있는 항체 라이브러리가 구축되었음을 확인하였다.

[0221] 형질전환된 대장균 라이브러리를 배양하고 VCSM13 보조 파지(helper phage)로 감염시켜 scFv 클론이 표면제시된 항체파지 라이브러리를 얻었다 (H.Y.Yang et al., Mol. Cells 2009, 27, 225-235). 이 라이브러리는 1 ㎖ 당 10^{12} 이상의 콜로니형성단위 (colony forming unit, CFU)를 포함한다. 이를 사용하여 라이소자임(Hen egg white lysozyme, HEWL) 및 AIMP1(101-192) 항원에 대하여 항원특이적 항체를 선별하는 패닝 실험을 통해 아래와 같이

항체파지 라이브러리의 기능성을 검증하였다. 이후 면역시험관(immunotube)에 10 μg/ml 농도의 항원 용액을 첨가하여 1시간동안 시험관 표면에 단백질을 흡착시킨 후 탈지분유 3% 완충용액을 시험관에 첨가하여 항원이 흡착되지 않은 표면을 보호하였다.

[0222] 상기 시험관을 비운 후 여기에 탈지분유 3% 완충용액에 분산된 10^{12} CFU의 항체파지 라이브러리를 넣어 1-2 시간 동안 항원과 결합시켰다. 비특이적으로 결합한 파지를 TBST (tris buffered saline - tween20) 용액으로 3회 씻어낸 후, 남아있는 항원특이적 파지 항체를 100 mM 트리에틸아민 용액을 이용하여 용리하였다.

[0223] 상기 용리된 파지를 1.0M 농도의 Tris-HCl 버퍼 (pH 7.8)로 중화시킨 후 ER2537 대장균에 37 ℃에서 1시간 감염시키고, 감염된 대장균을 암피실린과 2% 포도당을 함유하는 LB (Luria-Bertani) 한천배지에 도포하여 37 ℃에서 배양하였다.

[0224] 다음날 배양된 대장균을 5 mL의 SB (super broth) 배양액에 현탁하고 15% 글리세롤을 첨가하여 -80 ℃에 보관하고, 50 ㎕를 20 mL의 SB-암피실린 용액에 접종하여 37 ℃에서 배양하였다.

[0225] 상기 배양액의 600 나노미터 광선의 흡광도가 0.5가 되면 10^{11} PFU (플라크 형성단위)의 VCSM13 보조파지를 넣고 서서히 교반하며 37 ℃에서 배양하였다. 1시간 후 카나마이신 70 μg/ml을 첨가하고 30 ℃에서 빠르게 교반하며 (220 rpm) 밤새 배양하였다. 다음날 배양액을 원심분리한 후, 상층액에 5 ㎖의 5x PEG 침전용액 (20% w/v PEG8000, 15% w/v NaCl)을 넣어 혼합하였다. 혼합액을 얼음에서 30분 이상 두어 파지를 침전시키고, 이를 원심분리하여 300 ㎕의 PBS에 분산시켰다. 다시 파지 용액을 원심분리하여 맑은 상층액만 취하고, 여기에 700 마이크로리터의 3% 탈지분유 완충용액을 섞고 이를 이용하여 위의 패닝 과정을 반복함으로써 항원특이적 클론을 농축시켰다. 3 내지 4회 정도의 반복적인 패닝 후 항체유전자를 포함하는 대장균을 암피실린과 2% 포도당을 함유하는 LB 한천배지에 도포, 배양하여 단일 콜로니들을 얻고, 이를 200 ㎕ SB-암피실린 용액에 접종, 배양한 후 IPTG로 유도하여 scFv 단백질을 대장균의 페리플라즘(periplasm)에서 발현유도하였다.

[0226] 대장균으로부터 수크로즈 완충용액을 이용한 삼투 충격법(osmotic shock)으로 페리플라즘 추출물(periplasmic extract)를 얻은 후, 이를 ELISA 기법을 사용하여 항원과 scFv와의 결합을 확인하는 데 사용하였다 (H.Y.Yang et al., Mol. Cells 2009, 27, 225-235). 결합한 scFv는 항-HA 항체-HRP와 테트라메틸벤지딘(TMB) 기질을 이용하여 검출하였다. 이로부터 확인된 항원특이적 항체 클론은 염기서열 분석법을 통해 분석하였다.

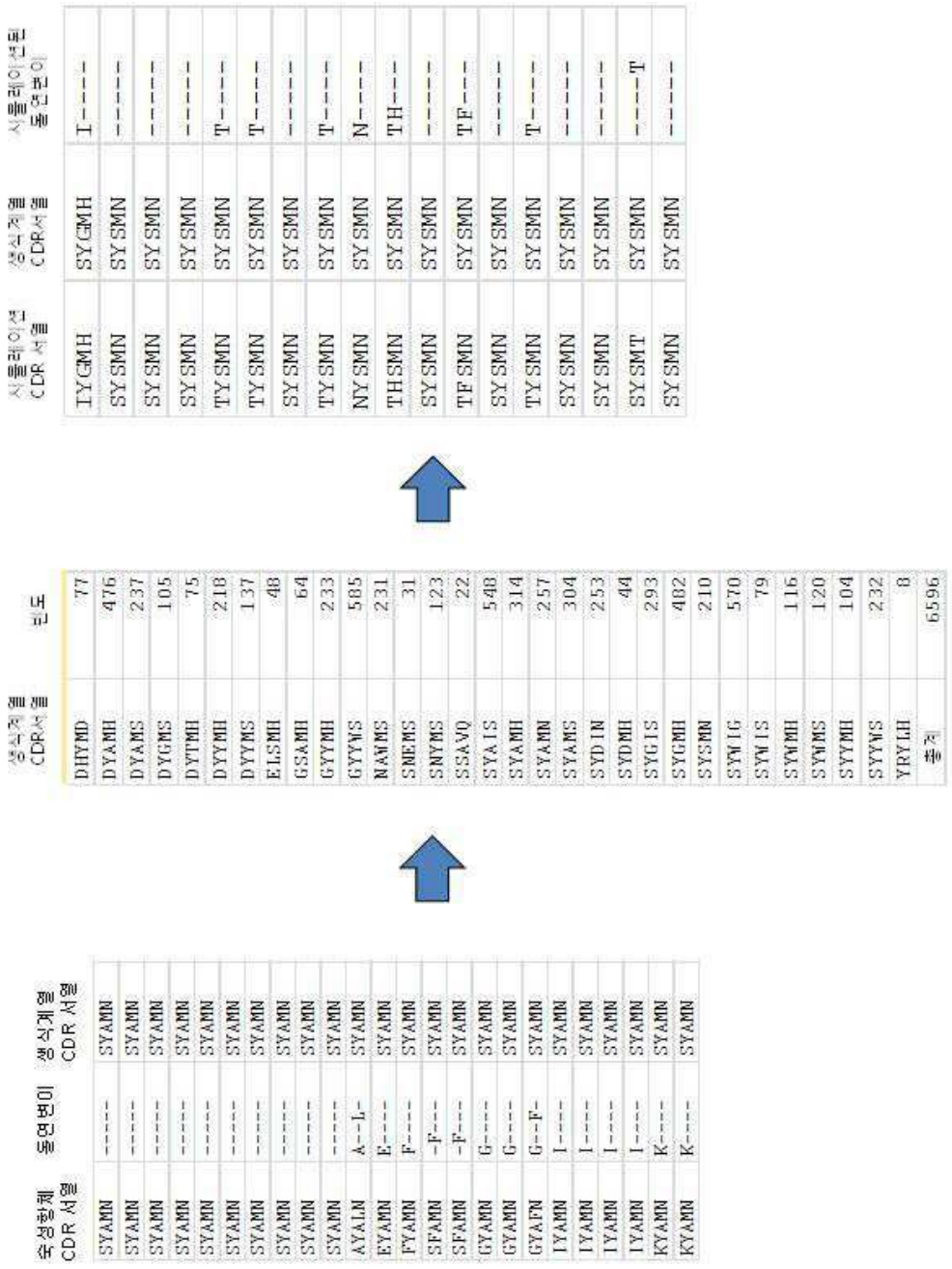
[0227] HEWL 항원에 대하여 188개의 클론들을 ELISA로 스크리닝하여 배경신호 대비 3배 이상의 결합신호를 보이는 ELISA 양성 클론들을 150개 확인하였으며 이 중 16개를 서열분석하여 5개의 고유 서열을 발견하였다. AIMP1(101-192)에 대하여 94개의 클론들을 ELISA로 스크리닝하여 배경신호 대비 3배 이상의 결합신호를 보이는 ELISA양성 클론들을 18개 확인하였으며 이 중 5개를 서열분석하여 2개의 고유 서열을 발견하였다.

[0228] 상기 내용을 종합하면, 본 발명에 의해 제조된 항체 라이브러리는 다수의 항원에 대해 우수한 물성을 가지는 항체들을 포함하여, 기능적 다양성을 가질 뿐 아니라 고유서열을 다수 포함하는 등 항체 라이브러리로써 유용하게 사용될 수 있음을 확인하였다.

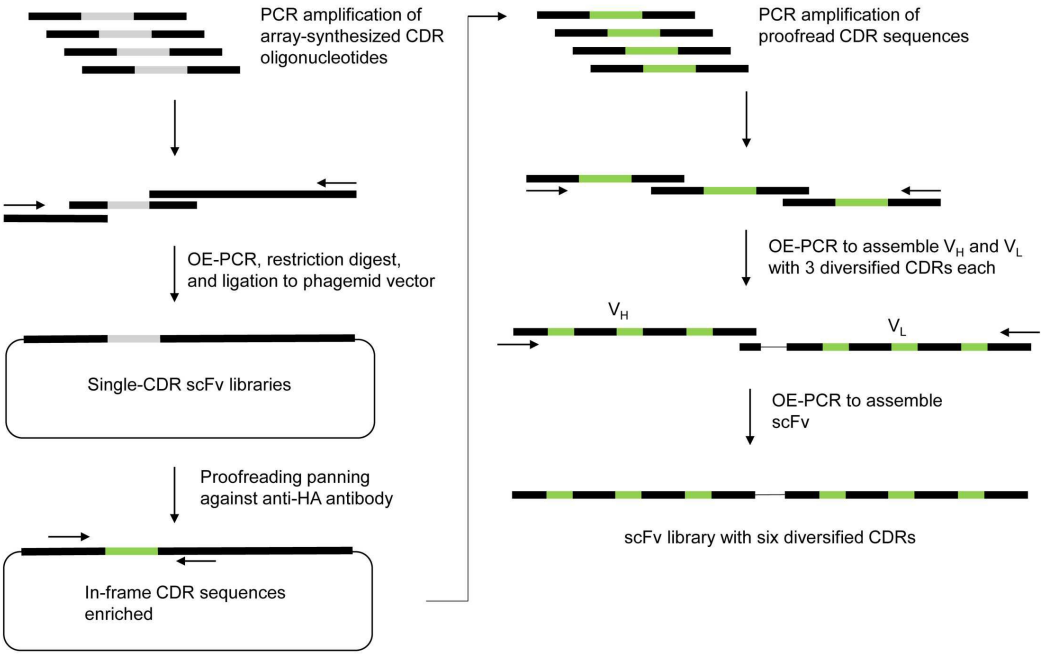
[0229] 이상의 설명으로부터, 본 발명이 속하는 기술분야의 당 업자는 본 발명이 그 기술적 사상이나 필수적 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 이와 관련하여, 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적인 것이 아닌 것으로서 이해해야만 한다. 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허 청구범위의 의미 및 범위 그리고 그 등가 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면

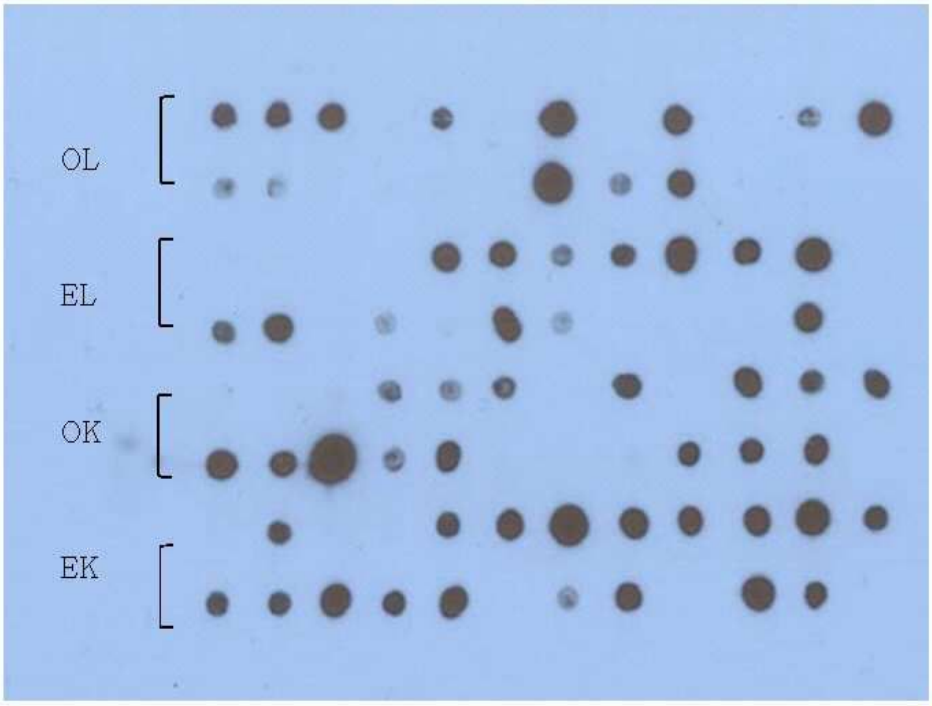
도면1


숙성항체 CDR 서열
돌연변이
생식계열 CDR 서열
빈도
생식계열 CDR서열
시뮬레이션 CDR 서열
생식계열 CDR서열
시뮬레이션된 돌연변이
총계
6596

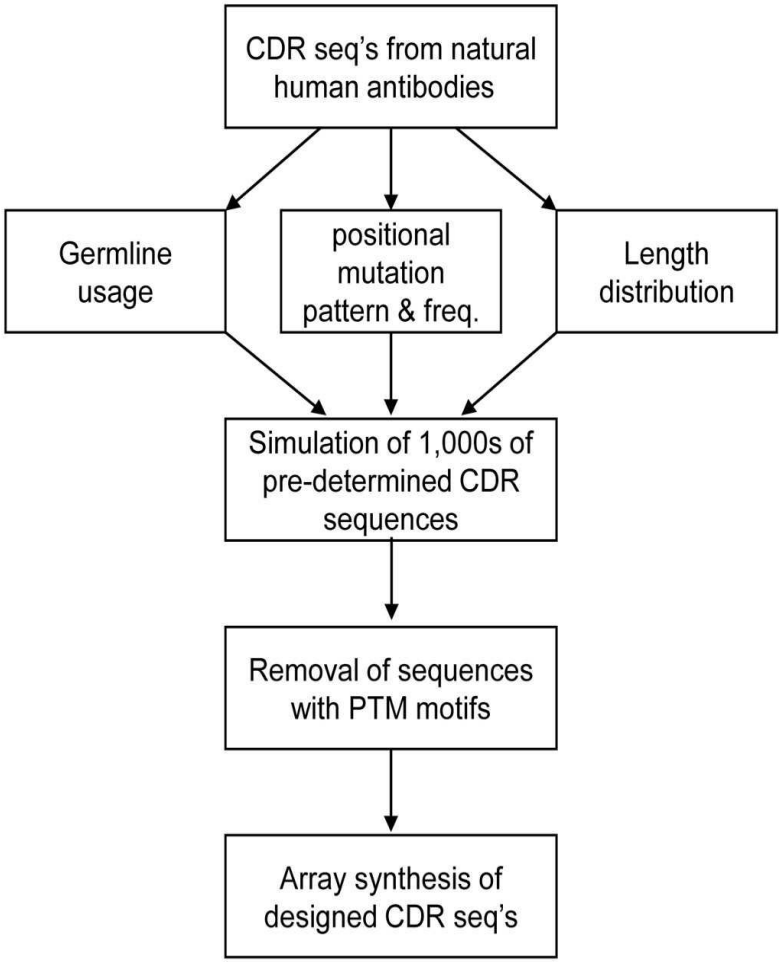
도면2


PCR amplification of array-synthesized CDR oligonucleotides
OE-PCR, restriction digest, and ligation to phagemid vector
Single-CDR scFv libraries
Proofreading panning against anti-HA antibody
In-frame CDR sequences enriched
PCR amplification of proofread CDR sequences
OE-PCR to assemble V_H and V_L with 3 diversified CDRs each
V_H
V_L
OE-PCR to assemble scFv
scFv library with six diversified CDRs

도면3

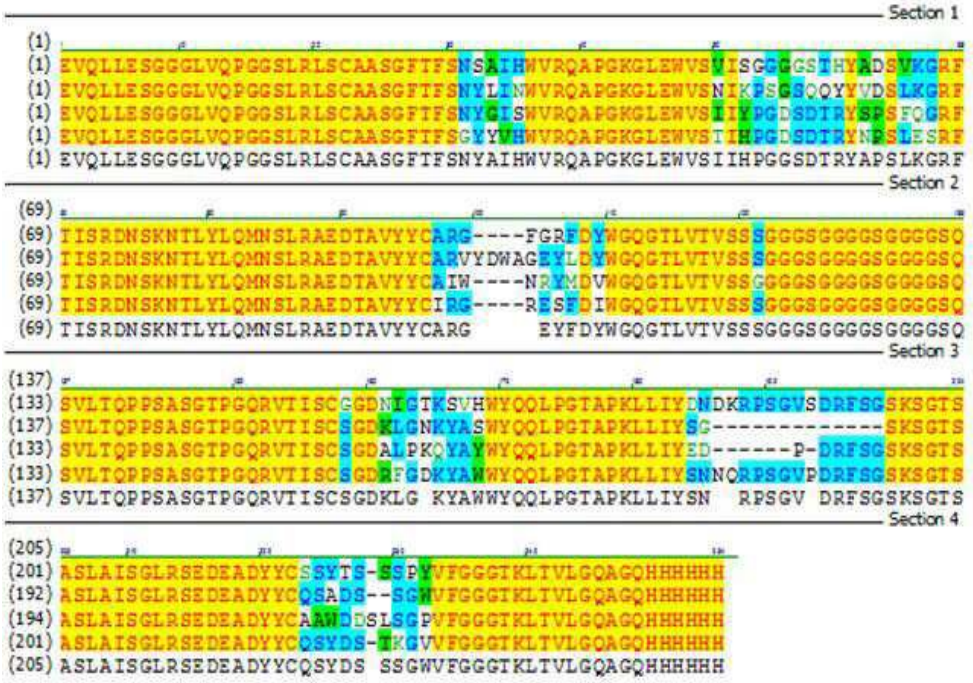

OL
EL
OK
EK

도면4

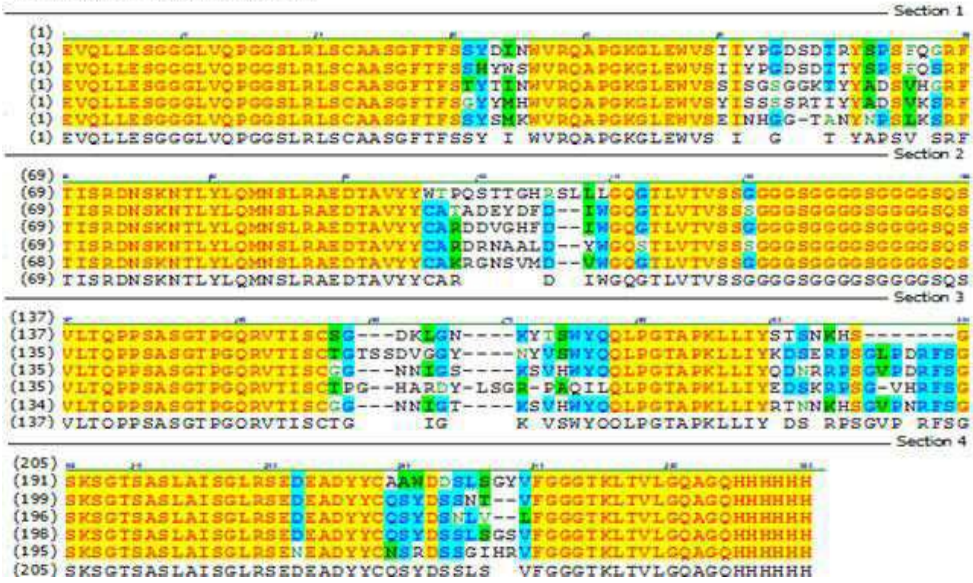

CDR seq's from natural human antibodies
Germline usage
positional mutation pattern & freq.
Length distribution
Simulation of 1,000s of pre-determined CDR sequences
Removal of sequences with PTM motifs
Array synthesis of designed CDR seq's

도면5a

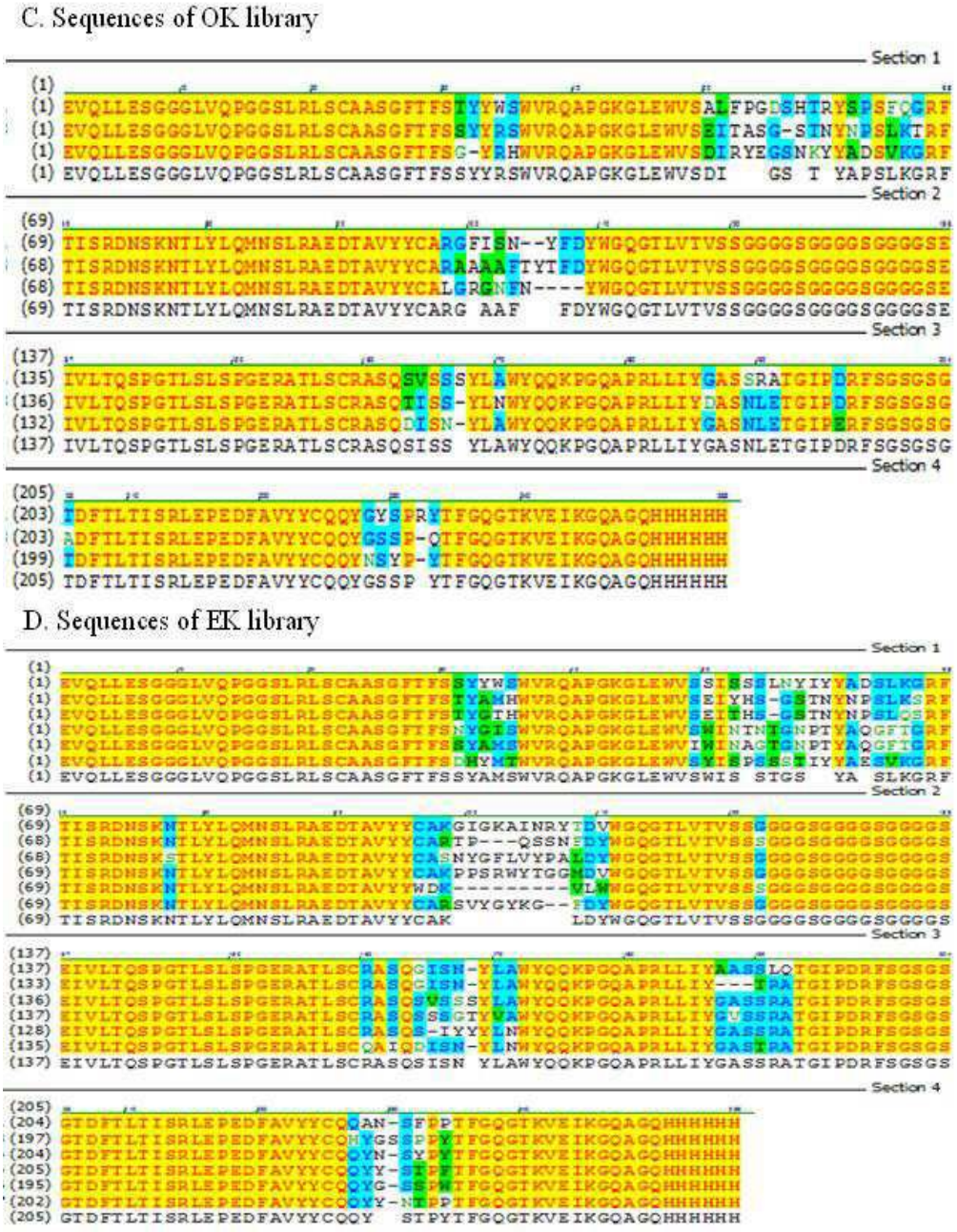
A. Sequences of OL library



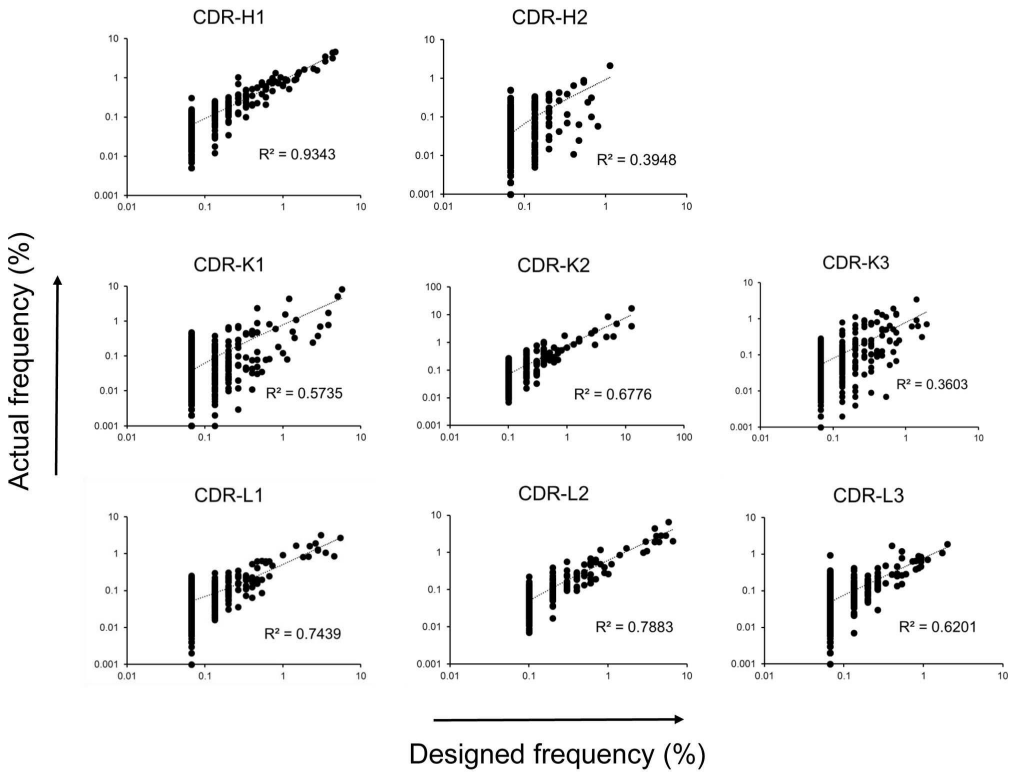
B. Sequences of EL library



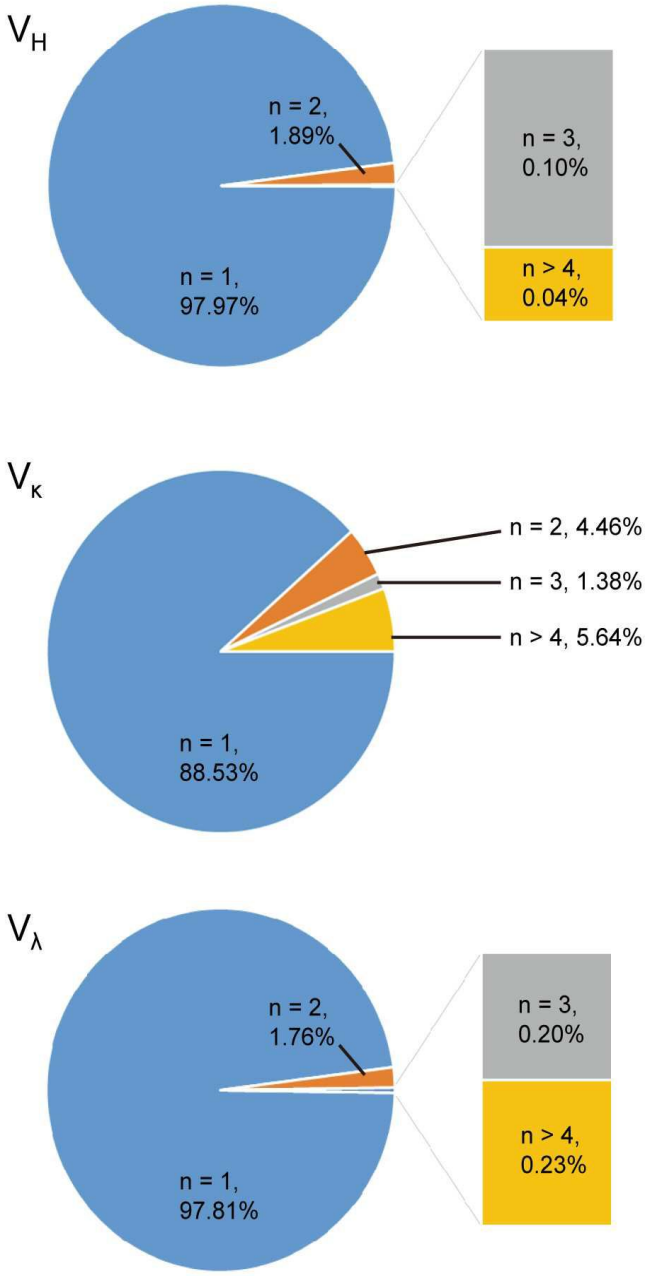
도면5b


C. Sequences of OK library
Section 1
Section 2
Section 3
Section 4
D. Sequences of EK library
Section 1
Section 2
Section 3
Section 4

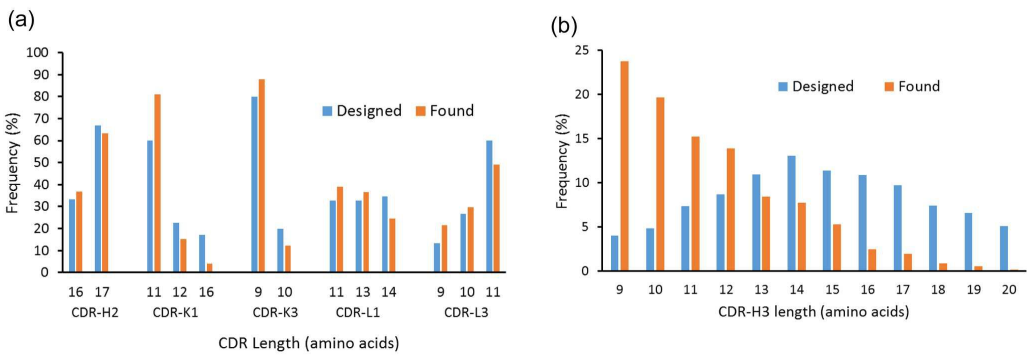
도면6


CDR-H1
R² = 0.9343
CDR-H2
R² = 0.3948
CDR-K1
R² = 0.5735
CDR-K2
R² = 0.6776
CDR-K3
R² = 0.3603
CDR-L1
R² = 0.7439
CDR-L2
R² = 0.7883
CDR-L3
R² = 0.6201
Actual frequency (%)
Designed frequency (%)

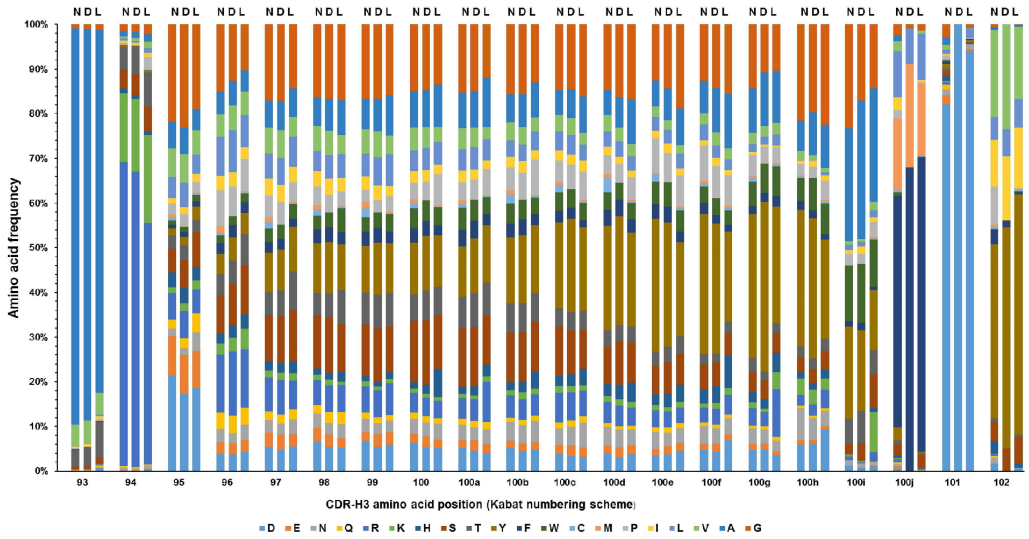
도면7


V_H
n = 2,
1.89%
n = 3,
0.10%
n > 4,
0.04%
n = 1,
97.97%
V_K
n = 2, 4.46%
n = 3, 1.38%
n > 4, 5.64%
n = 1,
88.53%
V_λ
n = 2,
1.76%
n = 3,
0.20%
n > 4,
0.23%
n = 1,
97.81%

도면8


(a)
Frequency (%)
Designed
Found
16 17
CDR-H2
11 12 16
CDR-K1
9 10
CDR-K3
11 13 14
CDR-L1
9 10 11
CDR-L3
CDR Length (amino acids)
(b)
Frequency (%)
Designed
Found
9 10 11 12 13 14 15 16 17 18 19 20
CDR-H3 length (amino acids)

도면9



서 열 목 록

<110> Ewha University - Industry Collaboration Foundation

<120> Novel method for generating a antibody library and the generated library therefrom

<130> KPA141431-KR-P1

<150> KR 10-2015-0006358

<151> 2015-01-13

<160> 252

<170> KoPatentIn 3.0

<210> 1

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-h1-f

<400> 1

cagcggattc accttcagc 19

<210> 2

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-h1-b primer

<400> 2

aggtgcttgg cgaaccca 18

<210> 3

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-h2-f primer

<400> 3

ggcctggaat gggtgagc 18

<210> 4

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-h2-b primer

<400> 4

gcggctgatg gtaaagcg 18

<210> 5

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-h3-f primer

<400> 5

ggacaccgca gtctactact 20

<210> 6

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-h3-b primer

<400> 6

caccagagta ccttgtccc 19

<210> 7

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-k1-f primer

<400> 7

cgcgcaacac tgtcatgc 18

<210> 8

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-k1-b primer

<400> 8

tggagcctga cctggtttc 19

<210> 9

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-k2-f primer

<400> 9

ccaggtcagg ctccacgt 18

<210> 10

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-k2-b primer

<400> 10

accgcttcca gatcctgag 19

<210> 11

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-k3-f primer

<400> 11

ctggaacctg aggactttg 19

<210> 12

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-k3-b primer

<400> 12

actttagtgc cctgaccg 18

<210> 13

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-l1-f primer

<400> 13

gcgcgtgact attagctgt 19

<210> 14

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-l1-b primer

<400> 14

aggtgcagtt ccaggcagt 19

<210> 15

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-l2-f primer

<400

> 15

gcctggaact gcacctaag 19

<210> 16

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-l2-b primer

<400> 16

gcctgatttg ctaccgcta 19

<210> 17

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-13-f primer

<400> 17

cttcgctccg aagatgaag 19

<210> 18

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-13-b primer

<400> 18

gtcagcttgg taccgcca 18

<210> 19

<211> 717

<212> DNA

<213> Artificial Sequence

<220><223> scFv VH3-23_linker_VK3-A27

<400> 19

gaagtgcagc tgctggaaag tggaggtgga ctggtgcagc ctggcggcag cctgcgcctg 60

agctgtgccg ccagcggatt caccttcagc agctatgcca tgagctgggt tcgccaagca 120

cctggcaaag gcctggaatg ggtgagcgcc atcagcggca gcggcggcag cacctattat 180

gccgatagcg tgaaaggccg ctttaccatc agccgcgata acagcaaaaa caccctgtat 240

ctgcagatga acagcctgcg cgccgaggac accgcagtct actactgtgc caaaaacaaa 300

ctgtggtttg attattgggg acaaggtact ctggtgaccg tgagcagcgg tggaggaggt 360

agcggaggtg gtggatctgg aggtggaggt agtgaaatcg tgctgaccca gagccctggc 420

accctgagcc tgagccctgg cgaacgcgca acactgtcat gccgcgccag ccagagcgtg 480

agcagcagct atctggcctg gtatcagcag aaaccaggtc aggctccacg tctgctgatc 540

tatggcgcca gcagccgcgc caccggcatc cctgatcgct tctcaggatc tggaagcggt 600

accgatttta ccctgaccat cagccgcctg gaacctgagg actttgccgt gtattattgt 660

cagcagtatg gcagcagccc ttggaccttc ggtcagggca ctaaagtgga aatcaaa 717

<210> 20

<211> 723

<212> DNA

<213> Artificial Sequence

<220><223> scFv VH3-23_linker_VL1g

<400> 20

gaagtgcagc tgctggaaag tggaggtgga ctggtgcagc ctggcggcag cctgcgcctg 60

agctgtgccg ccagcggatt caccttcagc agctatgcca tgagctgggt tcgccaagca 120

cctggcaaag gcctggaatg ggtgagcgcc atcagcggca gcggcggcag cacctattat 180

gccgatagcg tgaaaggccg ctttaccatc agccgcgata acagcaaaaa caccctgtat 240

ctgcagatga acagcctgcg cgccgaggac accgcagtct actactgtgc caaaaacaaa 300

ctgtggtttg attattgggg acaaggtact ctggtgaccg tgagcagcgg tggaggaggt 360

agcggaggtg gtggatctgg aggtggaggt agtcagagcg tgctgaccca gcctcctagc 420

gcctccggta caccaggaca gcgcgtgact attagctgta gcggcagcag cagcaacatc 480

ggcagcaact atgtgtattg gtaccagcaa ctgcctggaa ctgcacctaa gctgctgatc 540

tatcgcaaca accagcgccc tagcggcgtg cctgatcgct ttagcggtag caaatcaggc 600

accagcgcca gcctggccat cagcggcctt cgctccgaag atgaagccga ttattattgt 660

gccgcctggg atgatagcct gagcggctat gtgtttggtg gcggtaccaa gctgaccgtg 720

ctg 723

<210> 21

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-H1-rc-f primer

<400> 21

tgggttcgcc aagcacct 18

<210> 22

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-H2-rc-f primer

<400> 22

cgctttacca tcagccgc 18

<210> 23

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-H3-rc-f primer

<400> 23

gggacaaggt actctggtg 19

<210> 24

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-L1-rc-f primer

<400> 24

actgcctgga actgcacct 19

<210> 25

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-L2-rc-f primer

<400> 25

tagcggtagc aaatcaggc 19

<210> 26

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-L3-rc-f primer

<400> 26

tggcggtacc aagctgac 18

<210> 27

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-K1-rc-f primer

<400> 27

gaaaccaggt caggctcca 19

<210> 28

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-K2-rc-f primer

<400> 28

ctcaggatct ggaagcggt 19

<210> 29

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-K3-rc-f primer

<400> 29

cggtcagggc actaaagt 18

<210> 30

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-H1-rc-b primer

<400> 30

gctgaaggtg aatccgctg 19

<210> 31

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-H2-rc-b primer

<400> 31

gctcacccat tccaggcc 18

<210> 32

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-H3-rc-b primer

<400> 32

agtagtagac tgcggtgtcc 20

<210> 33

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-L1-rc-b primer

<400> 33

acagctaata gtcacgcgc 19

<210> 34

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-L2-rc-b primer

<400> 34

cttaggtgca gttccaggc 19

<210> 35

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-L3-rc-b primer

<400> 35

cttcatcttc ggagcgaag 19

<210> 36

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-K1-rc-b primer

<400> 36

gcatgacagt gttgcgcg 18

<210> 37

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-K2-rc-b primer

<400> 37

acgtggagcc tgacctgg 18

<210> 38

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-K3-rc-b primer

<400> 38

caaagtcctc aggttccag 19

<210> 39

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> pC3x-b primer

<400> 39

aaccatcgat agcagcaccg 20

<210> 40

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> pUC57-b primer

<400> 40

ttcgccattc aggctgcg 18

<210> 41

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> pC3x-f primer

<400> 41

gcacgacagg tttcccgac 19

<210> 42

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> lFR3-f primer

<400> 42

ctggccatca gcggccttc 19

<210> 43

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> lFR3-b primer

<400> 43

ctgggtcagc acgatttc 18

<210> 44

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> kFR1-f primer

<400> 44

gaaatcgtgc tgacccag 18

<210> 45

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> kFR1-b primer

<400> 45

ctgggtcagc acgatttc 18

<210> 46

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-FR4-f primer

<400> 46

ctggtgaccg tgagcagc 18

<210> 47

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> OPALS-FR4-b primer

<400> 47

gctgctcacg gtcaccag 18

<210> 48

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> pC3-seq primer

<400

> 48

gtgagcggat aacaattga 19

<210> 49

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> dp-seq primer

<400> 49

agaagcgtag tccggaacg 19

<210> 50

<211> 627

<212> DNA

<213> Artificial Sequence

<220><223> (VH3-23)-linker-(Vk3-A27) frame work

<400> 50

gaagtgcagc tgctggaaag tggaggtgga ctggtgcagc ctggcggcag cctgcgcctg 60

agctgtgccg ccagcggatt caccttcagc nnnnnnnnnn nnnnntgggt tcgccaagca 120

cctggcaaag gcctggaatg ggtgnnnnnn nnnnnnnnnc gctttaccat cagccgcgat 180

aacagcaaaa acaccctgta tctgcagatg aacagcctgc gcgccgagga caccgcagtc 240

tactactgtn nnnnnnnnnn nnnntgggga caaggtactc tggtgaccgt gagcagcggt 300

ggaggaggta gcggaggtgg tggatctgga ggtggaggta gtgaaatcgt gctgacccag 360

agccctggca ccctgagcct gagccctggc gaacgcgcaa cactgtcatg cnnnnnnnnn 420

nnnnnntggt atcagcagaa accaggtcag gctccacgtc tgctgatcta tnnnnnnnnn 480

nnnnnnggca tccctgatcg cttctcagga tctggaagcg gtaccgattt taccctgacc 540

atcagccgcc tggaacctga ggactttgcc gtgtattatt gtnnnnnnnn nnnnnnnttc 600

ggtcagggca ctaaagtgga aatcaaa 627

<210> 51

<211> 211

<212> PRT

<213> Artificial Sequence

<220><223> (VH3-23)-linker-(Vk3-A27) frame work a.a.

<400> 51

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Xaa Xaa
20 25 30

Xaa Xaa Xaa Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Xaa Xaa Xaa Xaa Xaa Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn
50 55 60

Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
65 70 75 80

Tyr Tyr Cys Ala Lys Xaa Xaa Xaa Xaa Xaa Trp Gly Gln Gly Thr Leu
85 90 95

Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

100 105 110

Gly Gly Gly Ser Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser
115 120 125

Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Xaa Xaa Xaa Xaa Xaa
130 135 140

Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile Tyr Xaa

145 150 155 160

Xaa Xaa Xaa Xaa Gly Ile Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly

165 170 175

Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu Pro Glu Asp Phe Ala

180 185 190

Val Tyr Tyr Cys Xaa Xaa Xaa Xaa Xaa Phe Gly Gln Gly Thr Lys Val

195 200 205

Glu Ile Lys

210

<210> 52

<211> 624

<212> DNA

<213> Artificial Sequence

<220><223> (VH3-23)-linker-(Vl1g) frame work

<400> 52

gaagtgcagc tgctggaaag tggaggtgga ctggtgcagc ctggcggcag cctgcgcctg 60

agctgtgccg ccagcggatt caccttcagc nnnnnnnnnn nnnnntgggt tcgccaagca 120

cctggcaaag gcctggaatg ggtgnnnnnn nnnnnnnnnc gctttaccat cagccgcgat 180

aacagcaaaa acaccctgta tctgcagatg aacagcctgc gcgccgagga caccgcagtc 240

tactactgtn nnnnnnnnnn nnnntgggga caaggtactc tggtgaccgt gagcagcggt 300

ggaggaggta gcggaggtgg tggatctgga ggtggaggta gtcagagcgt gctgacccag 360

cctcctagcg cctccggtac accaggacag cgcgtgacta ttagctgtnn nnnnnnnnnn 420

nnntggtacc agcaactgcc tggaactgca cctaagctgc tgatctatnn nnnnnnnnnn 480

nnnggcgtgc ctgatcgctt tagcggtagc aaatcaggca ccagcgccag cctggccatc 540

agcggccttc gctccgaaga tgaagccgat tattattgtn nnnnnnnnnn nnnnnttggt 600

ggcggtacca agctgaccgt gctg 624

<210> 53

<211> 210

<212> PRT

<213> Artificial Sequence

<220><223> (VH3-23)-linker-(Vl1g) frame work a.a.

<400> 53

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Xaa Xaa
20 25 30

Xaa Xaa Xaa Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45
Xaa Xaa Xaa Xaa Xaa Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn
50 55 60
Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val
65 70 75 80
Tyr Tyr Cys Ala Lys Xaa Xaa Xaa Xaa Xaa Trp Gly Gln Gly Thr Leu
85 90 95
Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly

100 105 110
Gly Gly Gly Ser Gln Ser Val Leu Thr Gln Pro Pro Ser Ala Ser Gly
115 120 125
Thr Pro Gly Gln Arg Val Thr Ile Ser Cys Xaa Xaa Xaa Xaa Xaa Trp
130 135 140
Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu Ile Tyr Xaa Xaa
145 150 155 160
Xaa Xaa Xaa Gly Val Pro Asp Arg Phe Ser Gly Ser Lys Ser Gly Thr
165 170 175

Ser Ala Ser Leu Ala Ile Ser Gly Leu Arg Ser Glu Asp Glu Ala Asp
180 185 190
Tyr Tyr Cys Xaa Xaa Xaa Xaa Xaa Phe Gly Gly Gly Thr Lys Leu Thr
195 200 205
Val Leu
210
<210> 54
<211> 5
<212> PRT
<213> Homo sapiens

<400> 54

Asp His Tyr Met Asp

1 5

<210> 55

<211> 5

<212> PRT

<213> Homo sapiens

<400> 55

Asp Tyr Ala Met His

1 5

<

210> 56

<211> 5

<212> PRT

<213> Homo sapiens

<400> 56

Asp Tyr Ala Met Ser

1 5

<210> 57

<211> 5

<212> PRT

<213> Homo sapiens

<400> 57

Asp Tyr Gly Met Ser

1 5

<210> 58

<211> 5

<212> PRT

<213> Homo sapiens

<400> 58

Asp Tyr Thr Met His

1 5

<210> 59

<211> 5

<212> PRT

<213> Homo sapiens

<400> 59

Asp Tyr Tyr Met His

1 5

<210> 60

<211

> 5

<212> PRT

<213> Homo sapiens

<400> 60

Asp Tyr Tyr Met Ser

1 5

<210> 61

<211> 5

<212> PRT

<213> Homo sapiens

<400> 61

Glu Leu Ser Met His

1 5

<210> 62

<211> 5

<212> PRT

<213> Homo sapiens

<400> 62

Gly Ser Ala Met His

1 5

<210> 63

<211> 5

<212> PRT

<213> Homo sapiens

<400> 63

Gly Tyr Tyr Met His

1 5

<210> 64

<211> 5

<212> PRT

<213> Homo sapiens

<400> 64

Gly Tyr Tyr Trp Ser
1 5

<210> 65

<211> 5

<212> PRT

<213> Homo sapiens

<400> 65

Asn Ala Trp Met Ser
1 5

<210> 66

<211> 5

<212> PRT

<213> Homo sapiens

<400> 66

Ser Asn Glu Met Ser
1 5

<210> 67

<211> 5

<212> PRT

<213> Homo sapiens

<400> 67

Ser Asn Tyr Met Ser
1 5

<210> 68

<211> 5

<212> PRT

<213> Homo sapiens

<400> 68

Ser Ser Ala Val Gln
1 5

<210> 69

<211> 5

<212> PRT

<213> Homo sapiens

<400> 69

Ser Tyr Ala Ile Ser

1 5

<210> 70

<211> 5

<212> PRT

<213> Homo sapiens

<400> 70

Ser Tyr Ala Met His

1 5

<210> 71

<211> 5

<212> PRT

<213> Homo sapiens

<400> 71

Ser Tyr Ala Met Asn

1 5

<210> 72

<211> 5

<212> PRT

<213> Homo sapiens

<400> 72

Ser Tyr Ala Met Ser

1 5

<210> 73

<211> 5

<212> PRT

<213> Homo sapiens

<400> 73

Ser Tyr Asp Ile Asn

1 5

<210> 74

<211> 5

<212> PRT

<213> Homo sapiens

<400> 74

Ser Tyr Asp Met His

1 5

<210> 75

<211> 5

<212> PRT

<213> Homo sapiens

<400> 75

Ser Tyr Gly Ile Ser

1 5

<210> 76

<211> 5

<212> PRT

<213> Homo sapiens

<400> 76

Ser Tyr Gly Met His

1 5

<210> 77

<211> 5

<212> PRT

<213> Homo sapiens

<400> 77

Ser Tyr Ser Met Asn

1 5

<210> 78

<211> 5

<212> PRT

<213> Homo sapiens

<400> 78

Ser Tyr Trp Ile Gly

1 5

<210> 79

<211> 5

<212> PRT

<213> Homo sapiens

<400> 79

Ser Tyr Trp Ile Ser

1 5

<210> 80

<211> 5

<212> PRT

<213> Homo sapiens

<400> 80

Ser Tyr Trp Met His

1 5

<210> 81

<211

> 5

<212> PRT

<213> Homo sapiens

<400> 81

Ser Tyr Trp Met Ser

1 5

<210> 82

<211> 5

<212> PRT

<213> Homo sapiens

<400> 82

Ser Tyr Tyr Met His

1 5

<210> 83

<211> 5

<212> PRT

<213> Homo sapiens

<400> 83

Ser Tyr Tyr Trp Ser

1 5

<210> 84

<211> 5

<212> PRT

<213> Homo sapiens

<400> 84

Tyr Arg Tyr Leu His

1 5

<210> 85

<211> 16

<212> PRT

<213> Homo sapiens

<400> 85

Ala Ile Gly Thr Ala Gly Asp Thr Tyr Tyr Pro Gly Ser Val Lys Gly

1 5 10 15

<210> 86

<211> 16

<212> PRT

<213> Homo sapiens

<400> 86

Glu Ile Asn His Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys Ser

1 5 10 15

<210> 87

<211> 16

<212> PRT

<213> Homo sapiens

<400> 87

Glu Ile Tyr His Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys Ser

1 5 10 15

<210> 88

<211> 16

<212> PRT

<213> Homo sapiens

<400> 88

His Ile Phe Ser Asn Asp Glu Lys Ser Tyr Ser Thr Ser Leu Lys Ser

1 5 10 15

<210> 89

<211> 16

<212> PRT

<213> Homo sapiens

<400> 89

Leu Ile Tyr Trp Asn Asp Asp Lys Arg Tyr Ser Pro Ser Leu Lys Ser

1 5 10 15

<210> 90

<211> 16

<212>

PRT

<213> Homo sapiens

<400> 90

Arg Ile Asp Trp Asp Asp Asp Lys Phe Tyr Ser Thr Ser Leu Lys Thr

1 5 10 15

<210> 91

<211> 16

<212> PRT

<213> Homo sapiens

<400> 91

Ser Ile Tyr His Ser Gly Ser Thr Tyr Tyr Asn Pro Ser Leu Lys Ser

1 5 10 15

<210> 92

<211> 16

<212> PRT

<213> Homo sapiens

<400> 92

Ser Ile Tyr Tyr Ser Gly Ser Thr Tyr Tyr Asn Pro Ser Leu Lys Ser

1 5 10 15

<210> 93

<211> 16

<212> PRT

<213> Homo sapiens

<400> 93

Val Ile Tyr Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val Lys Gly

1 5 10 15

<210> 94

<211> 16

<212> PRT

<213> Homo sapiens

<400> 94

Tyr Ile Tyr His Ser Gly Ser Thr Tyr Tyr Asn Pro Ser Leu Lys Ser

1 5 10 15

<210> 95

<211> 16

<212>

PRT

<213> Homo sapiens

<400> 95

Tyr Ile Tyr Tyr Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys Ser

1 5 10 15

<210> 96

<211> 16

<212> PRT

<213> Homo sapiens

<400> 96

Tyr Ile Tyr Tyr Ser Gly Ser Thr Tyr Tyr Asn Pro Ser Leu Lys Ser

1 5 10 15

<210> 97

<211> 17

<212> PRT

<213> Homo sapiens

<400> 97

Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 98

<211> 17

<212> PRT

<213> Homo sapiens

<400> 98

Ala Ile Ser Ser Asn Gly Gly Ser Thr Tyr Tyr Ala Asn Ser Val Lys

1 5 10 15

Gly

<210> 99

<211> 17

<212> PRT

<213> Homo sapiens

<400> 99

Gly Phe Asp Pro Glu Asp Gly Glu Thr Ile Tyr Ala Gln Lys Phe Gln

1 5 10 15

Gly

<210> 100

<211> 17

<212> PRT

<213> Homo sapiens

<400> 100

Gly Ile Ile Pro Ile Phe Gly Thr Ala Asn Tyr Ala Gln Lys Phe Gln

1 5 10 15

Gly

<210> 101

<211> 17

<212> PRT

<213> Homo sapiens

<400> 101

Gly Ile Asn Trp Asn Gly Gly Ser Thr Gly Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 102

<211> 17

<212> PRT

<213> Homo sapiens

<400> 102

Gly Ile Ser Trp Asn Ser Gly Ser Ile Gly Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 103

<211> 17

<212> PRT

<213> Homo sapiens

<400> 103

Ile Ile Asn Pro Ser Gly Gly Ser Thr Ser Tyr Ala Gln Lys Phe Gln

1 5 10 15

Gly

<210> 104

<211> 17

<212> PRT

<213> Homo sapiens

<400> 104

Ile Ile Tyr Pro Gly Asp Ser Asp Thr Arg Tyr Ser Pro Ser Phe Gln

1 5 10 15

Gly

<210> 105

<211> 17

<212> PRT

<213> Homo sapiens

<400> 105

Leu Ile Ser Trp Asp Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 106

<211> 17

<212> PRT

<213> Homo sapiens

<400> 106

Leu Val Asp Pro Glu Asp Gly Glu Thr Ile Tyr Ala Glu Lys Phe Gln

1 5 10 15

Gly

<210> 107

<211> 17

<212> PRT

<213> Homo sapiens

<400> 107

Asn Ile Lys Gln Asp Gly Ser Glu Lys Tyr Tyr Val Asp Ser Val Lys

1 5 10 15

Gly

<210> 108

<211> 17

<212> PRT

<213> Homo sapiens

<400> 108

Arg Ile Asp Pro Ser Asp Ser Tyr Thr Asn Tyr Ser Pro Ser Phe Gln

1 5 10 15

Gly

<210> 109

<211> 17

<212> PRT

<213> Homo sapiens

<400> 109

Arg Ile Asn Ser Asp Gly Ser Ser Thr Ser Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 110

<211> 17

<212> PRT

<213> Homo sapiens

<400> 110

Ser Ile Ser Ser Ser Ser Ser Tyr Ile Tyr Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 111

<211> 17

<212> PRT

<213> Homo sapiens

<400> 111

Val Ile Ser Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 112

<211> 17

<212> PRT

<213> Homo sapiens

<400> 112

Val Ile Trp Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser Val Lys
1 5 10 15

Gly

<210> 113

<211> 17

<212> PRT

<213> Homo sapiens

<400> 113

Trp Ile Asn Ala Gly Asn Gly Asn Thr Lys Tyr Ser Gln Lys Phe Gln
1 5 10 15

Gly

<210> 114

<211> 17

<212> PRT

<213> Homo sapiens

<400> 114

Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe Gln
1 5 10 15

Gly

<210> 115

<211> 17

<212> PRT

<213> Homo sapiens

<400> 115

Trp Ile Asn Thr Asn Thr Gly Asn Pro Thr Tyr Ala Gln Gly Phe Thr
1 5 10 15

Gly

<210> 116

<211> 17

<212> PRT

<213> Homo sapiens

<400> 116

Trp Ile Ser Ala Tyr Asn Gly Asn Thr Asn Tyr Ala Gln Lys Leu Gln

1 5 10 15

Gly

<210> 117

<211> 17

<212> PRT

<213> Homo sapiens

<400> 117

Trp Ile Thr Pro Phe Asn Gly Asn Thr Asn Tyr Ala Gln Lys Phe Gln

1 5 10 15

Asp

<210> 118

<211> 17

<212> PRT

<213> Homo sapiens

<400> 118

Trp Ile Val Val Gly Ser Gly Asn Thr Asn Tyr Ala Gln Lys Phe Gln

1 5 10 15

Glu

<210> 119

<211> 17

<212> PRT

<213> Homo sapiens

<400> 119

Trp Met Asn Pro Asn Ser Gly Asn Thr Gly Tyr Ala Gln Lys Phe Gln

1 5 10 15

Gly

<210> 120

<211> 17

<212> PRT

<213> Homo sapiens

<400> 120

Tyr Ile Ser Ser Ser Gly Ser Thr Ile Tyr Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 121

<211> 17

<212> PRT

<213> Homo sapiens

<400> 121

Tyr Ile Ser Ser Ser Ser Ser Thr Ile Tyr Tyr Ala Asp Ser Val Lys

1 5 10 15

Gly

<210> 122

<211> 11

<212> PRT

<213> Homo sapiens

<400> 122

Lys Ala Ser Gln Asp Ile Asp Asp Asp Met Asn

1 5 10

<210> 123

<211> 11

<212> PRT

<213> Homo sapiens

<400> 123

Gln Ala Ser Glu Gly Ile Gly Asn Tyr Leu Tyr

1 5 10

<210> 124

<211> 11

<212> PRT

<213> Homo sapiens

<400> 124

Gln Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn

1 5 10

<210> 125

<211> 11

<212> PRT

<213> Homo sapiens

<400> 125

Arg Ala Arg Gln Gly Ile Ser Asn Tyr Leu Ala

1 5 10

<210> 126

<211> 11

<212> PRT

<213> Homo sapiens

<400> 126

Arg Ala Ser Gln Gly Ile Arg Asn Asp Leu Gly

1 5 10

<210> 127

<211> 11

<212> PRT

<213> Homo sapiens

<400> 127

Arg Ala Ser Gln Gly Ile Ser Asn Tyr Leu Ala

1 5 10

<210> 128

<211> 11

<212> PRT

<213> Homo sapiens

<400> 128

Arg Ala Ser Gln Gly Ile Ser Ser Ala Leu Ala

1 5 10

<210> 129

<211> 11

<212> PRT

<213> Homo sapiens

<400> 129

Arg Ala Ser Gln Gly Ile Ser Ser Trp Leu Ala

1 5 10

<210> 130

<211> 11

<212> PRT

<213> Homo sapiens

<400> 130

Arg Ala Ser Gln Gly Ile Ser Ser Tyr Leu Ala

1 5 10

<210> 131

<211> 11

<212> PRT

<213> Homo sapiens

<400> 131

Arg Ala Ser Gln Gly Val Ser Ser Tyr Leu Ala

1 5 10

<210> 132

<211> 11

<212> PRT

<213> Homo sapiens

<400> 132

Arg Ala Ser Gln Ser Ile Gly Ser Ser Leu His

1 5 10

<210> 133

<211> 11

<212> PRT

<213> Homo sapiens

<400> 133

Arg Ala Ser Gln Ser Ile Ser Ser Trp Leu Ala

1 5 10

<210> 134

<211> 11

<212> PRT

<213> Homo sapiens

<400> 134

Arg Ala Ser Gln Ser Ile Ser Ser Tyr Leu Asn

1 5 10

<210> 135

<211> 11

<212> PRT

<213> Homo sapiens

<400> 135

Arg Ala Ser Gln Ser Val Ser Ser Asn Leu Ala

1 5 10

<210> 136

<211> 11

<212> PRT

<213> Homo sapiens

<400> 136

Arg Ala Ser Gln Ser Val Ser Ser Tyr Leu Ala

1 5 10

<210> 137

<211> 11

<212> PRT

<213> Homo sapiens

<400> 137

Arg Met Ser Gln Gly Ile Ser Ser Tyr Leu Ala

1 5 10

<210> 138

<211> 11

<212> PRT

<213> Homo sapiens

<400> 138

Trp Ala Ser Gln Gly Ile Ser Ser Tyr Leu Ala

1 5 10

<210> 139

<211> 12

<212> PRT

<213> Homo sapiens

<400> 139

Gly Ala Ser Gln Ser Val Ser Ser Ser Tyr Leu Ala

1 5 10

<210> 140

<211> 12

<212> PRT

<213> Homo sapiens

<400> 140

Arg Ala Ser Gln Ser Val Ser Ser Ser Tyr Leu Ala

1 5 10

<210> 141

<211> 12

<212> PRT

<213> Homo sapiens

<400> 141

Arg Ala Ser Gln Ser Val Ser Ser Ser Tyr Leu Ser

1 5 10

<210> 142

<211> 16

<212> PRT

<213> Homo sapiens

<400> 142

Lys Ser Ser Gln Ser Leu Leu His Ser Asp Gly Lys Thr Tyr Leu Tyr

1 5 10 15

<210> 143

<211> 16

<212> PRT

<213> Homo sapiens

<400> 143

Arg Ser Ser Gln Ser Leu Leu His Ser Asn Gly Tyr Asn Tyr Leu Asp

1 5 10 15

<210> 144

<211> 16

<212> PRT

<213> Homo sapiens

<400> 144

Arg Ser Ser Gln Ser Leu Val His Ser Asp Gly Asn Thr Tyr Leu Ser

1 5 10 15

<210> 145

<211> 16

<212> PRT

<213> Homo sapiens

<400> 145

Arg Ser Ser Gln Ser Leu Val Tyr Ser Asp Gly Asn Thr Tyr Leu Asn

1 5 10 15

<210> 146

<211> 7

<212> PRT

<213> Homo sapiens

<400> 146

Ala Ala Ser Ser Leu Gln Ser

1 5

<210> 147

<211> 7

<212> PRT

<213> Homo sapiens

<400> 147

Ala Ala Ser Thr Leu Gln Ser

1 5

<210> 148

<211> 7

<212> PRT

<213> Homo sapiens

<400> 148

Asp Ala Ser Asn Leu Glu Thr

1 5

<210> 149

<211> 7

<212> PRT

<213> Homo sapiens

<400> 149

Asp Ala Ser Asn Arg Ala Thr

1 5

<210> 150

<211> 7

<212> PRT

<213> Homo sapiens

<400> 150

Asp Ala Ser Ser Leu Glu Ser

1 5

<210> 151

<211> 7

<212> PRT

<213> Homo sapiens

<400> 151

Asp Ala Ser Ser Arg Ala Thr

1 5

<210> 152

<211> 7

<212> PRT

<213> Homo sapiens

<400> 152

Glu Ala Thr Thr Leu Val Pro

1 5

<210> 153

<211> 7

<212> PRT

<213> Homo sapiens

<400> 153

Glu Val Ser Asn Arg Phe Ser

1 5

<210> 154

<211> 7

<212> PRT

<213> Homo sapiens

<400> 154

Glu Val Ser Ser Arg Phe Ser

1 5

<210> 155

<211> 7

<212> PRT

<213> Homo sapiens

<400> 155

Gly Ala Ser Ser Arg Ala Thr

1 5

<210> 156

<211> 7

<212> PRT

<213> Homo sapiens

<400> 156

Gly Ala Ser Thr Arg Ala Thr

1 5

<210> 157

<211> 7

<212> PRT

<213> Homo sapiens

<400> 157

Lys Ile Ser Asn Arg Phe Ser

1 5

<210> 158

<211> 7

<212> PRT

<213> Homo sapiens

<400> 158

Lys Val Ser Asn Arg Asp Ser

1 5

<210> 159

<211> 7

<212> PRT

<213> Homo sapiens

<400> 159

Lys Val Ser Asn Trp Asp Ser

1 5

<210> 160

<211> 7

<212> PRT

<213> Homo sapiens

<400> 160

Leu Gly Ser Asn Arg Ala Ser

1 5

<210> 161

<211> 7

<212> PRT

<213> Homo sapiens

<400> 161

Thr Leu Ser Tyr Arg Ala Ser

1 5

<210> 162

<211> 7

<212> PRT

<213> Homo sapiens

<400> 162

Trp Ala Ser Thr Arg Glu Ser

1 5

<210> 163

<211> 7

<212> PRT

<213> Homo sapiens

<400> 163

Tyr Ala Ser Gln Ser Phe Ser

1 5

<210> 164

<211> 7

<212> PRT

<213> Homo sapiens

<400> 164

Tyr Ala Ser Gln Ser Ile Ser

1 5

<210> 165

<211> 7

<212> PRT

<213> Homo sapiens

<400> 165

Tyr Ala Ser Ser Leu Gln Ser

1 5

<210> 166

<211> 11

<212> PRT

<213> Homo sapiens

<400> 166

Gly Gly Asn Asn Ile Gly Ser Lys Asn Val His

1 5 10

<210> 167

<211> 11

<212> PRT

<213> Homo sapiens

<400> 167

Gly Gly Asn Asn Ile Gly Ser Lys Ser Val His

1 5 10

<210> 168

<211> 11

<212> PRT

<213> Homo sapiens

<400> 168

Gln Gly Asp Ser Leu Arg Ser Tyr Tyr Ala Ser

1 5 10

<210> 169

<211> 11

<212> PRT

<213> Homo sapiens

<400> 169

Ser Gly Asp Ala Leu Pro Lys Lys Tyr Ala Tyr

1 5 10

<210> 170

<211> 11

<212> PRT

<213> Homo sapiens

<400> 170

Ser Gly Asp Ala Leu Pro Lys Gln Tyr Ala Tyr

1 5 10

<210> 171

<211> 11

<212> PRT

<213> Homo sapiens

<400> 171

Ser Gly Asp Lys Leu Gly Asp Lys Tyr Ala Cys

1 5 10

<210> 172

<211> 11

<212> PRT

<213> Homo sapiens

<400> 172

Ser Gly Asp Val Leu Ala Lys Lys Tyr Ala Arg

1 5 10

<210> 173

<211> 11

<212> PRT

<213> Homo sapiens

<400> 173

Ser Gly Asp Val Leu Gly Glu Asn Tyr Ala Asp

1 5 10

<210> 174

<211> 11

<212> PRT

<213> Homo sapiens

<400> 174

Ser Gly Glu Ala Leu Pro Lys Lys Tyr Ala Tyr

1 5 10

<210> 175

<211> 13

<212> PRT

<213> Homo sapiens

<400> 175

Ser Gly Ser Ser Ser Asn Ile Gly Asn Asn Ala Val Asn

1 5 10

<210> 176

<211> 13

<212> PRT

<213> Homo sapiens

<400> 176

Ser Gly Ser Ser Ser Asn Ile Gly Asn Asn Tyr Val Ser

1 5 10

<210> 177

<211> 13

<212> PRT

<213> Homo sapiens

<400> 177

Ser Gly Ser Ser Ser Asn Ile Gly Ser Asn Thr Val Asn

1 5 10

<210> 178

<211> 13

<212> PRT

<213> Homo sapiens

<400> 178

Ser Gly Ser Ser Ser Asn Ile Gly Ser Asn Tyr Val Tyr

1 5 10

<210> 179

<211> 13

<212> PRT

<213> Homo sapiens

<400> 179

Thr Arg Ser Ser Gly Ser Ile Ala Ser Asn Tyr Val Gln

1 5 10

<210> 180

<211> 14

<212> PRT

<213> Homo sapiens

<400> 180

Ala Ser Ser Thr Gly Ala Val Thr Ser Gly Tyr Tyr Pro Asn

1 5 10

<210> 181

<211> 14

<212> PRT

<213> Homo sapiens

<400> 181

Gly Leu Ser Ser Gly Ser Val Ser Thr Ser Tyr Tyr Pro Ser

1 5 10

<210> 182

<211> 14

<212> PRT

<213> Homo sapiens

<400> 182

Gly Ser Ser Thr Gly Ala Val Thr Ser Gly His Tyr Pro Tyr

1 5 10

<210> 183

<211> 14

<212> PRT

<213> Homo sapiens

<400> 183

Met Leu Ser Ser Gly Phe Ser Val Gly Asp Phe Trp Ile Arg

1 5 10

<210> 184

<211> 14

<212> PRT

<213> Homo sapiens

<400> 184

Thr Gly Ser Ser Ser Asn Ile Gly Ala Gly Tyr Asp Val His

1 5 10

<210> 185

<211> 14

<212> PRT

<213> Homo sapiens

<400> 185

Thr Gly Thr Ser Ser Asp Val Gly Gly Tyr Asn Tyr Val Ser

1 5 10

<210> 186

<211> 14

<212> PRT

<213> Homo sapiens

<400> 186

Thr Gly Thr Ser Ser Asp Val Gly Ser Tyr Asn Leu Val Ser

1 5 10

<210> 187

<211> 14

<212> PRT

<213> Homo sapiens

<400> 187

Thr Gly Thr Ser Ser Asp Val Gly Ser Tyr Asn Arg Val Ser

1 5 10

<210> 188

<211> 14

<212> PRT

<213> Homo sapiens

<400> 188

Thr Leu Pro Ser Asp Ile Asn Val Gly Ser Tyr Asn Ile Tyr

1 5 10

<210> 189

<211> 14

<212> PRT

<213> Homo sapiens

<400> 189

Thr Leu Arg Ser Gly Ile Asn Val Gly Thr Tyr Arg Ile Tyr

1 5 10

<210> 190

<211> 7

<212> PRT

<213> Homo sapiens

<400> 190

Asp Asn Asn Lys Arg Pro Ser

1 5

<210> 191

<211> 7

<212> PRT

<213> Homo sapiens

<400> 191

Asp Thr Ser Asn Lys His Ser

1 5

<210> 192

<211> 7

<212> PRT

<213> Homo sapiens

<400> 192

Asp Val Ser Lys Arg Pro Ser

1 5

<210> 193

<211> 7

<212> PRT

<213> Homo sapiens

<400> 193

Glu Asp Asn Gln Arg Pro Ser

1 5

<210> 194

<211> 7

<212> PRT

<213> Homo sapiens

<400> 194

Glu Asp Ser Glu Arg Tyr Pro

1 5

<210> 195

<211> 7

<212> PRT

<213> Homo sapiens

<400> 195

Glu Asp Ser Lys Arg Pro Ser

1 5

<210> 196

<211> 7

<212> PRT

<213> Homo sapiens

<400> 196

Glu Val Ser Lys Arg Pro Ser

1 5

<210> 197

<211> 7

<212> PRT

<213> Homo sapiens

<400> 197

Glu Val Ser Asn Arg Pro Ser

1 5

<210> 198

<211> 7

<212> PRT

<213> Homo sapiens

<400> 198

Gly Lys Asn Asn Arg Pro Ser

1 5

<210> 199

<211> 7

<212> PRT

<213> Homo sapiens

<400> 199

Gly Asn Ser Asn Arg Pro Ser

1 5

<210> 200

<211> 7

<212> PRT

<213> Homo sapiens

<400> 200

Lys Asp Ser Glu Arg Pro Ser

1 5

<210> 201

<211> 7

<212> PRT

<213> Homo sapiens

<400> 201

Gln Asp Ser Lys Arg Pro Ser

1 5

<210> 202

<211> 7

<212> PRT

<213> Homo sapiens

<400> 202

Arg Asp Ser Asn Arg Pro Ser

1 5

<210> 203

<211> 7

<212> PRT

<213> Homo sapiens

<400> 203

Arg Asn Asn Asn Arg Pro Ser

1 5

<210> 204

<211> 7

<212> PRT

<213> Homo sapiens

<400> 204

Arg Asn Asn Gln Arg Pro Ser

1 5

<210> 205

<211> 7

<212> PRT

<213> Homo sapiens

<400> 205

Ser Asn Asn Gln Arg Pro Ser

1 5

<210> 206

<211> 7

<212> PRT

<213> Homo sapiens

<400> 206

Ser Thr Asn Thr Arg Ser Ser

1 5

<210> 207

<211> 7

<212> PRT

<213> Homo sapiens

<400> 207

Ser Thr Ser Asn Lys His Ser

1 5

<210> 208

<211> 7

<212> PRT

<213> Homo sapiens

<400> 208

Tyr Asp Asp Leu Leu Pro Ser

1 5

<210> 209

<211> 7

<212> PRT

<213> Homo sapiens

<400> 209

Tyr Asp Ser Asp Arg Pro Ser

1 5

<210> 210

<211> 7

<212> PRT

<213> Homo sapiens

<400> 210

His Gln Ser Ser Ser Leu Pro

1 5

<210> 211

<211> 7

<212> PRT

<213> Homo sapiens

<400> 211

Leu Gln Asp Tyr Asn Tyr Pro

1 5

<210> 212

<211> 7

<212> PRT

<213> Homo sapiens

<400> 212

Leu Gln His Asp Asn Phe Pro

1 5

<210> 213

<211> 7

<212> PRT

<213> Homo sapiens

<400> 213

Leu Gln His Asn Ser Tyr Pro

1 5

<210> 214

<211> 7

<212> PRT

<213> Homo sapiens

<400> 214

Met Gln Ala Leu Gln Thr Pro

1 5

<210> 215

<211> 7

<212> PRT

<213> Homo sapiens

<400> 215

Met Gln Ala Thr Gln Phe Pro

1 5

<210> 216

<211> 7

<212> PRT

<213> Homo sapiens

<400> 216

Met Gln Gly Ile His Leu Pro

1 5

<210> 217

<211> 7

<212> PRT

<213> Homo sapiens

<400> 217

Met Gln Gly Thr His Trp Pro

1 5

<210> 218

<211> 7

<212> PRT

<213> Homo sapiens

<400> 218

Met Gln Arg Ile Glu Phe Pro

1 5

<210> 219

<211> 7

<212> PRT

<213> Homo sapiens

<400> 219

Met Gln Ser Ile Gln Leu Pro

1 5

<210> 220

<211> 7

<212> PRT

<213> Homo sapiens

<400> 220

Gln Lys Tyr Asn Ser Ala Pro

1 5

<210> 221

<211> 7

<212> PRT

<213> Homo sapiens

<400> 221

Gln Gln Ala Asn Ser Phe Pro

1 5

<210> 222

<211> 7

<212> PRT

<213> Homo sapiens

<400> 222

Gln Gln Asp Tyr Asn Leu Pro

1 5

<210> 223

<211> 7

<212> PRT

<213> Homo sapiens

<400> 223

Gln Gln Phe Asn Ser Tyr Pro

1 5

<210> 224

<211> 7

<212> PRT

<213> Homo sapiens

<400> 224

Gln Gln Gly Asn Lys His Pro

1 5

<210> 225

<211> 7

<212> PRT

<213> Homo sapiens

<400> 225

Gln Gln Leu Asn Ser Tyr Pro

1 5

<210> 226

<211> 7

<212> PRT

<213> Homo sapiens

<400> 226

Gln Gln Arg Ser Asn Trp His

1 5

<210> 227

<211> 7

<212> PRT

<213> Homo sapiens

<400> 227

Gln Gln Arg Ser Asn Trp Pro

1 5

<210> 228

<211> 7

<212> PRT

<213> Homo sapiens

<400> 228

Gln Gln Ser Tyr Ser Thr Pro

1 5

<210> 229

<211> 7

<212> PRT

<213> Homo sapiens

<400> 229

Gln Gln Tyr Asp Asn Leu Pro

1 5

<210> 230

<211> 7

<212> PRT

<213> Homo sapiens

<400> 230

Gln Gln Tyr Gly Ser Ser Pro

1 5

<210> 231

<211> 7

<212> PRT

<213> Homo sapiens

<400> 231

Gln Gln Tyr Asn Asn Trp Pro

1 5

<210> 232

<211> 7

<212> PRT

<213> Homo sapiens

<400> 232

Gln Gln Tyr Asn Ser Tyr Pro

1 5

<210> 233

<211> 7

<212> PRT

<213> Homo sapiens

<400> 233

Gln Gln Tyr Asn Ser Tyr Ser

1 5

<210> 234

<211> 7

<212> PRT

<213> Homo sapiens

<400> 234

Gln Gln Tyr Tyr Ser Phe Pro

1 5

<210> 235

<211> 7

<212> PRT

<213> Homo sapiens

<400> 235

Gln Gln Tyr Tyr Ser Thr Pro

1 5

<210> 236

<211> 7

<212> PRT

<213> Homo sapiens

<400> 236

Gln Gln Tyr Tyr Ser Tyr Pro

1 5

<210> 237

<211> 9

<212> PRT

<213> Homo sapiens

<400> 237

Ala Ala Trp Asp Asp Ser Leu Asn Gly

1 5

<210> 238

<211> 9

<212> PRT

<213> Homo sapiens

<400> 238

Ala Ala Trp Asp Asp Ser Leu Ser Gly

1 5

<210> 239

<211> 9

<212> PRT

<213> Homo sapiens

<400> 239

Cys Ser Tyr Ala Gly Ser Ser Thr Phe

1 5

<210> 240

<211> 9

<212> PRT

<213> Homo sapiens

<400> 240

Cys Ser Tyr Ala Gly Ser Tyr Thr Phe

1 5

<210> 241

<211> 9

<212> PRT

<213> Homo sapiens

<400> 241

Gly Thr Trp Asp Ser Ser Leu Ser Ala

1 5

<210> 242

<211> 9

<212> PRT

<213> Homo sapiens

<400> 242

Leu Ser Ala Asp Ser Ser Gly Thr Tyr

1 5

<210> 243

<211> 9

<212> PRT

<213> Homo sapiens

<400> 243

Asn Ser Arg Asp Ser Ser Gly Asn His

1 5

<210> 244

<211> 8

<212> PRT

<213> Homo sapiens

<400> 244

Gln Ala Trp Asp Ser Ser Thr Ala

1 5

<210> 245

<211> 9

<212> PRT

<213> Homo sapiens

<400> 245

Gln Ser Ala Asp Ser Ser Gly Thr Tyr

1 5

<210> 246

<211> 9

<212> PRT

<213> Homo sapiens

<400> 246

Gln Ser Tyr Asp Ser Ser Leu Ser Gly

1 5

<210> 247

<211> 9

<212> PRT

<213> Homo sapiens

<400> 247

Gln Val Trp Asp Ser Ser Ser Asp His

1 5

<210> 248

<211> 8

<212> PRT

<213> Homo sapiens

<400> 248

Gln Val Trp Asp Ser Ser Thr Ala

1 5

<210> 249

<211> 9

<212> PRT

<213> Homo sapiens

<400> 249

Ser Leu Tyr Thr Ser Ser Ser Thr Phe

1 5

<210> 250

<211> 9

<212> PRT

<213> Homo sapiens

<400> 250

Ser Ser Tyr Ala Gly Ser Asn Asn Phe

1 5

<210> 251

<211> 9

<212> PRT

<213> Homo sapiens

<400> 251

Ser Ser Tyr Thr Ser Ser Ser Thr Leu

1 5

<210> 252

<211> 9

<212> PRT

<213> Homo sapiens

<400> 252

Tyr Ser Thr Asp Ser Ser Gly Asn His

1 5