



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0113948
(43) 공개일자 2022년08월17일

(51) 국제특허분류(Int. Cl.)
C12Q 1/6848 (2018.01) C12Q 1/6853 (2018.01)
G01N 33/68 (2006.01)
(52) CPC특허분류
C12Q 1/6848 (2018.05)
C12Q 1/6853 (2018.05)
(21) 출원번호 10-2022-7019839
(22) 출원일자(국제) 2020년11월11일
심사청구일자 없음
(85) 번역문제출일자 2022년06월10일
(86) 국제출원번호 PCT/GB2020/052866
(87) 국제공개번호 WO 2021/094746
국제공개일자 2021년05월20일
(30) 우선권주장
1916379.9 2019년11월11일 영국(GB)

(71) 출원인
바이오크루시블 리미티드
영국 씨비22 3에프에이치 케임브리지셔 케임브리지 바브라함 바브라함 리서치 캠퍼스 더 글렌 버지 빌딩
(72) 발명자
아미즈, 나이얼
영국 씨비22 3에프에이치 케임브리지셔 케임브리지 바브라함 바브라함 리서치 캠퍼스 더 글렌 버지 빌딩 바이오크루시블 리미티드 내
윌리엄스, 한나
영국 씨비22 3에프에이치 케임브리지셔 케임브리지 바브라함 바브라함 리서치 캠퍼스 더 글렌 버지 빌딩 바이오크루시블 리미티드 내
(뒷면에 계속)
(74) 대리인
공병욱

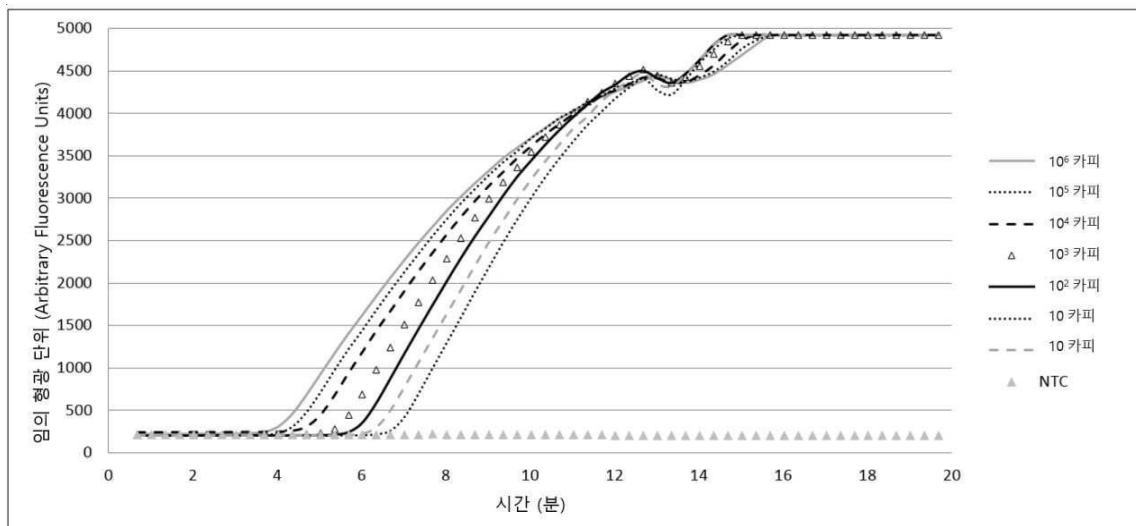
전체 청구항 수 : 총 68 항

(54) 발명의 명칭 IDR(INTRINSICALLY DISORDERED REGIONS)을 포함하는 시약 및 생화학 반응 방법

(57) 요약

본 발명은 수성(aqueous) 시험관 내(*in vitro*) 반응 시스템에서와 같은 생화학 반응을 수행하는 공정(process)에 관한 것이다. 이 공정은 1 이상의 기능적 IDR(Intrinsically Disordered Regions)을 포함하는 거대분자(macromolecule), 특히 폴리펩타이드를 포함한다. 본 발명은 또한 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 태깅된 아미노산 서열을 포함하는, 거대분자 또는 폴리펩타이드를 포함하는, IDR-폴리펩타이드를 포함하는 IDR-거대분자에 관한 것이다. 이러한 기능적 IDR은 생화학 반응의 효율을 증가시킬 수 있다. 본 발명은 이러한 거대분자 및 폴리펩타이드를 포함하는 키트에 관한 것이다. 본 발명은 또한 임의의 이러한 거대분자 및 폴리펩타이드를 다가 금속 이온과 조합하여 사용하여 용액에서 액체-액체 탈혼합(liquid-liquid demixing)을 자극 또는 향상시켜 생화학 반응의 효율을 증가시킬 수 있는 시약을 제공하는 공정에 관한 것이다.

대표도



(52) CPC특허분류

G01N 33/68 (2013.01)

C12Q 2521/507 (2013.01)

C12Q 2522/101 (2013.01)

C12Q 2527/125 (2019.08)

(72) 발명자

포레스트, 매튜

영국 씨비22 3에프에이치 케임브리지셔 케임브리지
바브라함 바브라함 리서치 캠퍼스 더 글렌 버지 빌
딩 바이오크루시블 리미티드 내

파커, 매튜

영국 씨비22 3에프에이치 케임브리지셔 케임브리지
바브라함 바브라함 리서치 캠퍼스 더 글렌 버지 빌
딩 바이오크루시블 리미티드 내

리우 시동

영국 씨비22 3에프에이치 케임브리지셔 케임브리지
바브라함 바브라함 리서치 캠퍼스 더 글렌 버지 빌
딩 바이오크루시블 리미티드 내

파커, 로렌

영국 씨비22 3에프에이치 케임브리지셔 케임브리지
바브라함 바브라함 리서치 캠퍼스 더 글렌 버지 빌
딩 바이오크루시블 리미티드 내

명세서

청구범위

청구항 1

다음에 포함하는 수성 시험관 내 반응 시스템(aqueous *in vitro* reaction system)에서 생화학 반응을 수행하는 공정에 있어서, 상기 생화학 반응은 적어도 하나의 거대분자, 선택적으로는 적어도 하나의 반응 폴리펩타이드의 기능에 의존적인, 공정:

반응을 수행하기에 적합한 조건 하에서 적어도 하나의 IDR-거대분자를 시험관 내 반응 시스템 내로 도입하는 단계, 상기 적어도 하나의 IDR-거대분자는 1개 이상의 IDR(intrinsically disordered region)을 포함하고, 적어도 하나의 IDR-거대분자가 시험관 내 반응 시스템으로 도입시 생화학 반응의 효율이 상기 적어도 하나의 IDR-거대분자에 의해 증가되고; 바람직하게는 상기 적어도 하나의 IDR-거대분자는 적어도 하나의 IDR-폴리펩타이드임.

청구항 2

제1항에 있어서, 상기 생화학 반응이 적어도 하나의 IDR-고분자, 임의로 하나 이상의 IDR-폴리펩타이드의 기능에 의존하고, 시스템 내로 도입시 적어도 하나의 IDR-거대분자 또는 적어도 하나의 IDR-폴리펩타이드는 생화학 반응에서 반응 기능을 수행하고 시스템에서 반응의 효율을 증가시키는 것인, 공정.

청구항 3

제1항 또는 제2항에 있어서, IDR-거대분자 또는 IDR-폴리펩타이드를 시스템 내에 유지하여 상기 시스템 내에서 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 액체-액체 탈혼합 및 복수의 상 분리된 수성 구획 형성을 야기하는 단계를 추가적으로 포함하고, 그렇게 함으로써 상기 시스템 내에서 생화학 반응의 효율을 증가시키는 것인, 공정.

청구항 4

제3항에 있어서, IDR-거대분자 또는 IDR-폴리펩타이드를 시스템 내에 유지하여, 상기 반응 수행에 필요한 분자를 복수의 상 분리된 수성 구획 내에서 IDR-거대분자 또는 IDR-폴리펩타이드와 공동국소화(co-localisation) 하는 단계, 그렇게 함으로써 상기 시스템 내에서 생화학 반응의 효율을 증가시키는 것인, 공정.

청구항 5

제3항 또는 제4항에 있어서, 상기 복수의 상 분리된 수성 구획은 복수의 검출가능한 상에 의해 분리된 수성 입자인, 공정.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 공정은 시험관 내 반응 시스템에서 핵산 분자를 합성하기 위한 생화학 반응인, 다음 단계를 포함하는 공정:

- (a) 적어도 하나의 핵산 프라이머를 제공하는 단계;
- (b) 적어도 하나의 타겟 가닥(target strand)을 포함하는 타겟 핵산 분자를 제공하고, 및 적어도 하나의 핵산 프라이머를 타겟 가닥과 접촉시키는 단계, 그렇게 함으로써 이중 가닥 구조를 형성함;
- (c) IDR-거대분자를 IDR-폴리펩타이드로 제공하는 단계, 상기 IDR-폴리펩타이드는 폴리머레이즈 또는 1 이상의

폴리펩타이드 보조인자(co-factor)임; 및

(d) 상기 반응이 진행되도록하는 단계, 그렇게 함으로써 폴리머레이즈 및 dNTP로 적어도 하나의 핵산 프라이머의 3' 말단을 연장하고, 선택적으로는 1 이상의 폴리펩타이드 보조인자의 존재 하에서 이중 가닥 핵산을 제조하는 단계, 제1 가닥은 타겟 가닥의 서열을 포함하고, 제2가닥은 타겟 가닥에 상보적인 서열을 포함함.

청구항 7

제6항에 있어서, 상기 공정은 시험관 내 반응 시스템에서 단일 가닥 타겟 핵산 분자 또는 이중 가닥 타겟 핵산 분자를 증폭하기 위한 생화학 반응이고, 바람직하게는 상기 타겟 핵산 분자는 DNA분자인, 공정.

청구항 8

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 공정은 시험관 내 반응 시스템 내에서 이중 가닥 타겟 핵산 분자를 증폭하기 위한 생화학 반응인, 다음 단계를 포함하는 공정:

- (a) 제1 및 제2 핵산 프라이머를 제공하는 단계;
- (b) 제1 가닥 및 제2가닥을 포함하는 이중 가닥 타겟 핵산 분자를 제공하고, 제1 및 제2 핵산 프라이머를 타겟 핵산 분자와 접촉하는 단계, 그렇게 함으로써 제1 가닥을 가진 제1 이중가닥 구조 및 제2 가닥을 가진 제2 이중가닥 구조를 형성함;
- (c) IDR-거대분자를 IDR-폴리펩타이드로 제공하는 단계, 상기 IDR-폴리펩타이드는 폴리머레이즈 또는 1 이상의 단백질 보조인자임;
- (d) 상기 반응이 진행되도록하는 단계, 그렇게 함으로써 폴리머레이즈 및 dNTP로 제1 및 제2 핵산 프라이머의 3' 말단을 연장하고, 선택적으로는 1 이상의 폴리펩타이드 보조인자의 존재 하에서 제1 및 제2 이중 가닥 핵산을 제조하는 단계; 및
- (e) 원하는 증폭 정도에 이를 때까지 (b) 내지 (d) 단계를 반복하는 단계.

청구항 9

제8항에 있어서, 상기 공정은 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자의 증폭을 위한 리코비네이즈 폴리머레이즈 증폭 공정(recombinase polymerase amplification process)이고, 다음 단계를 포함하는 공정:

- (a) 리코비네이즈 제제를 제공하는 단계, 선택적으로 리코비네이즈 로딩 단백질(recombinase loading protein), 단일 가닥 안정화제, 폴리머레이즈, 제1 및 제2 핵산 프라이머, 제1 가닥 및 제2가닥을 포함하는 이중 가닥 표적 핵산, 및 선택적으로 엑소뉴클리에이즈(exonuclease) III와 같은 엑소뉴클리에이즈;
- (b) 제1 및 제2 핵산 프라이머와 리코비네이즈 제제를 접촉하는 단계, 선택적으로 상기 리코비네이즈 로딩 단백질과 접촉하여, 이들의 3' 말단에 단일 가닥 영역을 포함하는 제1 및 제2 핵단백질 프라이머를 생성함;
- (c) 제1 및 제2 핵단백질 프라이머를 표적 핵산 분자와 접촉시킴으로써 제1 가닥을 가진 제1 이중 가닥 구조 및 제2 가닥을 가진 제2 이중 가닥 구조를 형성하는 단계;
- (d) 상기 반응이 진행되도록하는 단계, 그렇게 함으로써 폴리머레이즈 및 dNTP로 제1 및 제2 핵산 프라이머의 3' 말단을 연장하여, 제1 및 제2 이중 가닥 핵산, 제1 및 제2 변위 핵산 가닥을 형성하는 단계, 상기 단일 가닥 안정화제는 제1 및 제2 변위 가닥을 안정화하는 것임; 및
- (e) 원하는 증폭 정도에 이를 때까지 (b) 내지 (d) 단계를 반복하는 단계;

여기에서 상기 리코비네이즈 제제, 및/또는 상기 리코비네이즈 로딩 단백질, 및/또는 단일 가닥 안정화제, 및/또는 상기 폴리머레이즈는 IDR-폴리펩타이드로서 제공됨.

청구항 10

제9항에 있어서, 상기 리콤비네이즈 제제는 UvsX, T4 UvsX, T6 UvsX, RB18 UvsX, *E.coli* phage wV7 UvsX, *Shigella* phage CB8 UvsX, *Shigella* phage Shf12 UvsX, *E.coli* phage AR1 UvsX, phage vB_EcoM_G4507 UvsX, *Shigella* phage SHFML-11 UvsX, *Escherichia* phage vB_EcoM_DalCa UvsX, *E. coli* RecA, *E. coli* RadA, *E. coli* RadB, *E. coli* Rad 51 또는 이들의 기능적 유사체(functional analog), 상동체(homolog), 또는 유도체(derivative), 및 이들의 조합으로 이루어진 군으로부터 선택되는 것이고, 바람직하게는 상기 리콤비네이즈 제제는 UvsX이고, 더욱 바람직하게는 *Escherichia* phage vB_EcoM_DalCa UvsX인, 리콤비네이즈 폴리머레이즈 증폭 공정.

청구항 11

제9항 또는 제10항에 있어서, 상기 공정은 리콤비네이즈 로딩 단백질을 포함하고, 상기 리콤비네이즈 로딩 단백질을 포함하고, 상기 리콤비네이즈 로딩 단백질은 UvsY, 대장균 RecO (*E. coli* RecO), 대장균 RecR (*E. coli* RecR) 또는 이들의 기능적 유사체, 상동체, 또는 유도체, 및 이들의 임의의 조합으로 이루어진 군으로부터 선택되고, 바람직하게는 리콤비네이즈의 로딩 단백질은 UvsY이며, 더욱 바람직하게는 대장균 파지 STO UvsY(*Escherichia* phage STO UvsY)인, 리콤비네이즈 폴리머레이즈 증폭 공정.

청구항 12

제6항 내지 제 11항 중 어느 한 항에 있어서, 상기 폴리머레이즈는 pol- α , pol- β , pol- δ , pol- ϵ , 또는 이들의 기능적 유사체, 상동체, 또는 유도체, 및 이들의 조합으로 이루어진 군으로부터 선택된 진핵생물 폴리머레이즈인, 리콤비네이즈 폴리머레이즈 증폭 공정.

청구항 13

제6항 내지 제11항 중 어느 한 항에 있어서, 상기 폴리머레이즈는 바실러스 스테아로써모필루스 폴리머레이즈 I 대형 단편(*Bacillus stearothermophilus* polymerase I large fragment), 바실러스 섭틸리스 폴리머레이즈 I 대형 단편(*Bacillus subtilis* Pol I large fragment (Bsu polymerase)), 리스테리아 모노사이토제네스 DNA 폴리머레이즈 I (*Listeria monocytogenes* DNA polymerase I), 황색포도상구균 DNA 폴리머레이즈 I (*S. aureus* DNA polymerase I (Sau polymerase)), 대장균 DNA 폴리머레이즈 I 클레나우 단편 (*E. coli* DNA polymerase I Klenow fragment), 대장균 DNA 폴리머레이즈 I (*E. coli* DNA polymerase I), 대장균 DNA 폴리머레이즈 II (*E. coli* DNA polymerase II), 대장균 DNA 폴리머레이즈 III (*E. coli* DNA polymerase III), 대장균 DNA 폴리머레이즈 IV (*E. coli* DNA polymerase IV), 대장균 DNA 폴리머레이즈 V (*E. coli* DNA polymerase V), 또는 이들의 기능적 유사체, 상동체, 또는 유도체, 및 이들의 조합으로 이루어진 군으로부터 선택되고, 바람직하게는 상기 폴리머레이즈는 황색포도상구균 DNA 폴리머레이즈 I (*S. aureus* DNA polymerase I (Sau polymerase)) 또는 바실러스 섭틸리스 폴리머레이즈 I 대형 단편(*Bacillus subtilis* Pol I large fragment (Bsu polymerase))인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 14

제6항 내지 제11항 중 어느 한 항에 있어서, 상기 폴리머레이즈는 박테리오파지 T4 gp43 DNA 폴리머레이즈 (bacteriophage T4 gp43 DNA polymerase), T7 DNA 폴리머레이즈(T7 DNA polymerase) 및 파이-29 DNA 폴리머레이즈 (Phi-29 DNA polymerase), 또는 이들의 유사체, 상동체 또는 유도체, 및 이들의 조합으로 이루어진 군으로부터 선택된 박테리오파지 폴리머레이즈인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 15

제9항 내지 제14항 중 어느 한 항에 있어서, 상기 단일 가닥 안정화제는 Gp32, 대장균 SSB 단백질, 파지 T4 Gp32 단백질, 파지 Rb69 Gp32, 파지 vB_EcoM_NBG1 Gp32, 또는 이들의 기능적 유사체, 상동체 또는 유도체, 및 이들의 조합으로 이루어진 군으로부터 선택되고, 바람직하게는 상기 단일가닥 안정화제는 Gp32 또는 파지 vB_EcoM_NBG1 Gp32인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 16

제9항 내지 제15항 중 어느 한 항에 있어서, 상기 유일한 리콤비네이즈 제제는 IDR-폴리펩타이드로 제공되거나, 또는 상기 유일한 리콤비네이즈 로딩 단백질은 IDR-폴리펩타이드로 제공되거나, 또는 상기 유일한 단일 가닥 안정화제는 IDR-폴리펩타이드로 제공되거나, 또는 상기 유일한 폴리머레이즈는 IDR-폴리펩타이드로 제공되거나, 또는 상기 유일한 엑소뉴클레이즈(exonuclease)는 IDR-폴리펩타이드로 제공되는 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, 상기 IDR-폴리펩타이드의 1 이상의 기능적 IDR은, IDR-폴리펩타이드가 유전적으로 조작된 융합단백질인 것과 같은 1 이상의 IDR을 포함하거나 또는 이로 이루어진 아미노산 서열로서 태깅되고, 상기 1 이상의 기능적 IDR은 IDR-폴리펩타이드의 C-말단, N-말단, 또는 C-말단과 N-말단의 양쪽, 또는 폴리펩타이드의 길이를 따른 아미노산 위치에 위치하는 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서, 상기 IDR 거대분자 또는 IDR 폴리펩타이드의 1 이상의 기능적 IDR은 MetaDisorder 알고리즘으로 분석했을 때 0.5 이상의 스코어를 기록한 아미노산의 서열의 특징을 가진, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, 상기 IDR-거대분자 또는 IDR 폴리펩타이드의 1 이상의 기능적 IDR은 RGG 트리펩타이드의 1 이상의 반복을 포함하는 아미노산 서열을 포함하거나, 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 20

제19항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 FG 다이 펩타이드 서열의 1 이상의 반복을 추가적으로 포함하는 아미노산 서열을 포함하거나, 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 21

제19항 또는 제20항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 티로신(tyrosine) 또는 페닐알라닌(phenylalanine)으로 이루어진 적어도 하나의 방향족 아미노산 잔기를 추가적으로 포함하는 아미노산 서열을 포함하거나, 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 22

제1항 내지 제21항 중 어느 한 항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR

은 다음 아미노산 서열을 포함하거나, 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정:

- i. (YNPQGGYQQ)_n, 상기 n은 1 내지 10 사이의 양의 정수이고, 선택적으로 상기 n은 1, 2, 또는 3; 또는
- ii. (YSPTSPS)_n, 상기 n은 1 내지 10 사이의 양의 정수이고, 선택적으로 상기 n은 1, 2, 또는 3; 또는
- iii. (FSPTSPT)_n, 상기 n은 1 내지 10 사이의 양의 정수이고, 선택적으로 상기 n은 1, 2, 또는 3; 또는
- iv. (YSPTSP-A/N/G)_n, 상기 n은 1 내지 10 사이의 양의 정수이고, 선택적으로 상기 n은 1, 2, 또는 3; 또는 r
- v. (YSPGSPA)_n, 상기 n은 1 내지 10 사이의 양의 정수이고, 선택적으로 상기 n은 1, 2, 또는 3.

청구항 23

제1항 내지 제22항 중 어느 한 항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 글루타민이 풍부한 아미노산 서열, 선택적으로 상기 아미노산 서열은 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 또는 적어도 10의 연속된 글루타민 잔기를 포함하는 아미노산 서열을 포함하거나, 또는 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 24

제23항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은QQQ 트리펩타이드 서열의 1 이상의 반복을 포함하는 아미노산 서열을 포함하거나 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 25

제23항 또는 제24항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 (QQQPQY)_n의 아미노산 서열을 포함하거나, 이로 이루어진 것이고, 상기 n은 1 내지 10 사이의 양의 정수이고, 선택적으로 상기 n은 1, 2, 또는 3인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 26

제1항 내지 제25항 중 어느 한 항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 적어도 5회 연속되는 서열번호 1의 아미노산 서열을 포함하거나, 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 적어도 5회 연속되는 서열번호 9의 아미노산 서열을 포함하거나, 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 28

제1항 내지 제27항 중 어느 한 항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은, 금속 이온, 바람직하게는 2가 금속 이온과 방향족 양이온-파이 상호작용에 관여할 수 있는, 1 이상의 방향족 티로신 잔기 및 1 이상의 페닐알라닌 잔기를 포함하는 아미노산 서열을 포함하는, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 29

제1항 내지 제28항 중 어느 한 항에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR 은, 다가 금속 이온, 바람직하게는 2가 금속 이온과 구아니딘-금속 상호작용에 관여할 수 있는, 1 이상의 아르 기닌 잔기를 포함하는 아미노산 서열을 포함하는, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서, 상기 IDR-폴리펩타이드는 서열번호 1 내지 43 중 어느 하나의 아미 노산 서열, 또는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이 아미노산 서열, 예컨대 서열 번호 1 내지 43 중 어느 하나의 서열과 80% 이상의 상동성을 가진 아미노산 서열을 포함하거나 이로 이루어진 아미노산 서열로 태깅된 폴리펩타이드를 포함하거나, 이로 이루어진 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 31

제9항 내지 제30항 중 어느 한 항에 있어서, 상기 IDR-폴리펩타이드는 단일 가닥 안정화제 Gp32이고, 서열번호 65 내지 88 중 어느 하나의 아미노산 서열을 가지거나, 또는 이들의 기능적 변이체, 예컨대 서열번호 65 내지 88 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 32

제9항 내지 제30항 중 어느 한 항에 있어서, 상기 IDR-폴리펩타이드는 UvsX인 리콤비네이즈 제제고 서열번호 44 내지 59 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨 대 서열번호 44 내지 59 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 33

9항 내지 제30항 중 어느 한 항에 있어서, 상기 IDR-폴리펩타이드는 UvsY인 리콤비네이즈 로딩 단백질이고 서열번호 60 내지 64 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 60 내지 64 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서 열을 가지는 IDR-폴리펩타이드인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 34

제1항 내지 제33항 중 어느 한 항에 있어서, 시험관 내 반응 시스템 내에서 다가 금속 이온을 IDR-거대분자 또는 IDR-폴리펩타이드에 제공함으로써, 시험관 내 반응 시스템 내에서 액체-액체 탈혼합(liquid-liquid demixing)을 자극 또는 향상시키고, 생화학 반응의 효율을 증가시키는 단계; 바람직하게는 상기 액체-액체 탈혼합(liquid-liquid demixing)은 시험관 내 반응 시스템 내에서 검출가능한 상에 의해 분리된 입자(phase-separated particles)들의 생성을 촉진함으로써 생화학 반응의 효율을 증가시키는 단계; 선택적으로 상기 다가 금속 이온 은 약 22 mM 이상의 농도에서 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM 사이의 농도 에서 제공되는 것인, 리콤비네이즈 폴리머레이즈 증폭공정.

청구항 35

제34항에 있어서, 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 이고, 바람직하게는 Mg^{2+} , Mn^{2+} , 또는 Ca^{2+} 이고, 더욱 바람직하게는 Mg^{2+} 인, 리콤피네이즈 폴리머레이즈 증폭공정.

청구항 36

제1항 내지 제35항 중 어느 한 항에 있어서, 상기 생화학 반응은 표면(surface)을 포함하는 고상 반응 시스템 내에서 수행되는 것인, 리콤피네이즈 폴리머레이즈 증폭공정.

청구항 37

제36항에 있어서, 상기 생화학 반응은 제3항 또는 제4항에 정의된 시험관 내 반응 시스템에서 단일 가닥 표적 핵산 분자 또는 이중 가닥 표적 핵산 분자를 증폭시키는 공정이고, 상기 적어도 하나의 핵산 프라이머 및/또는 IDR-거대분자 및/또는 1 이상의 폴리펩타이드 보조-인자(polypeptide co-factor)는 상기 표면(surface)에 부착된 것인, 리콤피네이즈 폴리머레이즈 증폭공정.

청구항 38

제9항 내지 제34항 중 어느 한 항에 있어서, 상기 생화학 반응은 시험관 내 반응 시스템 내에서 이중 가닥 표적 핵산 분자를 증폭시키는 리콤피네이즈 폴리머레이즈 증폭 공정이고, 상기 반응은 표면을 포함하는 고상 반응 시스템에서 수행되고, 상기 리콤피네이즈 시약 및/또는 상기 리콤피네이즈 로딩 단백질 및/또는 상기 단일 가닥 안정화제 및/또는 상기 폴리머레이즈 및/또는 상기 엑소뉴클레이즈 및/또는 상기 제1 핵산 프라이머 및/또는 상기 제2 핵산 프라이머는 상기 표면에 부착된 것이고, 바람직하게는: (i) 상기 제1 핵산 프라이머 또는 제2 핵산 프라이머는 상기 표면에 부착된 것이거나; 또는 (ii) 제1 및 제2 핵산 프라이머가 모두 상기 표면에 부착된 것인, 리콤피네이즈 폴리머레이즈 증폭공정.

청구항 39

제36항 내지 제38항 중 어느 한 항에 있어서, 상기 표면은 평면, 또는 마이크로비드, 바람직하게는 상기 표면은 실리콘, 유리, 젤-기반 물질 및/또는 폴리스티렌과 같은 폴리머 물질을 포함하고, 더욱 바람직하게는 상기 표면은 폴리스티렌과 같은 폴리머 물질을 포함하는 마이크로비드인, 리콤피네이즈 폴리머레이즈 증폭공정.

청구항 40

제39항에 있어서, 상기 표면은 서브스트레이트(substrate)에 부착된 것이고, 바람직하게는 상기 표면은 평면 및/또는 유리를 포함하는 서브스트레이트(substrate)이고, 선택적으로 상기 표면 및/또는 서브스트레이트(substrate)은 플로우-셀로 제공되는 것인, 리콤피네이즈 폴리머레이즈 증폭공정.

청구항 41

거대분자 및 태그 아미노산 서열을 포함하는 비-자연적으로 발생하는 IDR-거대분자에 있어서, 상기 태그 서열은 1 이상의 기능적 IDR(intrinsically disordered region)을 포함하거나 이로 이루어지고, 상기 IDR-거대분자는 용액에서 액체-액체 탈혼합을 촉진할 수 있는 것인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 42

제41항에 있어서, 상기 액체-액체 탈혼합은 시험관 내 반응 시스템 내에서 검출 가능한 상에 의해 분리된 입자

의 생성을 촉진하는 것인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 43

제41항 또는 제42항에 있어서, 상기 액체-액체 탈혼합 및 상에 의해 분리된 수용성 분획을 생성함으로써 상기 시스템 내에서 생화학 반응의 효율을 증가시키는 것이거나, 또는 상기 검출 가능한 상에 의해 분리되는 수용성 입자를 생성함으로써 상기 시스템 내에서 생화학 반응의 효율을 증가시키는 것인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 44

제41항 내지 제43항 중 어느 한 항에 있어서, 상기 IDR-거대분자는 비자연적으로 발생하고, 인공적, 또는 유전적으로 조작된, 폴리펩타이드 및 태그 아미노산 서열을 포함하는 IDR-폴리펩타이드인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 45

제44항에 있어서, 상기 태그 아미노산 서열은 폴리펩타이드의 C-말단, 폴리펩타이드의 N-말단, 또는 폴리펩타이드의 C-말단 및 폴리펩타이드의 N-말단 둘 다, 또는 폴리펩타이드의 길이에 따른 임의의 아미노산 위치에 위치하는 것인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 46

제45항에 있어서, 상기 태그 아미노산 서열의 1 이상의 기능적 IDR은 제14항 내지 제28 항 중 어느 한 항에 정의된 기능적 IDR인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 47

제41항 내지 제46항 중 어느 한 항에 있어서, 상기 태그 서열은 다가 금속 양이온, 바람직하게는 2가 금속 양이온, 보다 바람직하게는 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 이온, 보다 더 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 이온, 보다 더 바람직하게는 Mg^{2+} 이온과의 방향족 양이온-파이 상호작용에 관여할 수 있는 아미노산 잔기를 포함하는 것인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 48

제44항 내지 제47항 중 어느 한 항에 있어서, 여기에서 1 이상의 기능적 IDR을 포함하거나, 이로 이루어진 서열이 태깅된 폴리펩타이드는 헬리케이스, 자이레이스, 리코비네이스, 예컨대 RPA 리코비네이스 제제; 뉴클레이스, 예컨대 엑소뉴클레이스 및 엔도뉴클레이스; 라이게이스(ligase), 글라이코라이에이스(glycolyase), 메틸레이스(methylase), 메틸트랜스퍼레이스(methyltransferase), 글루코실트랜스퍼레이스, 폴리머레이스, 카이네이스, 포스파타이스(phosphatase)와 같은 효소; CRISPR 효소와 같은 유전자 편집 효소, 예컨대 Cas9 효소; 보조-인자, 예컨대 RPA 리코비네이스 로딩 단백질 및 RPA 단일 가닥 안정화제인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 49

제48항에 있어서, 여기에서 1 이상의 기능적 IDR을 포함하거나, 이로 이루어진 서열이 태깅된 폴리펩타이드는 RNA 단일 가닥 안정화제, 바람직하게는 Gp32; 선택적으로 상기 IDR-폴리펩타이드는 서열번호 65 내지 88 중 어

는 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 65 내지 88 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 50

제48항에 있어서, 여기에서 1 이상의 기능적 IDR을 포함하거나, 이로 이루어진 서열이 태깅된 폴리펩타이드는 RPA 리콥비네이즈 제제, 바람직하게는 UvsX; 선택적으로 상기 IDR-폴리펩타이드는 서열번호 44 내지 59 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 44 내지 59 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 51

제48항에 있어서, 상기 여기에서 1 이상의 기능적 IDR을 포함하거나, 이로 이루어진 서열이 태깅된 폴리펩타이드는 RPA 리콥비네이즈 제제, 바람직하게는 UvsY; 선택적으로 상기 IDR-폴리펩타이드는 서열번호 60 내지 64 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 60 내지 64 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드인, 비-자연적으로 발생하는 IDR-거대분자.

청구항 52

제41항 내지 제51항 중 어느 한 항에 따른 IDR-폴리펩타이드를 인코딩하는 제1 핵산 서열을 포함하는 단리된 핵산 분자에 있어서; 선택적으로 프로모터를 인코딩하는 제2 핵산 서열을 포함하고, 상기 제1 핵산 서열은 상기 제2 핵산 서열과 작동가능하게 연결된 것인, 단리된 핵산 분자.

청구항 53

제52항에 따른 핵산 분자를 포함하는 재조합 폴리뉴클레오타이드 발현 벡터.

청구항 54

제52항에 정의된 핵산 분자, 또는 제53항에 정의된 재조합 폴리뉴클레오타이드 발현 벡터를 포함하는 숙주세포.

청구항 55

성장 배지 및 숙주세포의 집단을 포함하는 세포 배양물에 있어서, 상기 집단은 제54항에 따라 정의된 숙주 세포를 포함하는 것인, 세포 배양물.

청구항 56

제41항 내지 제51항 중 어느 한 항에 따른 비-자연적으로 발생하는 IDR-거대분자 또는 IDR-폴리펩타이드를 포함하는 키트.

청구항 57

제56항에 있어서, RPA 리콤비네이즈 제제, 및/또는 RPA 리콤비네이즈 로딩 단백질, 및/또는 폴리머레이즈, 및/또는 제1 및 제2 핵산 프라이머, 및/또는 엑소뉴클레이즈, 및/또는 버퍼, 및/또는 다가 금속 이온원(source of multivalent metal ions), 바람직하게는 2가 금속 양이온원을 포함하는, 추가적 RPA 구성을 추가적으로 포함하는, 키트

청구항 58

제56항 또는 제57항에 있어서, 상기 모든 구성은 동결건조 형태로 제공되는 것인, 키트.

청구항 59

용액에서 액체-액체 탈혼합을 자극 또는 향상시키는 공정으로서, 상기 공정은 제41항 내지 제51항 중 어느 한 항에 따른 IDR-거대분자 또는 IDR-폴리펩타이드를 포함하는 용액을 제공하는 단계, 및 용액 중 IDR-거대분자 또는 IDR-폴리펩타이드를 다가 금속 이온과 접촉시켜 용액 내 액체-액체 탈혼합을 자극 또는 향상시키는 단계를 포함하는 공정.

청구항 60

제59항에 있어서, 상기 액체-액체 탈혼합은 용액 내에서 검출가능한 상에 의해 분리된 입자를 형성하는 것인, 공정.

청구항 61

제59항 또는 제60항에 있어서, 상기 다가 금속 이온은 다가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} , 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} , 보다 바람직하게는 Mg^{2+} 인, 공정.

청구항 62

용액 내 액체-액체 탈혼합을 자극 또는 향상시키는, 다가 금속 이온의 용도에 있어서, 상기 탈혼합은 제41항 내지 제51항 중 어느 한 항의 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 매개되는 것인, 용도.

청구항 63

제62항에 있어서, 상기 액체-액체 탈혼합은 용액 내에서 검출가능한 상에 의해 분리된 입자를 형성하는 것인, 용도.

청구항 64

제62항 또는 제63항에 있어서, 상기 다가 금속 이온은 2가 금속 이온, 선택적으로는 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} , 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} , 보다 바람직하게는 Mg^{2+} 인, 용도.

청구항 65

제62항 내지 제64항 중 어느 한 항에 있어서, 상기 다가 금속 이온은 1 이상의 기능적 IDR 아미노산 서열에서 아미노산 잔기와 방향족 양이온-파이 상호작용에 관여함으로써, 액체-액체 탈혼합을 촉진하는 것인, 용도.

청구항 66

치료 또는 진단제 용도의, 제41항 내지 제48항 중 어느 한 항에 따른 비-자연적으로 발생하는 IDR-거대분자 또는 IDR-폴리펩타이드.

청구항 67

1 이상의 표적 폴리뉴클레오타이드 분자의 뉴클레오타이드 서열을 결정하는 방법에 있어서, 상기 방법은 다음 단계를 포함하는 것인 방법:

- (i) 제1항 내지 제40항 중 어느 한 항에 따른 공정을 수행하여 1 이상의 표적 폴리뉴클레오타이드 분자를 증폭함으로써 1 이상의 표적 폴리뉴클레오타이드 분자의 복수의 카피를 포함하는 집단을 얻는 단계; 및
- (ii) 표적 폴리뉴클레오타이드 분자의 복수의 카피를 포함하는 집단에 대해 1 이상의 핵산 시퀀싱 반응을 수행하는 단계, 바람직하게는 상기 방법은 표면을 포함하는 고상 반응 시스템에서 수행됨.

청구항 68

1 이상의 표적 폴리뉴클레오타이드 분자, 바람직하게는 DNA 분자의 뉴클레오타이드 서열을 결정하기 위한 방법에서 제41항 내지 제51항 중 어느 한 항에 따른 IDR-거대분자 또는 IDR-폴리펩타이드의 용도에 있어서, 바람직하게는 상기 방법은 제67항에 정의된 바와 같은, 용도.

발명의 설명

기술 분야

[0001] 본 발명은 수성(aqueous) 시험관 내(*in vitro*) 반응 시스템에서와 같은 생화학 반응을 수행하는 공정(process)에 관한 것이다. 이 공정은 1 이상의 기능적 IDR(Intrinsically Disordered Regions)을 포함하는 거대분자(macromolecule), 특히 폴리펩타이드를 포함한다. 본 발명은 또한 1 이상의 기능적 IDR을 포함하거나 이로 이루어진(comprising or consisting of) 태깅된 아미노산 서열을 포함하는, 거대분자 또는 폴리펩타이드를 포함하는, IDR-폴리펩타이드를 포함하는 IDR-거대분자(IDR-macromolecule)에 관한 것이다. 이러한 기능적 IDR은 생화학 반응의 효율을 증가시킬 수 있다. 본 발명은 이러한 거대분자 및 폴리펩타이드를 포함하는 키트에 관한 것이다. 본 발명은 또한 임의의 이러한 거대분자 및 폴리펩타이드를 다가 금속 이온과 조합하여 사용하여 용액에서 액체-액체 탈혼합(liquid-liquid demixing)을 자극 또는 향상시켜 생화학 반응의 효율을 증가시킬 수 있는 시약을 제공하는 공정에 관한 것이다.

배경 기술

[0002] 생화학 반응, 특히 시험관 내(*in vitro*) 생화학 반응의 성능은 생물학에서 근본적으로 중요하다. 많은 생화학 반응은 환자 진료 시점(at the point of care)이나 현장과 같이 실험실 외부에서 수행해야 할 수 있다. 이러한 환경에서는 실험실 환경에서 제공하는 정확한 방식으로 생화학 반응을 제어하는 것이 불가능할 수 있다. 이러한 환경에서 수행되는 생화학 반응의 효율을 높이는 것이 요구된다. 실제로, 시험관 내(*in vitro*) 및 생체 내(*in vivo*) 생화학 반응을 포함한 생화학 반응에 있어서, 정확한 설정에 관계없이 반응의 효율을 높이는 것이 요구된다. 본 발명은 이러한 문제를 다룬다.

[0003] 많은 생화학 반응은 성능 효율(driving performance efficiency)을 촉진하기 위해 보조인자(co-factor)를 사용해야 한다. 이러한 보조인자의 한 가지 특정한 예로는 거대분자 크라우딩 제제(macromolecular crowding agent)가 있다. 크라우딩 제제는 많은 생화학 반응의 수행에 필수적이다. 주목할 만한 예로는 핵산 증폭을 위한 리콤비네이즈 폴리머레이즈 증폭(Recombinase polymerase amplification, RPA) 시스템이 있다. 크라우딩 제제의 사용은 RPA 성능 효율을 높이는 데 필수적인 것으로 간주되어 오고 있다. 다만, 크라우딩 제제에는 단점이 있을 수 있다. 따라서, RPA를 포함한 생화학 반응의 성능 효율을 높이고 첨가적(added)/외인적(exogenous) 크라우딩

제제의 필요성을 제거할 대체 수단이 사용되어야 한다. 본 발명은 이러한 문제도 다룬다.

발명의 내용

해결하려는 과제

- [0004] 본 발명은 수성 시험관 내 반응 시스템에서 생화학 반응을 수행하는 공정을 제공하며, 상기 생화학 반응은 적어도 하나의 반응 거대분자, 선택적으로는 적어도 하나의 반응 폴리펩타이드의 기능에 의존적이고, 상기 공정은 다음의 단계를 포함한다: 반응을 수행하기에 적합한 조건 하에서 적어도 하나의 IDR-거대분자를 시험관 내 반응 시스템에 도입하는 단계로서, 상기 적어도 하나의 IDR-거대분자는 적어도 하나의 기능적 IDR을 포함하고, 상기 적어도 하나의 IDR-거대분자의 도입으로 인하여 시험관 내 반응 시스템의 생화학 반응 효율이 증가되며; 바람직하게는 상기 적어도 하나의 IDR-거대분자는 적어도 하나의 IDR-폴리펩타이드인, 단계.
- [0005] 상술한 공정에 있어서, 생화학 반응은 1 이상의 IDR-거대분자, 선택적으로는 1 이상의 IDR-폴리펩타이드의 기능에 의존할 수 있으며, 상기 적어도 하나의 IDR-거대분자 또는 상기 적어도 하나의 IDR-폴리펩타이드는 시험관 내 반응 시스템에 도입됨으로써 생화학 반응에서의 반응 기능을 수행하며 반응의 효율을 증가시킨다.
- [0006] 본원에 기재된 임의의 공정은 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 시스템 내에서 액체-액체 탈혼합을 야기하고 복수의 상-분리된 수성 구획(compartments) 형성을 야기하여, 이로써 시스템 내 생화학 반응의 효율을 증가시키기 위하여 IDR-거대분자 또는 IDR-폴리펩타이드를 시스템 내에 유지하는 단계를 추가적으로 포함할 수 있다.
- [0007] 본원에 기재된 임의의 공정은 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-거대분자 또는 IDR-폴리펩타이드와 공동국소화(co-localise)하도록, 또는 반응 수행에 필수적인 분자들의 IDR-거대분자 또는 IDR-폴리펩타이드와의 공동국소화를 추가적으로 자극 또는 향상시키도록 야기하여, 이로써 시스템 내 생화학 반응의 효율을 증가시키기 위하여 IDR-거대분자 또는 IDR-폴리펩타이드를 시스템 내에 유지하는 단계를 추가적으로 포함할 수 있다.
- [0008] 본원에 기재된 임의의 공정에 있어서, 복수의 상-분리된 수성 구획은 복수의 검출가능한 상-분리 수성 입자(particle)일 수 있다.
- [0009] 본 발명의 다른 양태에서, 본 발명은 수성 시험관 내 반응 시스템에서 생화학 반응을 수행하는 공정을 제공하며, 상기 생화학 반응은 적어도 하나의 반응 거대분자, 선택적으로는 적어도 하나의 반응 폴리펩타이드의 기능에 의존하고, 상기 공정은 다음의 단계를 포함한다: 반응을 수행하기에 적합한 조건 하에 적어도 하나의 기능적 IDR(Intrinsically Disordered Region)을 포함하거나 이로 이루어진 아미노산 서열로 태깅된 적어도 하나의 폴리펩타이드(IDR-폴리펩타이드)를 시험관 내 반응 시스템에 도입하는 단계, 및 IDR-폴리펩타이드에 의해 시스템 내에서 액체-액체 탈혼합을 야기하고 복수의 상-분리된 수성 구획(compartments), 바람직하게는 복수의 검출가능한 상-분리 수성 입자(particle)를 형성을 야기하며, 반응 수행에 필수적인 분자들이 상기 구획 내에서 IDR-폴리펩타이드와 공동국소화하도록 야기하여, 이로써 시스템 내 생화학 반응의 효율을 증가시키기 위하여 IDR-폴리펩타이드를 시스템 내에 유지하는 단계.
- [0010] 선택적으로, 이러한 다른 양태에 따른 공정에 있어서, 생화학 반응은 적어도 하나의 반응 폴리펩타이드의 기능에 의존하고, 상기 반응 폴리펩타이드는 적어도 하나의 IDR-폴리펩타이드이며, 상기 적어도 하나의 IDR-폴리펩타이드는 시험관 내 반응 시스템에 도입됨으로써 생화학 반응에서의 반응 기능을 수행하며 반응의 효율을 증가시킨다.
- [0011] 이러한 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 다가 금속 이온을 IDR-폴리펩타이드에 제공함으로써 액체-액체 탈혼합 및 IDR-폴리펩타이드에 의해 야기되는 복수의 상-분리된 수성 구획의 형성을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있다; 선택적으로, 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공된다. 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.
- [0012] 이러한 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-폴리펩타이드에 ATP를 제공함으로써 액체-액체 탈혼합 및 IDR-폴리펩타이드에 의해 야기되는 복수의 상-분

리된 수성 구획의 형성을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있고, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.

[0013] 이러한 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 다가 금속 이온을 IDR-폴리펩타이드에 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 생화학 반응의 효율을 더욱 증가시키는 것을 포함할 수 있다; 선택적으로, 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공된다. 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.

[0014] 이러한 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-폴리펩타이드에 ATP를 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있고, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.

[0015] 이러한 다른 양태에 따른 임의의 공정에 있어서, IDR-폴리펩타이드를 도입한 경우의 시스템 내 반응 효율은, 적어도 하나의 폴리펩타이드가 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 아미노산 서열로 태깅되지 않은 점을 제외하고는 동일한 반응 조건 하에서 적어도 하나의 폴리펩타이드를 도입했을 경우의 시스템 내 반응 효율과 비교하여 증가될 수 있다.

[0016] 본 발명은 또한 수성 시험관 내 반응 시스템에서 생화학 반응을 수행하는 공정을 제공하며, 상기 생화학 반응은 적어도 하나의 반응 거대분자, 선택적으로는 적어도 하나의 반응 폴리펩타이드의 기능에 의존하고, 상기 공정은 다음의 단계를 포함한다:

[0017] i. 반응을 수행하기에 적합한 조건 하에 적어도 하나의 IDR-거대분자를 포함하는 분자들을 시스템 내에 도입하는 단계로서, 상기 적어도 하나의 IDR-거대분자는 적어도 하나의 기능적 IDR(intrinsically disordered region)을 포함하고, 바람직하게는 상기 적어도 하나의 IDR-거대분자는 적어도 하나의 IDR-폴리펩타이드인 것인 단계;

[0018] ii. IDR-거대분자 또는 IDR-폴리펩타이드를 시스템 내에 유지하여 시스템 내에서 액체-액체 탈혼합을 야기하는 단계로서, 상기 액체-액체 탈혼합은 IDR-거대분자 또는 IDR-폴리펩타이드에 의하여 야기되고, 상기 액체-액체 탈혼합은 복수의 상-분리된 수성 구획을 시스템 내에 형성하는 것인 단계;

[0019] iii. IDR-거대분자 또는 IDR-폴리펩타이드를 시스템 내에 유지하여 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-거대분자 또는 IDR-폴리펩타이드와 공동국소화(co-localise)하는 단계; 및

[0020] iv. 상기 구획 내에서 반응이 진행되도록 하는 단계로서, 적어도 하나의 IDR-거대분자의 존재에 의하여 시스템 내 생화학 반응의 효율이 증가하는 것인 단계.

[0021] 상술한 공정에 있어서, 생화학 반응은 적어도 하나의 IDR-거대분자, 선택적으로는 적어도 하나의 IDR-폴리펩타이드의 기능에 의존할 수 있으며, 상기 적어도 하나의 IDR-거대분자 또는 상기 적어도 하나의 IDR-폴리펩타이드는 시험관 내 반응 시스템에 도입됨으로써 생화학 반응에서의 반응 기능을 수행하며 반응의 효율을 증가시킨다. 복수의 상-분리된 수성 구획은 복수의 검출가능한 상-분리 수성 입자(phase-separated particle)일 수 있다.

[0022] 본 발명의 또 다른 양태에 있어서, 본 발명은 수성 시험관 내 반응 시스템에서 생화학 반응을 수행하는 공정을 제공하며, 상기 생화학 반응은 적어도 하나의 반응 거대분자, 선택적으로는 적어도 하나의 반응 폴리펩타이드의 기능에 의존하고, 상기 공정은 다음의 단계를 포함한다:

[0023] i. 반응을 수행하기에 적합한 조건 하에 적어도 하나의 기능적 IDR(Intrinsically Disordered Region)을 포함하거나 이로 이루어진 아미노산 서열로 태깅된 적어도 하나의 폴리펩타이드(IDR-폴리펩타이드)를 시험관 내 반응 시스템에 도입하는 단계;

[0024] ii. IDR-폴리펩타이드를 시스템 내에 유지하여 시스템 내에서 액체-액체 탈혼합을 야기하고 복수의 상-분리된 수성 구획(compartments), 바람직하게는 복수의 검출가능한 상-분리 수성 입자(particle)를 형성을 야기하는 단

계로서, 상기 액체-액체 탈혼합은 IDR-폴리펩타이드에 의해 야기되는 것인 단계;

- [0025] iii. IDR-폴리펩타이드를 시스템 내에 유지하여 반응 수행에 필수적인 분자들이 상기 구획 내에서 IDR-폴리펩타이드와 공동국소화하도록 야기하는 단계; 및
- [0026] iv. 상기 구획 내에서 반응이 진행되도록 하는 단계로서, 적어도 하나의 IDR-폴리펩타이드의 존재에 의하여 생화학 반응의 효율이 증가하는 것인 단계.
- [0027] 선택적으로, 이러한 또 다른 양태에 따른 공정에 있어서, 생화학 반응은 적어도 하나의 반응 폴리펩타이드의 기능에 의존하고, 상기 반응 폴리펩타이드는 상기 적어도 하나의 IDR-폴리펩타이드이며, 상기 적어도 하나의 IDR-폴리펩타이드는 시험관 내 반응 시스템에 도입됨으로써 생화학 반응에서의 반응 기능을 수행하며 반응의 효율을 증가시킨다.
- [0028] 이러한 또 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 다가 금속 이온을 IDR-폴리펩타이드에 제공함으로써 액체-액체 탈혼합 및 IDR-폴리펩타이드에 의해 야기되는 복수의 상-분리된 수성 구획의 형성을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있다; 선택적으로, 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공된다. 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.
- [0029] 이러한 또 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-폴리펩타이드에 ATP를 제공함으로써 IDR-폴리펩타이드에 의해 야기되는 액체-액체 탈혼합 및 복수의 상-분리된 수성 구획의 형성을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 포함할 수 있으며, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.
- [0030] 이러한 또 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 다가 금속 이온을 IDR-폴리펩타이드에 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 포함할 수 있다; 선택적으로, 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공된다. 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.
- [0031] 이러한 또 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-거대분자 또는 IDR-폴리펩타이드에 ATP를 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 생화학 반응의 효율을 더욱 증가시키는 것을 포함할 수 있고, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.
- [0032] 이러한 또 다른 양태에 따른 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 다가 금속 이온을 IDR-폴리펩타이드에 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-거대분자 또는 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 생화학 반응의 효율을 더욱 증가시키는 것을 포함할 수 있다; 선택적으로, 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공된다. 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.
- [0033] 이러한 또 다른 양태에 따른 임의의 공정에 있어서 시스템 내 반응 효율은, 적어도 하나의 폴리펩타이드가 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 아미노산 서열로 태깅되지 않은 점을 제외하고는 동일한 반응 조건에서 적어도 하나의 폴리펩타이드를 도입한 경우의 시스템 내 반응 효율과 비교하여, IDR-폴리펩타이드에 의하여 증가될 수 있다.

- [0034] 상술한 임의의 공정에 있어서, 상기 공정은 다음 단계를 포함하는 시험관 내 반응 시스템에서 핵산 분자를 합성하기 위한 생화학 반응일 수 있다:
- [0035] (a) 적어도 하나의 핵산 프라이머를 제공하는 단계;
- [0036] (b) 적어도 하나의 표적 가닥을 포함하는 표적 핵산 분자를 제공하고, 상기 적어도 하나의 핵산 프라이머를 상기 표적 가닥과 접촉시켜 이중 가닥 구조를 형성하는 단계;
- [0037] (c) IDR-폴리펩타이드로서의 IDR-거대분자를 제공하는 단계로서, 상기 IDR-폴리펩타이드는 폴리머레이즈 또는 1 이상의 폴리펩타이드 보조인자인 것인 단계; 및
- [0038] (d) 반응이 진행되도록 하는 단계로서, 반응이 진행됨으로써 폴리머레이즈 및 dNTP를 사용하여, 선택적으로는 1 이상의 폴리펩타이드 보조인자의 존재 하에서, 상기 적어도 하나의 핵산 프라이머의 3' 말단을 연장하여 이중 가닥 핵산을 형성하고, 제1 가닥은 표적 가닥의 서열을 포함하고, 제2 가닥은 이에 상보적인 서열을 포함하는 것인 단계.
- [0039] 대안적으로, 상술한 임의의 공정에 있어서, 상기 공정은 시험관 내 반응 시스템에서 단일 가닥 표적 핵산 분자 또는 이중 가닥 표적 핵산 분자를 증폭하기 위한 생화학 반응일 수 있으며, 바람직하게는 상기 표적 핵산 분자는 DNA 분자이다.
- [0040] 상기 공정은 다음 단계를 포함하는 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하기 위한 생화학 반응일 수 있다:
- [0041] (a) 제1 핵산 프라이머 및 제2 핵산 프라이머를 제공하는 단계;
- [0042] (b) 제1 가닥 및 제2 가닥을 포함하는 이중 가닥 표적 핵산 분자를 제공하고, 제1 핵산 프라이머 및 제2 핵산 프라이머를 표적 핵산 분자와 접촉시켜 상기 제1 가닥으로 제1 이중 가닥 구조를 형성하고 상기 제2 가닥으로 제2 이중 가닥 구조를 형성하는 단계;
- [0043] (c) IDR-폴리펩타이드로서의 IDR-거대분자를 제공하는 단계로서, 상기 IDR-폴리펩타이드는 폴리머레이즈 또는 1 이상의 단백질 보조인자인 것인 단계;
- [0044] (d) 반응이 진행되도록 하는 단계로서, 반응이 진행됨으로써 폴리머레이즈 및 dNTP를 사용하여, 선택적으로는 1 이상의 단백질 보조인자의 존재 하에서, 상기 제1 핵산 프라이머 및 제2 핵산 프라이머의 3' 말단을 연장하여 제1 이중 가닥 및 제2 이중 가닥을 형성하는 것인 단계; 및
- [0045] (e) 원하는 증폭 정도에 도달할 때까지 상기 단계 (b) 내지 상기 단계 (d)를 반복하는 단계.
- [0046] 상술한 임의의 이중 가닥 표적 핵산 분자를 증폭하기 위한 공정에 있어서, 상기 공정은 다음 단계를 포함하는 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA(recombinase polymerase amplification) 공정일 수 있다:
- [0047] (a) 리콤비네이즈 제제, 선택적으로는 리콤비네이즈 로딩 단백질(recombinase loading protein), 단일 가닥 안정화제, 폴리머레이즈, 제1 핵산 프라이머 및 제2 핵산 프라이머, 제1 가닥 및 제2 가닥을 포함하는 이중 가닥 표적 핵산, 및 선택적으로 엑소뉴클레이즈 III과 같은 엑소뉴클레이즈를 제공하는 단계;
- [0048] (b) 리콤비네이즈 제제를 제1 핵산 프라이머 및 제2 핵산 프라이머, 선택적으로는 리콤비네이즈 로딩 단백질과 접촉시켜 3' 말단에 단일 가닥 영역(single stranded region)을 포함하는 제1 핵단백질 프라이머(nucleoprotein primer) 및 제2 핵단백질 프라이머를 형성하는 단계;
- [0049] (c) 제1 핵단백질 프라이머 및 제2 핵단백질 프라이머를 표적 핵산 분자와 접촉시켜 상기 제1 가닥으로 제1 이중 가닥 구조를 형성하고 상기 제2 가닥으로 제2 이중 가닥 구조를 형성하는 단계;
- [0050] (d) 반응이 진행되도록 하는 단계로서, 반응이 진행됨으로써 폴리머레이즈 및 dNTP를 사용하여 상기 제1 핵단백질 프라이머 및 제2 핵단백질 프라이머의 3' 말단을 연장하여 제1 이중 가닥 및 제2 이중 가닥을 형성하고, 상기 단일 가닥 안정화제가 전치된 제1 핵산 가닥 및 전치된 제2 핵산 가닥을 안정화시키는 것인 단계; 및
- [0051] (e) 원하는 증폭 정도에 도달할 때까지 상기 단계 (b) 내지 상기 단계 (d)를 반복하는 단계;
- [0052] 상기 리콤비네이즈 제제 및/또는 상기 리콤비네이즈 로딩 단백질, 및/또는 상기 단일 가닥 안정화제, 및/또는 상기 폴리머레이즈는 IDR-폴리펩타이드로서 제공되는 것임.

- [0053] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 상기 리콤비네이즈 제제는 UvsX, T4 UvsX, T6 UvsX, RB18 UvsX, 대장균 파지(*E. coli* phage) wV7 UvsX, 시겔라 파지(*Shigella* phage) CB8 UvsX, 시겔라 파지(*Shigella* phage) Shf12 UvsX, 대장균 파지(*E. coli* phage) AR1 UvsX, 파지 vB_EcoM_G4507 UvsX, 시겔라 파지(*Shigella* phage) SHFML-11 UvsX, 대장균 파지(*Escherichia* phage) vB_EcoM_DalCa UvsX, *E. coli* RecA, 대장균(*E. coli*) RadA, 대장균(*E. coli*) RadB, 대장균(*E. coli*) Rad 51 또는 이들의 기능적 유사체(functional analog), 상동체(homolog) 또는 유도체(derivative), 및 이들의 조합으로 이루어진 군으로부터 선택된 것일 수 있으며, 바람직하게는 상기 리콤비네이즈 제제는 UvsX이고, 더욱 바람직하게는 대장균 파지(*Escherichia* phage) vB_EcoM_DalCa UvsX이다.
- [0054] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 상기 공정은 리콤비네이즈 로딩 단백질을 포함할 수 있고, 상기 리콤비네이즈 로딩 단백질은 UvsY, 대장균(*E. coli*) RecO, 대장균(*E. coli*) RecR 또는 이들의 기능적 유사체(functional analog), 상동체(homolog) 또는 유도체(derivative) 및 이들의 조합으로 이루어진 군으로부터 선택된 것일 수 있으며, 바람직하게는 상기 리콤비네이즈 로딩 단백질은 UvsY이고, 더욱 바람직하게는 대장균 파지(*Escherichia* phage) STO UvsY이다.
- [0055] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 상기 폴리머레이즈는 pol- α , pol- β , pol- δ , pol- ϵ , 또는 이들의 기능적 유사체(functional analog), 상동체(homolog) 또는 유도체(derivative), 및 이들의 조합으로 이루어진 군으로부터 선택된 진행 폴리머레이즈일 수 있다. 상기 폴리머레이즈는 바실러스 스테아로써모필루스 폴리머레이즈 I 대형 단편(*Bacillus stearothermophilus* polymerase I large fragment), 바실러스 썩틸리스 폴리머레이즈 I 대형 단편(*Bacillus subtilis* Pol I large fragment, Bsu polymerase), 리스테리아 모노사이토제네스 DNA 폴리머레이즈 I(*Listeria monocytogenes* DNA polymerase I), 황색 포도상구균 DNA 폴리머레이즈 I(*S. aureus* DNA polymerase I, Sau polymerase), 대장균 DNA 폴리머레이즈 I 클레나우 단편(*E. coli* DNA polymerase I Klenow fragment), 대장균 DNA 폴리머레이즈 I(*E. coli* DNA polymerase I), 대장균 DNA 폴리머레이즈 II(*E. coli* DNA polymerase II), 대장균 DNA 폴리머레이즈 III(*E. coli* DNA polymerase III), 대장균 DNA 폴리머레이즈 IV(*E. coli* DNA polymerase IV), 대장균 DNA 폴리머레이즈 V(*E. coli* DNA polymerase V), 또는 이들의 기능적 유사체(functional analog), 상동체(homolog) 또는 유도체(derivative), 및 이들의 조합으로 이루어진 군으로부터 선택된 원핵 폴리머레이즈일 수 있고, 바람직하게는 상기 폴리머레이즈는 황색 포도상구균 DNA 폴리머레이즈 I(*S. aureus* DNA polymerase I, Sau polymerase) 또는 바실러스 썩틸리스 폴리머레이즈 I 대형 단편(*Bacillus subtilis* Pol I large fragment, Bsu polymerase)이다. 상기 폴리머레이즈는 박테리오파지 T4 gp43 DNA 폴리머레이즈(T4 gp43 DNA polymerase), T7 DNA 폴리머레이즈(DNA polymerase) 및 파이-29 DNA 폴리머레이즈(Phi-29 DNA polymerase), 또는 이들의 기능적 유사체(functional analog), 상동체(homolog) 또는 유도체(derivative), 및 이들의 조합으로 이루어진 군으로부터 선택된 박테리오파지 폴리머레이즈일 수 있다.
- [0056] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 상기 단일 가닥 안정화제는 Gp32, 대장균 SSB 단백질(*E. coli* SSB protein), 파지 T4 Gp32 단백질(phage T4 Gp32 protein), 파지 Rb69 Gp32(phage Rb69 Gp32), 파지 vB_EcoM_NBG1 Gp32(phage vB_EcoM_NBG1 Gp32), 또는 이들의 기능적 유사체(functional analog), 상동체(homolog) 또는 유도체(derivative), 및 이들의 조합으로 이루어진 군으로부터 선택된 것일 수 있고, 바람직하게는 상기 단일 가닥 안정화제는 Gp32 또는 파지 vB_EcoM_NBG1 Gp32(phage vB_EcoM_NBG1 Gp32)이다.
- [0057] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 리콤비네이즈 제제만이 IDR-폴리펩타이드로서 제공될 수 있거나, 리콤비네이즈 로딩 단백질만이 IDR-폴리펩타이드로서 제공될 수 있거나, 단일 가닥 안정화제만이 IDR-폴리펩타이드로서 제공될 수 있거나, 폴리머레이즈만이 IDR-폴리펩타이드로서 제공될 수 있거나, 또는 엑소뉴클레이즈만이 IDR-폴리펩타이드로서 제공될 수 있다.
- [0058] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, IDR-폴리펩타이드의 적어도 하나의 기능적 IDR은 IDR-폴리펩타이드가 유전적으로 조작된 융합 단백질이 되도록 적어도 하나의 IDR을 포함하거나 이로 이루어진 아미노산 서열로서 태깅될 수 있고, 상기 1 이상의 기능적 IDR은 IDR-폴리펩타이드의 C-말단, IDR-폴리펩타이드의 N-말단, 또는 IDR-폴리펩타이드의 C-말단 및 IDR-폴리펩타이드의 N-말단 모두, 또는 폴리펩타이드의 길이에 따른 임의의 아미노산 위치에 위치한다.
- [0059] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 알고리즘 MetaDisorder에 의해 분석될 때 0.5보다 큰 스코어(score)를 기록한 아미노산의 서열의 특징을 가질 수 있다.

- [0060] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 트리펩타이드 서열 RGG의 1 이상의 반복부를 포함하는 아미노산 서열을 포함하거나 이로 이루어질 수 있다. 임의의 이러한 공정에, IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 다이펩타이드 서열 FG의 1 이상의 반복부를 추가적으로 포함하는 아미노산 서열을 포함하거나 이로 이루어질 수 있다. 임의의 이러한 공정에, IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 티로신 혹은 페닐알라닌으로 이루어진 적어도 하나의 방향족 아미노산 잔기를 추가적으로 포함하는 아미노산 서열을 포함하거나 이로 이루어질 수 있다.
- [0061] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 다음의 아미노산 서열을 포함하거나 이로 이루어질 수 있다:
- [0062] i. (YNPQGGYQQ)_n (서열번호 19), 상기 n은 1 내지 10의 양의 정수이고, 선택적으로 n은 1, 2, 또는 3; 또는
- [0063] ii. (YSPTSPS)_n (서열번호 124), 상기 n은 1 내지 10의 양의 정수이고, 선택적으로 n은 1, 2, 또는 3; 또는
- [0064] iii. (FSPTSP)_n (서열번호 125), 상기 n은 1 내지 10의 양의 정수이고, 선택적으로 n은 1, 2, 또는 3; 또는
- [0065] iv. (YSPTSP-A/N/G)_n (서열번호 126), 상기 n은 1 내지 10의 양의 정수이고, 선택적으로 n은 1, 2, 또는 3; 또는
- [0066] v. (YSPGSPA)_n (서열번호 127), 상기 n은 1 내지 10의 양의 정수이고, 선택적으로 n은 1, 2, 또는 3.
- [0067] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 글루타민이 풍부(rich)한 아미노산 서열, 선택적으로 상기 아미노산 서열은 적어도 4, 적어도 5, 적어도 6, 적어도 7, 적어도 8, 적어도 9, 또는 적어도 10의 연속된 글루타민 잔기를 포함하는 아미노산 서열을 포함하거나, 또는 이로 이루어질 수 있다. 임의의 이러한 공정에, IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 QQQ 트리펩타이드 서열의 1 이상의 반복을 포함하는 아미노산 서열을 포함하거나 이로 이루어질 수 있다. 임의의 이러한 공정에, IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 (QQQPQY)_n(서열번호 128)의 아미노산 서열을 포함하거나, 이로 이루어질 수 있으며, 상기 n은 1 내지 10의 양의 정수이고, 선택적으로 상기 n은 1, 2, 또는 3이다.
- [0068] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 적어도 5회 연속되는 서열번호 1의 아미노산 서열을 포함하거나, 이로 이루어질 수 있다.
- [0069] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 적어도 5회 연속되는 서열번호 9의 아미노산 서열을 포함하거나, 이로 이루어질 수 있다.
- [0070] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 다가 금속 이온, 바람직하게는 2가 금속 이온과의 방향족 양이온-파이 상호작용에 관여할 수 있는 1 이상의 방향족 티로신 잔기 및 1 이상의 페닐알라닌 잔기를 포함하는 아미노산 서열을 포함할 수 있다.
- [0071] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드의 1 이상의 기능적 IDR은 다가 금속 이온, 바람직하게는 2가 금속 이온과 구아니딘-금속 상호작용에 관여할 수 있는, 1 이상의 아르기닌 잔기를 포함하는 아미노산 서열을 포함할 수 있다.
- [0072] 상술한 임의의 공정에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열, 또는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이 아미노산 서열, 예컨대 서열번호 1 내지 43 중 어느 하나의 서열과 80% 이상의 상동성을 가진 아미노산 서열을 포함하거나 이로 이루어진 아미노산 서열로 태깅된 거대분자 또는 폴리펩타이드를 포함하거나, 이로 이루어질 수 있다.
- [0073] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 상기 IDR-폴리펩타이드는 단일 가닥 안정화제 Gp32일 수 있고, 서열번호 65 내지 88 중 어느 하나의 아미노산 서열을 가지거나, 또는 이들의 기능적 변이체, 예컨대, 서열번호 65 내지 88 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드일 수 있다.
- [0074] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 상기 IDR-폴리펩타이드는 UvsX인 리컴비네이즈 제체일 수 있고, 서열번호 44 내지 59 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 44 내지 59 중 어느 하나의 아미

노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드일 수 있다.

- [0075] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 상기 IDR-폴리펩타이드는 UvsY인 리코비네이즈 로딩 단백질일 수 있고, 서열번호 60 내지 64 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 60 내지 64 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드일 수 있다.
- [0076] 상술한 임의의 공정에 있어서, 상기 공정은 시험관 내 반응 시스템 내에서 다가 금속 이온을 IDR-거대분자 또는 IDR-폴리펩타이드에 제공함으로써, 시험관 내 반응 시스템 내에서 액체-액체 탈혼합(liquid-liquid demixing)을 자극 또는 향상시키고, 생화학 반응의 효율을 증가시키는 단계로서, 상기 다가 금속 이온이 시스템 내에서 복수의 상-분리된 수성 구획의 형성을 추가적으로 자극 또는 향상시켜 시스템 내에서 생화학 반응의 효율을 더욱 증가시키고, 바람직하게는 상기 다가 금속 이온은 추가적으로 검출가능한 상에 의해 분리된 입자(phase-separated particles)들의 생성을 자극 또는 향상시키고; 선택적으로 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공되는 단계를 추가적으로 포함할 수 있다. 임의의 이러한 공정에 있어서, 상기 다가 금속 이온은 2가 금속 이온, 선택적으로, Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} , 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.
- [0077] 상술한 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-거대분자 또는 IDR-폴리펩타이드에 ATP를 제공함으로써 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 야기되는 액체-액체 탈혼합 및 복수의 상-분리된 수성 구획의 형성을 추가로 자극 또는 향상시키고 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.
- [0078] 상술한 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 다가 금속 이온을 IDR-거대분자 또는 IDR-폴리펩타이드에 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-거대분자 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며; 선택적으로 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공된다. 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.
- [0079] 상술한 임의의 공정에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-거대분자 또는 IDR-폴리펩타이드에 ATP를 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.
- [0080] 상술한 임의의 공정에 있어서, 상기 생화학 반응은 표면(surface)을 포함하는 고상(solid phase) 반응 시스템 내에서 수행될 수 있다. 임의의 이러한 공정에서, 상기 생화학 반응은 상술한 시험관 내 반응 시스템에서 단일 가닥 표적 핵산 분자 또는 이중 가닥 표적 핵산 분자를 증폭시키는 공정일 수 있으며, 상기 적어도 하나의 핵산 프라이머 및/또는 IDR-거대분자 및/또는 1 이상의 폴리펩타이드 보조인자는 상기 표면에 부착된 것일 수 있다.
- [0081] 상술한 임의의 시험관 내 반응 시스템에서 이중 가닥 표적 핵산 분자를 증폭하는 RPA 공정에 있어서, 상기 반응은 표면을 포함하는 고상 반응 시스템 내에서 수행될 수 있으며, 상기 리코비네이즈 제제 및/또는 상기 리코비네이즈 로딩 단백질 및/또는 상기 제1 핵산 프라이머 및/또는 상기 제2 핵산 프라이머는 상기 표면에 부착된 것일 수 있고, 바람직하게는: (i) 상기 제1 핵산 프라이머 또는 제2 핵산 프라이머는 상기 표면에 부착된 것이거나; (ii) 상기 제1 핵산 프라이머 및 제2 핵산 프라이머가 모두 상기 표면에 부착된 것이다.
- [0082] 상술한 임의의 표면을 포함하는 고상 반응 시스템 내에서 수행되는 공정에 있어서, 상기 평면은 평면, 또는 마이크로비드(microbead)일 수 있고, 바람직하게는 상기 표면은 실리콘, 유리, 젤-기반 물질(gel-based material) 및/또는 폴리스티렌과 같은 폴리머 물질을 포함하고, 더욱 바람직하게는 상기 표면은 폴리스티렌과 같은 폴리머 물질을 포함하는 마이크로비드일 수 있다. 임의의 이러한 공정에서, 상기 표면은 서브스트레이트(substrate)에 부착된 것일 수 있고, 바람직하게는 상기 표면은 평면 및/또는 유리를 포함하는 서브스트레이트(substrate)이다. 상기 표면, 예컨대 평면 표면 및/또는 서브스트레이트는 플로우-셀로 제공될

수 있다.

- [0083] 본 발명은 배양된(cultured) 세포 내에서의 생화학 반응의 효율을 증가시키기 위하여, 적어도 하나의 상술한 IDR-거대분자 또는 적어도 하나의 상술한 IDR-폴리펩타이드를 배양된 숙주 세포 내로 도입함으로써, 또는 적어도 하나의 상술한 IDR-폴리펩타이드를 배양된 숙주 세포 내에서 발현함으로써 배양 중(in culture)인 세포 내에서 생화학 반응을 수행하는 공정을 제공한다.
- [0084] 상술한 임의의 시험관 내 생화학 반응을 수행하기 위한 공정은 배양된 세포 내에서의 생화학 반응의 효율을 증가시키기 위하여 적어도 하나의 상술한 IDR-거대분자 또는 적어도 하나의 상술한 IDR-폴리펩타이드를 배양된 숙주 세포 내로 도입함으로써, 또는 적어도 하나의 상술한 IDR-폴리펩타이드를 배양된 숙주 세포 내에서 발현함으로써와 같은 배양 중인 세포 내에서 수행되는 생화학 반응을 포함할 수 있다.
- [0085] 생화학 반응은 배양된 숙주 세포 내에서 핵산 분자의 조작을 유도하거나, 또는 핵산 분자의 뉴클레오타이드 서열의 변경(change) 등의 핵산 분자 구조의 변경(change)과 같은 배양된 숙주 세포 내에서 핵산 분자의 변경(alteration)을 유도하는 임의의 반응일 수 있다. 생화학 반응은 배양된 숙주 세포 내에서 핵산 분자의 합성을 유도하는 임의의 반응일 수 있다. 생화학 반응은 핵산 분자로부터 폴리펩타이드의 발현을 유도하는 임의의 반응일 수 있다. 생화학 반응은 배양된 숙주 세포 내에서 핵산 서열의 편집을 유도하는 임의의 반응일 수 있으며, 예컨대 상기 IDR-폴리펩타이드는 Cas9 폴리펩타이드를 포함하는 Cas 폴리펩타이드와 같은 CRISPR 폴리펩타이드일 수 있다. 생화학 반응은 배양된 숙주 세포 내에서 핵산의 절단을 유도하는 임의의 반응일 수 있다. 생화학 반응은 배양된 숙주 세포 내에서 핵산의 상동 재조합(homologous recombination)을 유도하는 임의의 반응일 수 있다. 생화학 반응은 배양된 숙주 세포 내에서 1 이상의 관심 생물학적 산물(biological products of interest)을 생산하기 위한, 또는 배양된 숙주 세포로부터 분비(secret)되거나 다른 방식으로 배양된 숙주 세포로부터 배양 배지로 방출(release)되는 1 이상의 관심 생물학적 산물을 생산하기 위한 배양된 숙주 세포 내 대사 반응일 수 있다.
- [0086] 상술한 임의의 공정에 있어서, 생화학 반응의 효율이 증가되는 것은, 관련된 적어도 하나의 거대분자 또는 적어도 하나의 폴리펩타이드가 1 이상의 기능적 IDR 폴리펩타이드 서열을 포함하지 않거나 이로 태깅되지 않은 점을 제외하고는 동일한 조건에서 반응을 수행한 경우의 반응 효율과 비교하여, 적어도 하나의 IDR-거대분자 또는 적어도 하나의 IDR-폴리펩타이드를 사용한 경우의 반응 효율이 증가되는 것을 포함할 수 있으며, 선택적으로 상기 반응은 외인적으로(exogenously) 첨가된 크라우딩 제제의 부존재 하에서 수행되는 것이다.
- [0087] 상술한 임의의 RPA 공정에 있어서, RPA 생화학 반응의 효율 또는 성능이 증가 또는 향상되는 것은, 관련된 적어도 하나의 폴리펩타이드가 1 이상의 기능적 IDR 폴리펩타이드 서열을 포함하지 않거나 이로 태깅되지 않은 점을 제외하고는 동일한 조건에서 반응을 수행한 경우의 증폭된 산물의 양과 비교하여, 적어도 하나의 IDR-폴리펩타이드를 사용한 경우의 증폭된 산물의 양이 증가되는 것을 포함할 수 있으며, 선택적으로 상기 반응은 외인적으로 첨가된 크라우딩 제제의 부존재 하에서 수행되는 것이다.
- [0088] 적어도 하나의 IDR-거대분자 또는 IDR-폴리펩타이드를 시험관 내 반응 시스템에 도입하는 것을 포함하는 상술한 임의의 공정에 있어서, 시스템 내 반응 효율은, 적어도 하나의 거대분자 또는 폴리펩타이드가 적어도 하나의 기능적 IDR을 포함하지 않는 것을 제외하고는 동일한 반응 조건에서 적어도 하나의 거대분자 또는 폴리펩타이드를 도입한 경우의 반응 효율과 비교하여, IDR-거대분자 또는 IDR-폴리펩타이드에 의하여 증가된다.
- [0089] 적어도 하나의 기능적 IDR을 포함하거나 이로 이루어진 아미노산 서열로 태깅된 적어도 하나의 폴리펩타이드(IDR-폴리펩타이드)를 시험관 내 반응 시스템에 도입하는 것을 포함하는 상술한 임의의 공정에 있어서, 시스템 내 반응 효율은, 적어도 하나의 폴리펩타이드가 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 아미노산 서열로 태깅되지 않은 점을 제외하고는 동일한 반응 조건에서 적어도 하나의 폴리펩타이드를 도입한 경우의 시스템 내 반응 효율과 비교하여, IDR-폴리펩타이드에 의하여 증가한다.
- [0090] 본 발명은 또한 거대분자 및 태그 아미노산 서열을 포함하는 비-자연적으로 발생하는(non-naturally occurring) IDR-거대분자를 제공하며, 상기 태그 아미노산 서열은 1 이상의 기능적 IDR을 포함하거나 이로 구성되고, 상기 IDR-거대분자는 수성 시험관 내 반응 시스템에서 액체-액체 탈혼합 반응을 야기할 수 있다. 임의의 이러한 IDR-거대분자는 액체-액체 탈혼합 및 시스템 내에서의 복수의 상-분리된 수성 구획, 바람직하게는 검출가능한 상에 의해 분리되는 수성 입자의 형성을 야기할 수 있다. 임의의 이러한 비-자연적으로 발생하는 IDR-거대분자에 의해 야기된 시험관 내 반응 시스템 내에서의 임의의 액체-액체 탈혼합은 이에 의해 생화학 반응의 효율을 증가시킬 수 있다.

- [0091] 상술한 임의의 IDR-거대분자는 거대분자 또는 폴리펩타이드 및 태그 아미노산 서열을 포함하는 비-자연적으로 발생하거나, 인공적이거나, 또는 유전적으로 조작된 IDR-거대분자 또는 IDR-폴리펩타이드일 수 있다. IDR-폴리펩타이드의 경우, 상기 태그 아미노산 서열은 폴리펩타이드의 C-말단, 폴리펩타이드의 N-말단, 또는 폴리펩타이드의 C-말단 및 폴리펩타이드의 N-말단 모두, 또는 폴리펩타이드의 길이에 따른 임의의 아미노산에 위치할 수 있다.
- [0092] 상술한 임의의 IDR-거대분자 또는 IDR-폴리펩타이드에 있어서, 태그 아미노산 서열의 1 이상의 기능적 IDR은 상술한 임의의 공정에서 정의된 기능적 IDR이다.
- [0093] 상술한 임의의 IDR-거대분자 또는 IDR-폴리펩타이드에 있어서, 상기 태그 서열은 다가 금속 양이온, 바람직하게는 2가 금속 양이온, 보다 바람직하게는 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 이온, 보다 더 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 이온, 보다 더 바람직하게는 Mg^{2+} 이온과의 방향족 양이온-파이 상호작용에 관여할 수 있는 아미노산 잔기를 포함한다.
- [0094] 상술한 임의의 IDR-거대분자 또는 IDR-폴리펩타이드에 있어서, 상기 IDR-거대분자 또는 IDR-폴리펩타이드는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열, 또는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이 아미노산 서열, 예컨대 서열번호 1 내지 43 중 어느 하나의 서열과 80% 이상의 상동성을 가진 아미노산 서열을 포함하거나 이로 이루어진 아미노산 서열로 태깅된 거대분자 또는 폴리펩타이드를 포함하거나, 이로 이루어진다.
- [0095] 상술한 임의의 IDR-폴리펩타이드에 있어서, 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 서열이 태깅된 상기 폴리펩타이드는 헬리카이즈(helicase), 자이레이즈(gyrase), 리콤비네이즈, 예컨대 RPA 리콤비네이즈 제제; 뉴클레이즈(nuclease), 예컨대 엑소뉴클레이즈(exonuclease) 및 엔도뉴클레이즈(endonuclease); 라이게이즈(ligase), 글라이코라이에이즈(glycolyase), 메틸레이즈(methylase), 메틸트랜스퍼레이즈(methyltransferase), 글루코실트랜스퍼레이즈(glucosyltransferase), 폴리머레이즈, 카이네이즈(kinase), 포스파타이즈(phosphatase)와 같은 효소; CRISPR 효소와 같은 유전자 편집 효소, 예컨대 Cas9 효소; 보조인자, 예컨대 RPA 리콤비네이즈 로딩 단백질 및 RPA 단일 가닥 안정화제일 수 있다. 상기 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 서열이 태깅된 폴리펩타이드는 라이게이즈, 선택적으로는 RB69 라이게이즈-His2(서열번호: 112)와 같은 RB69 라이게이즈일 수 있다. 상기 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 서열이 태깅된 폴리펩타이드는 RPA 단일 가닥 안정화제, 바람직하게는 Gp32; 선택적으로 상기 IDR-폴리펩타이드는 서열번호 65 내지 88 및 120 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 65 내지 88 및 120 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드일 수 있다. 상기 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 서열이 태깅된 폴리펩타이드는 RPA 리콤비네이즈 제제, 바람직하게는 UvsX; 선택적으로 상기 IDR-폴리펩타이드는 서열번호 44 내지 59 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 44 내지 59 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드일 수 있다. 상기 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 서열이 태깅된 폴리펩타이드는 RPA 리콤비네이즈 로딩 단백질, 바람직하게는 UvsY; 선택적으로 상기 IDR-폴리펩타이드는 서열번호 60 내지 64 중 어느 하나의 아미노산 서열을 가지거나, 또는 상기 IDR-폴리펩타이드는 이들의 기능적 변이체, 예컨대 서열번호 60 내지 64 중 어느 하나의 아미노산 서열과 80% 이상의 상동성을 가지는 아미노산 서열을 가지는 IDR-폴리펩타이드일 수 있다.
- [0096] 본 발명은 또한 상술한 임의의 IDR-폴리펩타이드를 인코딩하는 제1 핵산 서열을 포함하는 단리된(isolated) 핵산 분자; 선택적으로 프로모터를 인코딩하는 제2 핵산 서열을 포함하고, 상기 제1 핵산 서열은 상기 제2 핵산 서열과 작동가능하게 연결(operably linked)된 단리된 핵산 분자를 제공한다. 본 발명은 또한 임의의 이러한 핵산 분자를 포함하는 재조합 폴리뉴클레오타이드 발현 벡터를 제공한다. 본 발명은 또한 임의의 이러한 핵산 분자 또는 임의의 이러한 재조합 폴리뉴클레오타이드 발현 벡터를 포함하는 숙주 세포를 제공한다. 본 발명은 또한 성장 배지(growth medium) 및 상술한 임의의 숙주 세포의 집단(population)을 포함하는 세포 배양물(cell culture)을 제공한다.
- [0097] 본 발명은 또한 상술한 임의의 비-자연적으로 발생하는 IDR-거대분자 또는 IDR-폴리펩타이드를 포함하는 키트를 제공한다. 임의의 이러한 키트는 RPA 리콤비네이즈 제제, 및/또는 RPA 리콤비네이즈 로딩 단백질, 및/또는 폴리머레이즈, 및/또는 제1 및 제2 핵산 프라이머, 및/또는 엑소뉴클레이즈, 및/또는 버퍼, 및/또는 다가 금속 이온

원(source of multivalent metal ions), 바람직하게는 2가 금속 양이온원을 포함하는, 추가적 RPA 구성을 추가적으로 포함할 수 있다. 임의의 이러한 키트에 있어서, 상기 모든 구성은 동결건조 형태로 제공될 수 있다.

[0098] 본 발명은 또한 용액 내 액체-액체 탈혼합을 자극 또는 향상시키는 공정을 제공하며, 상기 공정은 상술한 임의의 IDR-거대분자 또는 IDR-폴리펩타이드를 포함하는 용액을 제공하는 단계, 및 용액 중 IDR-거대분자 또는 IDR-폴리펩타이드를 다가 금속 이온과 접촉시켜 용액 내 액체-액체 탈혼합을 자극 또는 향상시키는 단계를 포함한다. 본 발명은 또한 수성 시험관 내 반응 시스템에서 IDR-거대분자 또는 IDR-폴리펩타이드에 의한 액체-액체 탈혼합을 자극 또는 향상시키는 추가적인 공정을 제공하며, 상기 공정은 상술한 임의의 IDR-거대분자 또는 상술한 임의의 IDR-폴리펩타이드를 시스템 내로 제공하는 단계, 다가 금속 이온을 시스템 내로 제공하는 단계, 및 상기 IDR-거대분자 또는 IDR-폴리펩타이드가 접촉하도록 하여 용액 내 액체-액체 탈혼합을 자극 또는 향상시키는 단계를 포함한다. 임의의 이러한 공정에 있어서, 상기 액체-액체 탈혼합으로 인하여 용액 내에 복수의 상-분리된 수성 구획, 바람직하게는 복수의 검출가능한 상에 의해 분리된 입자가 형성될 수 있다. 임의의 이러한 공정에 있어서, 상기 다가 금속 이온은 2가 금속 이온, 선택적으로는 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} , 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} , 보다 바람직하게는 Mg^{2+} 일 수 있다. 임의의 이러한 추가적인 공정에 있어서, 상기 다가 금속 이온은 1 이상의 기능적 IDR 아미노산 서열에서 아미노산 잔기와의 방향족 양이온-파이 상호작용에 관여함으로써, 액체-액체 탈혼합을 촉진할 수 있다.

[0099] 임의의 이러한 추가적인 공정에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-거대분자 또는 IDR-폴리펩타이드에 ATP를 제공함으로써 액체-액체 탈혼합 및 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 야기되는 복수의 상-분리된 수성 구획의 형성을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.

[0100] 임의의 이러한 추가적인 공정에 있어서, 반응을 수행하기에 적합한 조건은 다가 금속 이온을 IDR-폴리펩타이드에 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-거대분자 또는 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며; 선택적으로 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공된다. 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.

[0101] 임의의 이러한 추가적인 공정에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-거대분자 또는 IDR-폴리펩타이드에 ATP를 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.

[0102] 본 발명은 또한 용액 내 액체-액체 탈혼합을 자극 또는 향상시키는 다가 금속 이온의 용도를 제공하며, 상기 탈혼합은 상술한 임의의 IDR-거대분자 또는 상술한 임의의 IDR-폴리펩타이드에 의해 매개된다. 본 발명은 또한 시험관 내로 도입된 IDR-거대분자 또는 IDR-폴리펩타이드에 의하여 야기되는 액체-액체 탈혼합을 수성 시험관 내 반응 시스템에서 자극 또는 향상시키는 다가 금속 이온의 용도를 제공하며, 상기 IDR-거대분자 또는 IDR-폴리펩타이드는 상술한 임의의 IDR-거대분자 또는 상술한 임의의 IDR-폴리펩타이드이다. 임의의 이러한 용도에 있어서, 상기 액체-액체 탈혼합으로 인하여 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 야기되는 복수의 상-분리된 수성 구획, 바람직하게는 용액 내 복수의 검출가능한 상에 의해 분리된 수성 입자가 형성될 수 있다. 임의의 이러한 용도에 있어서, 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다. 임의의 이러한 용도에 있어서, 다가 금속 이온은 1 이상의 기능적 IDR 아미노산 서열에서 아미노산 잔기와의 방향족 양이온-파이 상호작용에 관여함으로써, 액체-액체 탈혼합을 촉진할 수 있다.

[0103] 임의의 이러한 용도에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-거대분자 또는 IDR-폴리펩타이드에 ATP를 제공함으로써 액체-액체 탈혼합 및 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 야기되는 복수의 상-분리된 수성 구획의 형성을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효

율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.

- [0104] 임의의 이러한 용도에 있어서, 반응을 수행하기에 적합한 조건은 다가 금속 이온을 IDR-폴리펩타이드에 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-거대분자 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며; 선택적으로 상기 다가 금속 이온은 약 22 mM 이상의 농도로 제공되고, 바람직하게는 상기 다가 금속 이온은 약 22 mM 내지 50 mM의 농도로 제공된다. 상기 다가 금속 이온은 2가 금속 이온, 선택적으로 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 일 수 있으며, 바람직하게는 Mg^{2+} , Mn^{2+} 또는 Ca^{2+} 일 수 있고, 더욱 바람직하게는 Mg^{2+} 일 수 있다.
- [0105] 임의의 이러한 용도에 있어서, 반응을 수행하기에 적합한 조건은 시험관 내 반응 시스템에서 IDR-거대분자 또는 IDR-폴리펩타이드에 ATP를 제공함으로써 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-폴리펩타이드와 공동국소화하는 것을 추가로 자극 또는 향상시키고, 이로써 시스템 내 생화학 반응의 효율을 더욱 증가시키는 것을 추가적으로 포함할 수 있으며, 상기 ATP는 1 mM 내지 3.5 mM의 농도로, 선택적으로 1 mM 내지 2 mM의 농도로, 바람직하게는 1 mM의 농도로 시스템에 제공된다.
- [0106] 본 발명은 또한 치료(therapy)에 사용하기 위한 용도, 의약(medicament)으로 사용하기 위한 용도, 약제(pharmaceutical)로 사용하기 위한 용도, 진단 방법에 사용하기 위한 용도, 또는 진단제로 사용하기 위한 용도의 상술한 임의의 비-자연적으로 발생하는 IDR-거대분자 또는 상술한 임의의 IDR-폴리펩타이드를 제공한다.
- [0107] 본 발명은 또한 상술한 임의의 비-자연적으로 발생하는 IDR-거대분자 또는 상술한 임의의 IDR-폴리펩타이드를 제조하는 공정을 제공하며, 상기 공정은 거대분자, 선택적으로 폴리펩타이드를 제공하는 단계, 및 상기 거대분자 또는 폴리펩타이드를 1 이상의 기능적 IDR 폴리펩타이드 서열로 태깅하는 단계를 포함한다. 상기 태깅은 본원에 기재되고 정의된 임의의 방법으로 수행될 수 있다. 상기 1 이상의 기능적 IDR 폴리펩타이드 서열은 본원에 기재되고 정의된 임의의 것과 동일할 수 있다. 상기 거대분자 또는 폴리펩타이드는 본원에 기재되고 정의된 임의의 거대분자 또는 폴리펩타이드를 포함하는 임의의 적합한 거대분자 또는 폴리펩타이드일 수 있다.
- [0108] 상술한 임의의 IDR-거대분자 또는 IDR-폴리펩타이드는 생화학 반응의 효율을 증가시킬 수 있다. 생화학 반응의 효율이 증가되는 것은, 관련된 적어도 하나의 거대분자 또는 적어도 하나의 폴리펩타이드가 1 이상의 기능적 IDR 폴리펩타이드 서열을 포함하지 않거나 이로 태깅되지 않은 점을 제외하고는 동일한 조건에서 반응을 수행한 경우의 반응 효율과 비교하여, 적어도 하나의 IDR-거대분자 또는 적어도 하나의 IDR-폴리펩타이드를 사용한 경우의 반응 효율이 증가되는 것을 포함할 수 있으며, 선택적으로 상기 반응은 외인적으로 첨가된 크라우딩 제제의 부존재 하에서 수행되는 것이다.
- [0109] 상술한 임의의 IDR-거대분자 또는 IDR-폴리펩타이드는 생화학 반응의 효율을 증가시킬 수 있으며, 상기 반응은 리컴비네이즈 폴리머레이즈 증폭(RPA) 반응이다. RPA 생화학 반응의 성능이 증가되는 것은, 관련된 적어도 하나의 폴리펩타이드가 1 이상의 기능적 IDR 폴리펩타이드 서열을 포함하지 않거나 이로 태깅되지 않은 점을 제외하고는 동일한 조건에서 반응을 수행한 경우의 증폭된 산물의 양과 비교하여, 적어도 하나의 IDR-폴리펩타이드를 사용한 경우의 증폭된 산물의 양이 증가되는 것을 포함할 수 있으며, 선택적으로 상기 반응은 외인적으로 첨가된 크라우딩 제제의 부존재 하에서 수행되는 것이다.
- [0110] 본 발명은 또한 1 이상의 표적 폴리뉴클레오타이드 분자의 뉴클레오타이드 서열을 결정하는 방법으로서, 상기 방법은 다음 단계를 포함하는 것인 방법을 제공한다:
- [0111] (i) 상술한 임의의 공정을 수행하여 1 이상의 표적 폴리뉴클레오타이드를 증폭함으로써 1 이상의 표적 폴리뉴클레오타이드 분자의 복수의 카피를 포함하는 집단(population)을 얻는 단계; 및
- [0112] (ii) 표적 폴리뉴클레오타이드 분자의 복수의 카피를 포함하는 집단에 대해 1 이상의 핵산 시퀀싱 반응을 수행하는 단계, 바람직하게는 상기 방법은 표면을 포함하는 고상 반응 시스템에서 수행됨.
- [0113] 본 발명은 또한 1 이상의 폴리뉴클레오타이드 분자의 뉴클레오타이드 서열을 결정하는 방법에 있어서의 상술한 임의의 IDR-거대분자 또는 상술한 임의의 IDR-폴리펩타이드의 용도를 제공하며, 바람직하게는 상기 방법은 상술한 방법과 같다.
- [0114] 본 발명은 또한 서열번호 1 내지 43 중 어느 하나의 아미노산 서열, 또는 서열번호 1 내지 43 중 어느 하나의 아

미노산 서열의 기능적 변이 아미노산 서열, 예컨대 서열번호 1 내지 43 중 어느 하나의 서열과 80% 이상의 상동성을 가진 아미노산 서열을 포함하거나 이로 이루어진 폴리펩타이드 또는 단리된 폴리펩타이드를 제공한다. 임의의 이러한 폴리펩타이드는 거대분자 또는 폴리펩타이드에 부착/태깅되어 본원에 추가로 기재된 바와 같이 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드를 형성할 수 있다. 태깅된 상기 거대분자 또는 폴리펩타이드는 수성 반응 시스템 내 생화학 반응의 수행에 필요한 거대분자 또는 폴리펩타이드일 수 있다. 임의의 이러한 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는, 생화학 반응을 수행하기 위한 조건 하에 수성 반응 시스템 내에 유지되는 경우, 서열번호 1 내지 43 중 어느 하나의 아미노산 서열 또는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이 아미노산 서열에 의하여 야기되는 액체-액체 탈혼합 및 복수의 상-분리된 수성 구획, 바람직하게는 검출가능한 상에 의해 분리된 입자의 형성을 시스템 내에서 야기할 수 있으며, 이로써 시스템 내 생화학 반응의 효율을 증가시킬 수 있다. 임의의 이러한 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는, 생화학 반응을 수행하기 위한 조건 하에 수성 반응 시스템 내에 유지되는 경우, 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-거대분자 또는 IDR-폴리펩타이드와 공동국소화 하도록 야기하며, 이로써 시스템 내 생화학 반응의 효율을 증가시킨다.

[0115] 상기 수성 반응 시스템은 수성 시험관 내 반응 시스템일 수 있다.

[0116] 본 발명은 또한 서열번호 1 내지 43 중 어느 하나의 아미노산 서열, 또는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이 아미노산 서열, 예컨대 서열번호 1 내지 43 중 어느 하나의 서열과 80% 이상의 상동성을 가진 아미노산 서열을 포함하거나 이로 이루어진 폴리펩타이드를 인코딩하는 핵산 서열을 포함하는 단리된 핵산 분자를 제공한다.

[0117] 본 발명은 또한 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드를 생산함에 있어서의 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 폴리펩타이드인 IDR 모이어티의 용도를 제공하며, 상기 IDR 모이어티는 거대분자 또는 폴리펩타이드에 부착/태깅되어 있고; 상기 태깅된 거대분자 또는 폴리펩타이드는 수성 반응 시스템 내 생화학 반응에 필요한 거대분자 또는 폴리펩타이드며, 상기 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는, 생화학 반응을 수행하기 위한 조건 하에 수성 반응 시스템 내에 유지되는 경우, IDR 모이어티에 의하여 야기되는 액체-액체 탈혼합 및 복수의 상-분리된 수성 구획, 바람직하게는 검출가능한 복수의 상에 의해 분리된 입자의 형성을 시스템 내에서 야기하며, 이로써 시스템 내 생화학 반응의 효율을 증가시킨다.

[0118] 바람직하게는, 상기 IDR 모이어티는 폴리펩타이드에 부착/태깅됨으로써 IDR-태깅된 폴리펩타이드를, 바람직하게는 제조함 유전적 융합 단백질로서 생산한다.

[0119] 바람직하게는, 상기 IDR 모이어티는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열, 또는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이 아미노산 서열, 예컨대 서열번호 1 내지 43 중 어느 하나의 서열과 80% 이상의 상동성을 가진 아미노산 서열을 포함하거나 이로 이루어진 폴리펩타이드이다.

[0120] 상술한 임의의 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는 비-자연적으로 발생되거나, 인공적이거나, 또는 유전적으로 조작된 거대분자 또는 폴리펩타이드로 정의될 수 있다.

[0121] 상술한 임의의 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는 본원에서 기재되고 정의된 임의의 1 이상의 IDR-거대분자 또는 IDR-폴리펩타이드의 특유의 특징을 추가적으로 가질 수 있다.

[0122] 상기 수성 반응 시스템은 수성 시험관 내 반응 시스템일 수 있다.

[0123] 본 발명은 상술한 임의의 용도에 따라 얻어진 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드를 추가적으로 제공한다.

[0124] 본 발명은 또한 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드를 생산하는 방법을 제공하며, 상기 방법은 거대분자 또는 폴리펩타이드를 제공하는 단계 및 이에 IDR 모이어티를 부착/태깅하는 단계를 포함하고, 상기 IDR 모이어티는 1 이상의 기능적 IDR을 포함하거나 이로 이루어진 폴리펩타이드이며; 태깅된 상기 거대분자 또는 폴리펩타이드는 수성 반응 시스템 내 생화학 반응의 수행에 필요한 거대분자 또는 폴리펩타이드이고, 상기 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는, 생화학 반응을 수행하기 위한 조건 하에 수성 반응 시스템 내에 유지되는 경우, IDR 모이어티에 의하여 야기되는 액체-액체 탈혼합 및 복수의 상-분리된 수성 구획, 바람직하게는 검출가능한 상에 의해 분리된 입자의 형성을 시스템 내에서 야기하며, 이로써 시스템 내 생화학 반응의 효율을 증가시킨다. 임의의 이러한 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는, 생화학 반응을 수행하기 위한 조건 하에 수성 반응 시스템 내에 유지되는 경우, 반응 수행에 필수적인 분자들이 복수의 상-분리된 수성 구획 내에서 IDR-거대분자 또는 IDR-폴리펩타이드와 공동국소화 하도록 야기하며, 이로써 시스템 내

생화학 반응의 효율을 증가시킨다.

- [0125] 바람직하게는, 상기 방법은 폴리펩타이드를 제공하는 단계 및 이에 IDR 모이어티를 부착/태깅하여 IDR-태깅된 폴리펩타이드를, 바람직하게는 재조합 유전적 융합 단백질로서 생산하는 단계를 포함한다.
- [0126] 바람직하게는, 상기 IDR 모이어티는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열, 또는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이 아미노산 서열, 예컨대 서열번호 1 내지 43 중 어느 하나의 서열과 80% 이상의 상동성을 가진 아미노산 서열을 포함하거나 이로 이루어진 폴리펩타이드이다.
- [0127] 상술한 임의의 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는 비-자연적으로 발생되거나, 인공적이거나, 또는 유전적으로 조작된 거대분자 또는 폴리펩타이드로 정의될 수 있다.
- [0128] 상술한 임의의 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드는 본원에서 기재되고 정의된 임의의 1 이상의 IDR-거대분자 또는 IDR-폴리펩타이드의 특유의 특징을 추가적으로 가질 수 있다.
- [0129] 상기 수성 반응 시스템은 수성 시험관 내 반응 시스템일 수 있다.
- [0130] 본 발명은 상술한 임의의 방법에 따라 얻어진 IDR-태깅된 거대분자 또는 IDR-태깅된 폴리펩타이드를 추가적으로 제공한다.

과제의 해결 수단

- [0131] 리컴비네이즈 폴리머레이즈 증폭(recombinase polymerase amplification, RPA)은 핵산 분자를 증폭하는 기술이다. 시스템은 그 중에서도, 리컴비네이즈 효소 및 바람직하게는 리컴비네이즈 로딩 단백질을 이용한다. 이러한 단백질 성분은 증폭 프라이머와 복합체를 형성한다. 증폭될 표적 핵산 분자에 결합한 후, 복합체는 표적 핵산 분자를 "스캔"하고 표적 및 프라이머 서열 사이의 상보성 영역을 "검색"한다. 상보적 영역이 발견되면 복합체는 프라이머의 표적 서열 결합을 촉진한다. 그러면, 폴리머레이즈 효소는 프라이머를 확장하여 표적 서열의 사본을 생성할 수 있다. 리컴비네이즈 복합체의 사용은 PCR과 같은 다른 핵산 증폭 방법에 비해 중요한 차이점을 제공한다. RPA에서는 리컴비네이즈 복합체가 프라이머 결합 문제에서 완전히 효소-기반의 솔루션을 제공하기 때문에, 열 순환에 의해 구동되는 용융(melting) 및 어닐링(annealing) 단계에 대한 요구 사항이 없다. 따라서 RPA는 등온 기술이다. 극도의 열적 주기성(extreme thermal cycling)에 대한 요구 사항이 없다는 것은 RPA가 PCR과 같은 기술에 비해 많은 명백한 이점이 있음을 의미한다.
- [0132] RPA 기법에서 잘 문서화된 요구 사항은 기술 분야에서 일반적으로 "거대분자 크라우딩 제제(macromolecular crowding agent)"라고도 하는 "크라우딩 제제(crowding agent)"의 존재이다. 이러한 제제는 기술 분야에서 잘 알려져 있으며 당업계에서 이해되는 의미를 가지고 있다. 크라우딩 제제는 여기에서 더 자세히 논의된다. RPA 기법에서 가장 일반적으로 사용되는 크라우딩 제제 중 하나는 폴리에틸렌 글리콜(polyethylene glycol, PEG)이지만, 다른 크라우딩 제제를 사용할 수 있다. 본 발명 이전에는 크라우딩 제제의 사용이 RPA 방법의 필수 요건으로 간주되었다.
- [0133] 본 발명자들은 놀랍게도 RPA 기법에서 크라우딩 제제에 대한 지금까지 인식된 중요한 요구 사항을 우회하는 것이 가능하다는 것을 발견했다. 본 발명은 이러한 발견에 기초한 것이다.
- [0134] 본 발명자들은 놀랍게도 RPA 기법에 필요한 단백질 성분과 같은 거대분자를 하나 이상의 기능적인 "IDR (intrinsically disordered regions)"을 포함하거나 구성하는 아미노산 서열로 "태깅"함으로써, IDR 아미노산 서열 태그가 크라우딩 제제가 전혀 없는 상태에서 효율적인 RPA를 촉진할 수 있음을 발견하였다. 이와 같이 크라우딩 제제에 의존하지 않고 RPA 시스템에서 효율적인 증폭을 달성할 수 있으므로 RPA 반응의 복잡성을 줄일 수 있다.
- [0135] 본 발명자들은 또한 놀랍게도, 크라우딩 제제 없이 IDR-태깅된 거대분자 성분을 포함하는 RPA 기법의 증폭 효율은 생화학 반응 시스템에서 액체-액체 탈혼합을 촉진해 상 분리를 유도하는 IDR 아미노산 태그 서열의 기능적 능력과 상관될 수 있음을 발견하였다. 상 분리는 생화학 반응 환경에서 상 분리된 수성 구획, 특히 본 명세서에 추가 설명된 바와 같이 현미경 관찰을 비롯한 표준 방법으로 검출할 수 있는 구형-유사 수성 구획의 입자(spherical-like aqueous globular foci) 또는 상-분리된 수성 입자의 형성에 의해 평가될 수 있다.
- [0136] 더욱이, 본 발명자들은 또한 놀랍게도, IDR-태깅된 거대분자 성분 및 크라우딩 제제의 제공이 RPA 기법의 증폭 효율성과 관련하여 부가적 및 심지어 상승적 효과를 제공할 수 있다는 것을 발견하였다.
- [0137] 또한, 본 발명자들은 놀랍게도, 크라우딩 제제 없이 IDR 아미노산-태깅된 거대분자 성분을 포함하는 RPA 기법에

서 증폭 효율이 생화학적 반응 환경에 도입된 다가 금속 양이온의 농도와 상관될 수 있음을 발견하였다. 따라서 다가 금속 양이온은 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 유발된 액체-액체 탈혼합을 더욱 자극하거나 향상시켜 반응 효율을 더욱 증가시킬 수 있다.

[0138] 본 발명자들은 또한 놀랍게도, 본 명세서에 추가로 기재된 바와 같이 특정 농도의 ATP가 IDR-거대분자 또는 IDR-폴리펩타이드에 의한 액체-액체 탈혼합을 추가로 자극하거나 향상시킬 수 있고, 그에 따라 반응 효율을 추가로 증가시킬 수 있다는 것을 발견하였다.

[0139] 이러한 놀라운 발견에 기초하여, 본 발명은 본 명세서에 추가로 기재된 바와 같이 RPA 반응을 포함하는 효소-기반 시험관 내 생화학 반응의 효율을 증가시키는 방법 및 체제를 제공한다.

[0140] 본 명세서에 서술되고 정의된 IDR 아미노산 서열 및 IDR 체제는 폴리펩타이드와 같은 생화학 반응의 임의의 적합한 거대분자 성분에 적용되는 유용한 체제로서 광범위한 적용 가능성을 가지며, 이에 따라 특히 IDR 아미노산 서열이 다가 금속 양이온과 함께 사용되는 생화학적 반응 환경에서 거대분자 크라우딩 체제에 의존하지 않는 상 분리를 포함한 액체-액체 탈혼합을 촉진한다. 본 발명은 또한 거대분자 크라우딩제에 의존하지 않고 생화학적 반응 환경에서 IDR 아미노산 서열-매개 상 분리를 촉진하는데 있어서 2가 금속 양이온 또는 그의 임의의 기능적 등가물과 같은 다가 금속 양이온의 사용을 포함한다.

[0141] 따라서, 본 발명은 본 명세서에 추가로 서술되고 정의된 바와 같이 IDR-기반 공정, 거대분자, 폴리펩타이드, 핵산, 벡터, 숙주 세포 및 용도를 제공한다.

[0142] 본 발명의 요소들은 아래에서 차례로 설명된다.

[0143] 생화학 반응

[0144] 위에서 설명된 바와 같이, 본 발명자들은 놀랍게도 이전에 고려된 RPA 및 기타 반응의 필수 구성요소인 크라우딩 체제에 대한 요구 사항을 우회하는 것이 가능하다는 것을 발견했다. 본 명세서에 상세히 기재된 바와 같이, 이것은 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하는 아미노산 서열을 RPA 반응에 필요한 단백질 성분에 부착(attach)/테더링(tethering)/태깅(tagging)함으로써 달성될 수 있다. 본 발명자들은 또한 놀랍게도 라이게이즈 효소에 부착된 기능적 IDR(intrinsically disordered region)이 라이게이즈 반응의 효율을 증가시킬 수 있음을 보여주었다. 본 발명자들은 IDR 아미노산 서열에 의해 유도된 상 분리의 정도가, 예를 들어 크라우딩 체제가 없는 조건에서의 증폭처럼 반응의 효율과 상관될 수 있으며, 다가 금속 양이온으로 향상될 수 있음을 보여주었다. 이러한 놀라운 관찰에 기초하여 생화학 반응의 거대분자 또는 단백질 성분과 관련된 그러한 IDR 아미노산 서열이 시험관 내 또는 생체 내의, 특히 추가적/외인성 크라우딩 체제가 없는, 생화학 반응 환경에서의 반응 효율을 향상시킬 것으로 쉽게 예상된다.

[0145] 따라서, 본 발명은 시험관내 또는 생체내 생화학적 반응의 임의의 적합한 거대분자 또는 폴리펩타이드 성분에 적용되는 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열의 사용을 포함하고, 따라서 생화학 반응 환경에서 액체-액체 탈혼합을 촉진하고 생화학 반응의 효율을 증가시킬 수 있는 IDR 체제를 제공한다. 생화학 반응 환경에서 이러한 액체-액체 탈혼합은 생화학 반응 환경에서 상 분리를 초래할 수 있다. 생화학 반응 환경에서 이러한 액체-액체 탈혼합은 상 분리를 야기할 수 있으며, 이는 본 명세서에 추가로 기술되는 바와 같이 생화학 반응 환경에서 검출 가능한 상-분리된 수성 입자를 포함하는 상-분리된 수성 구획의 형성을 야기하거나 이를 촉진할 수 있다. 본 명세서에 추가로 기재되고 정의되는 이러한 IDR 체제, 또는 IDR-기반 체제는 IDR-거대분자 또는 IDR-태그된 거대분자, 또는 IDR-폴리펩타이드 또는 IDR-태그된 폴리펩타이드 중 임의의 하나 이상을 설명하기 위해 상호교환가능하게 지칭될 수 있다.

[0146] 방법, 공정 및 용도 중 어느 하나에 있어서, 또는 본 명세서에 기술되고 정의된 IDR-거대분자, IDR-융합 거대분자 또는 이를 인코딩하는 단리된 핵산 분자, 재조합 폴리뉴클레오타이드 발현 벡터 또는 숙주 세포 중 어느 하나에 있어서, 생화학적 반응의 효율 또는 성능을 증가시키거나 향상시키는 것은 동일한 조건 하에 반응을 수행함으로써 수득된 효율과 비교하여 본 명세서에 기재된 IDR-기반 거대분자 또는 폴리펩타이드 중 임의의 하나 이상을 사용하여 반응 효율을 증가시키는 것을 포함할 수 있지만, 여기서 관련 거대분자 또는 폴리펩타이드는 하나 이상의 기능적인 IDR 폴리펩타이드 서열을 포함하지 않거나 태그되지 않았으며, 선택적으로 반응은 외인성 첨가된 크라우딩 체제 없이 수행된다.

[0147] 생화학적 반응의 효율성 또는 성능을 증가시키거나 향상시키는 것은 일반적으로 허용되는 개념에 따라 이해되어야 한다. 예를 들어, RPA 반응 또는 기타 핵산 증폭 반응에서의 반응 효율은 비교적 적은 시작 표적 핵산(starting target nucleic acid)을 사용하여 동등한 총 앰플리콘(amplicon) 집단을 제공하거나, 동일한 양의

시작 표적 핵산을 사용하여 비교적 빠른 검출 시간 또는 비교적 빠른 증폭 속도를 제공하는 것으로 이해될 수 있다.

- [0148] RPA 생화학적 반응의 효율 또는 성능을 증가 또는 향상시키는 것은, 동일한 조건이지만 관련된 거대분자 또는 폴리펩타이드가 하나 이상의 기능적인 IDR(intrinsically disordered region) 폴리펩타이드 서열을 포함하지 않거나 그와 태그되지 않은 것인 반응의 수행에서 얻은 증폭된 산물의 양과 비교해, 하나 이상의 본 명세서에 기술된 IDR-기반 거대분자 또는 폴리펩타이드를 사용한 RPA 반응에서 얻은 증폭된 산물의 총량을 증가시키는 것을 포함할 수 있으며, 선택적으로 반응은 외인성 첨가된 크라우딩 제제의 부재 하에 수행된다.
- [0149] 시험관내 반응 시스템과 같은 반응 시스템에서 생화학적 반응의 효율을 증가시키는 것은, 일정 기간 동안의 반응 속도, 일정 기간 동안 소모된 기질의 양, 일정 기간 동안 생산된 산물의 양 등과 같은 특정 기간 동안 시스템에서 반응의 측정 가능한 파라미터를 증가시키는 것을 포함할 수 있다.
- [0150] 시험관내 반응 시스템과 같은 반응 시스템에서 생화학적 반응의 효율을 증가시키는 것은, 검출가능한 상-분리된 수성 입자와 같은 검출가능한 상-분리 구획 내에서 반응의 파라미터를 증가시키는 것을 포함할 수 있다. 이것은 예를 들어, 반응의 파라미터를 측정하고, 검출 가능한 상-분리 수성 입자 및/또는 반응 분자의 검출가능한 상-분리 수성 입자로의 검출가능한 공동국소화(co-localisation)형성의 증가를 상관시킴으로써 간접적으로 추론할 수 있다.
- [0151] 본 명세서에 설명된 것은 임의의 IDR 아미노산 서열이 주어진 거대분자 또는 단백질과 함께 사용될 때 원하는 생화학 반응 환경에서 액체-액체 탈혼합 및 상 분리를 촉진하는 데 필요한 방식으로 기능할 수 있는지 여부를 확립하는 데 사용할 수 있고, 원하는 시험관내 생화학 반응환경에 포함될 수 있는 간단한 생물정보학적 방법 및 상 분리 분석이다. 더욱이 주어진 보조인자, 특히 다가의 금속 양이온(예를 들면, 2가 금속 양이온)은 이러한 분석에서 매우 간단한 방식으로 확립될 수 있다.
- [0152] 따라서, 본 발명은 임의의 주어진 원하는 시험관내 또는 생체내 생화학적 반응 환경에 유용하게 적용될 수 있는 본 명세서에 기술되고 정의된 IDR 체제를 제공한다.
- [0153] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 본 명세서에 기술된 임의의 반응과 같은 시험관내 또는 생체내 생화학적 반응의 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다.
- [0154] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 핵산 합성 반응의 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다.
- [0155] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 핵산 합성 반응의 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있으며, 여기서 폴리머레이즈는 프라이머 핵산 분자를 연장함으로써 새로운 핵산 분자를 합성하는 데 사용된다.
- [0156] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 핵산 증폭 반응의 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다. 핵산 증폭 반응은 열 순환을 수반하는 반응일 수 있다. 핵산 증폭 반응은 등온 증폭 반응일 수 있다. 핵산 증폭 반응은 폴리머레이즈 연쇄반응(polymerase chain reaction, PCR), 폴리머레이즈 나선반응(polymerase spiral reaction, PSR), 루프 매개 등온 증폭(loop mediated isothermal amplification, LAMP), 핵산 서열 기반 증폭(nucleic acid sequence based amplification, NASBA), 자가-지속 서열 복제(self-sustained sequence replication, 3SR), 롤링 서클 증폭(rolling circle amplification, RCA), 가닥 치환 증폭(strand displacement amplification, SDA), 다중 치환 증폭(multiple displacement amplification, MDA), 라이게이즈 연쇄 반응(ligase chain reaction, LCR), 헬리케이즈 의존성 증폭(helicase dependant amplification, HDA), 분지 증폭법(ramification amplification method, RAM), 리콤비네이즈 폴리머레이즈 증폭(recombinase polymerase amplification, RPA), 전사 매개 증폭(transcription-mediated amplification, TMA) 또는 NEAR(nicking enzyme amplification reaction)일 수 있다.
- [0157] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 유전자 편집 반응의 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다.
- [0158] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 CRISPR 반응의 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다.
- [0159] 본 명세서에 기술되고 정의된 바와 같은 임의의 IDR 아미노산 서열은 프라임 편집 유전자 편집 반응(prime editing gene editing reaction)의 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있으며,

여기서 CRISPR 효소, 예를 들어 Cas 효소, 예를 들어, Cas9은 적어도 역전사효소 효소, 선택적으로 추가로 프라임 편집 가이드 RNA(prime editing guide RNA, pegRNA)와 함께 복합체로 제공되며, 여기서 프라임 편집 복합체의 임의의 구성요소는 하나 이상의 기능적인 IDR(intrinsically disordered region) 폴리펩타이드 서열과 태그되어 제공될 수 있으며, 예를 들어 여기서 CRISPR 효소는 하나 이상의 기능적 IDR 폴리펩타이드 서열로 태그되거나 역전사효소 효소는 하나 이상의 기능적 IDR 폴리펩타이드 서열로 태그된다.

- [0160] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 라이게이션 반응 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다.
- [0161] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 엑소뉴클레이즈 반응 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다.
- [0162] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 엔도뉴클레이즈 반응, 전사 반응, DNA 메틸화 반응, DNA 글리코실화 반응, 항체-항원 반응, 약물-표적 반응 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다.
- [0163] 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열은 단백질:단백질 상호작용을 수반하는 반응의 수행에 필요한 임의의 거대분자 또는 단백질 성분과 함께 사용될 수 있다.
- [0164] 본 명세서에 사용된 시험관내 생화학 반응을 수행하는 공정은 본 명세서에 추가로 기재된 RPA 반응과 같이 반응 용기 내의 용액에서 직접 수행되는 생화학 반응을 포함하도록 의도된다.
- [0165] 본 명세서에 사용된 시험관내 생화학적 반응을 수행하는 공정은 배양된 숙주 세포 내에서 생화학적 반응의 효율을 증가시키기 위해 배양된 숙주 세포에서 본 명세서에 정의된 IDR 체제를 발현시키는 것과 같이 배양물 내에서 수행되는 생화학적 반응을 또한 포함한다.
- [0166] 본 명세서에 사용된 시험관내 생화학적 반응을 수행하는 공정은 배양된 숙주 세포 내에서 생화학적 반응의 효율을 증가시키기 위해 본 명세서에 정의된 IDR 체제를 배양된 숙주 세포에 도입하거나 배양된 숙주 세포에서 본 명세서에 정의된 바와 같은 IDR 체제를 발현시키는 배양된 숙주 세포 내에서 수행되는 생화학적 반응을 포함하며, 상기 생화학적 반응은 배양된 숙주 세포 내에서 핵산 분자의 조작을 유도하거나, 또는 핵산 분자의 뉴클레오타이드 서열의 변화, 핵산 분자의 구조 변화와 같은 배양된 숙주 세포 내에서 핵산 분자의 변형을 유도하는 임의의 반응이다.
- [0167] 본 명세서에 사용된 시험관내 생화학적 반응을 수행하는 공정은 배양된 숙주 세포 내에서 생화학적 반응의 효율을 증가시키기 위해 본 명세서에 정의된 IDR 체제를 배양된 숙주 세포에 도입하거나 배양된 숙주 세포에서 본 명세서에 정의된 IDR 체제를 발현시키는 배양된 세포 내에서 수행되는 생화학적 반응을 포함하며, 상기 생화학적 반응은 배양된 숙주 세포에서 핵산 분자의 합성을 유도하는 임의의 반응이다.
- [0168] 본 명세서에 사용된 시험관내 생화학적 반응을 수행하는 방법은 배양된 숙주 세포 내에서 생화학적 반응의 효율을 증가시키기 위해 본 명세서에 정의된 IDR 체제를 배양된 숙주 세포에 도입하거나 배양된 숙주 세포에서 본 명세서에 정의된 IDR 체제를 발현시키는 배양된 세포 내에서 수행되는 생화학적 반응을 포함하며, 상기 생화학적 반응은 핵산 분자로부터 폴리펩타이드의 발현을 유도하는 임의의 반응이다.
- [0169] 본 명세서에 사용된 시험관내 생화학적 반응을 수행하는 방법은 배양된 숙주 세포 내의 생화학적 반응의 효율을 증가시키기 위해 본 명세서에 정의된 IDR 체제를 배양된 숙주 세포에 도입하거나 배양된 숙주 세포에서 본 명세서에 정의된 IDR 체제를 발현시키는 배양된 세포 내에서 수행되는 생화학적 반응을 포함하며, 상기 생화학적 반응은 배양된 숙주 세포 내의 핵산 서열의 편집 (예를 들어, IDR-폴리펩타이드는 Cas9 폴리펩타이드를 포함하는 Cas 폴리펩타이드와 같은 CRISPR 폴리펩타이드 인 것, 또는 상기 IDR-폴리펩타이드는 CRISPR 폴리펩타이드와 복합체를 이루는 폴리펩타이드인 것, 상기 IDR-폴리펩타이드는 역전사효소와 같은 것), 배양된 숙주 세포 내 핵산의 절단 및 배양된 숙주 세포내에서 상동 재조합을 유도하는 임의의 반응이다.
- [0170] 본 명세서에 사용된 시험관내 생화학적 반응을 수행하는 방법은 배양된 숙주 세포 내의 생화학적 반응의 효율을 증가시키기 위해 본 명세서에 정의된 IDR 체제를 배양된 숙주 세포에 도입하거나 배양된 숙주 세포에서 본 명세서에 정의된 IDR 체제를 발현시키는 배양된 세포 내에서 수행되는 생화학적 반응을 포함하며, 상기 생화학적 반응은 배양된 숙주 세포 내에서 하나 이상의 관심있는 생물학적 산물을 생성하거나 또는 배양된 숙주 세포로부터 분리되거나 배양된 숙주 세포로부터 배양액으로 방출되는 하나 이상의 관심있는 생물학적 산물을 생산하는 대사 반응이다.

- [0171] 본 발명은 또한 생체의(ex vivo)에서 수행되는 생화학적 반응을 포함하는 것으로 의도되는데, 예를 들어, 조직 배양 또는 신체 외부에서 개발된 기타 적절한 복합 생물학적 시스템과 같이 본 명세서에 정의된 IDR 제제를 세포에서 발현시키는 것이다. 따라서, 본 명세서에 정의된 IDR 제제 중 임의의 것을 사용하여 본 명세서에 사용된 수성 시험관내 반응 시스템에서 생화학적 반응을 수행하기 위한 방법에 대한 임의의 언급은 대안적으로 본 명세서에 정의된 임의의 IDR 제제를 사용하는 수성 생체의 반응 시스템에서 생화학적 반응을 수행하기 위한 공정으로 정의될 수 있다.
- [0172] 본 발명은 또한 생체내 생화학적 반응의 효율을 증가시키기 위한 공정, 제제 및 방법을 제공한다. 따라서, 본 명세서에 정의된 임의의 IDR 제제를 사용하는 본 명세서에 사용된 수성 시험관내 반응 시스템에서 생화학적 반응을 수행하기 위한 방법에 대한 임의의 언급은 대안적으로 본 명세서에 정의된 임의의 IDR 제제를 사용하는 수성 생체내 반응 시스템에서 생화학적 반응을 수행하기 위한 공정으로 정의될 수 있다.
- [0173] 본 발명은 치료에 사용하기 위한, 치료제로 사용하기 위한, 의약으로 사용하기 위한, 약제로 사용하기 위한 또는 진단제로서 사용하기 위한, 본 명세서에 기술되거나 정의된 임의의 비자연적 발생 IDR-거대분자 또는 IDR-폴리펩타이드를 제공한다.
- [0174] 본 발명은 치료에 의한 인간 또는 동물 신체의 치료 방법에 사용하기 위한 본 명세서에 기술되거나 정의된 임의의 비자연적 발생 IDR-거대분자 또는 IDR-폴리펩타이드를 제공한다.
- [0175] 본 발명은 인간 또는 동물 신체에 대해 실시되는 진단 방법에 사용하기 위한 본 명세서에 기술되거나 정의된 임의의 비자연적 발생 IDR-거대분자 또는 IDR-폴리펩타이드를 제공한다.
- [0176] 본 발명은 치료에 의한 인간 또는 동물 신체의 치료용 약제의 제조에 사용하기 위한 본 명세서에 기술되거나 정의된 임의의 비자연적 발생 IDR-거대분자 또는 IDR-폴리펩타이드를 제공한다.
- [0177] 본 발명은 인간 또는 동물 신체에 대해 실행되는 진단 방법을 위한 진단제의 제조에 사용하기 위한 본 명세서에 기술되거나 정의된 임의의 비-천연 발생 IDR-거대분자 또는 IDR-폴리펩타이드를 제공한다.
- [0178] 본 발명은 치료적 유효량의 본 명세서에 기술되거나 정의된 임의의 비-천연 발생 IDR-거대분자 또는 IDR-폴리펩타이드를 이를 필요로 하는 인간 또는 동물에게 투여하는 것을 포함하는 인간 또는 동물의 치료 방법을 제공한다.
- [0179] 전술한 생화학적 반응의 효율을 증가시키기 위한 공정, 시약 및 방법 중 어느 하나에서, 비자연적 발생 IDR-거대분자 또는 IDR-폴리펩타이드는 액체-액체 탈혼합을 촉진할 수 있다. 상기 액체-액체 탈혼합 또는 상기 검출가능한 용액에서 상-분리된 구획 또는 입자의 형성 수성 입자를 포함하는 용액 내 상-분리된 수성 구획의 형성을 촉진할 수 있다. 상기 액체-액체 탈혼합 또는 상기 검출가능한 상-분리 구획 또는 입자의 형성은 그에 따라 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 야기되는 생화학적 반응의 효율을 증가시킨다.
- [0180] 본 명세서에 사용된 시험관 내 생화학적 반응을 수행하는 공정은 반응 용기 내의 용액 또는 배양액내 숙주세포에 본 명세서에 정의된 IDR 제제를 도입하거나 배양된 숙주세포에 IDR 제제를 도입 또는 발현시켜 용액 또는 배양된 숙주세포에서 액체-액체 탈혼합을 촉진하는 임의의 생화학적 반응을 포함한다. 임의의 그러한 생화학적 반응에서, 용액 또는 배양된 숙주 세포에서 액체-액체 탈혼합은 본 명세서에 기술되고 정의된 용액 또는 배양된 숙주 세포에서의 상 분리를 촉진한다.
- [0181] 이러한 임의의 생화학적 반응은, 본 명세서에 기술되고 정의된 임의의 IDR 아미노산 서열의, 용액 또는 배양된 숙주 세포에서의 액체-액체 탈혼합 촉진 및/또는 용액에서 상 분리 촉진 효능을 평가하기 위해 수행될 수 있다.
- [0182] 이러한 임의의 생화학적 반응은 약물, 폴리펩타이드 또는 기타 분자와 같은 시험 제제의, 용액 또는 배양된 숙주 세포에서 IDR 아미노산 서열에 의해 매개되는 액체-액체 탈혼합 자극 또는 향상, 및/또는 용액 또는 배양된 숙주 세포에서 IDR 아미노산 서열에 의해 매개되는 상 분리 자극 또는 향상 효능을 평가하기 위해 수행될 수 있으며, 바람직하게는 상기 시험 제제는 IDR 아미노산 서열과 상호작용한다.
- [0183] 이러한 생화학적 반응은 약물, 폴리펩타이드 또는 기타 분자와 같은 시험 제제의, 용액 또는 배양된 숙주 세포에서 IDR 아미노산 서열에 의해 매개되는 액체-액체 탈혼합 억제, 및/또는 용액 또는 배양된 숙주 세포에서 IDR 아미노산 서열에 의해 매개되는 상 분리 억제 효능을 평가하기 위해 수행될 수 있으며, 바람직하게는 여기서 시험 제제는 IDR 아미노산 서열과 상호작용한다.
- [0184] 시험관내, 생체내 또는 생체의 생화학적 반응을 수행하기 위한 본 명세서에 기재된 임의의 공정은 인간 복제 공

정을 배제할 수 있다.

- [0185] 시험관내, 생체내 또는 생체의 생화학적 반응을 수행하기 위한 본 명세서에 기재된 임의의 공정은 인간의 생식 계열 유전적 동일성(germ line genetic identity)을 변형시키는 공정을 배제할 수 있다.
- [0186] 시험관내, 생체내 또는 생체의 생화학적 반응을 수행하기 위한 본 명세서에 기재된 임의의 공정은 인간 배아의 사용 또는 전능성 인간 세포의 사용을 포함하는 공정을 배제할 수 있다.
- [0187] 본 명세서에 기재된 임의의 숙주 세포는 인간 배아, 전능성 인간 세포, 또는 인간 생식계열 세포를 배제할 수 있다.
- [0188] 본 발명의 일부 양태에 있어서 생체내 사용을 수용하지만, 본 발명의 일부 양태에 있어서 생체내 사용의 배제를 포함한다. 따라서, 수성 반응 시스템에서 생화학적 반응을 수행하기 위한 본 명세서에 기재된 임의의 공정, 용도 또는 방법 등은 생체내 수성 반응 시스템을 배제할 수 있다.
- [0189] 본 발명의 일부 양태에 있어서 생체의 사용을 수용하지만, 본 발명의 일부 양태에 있어서 생체의 사용의 배제를 포함한다. 따라서, 수성 반응 시스템에서 생화학적 반응을 수행하기 위한 본 명세서에 기재된 임의의 공정, 용도 또는 방법 등은 생체의 수성 반응 시스템을 배제할 수 있다.
- [0190] 본 명세서에 기술된 임의의 방법, 공정, 용도 또는 IDR 제제에 있어서, 하나 이상의 거대분자 또는 폴리펩타이드가 하나 이상의 기능적인 IDR을 포함하지 않는다는 점을 제외하고는 동일한 반응 조건에서 하나 이상의 거대분자 또는 폴리펩타이드 도입 후 시스템에서의 반응 효율과 비교할 때, 시스템에서의 반응 효율은 IDR-거대분자 또는 IDR-폴리펩타이드에 의해 증가된다.
- [0191] 본 명세서에 기술된 하나 이상의 IDR (IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드)를 포함하거나 구성하는 아미노산 서열을 포함하는 하나 이상의 폴리펩타이드를 포함하는 임의의 방법, 공정, 용도 또는 IDR 제제에 있어서, 하나 이상의 거대분자 또는 폴리펩타이드가 하나 이상의 IDR을 포함하거나 구성하는 아미노산 서열로 태그되지 않는다는 점을 제외하고는 동일한 반응 조건에서 하나 이상의 거대분자 또는 폴리펩타이드 도입 후 시스템에서의 반응 효율과 비교할 때, 시스템에서의 반응 효율은 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드에 의해 증가된다.
- [0192] 따라서, IDR-거대분자 또는 IDR-폴리펩타이드, 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드가 시스템에서 반응 효율을 증가시킬 수 있는지 여부는 하나 이상의 기능적 IDR이 있거나 없는 거대분자 또는 폴리펩타이드와 비교함으로써 매우 간단하게 확립될 수 있다. 숙련된 사람은 간단한 비교 테스트를 수행하여 관련 기능을 확립할 수 있다. 예시적인 테스트 분석은 본 명세서에 추가로 설명되어 있다.
- [0193] 유사하게, IDR-거대분자 또는 IDR-폴리펩타이드, 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드는 복수의 상-분리된 수성 구획 내에서 반응 수행에 필요한 분자가 IDR-거대분자 또는 IDR-폴리펩타이드, 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드와 공동국소화(co-localisation)되도록 할 수 있으며, 또는 복수의 상-분리된 수성 구획 내에서 반응의 수행에 필요한 분자의 공동국소화(co-localisation)를 추가로 자극하거나 향상시켜 시스템에서 생화학 반응의 효율을 증가시킬 수 있는지 여부는, 또한 하나 이상의 기능적 IDR이 있거나 없는 공동국소화(co-localisation)과 비교함으로써 매우 간단하게 확립될 수 있다. 다시 말하지만, 숙련된 사람은 간단한 비교 테스트를 수행하여 관련 기능을 확립할 수 있다. 예시적인 테스트 분석은 본 명세서에 추가로 설명되어 있다.
- [0194] 유사하게, 다가 금속 이온을 IDR-거대분자 또는 IDR-폴리펩타이드, 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드에 제공하여 액체-액체 탈혼합 및 복수의 상-분리된 수성 구획의 형성을 추가로 자극하거나 향상시켜 시스템에서 생화학 반응의 효율성을 더욱 증가시킬 수 있는지 여부는 다가 금속 이온을 제공하거나 제공하지 않는 액체-액체 탈혼합을 비교함으로써 매우 간단하게 확립될 수 있다. 다시 말하지만, 숙련된 사람은 간단한 비교 테스트를 수행하여 관련 기능을 확립할 수 있다. 예시적인 테스트 분석은 본 명세서에 추가로 설명되어 있다.
- [0195] 유사하게 IDR- 거대분자 또는 IDR-폴리펩타이드, 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드에 ATP를 제공하는 것이 액체-액체 탈혼합 및 복수의 상-분리된 수성 구획의 형성을 추가로 시뮬레이션하거나 향상시켜 시스템에서 생화학 반응의 효율성을 더욱 증가시킬 수 있는지 여부는 ATP를 제공하거나 제공하지 않는 액체-액체 탈혼합을 비교함으로써 매우 간단하게 확립될 수 있다. IDR- 거대분자 또는 IDR-폴리펩타이드, 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드에 ATP를 제공하는 것이 다수의 상-분리된 수성 구획 내에서

반응의 수행에 필요한 분자의 공동국소화(co-localisation)을 추가로 자극하거나 향상시켜 시스템에서 생화학적 반응의 효율성을 증가시키는지 여부는 또한 ATP를 제공하거나 제공하지 않는 공동국소화(co-localisation)을 비교함으로써 매우 간단하게 확립될 수 있다. 다시 말하지만, 숙련된 사람은 간단한 비교 테스트를 수행하여 관련 기능을 확립할 수 있다. 예시적인 테스트 분석은 본 명세서에 추가로 설명되어 있다.

[0196] 분석은 상-분리된 수성 입자의 형성을 유발하는 능력을 참조하여 액체-액체 탈혼합을 유발하는 능력을 확립하기 위해 본 명세서에 설명되어 있다(예를 들어, 본 명세서에 설명된 "상 분리 분석 방법" 참조). 동일한 분석을 사용하여 상 분리된 수성 구획(입자) 내에서 반응을 수행하는 데 필요한 분자의 공동국소화(co-localisation)를 일으키는 기능을 확립할 수 있다. 분석은 RPA 방법의 효율성을 증가시키는 능력을 참조하여 반응의 효율성을 증가시키는 능력을 확립하기 위해 본 명세서에 설명되어 있다(예를 들어, 본 명세서에 설명된 "RPA 분석 방법" 참조). 이러한 분석은 하나 이상의 기능적인 IDR(intrinsically disordered region)으로 구성되거나 이를 포함하는 아미노산 서열의 반응 효율을 증가시키는 능력의 평가 및/또는 추가로 반응의 효율성을 증가시키는 2가 금속 이온의 능력의 평가 및/또는 추가로 반응의 효율성을 증가시키는 ATP 능력의 평가에 사용할 수 있다.

[0197] 본 명세서에 기술된 간단한 분석을 사용하여 당업자는 5% 이상의 반응 효율 증가를 결정할 수 있으며, 반응 효율의 증가는 10% 이상, 15% 이상, 20% 또는 이상, 25% 이상, 30% 이상, 35% 이상, 40% 이상, 45% 이상, 50% 이상, 55% 이상, 60% 이상, 65% 이상, 70% 또는 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상 또는 100% 이상일 수 있다.

[0198] 크라우딩 제제

[0199] 크라우딩 제제는 일반적으로 단백질 또는 합성 블록 폴리머와 같은 분자량이 높은 거대분자이다. 크라우딩 제제는 본질적으로 생화학적 불활성인 것으로 간주된다. 즉, 특정 상호작용이나 촉매 작용에 기여하지 않는다.

[0200] 크라우딩 제제는 생체 외 시스템이든 생체 내 시스템이든 생물학적/생화학적 시스템에 영향을 미치며 용액에서 체적을 물리적으로 점유하여 입체적 장애를 일으키고 열린 용매 공간(open solvent space) 가용성을 감소시키는 것으로 널리 상정된다. 이러한 배제된 부피 메커니즘에 의해 크라우딩 제제는 해리 상수를 변경 및 특정 조직화된 복합체로 함께 모이는 다중 단백질과 같은 상호작용 거대분자 결합 선호의 특정 효과를 갖는 다른 거대분자의 효과 농도(effective concentration)를 증가시키는 것으로 보인다. 크라우딩 효과의 크기는 특히 포함된 분자의 분자 질량에 따라 달라지며 일반적으로 분자가 클수록 더 강하다. 따라서 일반적으로 거대분자 크라우딩은 다른 큰 분자의 특성에 큰 분자가 가하는 효과이다. 또한, 크라우딩 제제는 반응물이 마이크론 크기의 상-분리된 입자로 분리되는 바람직한 상으로 생물학적/생화학적 시스템의 형성을 촉진할 수 있는 것으로 널리 설명되었다. 이 효과는 부피가 많이 차지하는 벌크 용매로 쉽게 확산할 수 없기 때문에 상대적으로 "제한"되는 거대분자 복합체의 해리 상수에 대한 부피 배제 효과에서 실질적으로 발생한다. 추가적으로 및/또는 대안적으로, 폴리에틸렌 글리콜과 같은 블록 사슬 폴리머와 같은 일부 크라우딩 제제는 코스모트로픽(kosmotropic) 특성을 나타낼 수 있으며, 이는 이들이 일반적으로 물 밀도를 낮추어 벌크 물의 구조에 전반적인 변경을 가하도록 한다. 벌크 용매 특성의 이러한 변화는 표면이 물과 상호 작용해야 하는 다른 거대분자 및 그 어셈블리에 복잡한 영향을 미칠 수도 있다. 이것은 차례로 생물학적 성분이 상당히 풍부하고 동시에 주로 벌크 용매 상을 차지하는 크라우딩 제제가 고갈되는 대체 상으로 다른 거대분자의 분리(segregation)를 촉진할 수 있다.

[0201] 단순한 부피 점유 또는 용매 변경의 두가지 시나리오에서, 거대분자의 상-분리 응축을 자극하는 크라우딩 제제 개별 또는 조합의 효과는 응축물 요소의 관점에서 "인력(attractive)" 매커니즘이 아니라 "척력(repulsive)" 매커니즘에 의해 작동하는 것으로 보인다. 다시 말해서, 응축물에 매우 풍부한 성분의 관점에서 볼 때, 크라우딩 제제는 확산에 의해 쉽게 침투할 수 없는 벌크 상 환경을 생성 및/또는 분산이 어려운 순 엔탈피를 나타내는 방식으로 벌크 용매 특성이 변경되는 벌크 상 환경을 생성하도록 작용한다. 고농도 크라우딩 제제의 효과는 응축물 성분이 크라우딩 제제가 없을때처럼 쉽게 분산되지 못하기 때문에 현상이 발생하는 한, "방해적(obstructive)" 또는 "척력(repulsive)" 매커니즘을 통해 기능함으로써 상 분리를 자극하는 데 있어 본 명세서에서 언급된다. 그러나 동시에, 일반적으로 불활성인 특성의 관점에서 크라우딩 제제는 시스템의 다른 특정 분자에 직접적인 약화 효과가 거의 없거나 전혀 없는데, 예를 들어, 크라우딩 제제는 직접적인 방식으로 특정 분자 축쇄와 크게 상호 작용하거나 영향을 미치지 않는다.

[0202] 표준 RPA 반응에서 폴리에틸렌 글리콜(polyethylene glycol, PEG)은 재조합/DNA 합성에 중대한 영향을 미칠 수 있다. PEG는 예를 들어 RecA가 Gp32와 결합될 때 발생하는 다중 침습/신장 주기(invasion/extension cycle)의 수에 영향을 미칠 수 있다. PEG는 여러 가지 방식으로 구성된 증폭 반응을 자극할 수 있다. PEG 및 기타 유사한 크라우딩 제제는 Gp32 및 리콥비네이즈의 협동성에 영향을 미칠 수 있으며, 폴리머레이즈 진행성에 영향을 미칠

수 있으며 용액에서 올리고뉴클레오타이드의 혼성화 속도 및 거동에 영향을 미칠 수 있다. 폴리에틸렌글리콜의 사슬 길이는 결과에 영향을 줄 수 있다. PEG는 또한 리콤비네이즈-로드된(recombinase-loaded) 필라멘트의 안정성을 증가시킬 수 있으며 증가된 지속성은 RPA 효능을 증가시킬 수 있다.

- [0203] 시험관 내 생화학적 반응 환경에서 그 효과를 발휘하기 위해 첨가된 크라우딩 제제는 일반적으로 입체 배제/구속 효과(steric exclusion/confinement effect)가 발생할 것으로 예상되는 농도, 일반적으로 반응 부피 또는 반응 중량 기준으로 약 1% 이상으로 존재한다.
- [0204] 표준 RPA 반응에서 크라우딩 제제는 반응 부피 또는 반응 중량 기준으로 약 1% ~ 12%의 농도로 존재한다.
- [0205] 본 명세서의 용어 " 거대분자 크라우딩 제제" 또는 더 간단히 "크라우딩 제제"는 매우 잘 알려져 있고 당업계에 이해되는 용어이다. 이는 해당 용어가 광범위하게 사용되는 문헌상 명백하다. 예를 들어 Kuznetsova, I., M. *et al.*(What Macromolecular Crowding Can Do to a Protein, 2014, Int. J. Mol. Sci., 15, pp 23090-23140)은 320개의 논문이 있으며 해당 분야의 현재 지식에 대한 가장 포괄적인 개요 중 하나를 나타내는 것으로 제안된다. 본 명세서의 용어 "크라우딩 제제"는 이의 범용적 사용을 강조하는 텍스트 전체에서 광범위하게 사용된다(Mixed Macromolecular Crowding: A Protein and Solvent Perspective, Biswas, S. *et al.*, 2018, ACS Omega, 3(4), pp4316-4330 및 and Common Crowding Agents Have Only a Small Effect on Protein-Protein Interactions, Phillip Y. *et al.*, 2009, Biophysical Journal, 97 pp875-885 875참조).
- [0206] 화합물 또는 거대분자는 당업계에 공지된 수단에 의해 크라우딩 제제로서 식별될 수 있다. 예를 들어, 크라우딩 제제는 실험적으로 결정되고 계산된 유체역학적 반경을 통해 식별될 수 있다(Kuznetsova *et al.*, *supra*). 크라우딩 제제는 졸-겔 유리 캡슐화 분석을 통해 식별될 수 있다(Kuznetsova *et al.*, *supra*).
- [0207] 다음 화합물은 알려진 크라우딩 제제의 예시이다. 합성 블록 폴리머(synthetic block polymer), 폴리에틸렌 글리콜(polyethylene glycol, PEG), PEG 1450, PEG 2050, PEG3000, PEG 4600, PEG 6000, PEG 8000, PEG 10000, PEG 20000, PEG 35000, 분자량 15,000 ~ 20,000의 PEG 화합물(Carbowax 20M으로 알려짐), 덱스트란(dextran), 덱스트란 6, 덱스트란 40, 덱스트란 70, 덱스트란 670, 덱스트란 설페이트 10, 덱스트란 설페이트 500, 피콜(ficoll), 피콜 70, 피콜 400, Poly(sodium 4-styrene sulfonate) (PSS), 소 췌장 트립신 억제제(bovine pancreatic trypsin inhibitor, BPTI), 리보뉴클레아제 A, 리소자임, β -락토글로불린, 헤모글로빈, 소혈청 알부민(bovine serum albumin, BSA).
- [0208] 본 발명의 RPA 방법, 프로세스 및 용도 중 어느 하나를 포함하는 본 발명의 방법, 프로세스 및 용도 중 하나에서, 방법, 프로세스 및 용도는 크라우딩 제제의 부재 하에 수행될 수 있다.
- [0209] 본 발명의 RPA 방법, 공정 및 용도 중 어느 하나를 포함하는 본 발명의 방법, 공정 및 용도 중 어느 하나에서, 방법, 공정 및 용도는 크라우딩 제제의 존재 하에 수행될 수 있다.
- [0210] 본 발명의 RPA 방법, 공정 및 용도 중 어느 하나를 포함하는 본 발명의 방법, 공정 및 용도 중 어느 하나에서, 방법, 공정 및 용도는 크라우딩 제제의 존재 하에 수행될 수 있으며, 상기 크라우딩 제제는 IDR- 거대분자 또는 IDR-폴리펩타이드에 의해 제공되는 생화학적 반응의 효율성 증가의 향상을 제공하는 농도로 제공된다.
- [0211] 본 발명의 RPA 방법, 공정 및 용도를 포함하는 본 발명의 방법, 공정 및 용도 중 어느 하나에서, 방법, 공정 및 용도는 크라우딩 제제의 존재 하에 수행될 수 있으며, 상기 크라우딩 제제는 IDR- 거대분자 또는 IDR-폴리펩타이드에 의해 제공되는 생화학적 반응의 효율에 부가적인 효과를 제공하는 농도로 제공된다.
- [0212] 본 발명의 RPA 방법, 공정 및 용도 중 어느 하나를 포함하는 본 발명의 방법, 공정 및 용도 중 임의의 하나에서, 방법, 공정 및 용도는 크라우딩 제제의 존재 하에 수행될 수 있으며, 상기 크라우딩 제제는 IDR- 거대분자 또는 IDR-폴리펩타이드에 의해 제공되는 생화학적 반응의 효율에 상승 효과를 제공하는 농도로 제공된다.
- [0213] 본 발명의 RPA 방법, 공정 및 용도 중 어느 하나를 포함하는 본 발명의 방법, 공정 및 용도 중 임의의 하나에서, 방법, 공정 및 용도는 크라우딩 제제의 존재 하에 수행될 수 있으며, 상기 IDR- 거대분자 또는 IDR-폴리펩타이드의 생화학 반응 시스템 도입은 IDR- 거대분자의 또는 IDR-폴리펩타이드의 생화학 반응 시스템으로의 도입의 부재하에 생화학 반응의 효율을 동일한 수준으로 증가시키는 데 필요한 크라우딩 제제의 농도를 낮춘다.
- [0214] 크라우딩 제제의 존재하에 수행될 수 있는 전술한 방법, 공정 및 용도 중 어느 하나에서, 크라우딩 제제는 그의 정상적인 생물학적 효과(입체 배제(steric exclusion)/감금 효과(confinement effect))가 발생하는 농도 미만

으로 존재할 수 있다.

- [0215] 크라우딩 제제의 존재 하에 수행될 수 있는 전술한 방법, 공정 및 용도 중 어느 하나에서, 크라우딩 제제는 약 3% 미만의 반응 부피의 농도로 존재할 수 있으며, 약 2 %의 반응 부피 또는 반응 중량, 약 1 %의 반응 부피 또는 반응 중량 또는 약 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2% 또는 0.1%의 반응 부피 또는 반응 중량으로 존재할 수 있다.
- [0216] RPA 반응 방법을 포함하는 본 발명의 방법, 공정 및 용도 중 어느 하나에 사용되는 경우, 임의의 적합한 크라우딩 제제가 사용될 수 있다. 적합한 크라우딩 제제의 예가 본 명세서에 제공된다.
- [0217] **IDR (intrinsically disordered region)을 포함하는 거대분자 또는 폴리펩타이드**
- [0218] 본 발명의 방법, 공정 및 제제는 본 명세서에 기재된 "IDR-태그된 거대분자"를 비롯한 "IDR- 거대분자"를 포함한다. 본 발명의 방법, 공정 및 제제는 본 명세서에 기재된 "IDR-태그된 폴리펩타이드"를 비롯한 "IDR-폴리펩타이드"를 포함한다. 임의의 그러한 IDR- 거대분자, IDR-태그된 거대분자, IDR-폴리펩타이드 또는 IDR-태그된 폴리펩타이드는 본 명세서에서 IDR 제제 또는 IDR-기반 제제로 상호교환가능하게 지칭될 수 있다.
- [0219] 본 명세서에 사용된 IDR- 거대분자 또는 IDR-폴리펩타이드 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드는 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하는 임의의 거대분자, 또는 폴리펩타이드 또는 단백질이다.
- [0220] 본 명세서에 사용된 IDR- 거대분자 또는 IDR-폴리펩타이드 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드는 하나 이상의 기능적인 IDR(intrinsically disordered region)으로 구성되거나 이를 포함하는 아미노산 서열을 포함하는 임의의 거대분자, 또는 폴리펩타이드 또는 단백질이다.
- [0221] 따라서, 본 명세서에 언급된 IDR-거대분자 또는 IDR-폴리펩타이드는 하나 이상의 기능적인 IDR(intrinsically disordered region)으로 이루어진 아미노산 서열을 포함하는 거대분자 또는 폴리펩타이드; 또는 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하는 아미노산 서열을 포함하는 거대분자 또는 폴리펩타이드를 지칭할 수 있다.
- [0222] 또한, 본 명세서에서 언급된 바와 같이, 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하는 IDR-거대분자는 하나 이상의 기능적인 IDR(intrinsically disordered region)로 구성되거나 이를 포함하는 아미노산 서열로 태그가 부착된 관심 거대분자일 수 있다. 이러한 IDR-태그된 거대분자는 또한 본 명세서에 정의된 IDR 제제이다. 본 명세서에 언급된 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하는 IDR-태그된 폴리펩타이드는 하나 이상의 기능적인 IDR(intrinsically disordered region)로 구성되거나 이를 포함하는 아미노산 서열로 태그된 관심 폴리펩타이드일 수 있다. 이러한 IDR-태그된 폴리펩타이드는 또한 본 명세서에 정의된 바와 같은 IDR 제제이다.
- [0223] 본 명세서에 사용된 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드는 하나 이상의 기능적인 IDR(intrinsically disordered region)로 구성되거나 이를 포함하는 아미노산 서열로 "태그된" 임의의 거대분자, 또는 폴리펩타이드 또는 단백질이다.
- [0224] 따라서, 본 명세서에서 언급된 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드는 하나 이상의 기능적인 IDR(intrinsically disordered region)로 이루어진 아미노산 서열로 태그된 거대분자 또는 폴리펩타이드; 또는 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하는 아미노산 서열로 태그된 거대분자 또는 폴리펩타이드를 지칭할 수 있다.
- [0225] 하나 이상의 기능적인 IDR(intrinsically disordered region)으로 구성되거나 이를 포함하는 태그된 아미노산 서열은 태그된 거대분자 또는 폴리펩타이드 또는 단백질의 태그된 위치에서 자연적으로 또는 일반적으로 발견되지 않는다. 따라서, 하나 이상의 기능적인 IDR(intrinsically disordered region)로 구성되거나 이를 포함하는 태그된 아미노산 서열은 그것이 태그된 거대분자 또는 폴리펩타이드 또는 단백질과 비교하여 외인성 아미노산 서열인 것으로 간주될 수 있다. 태그된 거대분자 또는 폴리펩타이드 또는 단백질은 따라서 비자연적 발생, 인공 또는 유전적으로 조작된 거대분자, 폴리펩타이드 또는 단백질로 간주될 수 있다.
- [0226] 아미노산 서열이 폴리펩타이드 및 다른 거대분자에 "태그"될 수 있는 메커니즘은 본 명세서에서 추가로 설명된다.
- [0227] 임의의 하나 이상의 기능적인 IDR(intrinsically disordered region)은 본 명세서에 개시된 임의의 특정 IDR

아미노산 태그 서열 중 하나 이상, 또는 그의 임의의 하나 이상의 기능적 변이체, 유사체, 상동체 또는 그 유도체에 태그될 수 있다.

- [0228] 본 발명에서 사용하기 위해서는 IDR(intrinsically disordered region) 폴리펩타이드 서열 및 이의 도메인 모두가 "기능적"이어야 한다. 본 명세서의 용어 "기능적"은 임의의 IDR 아미노산 서열이 본 명세서에 추가로 약속된 기능적 특성 중 하나를 보유해야 함을 의미한다.
- [0229] 본 명세서의 용어 "IDR(intrinsically disordered region)"은 기술 분야에서 일반적으로 사용되는 당업계에서 이해되는 용어이다. 포괄적인 검토는 다음을 참조한다. Classification of Intrinsically Disordered Regions and Proteins, van der Lee et al., 2014, Chem. Rev. 114, pp 6589-6631.
- [0230] 본 발명은 특히 수성 시험관내 반응 시스템에서 생화학적 반응을 수행하는 공정을 제공하며, 상기 생화학적 반응은 하나 이상의 반응 거대분자, 선택적으로 하나 이상의 반응 폴리펩타이드의 기능에 의존하며, 상기 공정은: 반응을 수행하기에 적합한 조건 하에 시험관내 반응 시스템 내로 적어도 하나의 IDR- 거대분자를 도입하는 단계, 상기 적어도 하나의 IDR- 거대분자는 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하고, 상기 적어도 하나의 IDR- 거대분자의 시험관 내 반응 시스템으로 도입시 거대분자 생화학 반응의 효율은 적어도 하나의 IDR- 거대분자에 의해 증가된다. 생화학 반응의 효율은 IDR- 거대분자의 하나 이상의 기능적 IDR에 의해 증가된다. 임의의 그러한 공정에서 적어도 하나의 IDR- 거대분자는 적어도 하나의 IDR-폴리펩타이드일 수 있다. 임의의 그러한 공정에서, 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하는 IDR-거대분자 또는 IDR-폴리펩타이드는 그 기능에 생화학적 반응이 의존하는 "반응 거대분자" 또는 "반응 폴리펩타이드"가 아닐 수 있다. 따라서 이러한 모든 과정에서 IDR-거대분자 또는 IDR-폴리펩타이드는 생화학적 반응 자체에서 고유한 생화학적 역할을 하지 않을 수 있다. 그럼에도 불구하고 반응 시스템에 도입하면 생화학 반응의 효율성이 증가한다.
- [0231] 시험관내 반응 시스템에서 생화학적 반응을 수행하는 공정은, 상기 생화학적 반응이 하나 이상의 IDR-거대분자의 기능에 의존하는 과정일 수 있으며, 상기 시험관내 반응 시스템으로의 도입시 하나 이상의 IDR-거대분자는 생화학 반응에서 반응 기능을 수행하고 반응의 효율을 증가시킨다. 임의의 그러한 공정에서 적어도 하나의 IDR-거대분자는 적어도 하나의 IDR-폴리펩타이드일 수 있다. 이러한 과정에서 IDR-거대분자 또는 IDR-폴리펩타이드는 생화학적 반응 자체에서 고유한 생화학적 역할을 한다. 따라서, 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하는 적어도 하나의 IDR-거대분자 또는 IDR-폴리펩타이드는 그 기능에 생화학적 반응이 의존하는 "반응 거대분자" 또는 "반응 폴리펩타이드"이다.
- [0232] 적어도 하나의 IDR-거대분자 또는 IDR-폴리펩타이드는 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하거나 이로 이루어진 아미노산 서열을 포함한다. IDR-거대분자 또는 IDR-폴리펩타이드는 생화학 반응을 수행하기에 적합한 조건 하에서 생화학 반응 시스템에 도입된다. 하나 이상의 기능적인 IDR(intrinsically disordered region)의 존재로 인해 IDR-거대분자 또는 IDR-폴리펩타이드는 반응 효율을 증가시킨다.
- [0233] 반응 효율을 증가시킨다는 것은, IDR-거대분자 또는 IDR-폴리펩타이드가 하나 이상의 기능적인 IDR(intrinsically disordered region)을 포함하거나 이로 구성된 아미노산 서열 없이 제공되는 경우 관찰되는 반응의 효율과 비교해 반응의 효율이 개선됨을 의미한다. 이러한 개선은 IDR 아미노산 서열이 있거나 없는 반응 거대분자 또는 폴리펩타이드의 비교 시험에 의해 쉽게 확립될 수 있다.
- [0234] 본 발명의 방법 중 임의의 어느 하나에서, IDR-거대분자 또는 IDR-폴리펩타이드의 하나 이상의 기능적인 IDR(intrinsically disordered region)은 상 분리를 유도하는 반응 시스템에서 액체-액체 탈혼합을 촉진/유발한다. 반응 시스템에서 상 분리를 유도하는 액체-액체 탈혼합을 촉진하는 기능적 능력은 예를 들어 본 명세서에 기재된 바와 같은 상 분리 분석을 수행함으로써 쉽게 확립될 수 있다. 이러한 액체-액체 탈혼합은 상 분리를 촉진하고 이는 예를 들어 본 명세서에서 더 자세히 설명하는 것처럼 현미경 관찰 하에 검출 가능한 입자와 같은, 반응 시스템에서 분리된 구획의 형성으로 이어질 수 있다.
- [0235] IDR-거대분자 또는 IDR-폴리펩타이드는 촉매 활성을 나타내거나 나타내지 않을 수 있다. 예를 들어, IDR-폴리펩타이드는 본 명세서에 추가로 기재된 바와 같이 리코비네이즈 폴리머레이즈 증폭 반응에 사용되는 폴리머레이즈 효소와 같은 촉매 활성을 가질 수 있다. IDR-폴리펩타이드는 리코비네이즈 폴리머레이즈 증폭 반응에 사용되는, 본 명세서에 추가로 기재된 Gp32 같은 단일 가닥 안정화제와 같은 촉매 활성을 갖지 않을 수 있다.
- [0236] 본 명세서에서 추가로 논의된 바와 같이, IDR-거대분자 또는 IDR-폴리펩타이드가 촉매 활성을 갖고 있든 없든,

IDR-거대분자 또는 IDR-폴리펩타이드는 생화학 반응에 필요하거나 영향을 미치는 기능을 가질 수 있으므로 IDR-거대분자 또는 IDR-폴리펩타이드의 부재시 생화학 반응계의 생화학 반응이 진행되지 않거나 효율이 저하되어 진행된다. 대안적으로, 본 명세서에서 추가로 논의된 바와 같이, IDR-거대분자 또는 IDR-폴리펩타이드는 생화학적 반응 자체에 필요하거나 영향을 미치는 기능을 보유하지 않을 수 있다. 그럼에도 불구하고 IDR 아미노산 서열 때문에 생화학 반응 시스템에 IDR-거대분자 또는 IDR-폴리펩타이드를 도입하면 IDR-거대분자 또는 IDR-폴리펩타이드 부재시, 또는 IDR 아미노산 서열이 없는 동일한 거대분자 또는 폴리펩타이드의 존재 하에 관찰되는 효율에 비해 생화학 반응의 효율이 증가한다.

[0237] **IDR 폴리펩타이드의 구조적 특성(Structural properties of an IDR polypeptide)**

[0238] 아미노산 서열에서 IDR의 존재는 구조 분석에 의해 쉽게 결정할 수 있다. 폴리펩타이드 및 단백질 내 IDR의 존재 예측을 위해 많은 생물정보학 기반 플랫폼을 사용할 수 있다. 여기에는 ELM, MiniMotif, SLiMPrints, phylo-HMM, DiliMot, SLiMfinder, Phospho.ELM, PhosphoSite, PHOSIDA, ScanSite, NetPhorest, NetworkKIN, PhosphoNET, IDEAL, MoRFpred, ANCHOR, Pfam, FFPred, DisProt, D²P², and MetaDisorder가 포함된다. 이들 방법 중 임의의 것이 IDR 아미노산 서열을 확인하는 데 사용될 수 있다. 필요한 경우, 이러한 IDR 아미노산 서열은 본 명세서에 추가로 기재된 바와 같이 기능적 특성을 평가하기 위해 테스트될 수 있다.

[0239] IDR 아미노산 서열 식별을 위한 선호되는 생물정보학 기반 플랫폼은 MetaDisorder 소프트웨어 프로그램이다 (MetaDisorder: a meta-server for the prediction of internal disorder in protein. Kozlowski, L. P., et al., BMC Bioinformatics, 2012, 13(1)). : 111).

[0240] MetaDisorder 프로그램은 온라인 (<http://genesilico.pl/metadisorder/>) 에서 무료로 사용할 수 있다. 이 프로그램을 사용하여 관심 있는 아미노산 서열을 인터넷 브라우저 창에 붙여넣으면 프로그램이 시작된다. 온라인 문서에서 설명하는 것처럼 소프트웨어 패키지에서 점수가 >0.5인 모든 아미노산 영역은 IDR (intrinsically disordered region)을 포함하는 것으로 간주된다.

[0241] MetaDisorder 소프트웨어 플랫폼을 사용하여, 본 발명자들은 하나 이상의 IDR (intrinsically disordered region)을 포함하는 다수의 아미노산 서열을 확인하였다. 이는 표 1에 나타내었다.

[0242] 따라서, 방법, 공정 및 용도 중 임의의 하나, 또는 비-천연 발생 IDR-거대분자, IDR-융합 거대분자 또는 이를 암호화하는 단리된 핵산 분자, 재조합 폴리뉴클레오타이드 발현 벡터 또는 숙주 세포에서, IDR-거대분자 또는 IDR-폴리펩타이드의 하나 이상의 기능적 IDR은 알고리즘 MetaDisorder에 의해 분석될 때 0.5보다 큰 점수를 갖는 아미노산의 서열로서 특징될 수 있다. 아미노산의 서열은 Kozlowski, L.P., et al., BMC Bioinformatics, 2012, 13(1): 111의 방법에 따른 MetaDisorder 알고리즘에 의해 분석될 때 0.5보다 큰 점수를 받은 아미노산의 서열일 수 있다.

[0243] 본 발명은 서열번호 1 내지 43(표 1) 중 어느 하나의 아미노산 서열 및 그의 변이체를 포함하거나 이로 이루어진 바람직한 IDR 아미노산 서열을 제공하고 이에 관한 것이다. 모든 경우에 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 변이체는 본 명세서에 추가로 기재된 바와 같이 IDR 기능적 특성을 유지하는 기능적 변이체이다.

[0244] 또한, 본 명세서에 추가로 기재된 바와 같이, IDR-거대분자 또는 IDR-폴리펩타이드는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열을 포함하거나 이로 구성된 아미노산 서열로 태그된 거대분자 또는 폴리펩타이드를 포함하거나 이로 이루어질 수 있고, 또는 서열번호 1 내지 43의 기능적 변이체 아미노산 서열을 포함하거나 이로 구성된다.

[0245] 기능적 변이체는 본 명세서에 기재된 IDR 아미노산 서열과 비교하여 적어도 80% 서열 동일성을 가질 수 있다(표 1). 기능적 변이체는 본 명세서에 기재된 IDR 아미노산 서열(표 1)과 비교하여 적어도 81% 서열 동일성, 또는 82% 서열 동일성, 또는 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% 또는 99%의 서열 동일성을 가질 수 있다.

[0246] 본 발명의 목적을 위해, 퍼센트 동일성을 결정하기 위해, 예를 들어, 서열번호 1 내지 43 중 임의의 하나의 아미노산 서열과 서열번호 1 내지 43 중 임의의 하나의 아미노산 서열의 기능적 변이체 사이에 두 개의 각각의 아미노산 서열은 최적의 비교를 위해 정렬된다(예를 들어, gaps 제2 서열과의 최적 정렬을 위해 제1 서열에 도입될 수 있음). 그런 다음 뉴클레오타이드 위치의 뉴클레오타이드 잔기를 비교한다. 첫 번째 서열의 위치가 두 번째 서열의 상응하는 위치와 동일한 뉴클레오타이드 잔기에 의해 점유될 때, 뉴클레오타이드는 그 위치에서 동일

하다. 두 서열 간의 퍼센트 동일성은 서열이 공유하는 동일한 위치 수의 함수이다(즉, % 동일성 = 동일한 위치의 수/참조 서열의 총 위치 수 x 100).

[0247] 일반적으로 시퀀스 비교는 참조 시퀀스의 전장(full length)에 대해 수행된다. 예를 들어, 당업자가 주어진 ("변이체") 서열이 서열번호2와 80% 서열 동일성이 있는지 결정하기를 원하면, 서열번호 2는 참조 서열이 된다. 예를 들어, 변이 서열이 서열번호 2(참조 서열의 예)와 80% 이상 동일한지 여부를 평가하기 위해, 당업자는 서열번호2의 길이에 걸쳐 정렬을 수행하고 테스트 서열에서 얼마나 많은 위치가 서열번호2와 동일한지 확인한다. 위치의 80% 이상이 동일하면 테스트 시퀀스는 서열번호2와 80%이상 동일하다. 서열번호2 보다 서열이 짧은 경우, 공백 또는 누락된 위치는 동일하지 않은 위치로 간주되어야 한다.

[0248] 당업자는 2개의 서열 사이의 상동성 또는 동일성을 결정하는데 이용가능한 상이한 컴퓨터 프로그램을 알고 있다. 예를 들어, 수학적 알고리즘을 사용하여 서열 비교 및 두 서열 간의 동일성 백분율 결정을 수행할 수 있다. 2개의 아미노산 또는 핵산 서열 사이의 퍼센트 동일성은, 예를 들어 Blosun 62 매트릭스 또는 PAM250 매트릭스, 갭 가중치 16, 14, 12, 10, 8, 6 또는 4, 및 길이 가중치 1, 2, 3, 4, 5 또는 6을 사용하는 Accelrys GCG 소프트웨어 패키지(<http://www.accelrys.com/products/gcg/> 에서 이용가능)의 GAP 프로그램에 통합된 Needleman and Wunsch(1970) 알고리즘을 사용해 결정될 수 있다.

[0249] 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이체는 서열번호 1 내지 43 중 어느 하나의 아미노산 서열과 각각 비교하여 적은 수의 아미노산을 가짐으로써(즉, 기능적 변이체가 더 짧음), 각각 서열번호 1 내지 43 중 어느 하나의 아미노산 서열과 비교하여 더 많은 아미노산을 가짐으로써(즉, 기능적 변이체가 더 김) 상이한 아미노산 서열일 수 있다. 따라서, 기능적 변이체는 참조 아미노산 서열과 비교하여 하나 이상의 아미노산 결실 및/또는 하나 이상의 삽입을 포함할 수 있다. 변이가 참조 서열과 상이한 기능적 변이체 아미노산 서열 내의 아미노산의 수는 1개 이상, 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상, 10개 이상, 11개 이상, 12개 이상, 13개 이상, 14개 이상, 15개 이상, 16개 이상, 17개 이상, 18개 이상, 19개 이상 또는 20개 이상 일 수 있다.

[0250] 서열번호 1 내지 43 중 어느 하나의 아미노산 서열의 기능적 변이체는 예를 들어 표 1에 열거된 서열에 표시된 아미노산 잔기의 보존적 아미노산 치환을 포함할 수 있다. 보존적 치환은 예를 들어 일반적으로 허용되는 아미노산 그룹을 설명하는 다음 표에 따라 이루어질 수 있다. 따라서 기능적 변이체는 참조 아미노산 서열과 비교하여 보존적 아미노산 치환을 포함할 수 있다. 참조 서열과 비교하여 보존적 아미노산 치환인 기능적 변이체 아미노산 서열 내의 아미노산의 수는 1개 이상, 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상, 10개 이상, 11개 이상, 12개 이상, 13개 이상, 14개 이상, 15개 이상, 16개 이상, 17개 이상, 18개 이상, 19개 이상 또는 20개 이상일 수 있다.

Set		Sub-set	
Hydrophobic	F W Y H K M I L V A G C	Aromatic	F W Y H
		Aliphatic	I L V
Polar	W Y H K R E D C S T N Q	Charged	H K R E D
		Positively charged	H K R
		Negatively charged	E D
Small	V C A G S P T N D	Tiny	A G S

[0251] 20

[0252] 주어진 변이체가 IDR 기능적 특성을 유지하는지 여부는 예를 들어 여기에 추가로 설명된 방법에 의해 쉽게 확립될 수 있다.

표 1

Tag	Original protein	Accession ID	Organism	Position	Length	Comment	Sequence	SEQ ID NO
fib (or fib-1)	rRNA 2'-O-methyltransferase fibrillarlin	Uniprot: P22087	Homo sapiens	3-45	43	N/A	ProGlyPheSerProArgGlyGlyGlyPheGlyGlyArgGlyGlyPheGlyAspArgGlyGlyArgGlyGlyArgGlyGlyPheGlyGlyGlyArgGlyGlyPheArgGlyArgGlyArg	1
fib2	rRNA 2'-O-methyltransferase fibrillarlin	Uniprot: P22087	Homo sapiens	3-45	86	double of fib tag	ProGlyPheSerProArgGlyGlyGlyPheGlyGlyArgGlyGlyPheGlyGlyArgGlyGlyGlyGlyPheArgGlyArgGlyArgProGlyPheSerProArgGlyGlyGlyPheGlyGlyArgGlyGlyPheGlyAspArgGlyGlyArgGlyGlyPheGlyGlyArgGlyArgGlyGlyArg	2
Fib [short]	rRNA 2'-O-methyltransferase fibrillarlin	Uniprot: P22087	Homo sapiens	8-32	25	N/A	ArgGlyGlyGlyPheGlyGlyArgGlyGlyPheGlyAspArgGlyGlyArgGlyGlyArgGlyGlyPheGlyGly	3
fib-2 (clipper 1)	N/A	N/A	N/A	N/A	36	self-designed: 7 amino acids deletion at the C-terminus of fib (fib-1)	ProGlyPheSerProArgGlyGlyGlyPheGlyGlyArgGlyGlyPheGlyAspArgGlyGlyArgGlyGlyArgGlyGlyPheGlyGlyGlyArgGlyGly	4

[0253]

fib-3 (clipper 2)	N/A	N/A	N/A	N/A	27	self-designed: 16 amino acids deletion at the C-terminus of fib (fib-1)	ProGlyPheSerProArgGlyGlyGlyPheGlyGlyArgGlyGlyPheGlyAspArgGlyGlyArgGlyGlyArgGlyGly	5
fib-4 (clipper 3)	N/A	N/A	N/A	N/A	15	self-designed: 28 amino acids deletion at the C-terminus of fib (fib-1)	ProGlyPheSerProArgGlyGlyGlyPheGlyGlyArgGlyGly	6
PCF	Protein PCF11	Uniprot: P39081	Saccharomyces cerevisiae	208-287	67	modification: amino acids deletion from original positions 234-246	GlnValGlnMetGlnLeuArgGlnValPheSerGlnAspGlnGlnValLeuGlnGlnArgMetArgTyrHisGlnLeuGlnGlnGlnGlnGlnGlnTyrHisGlnThrLysAspMetValGlySerTyrThrGlnAsnSerAsnSerAlaIleProLeuPheGlyAsnAsnSerAspThrThrAsnGlnGlnAsnSer	7
ENT1	Ent2p	Uniprot: C7GIT6	Saccharomyces cerevisiae	530-588	59	N/A	GlnAsnProThrGlyIleSerTyrSerGlnProGlnGlnGlnGlnGlnProGlnGlnGlnProGlnTyrMetGlnAsnPheGlnGlnGlnGlnGlnProGlnTyrAlaGlnAsnPheGlnGlnGlnGlnProGlnTyrThrGlnAsnTyrGlnGlnGlnProGlnTyrIleGlnProHisGln	8
HRP1	Nuclear polyadenylated RNA-binding protein 4	Uniprot: Q99383	Saccharomyces cerevisiae	331-388	58	N/A	GlyGlyAsnAsnGlyGlyAsnAsnMetAsnArgArgGlyGlyAsnPheGlyAsnGlnGlyAspPheAsnGlnMetTyrGlnAsnProMetMetGlyGlyTyrAsnProMetMetAsnProGlnAlaMetThrAspTyrTyrGlnLysMetGlnGluTyrTyrGlnGlnMetGln	9

[0254]

HRP2	Nuclear polyadenylated RNA-binding protein 4	Uniprot: Q99383	Saccharomyces cerevisiae	331-388	116	double of HRP1 tag	GlyGlyAsnAsnGlyGlyAsnAsnMetAsnArgArgGlyGlyAsnPheGlyAsnGlnGlyAspPheAsnGlnMetTyrGlnAsnProMetMetGlyGlyTyrAsnProMetMetAsnProGlnAlaMetThrAspTyrTyrGlnLysMetGlnGluTyrTyrGlnGlnMetGln	10
DDX	Probable ATP-dependent RNA helicase DDX4	Uniprot: Q9NQ10	Homo sapiens	4-240	240	N/A	MetGlyAspGluAspTrpGluAlaGluIleAsnProHisMetSerSerTyrValProIlePheGluLysAspArgTyrSerGlyGluAsnGlyAspAsnPheAsnArgThrProAlaSerSerSerGluMetAspAspGlyProSerArgAspHisPheMetLysSerGlyPheAlaSerGlyArgAsnPheGlyAsnArgAspAlaGlyGluCysAsnLysArgAspAsnThrSerThrMetGlyGlyPheGlyValGlyLysSerPheGlyAsnArgGlyPheSerAsnSerArgPheGluAspGlyAspSerSerGlyPheTrpArgGluSerSerAsnAspCysGluAspAsnProThrArgAsnArgGlyPheSerLysArgGlyGlyTyrArgAspGlyAsnAsnSerGluAlaSerGlyProTyrArgArgGlyGlyArgGlySerPheArgGlyCysArgGlyGlyPheGlyLeuLysSerProAsnAsnAspLeuAspProAspGluCysMetGlnArgThrGlyGlyLeuPheGlySerArgArgProValLeuSerGlyThrGlyAsnGlyAspThrSerGlnSerArgSerGlySerGlySerGluArgGlyGlyTyrLysGlyLeuAsnGluGluValIleThrGlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyGluSerSerAspThrGln	11
HNRNPAl	Heterogeneous nuclear ribonucleoprotein A1	Uniprot: Q0VAC0	Homo sapiens	187-320	134	N/A	AlaSerAlaSerSerSerGlnArgGlyArgSerGlySerGlyAsnPheGlyGlyGlyArgGlyGlyPheGlyGlyAsnAspAsnPheGlyArgGlyGlyAsnPheSerGlyArgGlyGlyPheGlyGlySerArgGlyGlyGlyGlyTyrGlyGlySerGlyAspGlyTyrAsnGlyPheGlyAsnAspGlySerAsnPheGlyGlyGlyGlySerTyrAsnAspPheGlyAsnTyrAsnAsnGlnSerSerAsnPheGlyProMetLysGlyGlyAsnPheGlyGlyArgSerSerGlyProTyrGlyGlyGlyGlyGlnTyrPheAlaLysProGlnAsnGlnGlyTyrGlyValSerSerSerSerSerSerTyrGlySerGlyArgArgPhe	12

[0255]

[0256]

[illegible]

mimic2	N/A	N/A	N/A	N/A	49	self-designed: seven repeats of YSFTDP8 (SEQ ID NO: 130) motifs mimicking the c-terminus of RNA polymerase II	TyrSerProThrAspProSerTyrSerProThrAspProSerTyxerProThrAspProSerTyrSerProThrAspProSerTyxerProThrAspProSerTyrSerProThrAspProSerTyxerProThrAspProSer	16
C-PolCTD	DNA-directed RNA polymerase II subunit RPB1	Uniprot: P04050	Saccharomyces cerevisiae	1642-1662	21	N/A	ProThrSerProSerTyrSerProThrSerProSerTyrSerProSerTyrSerProAlaTyrSer	17
Sup	Eukaryotic peptide chain release factor GTP-binding subunit	Uniprot: P05453	Saccharomyces cerevisiae	1-114	115	insertion of QY between 61 and 62; mutation: D66Q(A67G; deletion of Q70)	MetSerAspSerAsnGlnGlyAsnAsnGlnGlnAsnTyrGlnGlnTyrSerGlnAsnGlyAsnGlnGlnGlnGlyAsnAsnArgTyrGlnGlyTyrGlnAlaIstyrAsnAlaGlnAlaGlnProAlaGlyGlyTyrTyrGlnAsnTyrGlnGlyTyrSerGlyTyrGlnGlnGlyGlyTyrGlnGlnTyrGlnTyrAsnProGlnGlyGlyTyrGlnGlnTyrAsnProGlnGlyGlyTyrGlnGlnTyrAsnProGlnGlyGlyTyrGlnGlnPheAsnProGlnGlyGlyArgGlyAsnTyrLysAsnPheAsnTyrAsnAsnLeuGlnGlyTyrGln	18
Sup1	N/A	N/A	N/A	N/A	9	self-designed: single repeat of YNPQGGVQQ (SEQ ID NO: 19)	TyrAsnProGlnGlyGlyTyrGlnGln	19

[illegible]

[0258]

His4	Forkhead box p rotein G1	Uniprot: P553 16	Homo sapien s	32-73	42	N/A	AsnHisHisAlaSerHisGlyHisHisAsnSerHisHisProGlnHis HisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis GlnProProProProProGlnGlnGlnGln	26
His5	Transcription factor MafA	Uniprot: Q8NB W3	Homo sapien s	145-220	76	N/A	SerGlyHisHisGlyAlaHisHisGlyAlaHisHisProAlaHisAla AlaAlaTyrGluAlaPheArgGlyProGlyPheAlaGlyGlyGlyGly AlaAspAspMetGlyAlaGlyHisHisHisHisHisHisHisHisHis HisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis GlyGlyAlaGlyHisGlyGlyGlyAlaGlyHisHis	27
His6	Transcriptiona l repressor pr oteins YY1	Uniprot: P254 90	Homo sapien s	54-80	27	N/A	GlyGlyGlyGlyAspHisGlyGlyGlyGlyGlyHisHisHisHisHis HisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis	28
His7	Disks large-as sociated protei n 3	Uniprot: Q958 86	Homo sapien s	214-238	25	N/A	ProGlySerGlyGlyProHisThrSerHisHisHisHisHisHisHis HisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis	29
His8	Nuclear recept or subfamily 4 group A member 3	Uniprot: Q825 70	Homo sapien s	91-115	25	N/A	AlaProSerTyrHisHisHisHisHisHisHisHisHisHisHisHis HisHisGlnGlnGlnHisGlnGlnPro	30
His9	E3 SUMO-protei n ligase CBX4	Uniprot: Q002 57	Homo sapien s	377-400	24	N/A	ThrHisProProSerHisHisProHisProHisProHisHisHisHis HisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis	31
His10	SKI/DACH domai n-containing p rotein 1	Uniprot: Q1XB 10	Homo sapien s	334-371	38	N/A	ProProProHisHisHisHisHisHisHisHisHisHisHisHisHis HisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis HisArgProGlnProHis	32
His11	G8 homeobox 2	Uniprot: Q9B2 M3	Homo sapien s	123-140	18	N/A	AsnHisAlaHisHisHisHisHisHisHisHisHisHisHisHisHis HisGln	33
His12	N/A	N/A	N/A	N/A	37	self-desi gned	AlaGlyHisHisHisHisHisHisHisHisHisHisHisHisHisHis GlyAlaGlyGlyAlaGlyGlyAlaHisHisHisHisHisHisHisHis HisHisGlyGlyAla	34

His13	N/A	N/A	N/A	N/A	37	self-desi gned	AlaGlyHisHisHisHisHisHisHisHisHisHisHisHisHisSerG lySerGlySerSerGlySerGlySerSerHisHisHisHisHisHisH isHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisH	35
His14	N/A	N/A	N/A	N/A	17	self-desi gned	AlaGlyHisHisHisHisHisHisHisHisHisHisHisHisHisHis yGlyAla	36
His15	N/A	N/A	N/A	N/A	24	self-desi gned	AlaGlyHisHisHisHisHisHisHisHisHisHisHisHisHisHis sHisHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis	37
super+ve	single strand ed DNA-bindin g protein	GenBank: AW M11745.1	Escherichia phage vB_Ec _EcoM_NBGI	256-299	74	self-desi gned: dou ble seque nce of ME AFSSAKTED DFMSSSSSD DSDLDLLA GL (SEQ I D NO: 131)and chan ge amino acids D/E to K	LysLysValAlaSerLysLeuLysPheLysLysLysMetGluAla PheSerSerAlaLysThrLysLysLysPheMetSerSerSerS erSerLysLysSerLysLeuLysLysLeuAlaGlyLeuMet GluAlaPheSerSerAlaLysThrLysLysLysPheMetSerSe rSerSerSerLysLysSerLysLeuLysLysLysLeuAlaGlyL eu	38
super-ve	single strand ed DNA-bindin g protein	GenBank: AW M11745.1	Escherichia phage vB_Ec _EcoM_NBGI	256-299	75	self-desi gned: dou ble seque nce of ME AFSSAKTED DFMSSSSSD DSDLDLLA GL (SEQ I D NO: 131)and chan ge amino acid K to D/E	AspAspValAlaSerAspLeuAspAspPheAspAspMetGlu AlaPheSerSerAlaLysThrLysLysLysPheMetSerSerS erSerSerAspAspSerAspLeuAspAspLeuAlaGlyLeu MetGluAlaPheSerSerAlaLysThrLysLysLysPheMetSe rSerSerSerAspAspSerAspLeuAspAspLeuAlaGlyL eu	39

superpositive	RecA-like reco mbination prot ein	GenBank: AYP 69747.1	Escherichia phage vB_Ec _EcoM_DelCa	348-392	90	self-desi gned: dou ble seque nce of EI VEAEVDELI NSHVERPKS FEKSKSAA DLETDLQL SDMEEFN S EQ ID NO: 132) and change am ino acids D/E to R/ F	ArgIleValIleAlaLysValIleArgLeuIleAsnSerLysValIle LysPheLysSerProLysSerLysSerLysSerAlaAlaLysLeu LysThrLysLeuLysGlnLeuSerLysMetLysPheAsnLysIle ValLysAlaLysValLysLysLeuIleAsnSerLysValArgLysPh eLysSerProLysSerLysSerLysSerAlaAlaLysLeuLysThrL ysLeuLysGlnLeuSerLysMetLysGluPheAsn	40
supernegative	RecA-like reco mbination prot ein	GenBank: AYP 69747.1	Escherichia phage vB_Ec _EcoM_DelCa	348-392	90	self-desi gned: dou ble seque nce of EI VEAEVDELI NSHVERPKS FEKSKSAA DLETDLQL SDMEEFN S EQ ID NO: 132) and change am ino acids R/K to D/ E	GluIleValGluAlaGluValAspGluLeuIleAsnSerLysValGlu LysPheAspSerProGluSerAspSerAspGluAlaAlaAspLeuG luThrAspLeuGluGlnLeuSerAspMetGluGluPheAsnGluIle ValGluAlaGluValAspGluLeuIleAsnSerLysValGluAspPh eAspSerProGluSerAspSerAspGluAlaAlaAspLeuGluThrA spLeuGluGlnLeuSerAspMetGluGluPheAsn	41
DoubleX	RecA-like reco mbination prot ein	GenBank: AYP 69747.1	Escherichia phage vB_Ec _EcoM_DelCa	348-392	90	self-desi gned: a i nner of FEF and d ouble se quence of EIVEAEVDE LNSHVERPK SFEKSKSAA DLETDLQL SDMEEFN SEQ ID NO : 132)	GluIleValGluAlaGluValAspGluLeuIleAsnSerLysValGlu LysPheLysSerProGluSerLysSerLysSerAlaAlaAspLeuG luThrAspLeuGluGlnLeuSerAspMetGluGluPheAsnGluIle ValGluAlaGluValAspGluLeuIleAsnSerLysValGluLysPh eLysSerProGluSerLysSerLysSerAlaAlaAspLeuGluThrA spLeuGluGlnLeuSerAspMetGluGluPheAsn	42

TripleXtail	RecA-like recombination protein	GenBank: AYP 69747.1	Escherichia phage vB_EcoH_DalCa	348-392	135	self-designed: a linker of DVAEF (SEQ ID NO: 133) and d triple sequence of EIVEAE VDELINSKV EKTSPESK SKSAADLET DLEQLSDME EEN (SEQ ID NO: 132)	GluIleValGluAlaGluValAspGluLeuIleAsnSerLysValGluLysPheLysSerProGluSerLysSerLysSerAlaAlaAspLeuGluThrAspLeuGluGlnLeuSerAspPheGluGluPheAsnGluIleValGluAlaGluValAspGluLeuIleAsnSerLysValGluLysPheLysSerProGluSerLysSerLysSerAlaAlaAspLeuGluThrAspLeuGluGlnLeuSerAspMetGluGluPheAsnGluIleValGluAlaGluValAspGluLeuIleAsnSerLysValGluLysPheLysSerProGluSerLysSerLysSerAlaAlaAspLeuGluThrAspLeuGluGlnLeuSerAspMetGluGluPheAsn	43
-------------	---------------------------------	----------------------	---------------------------------	---------	-----	---	---	----

IDR을 포함하는 아미노산 서열

아미노산의 경우 염기를 염기로, 산성을 산성으로, 극성을 극성으로 하는 등의 유사 치환(like-for-like substitution)이 이루어질 수 있다. 비-상동 치환(Non-homologous substitution)도 발생할 수 있으며, 즉, 잔기의 한 클래스에서 다른 클래스로, 또는 대안적으로 오르니틴(ornithine), 디아미노부티르산 오르니틴(diaminobutyric acid ornithine), 노르류신 오르니틴(norleucine ornithine), 피릴알라닌(pyriylalanine), 티에닐알라닌(thienylalanine), 나프틸알라닌(naphthylalanine) 및 페닐글리신(phenylglycine)과 같은 비자연적 아미노산의 포함이 포함된다.

본 명세서에 개시된 특정 IDR 아미노산 서열(표 1 참조)은 크게 4개의 그룹으로 분류될 수 있다. 일부 IDR 시퀀스는 둘 이상의 그룹으로 분류될 수 있다. RGG/RG 그룹에는 FG/YG가 풍부한 IDR 서열이 포함된다. 이 그룹에는 fib, hnrpnA1, DDX, HRP1 및 Sup이 포함된다. Poly Q 그룹은 Q/N이 풍부한 IDR서열이 포함된다. 이 그룹에는 PCF11, Ent-1, HRP1, Sup, His4, His8 및 His10이 포함된다. Poly P 그룹은 P가 풍부한 서열을 포함한다. 이 그룹에는 His4, His9 및 His10이 포함된다. Poly H 그룹은 H가 풍부한 서열을 포함한다. 이 그룹에는 His1-11이 포함된다. IDR 아미노산 서열의 일부 주요 특징은 양이온-파이 상호작용 및 파이-파이 상호작용을 나타내고 아미드 가교 및 염 가교를 형성할 수 있다는 것이다. 바람직한 IDR 아미노산 서열의 주요 특징 및 주요 분자간/분자내 상호작용이 하기 표 2 내지 20에 제시되어 있다.

표 2

IDR	fib
Sequence	PGFSPRGGGFGGRGGFGDRGGRGGRGGGFGRGGGFRGRGR (SEQ ID NO: 1)
Key feature	RGG/RG, FG
Key inter/intra molecular interactions	R-F/H (cation-pi interaction) F/H-F/H (pi-pi interaction) R-D (salt bridge)

표 3

IDR	HNRPNA1
Sequence	ASASSSQRGRSGSGNFGGGGRGGFGGNDNFGRGGNFSGRGGFGGSRGGGGYGGSGDGYNGFGNDGNSNFGGGGSYNDFGNYNQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPQNQGGYGVSSSSSYGSGRRF (SEQ ID NO: 12)
Key feature	RGG/RG, FG/YG
Key inter/intra molecular interactions	R/K-Y/F/H (cation-pi interaction) Y/F/H-Y/F/H (pi-pi interaction) R/K-D (salt bridge) Q/N-Q/N (amide interaction)

표 4

IDR	DDX
Sequence	MGDEDWEAEINPHMSSYVPIFEKDRYSGENGDNFNRTPASSSEMDDGPSRRDHFMKSGFASGRNFGNRDAGECNKRDNTSTMGGFGVGKSGFGRGFSNSRFEDGDSSGFWRSSNDCEDNPTRNRGFSKRGGYRDGNSEASGPYRRGGGFSFRGCRGGFGLGSPNNDLDPDECMQRTGGLFGSRRPVLSGTGNGDTSQSRSGSGSERGGYKGLNEEVITGSGKNSWKSEAEGGESDQTQ (SEQ ID NO: 11)
Key feature	RGG/RG, FG

[0269]

Key inter/intra molecular interactions	R/K-Y/F/H (cation-pi interaction)
	Y/F/H-Y/F/H (pi-pi interaction)
	R/K-D/E (salt bridge)
	Q/N-Q/N (amide interaction)

Ⅴ 5

IDR	HRP1
Sequence	GGNNGGNMNRGGNFGNQGFNQMYQNPMGGYNPMNPQAMTDYYQKMQEYYQQMQ (SEQ ID NO: 9)
Key feature	RGG/FG, Q/N rich, YYQ-K/Q-MQ repeat
Key inter/intra molecular interactions	Q/N-Q/N (amide bridge) R/K-Y/F/H (cation-pi interaction) Y/F/H-Y/F/H (pi-pi interaction) R/K-D/E (salt bridge)

II 6

IDR	Sup
Sequence	MSDSNQGNQNNYQQYSQNGNQQGNRRYQGYQAYNAQAQAPAGGYQNYQGYSGYQQGGYQQYQYNPQGGYQQYNPQGGYQQYNPQGGYQQQFNPOGGGRGNYNFNFNNNLQGYQ (SEQ ID NO: 18)
Key feature	YNPQGGYQQ (SEQ ID NO: 19) repeats, RG, Q/N rich
Key inter/intra molecular interactions	Q/N-Q/N (amide bridge) R/K-Y/F/H (cation-pi interaction) Y/F/H-Y/F/H (pi-pi interaction) R/K-D (salt bridge)

丑 7

[illegible]

丑 8

IDR	PCF11
Sequence	QVQMQLRQVFSQDQQVLQERMRYHELQQQQQQQYHETKDMVGSYTONSNSAIPLFGNNSDTTNQQNS (SEQ ID NO: 7)
Key feature	Poly Q
Key inter/intra molecular interactions	Q/N-Q/N (amide bridge) R/K-Y/F/H (cation-pi interaction) Y/F/H-Y/F/H (pi-pi interaction) R/K-D/E (salt bridge)

⦿ 9

IDR	ENT1
Sequence	QNPTGISYSQPP00000P000POYMONF0000POYAQNF000POYTONY000POYIOPHO (SEQ ID NO: 8)

Key feature	Poly Q, QQQPQY repeat (SEQ ID NO: 128)
Key inter/intra molecular interactions	Q/N-Q/N (amide bridge) Y/F/H-Y/F/H (pi-pi interaction)

표 10

[0274]

IDR	HIS1
Sequence	PQQNALHHHHGNSHHHHHHHHHHHGGQA (SEQ ID NO: 23)
Key feature	Poly H
Key inter/intra molecular interactions	Q/N-Q/N (amide bridge) H-H (pi-pi interaction)

표 11

[0275]

IDR	HIS2
Sequence	AGHHHHHPAHHPPLSQSSGHHHHHHHHHHHQQYGGSG (SEQ ID NO: 24)
Key feature	Poly H
Key inter/intra molecular interactions	Q-Q (amide bridge) Y/H-Y/H (pi-pi interaction)

표 12

[0276]

IDR	HIS3
Sequence	AGHHHHHHHHLPHLPPPHLHHHHHPQHHLHPGSAAAVHPVQQH (SEQ ID NO: 25)
Key feature	Poly H
Key inter/intra molecular interactions	Q-Q (amide bridge) H-H (pi-pi interaction)

표 13

[0277]

IDR	HIS4
Sequence	NHHASHGHNSHHHPQHSHHHHHHHHHHPPPPAPQPPPPQQQ (SEQ ID NO: 26)
Key feature	Poly H, poly Q, poly proline
Key inter/intra molecular interactions	Q/N-Q/N (amide bridge) H-H (pi-pi interaction)

표 14

[0278]

IDR	HIS5
Sequence	SGHHGAHHGAHHAAAAAYEAFRGPGFAGGGGADDMGAGHHGAHHAAHHHHHHHHHHGGAGHGGG AGHH (SEQ ID NO: 27)
Key feature	Poly H
Key inter/intra molecular interactions	R-Y/F/H(cation-pi interaction) Y/F/H-Y/F/H (pi-pi interaction) R-D/E (salt bridge)

표 15

[0279]

IDR	HIS6
Sequence	GGGGDHGGGGGHGAGHHHHHHHHHHH (SEQ ID NO: 28)
Key feature	Poly H
Key inter/intra molecular interactions	H-H (pi-pi interaction)

표 16

[0280]

IDR	HIS7
Sequence	PGSGGPHTSHHHHHHHHHHHHQRH (SEQ ID NO: 29)
Key feature	Poly H
Key inter/intra molecular interactions	H-H (pi-pi interaction) R-H(cation-pi interaction) Q-Q (amide bridge) R-D/E (salt bridge)

표 17

[0281]

IDR	HIS8
Sequence	APSYHHHHHHHHHHHHHHHHHHHQQHQQP (SEQ ID NO: 30)
Key feature	Poly H, poly Q
Key inter/intra molecular interactions	Y/H-Y/H (pi-pi interaction) Q-Q (amide bridge)

표 18

[0282]

IDR	HIS9
Sequence	THPPSHHPHPHHHHHHHHHHH (SEQ ID NO: 31)
Key feature	Poly H, P rich
Key inter/intra molecular interactions	H-H (pi-pi interaction)

표 19

[0283]

IDR	HIS10
Sequence	PPHHHHHHHHHHHHHHRAQPQQSHHPHHHRPQPH (SEQ ID NO: 32)
Key feature	Poly H, P and Q rich
Key inter/intra molecular interactions	H-H (pi-pi interaction) Q-Q (amide bridge) R-H (cation-pi interaction)

표 20

[0284]

IDR	HIS11
Sequence	NHAHHHHHPQHHHHHHQ (SEQ ID NO: 33)
Key feature	Poly H
Key inter/intra molecular interactions	H-H (pi-pi interaction) Q/N-Q/N (amide bridge)

[0285]

IDR-거대분자 또는 IDR-폴리펩타이드의 기능적 특성 (Functional properties of an IDR-macromolecule or IDR-polypeptide)

[0286]

본 명세서에 정의된 IDR-태그된 거대분자 또는 폴리펩타이드, 또는 IDR-거대분자 또는 IDR-폴리펩타이드는 본 발명의 공정에서 사용하기 위한 하나 이상의 기능적 IDR로 구성되거나 이를 포함하는 아미노산 서열을 보유해야 한다. 이러한 IDR 아미노산 서열 또는 도메인이 기능적인지 여부는 본 명세서에 기재된 것과 같은 일상적인 방법에 의해 확립될 수 있다.

[0287]

입자 형성 (Particle formation)

[0288]

본 발명자들은 놀랍게도 IDR-태그된 폴리펩타이드 또는 IDR-폴리펩타이드가 적절한 용액에서 입자를 형성할 수 있음을 발견하였다. 이것은 IDR 아미노산 서열에 의해 매개되는 용액 혼합물 내 유체의 상 분리를 유도하는 액체-액체 탈혼합에 의해 발생하는 것으로 믿어진다.

[0289]

IDR 아미노산 서열에 의해 매개되는 입자의 형성은 다음 실시예에서 추가로 설명된다. 입자는 회전 타원체 모양을 나타내며 "구체(globule)", "구형의 입자(globular foci)" 또는 "입자(particle)"로 설명할 수 있다.

[0290]

본 명세서의 용어 "입자", "구체" 또는 "구형의 입자"는 동의어로 의도되고 상호교환가능하게 사용될 수 있다. 입자의 관찰 및 검출을 허용하는 조건 및 방법은 다음 실시예를 포함하여 본 명세서에 제시되어 있다.

[0291]

본 명세서에 기재된 실시예에서, 입자 형성은 IDR-태그된 폴리펩타이드 및 2가 금속 양이온의 용액만을 포함하는 단순한 시스템에서 발생하는 것으로 관찰되었다. 입자 형성은 또한 RPA 단백질 성분(Gp32) 중 하나가 IDR 태그된 RPA에 필요한 성분을 포함하는 혼합물을 포함하는 더 복잡한 혼합물에서 발생하는 것으로 밝혀졌다. 이러한 상황에서, 반응 성분은 입자와 강하게 공동국소화되는 것으로 나타났으며, 예를 들어 입자는 모든 다른 RPA 반응 단백질 성분을 포함할 뿐만 아니라 그에 부착된 형광 표지된 프로브에 의해 검출된 바와 같이 올리고뉴클레오타이드에서 조밀한 것으로 나타났다.

[0292]

입자의 검출 및 모니터링은 다음 실시예에 기재된 방법 뿐만 아니라 임의의 적합한 방법을 사용하여 수행할 수 있다. 예시적인 방법에는 현미경, 광 산란, 유세포 분석 및 미세유체 방법이 포함된다.

[0293]

입자는 고배율에서 입자를 직접 관찰하기 위한 현미경, 예를 들어 미분 간섭 대조 또는 형광 현미경을 사용하여 검출할 수 있다. 컴퓨터의 도움으로 현미경 이미지를 자동으로 얻고 분석할 수 있다. 또한, 현미경 검사를 통해 입자를 포함하는 혼합물의 적어도 일부를 지속적으로 또는 빈번하게 모니터링할 수 있다.

[0294]

입자는 유세포 분석을 사용하여 감지할 수 있다. 유세포 분석에서 하나 이상의 빛의 빔, 예를 들어 각각의 단일 파장은 유체 역학적-초점을 맞춘 유체 흐름으로 향하게 된다. 빔을 통과하는 부유 입자는 빛을 산란시키고 입자에서 발견되거나 입자에 부착된 형광성 화학물질이 존재할 수 있다. 산란광 및/또는 형광등은 입자 크기 및 형광에 대한 정보를 결정할 수 있는 장치 내의 검출기에 의해 분석된다. 최신 유세포 분석기는 초당 수천 개의 입자를 "실시간"으로 분석할 수 있으며 지정된 특성을 가진 입자를 능동적으로 분리하고 분리할 수 있다.

[0295]

입자는 예를 들어 미국 특허 출원 공개 번호 US2009/0079963 및 US2010/0179068, 국제 특허 출원 공개 번호 WO2009/112594에 개시된 세포측정법, 장치 및 시스템을 사용하여 검출할 수 있다.

[0296]

입자는 미세 유체 방법, 장치 및 시스템을 사용하여 감지할 수 있다. 예를 들어, 입자는 랩-온-어-칩(lab-on-a-chip) 장치 또는 시스템 등을 사용하여 검출될 수 있다 (예를 들어, 미국 특허 출원 공개 번호 US2009/0326903 및 US2009/0297733 참조).

[0297]

입자는 크기가 약 0.5-20 μm 일 수 있으며, 예를 들어 0.5, 1, 1.5, 2, 2.5, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15,

18 및 20 μm 로 이루어진 군에서 선택되는 임의의 두가지 크기의 사이일 수 있다 (예: 크기가 약 1-10 μm).

- [0298] 입자의 농도는 대략 10 내지 5000개 입자/nl일 수 있고, 예를 들어, nl당 10, 20, 50, 100, 200, 500, 1000, 2000, 및 5000개 입자로 이루어진 군에서 선택된 임의의 2개의 입자 수 사이에서 검출될 수 있다 (예: 약 nl당 100-500개의 입자). 입자의 농도는 나노리터당 대략 200-400개의 입자일 수 있다.
- [0299] 이러한 상분리 입자의 크기는 약 0.5 μm 미만일 수 있다. 크기가 약 0.5 μm 미만인 것을 포함하여 상-분리된 입자는 용액의 탁도 변화에 의해 감지될 수 있다. 용액의 탁도 변화는 표준 수단으로 측정할 수 있으며 일반적으로 포르마진 탁도 단위(Formazin Turbidity Unit, FTU) 또는 포르마진 네펠로메트릭 단위(Formazin Nephelometric Unit, FNU)에 따라 정량화할 수 있다. 다른 방법에는 다각광 산란을 포함한 크기 배제 크로마토그래피(size exclusion chromatography, including multiangle light scattering, SEC-MALS)가 있다.
- [0300] IDR 기능의 실험적 결정 (Experimental determination of IDR function)
- [0301] 본 명세서에 정의된 IDR-거대분자 또는 IDR-폴리펩타이드, 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드는 기능적인 IDR (intrinsically disordered region) 아미노산 서열 및/또는 이의 도메인을 보유하는 것으로 결정될 수 있으며, 따라서 예를 들어, 다음에 기재된 상분리 분석법 또는 RPA 분석법 사용에 의해 본 발명의 공정 및 제제에 사용된다.
- [0302] 따라서, IDR-거대분자, IDR-폴리펩타이드 또는 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드는 하나 이상의 기능적인 IDR(intrinsically disordered region); 또는 하나 이상의 기능적 IDR(intrinsically disordered region)을 포함하는 아미노산 서열로 태그된 거대분자 또는 폴리펩타이드를 포함하는 거대분자 또는 폴리펩타이드이다. 모든 경우에 기능적인 IDR(intrinsically disordered region)은 다음에 기술된 상 분리 분석 방법 및/또는 다음에 기술된 RPA 분석 방법에서 기능적인 것으로 결정될 수 있는 것이다.
- [0303] 상분리 분석법(Phase separation assay method)
- [0304] 상 분리 분석 방법은 다음을 포함하는 방법이다:
- [0305] 1. 하나 이상의 IDR(intrinsically disordered region) 아미노산 서열을 폴리펩타이드에 태그하여 IDR-폴리펩타이드 용합 단백질질을 생성하고, 바람직하게는 재조합 파지 vB EcoM NBG1 Gp32 단백질에 태그하여 Gp32-IDR 용합 단백질, 및 정제된 IDR- 폴리펩타이드 용합 단백질질을 생성하는 단계;
- [0306] 2. IDR-폴리펩타이드 용합 단백질을 1000ng/ μl 의 최종 농도로 물 부피에 첨가하는 단계로, 여기서 혼합물의 최종 부피는 50 μl 이고, 바람직하게는 2가 금속 양이온을 2mM의 최종 농도로, 보다 바람직하게는 2가 금속 양이온이 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 인, 첨가하는 단계;
- [0307] 3. 혼합물을 볼텍싱한 후 혼합물을 펄스 원심분리하는 단계;
- [0308] 4. 혼합물 상층액의 샘플 10 μl 를 혈구계(hemocytometer) 슬라이드로 옮기는 단계;
- [0309] 5. x400 배율에서 현미경으로 혈구계 슬라이드를 보는 단계;
- [0310] 6. 혼합물에서 입자 형성을 관찰하는 단계;
- [0311] 7. 혼합물에 입자가 없는 경우, 단계 1 내지 6을 반복하고 혼합물에서 입자의 형성이 관찰될 때까지 2가 금속 양이온의 농도를 점진적으로 증가시키는 단계;
- [0312] 8. x400 배율에서 218 μm x 175 μm 의 배율 영역에서 형성되는 입자의 수를 계산하는 단계; 및
- [0313] 9. (i) 10개 이상의 입자, 바람직하게는 50개 이상의 입자, 더욱 바람직하게는 100개 이상의 입자가 확대 영역에서 계수될 때, 하나 이상의 IDR(intrinsically disordered region)로 구성되거나 이를 포함하는 아미노산 서열이 기능적임을 확인하는 단계; 또는
- [0314] (ii) 2가 금속 양이온의 농도가 100mM 이상으로 증가되고 확대 영역에서 입자가 형성되는 것으로 관찰되지 않는 경우, 또는 확대 영역에서 10개 미만의 입자가 계수되는 경우 하나 이상의 IDR(intrinsically disordered region)로 구성되거나 이를 포함하는 아미노산 서열이 비기능적임을 확인하는 단계.
- [0315] 상기 방법에 있어서, 입자 형성에 대한 2가 금속 양이온 제공의 효과를 조사하는 것이 바람직하다면, 단계 2는 2가 금속 양이온을 임의의 원하는 최종 농도로 첨가하는 단계를 포함할 수 있다. 따라서 다양한 농도의 2가 금속 양이온의 효과를 조사할 수 있다.

- [0316] 상기 방법에 있어서, 입자 형성에 대한 ATP 제공의 효과를 조사하는 것이 바람직하다면, 단계 2는 ATP를 원하는 최종 농도로 첨가하는 단계를 포함할 수 있다. 따라서 다양한 농도의 ATP의 효과를 조사할 수 있다. ATP는 예를 들어 1mM 내지 3.5mM의 농도로, 예를 들면 1mM 내지 2mM의 농도로, 제공될 수 있다.
- [0317] 단계 2는 검출가능한 핵산 분자를 첨가하는 단계를 포함할 수 있고, 단계 8은 검출 수단에 의해 입자의 수를 계수하는 단계를 포함한다. 예를 들어, 단계 2는 FAM(플루오레세인)으로 표지된 서열번호 104에 대해 설정된 핵산 서열을 갖는 프로브를 첨가하는 단계를 포함할 수 있고, 단계 8은 형광에 의해 입자를 검출하는 단계를 포함할 수 있다. 검출 가능한 핵산 분자는 0.5 μ M과 같은 적절한 최종 농도로 추가될 수 있다.
- [0318] 따라서, 상기 분석은 반응 효율, 액체-액체 탈혼합을 일으키는 능력 및 분자가 복수의 상-분리된 수성 구획(입자) 내에서 공동국소화 되도록 하는 능력을 조사하는데 사용될 수 있다.
- [0319] 상기 방법에서, 2가 금속 양이온이 Mg^{2+} 인 경우 양이온의 공급원은 바람직하게는 MgOAc이다. 2가 금속 양이온이 Ca^{2+} 인 경우, 양이온의 공급원은 바람직하게는 $CaCl_2$ 이다. 2가 금속 양이온이 Mn^{2+} 인 경우, 양이온의 공급원은 바람직하게는 $MnCl_2$ 이다.
- [0320] **RPA 분석 방법(RPA assay method)**
- [0321] RPA 분석 방법은 다음을 포함하는 방법이다:
- [0322] 1. 하나 이상의 IDR(intrinsically disordered region) 폴리펩타이드 서열을 Gp32 단백질, 바람직하게는 재조합 파지 vB EcoM NBG1 Gp32 단백질에 태그하여 Gp32-IDR 융합 단백질을 생성하고 정제된 Gp32-IDR 융합 단백질을 제공하는 단계;
- [0323] 2. 다음을 포함하는 반응 혼합물을 생성하는 단계:
- [0324] a. Tris HCl pH 8.3, 25 mM;
- [0325] b. KOAc, 7.5 mM;
- [0326] c. DTT, 1 mM;
- [0327] d. ATP, 2.5 mM;
- [0328] e. phosphocreatine, 20 mM;
- [0329] f. creatine kinase, 1 μ M;
- [0330] g. dNTPs, 1 mM;
- [0331] h. purified Gp32-IDR fusion protein, 20 μ M;
- [0332] i. purified UvsX, 4.8 μ M
- [0333] j. purified UvsY, 8.6 μ M
- [0334] k. *S. aureus* DNA polymerase 1 (Sau), 0.135 μ M
- [0335] l. Exonuclease III, 0.27 μ M
- [0336] m. forward primer, 0.4 μ M
- [0337] n. reverse primer, 0.4 μ M
- [0338] o. probe, 0.12 μ M
- [0339] 3. 반응 혼합물에 33mM MgOAc 및 10카피의 주형 핵산을 첨가하여 리콤비네이즈 폴리머레이즈 증폭 반응을 개시하는 단계;
- [0340] 4. 베어링 볼을 사용하여 자기 혼합(magnetic mixing)하여 형광계(fluorometer)에서 39°C에서 반응 혼합물을 인큐베이션하는 단계;
- [0341] 5. (i) 주형 의존적 방식의 기준선과 비교하여 증폭 산물의 2배 이상의 증가가 형광의 측정가능한 증가에 의해 15분 이내에 검출될 때 하나 이상의 IDR 폴리펩타이드 서열이 기능적임을 확립하는 단계; 또는

- [0342] (ii) 15분 후에 형광에 의해 증폭된 생성물이 검출되지 않을 때 하나 이상의 IDR 폴리펩타이드 서열이 비기능적임을 확인하는 단계.
- [0343] 상기 방법에서, 정방향 프라이머 서열은 CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98); 역방향 프라이머 서열은 CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99); 프로브 서열은 CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM][THF][BHQ-1] ATACAAGGATTGGA (서열번호 100)이며, 여기서 FAM은 플루오레세인(fluorescein)이고, THF는 테트라하이드로푸란(tetrahydrofuran)이고 BHQ는 블랙홀 퀀처(Black Hole Quencher)이다. 주형은 *Listeria monocytogenes* 지노믹 DNA이다.
- [0344] 상기 RPA 분석 방법의 단계 5는 주형 의존적 방식의 기준선과 비교하여 측정 가능한 형광의 증가에 의해 증폭된 산물의 5배 이상 증가가 15분 이내에 검출될 때 하나 이상의 IDR(intrinsically disordered region) 폴리펩타이드 서열이 기능적임을 확인하는 단계를 대안적으로 포함할 수 있으며, 또는 10배 이상의 증가, 또는 20배 이상의 증가, 또는 30배 이상의 증가, 또는 40배 이상의 증가가 검출가능한 경우, 또는 50배 이상 증가, 또는 100배 이상 증가, 또는 150배 이상 증가, 또는 200배 이상 증가, 또는 250배 이상 증가, 또는 300배 이상 증가, 또는 350배 이상 증가, 또는 400배 이상 증가, 또는 450배 이상 증가, 또는 500배 이상 증가, 1000배 이상, 2000배 이상, 3000배 이상, 4000배 이상 또는 5000배 이상 증가될 수 있다. 기준선에 비해 증폭 산물의 증가는 Gp32 단백질이 하나 이상의 IDR(intrinsically disordered region) 폴리펩타이드 서열로 태그되지 않은 경우를 제외하고 동일한 조건에서 반응을 수행하여 얻은 증폭 산물의 양과 비교하여 증폭된 산물의 증가를 의미한다.
- [0345] 상기 방법에서, 반응 효율에 대한 2가 금속 양이온 제공의 효과를 조사하는 것이 바람직하다면, 단계 3은 2가 금속 양이온을 임의의 원하는 최종 농도로 첨가하는 단계를 포함할 수 있다. 따라서 다양한 농도의 2가 금속 양이온의 효과를 조사할 수 있다.
- [0346] 상기 방법에서, 반응 효율에 대한 ATP 제공의 효과를 조사하는 것이 바람직하다면, 단계 2는 ATP를 원하는 최종 농도로 첨가하는 단계를 포함할 수 있다. 따라서 다양한 ATP 농도의 효과를 조사할 수 있다. ATP는 예를 들어 1mM 내지 3.5mM의 농도, 예를 들어 1mM ~ 2mM의 농도로 제공될 수 있다.
- [0347] 거대분자 및 폴리펩타이드에 대한 IDR 아미노산 서열의 태그 (Tagging of IDR amino acid sequences to macromolecule and polypeptides)
- [0348] 본 발명의 방법, 공정 및 제제는 특히 IDR-태그된 거대분자 및 IDR-태그된 폴리펩타이드를 포함하며, 여기서 IDR-태그된 거대분자 또는 IDR-태그된 폴리펩타이드는 하나 이상의 IDR(본 명세서에서 IDR 모이어티로 지칭될 수 있음)으로 구성되거나 이를 포함하는 아미노산 서열로 태그된 관심 거대분자 또는 폴리펩타이드이다.
- [0349] 본 명세서의 용어 "태그" 또는 "태깅"은 가장 넓은 의미로 이해되어야 한다. 이 용어는 IDR 모이어티(moiety), 즉 하나 이상의 기능적인 IDR로 구성되거나 이를 포함하는 아미노산 서열이 관심 대상 거대분자 또는 폴리펩타이드에 임의의 적절한 방법으로 부착되거나, 테더링되거나(tethered), 결합되거나 이와 달리 회합됨을 의미하는 것으로 이해되어야 한다.
- [0350] IDR 모이어티가 관심 폴리펩타이드에 태그되는 가장 바람직한 수단은 제조할 유전자 융합 단백질을 생성하는 것이며, 상기 관심 폴리펩타이드는 뉴클레오타이드 수준에서 유전적으로 조작되어, 전사되고 번역될 때 발현된 단백질이 IDR 모이어티와 함께 관심 폴리펩타이드를 포함하도록 한다.
- [0351] 원하는 경우 링커는 관심 폴리펩타이드와 IDR 모이어티 사이에 위치할 수 있다. 예를 들어, 유연하고 강직하며 절단 가능한 링커는 당업계에 잘 알려져 있으며 융합 단백질의 제조에 널리 사용된다(예: Fusion Protein Linkers: Property, Design and Functionality, Chen, X., *et al.* 2013, Adv. Drug Deliv. Rev., 15, 65(10), pp1357-1369 참조).
- [0352] 유전 공학을 위한 표준 방법은 단백질 발현 및 정제 방법과 같이 당업계에 잘 알려져 있다(예를 들어, Sambrook *et al.*, 2001, Molecular Cloning: a Laboratory Manual, 3rd edition, Cold Spring Harbor Laboratory Press; 및 Current Protocols in Molecular Biology, Greene Publishing and Wiley-Interscience, New York (1995) 참조).
- [0353] IDR 모이어티가 관심 거대분자 또는 폴리펩타이드에 태그될 수 있는 다른 수단은 하나 이상의 공유 결합 또는 친화성 상호작용에 의한 것이다.
- [0354] IDR 모이어티는 관심 폴리펩타이드의 N-말단, 관심 폴리펩타이드의 C-말단과 같은 임의의 적합한 방향으로 폴리펩타이드에 태그될 수 있으며, 또는 관심 폴리펩타이드는 N-말단 및 C-말단 양쪽 모두, 또는 폴리펩타이드의 길이

를 따라 임의의 아미노산 위치에서 IDR 모이어티를 포함할 수 있다.

- [0355] 펩타이드/올리고펩타이드/폴리펩타이드/단백질은 이 기술 분야에서 매우 알려진 방법을 사용하여 다른 펩타이드/올리고펩타이드/폴리펩타이드/단백질을 비롯한 다른 거대분자에 접합/테더링될 수 있다.
- [0356] 그러한 방법 중 하나는 "클릭 화학 반응(click chemistry)"이다. 본 명세서의 용어 "클릭 화학 반응"은 일반적으로 구리 촉매의 존재 하에 1,5-이치환된 1,2,3-트리아졸을 생성하는 알킨과 아지드의 반응을 설명하는 데 사용된다. 클릭 화학은 펩타이드/올리고펩타이드/폴리펩타이드/단백질이 예를 들어 탄수화물, 핵산, 폴리머, 약물, 앵타머, 하이드로겔 등 뿐만 아니라 다른 펩타이드/올리고펩타이드/폴리펩타이드/단백질을 포함하는 다양한 범위의 거대분자와 결합할 수 있게 한다. 이 방법은 "Cu 촉매 알킨 아지드 고리화 첨가(Cu catalyzed alkyne azide cycloaddition, CuAAC)"라고도 한다(예: "Click" reactions: a versatile toolbox for the synthesis of peptide-conjugates. Tang, W. *et al.*, 2014, Chem. Soc. Rev., 43, pp7013-7039 참조).
- [0357] 아민과 반응하는 말레이미드(maleimide), 설프하이드릴 반응성 그룹(sulfhydryl reactive group) 또는 숙신이미딜 에스테르(succinimidyl ester, 종종 NHS 에스테르로 지칭됨)를 함유하는 크로스링커(crosslinker)와 같은 다른 거대분자에 펩타이드/올리고펩타이드/폴리펩타이드/단백질을 접합하기 위해 많은 다른 링커/크로스링커 화학 반응이 이용 가능하다. 예를 들어, 숙신이미드(succinimide)는 단백질 또는 펩타이드 및 플라스틱 재료 사이에 공유 결합을 형성하는 데 사용할 수 있다.
- [0358] 항체-약물 접합체를 생성하기 위한 화학 반응과 같이 폴리펩타이드 및 비-폴리펩타이드 분자 사이의 접합체를 생성하기 위해 일반적으로 사용되는 표준 화학 반응을 사용할 수 있다. 이러한 많은 기술이 이 기술 분야에서 잘 알려져 있다.
- [0359] 친화성-기반 상호 작용도 사용할 수 있다. 예를 들어, 하나 이상의 기능적인 IDR(intrinsically disordered region)으로 구성되거나 이를 포함하는 아미노산 서열은 스트렙타비딘-비오틴, 수용체-리간드 상호작용 등과 같은 친화성 기반 상호작용에 의해 관심 거대분자 또는 폴리펩타이드에 부착/테더링될 수 있다.
- [0360] IDR 아미노산 서열 기능을 위한 다가 금속 양이온 (Multivalent metal cations for IDR amino acid sequence function)
- [0361] 본 명세서에 기술되고 정의된 IDR-거대분자 또는 IDR-폴리펩타이드가 시험관내 생화학 반응에서 사용되는 경우, 시험관내 생화학 반응 완충액은 바람직하게는 다가 금속 양이온, 바람직하게는 2가 금속 양이온을 포함한다.
- [0362] 반응 완충액에 다가/2가 금속 양이온이 존재시 액체-액체 탈혼합의 촉진 및 향상에 도움이 되어, IDR-거대분자 또는 IDR-폴리펩타이드 매개/유발 시험관 내 생화학 반응 환경에서의 상 분리가 유도된다.
- [0363] IDR-거대분자 또는 IDR-폴리펩타이드 매개/유발 시험관내 생화학적 반응 환경에서의 상 분리를 향상시키는 2가 금속 양이온의 기능적 능력은 예를 들어 본 명세서에 개시되고 정의된 기술에 의해 쉽게 확립될 수 있다. 특히, 이러한 기능적 능력은 IDR-의존적 방식으로 시험관내 생화학적 반응 환경에서 본 명세서에 추가로 기재되고 정의된 바와 같이 구형의 입자 또는 입자의 형성을 유도하는 다가/2가 금속 양이온의 능력에 의해 확립될 수 있으며, 예를 들어 본 명세서에 기술된 분석 방법에 의해 결정된 바와 같다.
- [0364] IDR 의존성 액체-액체 탈혼합을 촉진/향상시켜 상 분리를 유도하는 데 2가 금속 양이온을 사용하는 것이 바람직하다. 그러나, 임의의 다가 또는 임의의 2가 금속 양이온의 기능적 등가물이 예상된다. 본 명세서에 기술된 다가/2가 금속 양이온의 기능적 등가물은 IDR-의존성 액체-액체 탈혼합을 촉진하여, 예를 들어 본 명세서에 기술된 분석 방법에 의해 결정된 바와 같이 시험관 내 생화학적 반응 환경에서의 상 분리를 유도하는 데 있어 2가 금속 양이온을 대체할 수 있는 임의의 제제이다.
- [0365] 임의의 적합한 다가/2가 금속 양이온은 단일 제제 또는 제제의 조합으로, 및 선택적으로 에틸렌디아민테트라아세트산(ethylenediaminetetraacetic acid, EDTA), 에틸렌 글리콜-비스(β -아미노에틸에테르)-N, N, N', N'-테트라아세트산(ethylene glycol-bis(β -aminoethyl ether)-N, N, N', N'-tetraacetic acid, EGTA) 또는 니트릴로아세트산(nitriloacetic acid, NTA)의 존재하에 사용될 수 있다.
- [0366] 2가 금속 양이온은 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} , Ni^{2+} 또는 Cu^{2+} 일 수 있다. 이들 양이온 중 임의의 것이 단일 제제로서 사용될 수 있거나 양이온의 임의의 조합이 사용될 수 있다. 바람직하게는 단일 제제로 사용된다. 바람직한 2가 금속 양이온은 Mg^{2+} , Mn^{2+} 및 Ca^{2+} 이다.
- [0367] 사용되는 다가/2가 금속 양이온의 특정 농도뿐만 아니라 시험관 내 생화학적 반응 환경에서 IDR-매개 상 분리를

촉진하는 최적의 결과를 달성하는 특정 다가/2가 금속 양이온은, 관심 거대분자 또는 폴리펩타이드에 태그를 지정하는 데 사용되는 특정 IDR(intrinsically disordered region) 아미노산 서열에 따라 달라질 수 있다. 최적의 다가/2가 금속 양이온 및 최적의 농도는 일상적인 테스트를 사용하여 경험적으로 설정할 수 있다. 본 명세서에 추가로 설명된 상 분리 분석이 이러한 목적을 위해 사용될 수 있다.

[0368] 다가/2가 금속 양이온의 바람직한 농도 범위는 약 300 μM 내지 약 100mM, 약 300 μM 내지 약 50mM, 약 400 μM 내지 약 50mM, 약 400 μM 내지 약 20mM, 약 400 μM 내지 약 30mM, 약 500 μM 내지 약 10mM, 약 500 μM 내지 약 25mM 및 약 1mM 내지 약 35mM이다.

[0369] 시험관 내 생화학 반응 완충액에는 Mg^{2+} 이온이 포함될 수 있다. 바람직한 농도 범위는 약 300 μM 내지 약 100mM, 보다 바람직하게는 약 400 μM 내지 약 50mM, 더욱 더 바람직하게는 약 500 μM 내지 약 40mM, 훨씬 더 바람직하게는 33mM 같은 약 25mM 내지 약 35mM이다. 바람직하게는 완충액은 표시된 농도의 MgOAc 를 포함한다.

[0370] 시험관 내 생화학 반응 완충액에는 Ca^{2+} 이온이 포함될 수 있다. 바람직한 농도 범위는 약 300 μM 내지 약 100mM, 보다 바람직하게는 약 400 μM 내지 약 50mM, 더욱 더 바람직하게는 약 1mM 내지 약 40mM, 더욱 더 바람직하게는 33mM 같은 약 25mM 내지 약 35mM이다. 바람직하게는 완충액은 표시된 농도의 CaCl_2 를 포함한다.

[0371] 시험관 내 생화학 반응 완충액에는 Mn^{2+} 이온이 포함될 수 있다. 바람직한 농도 범위는 약 300 μM 내지 약 50mM, 보다 바람직하게는 약 400 μM 내지 약 50mM, 더욱 더 바람직하게는 약 500 μM 내지 약 40mM, 훨씬 더 바람직하게는 33mM같은 약 25mM 내지 약 35mM이다. 바람직하게는 완충액은 표시된 농도의 MnCl_2 를 포함한다.

[0372] 리코비네이즈 폴리머레이즈 증폭(recombinase polymerase amplification, RPA)

[0373] 리코비네이즈 폴리머레이즈 증폭(RPA)은 핵산의 등온 증폭 방법이다. 일반적으로, RPA의 첫 번째 단계에서, 리코비네이즈 제제는 제1 및 제2 핵산 프라이머 및 리코비네이즈 로딩 단백질과 접촉하여 제1 및 제2 핵단백질(nucleoprotein) 프라이머를 형성한다. 일반적으로, 두 번째 단계에서 제1 및 제2 핵단백질 프라이머는 이중 가닥 주형 핵산과 접촉하여 주형 핵산의 제1 가닥의 제1 부분에서 제1 이중 가닥 구조를 형성하고, 주형 핵산의 제2 가닥의 제2 부분에서 제2 이중 가닥 구조를 형성하여, 제1 핵산 프라이머와 제2 핵산 프라이머의 3' 말단이 주어진 핵산 분자에서 서로를 향하도록 한다. 일반적으로 세 번째 단계에서는 제1 및 제2 핵단백질 프라이머의 3' 말단이 폴리머레이즈에 의해 확장되어 제1 및 제2 이중 가닥 핵산과 핵산의 제1 및 제2 치환된 단일 가닥을 생성한다. 단일 가닥 안정화 제제는 핵산의 제1 및 제2 변위된 단일 가닥을 안정화하기 위해 사용된다. 일반적으로 원하는 증폭 정도에 도달할 때까지 두 번째 및 세 번째 단계를 반복할 수 있다.

[0374] RPA 방법은 예를 들어 미국 특허 번호 7,270,981; 미국 특허 번호 7,399,590; 미국 특허 번호 7,666,598; 미국 특허 번호 7,435,561 및 국제 특허 출원 공개 번호 W02010/141940에 개시되어 있다. 또한, 포괄적인 최근 검토는 다음을 참조한다: Review: a comprehensive summary of a decade development of the recombinase polymerase amplification, Li, J. *et al.*, 2019, *Analyst*, 144, pp31-67).

[0375] 리코비네이즈 제제(Recombinase agent)

[0376] 본 발명을 포함하는 RPA 방법은 리코비네이즈 제제를 사용한다.

[0377] 본 발명의 임의의 하나 이상의 IDR-폴리펩타이드는 임의의 리코비네이즈 제제에 부착/테더링/태그될 수 있다.

[0378] 리코비네이즈 제제는 단일 가닥 핵산, 일반적으로 DNA(ssDNA)인 단일 가닥 핵산을 코팅하여 핵단백질 필라멘트를 형성할 수 있는 분자, 일반적으로 효소이다. 그런 다음 이러한 필라멘트는 이중 가닥 핵산 분자, 일반적으로 서열 상동성/상보성 영역에 대해 DNA(dsDNA)를 "스캔"할 수 있다. 상보적인 서열이 위치할 때, 핵단백질 필라멘트(리코비네이즈 제제를 포함한다) 가닥은 이중 가닥 핵산 분자를 침범하여 D-루프라고 알려진 짧은 하이브리드 및 변위된 가닥 버블을 생성한다.

[0379] 임의의 적합한 리코비네이즈 제제는 본 명세서에 기재된 RPA 방법에 사용될 수 있고 본 명세서에 기재된 임의의 IDR 아미노산 서열로 태그될 수 있다.

[0380] 리코비네이즈 제제는 원핵생물, 진핵생물 또는 바이러스 유기체로부터 유래할 수 있다.

[0381] 리코비네이즈 제제는 RecA, UvsX, RadA, RadB, Rad 51 또는 임의의 이들 단백질의 임의의 기능적 변이체, 유사체, 상동체 또는 유도체일 수 있다.

- [0382] 이들 단백질의 임의의 조합이 사용될 수 있다.
- [0383] 적합한 리코비네이즈 제제는 *E. coli* RecA 단백질, T4 UvsX 단백질, 또는 임의의 문(phyla)으로부터의 임의의 상동성 단백질 또는 단백질 복합체를 포함한다.
- [0384] 진핵생물 RecA 상동체(homologue)는 확인된 이 그룹의 첫 번째 구성원의 이름을 따서 일반적으로 Rad51로 명명된다. 다른 비-상동성 리코비네이즈 제제, 예를 들어 RecT 또는 RecO가 RecA 대신에 사용될 수 있다.
- [0385] 예시적인 리코비네이즈 제제는 RecA 및 UvsX, 및 이들의 단편 또는 돌연변이체 및 이들의 조합을 포함한다. RecA 및 UvsX 단백질은 모든 종에서 얻을 수 있다. RecA 및 UvsX 단편 또는 돌연변이 단백질은 또한 이용 가능한 RecA 및 UvsS 단백질 및 핵산 서열, 및 분자 생물학 기술을 사용하여 생산할 수 있다. 예시적인 UvsX 단백질은 T4, T2, T6, Rb69, Aeh1, KVP40, 아시네토박터 파지 133(*Acinetobacter* phage 133), 아에로모나스 파지 65(*Aeromonas* phage 65), 시아노파지 P-SSM2(cyanophage P-SSM2), 시아노파지 PSSM4(cyanophage PSSM4), 시아노파지 S-PM2(cyanophage S-PM2), Rb14, Rb32, 아에로모나스 파지 25, 비브리오 파지 nt-1, phi-1, Rb16, Rb43, 파지 31, 파지 44RR2.81, Rb49, 파지 Rb3 및 파지 LZ2와 같은 미오바이러스과 파지(mioviridae phage)로부터 유래된 것을 포함한다. 추가의 예시적인 리코비네이즈 제제는 고세균 RADA 및 RADB 단백질 및 진핵생물(예를 들어, 식물, 포유동물 및 진균) Rad51 단백질(예를 들어, RAD51, RAD51B, RAD51C, RAD51D, DMC1, XRCC2, XRCC3, 및 recA)을 포함한다.
- [0386] 리코비네이즈 제제는 바람직하게는 UvsX, T4 UvsX, T6 UvsX, RB18 UvsX, *E. coli* 파지 wV7 UvsX, *Shigella* 파지 CB8 UvsX, *Shigella* 파지 Shf12 UvsX, *E. coli* 파지 AR1 UvsX, 파지 vB_EcoM_G4507 UvsX, *Shigella* 파지 SHFML-11 UvsX, *Escherichia* 파지 vB_EcoM_DalCa UvsX, *E. coli* RecA, *E. coli* RadA, *E. coli* RadB, *E. coli* Rad 51 또는 이들의 임의의 기능적 변이체, 유사체, 상동체 또는 유도체 또는 이들의 임의의 조합이다. 특히 바람직한 리코비네이즈 제제는 *Escherichia* 파지 vB_EcoM_DalCa UvsX이다.
- [0387] 리코비네이즈 제제는 또한 활성을 개선하기 위해 산성 잔기의 C-말단 결실을 포함할 수 있다.
- [0388] 상기 리코비네이즈 제제의 임의의 기능적 변이체, 유사체, 상동체 또는 유도체는 또한 그 자체로 리코비네이즈 제제로서 기능할 수 있고, 이들 기능적 변이체, 유사체, 상동체 또는 유도체는 또한 본 명세서에 기술되고 정의된 공정에서 사용되는 리코비네이즈 제제로서 고려된다.
- [0389] 예를 들어, RecA의 작은 펩타이드는 RecA의 재조합 특성의 일부 측면을 유지하는 것으로 나타났다. 이 펩타이드는 *E. coli* RecA의 잔기 193에서 212를 포함하고 단일 가닥 올리고뉴클레오타이드의 페어링을 매개할 수 있다.
- [0390] 리코비네이즈 제제(예를 들어, UvsX)는 돌연변이 또는 하이브리드 리코비네이즈 제제일 수 있다. UvsX의 돌연변이 형태는 미국 특허 번호 8,071,308에 기술되어 있다. 돌연변이 UvsX는 Rb69 UvsX 아미노산 서열에 적어도 하나의 돌연변이를 포함하는 Rb69 UvsX일 수 있고, 상기 돌연변이는 다음과 같이 이루어진 군에서 선택되는 것이다; (a) 위치 64에서 히스티딘이 아닌 아미노산, 위치 64에서 세린, C-말단에서 하나 이상의 글루탐산 잔기의 첨가, C-말단에서 하나 이상의 아스파르트산 잔기의 첨가, 및 이들의 조합.
- [0391] 돌연변이 UvsX는 T6 UvsX 아미노산 서열에 적어도 하나의 돌연변이를 갖는 T6 UvsX일 수 있으며, 상기 돌연변이는 다음과 같이 이루어진 군에서 선택되는 것이다; (a) 위치 66에서 히스티딘이 아닌 아미노산; (b) 위치 66의 세린; (c) C-말단에서 하나 이상의 글루탐산 잔기의 첨가; (d) C-말단에서 하나 이상의 아스파르트산 잔기의 첨가; 및 (e) 이들의 조합. 하이브리드 리코비네이즈 제제가 사용되는 경우, 하이브리드 단백질은 예를 들어 상이한 UvsX 종으로부터 유래된 아미노산 서열을 포함하는 적어도 하나의 영역을 포함하는 UvsX 단백질일 수 있다. 영역은 예를 들어 UvsX의 DNA-결합 루프-2 영역일 수 있다.
- [0392] 원하는 경우, 리코비네이즈 제제는 온도 민감성(temperature-sensitive)(본 명세서에서 "ts"로 지칭됨) 리코비네이즈 제제일 수 있다. ts 리코비네이즈 제제를 사용하는 경우, RPA 반응은 한 온도(허용 온도)에서 시작되고 다른 온도(비허용 온도)에서 종료될 수 있다. 허용 온도의 조합은 예를 들어 25° C/30° C, 30° C/37° C, 37° C/42° C 등일 수 있다. ts 단백질은 가역적일 수 있다. 가역적 ts 단백질의 활성은 비허용 온도에서 허용 온도로 전환될 때 복원된다.
- [0393] 임의의 리코비네이즈 제제 농도가 사용될 수 있지만, 바람직한 리코비네이즈 농도는, 예를 들어 0.2-12 μM, 6-12 μM, 4-12 μM 및 4-6 μM, 바람직하게는 약 5 μM, 보다 바람직하게는 약 4.8 μM 일 수 있다.
- [0394] 리코비네이즈 제제는 일반적으로 ATP, ATP γ S 또는 기타 뉴클레오시드 삼인산 또는 그 유사체의 존재를 필요로 한다. D-루프 자극 합성 라운드 직후 표적 부위의 재생이 일어날 수 있는 반응 환경에서 리코비네이즈 제제를

사용하는 것이 바람직하다. 리콤비네이즈 분해를 포함하는 완료된 재조합 이벤트는 한쪽 끝에서 다른 쪽 끝으로 진동하는 단면 합성으로 인한 ssDNA의 증폭 지연 또는 매우 비효율적인 선형 증폭을 방지한다.

[0395] IDR(intrinsically disordered regions) 을 포함하는 아미노산 태그 서열로 태그된 예시적인 UvsX 리콤비네이즈 제제는 아래 표 21에 제시되어 있다.

표 21

단백질 명	SEQ ID NO.	아미노산 서열
UvsX (7His)	44	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIPM LNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMVAAYL NKYPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSV EQLKIDMVNQLETIERGEKVIVFIDSIGNMASKKETEDA LNEKSVADMTRAKSLKSLFRIVTPYFSIKNIPCVAVNHTI ETIEMFSKTVMTGGTGVMYSADTVFIIGKRQIKDGSDLQ GYQFVLNVEKSRTVKEKSKFFIDVKFDGGIDPYSGLLD MALELGFVVKPKNGWYAREFLDEETGEMIREEKSWRA KDTNCTTFWGPLFKHQPFDAIKRAYQLGAIDSNEIVEA EVDELINSKVEKFKSPESKSKSAADLETDLQLSDMEEF NEHHHHHHH
UvsX-fib-1	45	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIPM LNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMVAAYL NKYPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSV EQLKIDMVNQLETIERGEKVIVFIDSIGNMASKKETEDA LNEKSVADMTRAKSLKSLFRIVTPYFSIKNIPCVAVNHTI ETIEMFSKTVMTGGTGVMYSADTVFIIGKRQIKDGSDLQ GYQFVLNVEKSRTVKEKSKFFIDVKFDGGIDPYSGLLD MALELGFVVKPKNGWYAREFLDEETGEMIREEKSWRA KDTNCTTFWGPLFKHQPFDAIKRAYQLGAIDSNEIVEA EVDELINSKVEKFKSPESKSKSAADLETDLQLSDMEEF NEPGFSPRGGGFGGRGGFGDRGGRGGRGGFGGGRGRG GGFRGRGRHHHHHHH

[0396]

UvsX-fib-2	46	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIPML NIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMVAAYLNK YPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSVEQLK IDMVNQLETIERGEKVIVFIDSIGNMASKKETEDALNEKSV ADMTRAKSLKSLFRIVTPYFSIKNIPCVAVNHTIETIEMFSK TVMTGGTGVMYSADTVFIIGKRQIKDGSDLQGYQFVLNV EKSRTVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVVK PKNGWYAREFLDEETGEMIREEKSWRAKDTNCTTFWGPL FKHQPFRAIKRAYQLGAIDSNEIVEAEVDELINSKVEKFK SPESKSKSAADLETDLQLSDMEEFNPGFSPRGGGFGGR GGFGDRGGRGGRGGFGGGRGRGGVEHHHHHH
UvsX-fib-3	47	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIPML NIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMVAAYLNK YPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSVEQLK IDMVNQLETIERGEKVIVFIDSIGNMASKKETEDALNEKSV ADMTRAKSLKSLFRIVTPYFSIKNIPCVAVNHTIETIEMFSK TVMTGGTGVMYSADTVFIIGKRQIKDGSDLQGYQFVLNV EKSRTVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVVK PKNGWYAREFLDEETGEMIREEKSWRAKDTNCTTFWGPL FKHQPFRAIKRAYQLGAIDSNEIVEAEVDELINSKVEKFK SPESKSKSAADLETDLQLSDMEEFNPGFSPRGGGFGGR GGFGDRGGRGGRGGVEHHHHHH

[0397]

UvsX-fib-4	48	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRT KIPMLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMS LTMVAAYLNKYPDAICLFYDSEFGITPAYLKSMGV DPERVIHTPIQSVEQLKIDMVNQLETIERGEKVIVFI DSIGNMASKKETEDALNEKSVADMTRAKSLKSLF RIVTPYFSIKNIPCVAVNHTIETIEMFSKTVMTGGTG VMYSADTVFIIGKRQIKDGSDLQGYQFVLNVEKSR TVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVV KPKNGWYAREFLDEETGEMIREEKSWRAKDTNCT TFWGPLFKHQPRDAIKRAYQLGAIDSNEIVEAEV DELINSKVEKFKSPESKSKSAADLETDLQLSDME EFNEPGFSPRGGGFGGRGGVEHHHHHH
UvsX-HNRN PA1	49	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRT KIPMLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMS LTMVAAYLNKYPDAICLFYDSEFGITPAYLKSMGV DPERVIHTPIQSVEQLKIDMVNQLETIERGEKVIVFI DSIGNMASKKETEDALNEKSVADMTRAKSLKSLF RIVTPYFSIKNIPCVAVNHTIETIEMFSKTVMTGGTG VMYSADTVFIIGKRQIKDGSDLQGYQFVLNVEKSR TVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVV KPKNGWYAREFLDEETGEMIREEKSWRAKDTNCT TFWGPLFKHQPRDAIKRAYQLGAIDSNEIVEAEV DELINSKVEKFKSPESKSKSAADLETDLQLSDME EFNEASASSQGRSGSGNFGGGGFGGNDNF GRGGNFSGRGGFGGSRGGGGYGGSGDGYNGFGN DGSNFGGGGSYNDFGNYNQSSNFGPMKGGNFG GRSSGPYGGGGQYFAKPQNQGGYGVSSSSSYGS GRRFHHHHHHH

[0398]

UvsX-DDX	50	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIPML NIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMVAAYLN KYPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSVEQ LKIDMVNQLETIERGEKVIVFIDSIGNMASKKETEDALNE KSVADMTRAKSLKSLFRIVTPYFSIKNIPCVAVNHTIETIE MFSKTVMTGGTGVMYSADTVFIIGKRQIKDGSDDLQGYQ FVLNVEKSRTVKEKSKFFIDVKFDGGIDPYSGLLDMALE LGFVVKPKNGWYAREFLDEETGEMIREEKSWRAKDTNC TTFWGPLFKHQPFDAIKRAYQLGAIDSNEIVEAEVDELI NSKVEKFKSPESKSKSAADLETDLQLSDMEEFNEMGDE DWEAEINPHMSSYVPIFEKDRYSGENGDNFNRTPASSSE MDDGPSRRDHFMKSGFASGRNFGNRDAGECNKRNTST MGGFGVGKSFGNRGFSNSRFEDGDSSGFWRESSNDCED NPTRNRGFSKRGGYRDGNNSEASGPYRRGGRGSFRGCR GGFGLGSPNNDLDPDECMQRTGGLFGSRRPVLSGTGNG DTSQSRSGSGSERGGYKGLNEEVITGSGKNSWKSEAEGG ESSDTQHHHHHHH
UvsX addPolC TD	51	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIPML NIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMVAAYLN KYPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSVEQ LKIDMVNQLETIERGEKVIVFIDSIGNMASKKETEDALNE KSVADMTRAKSLKSLFRIVTPYFSIKNIPCVAVNHTIETIE MFSKTVMTGGTGVMYSADTVFIIGKRQIKDGSDDLQGYQ FVLNVEKSRTVKEKSKFFIDVKFDGGIDPYSGLLDMALE LGFVVKPKNGWYAREFLDEETGEMIREEKSWRAKDTNC TTFWGPLFKHQPFDAIKRAYQLGAIDSNEIVEAEVDELI NSKVEKFKSPESKSKSAADLETDLQLSDMEEFNEHHHH HHHFEFSPFGAYGEAPTSPGFGVSSPGFSPTSPTYSPTSPA YSPTSPSYSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTS SPTSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTS TSPSYSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTS PSYSPTSPNYSPTSPTSPTSPTSPTSPTSPTSPTSPTSPTS LEHHHHHHH

[0399]

UvsX-fusPolII	52	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKI PMLNIAISGAIDGGMQSGLTIFAGPSKSFKNMSLTM VAAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERV IHTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNM ASKKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSI KNIPCVAVNHTIETIEMFSKTVMTGGTGVMYSADTV FIIGKRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFI DVKFDGGIDPYSGLLDMALELGFVVKPKNGWYARE FLDEETGEMIREEKSWRAKDTNCTTFWGPLFKHQPF RDAIKRAYQLGAIDSNEIVEAEVDELINSKVEKFSPE SKSKSAADLETDLQLSDMEEFNETSPSYSPSPSYSP TSPGYSPTSPAYSPTSPTYSPTSPTYSPPHHHHHHH
UvsX-PCF11	53	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKI PMLNIAISGAIDGGMQSGLTIFAGPSKSFKNMSLTM VAAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERV IHTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNM ASKKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSI KNIPCVAVNHTIETIEMFSKTVMTGGTGVMYSADTV FIIGKRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFI DVKFDGGIDPYSGLLDMALELGFVVKPKNGWYARE FLDEETGEMIREEKSWRAKDTNCTTFWGPLFKHQPF RDAIKRAYQLGAIDSNEIVEAEVDELINSKVEKFSPE SKSKSAADLETDLQLSDMEEFNEQVQMQLRQVFSQ DQQVLQERMRYHELQQQQQQYHETKDMVGSYTQ NSNSAIPFLGNNSDTTNQNSHHHHHHH

[0400]

UvsX-Sup	54	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIPM LNIAISGAIDGGMQSGLTIFAGPSKSFKNMSLTMVAAYL NKYPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSV EQLKIDMVNQLETIERGEKVIVFIDSIGNMASKKETEDA LNEKSVADMTRAKSLKSLFRIVTPYFSIKNIPCVAVNHTI ETIEMFSKTVMTGGTGVMYSADTVFIIGKRQIKDGSDLQ GYQFVLNVEKSRTVKEKSKFFIDVKFDGGIDPYSGLLD MALELGFVVKPKNGWYAREFLDEETGEMIREEKSWRA KDTNCTTFWGPLFKHQPFDAIKRAYQLGAIDSNEIVEA EVDELINSKVEKFKSPESKSKSAADLETDLQLSDMEEF NEMSDSNQGNNQQNYQQYSQNGNQQQGNNRYQGYQ AYNAQAQPAGGYYQNYQGYSGYQQGGYQQYQYNPQG GYQQYNPQGGYQQYNPQGGYQQQFNPQGGRGNYKNF NYNNNLQGYQH HHHHHHH
UvsX-DoubleX	55	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIPM LNIAISGAIDGGMQSGLTIFAGPSKSFKNMSLTMVAAYL NKYPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSV EQLKIDMVNQLETIERGEKVIVFIDSIGNMASKKETEDA LNEKSVADMTRAKSLKSLFRIVTPYFSIKNIPCVAVNHTI ETIEMFSKTVMTGGTGVMYSADTVFIIGKRQIKDGSDLQ GYQFVLNVEKSRTVKEKSKFFIDVKFDGGIDPYSGLLD MALELGFVVKPKNGWYAREFLDEETGEMIREEKSWRA KDTNCTTFWGPLFKHQPFDAIKRAYQLGAIDSNEIVEA EVDELINSKVEKFKSPESKSKSAADLETDLQLSDMEEF NE HHHHHHHHFEFEIVEAEVDELINSKVEKFKSPESKSKS AADLETDLQLSDMEEFN EIVEAEVDELINSKVEKFKSP ESKSKSAADLETDLQLSDMEEFN HHHHHHH

[0401]

Superpositive	56	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIPML NIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMVAAYLNK YPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSVEQLK IDMVNQLETIERGEKVIVFIDSIGNMASKKETEDALNEKSV ADMTRAKSLKSLFRIVTPYFSIKNIPCAVNHTIETIEMFSK TVMTGGTGVMYSADTVFIIGKRQIKDGSDDLQGYQFVLNV EKSRTVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVVK PKNGWYAREFLDEETGEMIREEKSWRAKDTNCTTFWGPL FKHQPFDAIKRAYQLGAIDSNRIVKAKVKRLINSKVKKF KSPKSKSKSAAKLKTCLKQLSKMKFNKIVKAKVKKLIN SKVRKFKSPKSKSKSAAKLKTCLKQLSKMKEFNHHHHHH H
Supernegative	57	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIPML NIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMVAAYLNK YPDAICLFYDSEFGITPAYLKSMGVDPERVIHTPIQSVEQLK IDMVNQLETIERGEKVIVFIDSIGNMASKKETEDALNEKSV ADMTRAKSLKSLFRIVTPYFSIKNIPCAVNHTIETIEMFSK TVMTGGTGVMYSADTVFIIGKRQIKDGSDDLQGYQFVLNV EKSRTVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVVK PKNGWYAREFLDEETGEMIREEKSWRAKDTNCTTFWGPL FKHQPFDAIKRAYQLGAIDSNEIVEAEVDELINSEVEEFD SPESDSDEAADLETDLQLSDMEEFNEIVEAEVDELINSEV EDFDSPESDSDEAADLETDLQLSDMEEFNHHHHHHH

[0402]

UvsX His2	58	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKNMSLTMVA AYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVIHT PIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMASK KETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNIP CVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIGK RQIKDGSDDLQGYQFVLNVEKSRTVKEKSKFFIDVKFD GGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEET GEMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNEAGHHHHHPAHHPPLSQSS GHHHHHHHHHHQGYGGSG
UvsX+PCF+His2	59	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKNMSLTMVA AYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVIHT PIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMASK KETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNIP CVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIGK RQIKDGSDDLQGYQFVLNVEKSRTVKEKSKFFIDVKFD GGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEET GEMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNEQVQMQLRQVFSQDQQVL QERMRYHELQQQQQQYHETKDMVGSYTQNSNSAI PLFGNNSDTTNQQNSAGHHHHHPAHHPPLSQSSGHH HHHHHHHHHHQGYGGSG

[0403]

[0404]

리콤비네이즈 로딩 단백질(Recombinase loading protein)

[0405]

본 발명의 방법을 포함하는 RPA 방법은, 리콤비네이즈 로딩 단백질을 추가적으로 포함/사용할 수 있다.

[0406]

임의의 적합한 리콤비네이즈 로딩 단백질은 본원에 기재된 RPA 방법에 사용될 수 있다.

[0407]

본 발명의 임의의 1 이상의 IDR-폴리펩타이드는 임의의 리콤비네이즈 로딩 단백질에 부착/테더링/태깅될 수 있다.

[0408]

리콤비네이즈 로딩 단백질은 원핵생물, 바이러스성 또는 진핵생물 유기체로부터 유래할 수 있다. 예시적인 리콤비네이즈 로딩 단백질은 대장균 rEco (*E. coli* RecO), 대장균 RecR (*E. coli* RecR), UvsY, 및 이들의 돌연변이체 또는 이들의 단편, 또는 이들의 조합을 포함한다.

[0409]

예시적인 UvsY 단백질은 마이오바이러스과 파지(myoviridae phage), 예를 들어 T4, T2, T6, Rb69, Aeh1, KVP40, 아시네토박터 파지 133, 아에로모나스 파지 65, 시아노파지 P-SSM2, 시아노파지 PSSM4, 시아노파지 S-PM2, Rb3에서 유래된 것을 포함한다. 파지 25, 비브리오 파지 nt-1, phi-1, Rb16, Rb43, 파지 31, 파지 44RR2.8t, Rb49, 파지 Rb3 및 파지 LZ2와 같은 것을 포함한다.

[0410]

바람직한 리콤비네이즈 로딩 단백질은 UvsY, 대장균 RecO, 대장균 RecR 또는 이들 단백질의 임의의 기능적 변이체, 유사체, 상동체 또는 유도체이다. 특히 바람직한 UvsY 리콤비네이즈 로딩 단백질은 대장균 파지 STO UvsY이다.

[0411] 이들 단백질의 임의의 조합이 사용될 수 있다.

[0412] 이들 단백질의 바람직한 농도는 0.1-24 μM , 6-24 μM , 4-24 μM 및 4-12 μM , 바람직하게는 약 10 μM , 더욱 바람직하게는 약 8.6 μM 이다. 리콤비네이즈 로딩 단백질은 리콤비네이즈 제제의 마이크로몰 농도의 약 0.5 내지 약 2 배 사이로 존재할 수 있다.

[0413] 표 22에 제시되어 있는 IDR(intrinsically disordered regions)을 포함하는 아미노산 태그 서열로 태깅된 예시적인 UvsY 리콤비네이즈 로딩 단백질은 하기 표 22에 제시되어 있다.

표 22

단백질 명	SEQ ID NO.	아미노산 서열
UvsY[Phage ST0]	60	MHHHHHHHMKLEDLQEELDADLAIDMSKLQYETANNVKLYSKWLRKHSFIRKEMLR IETQKKTAL KARLDYYSGRGDGDDEFSDRYEKSEMKTVLAADKDV LKIETTLQYWGILLEFCSGALDAVKSRSF ALKHIQDMREFEAGQ
UvsY C-Pol CTD	61	MHHHHHHHMKLEDLQEELDADLAIDMSKLQYETANNVKLYSKWLRKHSFIRKEMLR IETQKKTAL KARLDYYSGRGDGDDEFSDRYEKSEMKTVLAADKDV LKIETTLQYWGILLEFCSGALDAVKSRSF ALKHIQDMREFEAGQSGSGSGPTSPSYSPSPSPAYS
UvsY fib[short]	62	MHHHHHHHMKLEDLQEELDADLAIDMSKLQYETANNVKLYSKWLRKHSFIRKEMLR IETQKKTAL KARLDYYSGRGDGDDEFSDRYEKSEMKTVLAADKDV LKIETTLQYWGILLEFCSGALDAVKSRSF ALKHIQDMREFEAGQSGSGSGRGGFGDRGGRGGRGGFGG
UvsY Sup1	63	MHHHHHHHMKLEDLQEELDADLAIDMSKLQYETANNVKLYSKWLRKHSFIRKEMLR IETQKKTAL KARLDYYSGRGDGDDEFSDRYEKSEMKTVLAADKDV LKIETTLQYWGILLEFCSGALDAVKSRSF ALKHIQDMREFEAGQSGSGYNPQGGYQQNNL
UvsY Sup1 HIS2	64	MAGHHHHHPAHHP LSSSGHHHHHHHHHQQYGGSGKLEDLQEELDADLAIDMSKLQYETANNV KLYSKWLRKHSFIRKEMLR IETQKKTAL KARLDYYSGRGDGDDEFSDRYEKSEMKTVLAADKDV L KIETTLQYWGILLEFCSGALDAVKSRSFALKHIQDMREFEAGQSGSGYNPQGGYQQNNLQ

[0415] 단일가닥 안정화제(Single strand stabilizing agent)

[0416] 본 발명의 방법을 포함하는, RPA 방법은 단일 가닥 안정화제를 사용한다.

[0417] 임의의 적합한 단일 가닥 안정화제(단일 가닥 DNA 결합 단백질)가 본원에 기재된 RPA 방법에 사용될 수 있다.

[0418] 본 발명의 임의의 1 이상의 IDR-폴리펩타이드는 임의의 단일 가닥 안정화제에 부착/테더링/태그될 수 있다.

[0419] 단일 가닥 안정화제는 RPA 반응 동안 발생하는 다양한 교환 반응 동안 핵산을 안정화시키는 데 사용된다. 특히 단일 가닥 안정화제는 리콤비네이즈/ssDNA 핵단백질 필라멘트를 안정화하는 데 사용된다.

[0420] 단일 가닥 안정화제는 임의의 종, 예를 들어 원핵생물, 바이러스 또는 진핵생물 종으로부터 유래하거나 수득될 수 있다.

[0421] 단일 가닥 안정화제는 대장균(*E. coli*) 및 T4, T2, T6, Rb69, Aeh1, KVP40, 아시네토박터 파지 133(*Acinetobacter phage 133*), 아에로모나스 파지 65 (*Aeromonas phage 65*), 시아노파지(cyanophage) P-SSM2, 시아노파지 PSSM4, 시아노파지 S-PM2, Rb14, Rb32, 아에로모나스 파지 25, 비브리오 파지(*Vibrio phage*) nt-1, phi-1, Rb16, Rb43, 파지 31, 파지 44RR2.81, Rb49, 파지 Rb3, 및 파지 LZ2와 같은 마यो바이러스과 파지(myoviridae phage)에서 유래된 단일 가닥 DNA 결합 단백질을 포함한다. 단일 가닥 안정화제의 추가 예는 *A. denitrificans* Alide_2047, *Burkholderia thailandensis* BthaB_33951, *Prevotella pollens* HMPREF9144_0124 및 진핵생물 단일 가닥 DNA 결합 단백질 복제 단백질 A를 포함한다.

[0422] 바람직한 단일 가닥 안정화제는 Gp32, 대장균(*E. coli*) SSB 단백질, 파지 T4 Gp32 단백질, 파지 Rb69 Gp32, 파지 vB_EcoM_NBG1 Gp32, 및 이들의 유도체 및 이들의 임의의 조합으로 이루어진 군으로부터 선택된다. 특히 바람직한 단일 가닥 안정화제는 Gp32 및 특히 파아지 vB_EcoM_NBG1 Gp32이다.

[0423] 이들 단백질의 임의의 조합이 사용될 수 있다.

[0424] 단일 가닥 안정화제의 바람직한 농도는 대략 5-30 μM 사이, 예를 들어 대략 8.6 μM , 바람직하게는 대략 15-25 μM 사이, 보다 바람직하게는 대략 20 μM 이다.

[0425] IDR (intrinsically disordered regions)을 포함하는 아미노산 태그 서열로 태깅된 예시적인 Gp32 단일 가닥

안정화제는 하기 표 23에 제시되어 있다.

표 23

단백질 명	SEQ ID NO.	아미노산 서열
Gp32 (7His) from phage vB	65	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKLD ASGNQAVIRFLPAKTDDALPFTILVNHGFKKNGKWYI ETCSSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQKLRK TSYWANILVVKDPQAPDNEGKVFKYRFGKKIWDKINA MIAVDTEMGETPVDVTCPWEGANFVLKVKQVSGFSNY DESKFLNQSAIPNIDDESFKELFEQMVDLSEMTSKDK FKSFEELNTKFNQVLGTAALGGAAAAAASVADKVASD LDDFDKDMEAFSSAKTEDDFMSSSSSDSDLLDLAG LHHHHHHH
Gp32Super +ve	66	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKLD ASGNQAVIRFLPAKTDDALPFTILVNHGFKKNGKWYI ETCSSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQKLRK TSYWANILVVKDPQAPDNEGKVFKYRFGKKIWDKINA MIAVDTEMGETPVDVTCPWEGANFVLKVKQVSGFSNY DESKFLNQSAIPNIDDESFKELFEQMVDLSEMTSKDK FKSFEELNTKFNQVLGTAALGGAAAAAASVAKKVASK LKFKKKMEAFSSAKTKKKFMSSSSSKSKLKKLLAGL MEAFSSAKTKKKFMSSSSSKSKLKKLLAGLHHHHHH H

[0426]

Gp32Super -ve	67	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQL KRKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWD KINAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVS GFSNYDESKFLNQSAIPNIDDESFKELFEQMVDLSE MTSKDKFKSFEELNTKFNQVLGTAALGGAAAAAASV ADDVASDLDDFDDDMEAFSSADTEDDFMSSSSSDDS DLDDLLAGLMEAFSSADTEDDFMSSSSSDDSLDDL AGLHHHHHHH
Gp32-TripleXtail	68	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQL KRKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWD KINAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVS GFSNYDESKFLNQSAIPNIDDESFKELFEQMVDLSE MTSKDKFKSFEELNTKFNQVLGTAALGGAAAAAASV ADDVASEFEIVEAEVDELINSKVEKFKSPESKSKSAAD LETDLEQLSDMEEFNEIVEAEVDELINSKVEKFKSPES KSKSAADLETDLEQLSDMEEFNEIVEAEVDELINSKV EKFKSPESKSKSAADLETDLEQLSDMEEFNHHHHHH

[0427]

Gp32-fib	69	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQK RKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWDKI NAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVSGF SNYDESKFLNQSAIPNIDDESQKELFEQMVDLSEMTS KDKFKSFEELNTKFNQVLGTAALGGAAAAAASVADK VASDLDDFDKDMEAFSSAKTEDDFMSSSSSDSDLDD LLAGLPGFSPRGGGFGGRGGFGDRGGRGGRGGFGGG RGRGGGFRGRGRHHHHHHH
Gp32-PCF11	70	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQK RKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWDKI NAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVSGF SNYDESKFLNQSAIPNIDDESQKELFEQMVDLSEMTS KDKFKSFEELNTKFNQVLGTAALGGAAAAAASVADK VASDLDDFDKDMEAFSSAKTEDDFMSSSSSDSDLDD LLAGLNEQVQMQLRQVFSQDQQVLQERMRYHELQQ QQQQQYHETKDMVGSYTQNSNSAIPLFGNNSDTTNQ QNSHHHHHHH

[0428]

Gp32-Sup	71	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQLK RKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWDKI NAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVSGF SNYDESKFLNQSAIPNIDDESFKELFEQMVDLSEMTS KDKFKSFEELNTKFNQVLGTAALGGAAAAAASVADK VASDLDDFDKDMEAFSSAKTEDDFMSSSSSDSDLDD LLAGLMSDSNQGNNQNNYQQYSQNGNQQQGNNRYQ GYQAYNAQAQPAGGYQNYQGYSGYQQGGYQQYQ YNPQGGYQQYNPQGGYQQYNPQGGYQQQFNPQGGR GNYKNFNYNLQGYQHSHHHHHH
Gp32-Sup1	72	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQLK RKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWDKI NAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVSGF SNYDESKFLNQSAIPNIDDESFKELFEQMVDLSEMTS KDKFKSFEELNTKFNQVLGTAALGGAAAAAASVADK VASDLDDFDKDMEAFSSAKTEDDFMSSSSSDSDLDD LLAGLYNPQGGYQQSHHHHHH

[0429]

Gp32-Sup2	73	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQ MVDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGA AAAAASVADKVASDLDDFDKDMEAFSSAKTEDDFM SSSSSDSDLDLLAGLYNPQGGYQQYNPQGGYQQ HHHHHHH
Gp32-Sup3	74	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQ MVDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGA AAAAASVADKVASDLDDFDKDMEAFSSAKTEDDFM SSSSSDSDLDLLAGLYNPQGGYQQYNPQGGYQQ YNPQGGYQQHHHHHHH

[0430]

Gp32-Sup4	75	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKL KLDASGNGQAVIRFLPAKTDDALPFTILVNHGFKK NGKWYIETCSSTHGDYDSCPVCQYISKNDLYNTN KTEYSQLKRKTSYWANILVVKDPQAPDNEGKVFK YRFGKKIWDKINAMIAVDTEMGETPVDVTCPWEG ANFVLKVKQVSGFSNYDESKFLNQAIPNIDDES QKELFEQMVDLSEMTSKDKFKSFEELNTKFNQVL GTAALGGAAAAAASVADKVASDLDDFDKDMEAF SSAKTEDDFMSSSSSDSDLDLLAGLYNPQGGY QQYNPQGGYQQYNPQGGYQQYNPQGGYQQHHH HHHH
Gp32-DDX	76	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKL KLDASGNGQAVIRFLPAKTDDALPFTILVNHGFKK NGKWYIETCSSTHGDYDSCPVCQYISKNDLYNTN KTEYSQLKRKTSYWANILVVKDPQAPDNEGKVFK YRFGKKIWDKINAMIAVDTEMGETPVDVTCPWEG ANFVLKVKQVSGFSNYDESKFLNQAIPNIDDES QKELFEQMVDLSEMTSKDKFKSFEELNTKFNQVL GTAALGGAAAAAASVADKVASDLDDFDKDMEAF SSAKTEDDFMSSSSSDSDLDLLAGLHHHHHHH FEFMGDEDWEAEINPHMSSYVPIFEKDRYSGENG DNFNRTPASSEMDDGPSRRDHFMKSGFASGRNF GNRDAGECNKRDNTSTMGGFGVGKSFNGRGFSN SRFEDGDSSGFWRESSNDCEDNPTRNRGFSKRG YRDGNNSEASGPYRRGGRGSFRGCRGGFGLGSPN NDLDPDECMQRTGGLFGSRRPVLSGTGNGDTSQS RSGSGSERGGYKGLNEEVITGSGKNSWKSEAEGG ESSDTQLEHHHHHH

[0431]

Gp32 HRP1	79	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQL KRKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWD KINAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVS GFSNYDESKFLNQSAIPNIDDESFKELFEQMVDLSE MTSKDKFKSFEELNTKFNQVLGTAALGGAAAAAASV ADKVASDLDDFDKDMEAFSSAKTEDDFMSSSSSDDS DLDDLLAGLGGNNGGNNMNRGGNFGNQGDFNQM YQNPMMGGINPMMNPQAMTDYYQKMQEYYQQMQ HHHHHHH
Gp32 HRP2	80	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQL KRKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWD KINAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVS GFSNYDESKFLNQSAIPNIDDESFKELFEQMVDLSE MTSKDKFKSFEELNTKFNQVLGTAALGGAAAAAASV ADKVASDLDDFDKDMEAFSSAKTEDDFMSSSSSDDS DLDDLLAGLGGNNGGNNMNRGGNFGNQGDFNQM YQNPMMGGINPMMNPQAMTDYYQKMQEYYQQMQ GGNNGGNNMNRGGNFGNQGDFNQMYQNPMMG YNPMMNPQAMTDYYQKMQEYYQQMQHHHHHHH

[0433]

Gp32 HIS1	81	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKL KLDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKN GKWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKT EYSQLKRKTSYWANILVVKDPQAPDNEGKVFKYR FGKKIWDKINAMIAVDTEMGETPVDVTCPWEGAN FVLKVKQVSGFSNYDESKFLNQSAIPNIDDESQKE LFEQMVDLSEMTSKDKFKSFEELNTKFNQVLGTAA LGGAAAAAASVADKVASDLDDFDKDMEAFSSAKT EDDFMSSSSSDSDLDDLLAGLPQQNALHHHHGN SSHHHHHHHHHHHHHHHGQQA
GP32 HIS2	82	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKL KLDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKN GKWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKT EYSQLKRKTSYWANILVVKDPQAPDNEGKVFKYR FGKKIWDKINAMIAVDTEMGETPVDVTCPWEGAN FVLKVKQVSGFSNYDESKFLNQSAIPNIDDESQKE LFEQMVDLSEMTSKDKFKSFEELNTKFNQVLGTAA LGGAAAAAASVADKVASDLDDFDKDMEAFSSAKT EDDFMSSSSSDSDLDDLLAGLAGHHHHHPHAHH PLSQSSGHHHHHHHHHHHQQGYGGSG

[0434]

GP32 HIS3	83	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLAGHHHHHHHHLPHLPPPHLH HHHHPQHHLHPGSAAAVHPVQQH
GP32 HIS4	84	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLNHHASHGHHNSHHPQH HHHHHHHPPPPAPQPPPPPPQQQ

[0435]

GP32 HIS5	85	MFKRKSTADLAAQMAKLNKNGKFSSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQ KRKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWD KINAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVS GFSNYDESKFLNQSAIPNIDDESFKELFEQMVDLSE MTSKDKFKSFEELNTKFNQVLGTAALGGAAAAAASV ADKVASDLDDFDKDMEAFSSAKTEDDFMSSSSSDDS DLDDLLAGLSGHHGAHHGAHHHPAAAAAYEAFRGP AGGGGADDMGAGHHHGAHHHAHHHHHAHHHHHHH HHHHGGAGHGGGAGHH
Gp32 Mimic1	86	MFKRKSTADLAAQMAKLNKNGKFSSSEDKGEWKLKL DASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKW YIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQ KRKTSYWANILVVKDPQAPDNEGKVFKYRFGKKIWD KINAMIAVDTEMGETPVDVTCPWEGANFVLKVKQVS GFSNYDESKFLNQSAIPNIDDESFKELFEQMVDLSE MTSKDKFKSFEELNTKFNQVLGTAALGGAAAAAASV ADKVASDLDDFDKDMEAFSSAKTEDDFMSSSSSDDS DLDDLLAGLYDPTSPSYDPTSPSYDPTSPSYDPTSPSY DPTSPSYDPTSPSYDPTSPSHHHHHHH

[0436]

Gp32 Mimic2	87	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKLD ASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKWYI ETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQLKRK TSYWANILVVKDPQAPDNEGKVFKYRFGKKIWDKINA MIAVDTEMGETPVDVTCPWEGANFVLKVKQVSGFSNY DESKFLNQSAIPNIDDESFKELFEQMVDLSEMTSKDK FKSFEELNTKFNQVLGTAALGGAAAAAASVADKVASD LDDFDKDMEAFSSAKTEDDFMSSSSSDDSDLDLLAG LYSPTDPSYSPTDPSYSPTDPSYSPTDPSYSPTDPSYSPT DPSYSPTDPSHHHHHHH
Gp32 short pol ctd	88	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLKLD ASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNGKWYI ETCSSTHGDYDSCPVCQYISKNDLYNTNKTEYSQLKRK TSYWANILVVKDPQAPDNEGKVFKYRFGKKIWDKINA MIAVDTEMGETPVDVTCPWEGANFVLKVKQVSGFSNY DESKFLNQSAIPNIDDESFKELFEQMVDLSEMTSKDK FKSFEELNTKFNQVLGTAALGGAAAAAASVADKVASD LDDFDKDMEAFSSAKTEDDFMSSSSSDDSDLDLLAG LTSPSYSPTSPSYSPTSPGYSPTSPAYSPTSPTYSPTYS PHHHHHHHH

[0437]

[0438] 폴리머레이즈 (Polymerase)

[0439] 본 발명의 방법을 포함하는, RPA 방법은 폴리머레이즈를 사용한다.

[0440] 임의의 적합한 폴리머레이즈가 본 발명에 기재된 방법에 사용될 수 있다.

[0441] 본 발명의 임의의 하나 이상의 IDR-폴리펩타이드는 임의의 적합한 폴리머레이즈에 부착/테더링/태깅될 수 있다.

[0442] DNA의 합성이나 증폭을 위해서는 DNA 폴리머레이즈가 바람직하게 사용된다.

[0443] RPA 반응의 한 가지 장점은 사용할 수 있는 폴리머레이즈의 유형에 제한이 없다는 것이다. 예를 들어, 진핵생물, 원핵생물 및 박테리오파지 폴리머레이즈를 사용할 수 있다.

[0444] DNA 폴리머레이즈는 진핵생물 폴리머레이즈일 수 있다. 사용될 수 있는 진핵생물 폴리머레이즈의 예는 pol- α , pol- β , pol- δ , pol- ϵ 또는 이들의 임의의 기능적 변이체(variant), 유사체(analogue), 상동체(homologue) 또는 유도체(derivative) 및 이들의 임의의 조합을 포함한다.

[0445] DNA 폴리머레이즈는 원핵생물 폴리머레이즈일 수 있다. 사용될 수 있는 원핵생물 폴리머레이즈의 예는 대장균 (*E. coli*) DNA, 폴리머레이즈 I 클레나우 단편, 대장균(*E. coli*) DNA 폴리머레이즈 I, 대장균(*E. coli*) DNA 폴리머레이즈 II, *E. coli* DNA 폴리머레이즈 III, *E. coli* DNA 폴리머레이즈 IV, *E. coli* DNA 폴리머레이즈 V, 바실러스 스테아로써모필루스 폴리머레이즈 I 대형 단편, 바실러스 쉼틸리스 폴리머레이즈 I 대형 단편 (Bsu polymerase), 리스테리아 모노사이토제네스 DNA 폴리머레이즈 I, 황색포도상구균 DNA 폴리머레이즈 I (*S.*

aureus DNA polymerase I (Sau polymerase)), 또는 이들의 임의의 기능적 유사체, 상동체, 또는 유도체, 및 이들의 임의의 조합을 포함한다.

- [0446] DNA 폴리머레이즈는 박테리오파지 폴리머레이즈일 수 있다. 본 명세서에 기재된 방법에서 사용될 수 있는 박테리오파지 폴리머레이즈의 예는 Phi-29 DNA 폴리머레이즈, T7 DNA 폴리머레이즈, 박테리오파지 T4 gp43 DNA 폴리머레이즈, 또는 이들의 임의의 기능적 변이체, 유사체, 상동체 또는 유도체 및 이들의 임의의 조합을 포함한다.
- [0447] DNA 폴리머레이즈는 일반적으로 가닥 변위 특성(strand displacing properties)을 포함한다.
- [0448] DNA 폴리머레이즈는 침입하는 가닥의 유리 3'-하이드록실기를 사용하여 새로운 뉴클레오타이드를 통합하여 DNA 합성을 촉매할 수 있다. 다수의 폴리머레이즈는 침입 가닥의 3'-하이드록실기를 사용하여 합성을 촉매하고 동시에 합성이 일어날 때 다른 가닥을 변위(displace)시킬 수 있다. 예를 들어 *E. coli* 폴리머레이즈 II 또는 III를 사용하여 침범된 D-루프를 연장할 수 있다. 또한, *E. coli*에서 SOS-병변-표적 돌연변이에 보통 사용되는 *E. coli* 폴리머레이즈 V를 사용할 수 있다. 이러한 모든 폴리머레이즈는 단일 가닥 DNA 결합 단백질(single stranded DNA binding protein, SSB) 및 기타 구성 요소뿐만 아니라 β -다이머 클램프와의 상호 작용 및 협력을 통해 고도로 진행적으로 만들 수 있다. 원핵생물, 바이러스 및 진핵생물의 다른 폴리머레이즈도 침입 가닥을 연장하는 데 사용할 수 있다.
- [0449] 많은 DNA 폴리머레이즈는 3'-5' 엑소뉴클레이즈 활성을 가지고 있으며 일부는 5'-3' 엑소뉴클레이즈 활성도 가지고 있다. 이는 변위(displacement)가 아니라 앞으로 폴리머레이즈가 이동함에 따라 하나의 DNA 가닥을 점진적으로 분해(digest)시키는 결과를 낳기 때문에 RPA 반응에서 바람직하지 않다.
- [0450] 3'-5' 엑소뉴클레이즈는 잠재적인 장점과 명백한 단점이 있다. 한편으로 3'-5' 엑소뉴클레이즈 활성은 복제 반응(replication reaction)의 신뢰도(fidelity)를 증가시키고 잘못된 염기삽입(misincorporation) 지점에서 중합효소의 지연을 방지할 수도 있다. 고신뢰도 증폭(High fidelity amplification)은 많은 DNA 응용 분야에서 바람직하다. 3'-5' 엑소뉴클레이즈 활성은 더 큰 DNA 단편의 증폭에도 바람직할 수 있고, 여기에서 잘못된 염기삽입(misincorporation)으로 인한 지연이 효과적인 증폭을 억제할 수 있다.
- [0451] 3'-5' 엑소뉴클레이즈 활성의 이러한 분명한 장점에도 불구하고 몇 가지 단점이 있다. 유리 올리고뉴클레오타이드(free oligonucleotide)는 3'-5' 엑소뉴클레이즈를 보유하는 폴리머레이즈가 사용되는 경우 말단 의존적 분해(end-dependant degradation)의 대상이 될 수 있다.
- [0452] 3'-5' 엑소뉴클레이즈 활성이 없는 폴리머레이즈를 활용하여 반응 노이즈(reaction noise)를 줄일 수 있다. 이는 폴리머레이즈의 3'-5' 엑소뉴클레이즈 활성에 의해 단축된 올리고뉴클레오타이드로 인해 프라이밍 오류(mispriming)가 발생할 수 있음을 시사합니다. 결과적으로 3'-5' 엑소뉴클레이즈 편집 활성, 피로인산분해(pyrophosphorylysis) 또는 기타 유사한 편집 활동이 노이즈의 원인이 될 수 있다. 이것은 클레나우 단편(Klenow fragment)과 같은 일부 폴리머레이즈와 함께 상대적으로 협력적인 Gp32 단백질의 포화량을 사용함으로써 크게 억제될 수 있다. 그럼에도 불구하고, 3'-5' 엑소뉴클레이즈 활성이 결여된 본원에 기재된 방법에 사용하기 위한 폴리머레이즈가 제공될 수 있다.
- [0453] DNA 폴리머레이즈는 10,000 유닛/ml 내지 10 유닛/ml, 예컨대 5000 유닛/ml 내지 500 유닛/ml의 농도로 존재할 수 있다.
- [0454] 보조제(Accessory agents)
- [0455] 본 발명을 포함하여, RPA 반응은 추가로 보조제를 이용할 수 있다.
- [0456] 본 발명의 임의의 하나 이상의 IDR-폴리펩타이드는 임의의 보조제에 부착/테더링/태깅될 수 있다.
- [0457] 이러한 보조제는 단일 가닥 결합 단백질(single strand binding protein), 헬리케이스(helicase), 토포아이스머레이즈(topoisomerase), 레졸베이스(resolvase) 및 이들의 임의의 조합을 포함한다. 이러한 제제는 각각 핵산에 대한 풀기(unwinding), 이완(relaxing) 및 분해(resolving) 활성을 가질 수 있다.
- [0458] 보조제는 또한 RuvA, RuvB, RuvC, RecG, PriA, PriB, PriC, DnaT, DnaB, DnaC, DnaG, DnaX 클램프 로더, 폴리머레이즈 코어 복합체, DNA 라이케이스 및 슬라이딩 클램프 및 이들의 임의의 조합을 포함할 수 있다. 슬라이딩 클램프는 *E. coli* β -dimer 슬라이딩 클램프, 진핵생물 PCNA 슬라이딩 클램프, 또는 T4 슬라이딩 클램프 gp45 및 이들의 조합일 수 있다. 보조제는 또한 β -Clamp, DnaX 클램프 로더 및 폴리머레이즈 코어 복합체로 구성된 DNA Polymerase III 홀로엔자임 복합체(DNA Polymerase III holoenzyme complex)를 포함할 수 있다. 이러한 후

자의 보조제는 선행 및 후행 RPA(leading and lagging RPA)의 성능(performance)을 허용한다.

- [0459] RPA 반응은 DNA 합성 개시 후 리컴비네이즈 제제/dsDNA 복합체의 효율적인 분해를 촉진할 수 있는 하나 이상의 추가 효소로 수행될 수 있다. 이러한 효소에는 3'에서 5' 분해를 자극할 수 있는 효소와 5'에서 3' 분해를 지원할 수 있는 효소가 포함된다.
- [0460] 그러한 추가 효소는 RecA를 3'에서 5' 방향으로 변위시킬 수 있는 여러 폴리머레이즈를 포함하며, 리컴비네이즈 제제/dsDNA 복합체의 3'에서 5' 분해를 자극할 수 있다. 이러한 DNA 폴리머레이즈는 *E. coli* PolV 및 다른 종의 상동적 폴리머레이즈(homologous polymerase)를 포함한다. *E. coli* PolV 또는 이의 기능적 변이체, 유사체, 상동체 또는 유도체의 포함은 증폭 효율을 향상시킬 수 있다.
- [0461] 다른 효소들은 dsDNA에서 RecA의 분해를 촉진하는 데 사용할 수 있는 헬리케이스라고 하는 효소의 종류를 포함한다. 이들은 5'에서 3' 및 3'에서 5' 방향으로 분해를 촉진한다. 중간체(intermediate)로부터 RecA의 분해를 자극하기 위한 이상적인 헬리케이스 복합체는 *E. coli* 단백질 RuvA 및 RuvB로 이루어진다. RuvAB 복합체는 가지 이동(branch migration)을 촉진하고 RecA 단백질을 해리하여 RecA가 재활용되도록 한다. RPA 혼합물에 RuvAB를 통합하면 가닥 교환(strand exchange) 및 변위(displacement) 후 dsDNA로부터 RecA의 해리가 촉진되어 동일한 부위에서 복제된 템플릿을 새롭게 합성할 수 있다. 또한 RuvAB 콤플렉스는 RuvC와 함께 작동하여 Holliday 접합을 최종적으로 절단하고 해결할 수 있습니다. RPA 반응 혼합물에 RuvC를 추가하면 침입 사이트에 형성된 Holliday 접합과 같은 복잡한 구조를 해결할 수 있다.
- [0462] 또 다른 효소는 *E. coli* RecG 단백질을 포함한다. RecG는 가지 구조(branch structure)의 분해를 자극할 수 있다.
- [0463] RPA 반응 혼합물에 유용한 다른 효소는 ATP 및 단일 가닥 안정화제의 존재 하에 RecA 핵단백질 필라멘트의 연속 생성을 허용하는 효소이다. 따라서, RecO 및 RecR, 그리고 선택적으로 RecF 단백질이 사용될 수 있다.
- [0464] 엑소뉴클레이즈 효소는 종종 RPA 반응 혼합물에 포함된다. 이들은 절단 가능한 프로브의 효율적인 작동을 위해 포함된다. 일반적으로 사용되는 엑소뉴클레이즈 효소의 한 예는 엑소뉴클레이즈 III입니다. 본 발명의 임의의 IDR 폴리펩타이드는 임의의 엑소뉴클레이즈에 부착/테더링/태깅될 수 있다.
- [0465] 프라이머(Primers)
- [0466] RPA 방법은 폴리머레이즈를 사용하여 주형 핵산 분자(template nucleic acid)의 사본을 생성한다. 따라서, 본 발명의 방법을 포함하는 RPA 방법은 프라이머를 사용하여 폴리머레이즈에 의한 연장을 개시한다.
- [0467] 새로운 합성 부위에 인접한 이중 가닥 핵산의 짧은 스트레치의 말단 당이 유리 3'-하이드록실 모이어티를 필요로 한다는 것은 대부분의 핵산 중합효소의 필요성이다. 이중 가닥 핵산의 이러한 스트레치는 일반적으로 폴리머레이즈 합성 반응의 개시 부위 역할을 하는 프라이머라고 하는 상보적 서열을 일반적으로 갖는 짧은 올리고뉴클레오타이드에 의해 주형에 형성된다. 어떤 경우에는 sulfydryl과 같은 3' 변형을 사용하여 합성 반응을 시작할 수 있다. 주형과 염기쌍을 이루고 폴리머레이즈에 의해 연장되는 프라이머 핵산은 RNA 또는 DNA일 수 있다. 일반적으로 시험관 내 반응의 경우 프라이머는 짧고 종종 화학적으로 합성된 단일 가닥 DNA(또는 변형된 DNA 또는 RNA)로 공급되며 일반적으로 올리고뉴클레오타이드 프라이머라고 한다. 무작위 프라이머도 사용될 수 있지만, 프라이머는 종종 특정 서열을 가진다. 프라이머는 특정 염기쌍 능력 덕분에 상보적 서열을 표적으로 한다. 올리고뉴클레오타이드 프라이머와 표적 핵산 사이의 하이브리드 형성은 일반적으로 자발적 어닐링을 허용하는 염, pH 및 온도 조건 하에서 용액에서 둘을 인큐베이션함으로써 형성된다.
- [0468] RPA에 사용되는 프라이머는 리컴비네이즈 제제의 존재 하에 표적 DNA에 혼성화하기 위한 단일 가닥 영역을 가질 수 있다. 단일 가닥 영역은, 예를 들어, 약 10개 염기, 약 15개 염기, 약 20개 염기, 약 25개 염기, 약 30개 염기, 약 40개 염기, 및 약 50개 염기일 수 있다. 이론상으로는 약 75개 염기, 약 100개 염기, 약 150개 염기 또는 그 이상과 같은 더 긴 영역이 사용될 수 있다. 단일 가닥 영역의 선택은 시작 핵산의 복잡성에 따라 달라지므로, 예를 들어 인간 게놈은 더 긴 프라이머가 필요할 수 있지만 플라스미드는 훨씬 더 짧은 프라이머가 필요할 수 있다.
- [0469] 바람직한 프라이머 길이는 약 30 내지 약 50개 염기이다. 예를 들어, 30 내지 45개 염기, 30 내지 40개 염기, 30 내지 35개 염기, 35 내지 40개 염기, 40 내지 45개 염기, 및 45 내지 50개 염기일 수 있다. 상기 언급된 프라이머 길이가 표시되어 있지만, 최적 프라이머 길이가 30개 염기 미만인 리컴비네이즈 및/또는 단일 가닥 결합 단백질도 가능하고 구상된다.

- [0470] RPA에 사용되는 프라이머는 PNA이지만 DNA가 바람직하고, RNA도 프라이머로 사용하기에 적합하다. 사실, 자연 DNA 복제에서 DNA 폴리머레이즈는 RNA 프라이머로부터 연장하여 게놈 DNA를 연장한다.
- [0471] 프라이머는 표준 기술에 따라 합성될 수 있다. 변형된 염기 및/또는 링커 백본 화학이 일부 경우에 바람직하고 기능적일 수 있다. 또한 올리고뉴클레오타이드는 예를 들어 형광기(fluorescent groups), 소광제(quenchers), 보호(차단) 기(가역적이든 아니든), 자기 태그(magnetic tags), 단백질 등, 5' 또는 3'의 말단에서 다양한 목적을 제공하는 그룹으로 변형될 수 있다. 어떤 경우에는 단일 가닥 올리고뉴클레오타이드가 가닥 침입에 사용될 수 있고, 다른 경우에는 부분적으로 단일 가닥 핵산이 사용될 수 있으며, 침입하는 핵산의 5' 스트레치는 이미 올리고뉴클레오타이드에 혼성화되어 있다.
- [0472] 프라이머는 표적 핵산과 상동성이 아닌 5' 영역을 포함할 수 있다. 프라이머가 표적 핵산에 완전히 상보적이지 않더라도 증폭이 달성될 수 있다는 점에 유의해야 한다. 프라이머는 5' 말단에 추가 서열을 가짐으로써 비-상보적일 수 있다. 이러한 추가 서열은 예를 들어 제한 엔도뉴클레이즈 인식 부위에 대한 서열 또는 시퀀싱 프라이머에 상보적인 서열일 수 있다. 제한 엔도뉴클레이즈 인식 부위는 증폭된 서열의 후속 절단에 유용할 수 있다. 제한 엔도뉴클레이즈 인식 부위 외부에서 핵산을 절단하는 제한 엔도뉴클레이즈의 사용도 고려된다. 시퀀싱 프라이머에 대해 상보적인 서열은 시판되는 프라이머 또는 시판되는 시퀀싱 장치를 사용하여 증폭된 생성물의 신속한 DNA 시퀀싱을 허용할 수 있다.
- [0473] 시험관 내 DNA 합성 반응에 사용하기 위한 올리고뉴클레오타이드를 설계하는 소프트웨어는 특히 PCR에 사용하기 위해 잘 확립되어 있다. RPA 방법에 대한 고려 사항은 유사하며 올리고뉴클레오타이드의 용융 온도(melting temperature) 최적화, 올리고뉴클레오타이드 내 헤어핀 형성 방지 및 주어진 반응에 존재하는 다른 올리고뉴클레오타이드와의 상보성에 대한 선택을 포함한다. 따라서 바람직하지 않은 부반응을 피하기 위해 올리고뉴클레오타이드 프라이머 쌍을 설계하는 것이 중요하다.
- [0474] 올리고뉴클레오타이드 서열 설계를 최적화하는 것 외에도 프라이머 다이머 형성을 줄이거나 제거하기 위한 추가 접근 방식이 있다. 본 명세서의 다른 곳에서 언급했듯이, 반응 노이즈는 3'-5' 엑소뉴클레이즈 활성이 없는 폴리머레이즈를 활용하여 감소될 수 있다. 이는 폴리머레이즈의 3'-5' 엑소뉴클레이즈 활성에 의해 단축된 올리고뉴클레오타이드로 인해 프라이밍 오류가 발생할 수 있음을 시사한다. 결과적으로 3'-5' 엑소뉴클레이즈 편집 활동, 피로인산화분해(pyrophosphorylysis) 또는 기타 유사한 편집 활동이 노이즈의 원인이 될 수 있다. 엑소뉴클레이즈 활성이 없는 폴리머레이즈를 사용하고 피로인산을 피로포스파테이즈로 제거하는 것 외에도 최종 및/또는 끝에서 두 번째 연결에 비-가수분해성 백본이 있는 합성 올리고뉴클레오타이드를 사용하면 반응 노이즈를 줄이는 데 도움이 될 수 있다. 대체 백본은 포스포로티오레이트(phosphorothiorate), 모폴리노(morpholino), 잠금 핵산(locked nucleic acid) 또는 펩타이드 핵산(peptide nucleic acid)과 같은 상당한 범위의 화학 물질 중에서 선택될 수 있다.
- [0475] RPA 반응에 사용되는 시약(Reagents for use in RPA reactions)
- [0476] 본 발명을 포함하는 RPA 방법에 사용하기 위한 시약은 이하에 요약되어 있다.
- [0477] dNTPs
- [0478] dNTP, 예를 들어 dATP, dGTP, dCTP 및 dTTP, 및 이들의 유도체 및 유사체가 RPA 반응에 첨가될 수 있다. 선행 및 후행 가닥 RPA에는 RNA 프라이머 합성을 위해 ATP, GTP, CTP 및 UTP도 포함될 수 있다. 또한, ddNTP(ddATP, ddTTP, ddGTP 및 ddCTP 및 이들의 유도체 및 유사체)를 사용하여 단편 래더(fragment ladder)를 생성할 수 있다.
- [0479] dNTP는 각 NTP 종의 1 mM 내지 200 mM의 농도로 사용될 수 있다.
- [0480] dNTP와 ddNTP의 혼합물은 dNTP 농도(1 mM~200 mM)의 1/100~1/1000의 ddNTP 농도로 사용할 수 있다.
- [0481] RPA는 ATP, 가수분해성 ATP 유사체 또는 다른 뉴클레오타이드 삼인산의 존재 하에 수행될 수 있다. ATP 유사체는 예를 들어 dATP, ddATP, 또는 UTP와 같은 다른 뉴클레오타이드 트리포스페이트 유사체일 수 있다.
- [0482] 환원제(Reducing agents)
- [0483] RPA 반응에 사용될 수 있는 환원제는 DTT를 포함한다. DTT 농도는 1 mM 내지 10 mM, 바람직하게는 1mM일 수 있다.

- [0484] ATP
- [0485] ATP 또는 ATP 유사체가 RPA 반응에 사용될 수 있다.
- [0486] ATP 또는 ATP 유사체는 ATP, ATP- γ -S, ATP- β -S, ddATP 또는 이들의 조합 중 임의의 것일 수 있다. 바람직한 ATP 또는 ATP 유사체 농도는 1 mM 내지 10 mM, 바람직하게는 2.5 mM이다.
- [0487] ATP 재생을 위한 시스템(System for ATP regeneration)
- [0488] RPA 반응의 다른 구성 요소에는 ATP 재생 시스템(즉, ADP를 ATP로 변환하는 시스템)이 포함될 수 있다. 이러한 시스템은 예를 들어 포스포크레아틴 및 크레아틴 카이네이즈일 수 있다.
- [0489] ATP 재생 시스템은 리콤비네이즈가 핵산에 결합될 때 ATP 가수분해 속도가 매우 높기 때문에 지속적인 재조합 반응을 허용한다. 특히 UvsX 단백질은 RecA보다 10~20배 높은 가수분해율을 갖고 있으며, 모노머 당 1분에 200 분자의 ATP를 소모할 수 있다. 여러 시스템이 사용될 수 있다. 크레아틴 카이네이즈/포스포크레아틴 시스템이 바람직하다. UvsX를 사용하면 생성된 AMP가 ATP로 전환될 수 있다. 치킨 마이오카이네이즈가 추가로 사용될 수 있는데, 이는 AMP 분자와 ATP 중 하나를 ADP의 두 분자로 전환시킨다. 그런 다음 ADP는 크레아틴 카이네이즈/포스포크레아틴 시스템을 사용하여 ATP로 전환될 수 있다. ATP의 불량한 재생은 반응 속도를 감소시킬 수 있다.
- [0490] 본 명세서에 기재된 RPA 방법에서, 포스포크레아틴은 바람직하게는 15-25 mM, 더욱 바람직하게는 20 mM의 농도로 사용된다. 크레아틴 카이네이즈는 바람직하게는 약 0.25-5.0 μ M, 더욱 바람직하게는 1 μ M의 농도로 사용된다.
- [0491] 다가 금속 양이온(Multivalent metal cations)
- [0492] RPA 반응에서 완충액은 바람직하게는 다가 금속 양이온을 포함한다. 완충액은 다가 금속 양이온의 기능적 등가물을 함유할 수 있다.
- [0493] RPA 반응에서 완충액은 2가 금속 양이온을 포함하는 것이 보다 바람직하다. 완충액은 2가 금속 양이온의 기능적 등가물을 함유할 수 있다.
- [0494] 임의의 적합한 다가 또는 2가 금속 양이온 또는 그의 기능적 등가물이 단일 제제 또는 제제의 조합으로 사용될 수 있다.
- [0495] 사용되는 다가/2가 금속 양이온의 특정 농도뿐만 아니라 RPA 반응에서 IDR-매개 상 분리를 촉진/향상시키는 최적의 결과를 달성하는 특정 다가 또는 2가 금속 양이온 또는 그의 기능적 등가물은 특정 IDR 폴리펩타이드에 따라 달라질 수 있다. 최적 다가/2가 금속 양이온 또는 그의 기능적 등가물, 및 그의 최적 농도는 RPA 반응 자체 및/또는 여기에 추가로 설명되는 상 분리 분석을 포함하는 일상적인 테스트를 사용하여 경험적으로 확립될 수 있다.
- [0496] 2가 금속 양이온은 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} , Ni^{2+} 또는 Cu^{2+} 일 수 있다. 이들 양이온 중 임의의 것이 단일 제제로서 사용될 수 있거나 양이온의 임의의 조합이 사용될 수 있다. 바람직하게는 단일 제제로 사용된다. 바람직한 2가 금속 양이온은 Mg^{2+} , Mn^{2+} 및 Ca^{2+} 이다. 특히 바람직한 2가 금속 양이온은 Mg^{2+} 이다.
- [0497] 바람직한 농도 범위는 30 내지 40mM, 더욱 바람직하게는 33 내지 39mM이다.
- [0498] 완충액은 Mg^{2+} 이온을 바람직하게는 표시된 농도로 함유할 수 있다. 더욱 바람직하게는 완충액은 표시된 농도의 MgOAc를 함유한다.
- [0499] 완충액은 Ca^{2+} 이온을 바람직하게는 표시된 농도로 함유할 수 있다. 더욱 바람직하게는 완충액은 표시된 농도의 $CaCl_2$ 를 함유한다.
- [0500] 완충액은 Mn^{2+} 이온을 바람직하게는 표시된 농도로 함유할 수 있다. 더욱 바람직하게는 완충액은 표시된 농도의 $MnCl_2$ 를 함유한다.
- [0501] 완충액(Buffers)
- [0502] RPA 반응에서 완충액은 Tris-HCl 완충액, Tris-아세테이트 완충액, 또는 이들의 조합일 수 있다. 완충제는 약 10mM 내지 약 100mM의 농도로 존재할 수 있다. 바람직한 완충액은 약 20mM 내지 약 30mM, 가장 바람직하게는

25mM의 농도로 사용되는 Tris-HCl 완충액이다. 완충된 pH는 6.5 내지 9.0, 바람직하게는 pH 8.3일 수 있다.

[0503] 완충액은 약 5mM 내지 약 50mM, 바람직하게는 약 10mM 내지 약 40mM의 칼륨 아세테이트를 함유할 수 있다.

[0504] 반응 성분(Reaction components)

[0505] RPA 반응을 위한 바람직하지만 비-제한적인 반응 성분 세트는 하기와 같다.

Tris HCl pH 8.3	25 mM
KOAc	7.5 mM
DTT	1 mM
ATP	2.5 mM
Phosphocreatine	20 mM
Creatine kinase	1 μ M
dNTPs	1 mM
Gp32	20 μ M
UvsX	4.8 μ M
UvsY	8.6 μ M
<i>S. aureus</i> DNA polymerase 1 (Sau)	0.135 μ M
Or <i>B. subtilis</i> DNA polymerase 1 (Bsu)	
Exonuclease III	0.27 μ M
MgOAc	33 mM
Forward primer	0.4 μ M
Reverse primer	0.4 μ M
Exo probe	0.12 μ M

[0506]

[0507] RPA 반응 조건 (RPA reaction conditions)

[0508] 본 발명의 반응을 포함하는 RPA 반응은 임의의 적절한 시간 동안 인큐베이션될 수 있다.

[0509] 임의의 RPA 반응은 5분 내지 16시간 또는 그 이상, 예를 들어 15분 내지 3시간 또는 30분 내지 2시간 동안 인큐베이션 될 수 있다.

[0510] 인큐베이션은 원하는 증폭 정도가 달성될 때까지 수행될 수 있다. 원하는 증폭 정도는 10배, 100배, 1000배, 10,000배, 100,000배 또는 1,000,000배 증폭일 수 있다.

[0511] RPA의 한 가지 이점은 PCR과 같은 열 순환이 필요한 기술에 비해 낮은 온도에서 반응을 수행할 수 있다는 것이다. RPA의 또 다른 장점은 온도가 중요하지 않고 정밀한 제어가 선호되지만 절대적으로 필요한 것은 아니라는 점이다. 예를 들어, 현장 환경에서 RPA 반응을 실온 또는 체온에 가까운 온도(35° C ~ 38° C)에서, 예를 들어 신체의 틈새에 샘플을 배치함으로써, 배양하는 것으로 충분하다. 또한, 주형 핵산의 온도 유도 용융 없이 RPA 반응을 수행할 수 있다.

[0512] 따라서 임의의 RPA 반응은 임의의 적절한 온도에서 수행될 수 있다.

[0513] RPA 반응은 45° C 미만에서 수행될 수 있다. RPA 반응은 40° C 미만에서 수행될 수 있다. RPA 반응은 35° C 미만에서 수행될 수 있다. RPA 반응은 30° C 미만에서 수행될 수 있다.

[0514] RPA 반응은 20° C 내지 50° C, 20° C 내지 40° C, 예컨대 20° C 내지 30° C에서 수행될 수 있다.

[0515] RPA 반응 구성의 동결 건조 (Freeze drying of RPA reaction components)

[0516] RPA 반응의 한 가지 이점은 크라우딩제(사용되는 경우) 및 완충제를 제외하고 시약을 사용 전에 동결 건조(즉, 동결건조)할 수 있다는 것이다. 동결 건조된 시약은 활성을 유지하기 위해 냉장 보관할 필요가 없다는 이점을 제공한다. 예를 들어, RPA 시약 튜브는 실온에서 보관할 수 있다. 이 이점은 냉장에 대한 접근이 제한된 현장 조건에서 특히 유용하다.

[0517] RPA 시약은 튜브 바닥, 비드 또는 기타 적합한 유형의 고체 지지체에서 동결 건조될 수 있다. RPA 반응을 수행하기 위해 동결 건조된 시약은 동결 건조된 시약의 구성에 따라 완충 용액 및 크라우딩 에이전트(사용되는 경우) 또는 단순히 완충 용액 또는 물에서 재구성된다. 그런 다음 표적 핵산 또는 표적 핵산을 함유하는 것으로 의심되는 샘플이 추가된다. 재구성 액체는 또한 샘플 핵산을 포함할 수 있다. 재구성된 반응은 일정 기간 동안 배양되고 증폭된 핵산이 있는 경우 검출된다.

[0518] 본원에 기재된 RPA 방법 중 임의의 하나에서, 사용 전에 동결 건조될 수 있는 시약은 적어도, 재조합효소 제제, 재조합효소 로딩 단백질, 단일 가닥 안정화제, DNA 중합효소, dNTP 또는 dNTP 및 ddNTP의 혼합물, 환원제, ATP 또는 ATP 유사체, 프라이머 및 프로브를 포함한다.

[0519] 트레할로스 당과 같은 안정화제는 동결 건조 성능 및 저장 수명을 개선하기 위해 재구성된 반응에서, 예를 들어 20mM 내지 200mM, 가장 최적으로 40mM 내지 80mM으로 동결 건조된 혼합물에 포함될 수 있다. 원하는 경우, 동결 건조된 시약을 사용하기 전에 1일, 1주, 1개월 또는 1년 이상 보관할 수 있다.

[0520] RPA 시약과 같은 생화학적 반응 시약은 크라우딩 에이전트와 함께 동결 건조될 수 있다. 그러나, 동결건조된 혼합물에서 크라우딩 에이전트의 생략을 정당화할 수 있는 복잡한 상호 관련된 문제가 존재할 수 있다. 예를 들어, 사용자는 동결 건조 크라우딩 에이전트의 효과적인 재수화에 어려움을 겪을 수 있거나 더 큰 동결 건조 펠렛의 필요성을 포함하여 다른 해로운 영향을 경험할 수 있다. 따라서, 무엇보다도 펠렛 크기의 감소, 더 짧은 사이클 시간 및 더 쉬운 재수화를 포함하는 동결 건조 물질로부터 크라우딩제의 일부 또는 전부를 배제할 수 있다는 이점이 있을 수 있다. 그러나 이것은 사용되는 경우 생화학적 반응 혼합물이 재수화되고 사용 준비가 된 후 사용 전에 크라우딩제가 새로 추가되어야 한다는 결과적인 단점이 있다. 이는 현장 사용 또는 현장 사용과 같은 특정 상황에서 문제가 될 수 있다. 본 발명의 IDR 기반 시약의 장점은 동결건조된 환경에서 크라우딩 에이전트와 동일한 결점을 나타내지 않을 것으로 예상되며, 따라서 다른 생화학적 반응 성분과 함께 쉽게 동결 건조될 수 있으므로 사용하기 전에 새로운 시약을 추가할 필요가 없다.

[0521] RPA 반응 산물 검출 (Detection of RPA reaction products)

[0522] RPA 반응 산물의 검출은 임의의 적합한 방법을 사용하여 수행될 수 있다.

[0523] 예를 들어, 검출은 아가로스 또는 PAGE 겔상에서 전기영동에 이은 에티듐 브로마이드 염색을 사용하여 수행될 수 있다.

[0524] RPA 반응을 모니터링하는 단계는, 예를 들어, RPA 반응의 일부(fraction)를 제거하는 단계, 삽입되지 않은 일부(fraction)를 분리하는 단계, 및 삽입되지 않은 프라이머를 검출하는 단계를 포함할 수 있다. 삽입되지 않은 프라이머의 크기는 50 bp 이하, 40 bp 이하, 30 bp 이하, 또는 25 bp 이하일 수 있고, 증폭된 산물의 크기는 1 Kb 이상, 2 Kb 이상, 5 Kb 이상, 또는 10 Kb 이상일 수 있으며, 삽입된 프라이머와 삽입되지 않은 프라이머 사이에 큰 크기 차이가 있다. 삽입되지 않은 프라이머의 분리는, 예를 들어, 스핀 컬럼(spin column)과 같은 크기 배제 크로마토그래피(size exclusion chromatography)를 사용하여 신속하게 수행될 수 있다. 프라이머가 표지된(labeled) 경우, 스핀 컬럼 및 측정(예: 형광 또는 방사능)을 포함하는 모니터 절차는 1분 이내에 수행될 수 있다.

[0525] 신장되지 않은(unelongated) 프라이머로부터 신장된(elongated) 프라이머를 분리하기 위한 또 다른 대안은 PAGE의 사용을 포함한다. 예를 들어, 신장된 프라이머는 5분 이내에 겔 전기영동에 의해 신장되지 않은 프라이머로부터 분리될 수 있다.

[0526] 신장된 프라이머를 분리하기 위한 또 다른 대안은 고정된(immobilized) 올리고뉴클레오타이드(oligonucleotide)의 사용을 포함한다. 예를 들어, 증폭된 DNA 서열 내에서 독특하게 발견되는 서열에 상동성인 올리고뉴클레오타이드는 프라이머 신장에 의해 생성된 핵산을 특이적으로 포획(capture)하는 데에 사용될 수 있다. 이러한 포획 올리고뉴클레오타이드(capturing oligonucleotide)는 칩(chip), 또는 기타 서브스트레이트(substrate)에 고정될 수 있다. 포획 올리고뉴클레오타이드에 의한 신장된 올리고뉴클레오타이드의 포획은 RecA 단백질 매개된

방법에 의해, 또는 필요한 경우 전통적인 용액 혼성화(hybridization)에 의해 수행될 수 있다.

- [0527] 형광 프로브의 사용은 RPA 증폭 산물의 검출에 가장 통상적으로(commonly) 사용되고 선호되며 실시간 검출(real-time detection)을 제공하는 이점이 있다.
- [0528] 이러한 프로브는 플루오레세인(fluorescein, FAM)과 같은 형광단(fluorophore) 및 상기 형광단에 근접해 있는 블랙 홀 퀸처(Black Hole Quencher)와 같은 소광제(quencher)로 표지된다. 프로브는 3' 말단에 블로킹 그룹(blocking group)이 있어, 폴리머레이즈(polymerase)에 의한 프로브의 연장(extension)을 방지한다. 프로브가 절단될 때 및 소광제와 형광단이 분리될 때, 형광 신호가 검출됨으로써, 실시간 검출이 가능해진다. 프로브는 무염기 자리(abasic site), 일반적으로(typically) 테트라하이드로퓨란(tetrahydrofuran, THF) 또는 dR그룹을 포함하며, 무염기 자리에서 절단(cleavage)은 일반적으로 *E. coli* 엑소뉴클레아제 III(Exonuclease III)(THF에서 절단함) 또는 *E. coli* fpg (글라이코라이에아제(glycolase)/라이에아제(lyase))(dR 그룹에서 절단함)에 의해 발생한다.
- [0529] RPA 반응 성분을 포함하는 키트 (Kits comprising RPA reaction components)
- [0530] 본 발명은 또한 RPA 반응을 수행하기 위한 키트를 제공한다.
- [0531] 키트는 상술한 농도 중 어느 하나로 RPA에 대해 본원에 기재된 임의의 시약을 포함할 수 있다.
- [0532] 키트는 본원에 기재되고 정의된 IDR-태그된 거대분자(macromolecule) 및/또는 IDR-태그된 폴리펩타이드 중 어느 것을 포함할 수 있다. 바람직하게는 키트는 RPA 리코비네아제(recombinase) 제제, 및/또는 RPA 리코비네아제 로딩(loading) 단백질, 및/또는 폴리머레이즈(polymerase), 및/또는 제1 및 제2 핵산 프라이머(first and second nucleic acid primer), 및/또는 엑소뉴클레아제(exonuclease), 및/또는 버퍼(buffer), 및/또는 다가(multivalent) 금속 이온, 바람직하게는 Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} 또는 Ni^{2+} 와 같은 2가(divalent) 금속 양이온의 공급원으로부터 선택되는 추가적인 RPA 성분을 추가로 포함한다.
- [0533] 키트의 시약은 동결 건조(freeze drying)될 수 있으며, 이 경우 시약은 재용해(reconstitution)될 때 적절한 시약 농도가 달성되도록 임의의 적절한 양으로 제공될 수 있다.
- [0534] 폴리머레이즈(Polymerase)
- [0535] 상술한 바와 같이, 본원에 기재되고 정의된 임의의 IDR 아미노산 서열은 핵산 합성 반응의 수행에 필요한 임의의 단백질 성분에 태그될 수 있다.
- [0536] 본원에 기재되고 정의된 바와 같은 임의의 IDR 아미노산 서열은 핵산 합성 반응의 수행에 필요한 임의의 단백질 성분에 태그될 수 있으며, 폴리머레이즈는 프라이머 핵산 분자를 연장함으로써 새로운 핵산 분자를 합성하는 데 사용된다.
- [0537] 따라서, 임의의 적합한 폴리머레이즈는 본원에 기재되고 정의된 IDR 아미노산 서열로 태그될 수 있다. 폴리머레이즈는 프라이머 핵산 분자를 연장하여 새로운 핵산 분자를 합성하는 데 사용되는 임의의 반응과 양립 가능하고(compatible) 이러한 반응에서 사용될 수 있는 것일 수 있다.
- [0538] 폴리머레이즈는 임의의 핵산 증폭 반응과 양립 가능하고 이러한 반응에서 사용될 수 있는 것일 수 있다. 핵산 증폭 반응은 열 순환을 포함하는 반응일 수 있다. 핵산 증폭 반응은 등온 증폭 반응(isothermal amplification reaction)일 수 있다. 핵산 증폭 반응은 PCR(polymerase chain reaction), PSR(polymerase spiral reaction), LAMP(loop mediated isothermal amplification), NASBA(nucleic acid sequence based amplification), 3SR(self-sustained sequence replication), RCA(rolling circle amplification), SDA(strand displacement amplification), MDA(multiple displacement amplification), LCR(ligase chain reaction), HDA(helicase dependant amplification), RAM(ramification amplification method), RPA(recombinase polymerase amplification), TMA(transcription-mediated amplification) 또는 NEAR(nicking enzyme amplification reaction)일 수 있다.
- [0539] 서열 태그 (Sequence tags)
- [0540] RPA 반응에 관여하는 것을 포함하는, 본원에 기재된 바와 같은 생화학적 반응에 관여하는 임의의 IDR-거대분자 또는 임의의 IDR-폴리펩타이드는 하나 이상의 서열 태그를 포함할 수 있다. 사용되는 경우, 임의의 이러한 서열 태그는 바람직하게는 본원에 기재된 바와 같은 융합 단백질로서 폴리펩타이드에 부착된다. 서열 태그 및 서열

태그를 폴리펩타이드에 부착하기 위한 수단은 당업계에 잘 공지되어 있다.

- [0541] 서열 태그는 짧은 아미노산 서열 또는 단백질을 포함하는 더 큰 폴리펩타이드일 수 있다.
- [0542] 서열 태그는 폴리펩타이드의 C-말단(C-terminus), 폴리펩타이드 태그의 N-말단(N-terminus), 폴리펩타이드의 C- 및 N-말단 모두에, 또는 임의의 조합으로 폴리펩타이드의 길이를 따른 임의의 아미노산 위치에 부착될 수 있다.
- [0543] 적합한 아미노산 서열 태그의 비-제한적인 예에는 6-히스티딘(6X-His; HHHHHH; 서열번호 89), c-myc 에피토프(EQKLISEEDL; 서열번호 90), FLAG® 옥타펩타이드(DYKDDDDK; 서열번호 91), 단백질 C(EDQVDPRLIDGK; 서열번호 92), Tag-100(EETARFQPGYRS; 서열번호 93), V5 에피토프(GKPIPNPLLGLDST; 서열번호 94), VSV-G(YTDIEMNRLGK; 서열번호 95), Xpress(DLYDDDDK; 서열번호 96), 및 헤마글루티닌(YPY-DVPDYA; 서열번호 97)이 포함된다.
- [0544] 적합한 단백질 태그의 비-제한적 예에는 β -갈락토시데이즈(β -galactosidase), 티오레독신(thioredoxin), His-패치 티오레독신(His-patch thioredoxin), IgG-결합 도메인(IgG-binding domain), 인테인키티 결합 도메인(inteinichitin binding domain), T7 유전자 10(T7 gene 10), 글루타티온-S-트랜스퍼레이즈(glutathione-S-transferase, GST), 녹색 형광 단백질(green fluorescent protein, GFP) 및 말토스 결합 단백질(maltose binding protein, MBP)이 포함된다.
- [0545] 서열 태그 및 단백질 태그는, 예를 들어, 정제 및/또는 확인 목적을 위해 상호 교환적으로 사용될 수 있음을 당업자는 이해할 것이다.
- [0546] **고상 생화학적 반응 (Solid phase biochemical reactions)**
- [0547] 본 발명에 따른 공정(process)에서 수행되는 생화학적 반응은 고상(solid phase) 또는 가역적 고상(reversible solid phase) 기술을 사용하여 수행될 수 있다. 본원에 기재된 공정, 용도 및 방법을 수행하기에 적합한 고상 반응 시스템은 표면(surface)을 포함할 수 있다. 임의의 적합한 표면이 사용될 수 있다.
- [0548] 본원에 기재된 데이터는 생화학적 반응이 크라우딩 제제(crowding agent)의 부재에서 고상 기술을 사용하여 본 발명에 따른 IDR-기반 시약으로 수행될 수 있음을 입증한다. 하나의 특정 예는 프라이머가 고체 표면에 부착된 핵산의 RPA(recombinase polymerase amplification)이다. 고상 방법을 사용하여 처리할 수 있는 임의의 적합한 생화학적 반응은 본원에 기재되고 정의된 임의의 IDR-기반 시약을 포함하는 본 발명에 따른 공정을 사용하여 이러한 방법을 사용하여 수행될 수 있다.
- [0549] 이러한 다양한 고상 기술은 당업계에 공지되어 있고 사용될 수 있다.
- [0550] 핵산 증폭 프라이머, 펩타이드, 합텐(hapten), 호르몬, 약물 등을 포함하는 폴리뉴클레오타이드와 같은 거대분자가 표면에 고정될 수 있다.
- [0551] 본원에 기재되고 정의된 IDR-기반 시약을 포함하는, 생화학적 반응의 임의의 적합한 거대분자 성분이 표면에 고정될 수 있다.
- [0552] 증폭 반응에 사용하기 위한, 폴리뉴클레오타이드와 같은 거대분자, 예컨대, 프라이머는 표면에 직접적으로 또는 간접적으로 고정될 수 있다. 예를 들어, 화학적 결합에 의해 표면에 직접적으로 부착될 수 있다. 중간 표면(intermediate surface)을 통해 표면에 간접적으로 부착될 수 있다.
- [0553] 표면은, 예를 들어, 유리와 같은 평면 표면(planar surface), 겔-기반 물질(gel-based material), 또는 비드(bead) 또는 작용기화된 양자점(functionalised quantum dot)과 같은 마이크로입자(microparticle)의 표면일 수 있다. 표면을 구성하는 물질 자체가 서브스트레이트(substrate)에 결합될 수 있다. 서브스트레이트(substrate)는 유리, 플라스틱 또는 폴리머 물질(polymeric material)과 같은 임의의 적합한 물질을 포함할 수 있다.
- [0554] 본 발명의 공정에 따른 생화학적 반응에 관여하는 거대분자는, 예를 들어, 폴리아크릴아미드(polyacrylamide) 또는 하이드로겔(hydrogel)과 같은, 겔-기반 물질(gel-based material)에 고정될 수 있고, 겔-기반 물질은 그 자체가 유리 또는 플라스틱 또는 폴리머 물질과 같은 지지 서브스트레이트(supporting substrate)에 결합된다.
- [0555] 미리-형성된(Pre-formed) 폴리뉴클레오타이드는, 예를 들어, 핵산 마이크로어레이를 생성하기 위해 통상적으로(commonly) 사용되는 방법에 의해 표면에 고정될 수 있다. 예를 들어, 폴리뉴클레오타이드는 합성된 다음 표면상으로, 일반적으로는(typically) 평면 표면(planar surface)상으로 스폿팅(spotting)되거나 인쇄될 수 있다. 폴리뉴클레오타이드는 접촉 인쇄 기술(contact printing technique)을 사용하여 표면상으로 침착(deposition)

될 수 있다. 예를 들어, 중실(solid) 또는 중공(hollow) 팁(tip) 또는 핀(pin)은 미리-형성된 폴리뉴클레오타이드를 포함하는 용액에 담길 수 있고 표면에 접촉될 수 있다. 대안적으로, 폴리뉴클레오타이드는 마이크로-스탬프(micro-stamp)상에 흡착된 다음 물리적 접촉에 의해 표면으로 옮겨질 수 있다. 비-접촉 인쇄 기술에는 열 인쇄(thermic printing) 또는 압전 인쇄(piezoelectric printing)가 포함되며, 미리-형성된 폴리뉴클레오타이드를 포함하는 서브-나노리터(sub-nanolitre) 크기의 미세액적(microdroplet)은 잉크젯(inkjet) 및 버블젯(bubblejet) 인쇄에 사용되는 것과 유사한 방법을 사용하여 인쇄 팁(tip)에서 분출될 수 있다.

[0556] 폴리뉴클레오타이드는 핵산 마이크로어레이를 생성하는 데 사용되는 소위 "온-칩(on-chip)" 방법을 사용하는 것과 같이 표면상에서 직접적으로 합성될 수 있다. 폴리뉴클레오타이드를 생성하기 위한 온-칩 기술에는 보호된 뉴클레오타이드를 선택적으로 활성화하기 위해 포토리소그래피 마스크를 통해 지시된 UV 광의 사용을 포함하는 포토리소그래피(photolithography)가 포함되어 후속적으로 새로운 보호된 뉴클레오타이드를 통합할 수 있다. UV-매개 탈보호(UV-mediated deprotection)의 사이클 및 미리-결정된 뉴클레오타이드의 커플링은 원하는 서열을 갖는 폴리뉴클레오타이드의 현장(in situ) 생성을 가능하게 한다. 포토리소그래피 마스크 사용에 대한 대안으로, 잉크젯 인쇄 기술을 사용하여 핵염기(nucleobase)를 순차적으로 증착(deposition)하고 원하는 서열을 갖는 올리고뉴클레오타이드를 생성하기 위해 커플링, 산화 및 탈보호의 사이클을 사용함으로써 폴리뉴클레오타이드가 표면상에 생성될 수 있다(검토를 위해 Kosuri and Church, Nature Methods, 2014, 11, 499-507 참조).

[0557] 폴리뉴클레오타이드, 펩타이드, 합텐, 호르몬, 약물 등을 포함하는 거대분자의 부착을 위한 표면은 임의의 적합한 물질로 제조될 수 있다. 일반적으로 표면은 실리콘(silicon), 유리 또는 폴리스티렌과 같은 임의의 적합한 폴리머 물질을 포함할 수 있다. 표면은 폴리아크릴아미드 표면 또는 하이드로겔 표면과 같은, 겔 표면을 포함할 수 있다. 겔 표면은 차례로 고체 지지체(support) 또는 서브스트레이트(substrate)에 연결되거나 결합될 수 있으며, 상기 지지체 또는 서브스트레이트는 실리콘, 유리 또는 임의의 적합한 폴리머 물질과 같은 임의의 적합한 물질을 포함할 수 있다. 표면은 폴리스티렌 물질에 연결된 하이드로겔 물질을 포함할 수 있다.

[0558] 표면은 종종 마이크로스피어(microsphere) 또는 마이크로비드(microbead), 또는 간단히 비드(bead)로 지칭되는 마이크로입자(microparticle)의 표면일 수 있다.

[0559] 표면은 마이크로비드 형태의 폴리스티렌 물질에 연결된 하이드로겔 물질을 포함할 수 있다.

[0560] 마이크로비드와 같은 표면에 폴리뉴클레오타이드와 같은 거대분자를 고정하기 위해 다양한 표면 부착 방법 및 화학이 이용 가능하다. 표면은 부착을 용이하게 하기 위해 작용기화되거나(functionalised) 유도체화될(derivatized) 수 있다. 이러한 작용기화(functionalisation)는 당업계에 공지되어 있다. 예를 들어, 표면은 폴리히스티딘-태그(헥사 히스티딘-태그, 6xHis-태그, His6 태그 또는 His-태그®), Ni-NTA, 스트렙타비딘(streptavidin), 비오틴(biotin), 올리고뉴클레오타이드, 폴리뉴클레오타이드(예컨대, DNA, RNA, PNA, GNA, TNA 또는 LNA), 카르복실기, 4차 아민(quaternary amine)기, 티올(thiol)기, 아지드(azide)기, 알킨(alkyne)기, DIBO, 지질(lipid), FLAG-태그(FLAG 옥타펩타이드), 폴리뉴클레오타이드 결합 단백질(polynucleotide binding protein), 펩타이드, 단백질, 항체 또는 항체 단편(antibody fragment)으로 작용기화될 수 있다. 표면은 고정될 거대분자 또는 고정될 거대분자에 부착된 또 다른 모이어티(moiety)에 특이적으로 결합하는 분자 또는 기(group)로 작용기화될 수 있다. 표면에 거대분자의 공유결합 고정(Covalent immobilization)이 통상적으로 사용된다. 오직 예로서, 카르복실레이트-개질된 폴리스티렌 라텍스(carboxylate-modified polystyrene latex) 표면은, 예를 들어 EDAC-매개 커플링(EDAC-mediated coupling)을 통해, 예를 들어, 아민-말단 단백질(amine-terminated protein), DNA, 또는 다른 분자의 공유결합 부착(covalent attachment)에 적합하다. 다른 기술이 사용 가능하다. 일반적으로 거대분자는 화학적으로 부착되지만, 친화성 상호작용(affinity interaction)과 같은 간접적인 수단에 의해 표면에 부착될 수도 있다. 예를 들어, 고정될 거대분자는 비오틴으로 작용기화될 수 있고 아비딘 또는 스트렙타비딘으로 코팅된 표면에 결합될 수 있으며, 또는 그 반대의 경우도 가능하다.

[0561] 본원에 기재되고 정의된 임의의 공정, 용도 및 방법에서 거대분자는 하나 이상의 공유 결합을 통해 표면에 부착될 수 있다. 하나 이상의 공유 결합은 표면상의 작용기(functional group)와 거대분자(macromolecule)상의 작용기 사이에 형성될 수 있다. 거대분자상의 작용기는, 예를 들어, 아민기, 티올기, 티오포스페이트(thiophosphate)기 또는 티오아미드(thioamide)기일 수 있다. 표면상의 작용기는, 예를 들어, 브로모아세틸(bromoacetyl)기일 수 있고, 선택적으로 브로모아세틸기는 BRAPA (N-(5-bromoacetamidylpentyl) acrylamide)를 사용하여 유도된(derived) 폴리아크릴아미드 표면에 제공된다.

[0562] 본원에 기재되고 정의된 임의의 공정, 용도 및 방법에서 거대분자는 링커(linker)를 통해 직접적으로 또는 간접

적으로 표면에 부착될 수 있다. 본질적으로 생체적합성인 임의의 적합한 링커가 사용될 수 있다.

[0563] 링커는 선형(linear) 링커 또는 분지형(branched) 링커일 수 있다.

[0564] 링커는 탄화수소 사슬(hydrocarbon chain)을 포함할 수 있다. 탄화수소 사슬은 2개 내지 약 2000개 이상의 탄소 원자를 포함할 수 있다. 탄화수소 사슬은 알킬렌(alkylene)기, 예를 들어, C2 내지 약 2000개 이상의 알킬렌기를 포함할 수 있다. 탄화수소 사슬은 $-(CH_2)_n-$ 의 일반식을 가질 수 있으며, 상기 n은 2 내지 약 2000 이상이다. 탄화수소 사슬은 선택적으로 하나 이상의 에스테르(ester)기(즉, $-C(O)-O-$) 또는 하나 이상의 아미드(amide)기(즉, $-C(O)-N(H)-$)에 의해 중단될 수 있다.

[0565] 임의의 링커는 폴리아크릴아미드 (polyacrylamide), 폴리(2-하이드록시에틸 메타크릴레이트) (poly(2-hydroxyethyl methacrylate)), PMOXA (Poly-2-methyl-2-oxazoline), 쯔비터이온성 폴리머 (zwitterionic polymers), 예를 들어, PCBMA (poly(carboxybetaine methacrylate)), PSBMA (poly[N -(3-sulfopropyl)- N-methacryloxyethyl- N, N dimethyl ammonium betaine]), 글리코폴리머 (glycopolymer) 및 폴리펩타이드를 포함하는 그룹으로부터 선택될 수 있다.

[0566] 링커는 $-(CH_2-CH_2-O)_n-PO_2^--O]_m-$ 의 일반식을 갖는 올리고에틸렌 글리콜-포스페이트(oligoethylene glycol-phosphate) 단위를 포함할 수 있으며, 상기 n은 1 내지 약 600 이상이고 상기 m은 1 내지 200 이상일 수 있다.

[0567] 상술한 임의의 링커는 링커의 한쪽 말단에서 본원에 기재된 거대분자에 부착될 수 있고, 링커의 다른 말단에서 제1 작용기(first functional group)에 부착될 수 있으며, 상기 제1 작용기는 표면에 공유결합 부착을 제공할 수 있다. 제1 작용기는, 예를 들어, 본원에 추가로 기재된 바와 같은 아민기, 티올기, 티오포스페이트기 또는 티오아미드기일 수 있다. 표면은 추가 작용기로 작용기화되어 제1 작용기와 공유 결합을 제공할 수 있다. 추가 작용기는, 예를 들어, 본원에 추가로 기재된 바와 같은 2-브로모아세트아미도(2-bromoacetamido)기일 수 있다. 선택적으로 브로모아세틸(bromoacetyl)기는 BRAPA (N-(5-bromoacetamidylpentyl) acrylamide)를 사용하여 유도된 폴리아크릴아미드 표면에 제공된다. 표면상의 추가 작용기는 브로모아세틸기일 수 있으며, 선택적으로 브로모아세틸기는 BRAPA (N-(5-bromoacetamidylpentyl) acrylamide)를 사용하여 유도된 폴리아크릴아미드 표면에 제공되고 제1 작용기는, 예를 들어, 적절하게는, 아민기, 티올기, 티오포스페이트기 또는 티오아미드기일 수 있다. 폴리뉴클레오타이드가 부착되는 표면은 겔을 포함할 수 있다. 표면은 약 2% 폴리아크릴아미드와 같은 폴리아크릴아미드 표면을 포함할 수 있으며, 바람직하게는 폴리아크릴아미드 표면은 유리 와 같은 고체 지지체에 연결된다.

[0568] 가역적 고정을 용이하게 하는 마이크로입자 및 비드가 사용될 수 있다. SPRI (solid phase reversible immobilization) 방법 또는 변형된 방법이 당업계에 공지되어 있으며 이는 사용될 수 있다(예를 들어, DeAngelis M. M. et al. (1995) Solid-Phase Reversible Immobilization for Isolation of PCR Products, Nucleic Acids Research, 23(22): 4742-4743. 참조).

[0569] 표면은 예를 들어, 상자성 비드(paramagnetic bead)의 형태로 제공될 수 있다. 상자성 비드는 자기장의 영향하에서 뭉쳐질 수 있다. 예를 들어, 상자성 표면에는 카르복실기와 같은 화학기(chemical group)가 제공될 수 있고, 이는 적절한 부착 조건에서 핵산을 포함하는 거대분자에 대한 결합 모이머티로 작용할 것이다. 거대분자는 적절한 용리(elution) 조건에서 이러한 표면에서 용리될 수 있다. 마이크로입자 및 비드의 표면은 UV-민감성 폴리카보네이트로 제공될 수 있다. 예를 들어, 핵산은 적합한 고정 버퍼(immobilization buffer)의 존재에서 활성화된 표면에 결합될 수 있다.

[0570] 마이크로입자 및 비드는 반응 용액 내에서 자유롭게 이동하도록 허용된 다음, 표면에 에칭(etching)된 마이크로웰(microwell) 또는 구덩이(pit) 내에 비드를 고정함으로써, 가역적으로 고정될 수 있다. 비드는 예를 들어, 비드에 부착된 고유한 핵산 "바코드(barcode)"를 사용하거나 또는 색상-코딩을 사용함으로써, 어레이(array)의 일부로 국소화(localisation)될 수 있다.

[0571] 표면은 EWOD (electrowetting-on-dielectric system)의 일부일 수 있다. EWOD 시스템은 미세액적(microdroplet) 형태의 매우 작은 액체 부피의 미세유체(microfluidic) 조작을 용이하게 하는 유전체-코팅된(dielectric-coated) 표면을 제공한다(예를 들어, Chou, W-L., et al. (2015) Recent Advances in Applications of Applications of Droplet Microfluidics, Micromachines, 6: 1249-1271. 참조). 액적(droplet) 부피는 프로그램이 가능하게 일렉트로웨팅(electrowetting) 기술에 의해 온-칩상에서 생성, 이동, 분할 및 결합될 수 있다. 따라서 일렉트로웨팅 시스템은 거대분자를 표면에 가역적으로 고정시키고/고정시키거나

표면에 고정된 거대분자를 조작하기 위한 대체 수단을 제공한다.

- [0572] 따라서, 본원에 기재되거나 정의된 바와 같은 본 발명에 따른 공정 또는 용도 중 어느 하나에서, 생화학적 반응은 표면을 포함하는 고상 반응 시스템에서 수행될 수 있다.
- [0573] 본원에 기재되거나 정의된 바와 같은 본 발명에 따른 공정 또는 용도 중 어느 하나에서, 생화학적 반응은 표면을 포함하는 고상 반응 시스템에서 수행되고, 반응 수행에 필요한 임의의 거대분자가 표면에 부착될 수 있다. 예를 들어, 생화학적 반응이 본원에 기재된 시험관내(*in vitro*) 반응 시스템에서 단일 가닥의 타겟 핵산 분자 또는 이중 가닥의 타겟 핵산 분자를 증폭하는 공정인 하나의 이러한 과정에서, 적어도 하나의 핵산 프라이머, 및/또는 반응 거대분자, 및/또는 IDR-거대분자 및/또는 하나 이상의 폴리펩타이드 보조인자(co-factor)가 표면에 부착될 수 있다.
- [0574] 본원에 기재되거나 정의된 바와 같은 본 발명에 따른 공정 또는 용도 중 어느 하나에서, 생화학적 반응이 표면을 포함하는 고상 반응 시스템에서 수행되고, 반응 수행에 필요한 IDR-거대분자는 표면에 부착될 수 있다.
- [0575] 본원에 기재되거나 정의된 바와 같은 본 발명에 따른 공정 또는 용도 중 어느 하나에서, 생화학적 반응은 시험관내(*in vitro*) 반응 시스템에서 이중 가닥의 타겟 핵산 분자를 증폭하는 RPA(recombinase polymerase amplification) 공정이고, 반응은 표면을 포함하는 고상 반응 시스템에서 수행되며, 및 리코비네이즈 제제 및/또는 리코비네이즈 로딩 단백질 및/또는 단일 가닥 안정화제 및/또는 폴리머레이즈 및/또는 엑소뉴클레이즈 및/또는 제1 핵산 프라이머(first nucleic acid primer) 및/또는 제2 핵산 프라이머(second nucleic acid primer)는 표면에 부착될 수 있다. 이러한 하나의 공정 또는 용도에서, 제1 핵산 프라이머 또는 제2 핵산 프라이머는 표면에 부착될 수 있다. 대안적으로, 다른 이러한 공정 또는 용도에서, 제1 핵산 프라이머 및 제2 핵산 프라이머 모두가 표면에 부착될 수 있다.
- [0576] 본원에 기재되거나 정의된 바와 같은 본 발명에 따른 공정 또는 용도 중 어느 하나에서, 생화학적 반응은 고상 반응 시스템에서 수행되고, 거대분자가 부착된 표면은 마이크로비드일 수 있고, 바람직하게는 마이크로비드는 실리콘, 유리, 젤 또는 폴리스티렌과 같은 폴리머 물질, 또는 이들의 임의의 조합을 포함한다.
- [0577] 본원에 기재된 공정 또는 용도 중 어느 하나에서, 생화학적 반응은 표면 및/또는 서브스트레이트(substrate)를 포함하는 고상 반응 시스템에서 수행되고, 표면 및/또는 서브스트레이트는 플로우-셀(flow-cell)로서 제공될 수 있다. 수행되는 생화학적 반응과 양립 가능한 임의의 적합한 플로우-셀이 사용될 수 있다. 적합한 플로우-셀은 생화학 반응을 수행하는 데 사용되는 시약이 흐를 수 있는 복수의 유체 채널(fluidic channel)을 포함할 수 있다. 생화학적 반응을 수행하는 데 사용되는 임의의 하나 이상의 거대분자는 유체 채널을 라이닝(lining)하는 표면에 부착될 수 있다. 적합한 플로우-셀은 단일 가닥의 또는 이중 가닥의 타겟 핵산 분자의 증폭을 위한 생화학 반응을 수행하는 데 사용될 수 있다. 본원에 기재된 공정, 용도 및 방법을 사용하여 수행되는 시퀀싱(sequencing) 반응은 또한 적합한 플로우-셀을 사용하여 수행될 수 있다.

도면의 간단한 설명

- [0578] 도 1은 다양한 주형 핵산 농도에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HIS2)을 사용한 실시간 리코비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다.
- 도 2는 다양한 주형 핵산 농도에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HIS5)을 사용한 실시간 리코비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다.
- 도 3은 다양한 주형 핵산 농도에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HRP1)을 사용한 실시간 리코비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다.
- 도 4는 다양한 주형 핵산 농도에서 IDR-태깅된 Gp32 융합 단백질(Gp32-Sup1)을 사용한 실시간 리코비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다.
- 도 5는 다양한 주형 핵산 농도에서 IDR-태깅된 Gp32 융합 단백질(Gp32-Sup2)을 사용한 실시간 리코비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다. 도 5a, 도 5b, 도 5c 및 도 5d에 나타난 실험은 각각 1 개, 2 개, 3 개 및 4 개의 Sup2 IDR 반복을 갖는 Gp32 융합 단백질을 사용한 것이다.
- 도 6은 다양한 MgOAc 농도에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HIS5)을 사용한 실시간 리코비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다.
- 도 7은 다양한 포스포크레아틴 농도에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HIS2)을 사용한 실시간 리코비네이즈

즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다.

도 8은 다양한 KOAc 농도에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HRP1)을 사용한 실시간 리컴비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다.

도 9는 PEG(크라우딩 제제)의 존재 또는 부존재 하에서, 7 개의 히스티딘 잔기로 태깅된 Gp32(단백질 정제 목적을 위함, 즉, IDR 태깅되지 않음) 및 IDR-태깅된 Gp32 융합 단백질(Gp32-Sup1)의 실시간 리컴비네이즈 폴리머레이즈 증폭 추적(traces)을 비교하여 나타낸 것이다.

도 10은 크라우딩 제제의 부존재 하에서 IDR 아미노산 서열 태그에 의해 매개되는 상 분리(입자 형성) 촉진에 대한 다가 금속 양이온의 효과를 나타낸 것이다. IDR 아미노산 서열은 Gp32 단백질에 태깅되어 Gp32-HIS2 융합 단백질(도 10a), Gp32-HRP1 융합 단백질(도 10b), Gp32-Sup1 융합 단백질(도 10c) 및 Gp32-Fib 융합 단백질(도 10d)를 생성하였다. 각 경우에, 대표적인 농도의 2가 금속 양이온, 즉 마그네슘(MgOAc), 망간(MgCl₂) 및 칼슘(CaCl₂)의 효과가 테스트되었다.

도 11은 크라우딩 제제의 부존재 하에서 IDR 아미노산 서열 태그에 의해 매개되는 상 분리(입자 형성) 촉진에 대한 다가 금속 양이온의 효과를 나타낸 것이다. IDR 아미노산 서열은 Gp32 단백질에 태깅되어 Gp32-Fib 융합 단백질(도 11a), Gp32-Sup1 융합 단백질(도 11b), Gp32-HIS2 융합 단백질(도 11c), Gp32-HRP1 융합 단백질(도 11d) 및 Gp32-HIS5 융합 단백질(도 11e)을 생성하였다. 각 경우에, 대표적인 농도의 2가 금속 양이온, 즉 마그네슘(MgOAc), 망간(MgCl₂) 및 칼슘(CaCl₂)의 효과가 테스트되었다.

도 12는 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HRP1)의 상 분리(입자 형성) 촉진 능력에 대한 2가 금속 양이온, 즉 마그네슘(MgOAc)의 효과를 나타낸 것이다.

도 13은 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HIS2)의 상 분리(입자 형성) 촉진 능력에 대한 2가 금속 양이온, 즉 마그네슘(MgOAc)의 효과를 나타낸 것이다.

도 14는 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HRP1)의 상 분리(입자 형성) 촉진 능력에 대한 다양한 농도의 2가 양이온, 즉 마그네슘(MgOAc)의 효과를 나타낸 것이다.

도 15는 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HIS2)의 상 분리(입자 형성) 촉진 능력에 대한 다양한 농도의 2가 양이온, 즉 마그네슘(MgOAc)의 효과를 나타낸 것이다.

도 16은 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 Gp32 융합 단백질(Gp32-HRP1)의 상 분리 촉진 능력에 대한 2가 금속 양이온, 즉 마그네슘(MgOAc) 첨가의 효과를 나타낸 것이다. 상 분리는 입자 형성으로 인한 불투명한 용액(opaque solution)의 형성으로 알 수 있으며(도 16a), 입자 형성은 추가적으로 입자의 펠릿(pellet)화로 알 수 있다(도 16b). RPA 단백질 성분은 펠릿 물질의 SDS-PAGE 분석에 의해 밝혀진 바와 같이 입자와 관련된 것으로 나타났다(도 16c).

도 17은 크라우딩 제제의 존재 또는 부존재 하에서 천연 Gp32 융합 단백질을 사용한 실시간 리컴비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다. 실험을 통하여 IDR을 포함하는 아미노산 서열로 태깅되지 않은 Gp32는 크라우딩 제제가 부존재하는 경우 증폭을 매개할 수 없음을 알 수 있었다.

도 18. 도 18a는 이중 프라이머 비드(Bead)를 사용한 실시간 증폭을 위해 설정된 반응 혼합물을 묘사한 그림이다. 도 18b는 실시간 반응에서 증폭된 산물을 묘사한 그림이다. 도 18c는 엔드-포인트(end-point) 반응에서의 증폭 특성화를 묘사한 그림이다. 도 18d는 고체 표면에 부착된 프라이머를 사용한 경우 또는 용액 내 고정되지 않은(free) 프라이머를 사용한 경우의 IDR-태깅된 Gp32 융합 단백질(Gp32-HIS2)을 사용한 실시간 리컴비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다. 도 18e 및 도 18f는 표면에 부착된 프라이머를 사용한 경우의 IDR-태깅된 Gp32 융합 단백질(Gp32-HIS2)을 사용한 엔드-포인트(end-point) 리컴비네이즈 폴리머레이즈 증폭 추적(traces)을 나타낸 것이다.

도 19는 MetaDisorder 소프트웨어 프로그램을 사용하여 생성된 Gp32(도 19a), UvsY(도 19b) 및 UvsX(도 19c)의 무질서 프로파일(disorder profile)을 나타낸 것이다. (무질서 영역: Disorder regions, 질서 영역: Ordered

Regions)

도 20은 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 RB69 라이게이즈 융합 단백질(RB69 라이게이즈-HIS2)의 상 분리(입자 형성) 촉진 능력에 대한 다양한 농도의 2가 양이온, 즉 마그네슘($MgCl_2$)의 효과를 나타낸 것이다.

도 21은 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 RB69 라이게이즈 융합 단백질(RB69 라이게이즈-HIS2)의 라이게이즈 활성 성능을 나타낸 것이다.

도 22는 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 RB69 라이게이즈 융합 단백질(RB69 라이게이즈-HIS2)의 라이게이즈 활성 성능을 태깅되지 않은 RB69 라이게이즈 및 T4 DNA 라이게이즈와 비교하여 나타낸 것이다.

도 23은 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 RB69 라이게이즈 융합 단백질(RB69 라이게이즈-HIS2)의 라이게이즈 활성 성능을 NEBNext Ultra II Ligation Master Mix와 비교하여 나타낸 것이다.

도 24는 크라우딩 제제의 부존재 하에 예시적인 시험관 내 생화학 반응 환경에서 IDR-태깅된 RB69 라이게이즈 융합 단백질(RB69 라이게이즈-HIS2)의 상 분리(입자 형성) 촉진 능력에 대한 ATP의 효과를 나타낸 것이다.

도 25는 FlexWell™ 챔버의 대표적인 섹션에 대한 명시야 및 형광 이미지를 나타낸 것으로, FlexWell™ 챔버는 비드(bead) 당 0 개, 5 개, 10 개, 20 개, 40 개 또는 80 개의 단일 가닥 UP1-UP2'-TF1L 주형(template) 카피를 가졌으며, 주형 카피는 50° C에서 1시간 동안 어닐링된 후, PEG와 같은 크라우딩 제제의 부재 하에서 IDR-태깅된 Gp32 융합 단백질(Gp32-Hrp1)과 고상에 부착된 프라이머를 사용한 리콤비네이즈 폴리머레이즈 증폭으로 증폭되었다. 증폭은 앰플리콘을 Nt. BbvFI로 Nicking(nicking)하고 Nick(nick)을 Aminoallyl-dUTP-XX-ATTO-594로 연장하여 검출하였다. 주형이 첨가되지 않은 비드에서는 형광이 관찰되지 않았고, 증가하는 양의 주형이 어닐링된 비드에서는 증가하는 양의 형광이 관찰되었다. 이는 크라우딩 제제의 부재 하에서도 비드의 고체 표면에서 증폭이 일어났음을 나타낸다.

도 26a는 IDR-태깅된 Gp32 융합 단백질(Gp32-Hrp1)에 의해 매개된 상-분리된 수성 입자의 형성을 나타내는 명시야 및 형광 이미지를 나타낸 것이다. 도 26b는 Gp32-Hrp1에 의해 매개되는 상-분리된 수성 입자 형성 시 증가된 반응 효율(Cas12a에 의한 핵산 절단 속도)을 나타내는 플롯을 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

실시예

다음 실시예는 본 발명을 예시하기 위해 제공되며 본 발명을 제한하기 위함이 아니다.

실시예 1. 인간 Otx1으로부터 유래된 IDR 태그를 갖는 Gp32를 사용한 리스테리아 모노사이토제네스(*Listeria monocytogenes*) 유전자 hly의 RPA(recombinase polymerase amplification).

실험의 목적 및 요약

본 실험은 인간 호메오박스(homeobox) 단백질 Otx1의 IDR(intrinsically disordered region)에서 발견되는 히스티딘이-풍부한 아미노산 도메인 서열을 포함하는 태그(tag)를 포함하는 Gp32 융합 단백질 제제(preparation)의 성능을 평가하기 위해 수행되었다.

본 실시예는 크라우딩 제제(crowding agent)의 부재에서 히스티딘이 풍부한 IDR(intrinsically disordered region) 도메인(Otx1)으로 C-말단에 태깅된 Gp32를 사용하여 다양한 주형(template) 농도에 걸쳐 리스테리아 모노사이토제네스(*Listeria monocytogenes*) 유전자 hly의 RPA(recombinase polymerase amplification)를 입증한다.

재료 및 방법

사용된 IDR 도메인 태그의 특이적 아미노산 서열은 AGHHHHHPHAAHPLSQSSGHHHHHHHHHQGYGGSG (서열번호 24)이었다. 이는 파지(phage) vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질(recombinant fusion protein)은 테스트 중에서 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity

chromatography)를 사용하여 정제되었다. 융합 단백질을 Gp32-HIS2로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 82로 제시되어 있다(표 23).

[0587] 그런 다음, 재조합 파지 vB EcoM NBG1 Gp32 융합 단백질은 리스테리아 모노사이토제네스 지놈(genomic) DNA로부터 유래된 DNA 주형의 표시된 카피(copy)를 사용하여, PEG가-없는 증폭에서, 즉 크라우딩 제제(crowding agent)의 부재에서 테스트되었다. 테스트 주형은 도 1에 표시된 대로 카피 수(copy number)로 적정되었다(titrated).

[0588] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴(phosphocreatine), 1 μM 크레아틴 카이네이즈(creatine kinase), 1 mM dNTP, 0.4 μM 정방향 프라이머, 0.4 μM 역방향 프라이머, 0.12 μM 프로브, 20 μM Gp32 융합, 4.8 μM UvsX, 8.6 μM UvsY, 0.135 μM 황색포도상구균(*S. aureus*)의 DNA 폴리머레이즈 및 0.27 μM 엑소뉴클레이즈 III(Exonuclease III)를 혼합하여 설정되었다. 반응은 주어진 농도에서 33 mM MgOAc와 함께 주형을 첨가에 의해 시작되었다.

[0589] 관련 프라이머 및 프로브는 아래에 표기되어 있다.

[0590] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98).

[0591] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTA (서열번호 99).

[0592] 프로브: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM] [THF][BHQ-1]ATACAAGGATTGGA (서열번호 100), 상기 FAM은 플루오레세인(fluorescein)이고, THF는 테트라하이드로퓨란(tetrahydrofuran)이며 BHQ는 블랙 홀 퀸처(Black Hole Quencher)이다.

[0593] 그런 다음, 반응(reaction)은 39℃에서 인큐베이션되었고 베어링 볼(bearing ball)을 사용하여 자성 혼합(magnetic mixing)하면서 형광계(fluorometer)에 넣었다.

[0594] 결과 및 결론

[0595] 도 1에 나타난 바와 같이, 테스트 주형은 RPA 반응 시작(initiation) 7분 이내에 높은 민감도(sensitivity)로 쉽게 검출되었다. 앰플리콘(amplicon)은 타겟의 10개 카피 정도로 검출되었다.

[0596] 따라서, PEG와 같은 크라우딩 제제(crowding agent) 부재에서 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 증폭이 효율적으로 발생함을 발견하였다.

[0597] 실시예 2. 인간 MafA로부터 유래된 IDR 태그를 갖는 Gp32를 사용한 리스테리아 모노사이토제네스 유전자 hly의 RPA(Recombinase polymerase amplification)

[0598] 실험의 목적 및 요약

[0599] 본 실험은 인간 전사 인자(transcription factor) MafA의 IDR(intrinsically disordered region)에서 발견되는 히스티딘이-풍부한 도메인 서열을 포함하는 태그를 포함하는 Gp32 융합 단백질 제제(preparation)의 성능을 평가하기 위해 수행되었다.

[0600] 본 실시예는 크라우딩 제제의 부재에서 히스티딘이-풍부한 IDR(intrinsically disordered region) 도메인(MafA)으로 C-말단에 태그된 Gp32를 사용하여 다양한 주형(template) 농도에 걸쳐 리스테리아 모노사이토제네스 유전자 hly의 RPA(recombinase polymerase amplification)를 입증한다.

[0601] 재료 및 방법

[0602] 사용된 IDR 도메인 태그의 특이적인 아미노산 서열은 SGHHGAHHGAHHHPAAAAAYEAFRGPFGAGGGADDMGAGHHGAHHAAHHHHHHHHHHGGAGHGGGAGHH (서열번호 27)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질을 Gp32-HIS5로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 85에 제시되어 있다(표 23).

[0603] 그런 다음, 재조합 파지 vB EcoM NBG1 Gp32 융합 단백질은 리스테리아 모노사이토제네스 지놈 DNA로부터 유래된 DNA 주형의 표시된 카피를 사용하여, PEG가-없는 증폭에서, 즉 크라우딩 제제의 부재에서 테스트되었다. 테스트 주형(test template)은 도 2에 표시된 대로 카피 수로 적정되었다.

- [0604] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μM 크레아틴 카이네이즈, 1 mM dNTP, 0.4 μM 정방향 프라이머, 0.4 μM 역방향 프라이머, 0.12 μM 프로브, 20 μM Gp32 융합, 4.8 μM UvsX, 8.6 μM UvsY, 0.135 μM 황색포도상구균(*S. aureus*) DNA 폴리머레이즈 및 0.27 μM 엑소뉴클레이즈 III를 혼합하여 설정되었다. 반응은 주어진 농도에서 33 mM MgOAc와 함께 주형의 첨가에 의해 시작되었다.
- [0605] 관련 프라이머 및 프로브는 아래에 표기되어 있다.
- [0606] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98).
- [0607] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).
- [0608] 프로브: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM][THF][BHQ-1]ATACAAGGATTGGA (서열번호 100), 상기 FAM은 플루오레세인, THF는 테트라하이드로퓨란, BHQ는 블랙 홀 퀀처이다.
- [0609] 그런 다음, 반응은 39°C에서 인큐베이션되었고 베어링 볼을 사용하여 자성 혼합하면서 형광계에 넣었다.
- [0610] **결과 및 결론**
- [0611] 도 2에 나타난 바와 같이, 테스트 주형은 RPA 반응 시작 후 10분 이내에 높은 민감도로 쉽게 검출되었다. 앰플리콘은 타겟의 10개 카피 정도로 검출되었다.
- [0612] 따라서, PEG와 같은 크라우딩 제제 부재에서 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 증폭이 효율적으로 발생함을 발견하였다.
- [0613] **실시예 3. 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1으로부터 유래된 IDR 태그를 갖는 Gp32를 사용한 리스테리아 모노사이토제네스 유전자 hly의 RPA(Recombinase Polymerase Amplification).**
- [0614] **실험의 목적 및 요약**
- [0615] 본 실험은 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1 단백질의 IDR(intrinsically disordered region)을 포함하는 태그를 포함하는 Gp32 융합 단백질 제제의 성능을 평가하기 위해 수행되었다.
- [0616] 본 실시예는 크라우딩 제제의 부재에서 효모(yeast) Hrp1 단백질의 IDR(intrinsically disordered region)을 포함하는 서열로 C-말단에 태그된 Gp32를 사용하여 다양한 주형 농도에 걸쳐 리스테리아 모노사이토제네스 유전자 hly의 RPA(recombinase polymerase amplification)를 입증한다.
- [0617] **재료 및 방법**
- [0618] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 GGNGGNNMNRGGNFGNQGDFNQMYQNPMGGYNPMNPQAMTDYYQKMQEYYQQMQ (서열번호 9)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 제조한 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 태그 후 배치된, 추가적인 헵타히스티딘(heptahistidine) 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질은 Gp32-HRP1로 표기되었다. 융합 단백질의 전체 아미노산 서열은 서열번호 79에 제시되어 있다(표 23).
- [0619] 그런 다음, 제조한 파지 vB EcoM NBG1 Gp32 융합 단백질은 리스테리아 모노사이토제네스 지놈 DNA로부터 유래된 DNA 주형의 표시된 카피를 사용하여, PEG가-없는 증폭에서, 즉 크라우딩 제제의 부재에서 테스트되었다. 테스트 주형은 도 3에 표시된 대로 카피 수로 적정되었다.
- [0620] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μM 크레아틴 카이네이즈, 1 mM dNTP, 0.4 μM 정방향 프라이머, 0.4 μM 역방향 프라이머, 0.12 μM 프로브, 20 μM Gp32 융합, 4.8 μM UvsX, 8.6 μM UvsY, 0.135 μM 황색포도상구균(*S. aureus*) DNA 폴리머레이즈 및 0.27 μM 엑소뉴클레이즈 III를 혼합하여 설정되었다. 반응은 주어진 농도에서 33 mM MgOAc와 함께 주형의 첨가에 의해 시작되었다.
- [0621] 관련 프라이머 및 프로브는 아래에 표기되어 있다.
- [0622] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98).
- [0623] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).

- [0624] 프로브: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM] [THF][BHQ-1]ATACAAGGATTGGA (서열번호 100), 상기 FAM은 플루오레세인, THF는 테트라하이드로퓨란, BHQ는 블랙 홀 퀀터이다.
- [0625] 그런 다음, 반응은 39℃에서 인큐베이션되었고 베어링 볼을 사용하여 자성 혼합하면서 형광계에 넣었다.
- [0626] **결과 및 결론**
- [0627] 도 3에 나타난 바와 같이, 테스트 주형은 RPA 반응 시작 후 7분 이내에 높은 민감도로 쉽게 검출되었다. 앰플리콘은 타겟의 10개 카피 정도로 검출되었다.
- [0628] 따라서, PEG와 같은 크라우딩 제제 부재에서 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 증폭이 효율적으로 발생함을 발견하였다.
- [0629] **실시예 4. 사카로미세스 세레비시아 Sup2로부터 유래된 IDR 태그를 갖는 Gp32를 사용한 리스테리아 모노사이토제네스 유전자 hly의 RPA(Recombinase Polymerase Amplification).**
- [0630] **실험의 목적 및 요약**
- [0631] 본 실험은 사카로미세스 세레비시아 Sup2 단백질의 IDR(intrinsically disordered region) 도메인을 포함하는 태그를 포함하는 Gp32 융합 단백질을 제조하는 성능을 평가하기 위해 수행되었다.
- [0632] 본 실시예는 크라우딩 제제의 부재에서 효모 Sup2 단백질의 IDR(intrinsically disordered region) 도메인을 포함하는 서열로 C-말단에 태그된 Gp32를 사용하여 다양한 주형 농도에 걸쳐 리스테리아 모노사이토제네스 유전자 hly의 RPA(recombinase polymerase amplification)를 입증한다.
- [0633] **재료 및 방법**
- [0634] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 YNPQGGYQQ (서열번호 19)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 태그 후 배치된, 추가적인 헥사히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질을 Gp32- Sup1으로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 72에 제시되어 있다(표 23).
- [0635] 그런 다음, 재조합 파지 vB EcoM NBG1 Gp32 융합 단백질은 리스테리아 모노사이토제네스 지놈 DNA로부터 유래된 DNA 주형의 표시된 카피를 사용하여, PEG가-없는 증폭에서, 즉 크라우딩 제제의 부재에서 테스트되었다. 테스트 주형은 도 C에 표시된 대로 카피 수로 적정되었다.
- [0636] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μM 크레아틴 카이네이즈, 1 mM dNTP, 0.4 μM 정방향 프라이머, 0.4 μM 역방향 프라이머, 0.12 μM 프로브, 20 μM Gp32 융합, 4.8 μM UvsX, 8.6 μM UvsY, 0.135 μM 황색포도상구균(*S. aureus*) DNA 폴리머레이즈 및 0.27 μM 엑소 뉴클레이즈 III를 혼합하여 설정되었다. 반응은 주어진 농도에서 33 mM MgOAc와 함께 주형의 첨가에 의해 시작되었다.
- [0637] 관련 프라이머 및 프로브는 아래에 표기되어 있다.
- [0638] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98).
- [0639] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).
- [0640] 프로브: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM] [THF][BHQ-1]ATACAAGGATTGGA (서열번호 100), 상기 FAM은 플루오레세인, THF는 테트라하이드로퓨란, BHQ는 블랙 홀 퀀터이다.
- [0641] 그런 다음, 반응은 39℃에서 인큐베이션되었고 베어링 볼을 사용하여 자성 혼합하면서 형광계에 넣었다.
- [0642] **결과 및 결론**
- [0643] 도 4에 나타난 바와 같이, 테스트 주형은 RPA 반응 시작 후 7분 이내에 높은 민감도로 쉽게 검출되었다. 앰플리콘은 타겟의 10개 카피 정도로 검출되었다.
- [0644] 따라서, PEG와 같은 크라우딩 제제 부재에서 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 증폭이 효율적으로 발생함을 발견하였다.
- [0645] **실시예 5. 사카로미세스 세레비시아 Sup2로부터 유래된 IDR 태그를 갖는 Gp32를 사용한 인간 apoB 유전자의**

RPA(Recombinase Polymerase Amplification).

실험의 목적 및 요약

본 실험은 사카로미세스 세레비시아 Sup2 단백질의 IDR(intrinsically disordered region) 도메인 아미노산 서열을 포함하는 태그를 포함하는 다수의 Gp32 융합 단백질 제제의 성능을 평가하기 위해 수행되었다. IDR 도메인 반복 단위(repeat unit)의 다양한 수(number)가 평가되었고 융합 단백질의 다양한 농도에서 조사되었다.

본 실시예는 크라우딩 제제의 부재에서 효모 Sup2 단백질의 IDR(intrinsically disordered region) 도메인을 포함하는 서열로 C-말단에 태그된 Gp32를 사용하여 인간 apoB(apolipoprotein B) 유전자의 RPA(recombinase polymerase amplification)를 입증한다.

재료 및 방법

사용된 IDR 도메인 태그의 특이적 아미노산 서열은 YNPQGGYQQ (서열번호 19)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 단일 YNPQGGYQQ 단위가 부착되었거나, 또는 2개, 3개 또는 4개의 반복(repeat)이 부착되었다. 제조된 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질을 Gp32-Sup2 (2개 반복; 서열번호 20), Gp32-Sup3 (3개 반복; 서열번호 21) 및 Gp32-Sup4 (4개 반복; 서열번호 22)로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 73, 서열번호 74 및 서열번호 75에 각각 제시되어 있다(표 23).

그런 다음, 제조된 파지 vB EcoM NBG1 Gp32 융합 단백질은 인간 지놈 DNA로부터 유래된 DNA 주형을 사용하여, PEG가-없는 증폭에서, 즉 크라우딩 제제 부재에서 Gp32-Sup1과 함께 테스트되었다.

반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μM 크레아틴 카이네이즈, 1 mM dNTP, 0.4 μM 정방향 프라이머, 0.4 μM 역방향 프라이머, 0.12 μM 프로브, 도 5a 내지 도 5d에서 표시된 농도에서의 Gp32 융합 단백질, 4.8 μM UvsX, 8.6 μM UvsY, 0.135 μM 황색포도상구균(*S. aureus*) DNA 폴리머레이즈 및 0.27 μM 엑소뉴클레이즈 III를 혼합하여 설정되었다. 반응은 33 mM MgOAc와 함께 주형의 첨가에 의해 시작되었다. 각 경우에서 사용된 테스트 주형 카피 수는 10,000이었다.

관련 프라이머 및 프로브는 아래에 표기되어 있다.

정방향 프라이머: GCAGCTGTATAGCAAATTCCTGTTGAAAGCAG (서열번호 101).

역방향 프라이머: TCCTGGCTGTATTCATTGTTGTTAAATTGG (서열번호 102).

프로브: CACTGATGCT TTTCTAGACACGAGATGA[FAM-dT]G[THF]C[BHQ1-dT]TGTGGAGCCTTTGT (서열번호 103), 상기 FAM은 플루오레세인, THF는 테트라하이드로퓨란, BHQ는 블랙 홀 퀀터이다.

그런 다음, 반응은 39°C에서 인큐베이션되었고 베어링 볼을 사용하여 자성 혼합하면서 형광계에 넣었다.

결과 및 결론

결과는 도 5에 나타내었다. 도 5a는 단일 IDR 도메인 태그 단위를 사용한 결과를 나타낸다. 도 5b 내지 도 5d는 각각 2개, 3개 및 4개의 IDR 도메인 태그 단위 반복을 사용한 결과를 나타낸다. 테스트 주형은 RPA 반응 시작 약 10분 후에 검출되었다.

PEG와 같은 크라우딩 제제 부재에서 이러한 Gp32-IDR-태그된 융합 단백질을 사용하면 증폭이 효율적으로 발생함을 발견하였다. 단일 IDR 도메인 태그 단위 및 2개의 IDR 도메인 태그 단위에서 최고의 성능이 나타났다. 3개의 IDR 도메인 태그 단위 또한 양호한 성능을 보였다.

실시예 6. 인간 MafA로부터 유래된 IDR 태그를 갖는 Gp32를 사용한 리스테리아 모노사이토제네스 유전자 hly의 RPA(Recombinase polymerase amplification) - 마그네슘 이온 농도 비교.

실험의 목적

본 실험은 인간 전사 인자 MafA의 IDR(intrinsically disordered region)에서 발견되는 히스티딘이-풍부한 도메인 서열을 포함하는 태그를 포함하는 Gp32 융합 단백질 제제의 성능을 평가하기 위해 수행되었다. 실험은 다양한 마그네슘 농도에 걸쳐 성능을 평가하였다.

본 실시예는 크라우딩 제제의 부재에서 히스티딘이-풍부한 IDR(intrinsically disordered region) 도메인

(MafA)으로 C-말단에 태그된 Gp32를 사용하여 다양한 마그네슘 농도에 걸쳐 리스테리아 모노사이토제네스 유전자 hly의 RPA(recombinase polymerase amplification)를 입증한다.

[0665] **재료 및 방법**

[0666] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 SGHHGAHHGAHHAAAAAYEAFRGPGFAGGGGADDMGAGHHGAHHAAHHHHAAHHHHHHHHHGGAGHGGGAGHH (서열번호 27)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질을 Gp32-HIS5로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 85에 제시되어 있다(표 23).

[0667] 그런 다음, 재조합 파지 vB EcoM NBG1 Gp32 융합 단백질은 리스테리아 모노사이토제네스 지놈 DNA로부터 유래된 DNA 주형의 표시된 카피를 사용하여, PEG가-없는 증폭에서, 즉 크라우딩 제제의 부재에서 테스트되었다. 테스트 주형은 반응당 10,000 카피로 제공되었고 마그네슘 이온 농도는 5.6 mM 내지 44.8 mM로 다양하였다.

[0668] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μM 크레아틴 카이네이즈, 1 mM dNTP, 0.4 μM 정방향 프라이머, 0.4 μM 역방향 프라이머, 0.12 μM 프로브, 20 μM Gp32 융합, 4.8 μM UvsX, 8.6 μM UvsY, 0.135 μM 황색포도상구균(*S. aureus*) DNA 폴리머레이즈 및 0.27 μM 엑소 뉴클레이즈 III를 혼합하여 설정되었다. 반응은 5.6 mM 내지 44.8 mM MgOAc로 표시된 농도의 MgOAc 및 주형의 첨가에 의해 시작되었다.

[0669] 관련 프라이머 및 프로브는 아래에 표기되어 있다.

[0670] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98).

[0671] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).

[0672] 프로브: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM] [THF][BHQ-1]ATACAAGGATTGGA (서열번호 100), 상기 FAM은 플루오레세인, THF는 테트라하이드로퓨란, BHQ는 블랙 홀 퀀처이다.

[0673] 그런 다음, 반응은 39°C에서 인큐베이션되었고 베어링 볼을 사용하여 자성 혼합하면서 형광계에 넣었다.

[0674] **결과 및 결론**

[0675] PEG와 같은 크라우딩 제제 부재에서 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 증폭이 효율적으로 발생함을 발견하였다.

[0676] 도 6에 나타난 바와 같이, 28 mM 이상의 마그네슘이 존재하였을 때 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 양호한 증폭이 발생함을 발견하였다. 본 실험에서 최적 농도는 33.6 mM인 것으로 나타났고, 44.8 mM까지 추가로 증가시키면 검출과 유사한 시간이 산출되었다.

[0677] **실시예 7. 인간 ApoB 유전자 단편의 RPA(recombinase polymerase amplification)에 대한 포스포크레아틴 수준(level)의 영향.**

[0678] **실험의 목적 및 요약**

[0679] 본 실험은 인간 호메오박스 단백질 Otx1의 IDR(intrinsically disordered region)에서 발견되는 히스티딘이-풍부한 도메인 서열을 포함하는 태그를 포함하는 Gp32 융합 단백질 제제의 성능에 대한 다양한 포스포크레아틴 수준의 효과를 평가하기 위해 수행되었다.

[0680] 본 실시예는 크라우딩 제제의 부재에서 히스티딘이-풍부한 IDR(intrinsically disordered region) 도메인(Otx1)으로 C-말단에 태그된 Gp32를 사용하여 인간 apoB (apolipoprotein) 유전자 단편의 RPA(recombinase polymerase amplification)를 입증한다.

[0681] **재료 및 방법**

[0682] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 AGHHHHHPAHHPHSQSSGHHHHHHHHHHQGYGGSG (서열번호 24)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질을 Gp32-HIS2로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서

열번호 82에 제시되어 있다(표 23).

[0683] 그런 다음, 재조합 파지 vB EcoM NBG1 Gp32 융합 단백질은 PEG가-없는 증폭에서, 즉 크라우딩 제제의 부재에서 테스트되었다. 포스포크레아틴 적정(titration)은 인간 apoB 분석을 사용하여 수행되었다. 테스트 주형은 10^4 카피의 농도로 제공되었다.

[0684] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 도면에 표시된 포스포크레아틴의 수준(level), 1 μ M 크레아틴 카이네이즈, 1 mM dNTP, 0.4 μ M 정방향 프라이머, 0.4 μ M 역방향 프라이머, 0.12 μ M 프로브, 20 μ M Gp32 융합, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M 황색포도상구균(*S. aureus*) DNA 폴리머레이즈 및 0.27 μ M 엑소뉴클레이즈 III를 혼합하여 설정되었다. 반응은 33 mM MgOAc와 함께 반응 당 10^4 카피 주형의 첨가에 의해 시작되었다.

[0685] 정방향 프라이머: GCAGCTGTATAGCAAATTCCTGTTGAAAGCAG (서열번호 101).

[0686] 역방향 프라이머: TCCTGGCTGTATTCATTGTTGTTAAATTGG (서열번호 102).

[0687] 프로브: CACTGATGCT TTTCCTAGACACGAGATGA[FAM-dT]G[THF]C[BHQ1-dT]TGTGGAGCCTTTGT (서열번호 103), 상기 FAM은 플루오레세인, THF는 테트라하이드로퓨란, BHQ는 블랙 홀 퀀터이다.

[0688] 그런 다음, 반응은 39°C에서 인큐베이션되었고 베어링 볼을 사용하여 자성 혼합하면서 형광계에 넣었다.

[0689] 결과 및 결론

[0690] PEG와 같은 크라우딩 제제 부재에서 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 증폭이 발생함을 발견하였다. 도 7a, 도 7b 및 도 7c에 나타난 바와 같이, PEG 기반 RPA (50 mM)에 사용된 표준 포스포크레아틴 농도에서 20분 이내에 증폭 활성이 거의 나타나지 않았다. 포스포크레아틴을 20 mM으로 감소시키면 최적의 성능이 나타났지만, 15 mM 내지 25 mM에서도 양호한 성능이 관찰되었으며, 30 mM 내지 35 mM에서도 20분 이내에 더 낮은 수준의 증폭이 관찰되었다.

[0691] 실시예 8. 사카로미세스 세레비시아 Hrp1으로부터 유래된 IDR 태그를 갖는 Gp32를 사용한 리스테리아 모노사이토제네스 유전자 hly의 RPA(Recombinase Polymerase Amplification) - 염 농도 비교.

[0692] 실험의 목적 및 요약

[0693] 본 실험은 사카로미세스 세레비시아 Hrp1 단백질의 IDR(intrinsically disordered region)을 포함하는 태그를 포함하는 Gp32 융합 단백질 제제의 성능을 평가하기 위해 수행되었다. 본 실험에서는 다양한 염 농도에 걸쳐 성능을 평가하였고, 이 경우에는 아세트산칼륨(potassium acetate)을 사용하였다.

[0694] 본 실시예는 크라우딩 제제의 부재에서 사카로미세스 세레비시아 Hrp1 단백질의 IDR(intrinsically disordered region)로 C-말단에 태그된 Gp32를 사용하면 리스테리아 모노사이토제네스 유전자 hly의 RPA(recombinase polymerase amplification)가 다양한 염 농도에 걸쳐 최적화될 수 있음을 입증한다.

[0695] 재료 및 방법

[0696] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 GGNNGGNMNRGGNFGNQDFNQMYQNPMGGYNPMMNPQAMTDYYQKMQEYYQQMQ (서열번호 9)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질을 Gp32-HRP1로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 79에 제시되어 있다(표 23).

[0697] 그런 다음, 재조합 파지 vB EcoM NBG1 Gp32 융합 단백질은 리스테리아 모노사이토제네스 지놈 DNA로부터 유래된 DNA 주형의 100개 카피를 사용하여, PEG가-없는 증폭에서, 즉 크라우딩 제제의 부재에서 테스트되었다. 아세트산칼륨 농도는 10 mM 내지 100 mM로 다양하였다.

[0698] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μ M 크레아틴 카이네이즈, 1 mM dNTP, 0.4 μ M 정방향 프라이머, 0.4 μ M 역방향 프라이머, 0.12 μ M 프로브, 20 μ M Gp32 융합, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M 황색포도상구균(*S. aureus*) DNA 폴리머레이즈 및 0.27 μ M 엑소뉴클레이즈 III를 혼합하여 설정되었다. 반응은 33 mM MgOAc와 함께 주형의 첨가에 의해 시작되었다.

- [0699] 관련 프라이머 및 프로브는 아래에 표기되어 있다.
- [0700] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98).
- [0701] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).
- [0702] 프로브: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM] [THF][BHQ-1]ATACAAGGATTGGA (서열번호 100), 상기 FAM은 플루오레세인, THF는 테트라하이드로퓨란, BHQ는 블랙 홀 퀀처이다.
- [0703] 그런 다음, 반응은 39℃에서 인큐베이션되었고 베어링 볼을 사용하여 자성 혼합하면서 형광계에 넣었다.
- [0704] **결과 및 결론**
- [0705] PEG와 같은 크라우딩 제제 부재에서 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 증폭이 효율적으로 발생함을 발견하였다.
- [0706] 또한, 크라우딩 제제 부재에서 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 증폭은 다양한 염 농도에 걸쳐 최적화될 수 있으며, 그 중 아세트산칼륨이 대표적인 예임을 발견하였다.
- [0707] 도 8에 나타난 바와 같이, 10 mM 이상의 아세트산칼륨이 존재하였을 때 상기 Gp32 IDR-태그된 융합 단백질을 사용하면 양호한 증폭이 발생함을 발견하였다. 본 실험에서 최적 농도 범위는 10 mM 내지 40 mM인 것으로 나타났다. 40 mM 이상의 농도에서는 덜 효율적인 증폭이 관찰되었다.
- [0708] **실시예 9. 사카로미세스 세레비시아 Sup2로부터 유래된 IDR 태그를 갖는 Gp32를 사용한 인간 ApoB 유전자 단편의 RPA(Recombinase Polymerase Amplification) - 크라우딩 제제와의 시너지 효과.**
- [0709] **실험의 목적 및 요약**
- [0710] 본 실험은 효모 SUP2 유전자의 IDR(intrinsically disordered region)에서, 구체적으로 파지 vB EcoM NBG1 Gp32의 C-말단에 부착된, Sup1 서열 YNPQGGYQQ (서열번호 19)에서 발견되는 히스티딘이-풍부한 서열을 포함하는 Gp32 융합 단백질 제제의 반응 효율에 대한 저농도 크라우딩 제제, 이 경우에는 PEG의 영향을 평가하기 위해 수행되었다. 상기 융합 단백질의 성능은 인간 apoB (apolipoprotein) 유전자 단편의 RPA(recombinase polymerase amplification)에서 Sup1 IDR 태그가 없는 Gp32 단백질과 비교되었다.
- [0711] 낮은 농도의 크라우딩 제제는 Sup1 IDR-태그된 Gp32의 반응 효율을 향상시킬 수 있고 시너지 효과가 관찰될 수 있는 조건이 달성될 수 있음을 발견하였다.
- [0712] **재료 및 방법**
- [0713] Gp32-Sup1
- [0714] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 YNPQGGYQQ (서열번호 19)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 제조된 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 도메인 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질을 Gp32-Sup1로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 72에 제시되어 있다(표 23).
- [0715] Gp32 (7His)
- [0716] 파지 vB EcoM NBG1 Gp32는 테스트 중에서 단백질의 바로 그 C-말단에 배치된 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. 융합 단백질을 Gp32(7His)로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 65에 제시되어 있다(표 23).
- [0717] 제조된 파지 vB EcoM NBG1 Gp32 융합 단백질은 인간 apoB (apolipoprotein) 유전자의 단편을 포함하는 DNA 주형을 사용하여 크라우딩 제제의 존재에서 또는 부재에서 RPA 반응에서 테스트되었다.
- [0718] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μM 크레아틴 카이네이즈, 1 mM dNTP, 0.4 μM 정방향 프라이머, 0.4 μM 역방향 프라이머, 0.12 μM 프로브, 20 μM Gp32 융합 단백질, 4.8 μM UvsX, 8.6 μM UvsY, 0.135 μM 황색포도상구균(*S. aureus*) DNA 폴리머레이즈 및 0.27 μM 엑소뉴클레아제 III를 혼합하여 설정되었다. 반응은 33 mM MgOAc와 함께 주형의 첨가에 의해 시작되었다. 각 경우에서 사용된 테스트 주형 카피 수는 10,000이었다. PEG는 관련 도면에 표시된 최종 농도로 추가되었다.

사용된 PEG의 종(species)은 PEG 분자량 35,000이었다.

[0719]

관련 프라이머 및 프로브는 아래에 표기되어 있다.

[0720]

정방향 프라이머: GCAGCTGTATAGCAAATTCCTGTTGAAAGCAG (서열번호 101).

[0721]

역방향 프라이머: TCCTGGCTGTATTTCATTGTTGTTAAATTGG (서열번호 102).

[0722]

프로브: CACTGATGCT TTTCTAGACACGAGATGA[FAM-dT]G[THF]C[BHQ1-dT]TGTGGAGCCTTTGT (서열번호 103), 상기 FAM은 플루오레세인, THF는 테트라하이드로퓨란, BHQ는 블랙 홀 원치이다.

[0723]

그런 다음, 반응은 39℃에서 인큐베이션되었고 베어링 볼을 사용하여 자성 혼합하면서 형광계에 넣었다.

[0724]

결과 및 결론

[0725]

결과는 도 9에 나타내었다. 도 9는 Gp32-Sup1 IDR-태그된 융합 단백질이 크라우딩 제제 PEG의 부재에서 테스트되었을 때 테스트 주형이 효율적으로 검출되었음을 나타낸다.

[0726]

Sup1 IDR 태그가 없는 Gp32-7His 융합 단백질이 0.5% 내지 2%의 크라우딩 제제 PEG의 존재에서 테스트되었을 때, 작지만 그럼에도 불구하고 검출 가능한 양의 증폭된 산물이 관찰되었다.

[0727]

Gp32-Sup1 IDR-태그된 융합 단백질이 크라우딩 제제 PEG의 존재에서 테스트되었을 때 테스트 주형이 효율적으로 검출되었다. 이 경우에는 다음에서 관찰된 양(amount)과 비교할 때 합한 양을 초과하는 증폭된 산물의 양으로, 시너지 효과를 확인할 수 있었다: (i) PEG의 부재에서 Gp32-Sup1 IDR-태그된 융합 단백질 및 (ii) PEG의 부재에서 Sup1 IDR 태그가 없는 Gp32-7His 융합 단백질 (예를 들어, 도 9를 참조하고 Sup1 1% PEG를 Sup1 0% PEG + 정상 GP32 1% PEG와 비교함).

[0728]

이러한 결과는 생화학 반응의 성능 효율에 대한 향상된 효과가 반응의 IDR-태그된 거대분자 성분을 낮은 농도의 크라우딩 제제와 결합할 때 관찰될 수 있고, 반응의 IDR-태그된 거대분자 성분을 낮은 농도의 크라우딩 제제와 결합할 때 반응 효율에 대한 시너지 효과를 촉진하는 조건이 달성될 수 있음을 입증한다.

[0729]

실시예 10. 다가 금속 양이온(multivalent metal cation)의 존재에서 IDR 태그에 의한 상 분리(phase separation) 촉진.

[0730]

실험의 목적 및 요약

[0731]

본 실험은 수성(aqueous) 시험관내(in vitro) 생화학 시스템에서 IDR(intrinsically disordered region) 도메인 아미노산 서열을 포함하는 태그를 각각 갖는 여러 Gp32 융합 단백질에 의해 유도된/야기된 상 분리(phase separation) 촉진에 대한 다가 금속 양이온(multivalent metal cation)의 효과를 평가하기 위해 수행되었다.

[0732]

본 실시예는 IDR 도메인 아미노산 서열을 포함하는 태그가 놀랍게도 상 분리를 촉진할 수 있고, 더욱 놀랍게는 이 효과가 다가 금속 양이온의 존재에 의해 향상된다는 것을 입증한다.

[0733]

재료 및 방법

[0734]

Gp32-HIS2 융합 단백질

[0735]

사용된 IDR 도메인 태그의 특이적 아미노산 서열은 AGHHHHHPAHHPHSQSSGHHHHHHHHHHQGYGGSG (서열번호 24)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 제조된 융합 단백질은 테스트 중에서 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-HIS2 융합 단백질의 전체 아미노산 서열은 서열번호 82에 제시되어 있다(표 23).

[0736]

Gp32-HRP1 융합 단백질

[0737]

사용된 IDR 도메인 태그의 특이적 아미노산 서열은 GGNNGNNMNRGGNFGNQGDFNQMYQNPMGGYNPMNPQAMTDYYQKMQEYYQQMQ (서열번호 9)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 제조된 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-HRP1 융합 단백질의 전체 아미노산 서열은 서열번호 79에 제공되어 있다(표 23).

- [0738] Gp32-Sup1 융합 단백질
- [0739] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 YNPQGGYQQ (서열번호 19)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 도메인 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-Sup1 융합 단백질의 전체 아미노산 서열은 서열번호 72에 제공되어 있다(표 23).
- [0740] Gp32-Fib 융합 단백질
- [0741] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 PGFSPRGGGFGGRGGFGDRGGGRGGGFGGGGRGGGFRGRGR (서열번호 1) 이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 도메인 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-Fib 융합 단백질의 전체 아미노산 서열은 서열번호 69에 제공되어 있다(표 23).
- [0742] 상 분리 분석(Phase separation assay)
- [0743] 아래에 요약된(outlined) 방법은 테스트된 모든 융합 단백질에 적용된다. 사용된 융합 단백질 용액의 부피는 정제 후 단백질 농도에 따라 다르다.
- [0744] 각각의 경우에서 50 μ l의 용액은 1000 ng/ μ l 최종 농도로 태그된 융합 단백질과 본원에 제시된 관련 도면 및 아래에 표시된 타겟 농도로 아세테이트 또는 염화물 형태의 금속 이온을 포함하도록, 구성되었다.
- [0745] Gp32-HIS2 융합의 경우, 정제 후 단백질 농도는 48 mg/ml이었다. 이 융합 단백질 1.04 μ l는 각 50 μ l 반응에 사용되어, 용액에서 1000 ng/ μ l의 최종 농도를 달성하였다. Gp32-HRP1 융합의 경우, 정제 후 단백질 농도는 39 mg/ml이었다. 이 융합 단백질 1.28 μ l는 각 50 μ l 반응에 사용되어, 용액에서 1000 ng/ μ l의 최종 농도를 달성하였다. Gp32-Sup1 융합의 경우, 정제 후 단백질 농도는 36 mg/ml이었다. 이 융합 단백질 1.4 μ l는 각 50 μ l 반응에 사용되어, 용액에서 1000 ng/ μ l의 최종 농도를 달성하였다. Gp32-Fib 융합의 경우, 정제 후 단백질 농도는 20.2 mg/ml이었다. 이 융합 단백질 2.48 μ l는 각 50 μ l 반응에 사용되어, 용액에서 1000 ng/ μ l의 최종 농도를 달성하였다.
- [0746] 이러한 실험에서 검출 가능한 상 분리 항상 발생에 필요한 2가(divalent) 금속 양이온 농도는 다음의 대표적인 2가 금속 양이온으로 테스트되었다: 마그네슘(MgOAc), 망간(MgCl₂) 및 칼슘(CaCl₂). 망간과 칼슘의 아세테이트 형태는 단순히 용액에서 알려진 이들의 불안정성 때문에 사용되지 않았다. 망간은 아세테이트 용액에서 시간이 지남에 따라 산화되고 칼슘 아세테이트는 용액에서 일부 박테리아의 성장을 지원(support)하는 것처럼 보이지만 염화칼슘은 그렇지 않다.
- [0747] 물(water), IDR-태그된 단백질 및 2가 금속 양이온을 포함하는 혼합물을 구성한 후, 혼합물을 볼텍싱(vortexing)하였고 스핀 다운(spin down)하였고, 혼합물의 10 μ l 샘플을 DHC-B01 C-칩 혈구계 슬라이드(Chip haemocytometer slide)로 옮겼다. 그런 다음 슬라이드를 현미경으로 x400 배율로 이미지화하였다. 검출 가능한 상 분리는 확대(magnification)를 통해 시각적으로 확인할 수 있고 혈구계를 사용하여 계산(count)될 수 있는 구형-같은 구형의 입자/입자(spherical-like globular foci/particle)의 형성에 의해 평가되었다. 그런 다음, 단위 부피당 구형의 입자(globular foci) 계산(count)이 수행될 수 있다. 구형의 입자 계산은 x400 배율로 218 μ m x 175 μ m의 배율 영역에서 형성된 구형의 입자 수를 계산하여 수행되었다. 이는 확대된 이미지를 20개의 정사각형 세그먼트(segment)(이미지의 4 x 5)로 분할하고, 이러한 세그먼트 중 하나에서 구형의 입자를 계산한 다음 이 숫자에 20을 곱하여 수행되었다.
- [0748] **결과 및 결론**
- [0749] 본 분석에서 최소 검출 가능한 상 분리 농도(minimum detectible phase separating concentration; MPSC) 바로 아래와 MPSC 바로 위 사이에서 발생하는 전이(transition)는 수행된 모든 반응에서 매우 갑자기 발생하는 것으로 관찰되었다. MPSC 바로 아래에서는 검출 가능한 상-분리된 수성 입자가 전혀 관찰되지 않았으며, 용액에는 시각적으로 검출 가능한 입자(구형의 입자)가 없는 것으로 보였다. MPSC 위의 전이(transition)는 갑자기 형성되는 수백 개의 시각적으로 검출 가능한 입자(구형의 입자)와 함께 매우 분명하였다.
- [0750] 구형의 입자 크기는 다양하였으며 사용된 IDR 태그 및 2가 금속 양이온과 관련되어 있는 것으로 나타났다. 구형

의 입자의 구체적인 크기는 중요하지 않은 것으로 결정되었다.

- [0751] 구형의 입자는 모양이 광범위하게 구형인 입자와-같은 구조로 존재하였다. 주어진 임의의 IDR 태그 및 주어진 임의의 2가 금속 이온 조합에 대한 구형의 입자 집단의 평균 지름은 표준 방법을 사용하여 쉽게 결정될 수 있다.
- [0752] 개별 융합 단백질을 사용한 결과는 아래에 요약되어 있다.
- [0753] Gp32-HIS2 융합 단백질
- [0754] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 마그네슘의 최소 농도는 10 mM인 것으로 결정되었다 - 약 600개 입자(구형의 입자)가 필드(field) 내에서 계산되었다.
- [0755] 이러한 조건에서 검출 가능한 상-분리 수성 입자의 형성을 향상시키는 데 필요한 칼슘 이온의 최소 농도는 12 mM으로 결정되었다 - 약 500개 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0756] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 망간 이온의 최소 농도는 2 mM인 것으로 결정되었다 - 약 180개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0757] 대표적인 확대된 이미지는 도 10a에 나와 있다.
- [0758] Gp32-HRP1 융합 단백질
- [0759] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 마그네슘 이온의 최소 농도는 16 mM인 것으로 결정되었다. 이 농도에서 약 580개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0760] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 칼슘 이온의 최소 농도는 24 mM인 것으로 결정되었다. 이 농도에서 약 240개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0761] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 망간 이온의 최소 농도는 6 mM인 것으로 결정되었다. 이 농도에서 약 260개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0762] 대표적인 확대된 이미지는 도 10b에 나와 있다.
- [0763] Gp32-Sup1 융합 단백질
- [0764] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 마그네슘 이온의 최소 농도는 24 mM인 것으로 결정되었다. 이 농도에서 약 280개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0765] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 칼슘 이온의 최소 농도는 32 mM인 것으로 결정되었다. 이 농도에서 약 460개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0766] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 망간 이온의 최소 농도는 4 mM인 것으로 결정되었다. 이 농도에서 약 220개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0767] 대표적인 확대된 이미지는 도 10c에 나와 있다.
- [0768] Gp32-Fib 융합 단백질
- [0769] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 마그네슘 이온의 최소 농도는 500 μ M로 결정되었다. 이 농도에서 약 340개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0770] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 칼슘 이온의 최소 농도는 1 mM인 것으로 결정되었다. 이 농도에서 약 500개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0771] 이러한 조건에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 데 필요한 망간 이온의 최소 농도는 500 μ M로 결정되었다. 이 농도에서 약 360개의 입자(구형의 입자)가 필드 내에서 계산되었다.
- [0772] 대표적인 확대된 이미지는 도 10d에 나와 있다.
- [0773] 이러한 분석을 사용하면 단백질이 태그된 경우 시험관내(*in vitro*) 생화학적 환경에서 검출 가능한 상-분리된 수성 입자의 형성을 향상시키는 IDR 또는 IDR 도메인의 기능적 능력(functional capability)은 x400 배율로 218 μ m x 175 μ m의 배율 영역에서 10개 이상의 입자(구형의 입자)가 형성되었을 때 확립될 수 있다고 결정되었다. 단백질이 태그된 경우 시험관내(*in vitro*) 생화학적 환경에서 상 분리를 유도하는 IDR 또는 IDR 도메인의

기능적 능력(functional capability)은 바람직하게는 $x400$ 배율로 $218\ \mu\text{m} \times 175\ \mu\text{m}$ 의 배율 영역에서 50개 이상의 입자(구형의 입자)가 형성되었을 때, 보다 바람직하게는 100개 이상의 입자(구형의 입자)가 형성되었을 때 확립될 수 있다.

[0774] 본원에 사용된 용어 "구형의 입자(globular foci)"는 "구상 입자(globule)", "입자(particle)" 또는 "구형 입자(globular particle)"와 동의어이며 이러한 용어는 상호교환적으로 사용될 수 있다.

[0775] **실시예 11. 다가 금속 양이온의 존재에서 IDR 태그에 의한 구형의 입자 형성.**

[0776] **실험의 목적 및 요약**

[0777] 본 실험은 시험관내(*in vitro*) 생화학 반응 시스템에서 IDR(intrinsically disordered region) 도메인 아미노산 서열을 포함하는 태그를 각각 갖는 여러 Gp32 융합 단백질에 의해 유도된/야기된 상 분리 촉진에 대한 다가 금속 양이온의 효과를 평가하기 위해 수행되었다.

[0778] 본 실시예는 IDR 도메인 아미노산 서열을 포함하는 태그가 상 분리를 촉진/향상시킬 수 있고, 이 효과가 다양한 다가 금속 양이온의 존재에서 발생한다는 것을 입증한다.

[0779] **재료 및 방법**

[0780] Gp32-Fib 융합 단백질

[0781] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 PGFSPRGGGFGGRGGFGDRGGRGGRGGFGGGGRGGGFRGRGR (서열번호 1)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 도메인 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-Fib 융합 단백질의 전체 아미노산 서열은 서열번호 69에 제공되어 있다(표 23).

[0782] Gp32-Sup1 융합 단백질

[0783] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 YNPQGGYQQ (서열번호 19)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 도메인 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-Sup1 융합 단백질의 전체 아미노산 서열은 서열번호 72에 제공되어 있다(표 23).

[0784] Gp32-HIS2 융합 단백질

[0785] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 AGHHHHHPAHHPHSQSSGHHHHHHHHHHQGYGSG (서열번호 24)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-HIS2 융합 단백질의 전체 아미노산 서열은 서열번호 82에 제시되어 있다(표 23).

[0786] Gp32-HRP1 융합 단백질

[0787] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 GGNNGGNNMNRGGNFGNQDFNQMYQNPMGGYNPMMNPQAMTDYYQKMQEYYQQMQ (서열번호 9)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 바로 그 C-말단에 배치된, 즉 융합 단백질의 C-말단에 IDR 태그 후 배치된, 추가적인 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-HRP1 융합 단백질의 전체 아미노산 서열은 서열번호 79에 제공되어 있다(표 23).

[0788] Gp32-HIS5 융합 단백질

[0789] 사용된 IDR 도메인 태그의 특이적 아미노산 서열은 SGHHGAHHGAHPAAAAAYEAFRGPFGAGGGADDMGAGHHGAHHAAHHHHHHHHHHGGAGHGGGAGHH (서열번호 27)이었다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중에서 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피를 사용하여 정제되었다. Gp32-HIS5 융합 단백질의 전체 아미노산 서열은 서열번호 85에 제시되어 있다.

- [0790] 상 분리 분석(Phase separation assay)
- [0791] 아래에 요약된(outlined) 방법은 테스트된 모든 융합 단백질에 적용된다. 사용된 융합 단백질 용액의 부피는 정제 후 단백질 농도에 따라 다르다.
- [0792] 각각의 경우에서 50 μ l의 용액은 1000 ng/ μ l (29.4 μ M)의 최종 농도로 태그된 융합 단백질 및 2가 금속 양이온을 포함하도록 구성되었다. 테스트된 금속 이온은 Mg^{2+} (MgOAc), Mn^{2+} (MnCl₂) 및 Ca^{2+} (CaCl₂)이었으며 각각의 경우에서 이들은 최종 농도 20 mM로 사용되었다.
- [0793] 물, IDR-태그된 단백질 및 다가 금속 양이온을 포함하는 혼합물을 구성한 후, 혼합물을 볼텍싱(vortexing)하였고, 스핀 다운(spin down)하였고, 혼합물의 10 μ l 샘플을 DHC-B01 C-칩 혈구계 슬라이드로 옮겼다. 그런 다음, 슬라이드는 x400 배율에서 명시야 현미경(bright field microscopy)을 사용하여 이미지화되었다. 확대(magnification)를 통해 시각적으로 확인될 수 있고 혈구계를 사용하여 계산될 수 있는 구형-같은 구형의 입자(입자)(spherical-like globular foci (particles))의 형성에 의해 상 분리는 평가되었다.
- [0794] **결과 및 결론**
- [0795] 대표적인 확대된 이미지는 도 11a 내지 11e에 나와 있다.
- [0796] 테스트된 IDR-태그된 Gp32 융합 단백질 각각에 대해, 검출 가능한 구형-같은 상-분리된 입자(구형의 입자)의 형성에 의해 결정된 바와 같이 검출 가능한 상 분리가 관찰되었다. 각각의 경우에서 효과는 Mg^{2+} , Mn^{2+} 및 Ca^{2+} 2가 금속 이온의 존재에서 관찰되었다.
- [0797] 따라서, 상 분리를 유도/향상시키는 다가 금속 이온의 능력은 상당히 상이한 아미노산 서열을 갖는 광범위한 상이한 IDR 태그에 적용될 수 있는 일반적인 특성인 것으로 보인다.
- [0798] **실시예 12. 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1에서 유래된 IDR(Intrinsically Disordered Region) 태그를 갖는 Gp32에 의한 구형의 입자의 형성.**
- [0799] **실험의 목적 및 요약**
- [0800] 본 실험은 크라우딩 제제(crowding agent) 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 구형의 입자(globular foci)를 형성하는데 있어서, 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1 단백질의 IDR(Intrinsically Disordered Region)을 포함하는 태그를 갖는 Gp32 융합 단백질 제제의 능력을 평가하기 위해 수행되었다.
- [0801] 본 실시예는 크라우딩 제제 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 IDR 도메인 아미노산 서열을 포함하는 태그가, 검출 가능한 상-분리 수성 입자의 형성에 의해 결정된 바와 같이, 상 분리를 촉진/향상시킬 수 있음을 보여주었다.
- [0802] **재료 및 방법**
- [0803] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 GGNNGGNMNRGGNFGNQGDFFNQMYQNPMGGYNPMNPQAMTDYYQKMQEYYQQMQ (서열번호 9)이다. 이는 과거 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중인 융합 단백질의 바로 C-말단에 배치된, 즉, 융합 단백질의 C-말단에서 IDR 태그 뒤에 배치된, 추가 헵타히스티딘 태그에 의존하는 표준 -단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 Gp32-HRP1로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 79로 나타낸다 (표 23).
- [0804] IDR 도메인 서열 태그의 효과를 테스트하기 위해 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경이 생성되었다. 이 경우, 상기 환경은 RPA(recombinase polymerase amplification) 반응을 특징으로 하는 환경이다.
- [0805] 상기 반응은 다음 프로토콜에 따라 설정되었다. 반응 혼합물은 다음 성분으로 생성되었다: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴(phosphocreatine), 1 μ M 크레아틴 카이네이즈(creatine kinase), 1 mM dNTPs, 0.2 μ M 정방향 프라이머, 0.2 μ M 역방향 프라이머, 0.516 μ M 프로브, 22.6 μ M Gp32-HRP fusion, 8.4 μ M UvsX, 15.3 μ M UvsY, 0.135 μ M 황색포도상구균 DNA 폴리머레이즈(*S. aureus* DNA polymerase) (거대 서브유닛), 및 0.27 μ M 엑소뉴클레이즈III(Exonuclease III). Gp32, UvsX,

UvsY, 폴리머레이즈 및 엑소뉴클레이즈 III는 프라이머, 버퍼, 뉴클레오타이드 및 크레아틴 카이네이즈의 혼합물에 한 단계로 첨가되기 전에 예비혼합물로 제조되었다. 전체 볼륨은 44 μ l였다. 일단 결합되면, 6 μ l의 280 mM MgOAc를 혼합물에 첨가하여 33 mM의 최종 농도를 달성하였다. 그런 다음, 10 μ l의 반응 혼합물을 C-Chip 혈구계(haemocytometer) 슬라이드로 옮기고, 이를 명시야 광 조건 및 형광 조건에서 이미지가 촬영되는 현미경으로 관찰하기 전에 39°C로 설정된 가열된 스테이지 세트에 놓았다.

[0806] 관련 프라이머 및 프로브는 하기에 나타내었다.

[0807] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98)

[0808] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).

[0809] 프로브: FAM(플루오레세인)으로 표지된 CCGCAATGGTGCCTCTCAGTACAATCTGCTCTGATG (서열 번호 104).

[0810] 결과 및 결론

[0811] 도 12에 나타난 바와 같이, Gp32에 부착된 HRP 태그는 (형광 표지된 프로브에 의해 검출된 바와 같이) 올리고뉴클레오타이드에 밀집된 것으로 보이는 많은 검출가능한 상 분리 수성 입자(구형의 입자)의 형성을 촉진하였다.

[0812] Gp32 단백질이 HRP IDR 태그가 아닌 헵타히스티딘 서열로만 태그된 것을 제외하고 동일한 재료 및 조건으로 별도의 실험을 수행하였다. 이 실험에서 구형의 입자)은 형성되지 않았으며(데이터는 나타내지 않음), 구형 조점의 형성이 IDR 태그에 의해 특이적으로 유도되었고, 결과적으로 헵타히스티딘 서열은 본 명세서에 정의된 바와 같은 기능적 IDR이 아님을 나타낸다.

[0813] 상기 결과는 시험관 내(*in vitro*) 생화학 반응 환경(이 경우 RPA(recombinase polymerase amplification) 반응으로 특징 짓는 반응 혼합물 환경으로 나타냄)에서 크라우딩 제제 없이 검출가능한 상 분리를 촉진하는 IDR 도메인 태그(이 경우 상술한 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1 아미노산 서열 태그로 나타냄)의 기능적 능력을 보여주었다.

[0814] 실시예 13. 인간 Otx1에서 유래된 IDR 태그를 갖는 Gp32에 의한 구형의 입자의 형성.

[0815] 실험의 목적 및 요약

[0816] 본 실험은 크라우딩 제제(crowding agent) 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 구형의 입자(globular foci)를 형성하는데 있어서, 인간 Otx1 단백질의 IDR(Intrinsically Disordered Region)을 포함하는 태그를 갖는 Gp32 융합 단백질 제제의 능력을 평가하기 위해 수행되었다.

[0817] 본 실시예는 크라우딩 제제 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 IDR 도메인 아미노산 서열을 포함하는 태그가 검출 가능한 상 분리를 촉진시킬 수 있음을 보여주었다.

[0818] 재료 및 방법

[0819] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 AGHHHHHPAHHPLSQSSGHHHHHHHHHHQGYGGSG (서열번호 24)이다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 제조된 융합 단백질은 테스트 중인 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 Gp32-HIS2로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 82로 나타낸다 (표 23).

[0820] IDR 도메인 서열 태그의 효과를 테스트하기 위해 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경이 생성되었다. 이 경우, 상기 환경은 RPA(recombinase polymerase amplification) 반응을 특징으로 하는 환경이다.

[0821] 상기 반응은 다음 프로토콜에 따라 설정되었다. 반응 혼합물은 다음 성분으로 생성되었다: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μ M 크레아틴 카이네이즈, 1 mM dNTPs, 0.2 μ M 정방향 프라이머, 0.2 μ M 역방향 프라이머, 0.516 μ M 프로브, 22.6 μ M Gp32-HIS2 fusion, 8.4 μ M UvsX, 15.3 μ M UvsY, 0.135 μ M 황색포도상구균 DNA 폴리머레이즈 (거대 서브유닛), 및 0.27 μ M 엑소뉴클레이즈 III. Gp32-His2, UvsX, UvsY, 폴리머레이즈 및 엑소뉴클레이즈 III는 프라이머, 버퍼, 뉴클레오타이드 및 크레아틴 카이네이즈의 혼합물에 한 단계로 첨가되기 전에 예비혼합물로 제조되었다. 전체 볼륨은 44 μ l였다. 일단 결합되면, 6 μ l의 280 mM MgOAc를 혼합물에 첨가하여 33 mM의 최종 농도를 달성하였다. 그런 다음, 10 μ l의 반응 혼합물을 C-Chip 혈구계 슬라이드로 옮기고, 이를 명시야 광 조건 및 형광 조건에서 이미지가 촬영되는 현미경으로 관찰하기 전에 39°C로 설정된 가열된 스테이지 세트에 놓았다.

- [0822] 관련 프라이머 및 프로브는 하기에 나타내었다.
- [0823] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98)
- [0824] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).
- [0825] 프로브: FAM(플루오레세인)으로 표지된 CCGCAATGGTGACTCTCAGTACAATCTGCTCTGATG (서열 번호 104).
- [0826] **결과 및 결론**
- [0827] 도 13에 나타난 바와 같이, Gp32에 부착된 HIS IDR 태그는 (형광 표지된 프로브에 의해 검출된 바와 같이) 올리 고뉴클레오타이드에 밀집된 것으로 보이는 많은 검출가능한 상 분리 수성 입자(구형의 입자)의 형성을 촉진하였 다.
- [0828] 구상 입자(globule)는 본 명세서에 추가로 기재된 바와 같은 HRP IDR 태그가 Gp32에 부착될 때 형성된 구상 입 자에 비해 크기가 더 작게 나타남에 주목하였다.
- [0829] 상기 결과는 시험관 내(*in vitro*) 생화학 반응 환경(이 경우 RPA(recombinase polymerase amplification) 반응 으로 특징 짓는 반응 혼합물 환경으로 나타냄)에서 크라우딩 제제 없이 검출가능한 상 분리를 촉진하는 IDR 도 메인 태그(이 경우 상술한 HIS2 아미노산 서열 태그를 나타냄)의 기능적 능력을 보여주었다.
- [0830] **실시예 14. 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1에서 유래된 IDR 태그를 갖는 Gp32에 의 한 구형의 입자의 형성에 대한 다가 금속 양이온(multivalent metal cations)의 효과**
- [0831] **실험의 목적 및 요약**
- [0832] 본 실험은 크라우딩 제제(crowding agent) 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 구형의 입자(globular foci)를 형성하는데 있어서, 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1 단백질 의 IDR(Intrinsically Disordered Region)을 포함하는 태그를 갖는 Gp32 융합 단백질 제제의 능력에 대한 다가 금속 양이온의 효과를 평가하기 위해 수행되었다.
- [0833] 본 실시예는 IDR 도메인 아미노산 서열을 포함하는 태그가, 검출 가능한 상-분리 수성 입자의 형성에 의해 결정 된 바와 같이, 크라우딩 제제 없이 상 분리를 촉진/향상시킬 수 있고, 상기 상 분리는 다가 금속 양이온의 존재 하에 향상되고 상 분리를 촉진하기 위한 최적화된 농도가 결정 될 수 있음을 보여주었다.
- [0834] **재료 및 방법**
- [0835] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 GGNNGGNMNRGGNFGNGDFNQMYQNPMGGYNPMMNPQAMTDYYQKMQEYYQQMQ (서열번호 9)이다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중인 융합 단백질의 바로 C-말단에 배치된, 즉, 융 합 단백질의 C-말단에서 IDR 태그 뒤에 배치된, 추가 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니 켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하 여 정제되었다. 융합 단백질은 Gp32-HRP1로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 79로 나 타낸다 (표 23).
- [0836] 다양한 농도의 2가 금속 양이온의 존재 하에 IDR 도메인 서열 태그의 효과를 테스트하기 위해 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경이 생성되었다. 이 경우, 상기 환경은 RPA(recombinase polymerase amplification) 반응을 특징으로 하는 환경이다.
- [0837] 상기 반응은 다음 프로토콜에 따라 설정되었다. 반응 혼합물은 다음 성분으로 생성되었다: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μ M 크레아틴 카이네이즈, 1 mM dNTPs, 0.26 μ M 정방향 프라이머, 0.26 μ M 역방향 프라이머, 0.4 μ M 프로브, 22.6 μ M Gp32-HRP fusion, 8.4 μ M UvsX 및 15.3 μ M UvsY. Gp32, UvsX 및 UvsY는 프라이머, 버퍼, 뉴클레오타이드 및 크레아틴 카이네이즈의 혼합 물에 한 단계로 첨가되기 전에 예비혼합물로 제조되었다. MgOAc는 관련 도면에 나타난 바와 같이, 최종 농도를 달성하기 위해 혼합물에 첨가되었다. 그런 다음, 10 μ l의 반응 혼합물을 C-Chip 혈구계(haemocytometer) 슬라이 드로 옮기고, 이를 명시야 광 조건 및 형광 조건에서 이미지가 촬영되는 현미경으로 관찰하기 전에 39°C로 설정 된 가열된 스테이지 세트에 놓았다.
- [0838] 관련 프라이머 및 프로브는 하기에 나타내었다.

- [0839] 정방향 프라이머: CGCCTGCAAGTCTCTAAGACGCCAATCGAAAAGAAAC (서열번호 98)
- [0840] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).
- [0841] 프로브: FAM(플루오레세인)으로 표시된 CCGCAATGGTGCCTCTCAGTACAATCTGCTCTGATG (서열 번호 104).
- [0842] **결과 및 결론**
- [0843] 도 14a 및 b에 나타난 바와 같이, Gp32에 부착된 HRP IDR 태그는 (형광 표시된 프로브에 의해 검출된 바와 같이) 올리고뉴클레오타이드에 밀집된 것으로 보이는 많은 검출가능한 상 분리 수성 입자(구형의 입자)의 형성을 촉진하였다.
- [0844] 상기 구형의 입자는 22.4 mM MgOAc 에서 명확하게 볼 수 있었다. 구형의 입자의 최적 형성은 33 mM MgOAc에서 발생하였다. 33 mM 이상의 농도에서 구상 입자의 일부 덩어리가 관찰되기 시작하였다.
- [0845] 놀랍게도, 본 명세서에 기재된 바와 같이, 33 mM MgOAc는 크라우딩 제제 없이 IDR-태그된 Gp32를 사용한 RPA(recombinase polymerase amplification) 반응에서 최적 증폭 효율이 관찰되는 마그네슘 농도이다. 따라서 IDR 태그-매개 구형의 입자 형성의 효율성은 놀랍게도 크라우딩 제제 없이 시험관 내(*in vitro*) 시스템에서 예시적인 생화학적 반응(이 경우 테스트 생화학 시스템의 예로써 IDR-태그된 단백질을 사용한 RPA 반응에서의 증폭을 나타냄)의 효율성과 관련이 있다.
- [0846] 상기 결과는 크라우딩 제제 없이 생화학 반응의 효율을 유도/증가시키는 IDR 도메인 서열 태그의 성능이 상 분리의 효율성과 관련이 있다는 놀라운 결론을 뒷받침하고, 및 이는 차례로 IDR(Intrinsically Disordered Region) 또는 IDD(Intrinsically Disordered Domain)의 기능에 영향을 미치는 시스템에 포함된 다가 금속 양이온 또는 이의 기능적 등가물의 농도에 의해 향상되는 것으로 보인다.
- [0847] **실시예 15. 인간 Otx1에서 유래된 IDR 태그를 갖는 Gp32에 의한 구형의 입자의 형성에 대한 다가 금속 양이온(multivalent metal cations)의 효과.**
- [0848] **실험의 목적 및 요약**
- [0849] 본 실험은 크라우딩 제제(crowding agent) 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 입자/구형의 입자(globular foci)를 형성하는데 있어서, 인간 Otx1 단백질의 IDR(Intrinsically Disordered Region)을 포함하는 태그를 갖는 Gp32 융합 단백질 제제의 능력에 대한 다가 금속 양이온의 효과를 평가하기 위해 수행되었다.
- [0850] 본 실시예는 IDR 도메인 아미노산 서열을 포함하는 태그가, 크라우딩 제제 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 검출가능한 상 분리를 촉진시킬 수 있고, 상기 검출가능한 상 분리는 다가 금속 양이온의 존재에 의해 향상되며 검출가능한 상 분리를 촉진하기 위한 최적화된 농도가 결정 될 수 있음을 보여주었다.
- [0851] **재료 및 방법**
- [0852] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 AGHHHHHPAHHPHSQSSGHHHHHHHHHHQGYGGSG (서열번호 24)이다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 제조용 융합 단백질은 테스트 중인 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 Gp32-HIS2로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 82로 나타낸다 (표 23).
- [0853] 다양한 농도의 2가 금속 양이온의 존재 하에 IDR 도메인 서열 태그의 효과를 테스트하기 위해 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경이 생성되었다. 이 경우, 상기 환경은 RPA(recombinase polymerase amplification) 반응을 특징으로 하는 환경이다.
- [0854] 상기 반응은 다음 프로토콜에 따라 설정되었다. 반응 혼합물은 다음 성분으로 생성되었다: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴(phosphocreatine), 1 μM 크레아틴 카이네이스(creatine kinase), 1 mM dNTPs, 0.26 μM 정방향 프라이머, 0.26 μM 역방향 프라이머, 0.4 μM 프로브, 20 μM Gp32-HIS2 fusion, 8.4 μM UvsX 및 8.6 μM UvsY. Gp32, UvsX 및 UvsY는 프라이머, 버퍼, 뉴클레오타이드 및 크레아틴 카이네이스의 혼합물에 한 단계로 첨가되기 전에 예비혼합물로 제조되었다. MgOAc는 관련 도면에 나타난 바와 같이, 최종 농도를 달성하기 위해 혼합물에 첨가되었다. 그런 다음, 10 μl의 반응 혼합물을 C-Chip 혈구계(haemocytometer) 슬라이드로 옮기고, 이를 명시야 광 조건 및 형광 조건에서 이미지가 촬영되는 현미경

으로 관찰하기 전에 39℃로 설정된 가열된 스테이지 세트에 놓았다.

[0855] 관련 프라이머 및 프로브는 하기에 나타내었다.

[0856] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98)

[0857] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).

[0858] 프로브: FAM(플루오레세인)으로 표지된 CCGCAATGGTGACTCTCAGTACAATCTGCTCTGATG (서열 번호 104).

[0859] 결과 및 결론

[0860] 도 15a 및 b에 나타난 바와 같이, Gp32에 부착된 HIS2 태그는 (형광 표지된 프로브에 의해 검출된 바와 같이) 올리고뉴클레오타이드에 밀집된 것으로 보이는 많은 구형의 입자의 형성을 촉진하였다.

[0861] 상기 구형의 입자는 22.4 mM MgOAc 에서 명확하게 볼 수 있었다. 구형의 입자의 최적 형성은 33-39 mM의 MgOAc 에서 발생하였다. 39 mM 이상의 농도에서 구상 입자의 일부 덩어리가 관찰되기 시작하였다.

[0862] 놀랍게도, 본 명세서에 기재된 바와 같이, 33 mM 내지 39 mM의 MgOAc는 크라우딩 제제 없이 IDR-태그된 Gp32를 사용한 RPA(recombinase polymerase amplification) 반응에서 최적 증폭 효율이 관찰되는 마그네슘 농도이다. 따라서 IDR 태그-매개 구형의 입자 형성의 효율성은 놀랍게도 크라우딩 제제 없이 시험관 내(*in vitro*) 시스템에서 예시적인 생화학적 반응(이 경우 테스트 생화학 시스템의 예로써 IDR-태그된 단백질을 사용한 RPA 반응에서의 증폭을 나타냄)의 효율성과 관련이 있다.

[0863] 상기 결과는 크라우딩 제제 없이 생화학 반응의 효율을 유도/증가시키는 IDR 도메인 서열 태그의 성능이 상 분리의 효율성과 관련이 있다는 놀라운 결론을 뒷받침하고, 및 이는 차례로 IDR(Intrinsically Disordered Region) 또는 IDD(Intrinsically Disordered Domain)의 기능에 영향을 미치는 시스템에 포함된 다가 금속 양이온 또는 이의 기능적 등가물의 농도에 의해 향상되는 것으로 보인다.

[0864] 실시예 16. 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1에서 유래된 IDR 태그를 갖는 Gp32에 의한 구형의 입자의 형성에 대한 마그네슘 농도의 효과

[0865] 실험의 목적 및 요약

[0866] 본 실험은 크라우딩 제제(crowding agent) 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 구형의 입자(globular foci)를 형성하는데 있어서, 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1 단백질의 IDR(Intrinsically Disordered Region)을 포함하는 태그를 갖는 Gp32 융합 단백질을 제제의 능력에 대한 마그네슘의 효과를 평가하기 위해 수행되었다.

[0867] 본 실시예는 IDR 도메인 아미노산 서열을 포함하는 태그가, 검출 가능한 상-분리 수성 입자의 형성에 의해 결정된 바와 같이, 크라우딩 제제 없이 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 상 분리를 촉진/향상시킬 수 있고, 상기 상 분리는 마그네슘 이온의 존재에 의존하며, 반응 혼합물의 모든 단백질 성분은 벌크 상(bulk phase)이 아닌 상-분리된 입자와 관련된 것으로 밝혀진 것임을 보여주었다.

[0868] 재료 및 방법

[0869] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 GGNGGNNMNRGGNFGNQDFNQMYQNPMMGGYNPMNPQAMTDYYQKMQEYYQQMQ (서열번호 9)이다. 이는 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중인 융합 단백질의 바로 C-말단에 배치된, 즉, 융합 단백질의 C-말단에서 IDR 태그 뒤에 배치된, 추가 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 Gp32-HRP1로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 79로 나타낸다 (표 23).

[0870] 마그네슘 이온의 존재 또는 부재 하에(유무에 관계없이) IDR 도메인 서열 태그의 효과를 테스트하기 위해 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경이 생성되었다.

[0871] 상기 반응은 다음 프로토콜에 따라 설정되었다. 1 ml 반응 혼합물은 다음 성분으로 생성되었다: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μM 크레아틴 카이네이즈, 0.4 μM 정방향 프라이머, 0.4 μM 역방향 프라이머, 0.4 μM 프로브, 20.26 μM Gp32-HRP fusion, 5 μM UvsX, 8.67 μM

UvsY, 0.127 μ M 황색포도상구균 DNA 폴리머레이즈 (거대 서브유닛) 및 0 mM 또는 33.6 mM MgOAc.

[0872] 관련 프라이머 및 프로브는 하기에 나타내었다.

[0873] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98)

[0874] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (서열번호 99).

[0875] 프로브: FAM(플루오레세인)으로 표지된 CCGCAATGGTGACTCTCAGTACAATCTGCTCTGATG (서열 번호 104).

[0876] 완성된 혼합물의 사진을 찍었다.

[0877] 혼합물을 1분 동안 2,000 rcf에서 회전시켰다. MgOAc 혼합물에서 상층액을 제거하였다. 33.6 mM MgOAc와의 혼합물에는 작은 펠렛이 남아 있었고, 이는 상-분리된 구상 입자/입자로 구성된 것으로 추정되었다. MgOAc가 없는 혼합물에서는 유사한 펠렛이 관찰되지 않았다. 펠렛을 가용화하기 위하여 10 μ l의 1% SDS 용액을 첨가하였다. 펠렛의 볼륨은 4.5 μ l로 추정되어 총 볼륨은 14.5 μ l로 추정되었다. 각 샘플의 1 μ l를 SDS-PAGE로 분석하였다.

[0878] 결과 및 결론

[0879] 도 16a에 나타난 바와 같이, 1 ml의 RPA 혼합물에 마그네슘 아세테이트를 첨가하면 혼합물이 불투명해진다. 이는 마그네슘 아세테이트가 없는 경우에는 관찰되지 않았다. 상기 불투명한 효과는 등가적인 더 작은 반응에서 볼 수 있는 것과 같은 효과였으며, 현미경으로 볼 때 구형의 입자/상-분리 입자가 2-3 마이크로미터 범위로 추정되는 전형적인 직경의, 및 일반적으로 나노리터 당 약 200-400개의 입자로 형성되는 것이 관찰된 경우이다. 원심 분리를 하면 이러한 불투명한 혼합물이 제거되고 펠렛 또는 더 낮은 상이 튜브 바닥에 형성되는 것으로 나타났으며, 이는 단일 부피로 함께 강제된 입자 덩어리로 추정되었다(도 16b). 상기 펠렛화된 부분의 예상 볼륨은 약 4 μ l이며, 이는 평균 입자 직경이 3 μ m (따라서 약 13 펠토폴리터의 볼륨)이고 혈구계/현미경 필드 계산을 기반으로 나노리터당 약 400개 입자, 마이크로리터당 400,000개 입자의 풍부함을 가정했을 때 형성될 것으로 예상되는 입자의 대략적인 예상 총 볼륨이고, 이는 혼합물의 마이크로리터당 약 5 nL의 추정 부피를 생성하므로 혼합물의 1 ml당 약 5마이크로리터의 볼륨을 생성한다.

[0880] 마그네슘 아세테이트를 첨가하기 전후의 1 마이크로리터의 벌크 혼합물(또는 맑은 상(cleared phase))을 분석한 결과, 첨가하기 전에 다양한 단백질이 투명한 액체에서 예상대로 식별될 수 있음을 보여주고 - Gp32는 질량 기준으로 가장 두드러진 단백질이다. 응축 및 제거 후 상층액에서는 미량의 단백질만 발견될 수 있는 반면, 펠렛은 RPA 혼합물에 추가된 모든 단백질이 매우 풍부하다(도 16c). 추론에 의해 대략적으로 반응물의 약 200배 농도가 달성되었고 총 단백질 농도가 약 200 μ g/ μ l였다고 가정할 수 있다.

[0881] 상기 결과는 RPA 반응 혼합물의 모든 단백질 성분, 즉 크레아틴 카이네이즈, Gp32-HRP 융합, UvsX, UvsY 및 폴리머레이즈가 벌크 상(bulk phase)이 아니라 상-분리된 입자와 연관되어 있음을 보여주었다.

[0882] **실시예 17. 생화학적 반응의 효율을 증가시키는데 있어서 IDR(intrinsically disordered region)을 포함하는 아미노산 서열의 본질적 성질의 입증.**

[0883] 실험의 목적 및 요약

[0884] 본 실험은 크라우딩 제제(crowding agent)의 존재 또는 부재 하에 예시적인 시험관 내(*in vitro*) 생화학적 반응 환경에서 IDR(intrinsically disordered region)을 포함하는 아미노산 서열을 포함하는 태그가 결합된 Gp32 단백질의 성능을 평가하기 위하여 수행되었다.

[0885] 본 실시예는 IDR 도메인 아미노산 서열을 포함하는 태그가 없는 경우, Gp32가 크라우딩 제제 없이 RPA(recombinase polymerase amplification)를 효율적으로 매개할 수 없고 분석 기간 내에 본 분석 시스템에서 검출이 이루어진 지점까지 수행할 수 없음을 보여주었다. 실시예 1 내지 5와 같은 상기 기재된 다른 실시예와 비교하여, 이들 데이터는 IDR 도메인 아미노산 서열을 포함하는 태그가 크라우딩 제제 없이 생화학적 반응의 효율을 증가시키는데 필수적이라는 것을 보여준다.

[0886] 재료 및 방법

[0887] 파지 vB EcoM NBG1 Gp32 단백질은 외인성 IDR 태그 또는 히스티딘 태그가 없는 네이티브(native) 형태로 정제되었다. 상기 단백질은 헤파린 수지를 사용하여 정제되었고 NaCl 단계 구배로 용출되었다. 400 mM NaCl 분획의 네이티브 Gp32 단백질을 테스트하였다.

[0888] 크라우딩 제제의 존재 또는 부재 하에 네이티브 Gp32 단백질의 효과를 테스트하기 위해 예시적인 시험관 내(*in*

vitro) 생화학적 반응 환경이 생성되었다.

- [0889] 상기 반응은 다음 프로토콜에 따라 설정되었다: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μ M 크레아틴 카이네이즈, 1 mM dNTPs, 0.4 μ M 정방향 프라이머, 0.4 μ M 역방향 프라이머, 0.12 μ M 프로브, 20 μ M 네이티브 Gp32 단백질, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M 황색포도상구균 DNA 폴리머레이즈 (거대 서브유닛) 및 0.27 μ M 엑소뉴클레이즈 III.
- [0890] PEG-기반 반응 혼합물은 다음 성분으로 생성되었다: 50 mM Tris HCl pH 8.3, 100 mM KOAc, 1 mM DTT, 2.5 mM ATP, 50 mM 포스포크레아틴, 1 μ M 크레아틴 카이네이즈, 1 mM dNTPs, 0.4 μ M 정방향 프라이머, 0.4 μ M 역방향 프라이머, 0.12 μ M 프로브, 20 μ M 네이티브 Gp32 단백질, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M 황색포도상구균 DNA 폴리머레이즈 (거대 서브유닛), 0.27 μ M 엑소뉴클레이즈 III 및 관련 도면에서 나타낸 바와 같이 최종 농도의 PEG. 사용된 PEG의 종류는 PEG 분자량 35,000이었다.
- [0891] 모든 반응에서 관련 프라이머 및 프로브는 하기에 나타내었다.
- [0892] 정방향 프라이머: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (서열번호 98)
- [0893] 역방향 프라이머: CTGCATCTCCGTGGTATACTAATACATTGTTTTA (서열번호 99).
- [0894] 프로브: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM]
- [0895] [THF][BHQ-1]ATACAAGGATTGGA (서열 번호 100). 여기서 FAM은 플루오레세인(fluorescein), THF는 테트라하이드로푸란(tetrahydrofuran) 및 BHQ는 블랙홀 퀸처(Black Hole Quencher)이다.
- [0896] 모든 반응은 33 mM MgOAc와 *Listeria* 지놈 DNA로부터 유래한 100카피의 DNA 주형을 추가하여 시작되었다. 그런 다음 반응을 39°C에서 인큐베이션하고 베어링 볼을 사용하여 자기 혼합과 함께 형광계에 넣었다.
- [0897] **결과 및 결론**
- [0898] 도 17에 나타낸 바와 같이, 5.5% PEG 존재 하에서 네이티브 Gp32 단백질과의 급속 증폭이 관찰되었다. 그러나, PEG가 없을 경우, 네이티브 Gp32 단백질과의 증폭은 관찰되지 않았다.
- [0899] 실시예 1 내지 5와 같이, 상술한 다른 실시예에서는, Gp32 단백질이 IDR을 포함하는 아미노산 서열로 태그된 경우에만 PEG가 없는 상태에서 Gp32 매개 증폭이 관찰되었다.
- [0900] 따라서, 본 명세서에 기재된 다른 실시예에서 제시된 데이터와 함께, 이러한 데이터는 IDR을 포함하는 아미노산 서열을 포함하는 태그가 시험관 내(*in vitro*) 생화학적 반응의 기능에 필수적인 단백질 성분에 적용되어 반응에서 크라우딩 제제에 대한 요구를 우회할 수 있고 IDR 태그 서열 없이 관찰된 효율과 비교하여 생화학적 반응의 효율을 증가시킬 수 있음을 입증한다.
- [0901] **실시예 18. 인간 Otx1 에서 유래된 IDR 태그를 갖는 Gp32를 사용한 고체 표면에서의 RPA(Recombinase polymerase amplification)**
- [0902] **실험의 목적 및 요약**
- [0903] 본 실험은 인간 호메오박스(homeobox) 단백질 Otx1의 IDR에서 발견되는 히스티딘이 풍부한 아미노산 도메인 서열을 포함하는 태그를 포함한 Gp32 융합 단백질 제제의 성능을 평가하기 위해 수행되었다.
- [0904] 본 실시예는 실시간(real-time) 및 엔드-포인트(end-point) 분석에서 크라우딩 제제 없이 히스티딘이 풍부한 IDR 도메인(Otx1)으로 C-말단에 태그된 Gp32를 사용하여 고체 표면에서 인공 핵산 주형의 RPA를 보여주었다.
- [0905] **재료 및 방법**
- [0906] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG(서열번호 24)였다. 이것은 파지 vB EcoM NBG1 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중인 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 Gp32-HIS2로 명명하였다. 융합 단백질의 전체 아미노산 서열은 서열번호 82로 나타낸다 (표 23).
- [0907] 그런 다음, 재조합 파지 vB EcoM NBG1 Gp32 융합 단백질을 PEGrk 없는 증폭에서, 즉, 크라우딩 제제 없이 고체 표면에서 테스트하였다. 테스트는 서로 다른 비율로 비드 표면에 부착된 두 개의 올리고뉴클레오타이드 프라이머를 사용하여 수행되었다. 증폭은 절단 가능한 퀀칭된 형광 프로브를 사용하여 또는 어닐링된 형광 프로브의

엔드-포인트(end-point) 검출에 의한 형광에 의해 검출되었다. 실시간(Real-time) 및 엔드-포인트(End-point) RPA 반응에서 비드는 모두 같다. 비드는 Bangs Laboratories, Inc(<https://www.bangslabs.com/>)에서 공급되었으며, 카르복실화되고 그 위에서 올리고뉴클레오타이드가 공유적으로 부착되어 성장한 하이드로겔을 갖는 폴리스티렌 코어를 갖는다.

[0908] **실시간(Real-time) RPA 반응**

[0909] 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μ M 크레아틴 카이네이즈, 1 mM dNTPs, 120 nM 프로브, 20 μ M Gp32 융합, 4.9 μ M UvsX, 7.6 μ M UvsY, 0.146 μ M 황색포도상구균 DNA 폴리머레이즈(거대 서브유닛) 및 0.34 μ M 엑소뉴클레이즈 III을 혼합하여 반응을 설정하였다. 반응 혼합물은 또한 800,000 비드/ μ l을 포함하였으며, 각 비드는 PA30 정방향 프라이머 및 PB30 역방향 프라이머의 혼합으로 구성된 비드당 약 750,000 올리고뉴클레오타이드 프라이머를 갖는다.

[0910] 상기 반응은 33.6 mM MgOAc와 TF1L이라는 인공 DNA 주형을 반응 혼합물 μ l당 800,000 주형 카피의 최종 농도로 첨가하여 개시되었다.

[0911] 관련 프라이머, 프로브 및 주형은 하기에 나타내었다.

[0912] PA30 정방향 프라이머: CCATCTCATCCCTGCGTGTCTCCGACTCAG (서열번호 105)

[0913] PB30 역방향 프라이머: CCTATCCCCTGTGTGCCTTGGCAGTCTCAG (서열번호 106).

[0914] 프로브: AGCAGAAGCAATACCGCCAGCAATAGCA[dT-FAM]G[THF]G[dT-Quencher]AGAGCGAGCTGCC (서열번호 107), 여기서 FAM은 플루오레세인(fluorescein), THF는 테트라하이드로푸란(tetrahydrofuran) 및 BHQ는 블랙홀 퀀처(Black Hole Quencher)이다.

[0915] TF1L 주형 서열:

[0916] CCATCTCATCCCTGCGTGTCTCCGACTCAGTGTTTAGGGTCCCCGGGGTTAAAGGTTTCGAACCAACAGCTGTCTGGCAGCTCGCTCTACGCATGCTATTGCTGGCGGTATTGCTTCTGCTCTTGTGGTGGCGCATGTCTAAATTGTTGGAGCTGAGACTGCCAAGGCACACAGGGGATAGG(서열번호 108).

[0917] 그런 다음 반응을 T8 형광계에서 39°C에서 30분 동안 인큐베이션하고 FAM 채널의 형광을 기록하였다.

[0918] 도 18a은 이중-프라이머 비드를 사용한 실시간(real-time) 증폭을 위해 설정된 반응 혼합물을 묘사한 그림이다. 도 18b는 실시간 반응에서 증폭된 산물을 묘사한 그림이다.

[0919] **엔드-포인트(End-point) RPA 반응**

[0920] 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1 μ M 크레아틴 카이네이즈, 1 mM dNTPs, 20 μ M Gp32 융합, 4.9 μ M UvsX, 7.6 μ M UvsY 및 0.146 μ M 황색포도상구균 DNA 폴리머레이즈(거대 서브유닛)를 혼합하여 반응을 설정하였다. 반응 혼합물은 또한 800,000 비드/ μ l을 포함하였으며, 각 비드는 PA30 정방향 프라이머 및 PB30 역방향 프라이머의 혼합으로 구성된 비드당 약 750,000 올리고뉴클레오타이드 프라이머를 갖는다.

[0921] 상기 반응은 33.6 mM MgOAc와 TF1L이라는 인공 DNA 주형을 반응 혼합물 μ l당 800,000 주형 카피의 최종 농도로 첨가하여 개시되었다.

[0922] 관련 프라이머, 프로브 및 주형은 하기에 나타내었다.

[0923] PA30 정방향 프라이머: CCATCTCATCCCTGCGTGTCTCCGACTCAG (서열번호 105)

[0924] PB30 역방향 프라이머: CCTATCCCCTGTGTGCCTTGGCAGTCTCAG (서열번호 106).

[0925] TF1L 주형 서열:

[0926] CCATCTCATCCCTGCGTGTCTCCGACTCAGTGTTTAGGGTCCCCGGGGTTAAAGGTTTCGAACCAACAGCTGTCTGGCAGCTCGCTCTACGCATGCTATTGCTGGCGGTATTGCTTCTGCTCTTGTGGTGGCGCATGTCTAAATTGTTGGAGCTGAGACTGCCAAGGCACACAGGGGATAGG(서열번호 108).

[0927] 그런 다음 반응을 39°C에서 30분 동안 인큐베이션한 다음 SDS(Sodium Dodecyl Sulfate)를 1% 최종 농도로 첨가하고 10분 동안 65°C로 가열하여 단백질을 변성시켜 반응을 중단하였다.

[0928] SDS는 물로 10배 희석하고, 볼텍싱하고, ~18,000g에서 15분 동안 원심분리한 다음 상층액을 제거하여 제거하였다. 비드를 TE pH 8.0, 0.05% Triton X-100 버퍼에 재현탁하여 대략 800,000 비드/ μ l을 얻었다.

[0929] 2 개의 형광 올리고뉴클레오타이드 프로브(PB30' 프로브 (ROX-5'-CTGAGACTGCCAAGGCACACAGGGGATAGG; 서열번호 109) 및 TF1L 프로브 (FAM-5'-GGTTTCGAACCAACAGCTG; 서열번호 110), 여기에서 ROX는 카르복시로다민(carboxyrhodamine)이고 FAM은 플루오레세인(fluorescein)임을 TE pH 8.0, 0.05% Triton X-100, 100 mM NaCl 버퍼에서 비드에 혼성화했으며, 두 프로브 모두 1 μ M 및 80,000비드/ μ l의 최종 농도에서 사용하였다. 상기 혼성화는 2분 동안 95℃로 가열한 다음, 0.1℃/초로 25℃로 냉각하여 수행하였다. 양성 대조군은 TF1L 앰플리콘이 이미 부착된 비드를 사용하여 실시하였다. 그런 다음 혼성화 화합물을 TE pH 8.0, 0.05% Triton X-100 버퍼에 6 배 희석하고 대략 18,000g에서 15분 동안 원심분리함으로써 혼성화되지 않은 프로브를 제거하고자 비드를 세척한 다음, 가능한 많은 상층액을 제거하였다. 비드는 TE pH 8.0, 0.05% Triton X-100 버퍼에 재현탁하였다. 그런 다음 반응을 T8 형광계(FAM 수준은 17%로 설정, ROX 수준은 8%로 설정)에서 39℃에서 5분 동안 인큐베이션하고 FAM 및 ROX 채널의 형광을 기록하였다.

[0930] 도 18c는 앤드-포인트(End-point) 반응에서 증폭 특성화를 묘사한 그림이다.

[0931] **결과**

[0932] **실시간(Real-time) RPA 반응**

[0933] 도 18d는 특정 엑소뉴클레아제 절단 프로브를 사용한 TF1L 앰플리콘의 실시한 형광 검출을 나타낸다. 도면에 명시된 PA30 프라이머 백분율은 PA30 올리고뉴클레오타이드인 비드-결합된 올리고뉴클레오타이드의 백분율을 나타내며, 나머지 비드-결합된 올리고뉴클레오타이드는 PB30 올리고뉴클레오타이드이다. 증폭은 모든 PA30 및 PB30 올리고뉴클레오타이드 프라이머가 비드-결합된 경우 뿐만 아니라, PA30 및 PB30 올리고뉴클레오타이드 프라이머가 액상일 때 또는 PB30이 비드-결합되고 PA30이 액상인 경우, 검출된다. 비드에 PB30만 존재하고 PA30이 없는 경우에는 증폭이 검출되지 않았다.

[0934] **앤드-포인트(End-point) RPA 반응**

[0935] TF1L 앰플리콘의 앤드-포인트(end-point) 형광 검출은 PB30 올리고뉴클레오타이드 프라이머(ROX-표지됨, 도 18e) 및 TF1L 앰플리콘(FAM-표지됨, 도 18f)에 특이적인 프로브를 사용하여 관찰되었다. 도면에 명시된 백분율은 PA30 올리고뉴클레오타이드인 비드-결합된 올리고뉴클레오타이드의 백분율을 나타내고, 나머지 비드-결합된 올리고뉴클레오타이드는 PB30 올리고뉴클레오타이드이다. 하기 표는 각 비드 유형에 대한 형광 수준, TF1L 프로브 형광 대 PB30' 프로브 형광의 비율 및 불완전한 세척으로 인한 배경 형광을 설명하기 위해 직접 부착된 TF1L 앰플리콘을 사용하여 증폭되지 않은 대조군 비드로 정규화된 동일한 비율을 나타낸다.

Beads	TF1L fluorescence	PB30' fluorescence	TF1L:PB30'	Normalised ratio
0% PA30 beads	1249.75	4372	0.29	0.17
30% PA30 beads	3194.8125	4773.9375	0.67	0.39
40% PA30 beads	3017.1875	3424.625	0.88	0.52
50% PA30 beads	2857.1875	2456.6875	1.16	0.68
60% PA30 beads	1838.75	1264.75	1.45	0.85
70% PA30 beads	1471.1875	923.5	1.59	0.93
0% PA30 beads, 400nM free PA30	2355.8125	4060.9375	0.58	0.34
TF1L Control beads	1484.6875	868.8125	1.71	1

[0936]

[0937] **결론**

[0938] 실시간(real-time) 및 앤드-포인트(end-point) 분석 모두에서 Gp32-HIS2 융합 단백질을 사용하여 PEG와 같은 크라우딩 제제 없이 핵산 증폭을 효율적으로 발생하는 것으로 밝혀졌다.

[0939] **실시예 19. IDR(Intrinsically Disordered Regions)을 포함하는 아미노산 서열의 확인**

[0940] 파지 vB EcoM NBG1 Gp32, T4 UvsY 및 T4 UvsX의 아미노산 서열은 MetaDisorder 소프트웨어 프로그램

(MetaDisorder: a meta-server for the prediction of internal disorder in protein. Kozlowski, L. P., et al., BMC Bioinformatics, 2012, 13(1): 111)을 통해 평가되었다.

[0941] 도 19a, b 및 c에 각각 나타낸 바와 같이, 파지 vB EcoM NBG1 Gp32, T4 UvsY 및 T4 UvsX의 전장 아미노산 서열은 알고리즘에 의해 분석될 때 0.5보다 큰 스코어를 갖는 아미노산 서열 스트레치를 포함하고, 따라서 IDR(Intrinsically Disordered Regions) 서열을 포함한다.

[0942] 본 실시예는 IDR 서열 또는 이의 도메인이 표준 분석 방법을 사용하여 쉽게 식별될 수 있음을 보여주었다.

[0943] **실시예 20. 인간 Otx1 유래 IDR 태그를 갖는 RB69 라이게이즈(ligase) 및 RB69 라이게이즈의 상분리 촉진 활성의 비교**

[0944] 본 실험은 인간 호메오박스(homeobox) 단백질 Otx1(His 태그)의 IDR에서 발견되는 히스티딘이 풍부한 아미노산 도메인 서열을 포함하는 태그를 포함하는 라이게이즈 융합 단백질 제제의 상분리 촉진 활성을 평가하기 위해 수행되었다.

[0945] 상기 실험은 크라우딩 제제 없이 RB69 라이게이즈-His2 에 의한 상-분리된 수성 입자(구형의 입자(globular foci))의 형성이 Mg^{2+} 농도에 의해 강화된 반면, RB69 라이게이즈에 의한 구형의 입자의 형성은 Mg^{2+} 농도에 거의 관련이 없거나 전혀 관련이 없는 것을 보여주었다.

[0946] **재료 및 방법**

[0947] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 AGHHHHHPAHHPPLSQSSGHHHHHHHHHHQGYGGSG(서열번호 24)이다. 이는 RB69 DNA 라이게이즈의 C-말단에 부착되었다. 재조합 융합 단백질 및 IDR이 없는 단백질은 테스트 중인 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘과 IDR이 없는 단백질의 C-말단의 폴리-히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 RB69 라이게이즈-His2로 명명되었고 IDR이 없는 단백질은 RB69 라이게이즈로 명명되었다. 단백질의 전체 아미노산 서열은 하기 표 24에 각각 서열번호 111 및 서열번호 112로 나타낸다.

표 24

[0948] RB69 ligase 및 RB69 ligase-His2 단백질의 전체 아미노산 서열

단백질 명칭	서열번호	아미노산 서열
RB69 ligase	111	MILDILNQIAAIGSTKTKQEILKKNKDNKLLERVYRLTYARGIQYYIKKWPGPGRSQAYGLLELDDMLDFIEFTLATRKLTGNAAIKELMGYIADGKPDDVEVLRRVMMRDLEVGASVSIANKVWPGLIQLQPQLASAYDEKLITKNIKWPAFAQLKADGARCFAEVRDDGVQFFSRAGNEYHGLTLLADELMEMTKEARERHPNGVLIDGELVYHSFDIKKAVSSGNDLSFLFGDNEESEEVQVADRSTSNGLANKSLQGTISPKEAEGMVLQAWDYVPLDEVYSDGKIKGQKYDVRFAALENMAEGFKRIEPIENQLVRNLDEAKVYKYVDQGLEGIILKNRDSYWENKRSKNLIKFKVIDIALEVVGYYEHSKDPNKLGGVELVSRCRIITDCSGSGFKDTTHKTVDGVKVLIPLDERHDLDRERLMSEAREGKIGRIADCECNGWVHSGREGTVGIFLPIIKGFRFDKTEADSFEDVFGSWDQTGVSGHHHHHPAHHPPLSQSSGHHHHHHHHHHQGYGGSG
RB69 ligase-His2	112	MILDILNQIAAIGSTKTKQEILKKNKDNKLLERVYRLTYARGIQYYIKKWPGPGRSQAYGLLELDDMLDFIEFTLATRKLTGNAAIKELMGYIADGKPDDVEVLRRVMMRDLEVGASVSIANKVWPGLIQLQPQLASAYDEKLITKNIKWPAFAQLKADGARCFAEVRDDGVQFFSRAGNEYHGLTLLADELMEMTKEARERHPNGVLIDGELVYHSFDIKKAVSSGNDLSFLFGDNEESEEVQVADRSTSNGLANKSLQGTISPKEAEGMVLQAWDYVPLDEVYSDGKIKGQKYDVRFAALENMAEGFKRIEPIENQLVRNLDEAKVYKYVDQGLEGIILKNRDSYWENKRSKNLIKFKVIDIALEVVGYYEHSKDPNKLGGVELVSRCRIITDCSGSGFKDTTHKTVDGVKVLIPLDERHDLDRERLMSEAREGKIGRIADCECNGWVHSGREGTVGIFLPIIKGFRFDKTEADSFEDVFGSWDQTGVSGHHHHHPAHHPPLSQSSGHHHHHHHHHHQGYGGSG

[0949] 본 실험에 사용된 프로브 올리고는 FAM(플루오레세인)으로 표지된 CCGCAATGGTGCCTCTCAGTACAATCTGCTCTGATG(서열번호 104)였다.

[0950] 50 μ l 용액은 1 mg/ml의 최종 농도의 라이게이즈, 50 mM NaCl, 0.4 μ M FAM-올리고 및 관련 도면에 표시된 타겟 농도의 $MgCl_2$ 로 구성되었다. 그런 다음, 10 μ l의 반응 혼합물을 C-Chip 혈구계 슬라이드로 옮기고 명시야 광 조건 및 형광 조건에서 이미지를 촬영하였다.

[0951] **결과 및 결론**

- [0952] 도 20에 나타낸 바와 같이, RB69 라이게이즈-His2는 Mg^{2+} 의 존재 하에 올리고뉴클레오타이드 프로브(형광 표지에 의해 검출됨)에서 밀집된 것으로 보이는 많은 상-분리된 수성 입자(구형의 입자)의 형성을 향상시켰다. 태그가 되지 않은 RB69 라이게이즈는 20 mM Mg^{2+} 에서도 구형의 입자 형성을 향상시키는 데 거의 효과가 없었다.
- [0953] **실시예 21. 인간 Otx1 유래 IDR 태그를 갖는 RB69 라이게이즈(ligase)의 라이게이즈 활성 성능 평가.**
- [0954] **실험의 목적 및 요약**
- [0955] 본 실험은 인간 호메오박스(homeobox) 단백질 Otx1(His2 태그)의 IDR에서 발견되는 히스티딘이 풍부한 아미노산 도메인 서열을 포함하는 태그를 포함하는 라이게이즈 효소 융합 단백질 제제의 라이게이즈 활성 성능을 평가하기 위해 수행되었다.
- [0956] 상기 실험은 RB69 라이게이즈-His2의 농도가 증가할 때 이중 결합된 산물이 증가함을 보여주었다.
- [0957] **재료 및 방법**
- [0958] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 AGHHHHHPAHHPHSQSSGHHHHHHHHHQGYGGSG(서열번호 24)였다. 이는 RB69 DNA 라이게이즈의 C-말단에 부착되었다. 제조된 IDR 융합 단백질은 테스트 중인 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 RB69 라이게이즈-His2로 명명되었다. 단백질의 전체 아미노산 서열은 상기 표 24에 나타낸다.
- [0959] 결합 주형은 pUC19 벡터(New England Biolabs)에서 증폭된 170 bp 단편(Lig170)이었다. 50 μ l 증폭 반응은 25 μ l DreamTaq Green Master Mix(Thermo Fisher Scientific), 0.2 μ M Lig170_FW 프라이머, 0.2 μ M Lig170_RV 프라이머, 1 pg pUC19를 혼합하여 설정되었다. PCR 반응은 다음과 같이 수행되었다: 95°C에서 2분; 95°C에서 30초, 55°C에서 30초 및 72°C에서 1분의 35 사이클; 그 다음 72°C에서 5분간 최종 연장한다. 증폭 산물은 2% 아가로스 겔에서 실행되었다. 타겟 DNA의 밴드는 절단하고 Monarch DNA Gel Extraction Kit(New England Biolabs)에 의해 정제되었다. DNA는 T4 폴리뉴클레오타이드 카이네이즈(T4 PNK, Thermo Fisher Scientific)에 의해 5' 말단에서 추가로 인산화되었다. 50 μ l 인산화 반응은 1x 반응 버퍼 A, 1 mM ATP, 1 U T4 PNK 및 이전 단계의 DNA를 혼합하여 설정되었다. 상기 인산화 반응은 37°C에서 30분 동안 인큐베이션되었다. 5'-인산화된 이중 가닥 DNA는 Monarch PCR & DNA Cleanup Kit(New England Biolabs)로 정제하고 Qubit dsDNA HS assay kit(Thermo Fisher Scientific)로 정량화하였다.
- [0960] 관련 프라이머 및 주형 서열은 하기에 나타내었다.
- [0961] Lig170_FW 프라이머: 5'-GAGCGCAACGCAATTAA-3' (서열번호 113).
- [0962] Lig170_RV 프라이머: 5'-ATCCGCTCACAATTCCACAC-3' (서열번호 114).
- [0963] Lig170 주형: 5'-GAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
- [0964] AGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGAATTGTGAGCGGAT-3' (서열번호 115).
- [0965] 일루미나 어댑터(illumina adaptor)는 2개의 올리고, 1.5 μ M ILMN_AD_P5 및 1.5 μ M ILMN_AD_P7rc_IDX01을 천천히 어닐링하여 준비하였다. 상기 어닐링 공정은 올리고 혼합물을 95°C까지 가열하고 0.1°C/min의 속도로 14°C까지 냉각하는 것이다.
- [0966] ILMN_AD_P5: 5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3' (서열번호 116).
- [0967] ILMN_AD_P7rc_IDX01: 5'-PO4-GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCCGTCTTCTGCTTG-3' (서열번호 117).
- [0968] RB69 라이게이즈-His2는 35 mg/ml이었고 워킹 스톡(working stock)으로 1 mg/ml로 희석하였다. 20 μ l 용액은 최종 농도 1 ng/ μ l의 T4 PNK 처리된 Lig170, 187.5 nM 일루미나 어댑터(illumina adaptor), 5% PEG 35000, 1x T4 DNA 라이게이즈 반응 버퍼(New England Biolabs) 및 최종 농도 0.1/0.2/0.3/0.4 mg/ml의 RB69 라이게이즈-His2를 포함하여 구성되었다. 결합 반응은 16°C에서 20분, 65°C에서 15분간 수행하였다. 아가로스 겔에서 시각화하기 위해 각 반응 조건에 대해 8개의 병렬 반응을 설정하고 2% 아가로스 겔에 로딩하기 전에 조합하였다. 겔 이미지는 ImageJ(National Institutes of Health)에 의해 분석되었고 밴드의 광학 밀도는 Excel(Microsoft)에

의해 플로팅되었다.

- [0969] **결과 및 결론**
- [0970] 도 21에 나타낸 바와 같이, RB69 라이게이즈-His2의 농도가 증가할 때 이중 결찰 생성물이 증가하였다. 0.4 mg/ml의 RB69 라이게이즈-His2가 20 μ l 반응에 사용되었을 때(상기 도에서 0.8 μ g), 주형의 93.5%가 일루미나 어댑터에 의해 양쪽 끝에서 결찰될 수 있었다.
- [0971] **실시예 22. 인간 Otx1 유래 IDR 태그를 갖는 RB69 라이게이즈(ligase)와 RB69 라이게이즈의 라이게이즈의 활성 성능 비교.**
- [0972] **실험의 목적 및 요약**
- [0973] 본 실험은 인간 호메오박스(homeobox) 단백질 Otx1(His2 태그)의 IDR(intrinsically disordered region)에서 발견되는 히스티딘이 풍부한 아미노산 도메인 서열을 포함하는 태그를 포함하는 라이게이즈 효소 융합 단백질 제제의 활성 성능을 평가하기 위해 수행되었다.
- [0974] 상기 실험은 태그가 지정되지 않은 RB69 라이게이즈의 효율성과 비교하여 His2 태그가 RB69 라이게이즈의 TA 결찰 효율을 상당히 증가시킬 수 있음을 보여주었다.
- [0975] **재료 및 방법**
- [0976] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 AGHHHHHPAHHHPLSQSSGHHHHHHHHHHQYGGSG(서열번호 24; 표 1)였다. 이는 RB69 DNA 라이게이즈의 C-말단에 부착되었다. 재조합 융합 단백질 및 IDR이 없는 단백질은 테스트 중인 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘과 IDR이 없는 단백질의 C-말단의 폴리-히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 RB69 라이게이즈-His2로 명명되었고 IDR이 없는 단백질은 RB69 라이게이즈로 명명되었다. 단백질의 전체 아미노산 서열은 상기 표 24에 나타난다.
- [0977] 결찰 주형은 pUC19 벡터(New England Biolabs)에서 증폭된 170 bp 단편(Lig170)이었다. 50 μ l 증폭 반응은 25 μ l DreamTaq Green Master Mix(Thermo Fisher Scientific), 0.2 μ M Lig170_FW 프라이머, 0.2 μ M Lig170_RV 프라이머, 1 pg pUC19를 혼합하여 설정되었다. PCR 반응은 다음과 같이 수행되었다: 95°C에서 2분; 95°C에서 30초, 55°C에서 30초 및 72°C에서 1분의 35 사이클; 그 다음 72°C에서 5분간 최종 연장한다. 증폭 산물은 2% 아가로스 겔에서 실행되었다. 타겟 DNA의 밴드는 절단하고 Monarch DNA Gel Extraction Kit(New England Biolabs)에 의해 정제되었다. DNA는 T4 폴리뉴클레오타이드 카이네이즈(T4 PNK, Thermo Fisher Scientific)에 의해 5' 말단에서 추가로 인산화되었다. 50 μ l 인산화 반응은 1x 반응 버퍼 A, 1 mM ATP, 1 U T4 PNK 및 이전 단계의 DNA를 혼합하여 설정되었다. 상기 인산화 반응은 37°C에서 30분 동안 인큐베이션되었다. 5'-인산화된 이중 가닥 DNA는 Monarch PCR & DNA Cleanup Kit(New England Biolabs)로 정제하고 Qubit dsDNA HS assay kit(Thermo Fisher Scientific)로 정량화하였다.
- [0978] 관련 프라이머 및 주형 서열은 하기에 나타내었다.
- [0979] Lig170_FW 프라이머: 5'-GAGCGCAACGCAATTAA-3' (서열번호 113).
- [0980] Lig170_RV 프라이머: 5'-ATCCGCTCACAATTCACAC-3' (서열번호 114).
- [0981] Lig170 주형: 5'-GAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
- [0982] AGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGAATTGTGAGCGGAT-3' (서열번호 115).
- [0983] 일루미나 어댑터(illumina adaptor)는 2개의 올리고, 1.5 μ M ILMN_AD_P5 및 1.5 μ M ILMN_AD_P7rc_IDX01을 천천히 어닐링하여 준비하였다. 상기 어닐링 공정은 올리고 혼합물을 95°C까지 가열하고 0.1°C의 속도로 14°C까지 냉각하는 것이다.
- [0984] ILMN_AD_P5: 5'-AATGATACGGCGACACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCT-3' (서열번호 116).
- [0985] ILMN_AD_P7rc_IDX01: 5'-PO4-GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCCGTCTTCTGCTTG-3' (서열번호 117).
- [0986] RB69 라이게이즈-His2는 35 mg/ml이었고 RB69 라이게이즈는 27.75 mg/ml였다. 이들은 워킹 스톱(working

stock)으로 1 mg/ml로 희석하였다. 20 μ l 용액은 최종 농도 1 ng/ μ l의 T4 PNK 처리된 Lig170, 187.5 nM 일루미나 어댑터(illumina adaptor), 5% PEG 35000, 1x T4 DNA 라이게이즈 반응 버퍼(New England Biolabs), 및 최종 농도 0.075 mg/ml의 T4 DNA 라이게이즈/RB69 라이게이즈/RB69 라이게이즈-His2를 포함하여 구성되었다. 결찰 반응은 16°C에서 20분, 65°C에서 15분간 수행하였다.

[0987] DNA는 제조업체의 지침에 따라 PCR 비드(Beckman Coulter)에 대해 0.8x AMPure XP로 정제되었다. 정제된 DNA를 25 μ l DreamTaq Green Master Mix(Thermo Fisher Scientific), 0.2 μ M ILMN_P5 프라이머, 0.2 μ M ILMN_P7 프라이머와 혼합하였다. PCR 반응은 다음과 같이 수행되었다: 95°C에서 2분; 95°C에서 30초, 55°C에서 30초 및 72°C에서 1분의 10 사이클; 그 다음 72°C에서 5분간 최종 연장한다. 증폭 산물을 PCR 비드용 1x AMPure XP로 정제되었다. 정제된 DNA는 Qubit dsDNA HS assay kit에 의해 정량화되었고 DNA의 양은 Excel(Microsoft)에 의해 플롯팅되었다.

[0988] ILMN_P5: 5'-AATGATACGGCGACCACCGA-3'(서열번호 118)

[0989] ILMN_P7: 5'-CAAGCAGAAGACGGCATACG-3'(서열번호 119)

[0990] 결과 및 결론

[0991] 도 22에 나타낸 바와 같이, 이중 결찰 생성물만이 증폭될 수 있었다. RB69 라이게이즈와 RB69 라이게이즈-His2는 모두 결찰 효율을 크게 증가시킴을 보여주었다. 특히 RB69 라이게이즈-His2는 결찰 효율을 최대 3.1배까지 향상시킬 수 있었다.

[0992] 실시예 23. NEBNext Ultra II 라이게이즈(ligase)와 인간 Otx1 유래 IDR 태그를 갖는 RB69 라이게이즈의 라이게이즈 활성 성능 비교.

[0993] 실험의 목적 및 요약

[0994] 본 실험은 인간 호메오박스(homeobox) 단백질 Otx1(His2 태그)의 IDR(intrinsically disordered region)에서 발견되는 히스티딘이 풍부한 아미노산 도메인 서열을 포함하는 태그를 포함하는 라이게이즈 효소 융합 단백질 제제와 비교하여 NEBNext Ultra II Ligation Master Mix의 활성 성능을 평가하기 위해 수행하였다.

[0995] 상기 실험은 RB69 라이게이즈-His2가 NEBNext Ultra II Ligation Master Mix와 비교하여 상당히 향상된 결찰 효율을 가지고 있음을 보여주었다.

[0996] 재료 및 방법

[0997] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 AGHHHHHPAHHPHSQSSGHHHHHHHHHHQGYGGSG(서열번호 24)였다. 이는 RB69 DNA 라이게이즈의 C-말단에 부착되었다. 제조용 융합 단백질은 테스트 중인 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 RB69 라이게이즈-His2로 명명되었다. 단백질의 전체 아미노산 서열은 상기 표 24에 나타난다.

[0998] 결찰 주형은 pUC19 벡터(New England Biolabs)에서 증폭된 170 bp 단편(Lig170)이었다. 50 μ l 증폭 반응은 25 μ l DreamTaq Green Master Mix(Thermo Fisher Scientific), 0.2 μ M Lig170_FW 프라이머, 0.2 μ M Lig170_RV 프라이머, 1 pg pUC19를 혼합하여 설정되었다. PCR 반응은 다음과 같이 수행되었다: 95°C에서 2분; 95°C에서 30초, 55°C에서 30초 및 72°C에서 1분의 35 사이클; 그 다음 72°C에서 5분간 최종 연장한다. 증폭 산물은 2% 아가로스 겔에서 실행되었다. 타겟 DNA의 밴드는 절단하고 Monarch DNA Gel Extraction Kit(New England Biolabs)에 의해 정제되었다. DNA는 T4 폴리뉴클레오타이드 카이네이즈(T4 PNK, Thermo Fisher Scientific)에 의해 5' 말단에서 추가로 인산화되었다. 50 μ l 인산화 반응은 1x 반응 버퍼 A, 1 mM ATP, 1 U T4 PNK 및 이전 단계의 DNA를 혼합하여 설정되었다. 상기 인산화 반응은 37°C에서 30분 동안 인큐베이션되었다. 5'-인산화된 이중 가닥 DNA는 Monarch PCR & DNA Cleanup Kit(New England Biolabs)로 정제하고 Qubit dsDNA HS assay kit(Thermo Fisher Scientific)로 정량화하였다.

[0999] 관련 프라이머 및 주형 서열은 하기에 나타내었다.

[1000] Lig170_FW 프라이머: 5'-GAGCGCAACGCAATTAA-3' (서열번호 113).

[1001] Lig170_RV 프라이머: 5'-ATCCGCTCACAATTCACAC-3' (서열번호 114).

[1002] Lig170 주형: 5'-GAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT

- [1003] AGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGAT-3' (서열번호 115).
- [1004] 일루미나 어댑터(illumina adaptor)는 2개의 올리고, 1.5 μ M ILMN_AD_P5 및 1.5 μ M ILMN_AD_P7rc_IDX01을 천천히 어닐링하여 준비하였다. 상기 어닐링 공정은 올리고 혼합물을 95℃까지 가열하고 0.1℃/min의 속도로 14℃까지 냉각하는 것이다.
- [1005] ILMN_AD_P5: 5'-AATGATACGGCGACACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCT-3' (서열번호 116).
- [1006] ILMN_AD_P7rc_IDX01: 5'-PO4-GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCCGTCTTCTGCTTG-3' (서열번호 117).
- [1007] RB69 라이게이즈-His2는 35 mg/ml이었고 워킹 스톡(working stock)으로 1 mg/ml로 희석하였다. 93.5 μ l 용액은 10 ng의 T4 PNK 처리된 Lig170, 187.5 nM 일루미나 어댑터(illumina adaptor), 30 μ l NEBNext Ultra II Ligation Master Mix, 및 1 μ l NEBNext Ligation Enhancer를 포함하여 구성되었다. 결찰 반응은 20℃에서 15분간 수행하였다. 상기 결찰 반응은 6.5 μ l의 0.5 M EDTA를 첨가하여 종료되었고, 제조사의 지침에 따라 PCR 비드를 위한 0.8x AMPure XP(Beckman Coulter)로 DNA를 정제하였다. 93.5 μ l 용액은 10 ng의 T4 PNK 처리된 Lig170, 187.5 nM 일루미나 어댑터(illumina adaptor), 5%/7% PEG 35000, 1x T4 DNA 라이게이즈 반응 버퍼(New England Biolabs) 및 최종 농도 0.2/0.3/0.4 mg/ml의 RB69 라이게이즈-His2를 포함하여 구성되었다. 결찰 반응은 16℃에서 20분간 수행하였다. 상기 결찰 반응은 4.5 μ l의 0.5 M EDTA 및 2 μ l 프로테아제 K(Protease K)(New England Biolabs)를 첨가하여 종료되었고, 40℃에서 30분간 인큐베이션하였다. 제조사의 지침에 따라 PCR 비드를 위한 0.8x AMPure XP(Beckman Coulter)로 DNA를 정제하였다. 정제된 DNA를 25 μ l DreamTaq Green Master Mix(Thermo Fisher Scientific), 0.2 μ M ILMN_P5 프라이머, 0.2 μ M ILMN_P7 프라이머와 혼합하였다. PCR 반응은 다음과 같이 수행되었다: 95℃에서 2분; 95℃에서 30초, 55℃에서 30초 및 72℃에서 1분의 10 사이클; 그 다음 72℃에서 5분간 최종 연장한다. 증폭 산물은 PCR 비드를 위한 1x AMPure XP로 정제하였다. 정제된 DNA는 Qubit dsDNA HS assay kit에 의해 정량화되었고 DNA의 양은 Excel(Microsoft)에 의해 플롯팅되었다. DNA는 또한 2% 아가로스 젤에 의해 분석되었다.
- [1008] ILMN_P5: 5'-AATGATACGGCGACACCGA-3' (서열번호 118)
- [1009] ILMN_P7: 5'-CAAGCAGAAGACGGCATACG-3' (서열번호 119)
- [1010] **결과 및 결론**
- [1011] 이중 결찰된 제품만 증폭될 수 있었다. 도 23에서 나타난 바와 같이, R69 라이게이즈-His2는 NEBNext Ultra II Ligation Master Mix와 비교하여 최대 2.8배 향상된 결찰 효율을 가질 수 있었다.
- [1012] **실시예 24. 인간 Otx1 유래 IDR 태그를 갖는 RB69 라이게이즈(ligase)의 상 분리 성능에 대한 ATP의 영향 분석.**
- [1013] **실험의 목적 및 요약**
- [1014] 본 실험의 목적은 상 분리를 일으키는 라이게이즈 효소 융합 단백질 제제의 능력에 대한 ATP의 영향을 분석하는 것이었다. 라이게이즈 효소 융합 단백질은 인간 호메오박스(homeobox) 단백질 Otx1(His2 태그)의 IDR(intrinsically disordered region)에서 발견되는 히스티딘이 풍부한 아미노산 도메인 서열을 포함하는 태그를 갖는다.
- [1015] 본 실험은 ATP가 His2 태그에 의해 매개되는 상 분리를 상당히 향상시키는 것을 보여주었다.
- [1016] **재료 및 방법**
- [1017] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 AGHHHHHPAHHPPLSQSSGHHHHHHHHHQYGGSG(서열번호 24; 표 1)였다. 이는 RB69 DNA 라이게이즈의 C-말단에 부착되었다. 재조합 융합 단백질 및 IDR이 없는 단백질은 테스트 중인 융합 단백질의 IDR 도메인 태그에 자연적으로 존재하는 히스티딘과 IDR이 없는 단백질의 C-말단의 폴리-히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 RB69 라이게이즈-His2로 명명되었고 IDR이 없는 단백질은 RB69 라이게이즈로 명명되었다. 단백질의 전체 아미노산 서열은 상기 표 24에 나타난다. FAM-올리고는 FAM(플루오레세인)으로 표지된 CCGCAATGGTGCCTCTCAGTACAATCTGCTCTGATG(서열번호 104)였다.
- [1018] 50 μ l 용액은 최종 농도 1 mg/ml의 라이게이즈, 0.4 μ M FAM-올리고 및 0/20 mM MgCl₂ 및 0/1 mM ATP를 포함하

여 구성되었다. 그런 다음, 10 μ l의 반응 혼합물을 C-Chip 혈구계 슬라이드로 옮기고 명시야 광 조건 및 형광 조건에서 이미지를 촬영하였다.

[1019] **결과 및 결론**

[1020] 도 24에 나타난 바와 같이, IDR이 없는 RB69 라이게이즈를 사용하여 상 분리된 수성 입자(구형의 입자)가 거의 관찰되지 않았다. RB69 라이게이즈-His2는 1 mM ATP 존재 하에 많은 구형의 입자의 형성을 유의하게 촉진하였다. 20 mM Mg^{2+} 를 첨가한 후, 구형의 입자의 더 밝은 형광이 관찰되었고, 이는 구형의 입자 형성의 추가적인 향상 및/또는 더 많은 DNA가 구형의 입자로 공존하도록 하는 것을 나타낸다.

[1021] **실시예 25 . 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1에서 유래된 IDR 태그를 갖는 Gp32를 사용한 고체 표면에서의 RPA(Recombinase Polymerase Amplification)**

[1022] **실험의 목적 및 요약**

[1023] 본 실험은 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1 단백질의 IDR(intrinsically disordered region)을 포함하는 태그를 포함하는 Gp32 융합 단백질 제제의 고체 표면에서의 증폭 능력을 평가하기 위해 수행되었다.

[1024] 본 실시예는 실시간(real-time) 및 엔드-포인트(end-point) 분석에서 크라우딩 제제 없이 사카로미세스 세레비시아(*Saccharomyces cerevisiae*) Hrp1 단백질의 IDR으로 C-말단에 태그된 Gp32를 사용한 고체 표면에서 인공 핵산 주형의 RPA(Recombinase Polymerase Amplification)를 보여주었다.

[1025] **재료 및 방법**

[1026] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 GGNNGGNNMNRGGNFGNGDFNQMYQNPMGGYNPMNPQAMTDYYQKMQEYYQQMQ (서열번호 9)였다. 이는 T4 파지 Gp32의 C-말단에 부착되었다. 재조합 융합 단백질은 테스트 중인 융합 단백질의 바로 C-말단에 배치된, 즉, 융합 단백질의 C-말단에서 IDR 태그 뒤에 배치된, 추가 헵타히스티딘 태그에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 T4-Gp32-HRP1으로 명명되었다. 상기 융합 단백질의 전체 아미노산 서열은 하기 하기 서열번호 120으로 나타낸다.

[1027] MFKRKSTAEALAAQMAKLNKGFSSEDKGEWKLKLDNAGNGQAVIRFLPSKNDEQAPFAILVNHGFKKNGKWIETCSSTHGDYDSCPVCQYISKNDLYNTDNKEYSLVKRKTSYWANILVVKDPAAPENEGKVFYRFGKKIWDKINAMIAVDVEMGETPVDVTCPWEGANFVLKVKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQMVLDLSEMTSKDKFKSFEELNTKFGQVMGTAVMGAAATAAKKADKVADLDADFNVDDFNTKTEDDFMSSSSGSSSSADDTLDDLNDLGGNNGNNMNRGGNFGNGDFNQMYQNPMGGYNPMNPQAMTDYYQKMQEYYQQMQHHHHHHH(서열번호 120)

[1028] 그런 다음, 재조합 T4 파지 Gp32 융합 단백질을 PEG가 없는 증폭, 즉 크라우딩 제제 없이 고체 표면에서 테스트하였다. 비드 표면에 부착된 2개의 올리고뉴클레오타이드 프라이머를 사용하여 테스트를 수행하였다. 엔드-포인트(end-point) 반응을 사용하여 앰플리콘에서 발견된 니케이즈(nickase) 부위에 형광 뉴클레오타이드를 혼입함으로써 증폭을 검출하였다. 비드는 Bangs Laboratories, Inc(<https://www.bangslabs.com/>)에서 공급되었으며 카르복실화되고 올리고뉴클레오타이드가 공유적으로 부착된 공중합체로 그래프트된 폴리스티렌 코어를 갖는다. 포토리소그래피(photolithography), 소프트 리소그래피(soft lithography), 에칭(etching) 등과 같은 표준 미세 가공 기술(standard microfabrication technology)을 사용하여 개별 영역으로 패터닝된 유리 기반에 비드를 증착하였다. 결과 영역은 소수성 또는 친수성과 같은 특성을 가지며 분석할 샘플을 끌어당기거나 밀어낼 수 있다. Grace Bio-Labs FlexWell[®] 착탈식 인큐베이션 챔버를 사용하여 패터닝된 유리의 단일 조각을 6.5mm x 6.5mm 크기의 8개의 반응 챔버로 나누었고, 각각은 표면에 정렬된 배열로 1225만 개의 비드를 포함하는 것으로 추정된다. 모든 주형이 비드의 프라이머에 혼성화되는 경우 반응 챔버가 비드당 0, 5, 10, 20, 40 또는 80개의 주형 카피를 갖도록, 혼성화가 100% 효율보다 훨씬 낮다는 가정 하에 다양한 양의 증폭 주형을 다른 반응 챔버에 추가하였다. 단일 가닥 DNA 주형(UP1-UP2'_TF1L 주형 서열: AATGATACGGCGACACCGTGATCTACACTGTTTTACAACCTCAGCATGGAAAAAGGTTTCGAACCTAACAGCTGTCTGGCAGCTCGCTCTACGCATGCTAT TGCTGGCGGTATTGCTTCTGCTCTTGTGCTGGTGGCGCATGTCTAAATTGTCGATACATCTCGTATGCCGTCTTCTGCTTG (서열번호 121))을 50 μ l 버퍼(10 mM Tris HCl pH 8.0, 1 mM EDTA, 0.05% Triton X-100, 100 mM NaCl)의 반응 챔버에 추가하고 FlexWell[®] SealStrips[®]로 덮고 1시간 동안 50°C로 가열하여 주형이 비드의 상보적 올리고뉴클레오타이드에 어닐링되도록 한다. 그런 다음, 과량의 버퍼를 제거하고 비드를 TBM 버퍼(10 mM Tris HCl pH 8.0, 10 mM $MgCl_2$,

0.05% Triton X-100)로 2회 세척한 다음 반응 버퍼(25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT)로 2회 세척하여 어닐링되지 않은 주형을 제거하였다.

[1029] 반응은 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM 포스포크레아틴, 1.7 μ M 크레아틴 카이네이즈, 1 mM dNTP, 6.6 μ M Gp32 융합, 2.7 μ M UvsX, UvsY, 0.22 μ M 황색포도상구균 DNA 폴리머레이즈 및 23 mM MgOAc를 혼합하여 설정되었다. 반응 혼합물은 또한 290,000 비드/mm²를 포함했으며, 각 비드에는 0.59 μ l 반응 혼합물/mm²가 포함되어 있으며, 이는 하기 서열을 갖는 UP1 정방향 프라이머와 UP2-18 역방향 프라이머의 혼합으로 구성된 비드당 대략 600,000개의 올리고뉴클레오타이드 프라이머를 갖는다.

[1030] UP1 정방향 프라이머: AATGATAACGGCGACCACCGAGATCTACAC(서열번호 122).

[1031] UP2-18 역방향 프라이머: CAAGCAGAAGACGGCATA(서열 번호 123).

[1032] 그런 다음 반응을 43°C에서 60분 동안 인큐베이션한 다음 단백질을 변성시키기 위해 STTM 버퍼(10 mM Tris HCl pH 8.0, 10 mM MgCl₂, 0.05% Triton X-100 및 최종 농도 1%의 Sodium Dodecyl Sulfate(SDS))로 2회 세척(추가/제거)하여 반응을 중단하였다. 그런 다음 비드에서 SDS를 제거하기 위해 TTM 버퍼(10 mM Tris HCl pH 8.0, 1 mM EDTA, 0.05% Triton X-100)를 첨가/제거하여 2회 세척하였다.

[1033] 그런 다음 비드를 2.5 U 닉케이즈(nickase) Nt. BbvCI가 포함된 25 μ l 1x CutSmart 버퍼(NEB - 50 mM KOAc, 20 mM Tris-acetate, 10 mM MgOAc, 100 μ g/ml BSA, pH 7.9)로 덮었다. UP1-UP2'-TF1L주형은 DNA의 한 가닥에 닉(nick)을 도입하는 Nt. BbvCI(CC/TCAGC) 인식 부위의 단일 사본이 포함한다. 앰플리콘에 닉이 도입되는 것(were nicked)을 확인하기 위해 비드를 45분 동안 37°C로 가열하였다. 닉케이즈(nickase)를 변성시키기 위해 STTM 버퍼(10 mM Tris HCl pH 8.0, 1 mM EDTA, 0.05% Triton X-100 및 최종 농도 1%의 Sodium Dodecyl Sulfate(SDS))로 2회 세척(추가/제거)하여 닉킹(nicking)을 중단하였다. 그런 다음 비드에서 SDS를 제거하기 위해 TTM 버퍼(10 mM Tris HCl pH 8.0, 1 mM EDTA, 0.05% Triton X-100)를 첨가/제거하여 2회 세척하고 반응 버퍼(25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT)로 2회 세척하였다.

[1034] 프로토콜의 마지막 단계는 형광 표지된 dUTP를 앰플리콘에 통합하는 것이었다. 이는 0.11 μ M 황색포도상구균 DNA 폴리머레이즈, 160 pM Aminoallyl-dUTP-XX-ATTO-594(Jena Bioscience) 및 23 mM MgOAc가 포함된 반응 버퍼(25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT)에 비드를 담그고 43°C에서 45분 동안 가열하여 수행하였다. 닉케이즈(nickase)를 변성시키기 위해 STTM 버퍼(10 mM Tris HCl pH 8.0, 1 mM EDTA, 0.05% Triton X-100 및 최종 농도 1%의 Sodium Dodecyl Sulfate(SDS))로 2회 세척(추가/제거)하여 연장을 중단하였다. 그런 다음 비드에서 SDS를 제거하기 위해 TTM 버퍼(10 mM Tris HCl pH 8.0, 1 mM EDTA, 0.05% Triton X-100)를 첨가/제거하여 2회 세척하였다. 그 다음으로 Flexwells를 제거하고 글라스 커버슬립과 소량의 TTM 버퍼를 유리 웨이퍼(wafer)에 놓았다. 동일한 위치에서 촬영한 명시야 및 형광 사진과 함께 형광 현미경을 사용하여 웨이퍼를 검사하였다.

[1035] **결과**

[1036] UP1-UP2'-TF1L 앰플리콘의 앤드-포인트(end-point) 형광 검출은 dUTP로 표지된 ATTO-594를 앰플리콘의 Nt. BbvCI 닉킹 부위에 통합함으로써 관찰되었다. 도 25는 주형이 추가되지 않은 경우 비드가 어렵게 유지되었지만 주형이 증가함에 따라 비드의 비율이 증가하고 형광 수준이 증가했음을 나타낸다.

[1037] **결론**

[1038] PEG와 같은 크라우딩 제제(crowding agent) 없이, 핵산 증폭은 표면에 T4 Gp32-Hrp1 융합 단백질을 사용하여 효율적으로 발생하는 것으로 밝혀졌다.

[1039] **실시예 26. Cas12a 단백질 성능에 대한 상 분리에 연관된 Gp32-HRP1 ssDNA 결합 단백질의 효과 분석.**

[1040] **실험의 목적 및 요약**

[1041] 본 실험의 목적은 가이드 RNA와 관련하여, 형광 관독으로 모니터링되는 이중-가닥 DNA 표적에 결합하고 절단하기 위해 Cas12a 뉴클레이즈 단백질의 활성화에 대한 상 분리를 촉진하는 조건에서 폴리에틸렌 글리콜(Polyethylene Glycol, PEG)35K의 존재 하에 IDR-태그된 Gp32 ssDNA 결합 단백질의 효과를 분석하는 것이다. 사용된 Gp32 ssDNA 결합 단백질은 효모 HRP 단백질(Gp32-HRP1)의 IDR(intrinsically disordered region)에서 발견되는 아미노산 도메인 서열을 포함하는 태그를 갖는다. 상기 태그와 PEG의 존재 하에, 단백질의 낮은 농도에서 다른 요인이 없을 때 상 분리가 실질적으로 발생한다.

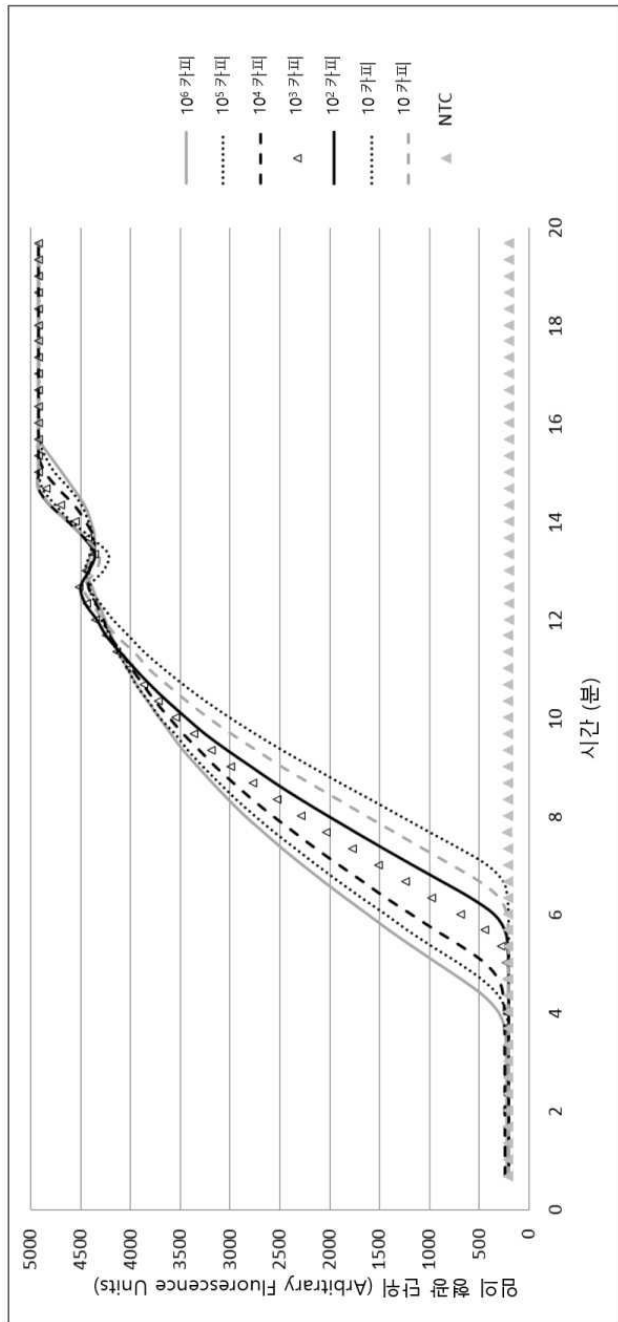
- [1042] 이 경우 Cas12a에 대한 이중 핵산 표적은 절단될 때 6-FAM 표지를 포함하는 뉴클레오타이드 단편을 생성하는 6-FAM/BHQ1 쌍을 가지며, 6-FAM은 본질적으로 어닐링된 하이브리드로부터 즉시 용해되어 측정 가능한 형광 증가를 유도한다. 이 주형은 추가로 연결할 단일-가닥 영역에 의해 Gp32-HRP1과 상호작용하도록 추가로 설계되었다.
- [1043] 상기 실험은 PEG35K의 존재 하에 Gp32-HRP1 ssDNA 결합 단백질의 사용이 상 분리된 수성 입자(구상 입자 또는 구형의 입자)의 형성을 유도하고 동시에 시험관 내(*in vitro*) 시스템에서 Cas12a가 이의 DNA 표적을 절단하는 속도를 상당히 향상시킨다는 것을 보여주었다.
- [1044] **재료 및 방법**
- [1045] 사용된 IDR 도메인 태그의 특정 아미노산 서열은 GGNGGNNMNRGGNFGNQGFNQMYQNPMGGYNPMMNPQAMTDYYQKMQEYYQQMQ(서열번호 9)였다. 이는 T4 Gp32 ssDNA 결합 단백질의 C-말단에 부착되었다. 제조된 IDR 융합 단백질은 IDR 태그의 C-말단에 추가된 7개의 추가 히스티딘에 의존하는 표준 1-단계 고정 금속(니켈) 친화성 크로마토그래피(standard 1-step immobilised metal (nickel) affinity chromatography)를 사용하여 정제되었다. 융합 단백질은 T4 Gp32-HRP1으로 명명되었다. 단백질의 전체 아미노산 서열은 서열번호 120으로 나타낸다.
- [1046] 가이드-RNA 서열은 5'- UAAUUUCUACUGUUGAGAUAAAGUGCUCAUCAUUGGAAAACG-3' (서열번호 134)이다.
- [1047] 이중-가닥/단일-가닥 DNA 표적은 2개의 올리고, 즉 5' 말단에 FAM(플루오레세인)으로 표지된 상부 올리고 5'-GAACGTTTCCAATGATGAGCACTTTTAAAGTCTATGTATCAAGCGGCCATTTGCGG-3' (서열번호 135) 및 3' 말단에 BHQ-1(소광제(quencher))로 표지된 하부 올리고 5'-AGAACTTTAAAGTGCTCATCTGGAAAACGTTC-3' (서열번호 136)를 어닐링하여 준비하였다. 어닐링 공정은 1 μ M 올리고 혼합물을 95°C까지 가열하고 0.1°C/min의 속도로 14°C까지 냉각하는 것이다. 이는 제공된 가이드 RNA와 함께 cas12a 뉴클레아제에 대한 이중 표적 부위를 제공하지만, 약 3개의 단백질 단량체의 예상되는 결합으로 Gp32-HRP1과 상호작용할 수 있는 추가 24개의 단일-가닥 잔기 또한 제공한다. 이러한 방식으로, 상-분리된 Gp32-HRP1 구상 입자(globules)가 발생할 경우, 어닐링된 표적의 많은 부분은 상-분리된 Gp32-HRP1 구상 입자에 위치할 것으로 예상되었다. 또한 어닐링될 때 근접해야 하지만 절단(하이브리드는 길이가 몇 개의 뉴클레오타이드인 결과로써) 후 분산되어야 하는 두 가닥의 형광단(fluorophore) 및 소광제(quencher)의 존재는 절단 속도를 평가하는 편리한 메커니즘을 제공한다. 예상대로, Cas12a-의존적 방식으로 형광은 일반적으로 낮은 수준에서 변화하고 시간이 지남에 따라 증가한다.
- [1048] EnGen Lba Cas12a 단백질은 New England Biolabs에서 구입하였다.
- [1049] 용액은 Cas12a 단백질, PEG35K 또는 T4 GP32 HRP1 단백질을 포함하거나 포함하지 않고 구성되었다. 상기 용액은 다음으로 구성되었다: 30 mM NaCl, 10 mM Tris Acetate pH8.3, 20 mM Mg 아세테이트, 0.1 mg/ml BSA, 33.3 nM 가이드 RNA, 50 nM dsDNA, 5% PEG35K. 반응에 포함될 때 다음 성분이 다음 농도로 포함되었다: 33.3 nM Cas12a 단백질, 333 ng/ μ l T4 GP32 HRP1 단백질. 반응 속도 양상을 평가하기 위해 30 μ l의 반응 용액을 0.2 ml 튜브에 옮기고 42°C의 실험 온도에서 Axxin T16 형광 판독기를 사용하여 분석하였다. 별도로 20 μ l의 반응 용액을 42°C에서 약 1분간 데운 후 C-Chip 혈구계 슬라이드로 옮기고 명시야 광 조건 및 형광 조건에서 이미지를 촬영하였다. 따라서 상기 이미지는 반응의 처음 몇 분 내에 시스템의 미시적 상태의 스냅샷을 나타냅니다.
- [1050] **결과 및 결론**
- [1051] 도 26a에 나타낸 바와 같이, Cas12a 단백질 없이 Gp32-HRP1 및 PEG의 존재 하에, 상-분리된 수성 입자(구형의 입자)가 관찰되었지만 Gp32-HRP1 및 PEG의 존재 하에(반응이 현미경으로 분석된 초기 시점에서) 표지되었지만 대부분 소광된(quenched) 표적 DNA가 존재할 때 약한 형광만을 보여주었다. 자연스럽게, 물론 형광단(fluorophore)/소광제(quencher) 프로브의 불완전한 어닐링 및 그에 따른 배경으로 인해 일부 형광은 관찰되지만, 형광 플롯의 평면 특성(도 26b)에서 나타낸 바와 같이 시간 경과에 따른 형광 변화의 최소 변화에 의해 뒷받침되는 Cas12a 단백질이 없으면 절단되지 않을 것으로 예상된다.
- [1052] Cas12a 단백질이 존재하지만, T4 Gp32 HRP1이 없는 경우에는, 구형의 입자가 관찰되지 않았으며, 이는 구상 입자의 형성을 가능하게 하는 T4 Gp32 HRP1에 대한 예측된 요구사항을 나타낸다. 카이네틱 분석(kinetic analysis)에 따르면 최대 10분까지 평가한 결과 표적 절단이 시간이 지남에 따라 꾸준히 증가했음을 나타내었다. 현미경적으로 전체 형광은 Cas12-마이너스 샘플에서보다 약간 높게 나타났는데, 이는 몇 분 이내에 일부 어닐링된 프로브가 카이네틱 연구(kinetic study)와 일치하게 처리되었음을 나타내었다.
- [1053] 그러나 뚜렷하고 현저한 대조에서, Cas12a 및 T4 Gp32-HRP1(및 PEG)의 존재 하에, 많은 구형의 입자가 관찰되었

고 일반적으로 현미경 이미지 전체에서 훨씬 더 강한 형광이 관찰되었으며, 이는 구상 입자 형성에 대한 T4 Gp32-HRP1의 필요성을 시사하지만, 이는 추가로 더 많은 처리를 유도하는 것을 나타낸다(일단 처리되면 작은 방출된 산물이 더 이상 구상 입자로 국소화될 것으로 예상되지 않음을 나타낸다). 카이네틱 그래프(kinetic graph)는 이러한 조건에서도 현저하고 놀라운 정도로 다를 뿐 아니라, 1분 전후(샘플을 관독기에 넣은 직후)에 피크까지 매우 빠른 형광 축적을 보였다가 나머지 분석 시간 동안 평탄한 상태(plateau)를 보여주었다. T4 Gp32-HRP1의 존재 하에 관찰된 DNA 절단 속도의 이러한 상당한 향상에 대해 본 발명자는 상-분리된 입자가 특이적 절단을 현저하게 촉진했다는 개념과 일치한다고 제안하며, 이는 아마도 Cas12a 및 그의 ds DNA 표적이 크게 증가된 반응 속도를 가능케 하는 구형의 입자 내부에서 공존함으로써 야기된다. 본 명세서에 설명된 증폭 시스템과 유사한 방식으로, 이는 단일 시스템 구성 요소만 상 분리를 유도하는 작용을 하는 경우에도, 다른 참가 요소가 국소적인 고농도 및 대규모의 가속화된 카이네틱(kinetics)를 유도하는 단계로 이끌 수 있음을 나타낸다.

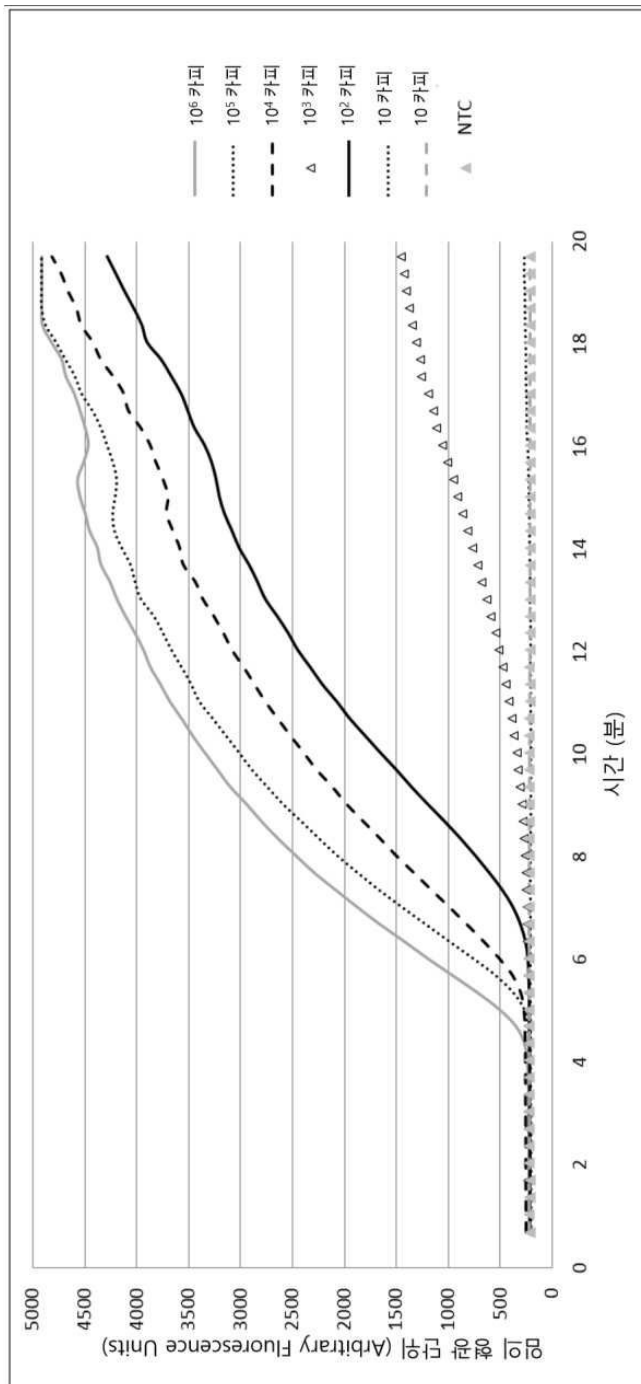
- [1054] 개시된 IDR-기반 방법, 공정, 거대분자(macromolecules), 폴리펩타이드 및 용도의 상이한 적용은 당업계의 구체적인 요구에 따라 맞추어질 수 있음을 이해해야 한다. 본 명세서에서 사용된 용어는 단지 본 발명의 특정 실시 예를 설명하기 위한 것일 뿐, 제한하려는 의도가 아님을 이해해야 한다.
- [1055] 본 명세서 및 첨부된 특허청구범위에 사용된 바와 같이, 단수형 "a", "an" 및 "the"는 내용이 명백하게 달리 지시하지 않는 한 복수의 지시 대상을 포함한다. 따라서, 예를 들어, "거대분자(macromolecules)", "폴리펩타이드", "폴리뉴클레오타이드", "세포", "숙주 세포" 등과 같은 실체에 대한 언급은 둘 이상의 이러한 실체를 포함한다.
- [1056] "약" 및 "대략"과 같은 용어는 내용이 달리 명시하지 않는 한 수치 값의 +/- 10% 또는 수치 값의 +/- 5% 관련 수치를 포함하는 것으로 이해해야 한다.
- [1057] 숫자의 범위가 하한값과 상한값 "사이"로 표시되는 경우 범위는 상한값과 하한값을 포함하는 것으로 해석되어야 한다. 예를 들어, 22 mM 내지 50 mM, 또는 약 22 mM 내지 약 50 mM의 범위는 22 mM 내지 50 mM의 값 또는 약 22 mM 내지 약 50 mM의 값을 포함하는 것으로 해석되어야 한다.
- [1058] 본 명세서에 인용된 모든 간행물, 특허 및 특허 출원은 그 전체가 참고로 여기에 포함된다.

도면

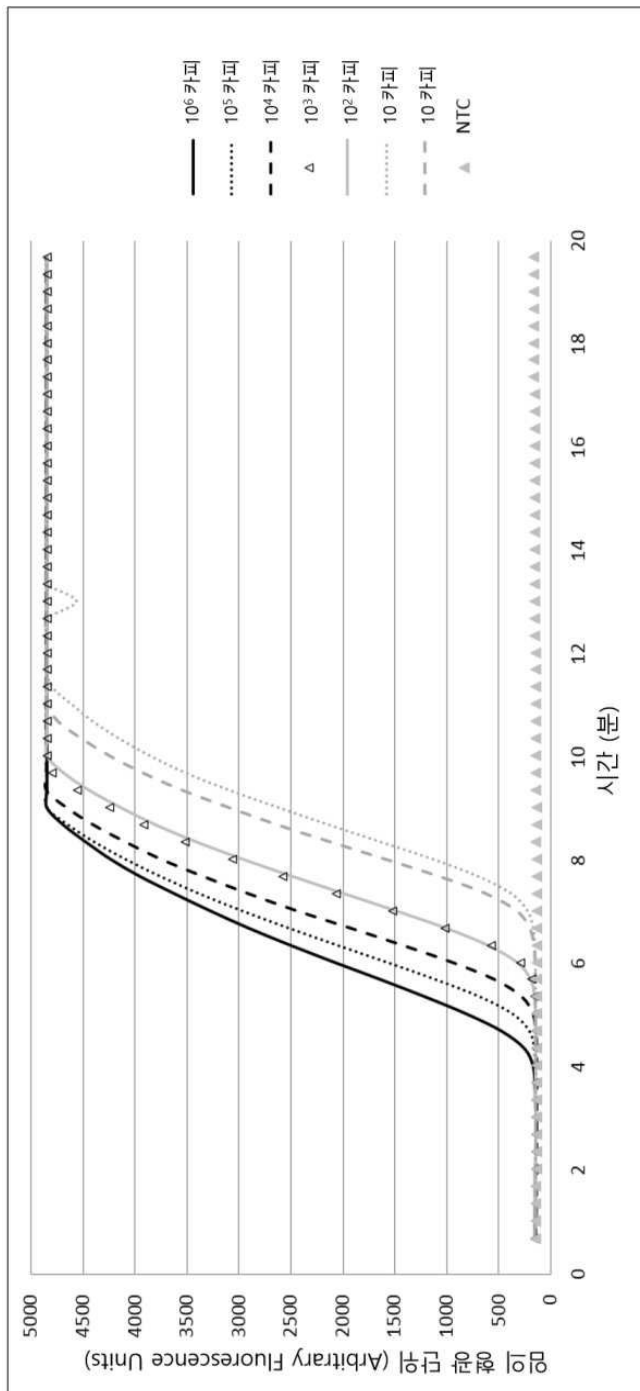
도면1



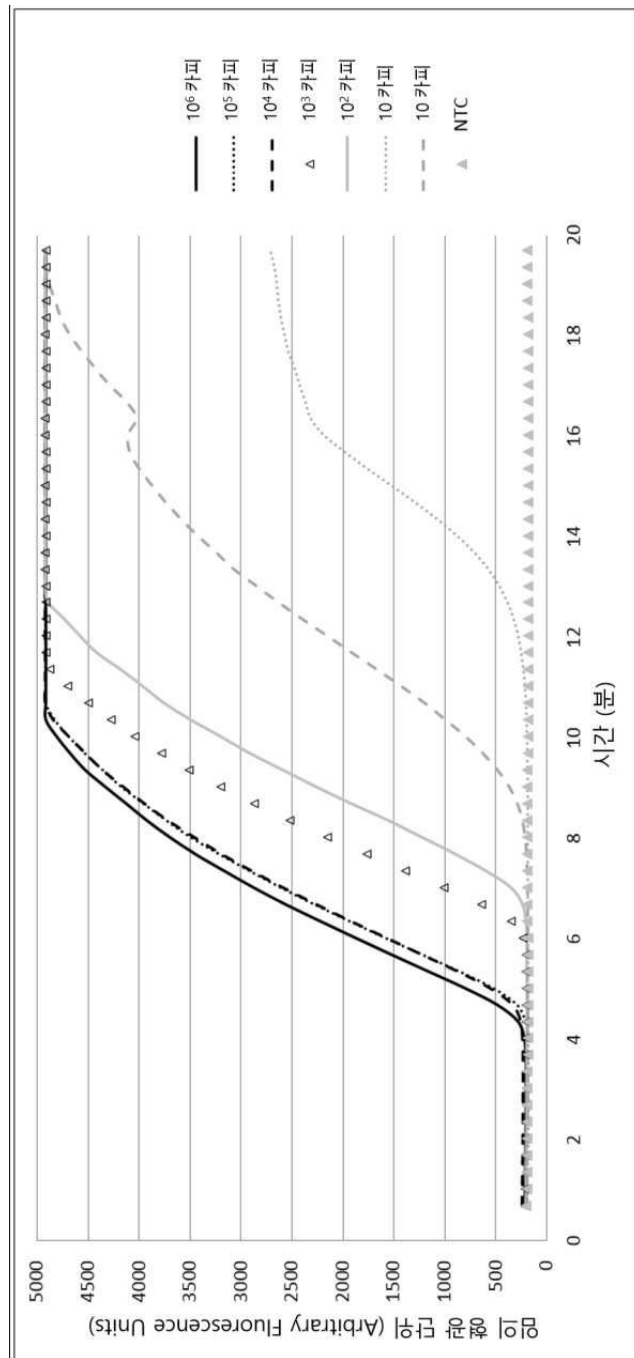
도면2



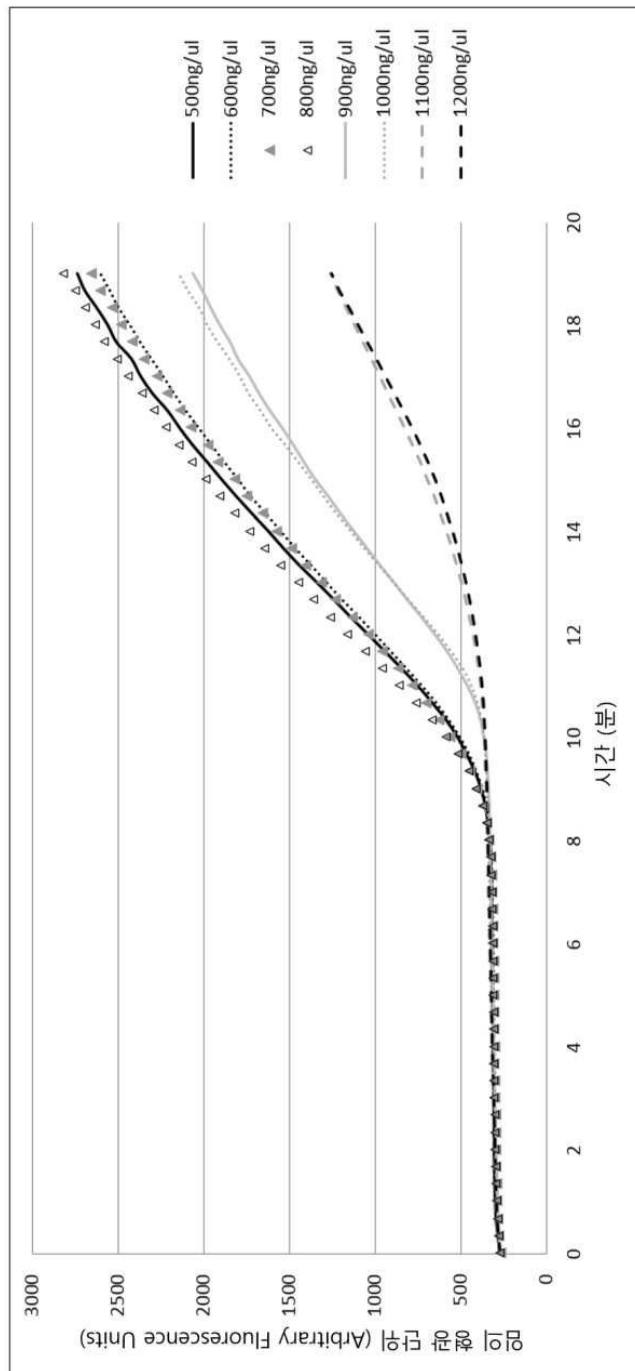
도면3



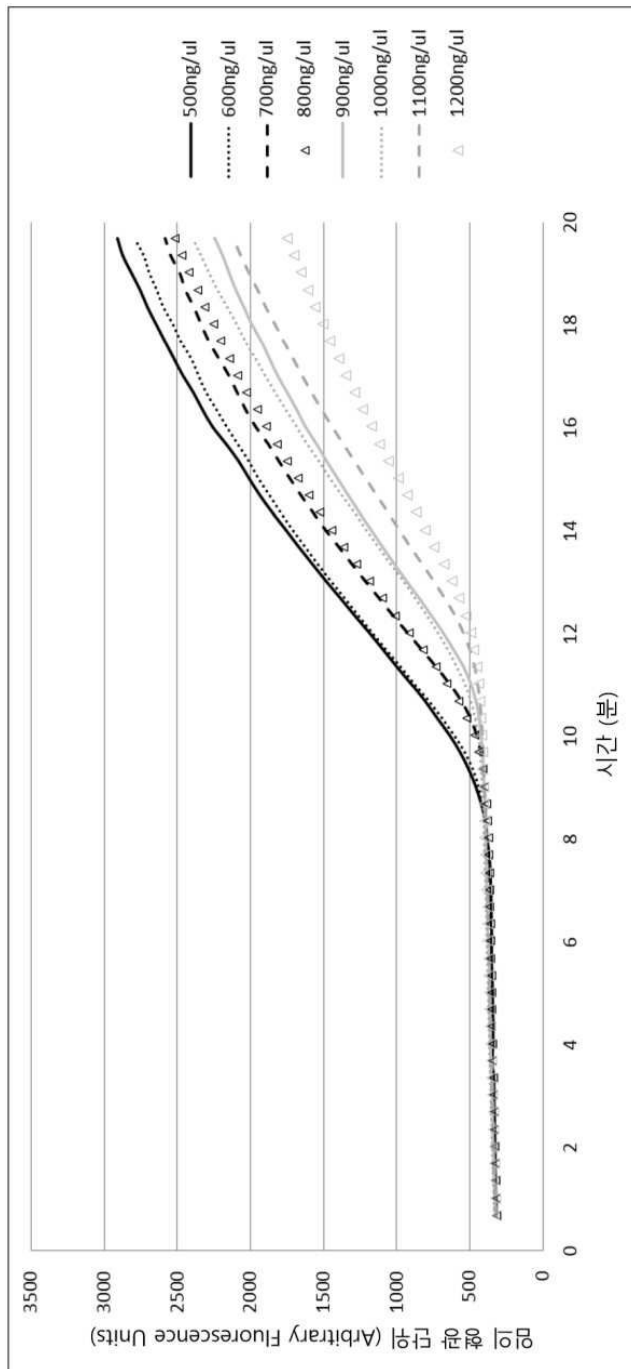
도면4



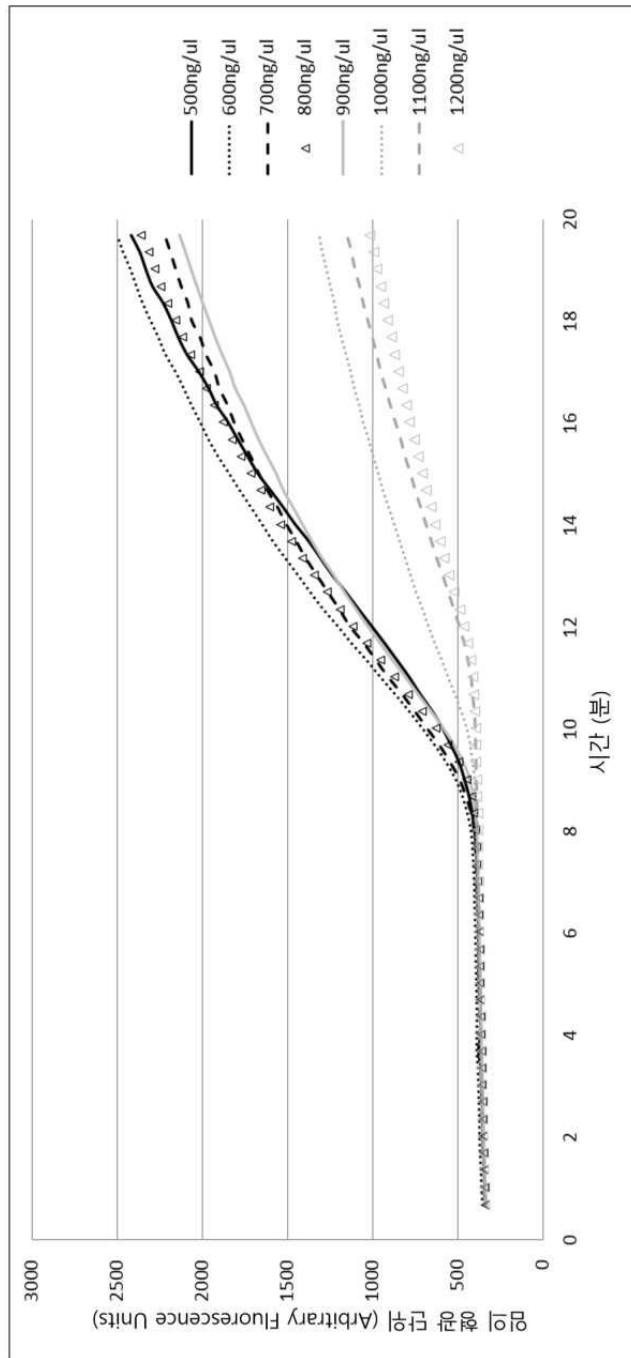
도면5a



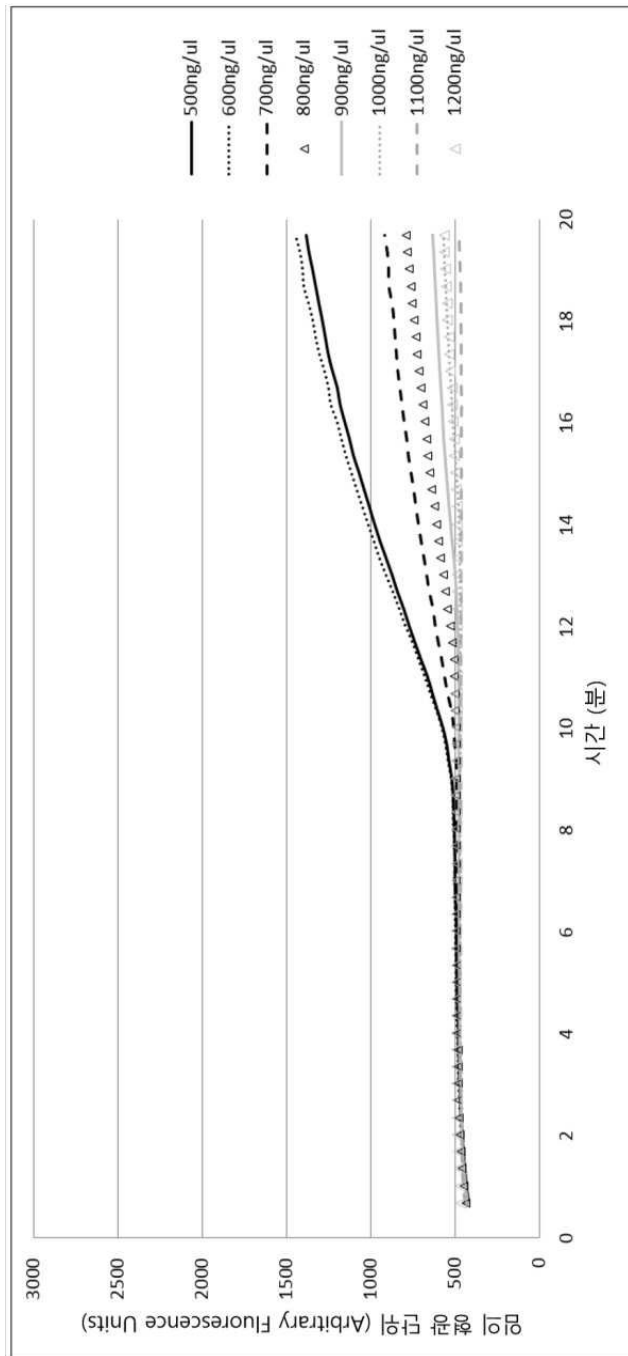
도면5b



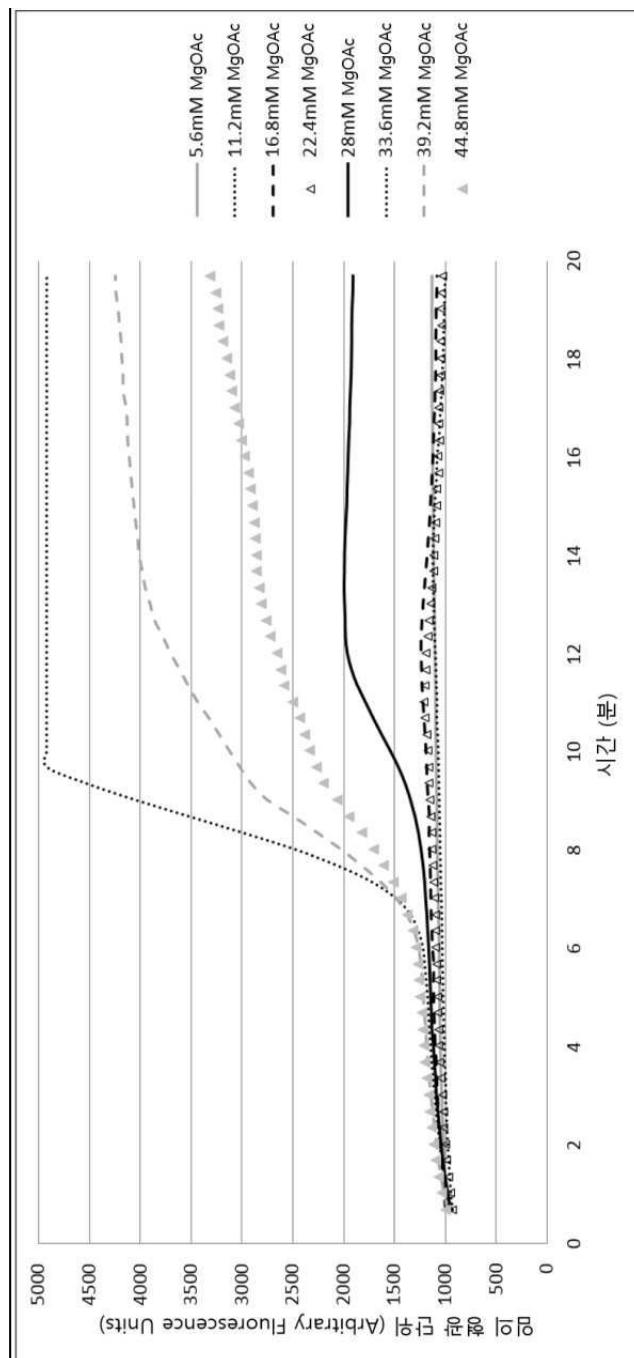
도면5c



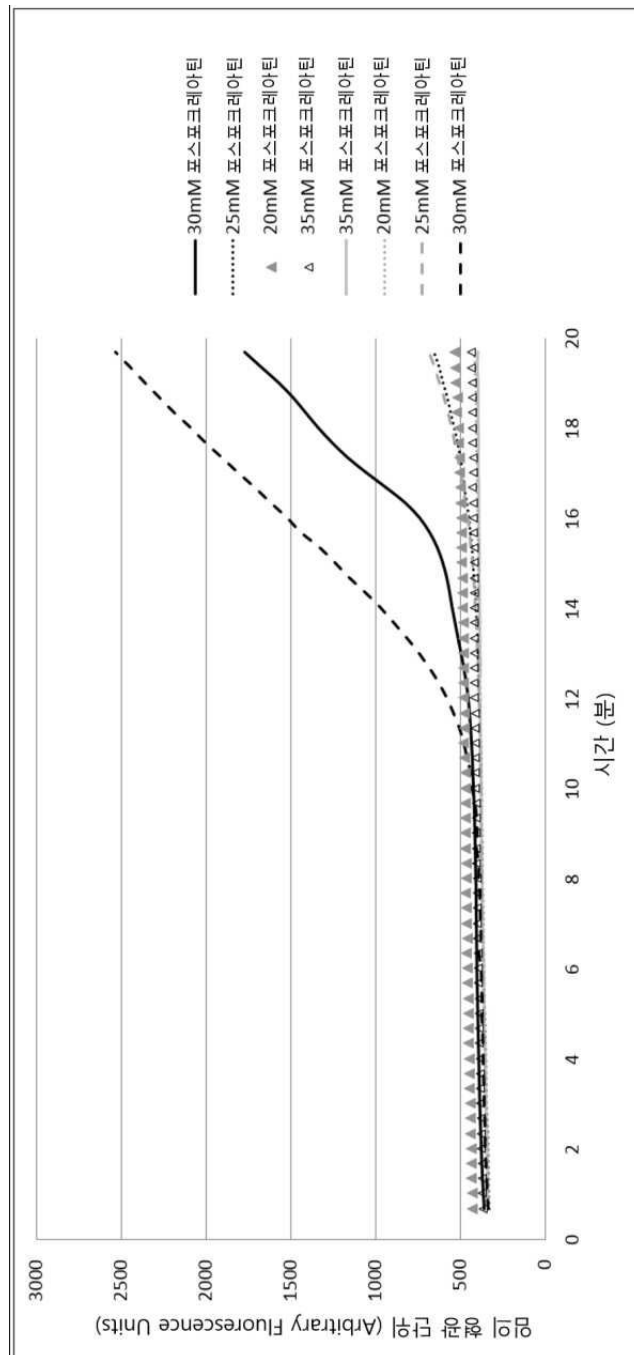
도면5d



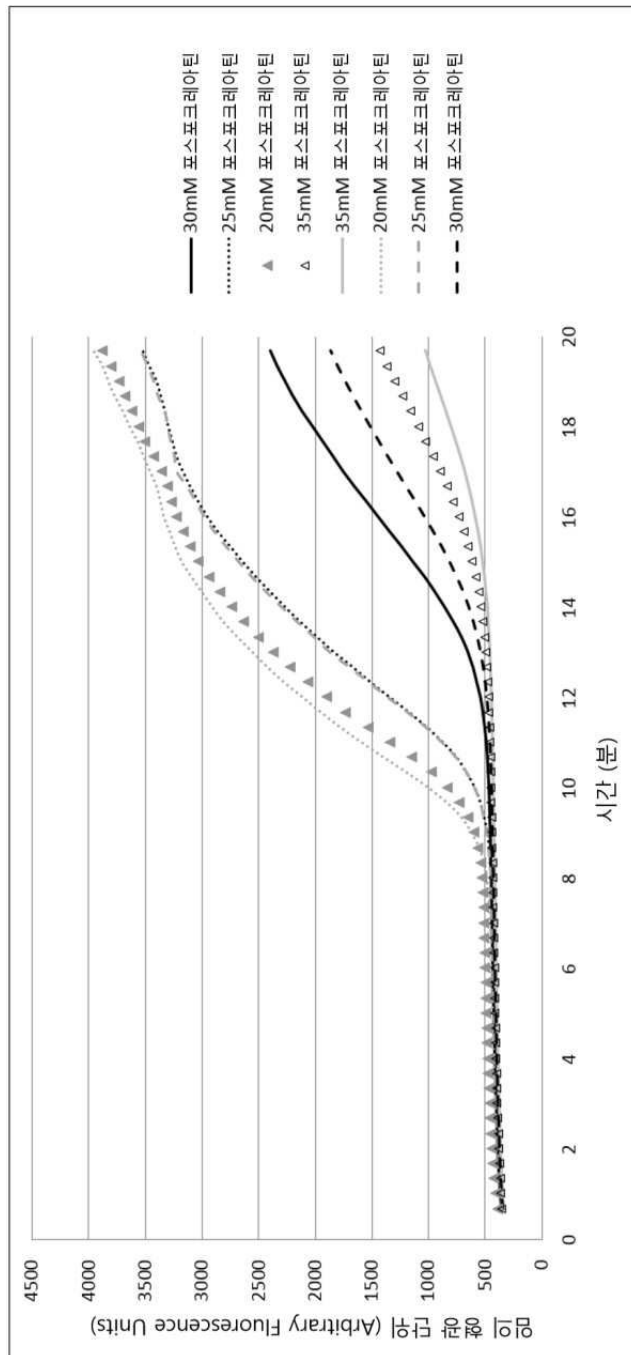
도면6



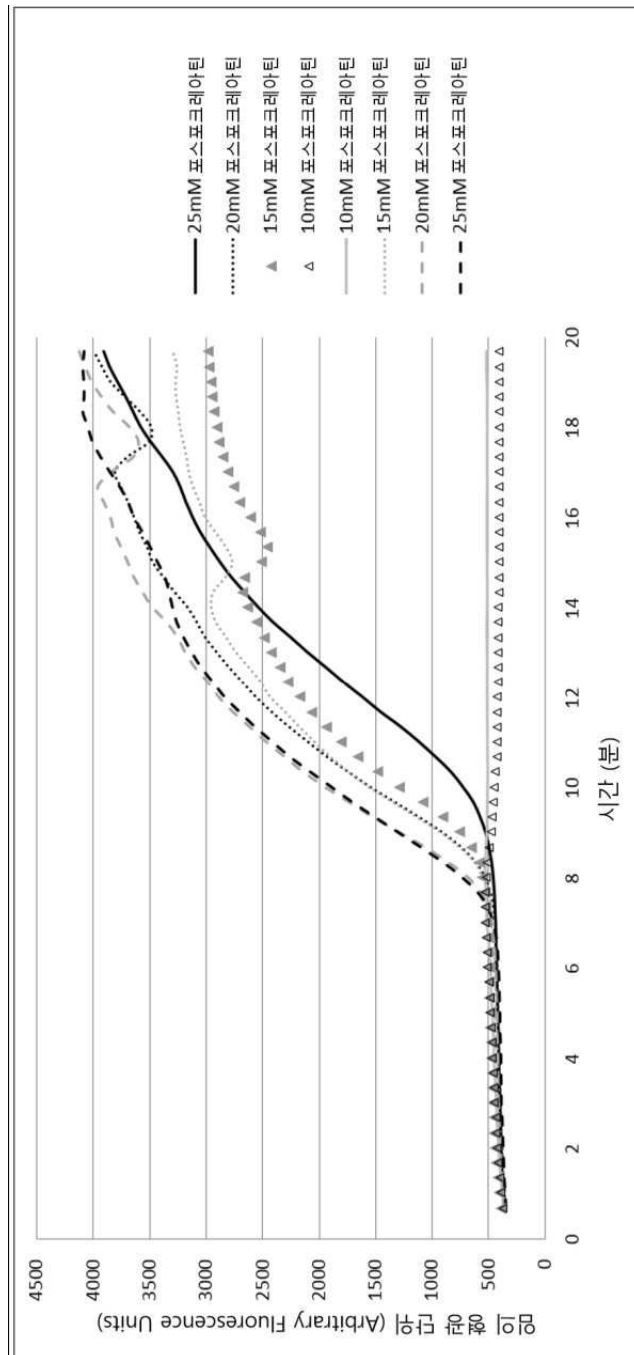
도면7a



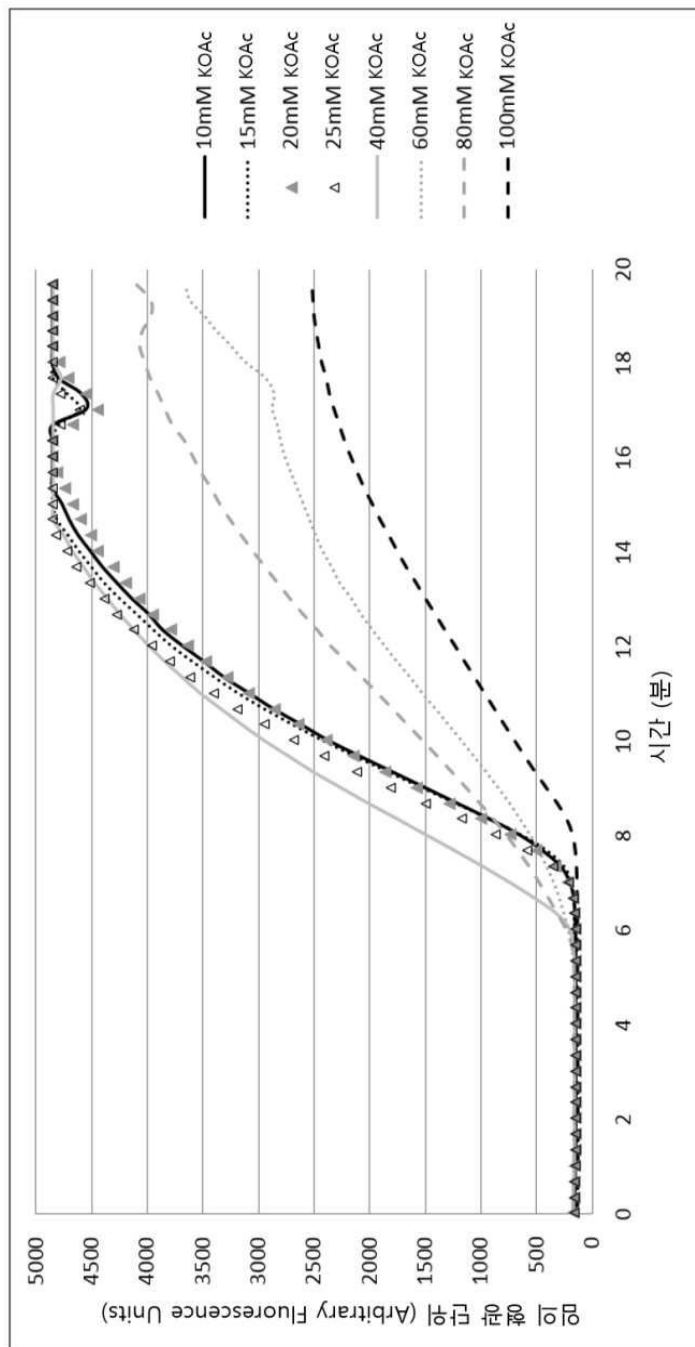
도면7b



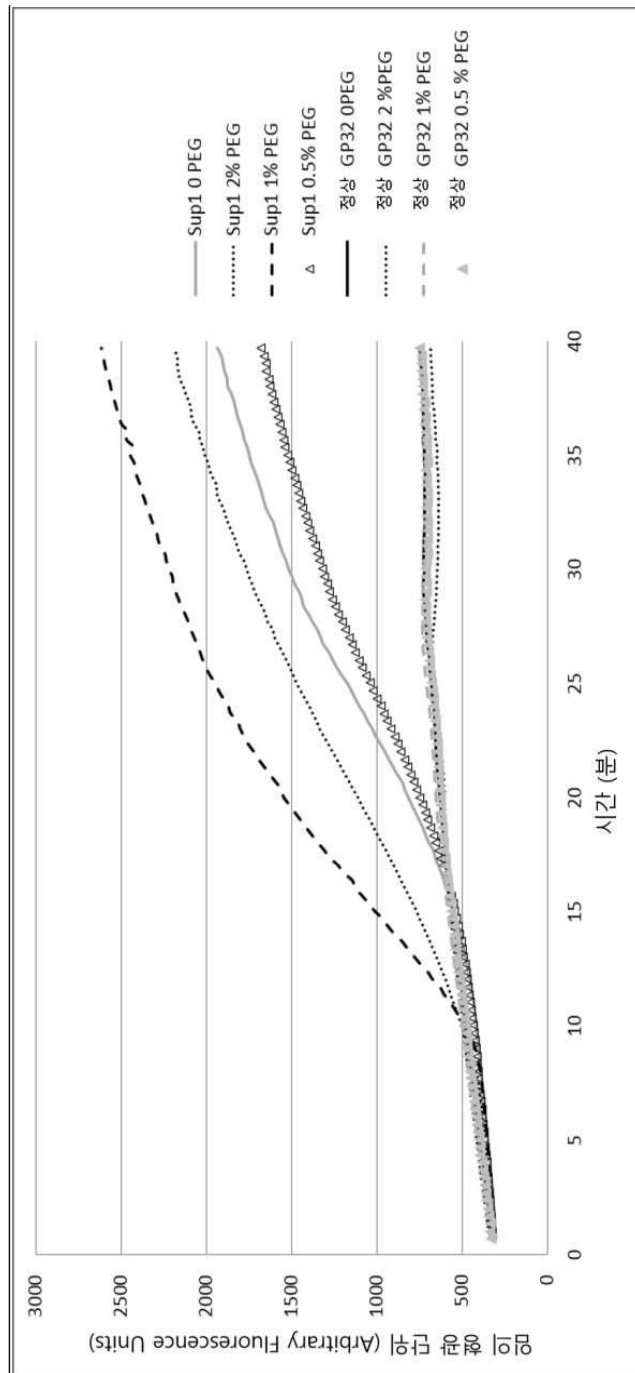
도면7c



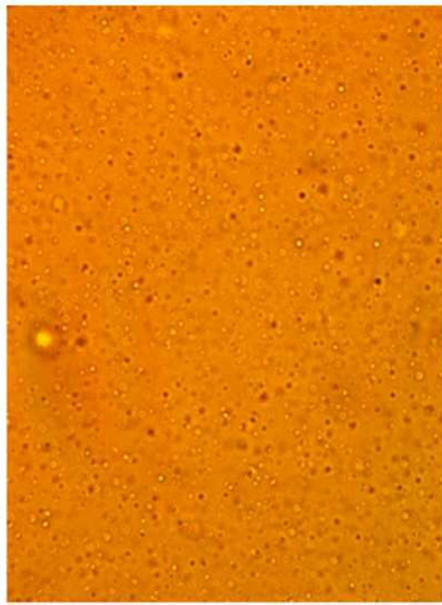
도면8



도면9



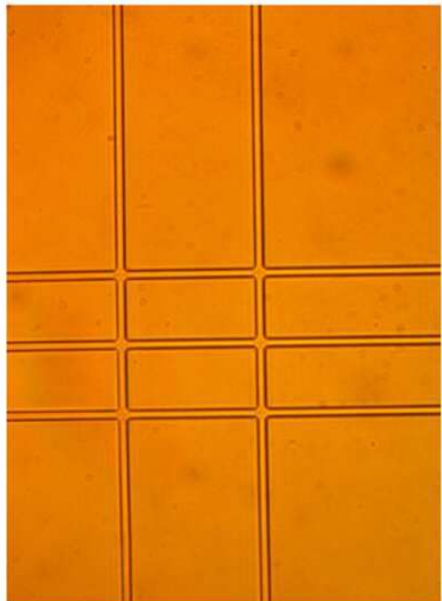
도면10a



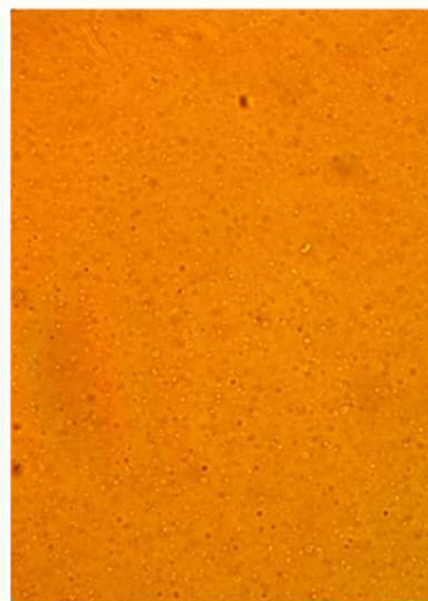
10mM MgOAc



2mM MnCl₂



이온
교환
수지

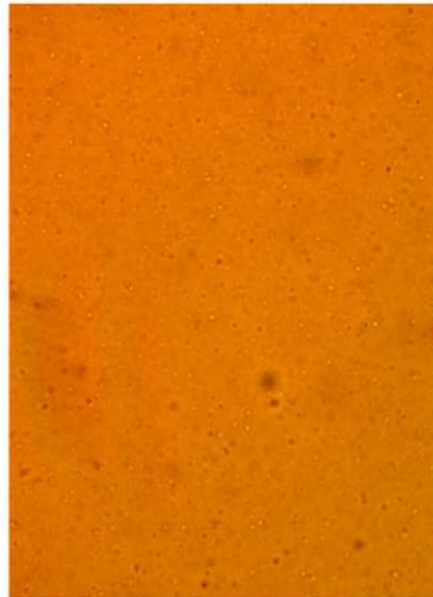


12mM CaCl₂

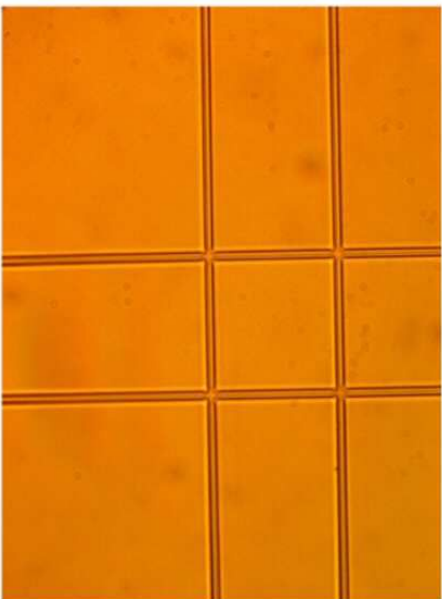
도면10b



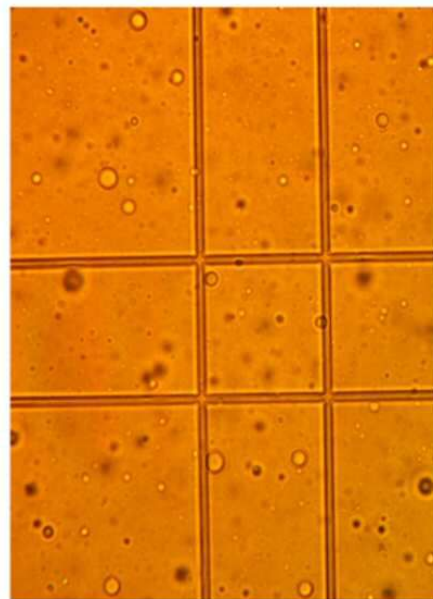
10mM MgOAc



2mM MnCl₂



이
영
상
의
미



12mM CaCl₂

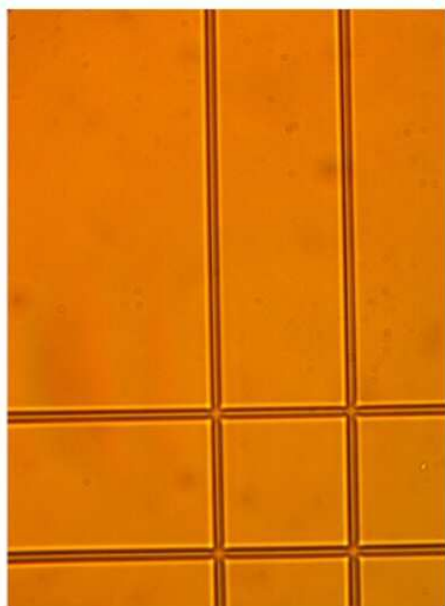
도면10c



10mM MgOAc



2mM MnCl₂



이온
교환
막



12mM CaCl₂

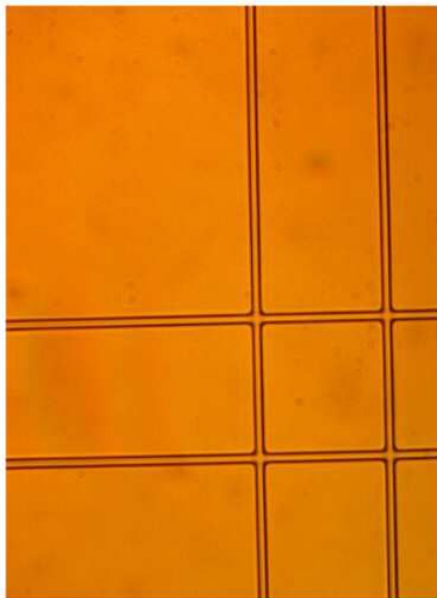
도면10d



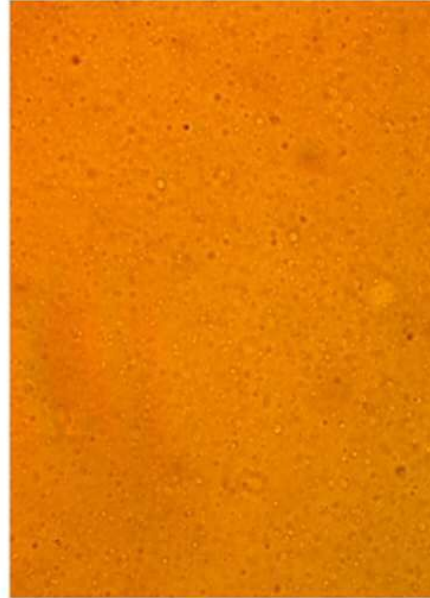
10mM MgOAc



2mM MnCl₂



이
공
식
상
의
도



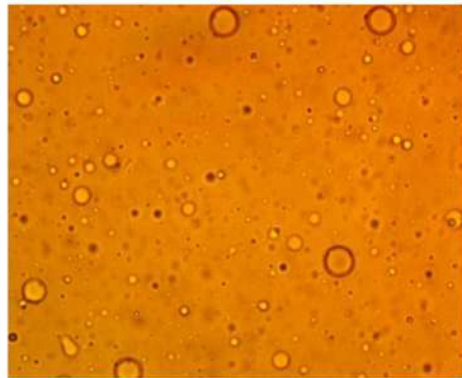
12mM CaCl₂

도면11a

GP32-Fib



20mM MgOAc



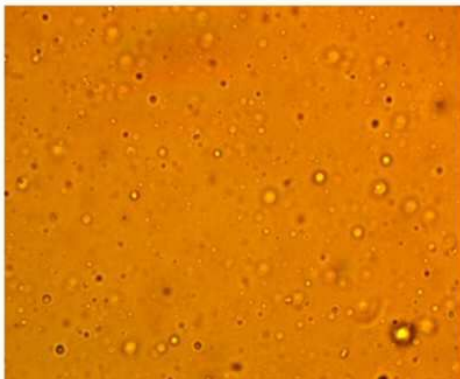
20mM MnCl₂



20mM CaCl₂

도면11b

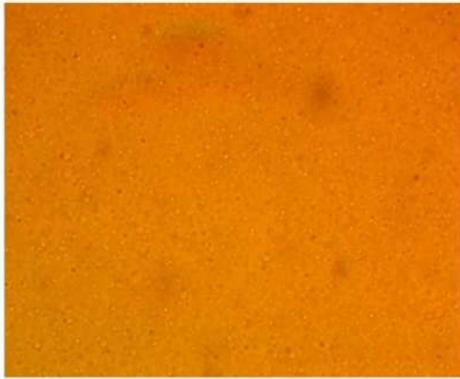
GP32-Sup1



20mM MgOAc



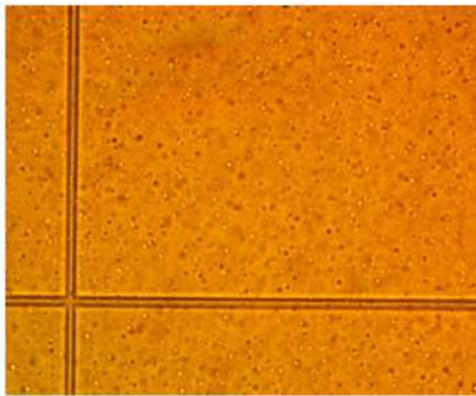
20mM MnCl₂



20mM CaCl₂

도면11c

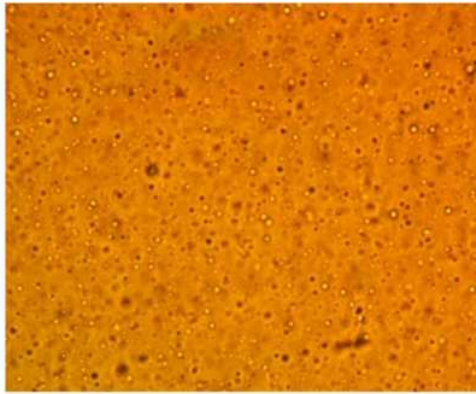
GP32-his2



20mM MgOAc



20mM MnCl₂



20mM CaCl₂

도면11d

GP32-hrp1



20mM MgOAc



20mM MnCl₂



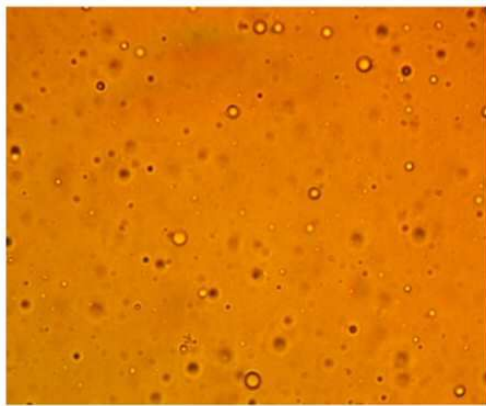
20mM CaCl₂

도면11e

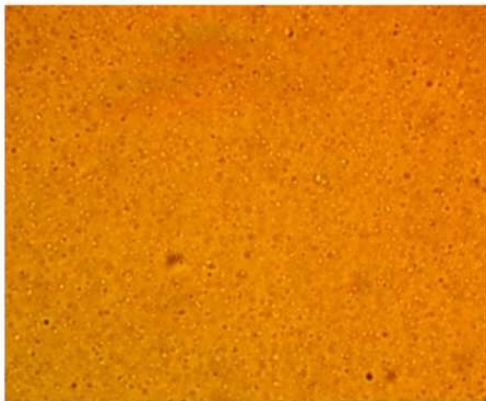
GP32-his5



20mM CaCl_2

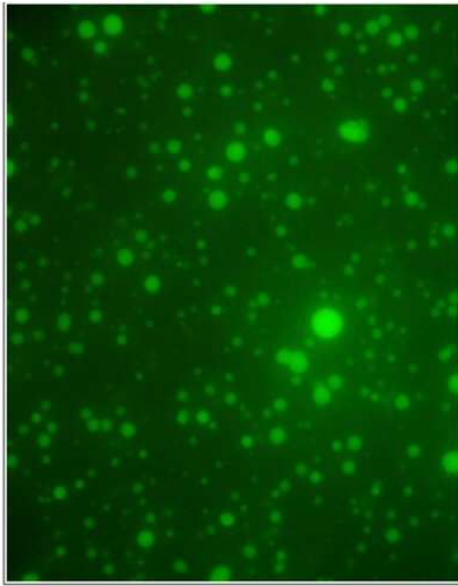


20mM MnCl_2

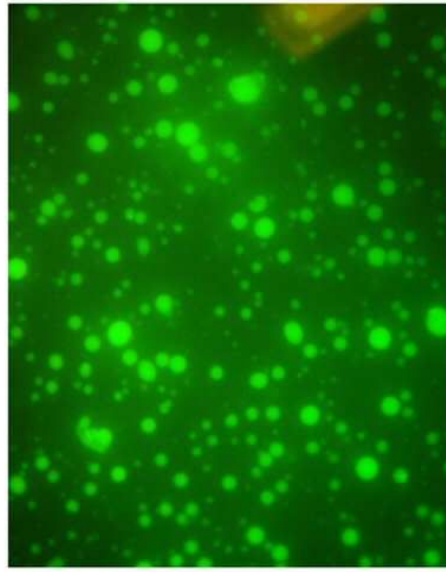


20mM MgOAc

도면12



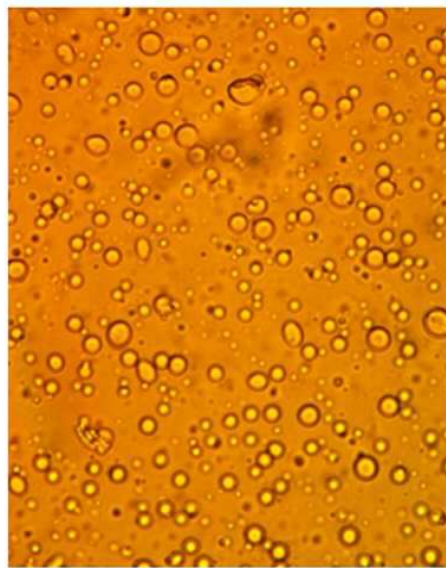
이미지 1- 글라스 커버슬립의 평광



이미지 2- C-Chip 혈구계의 평광

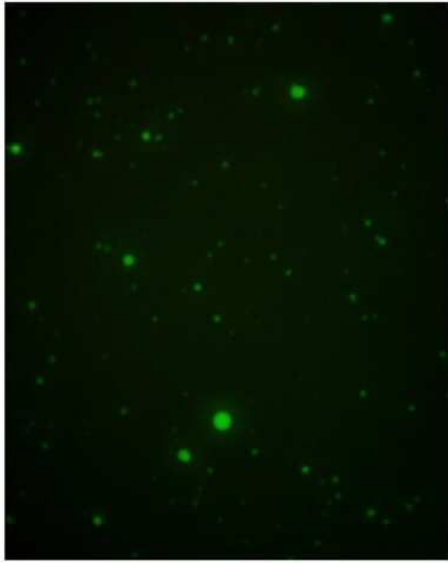


이미지 1- 글라스 커버슬립의 명시야

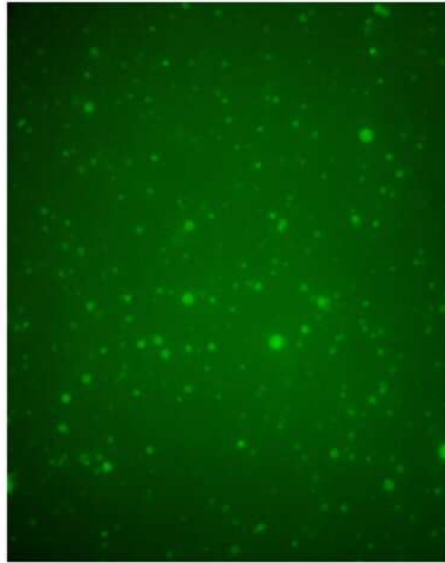


이미지 2- C-Chip 혈구계의 명시야

도면13



이미지 1- 글라스 커버슬립의 형광



이미지 2- C-Chip 혈구계의 형광

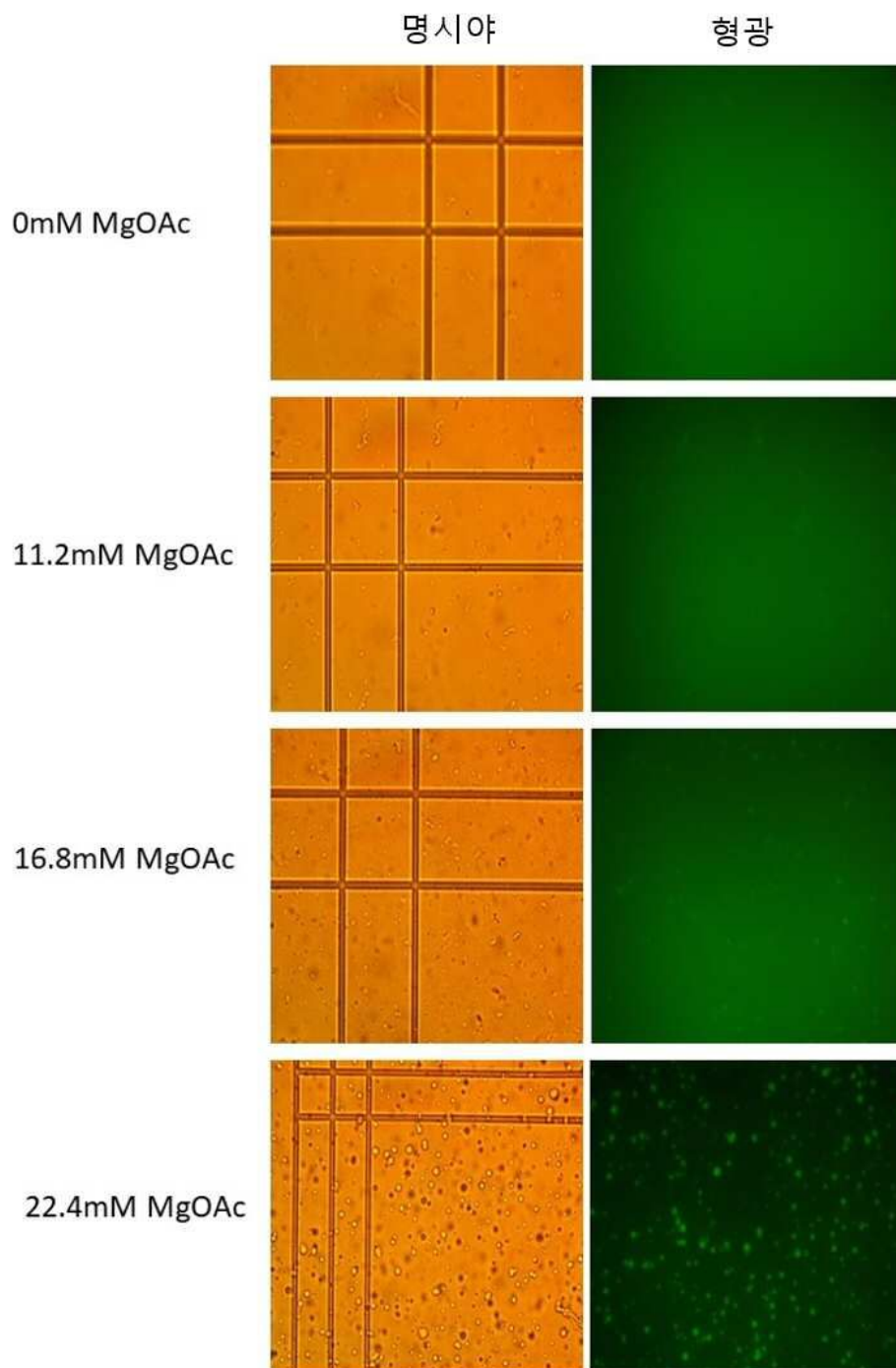


이미지 1- 글라스 커버슬립의 명시야

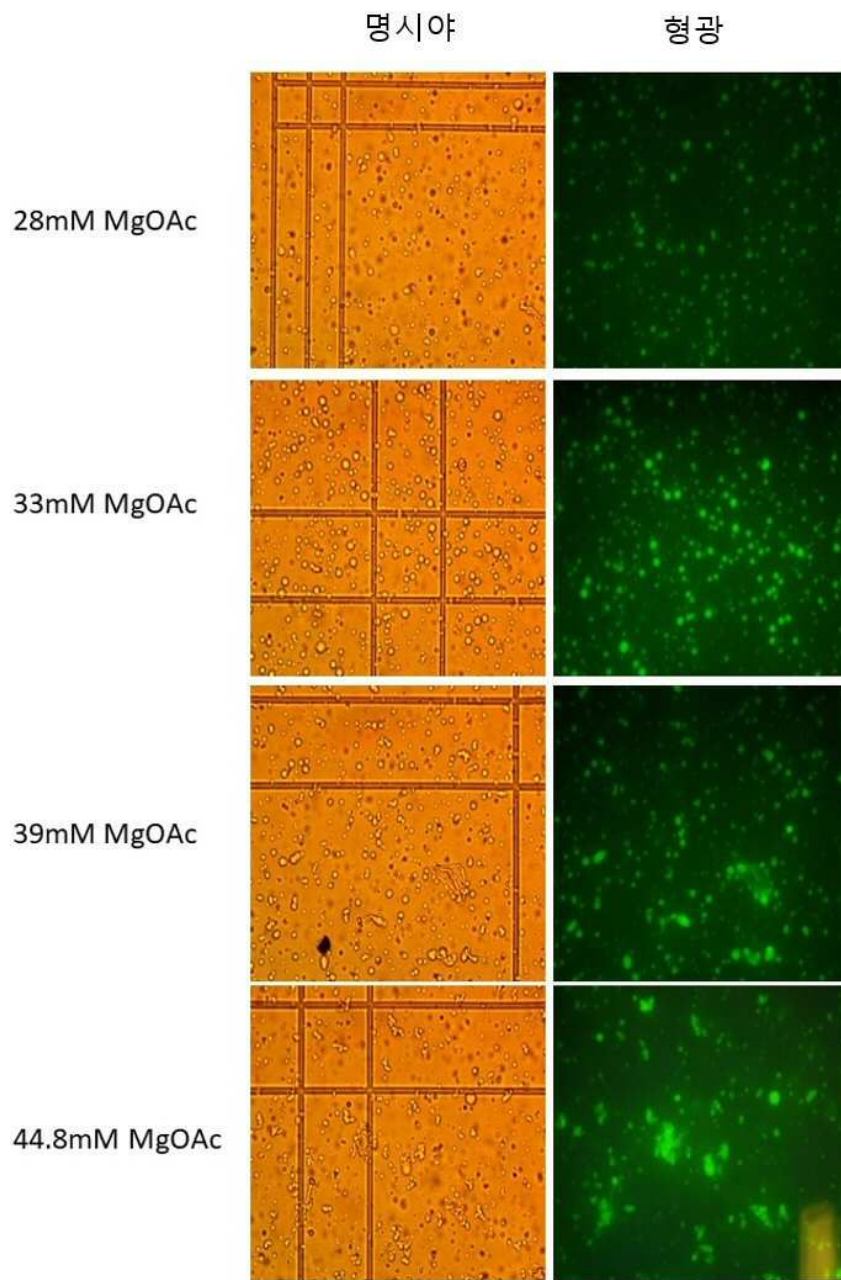


이미지 2- C-Chip 혈구계의 명시야

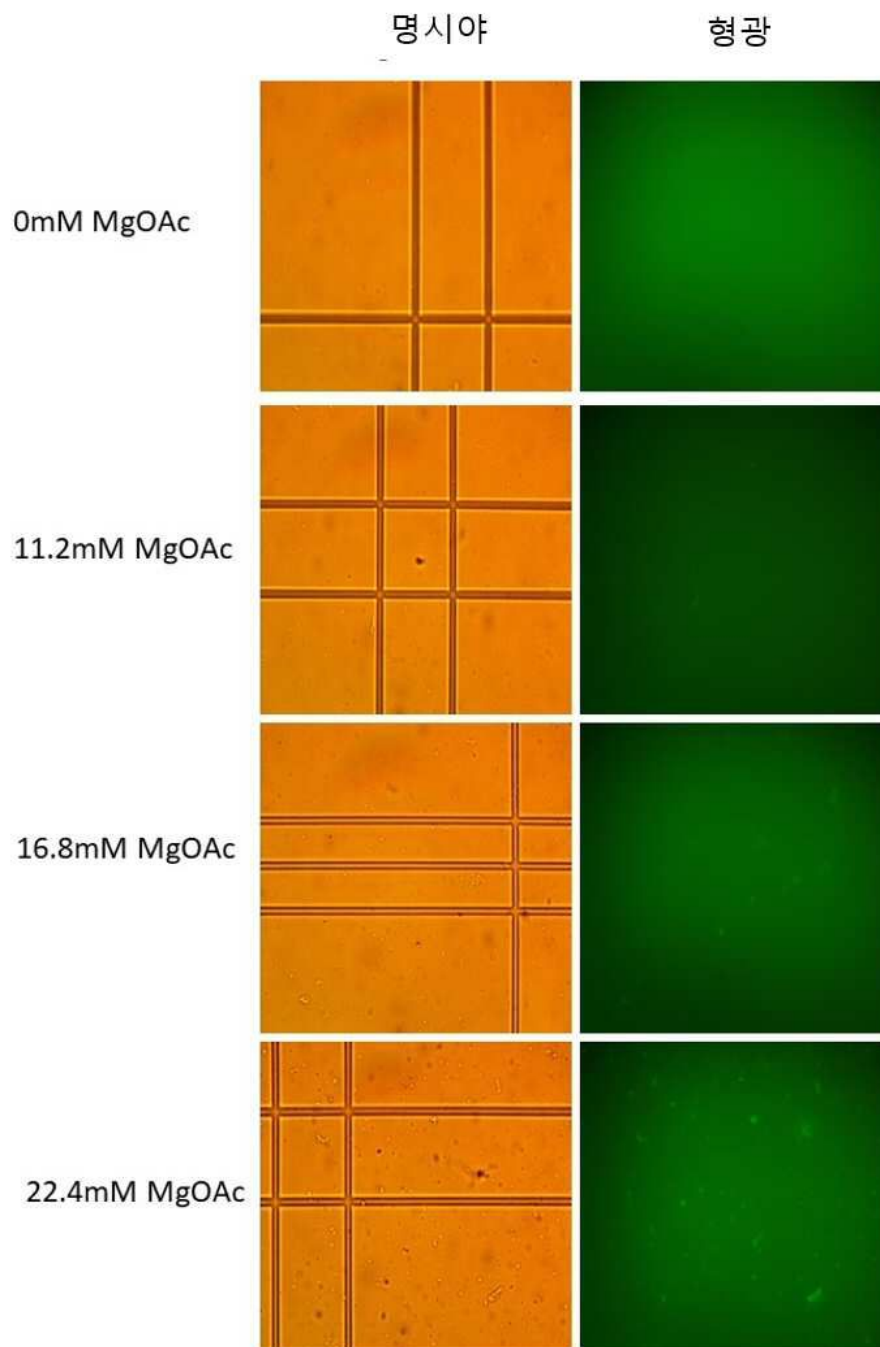
도면14a



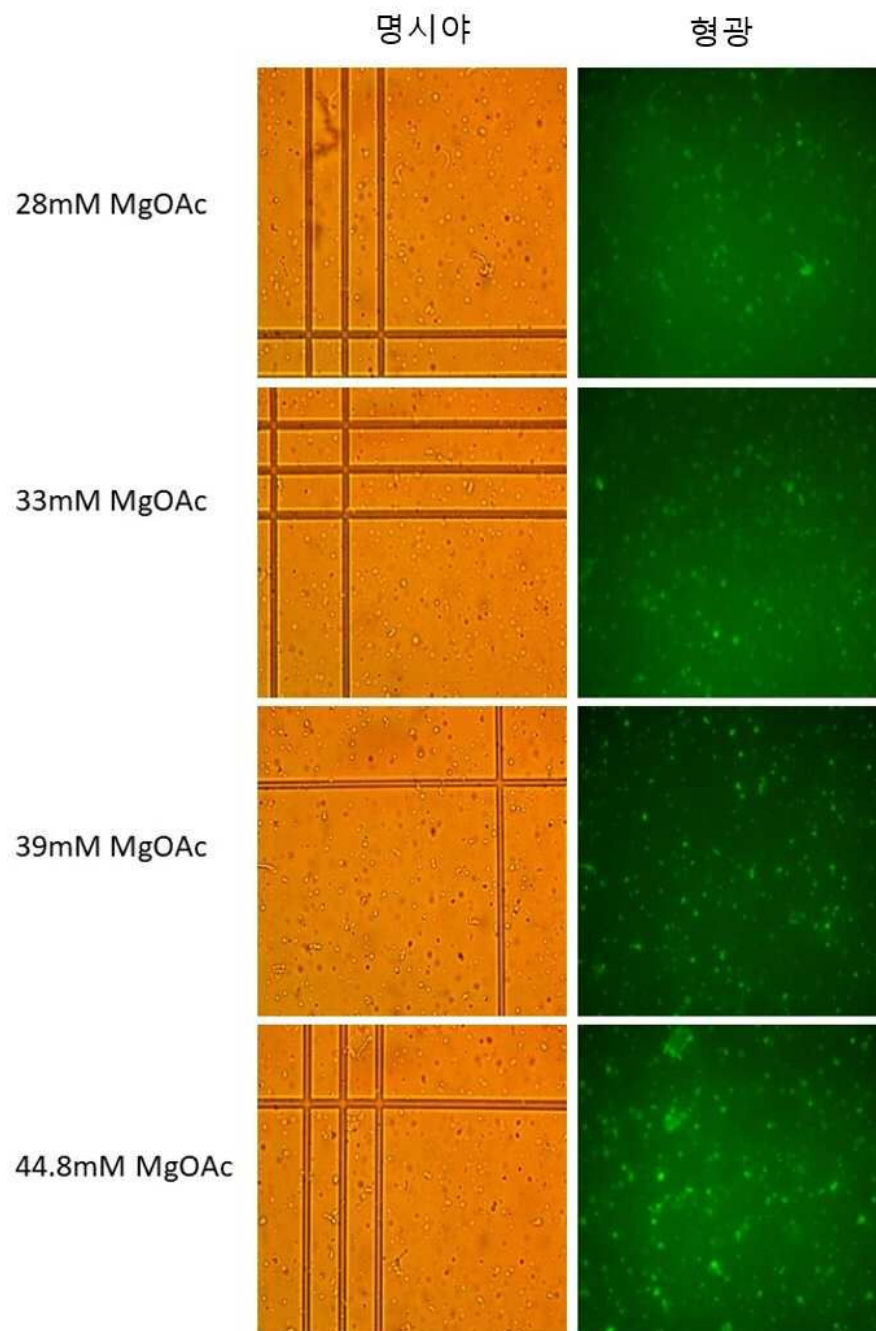
도면14b



도면 15a

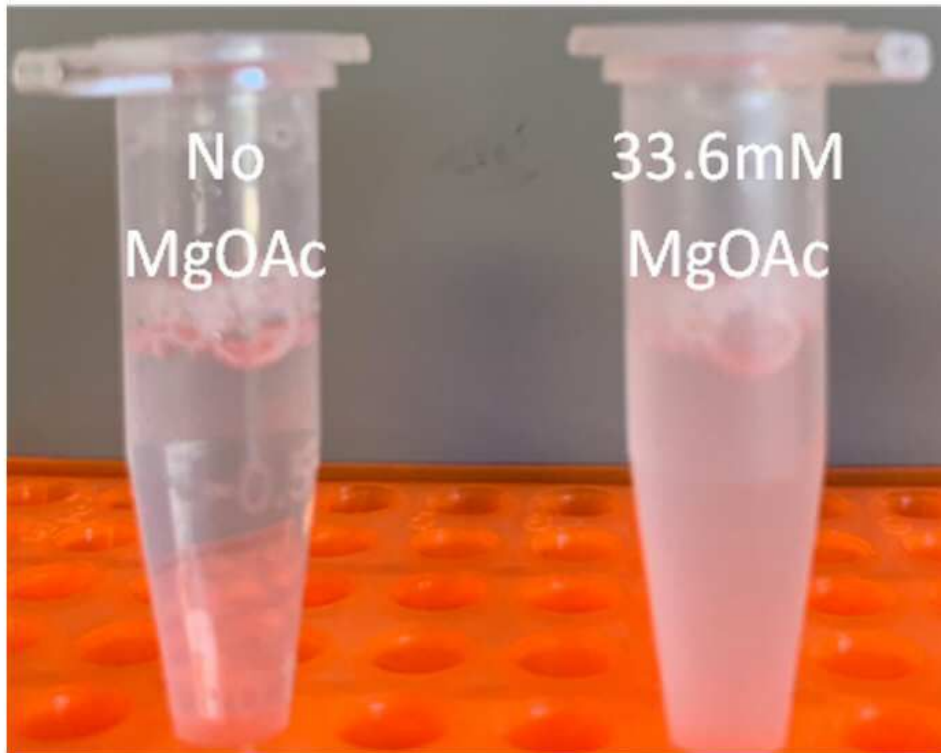


도면15b

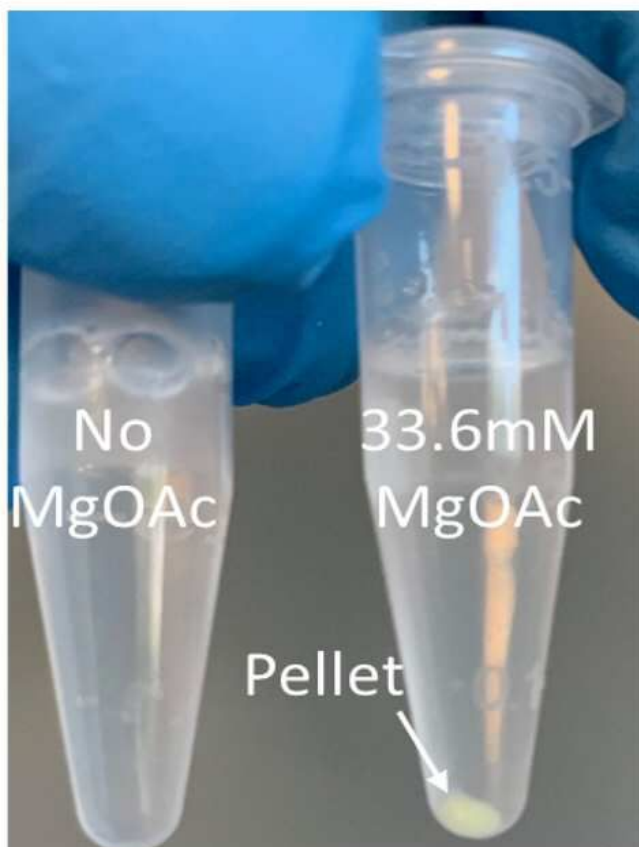


2011

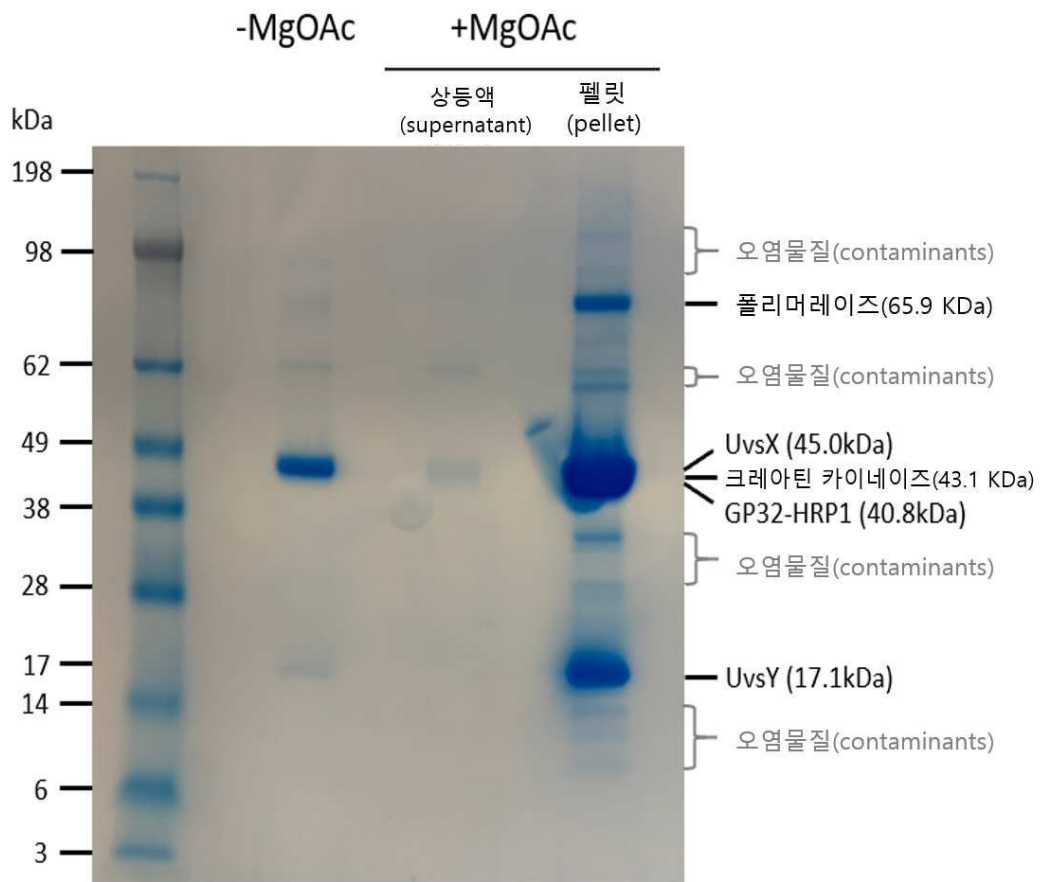
도면16a



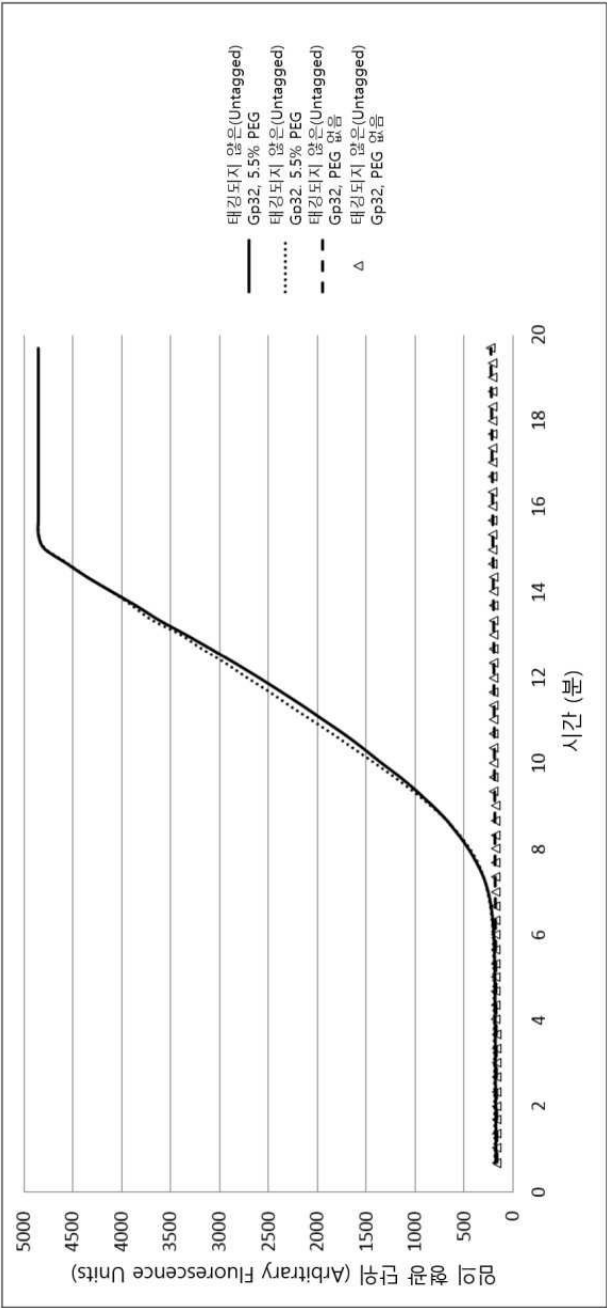
도면16b



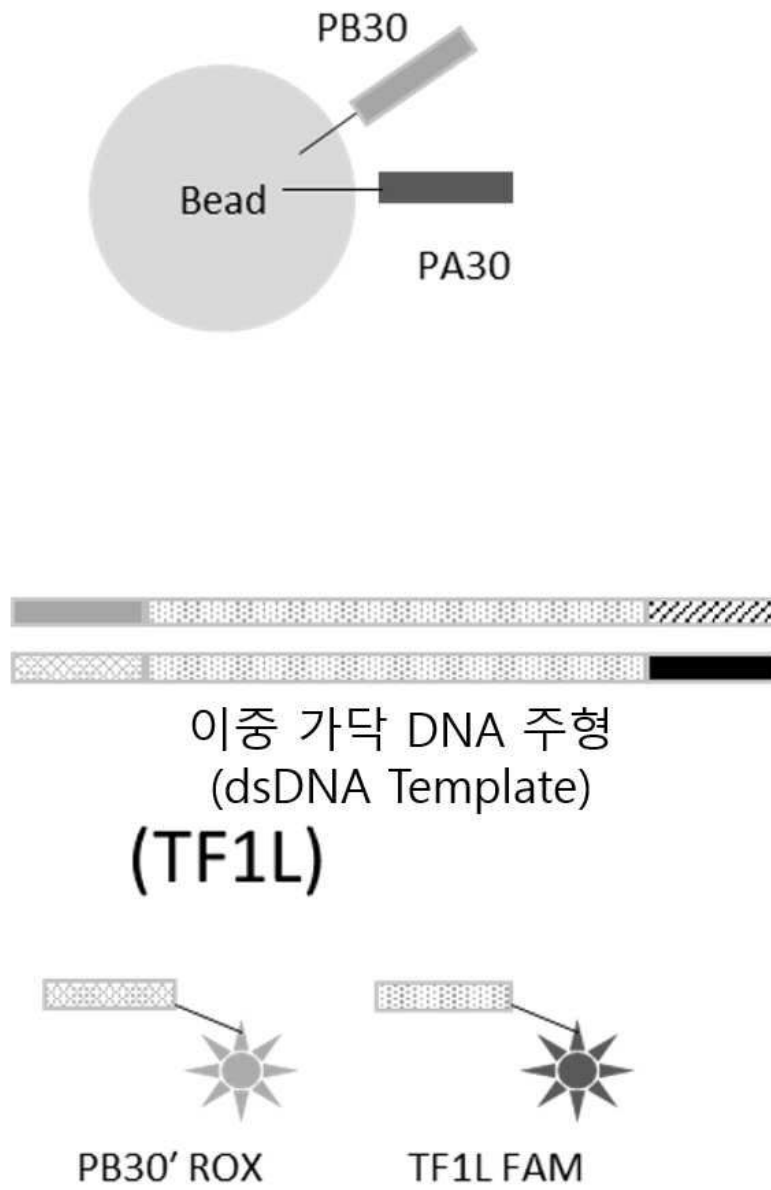
도면16c



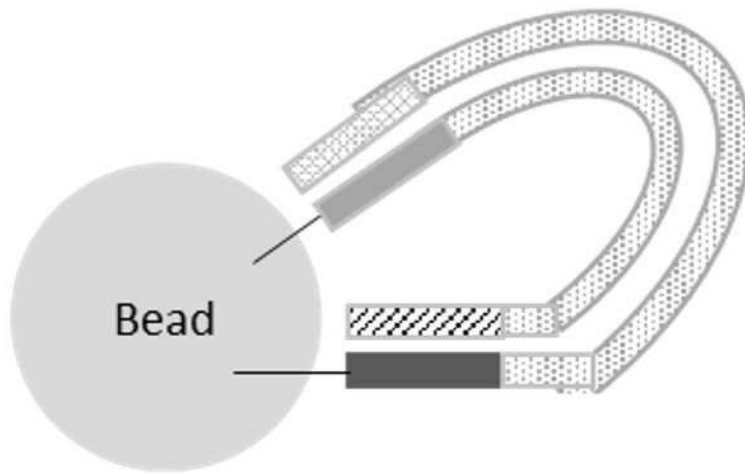
도면17



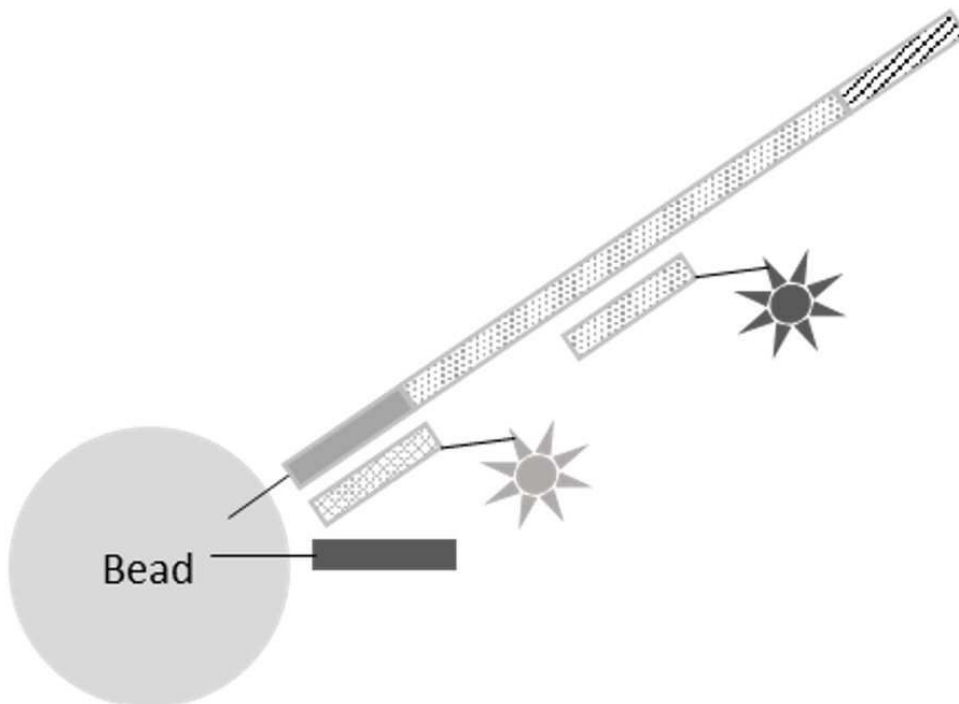
도면18a



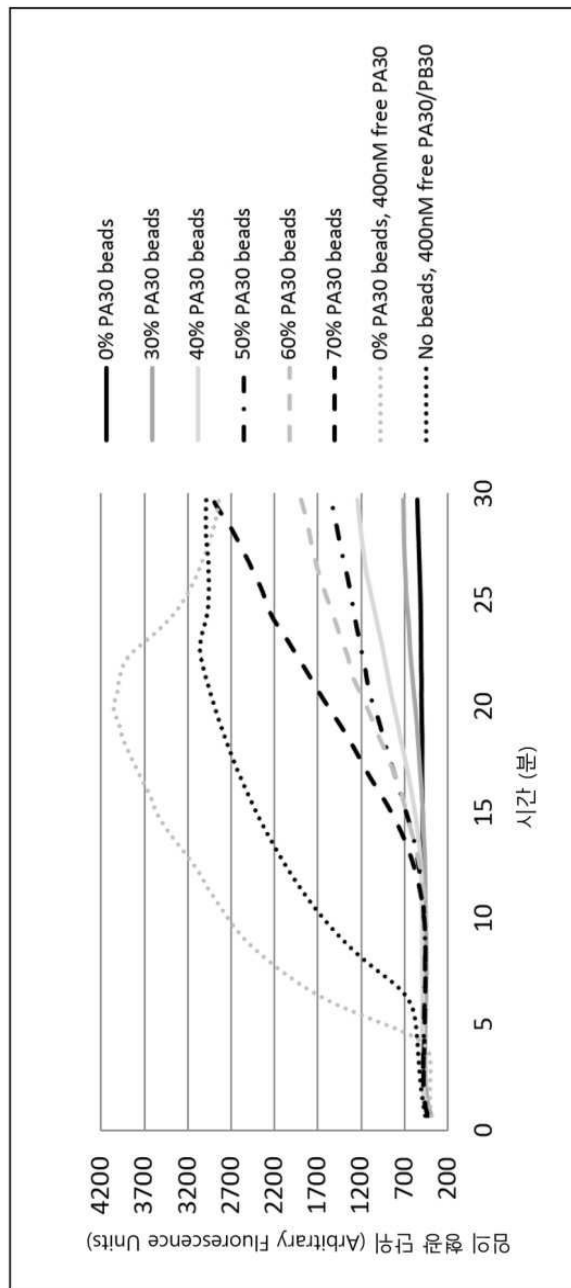
도면18b



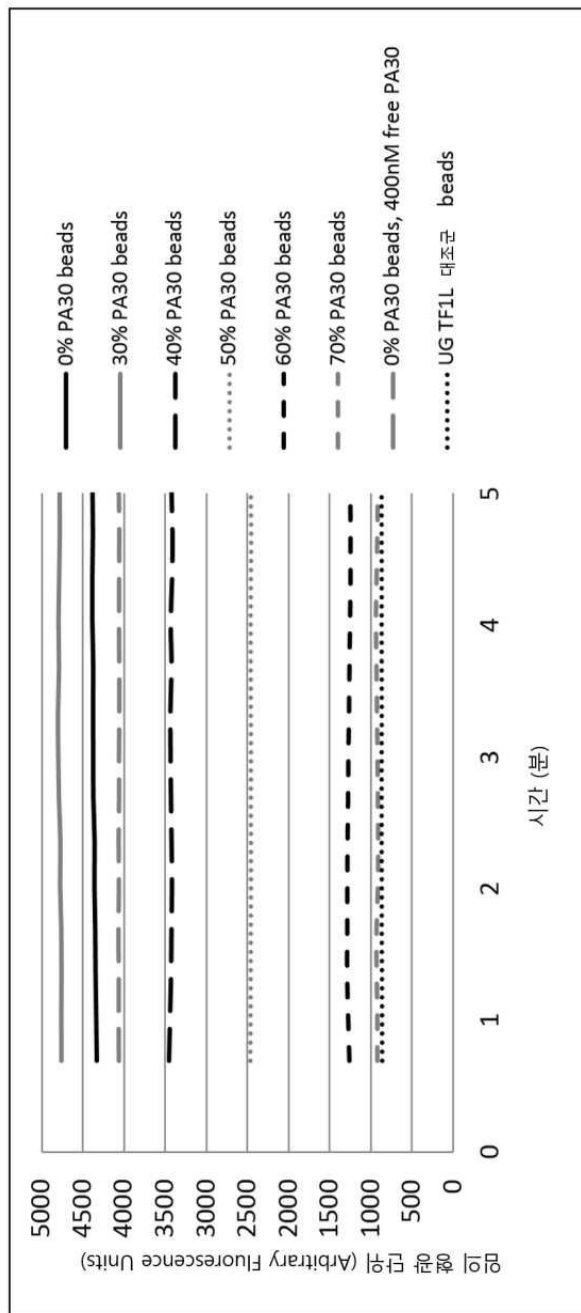
도면18c



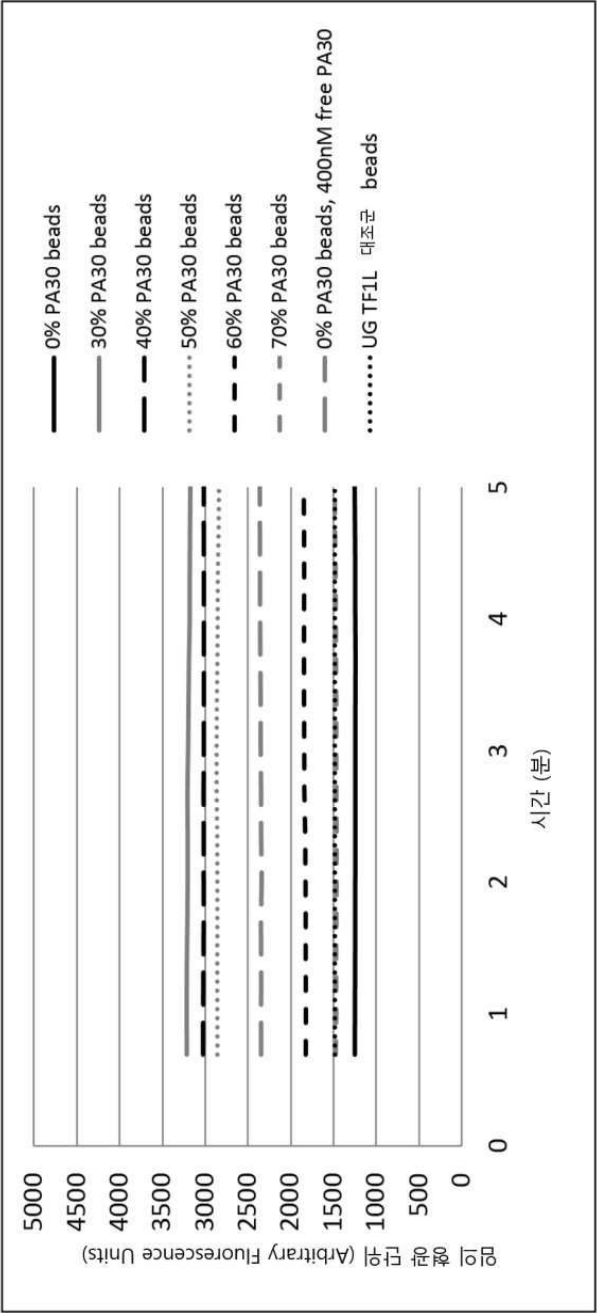
도면18d



도면18e



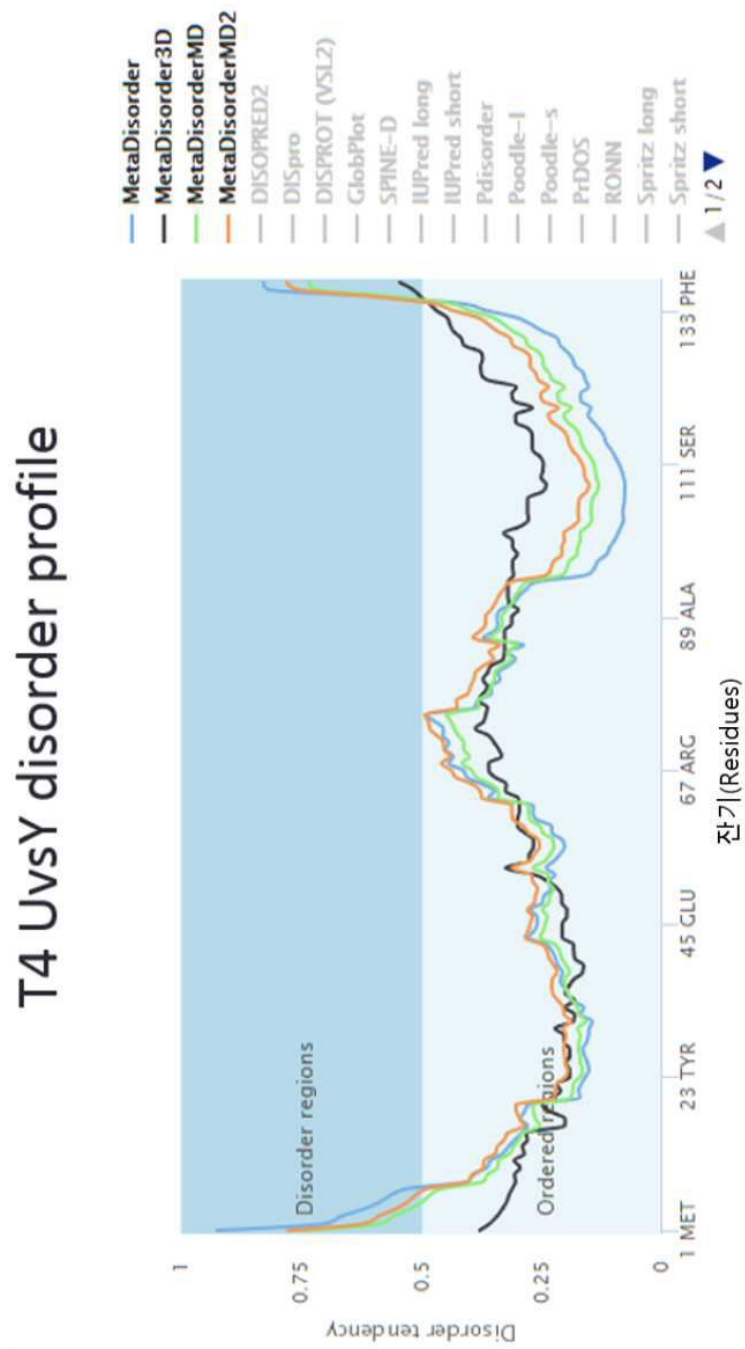
도면18f



도면19a

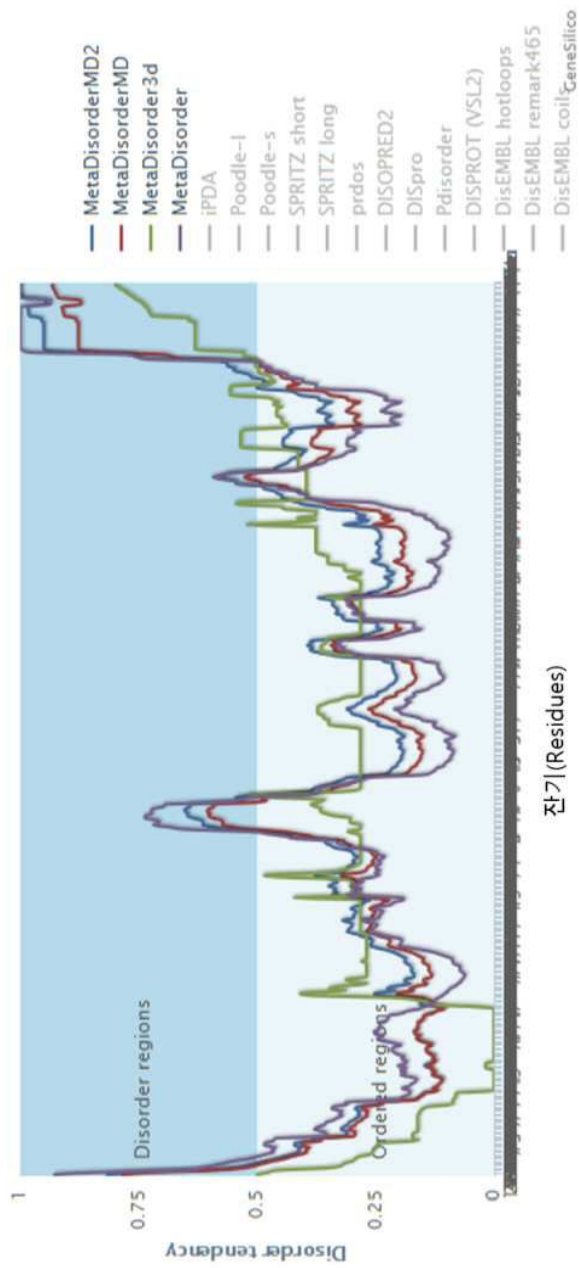


도면19b

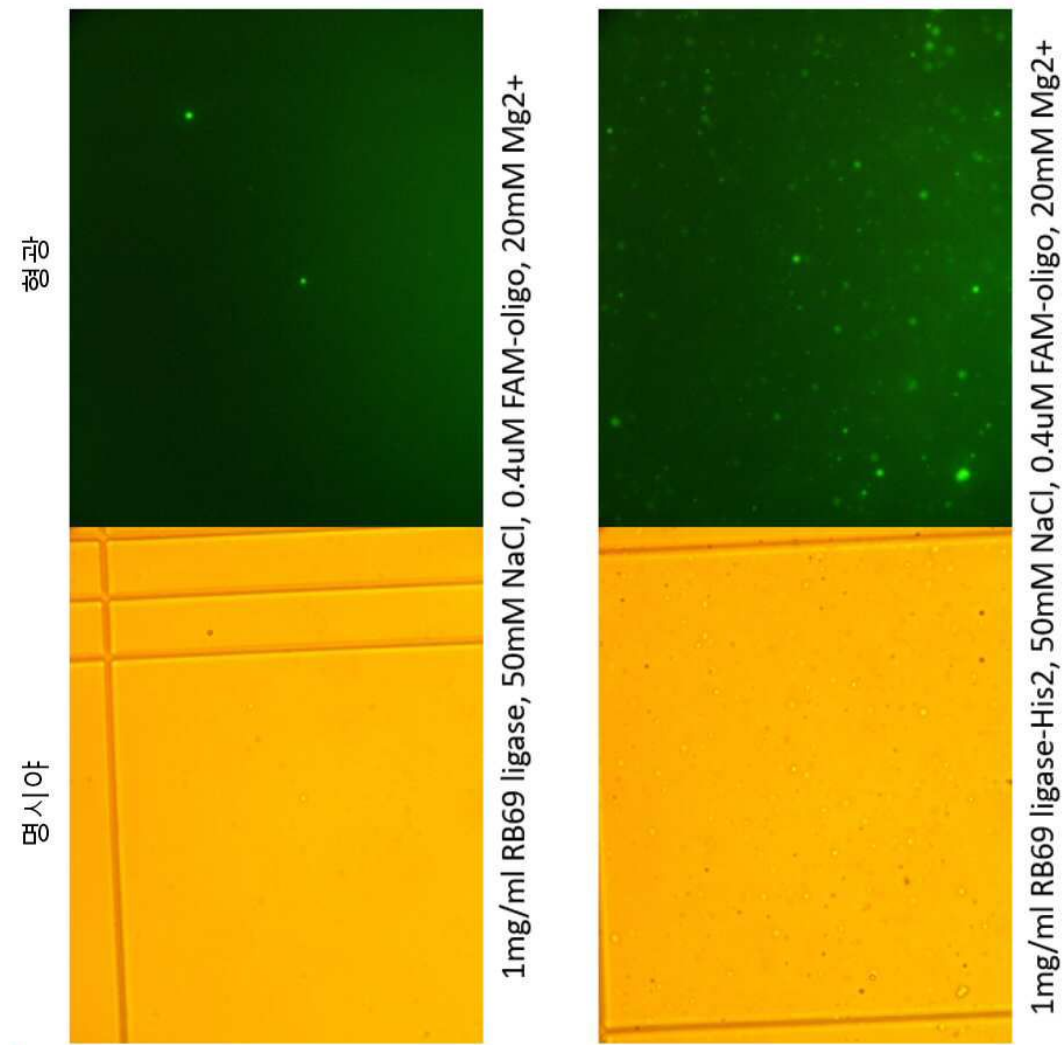


도면19c

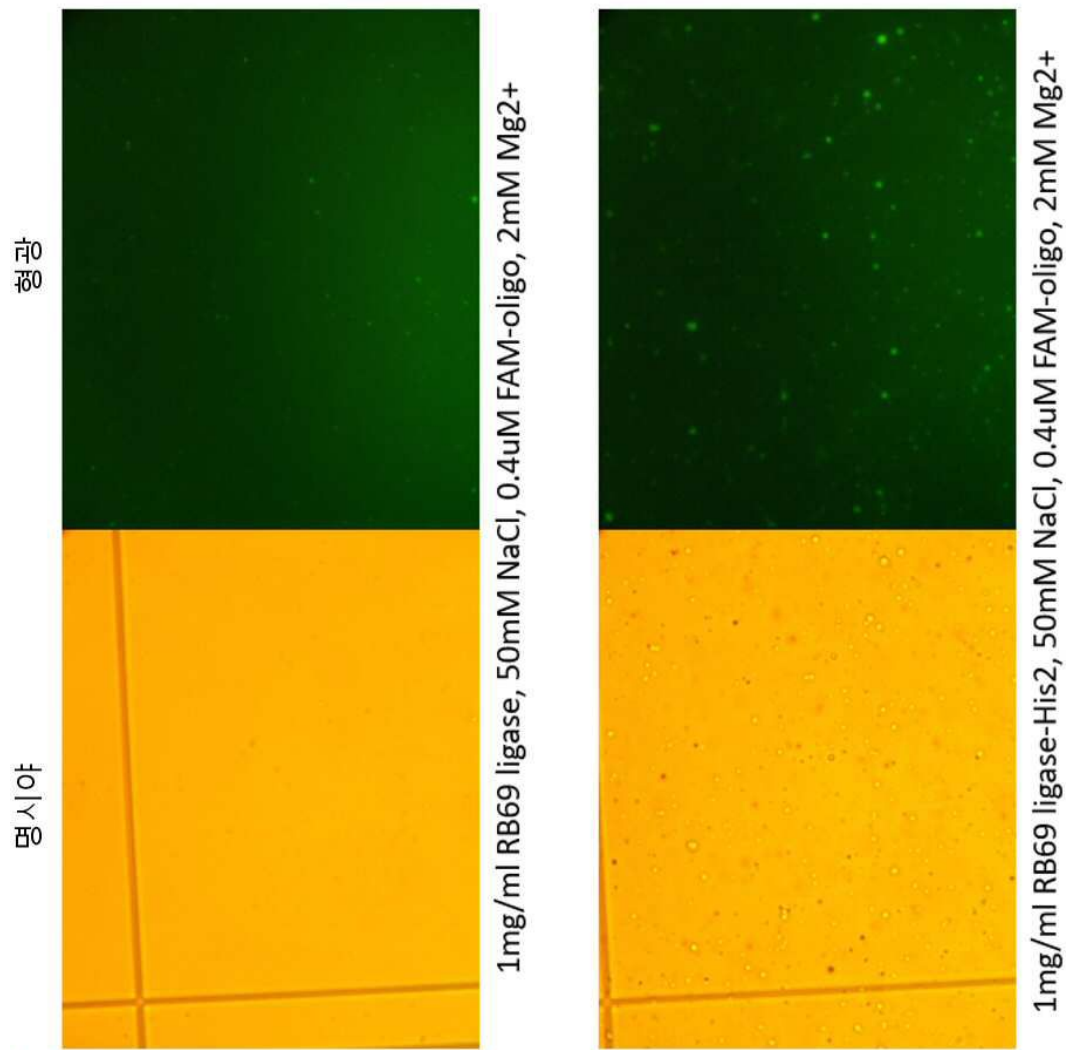
T4 UvsX disorder profile



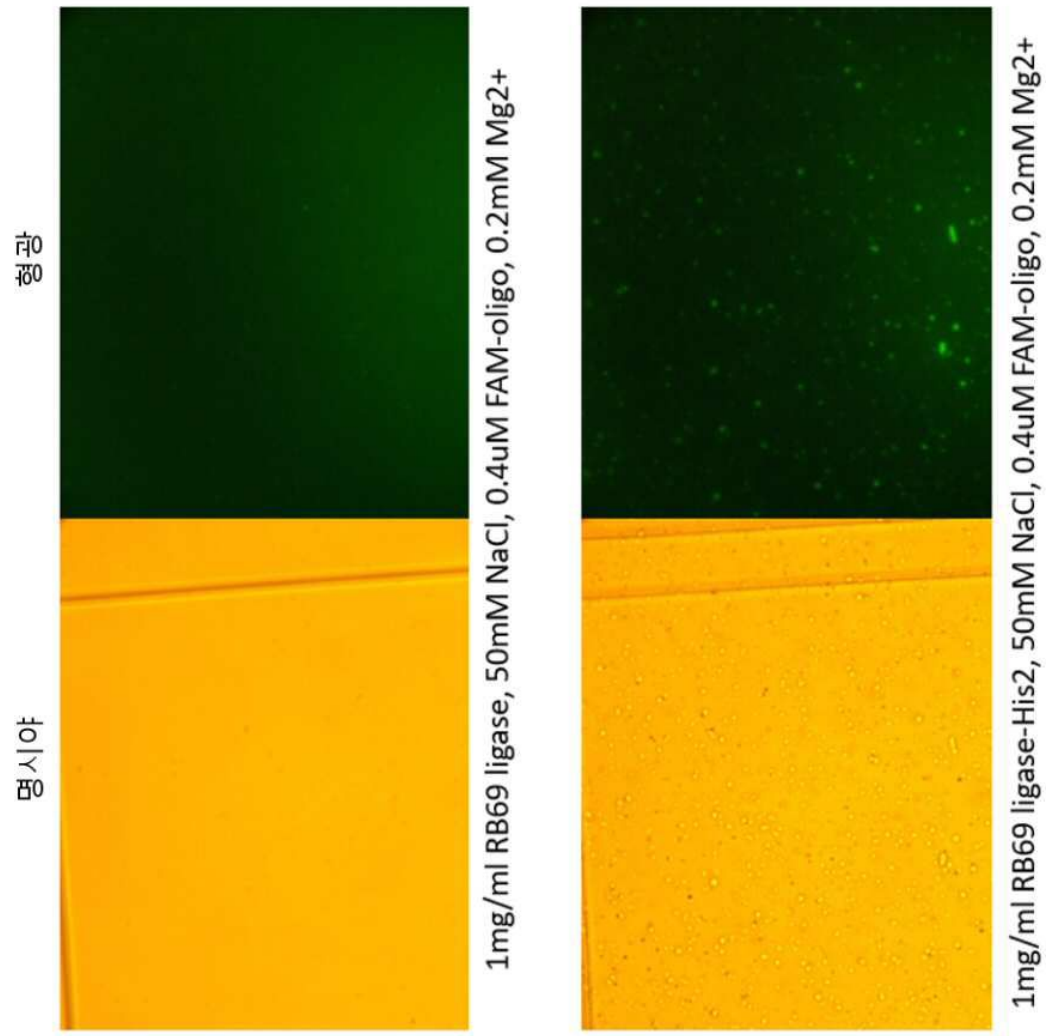
도면20a



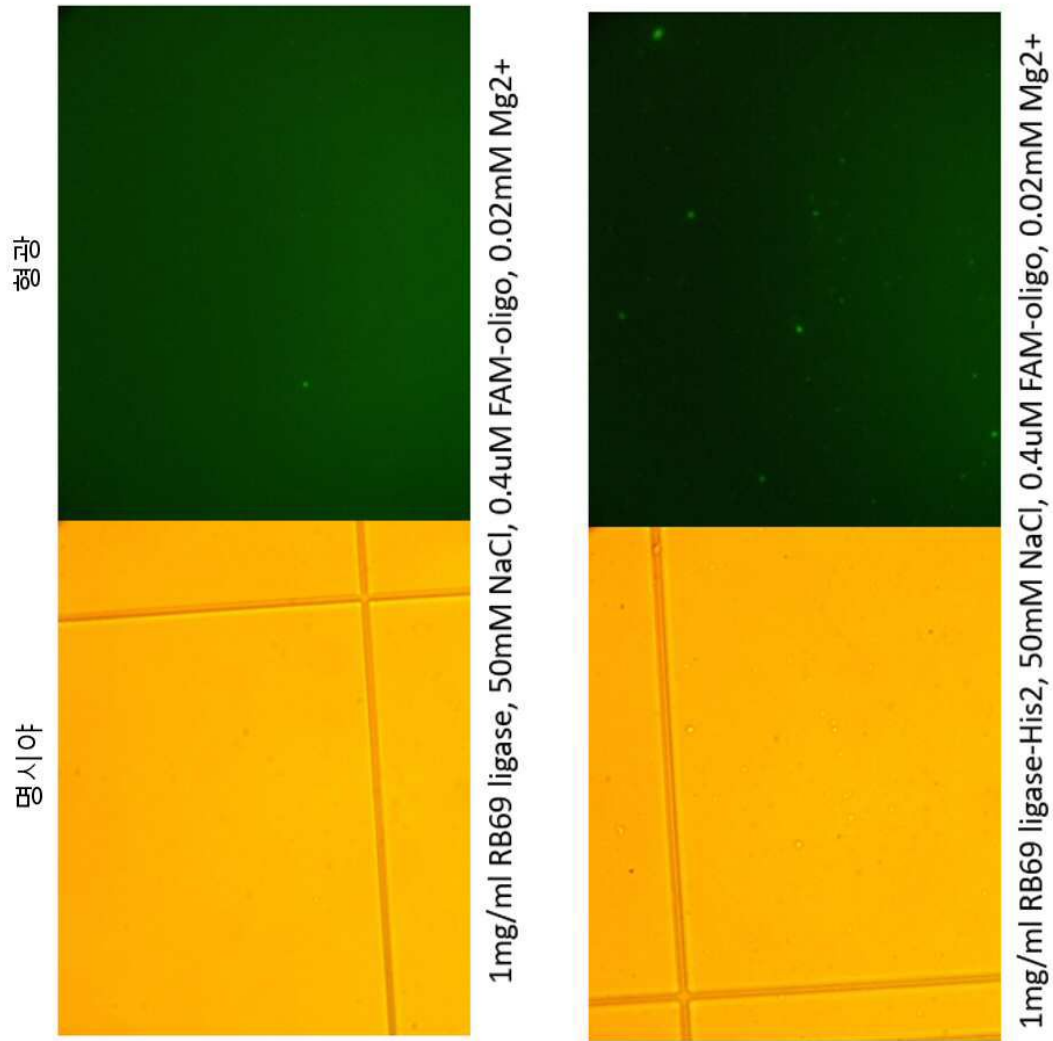
도면20b



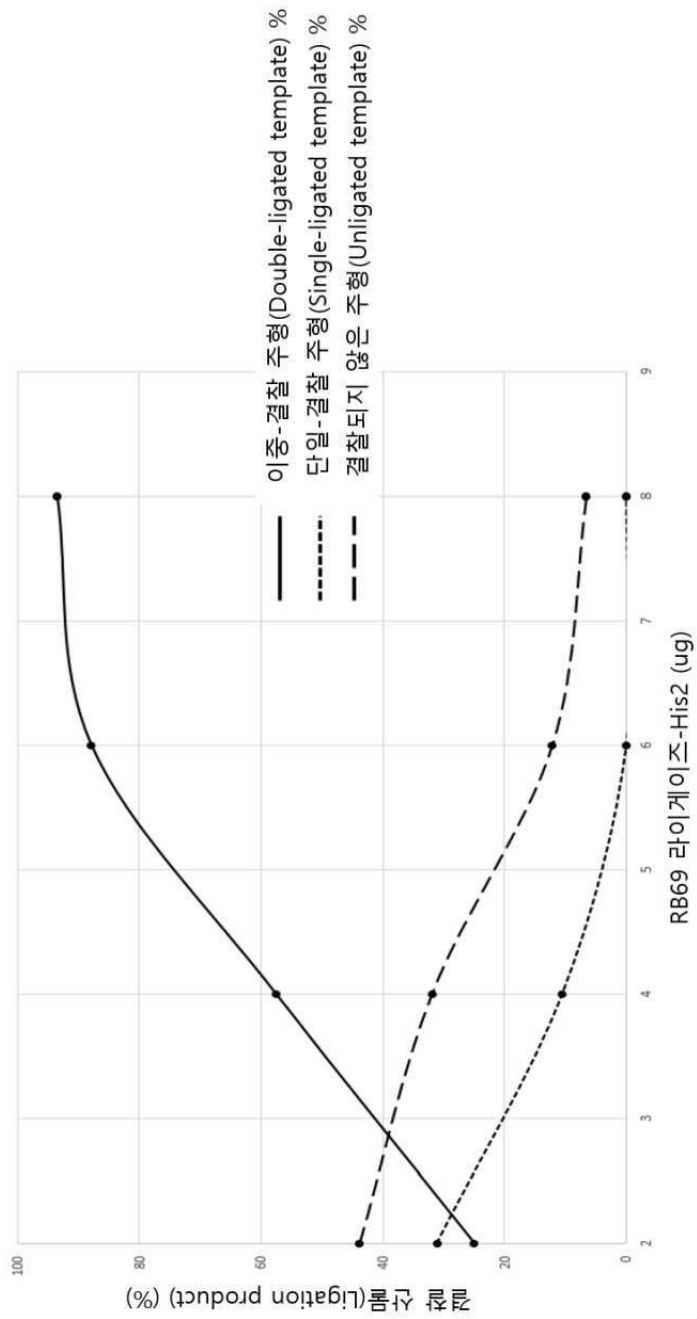
도면20c



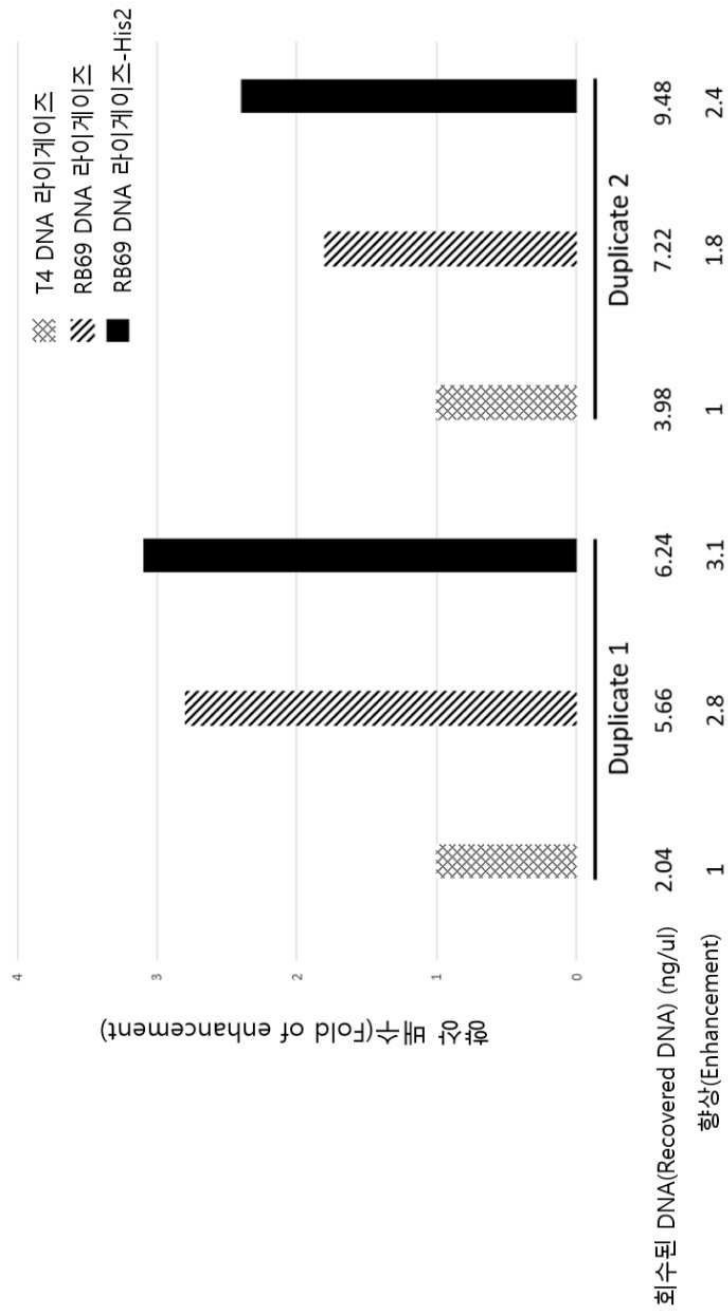
도면20d



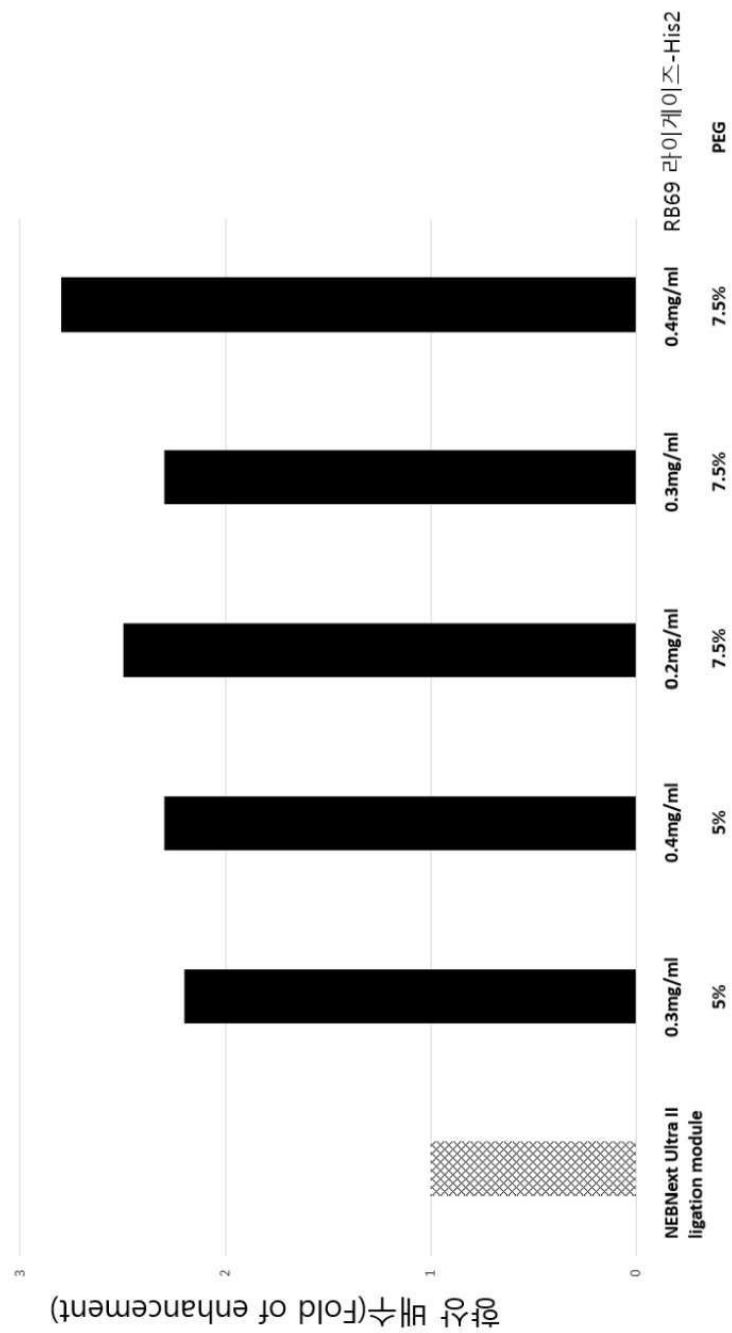
도면21



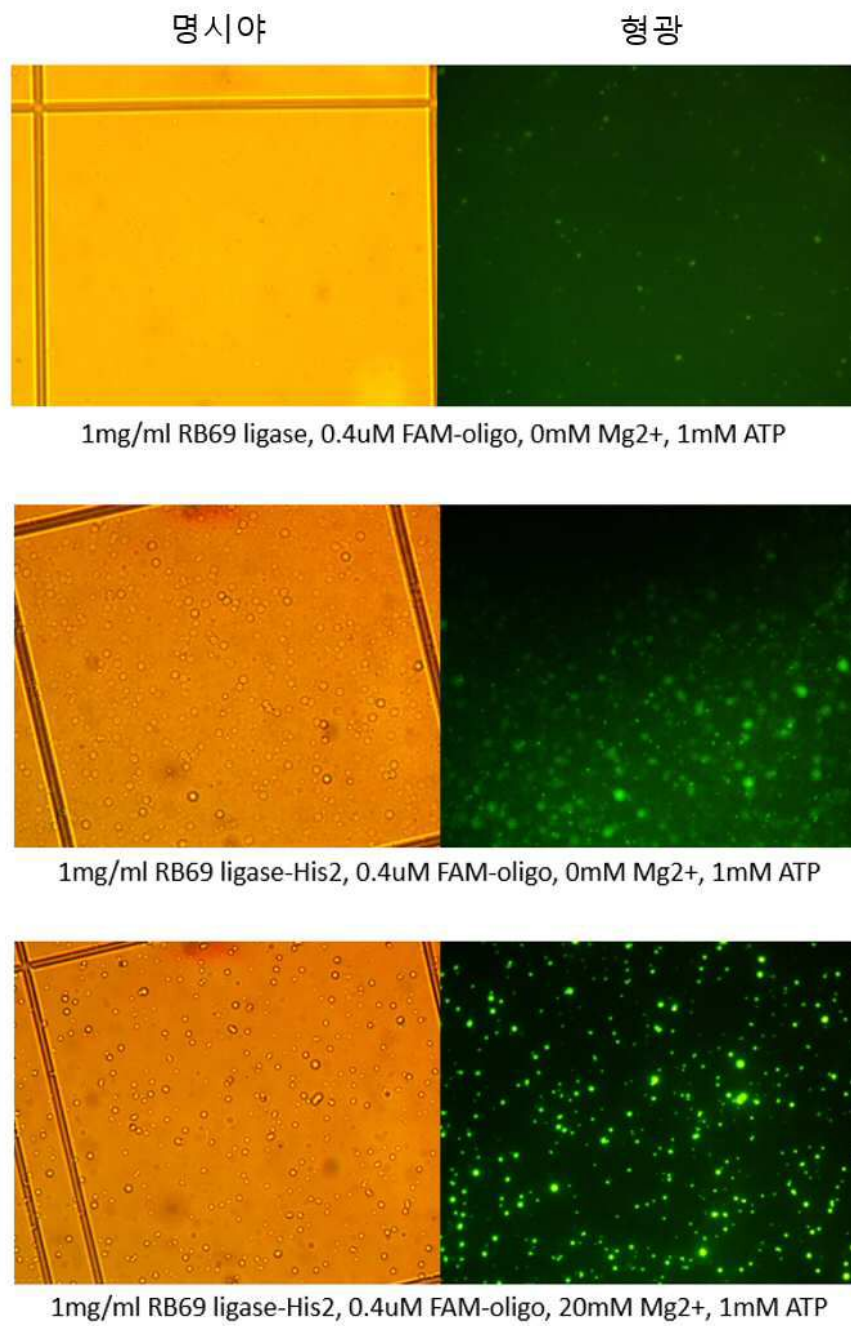
도면22



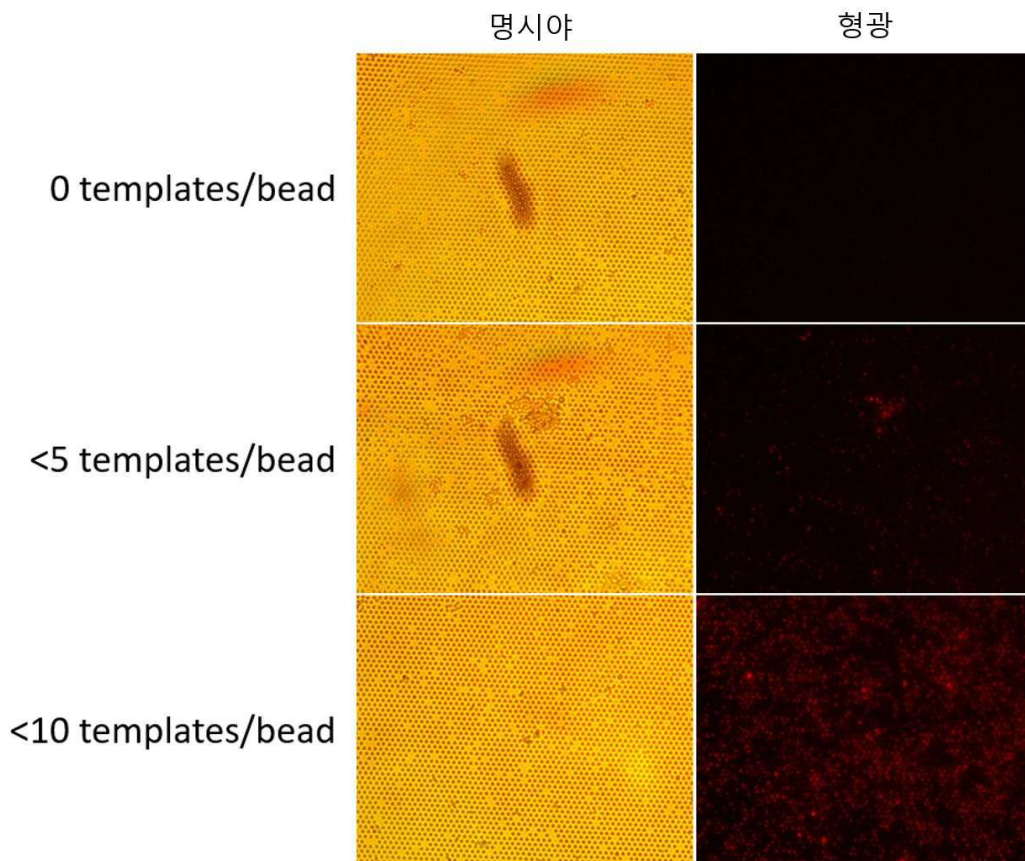
도면23



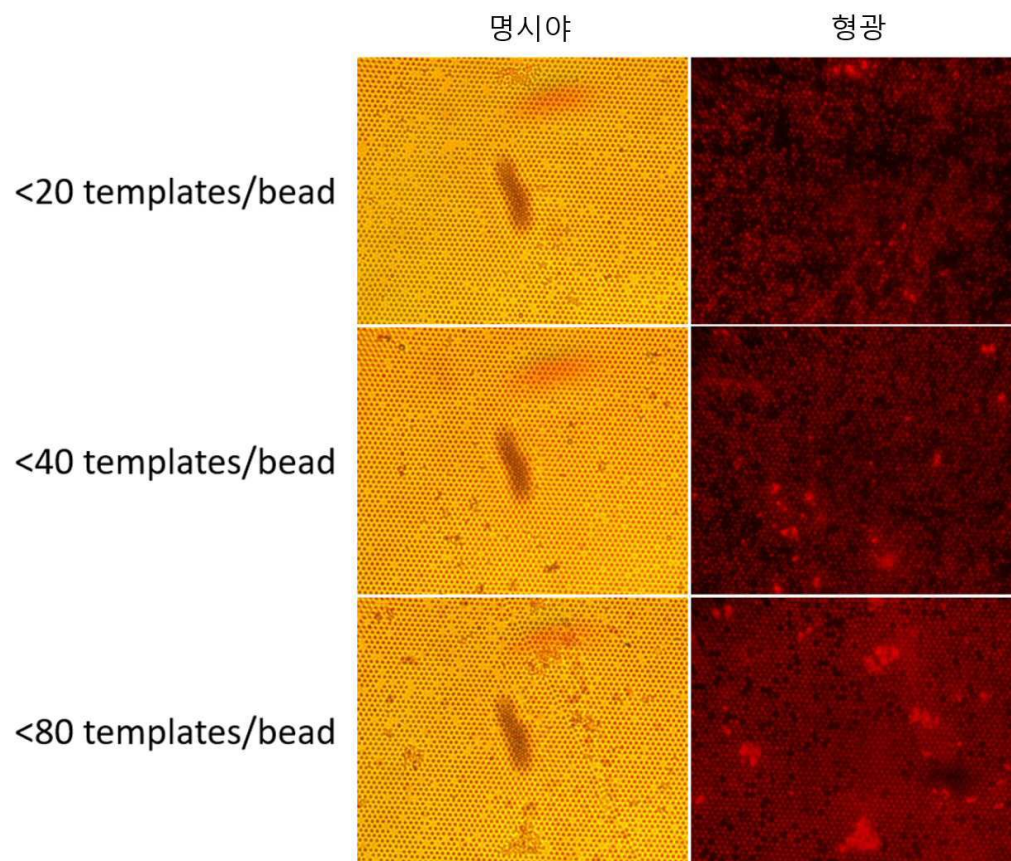
도면24



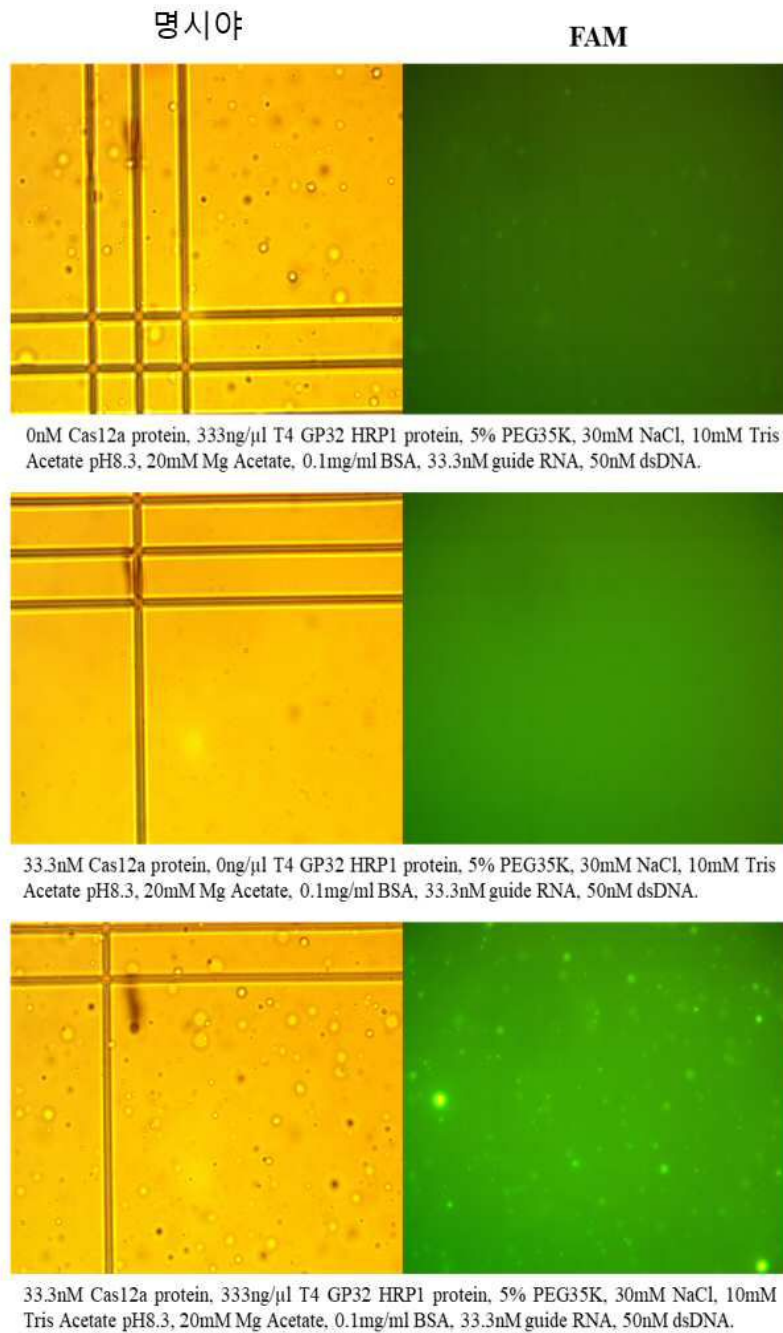
도면25a



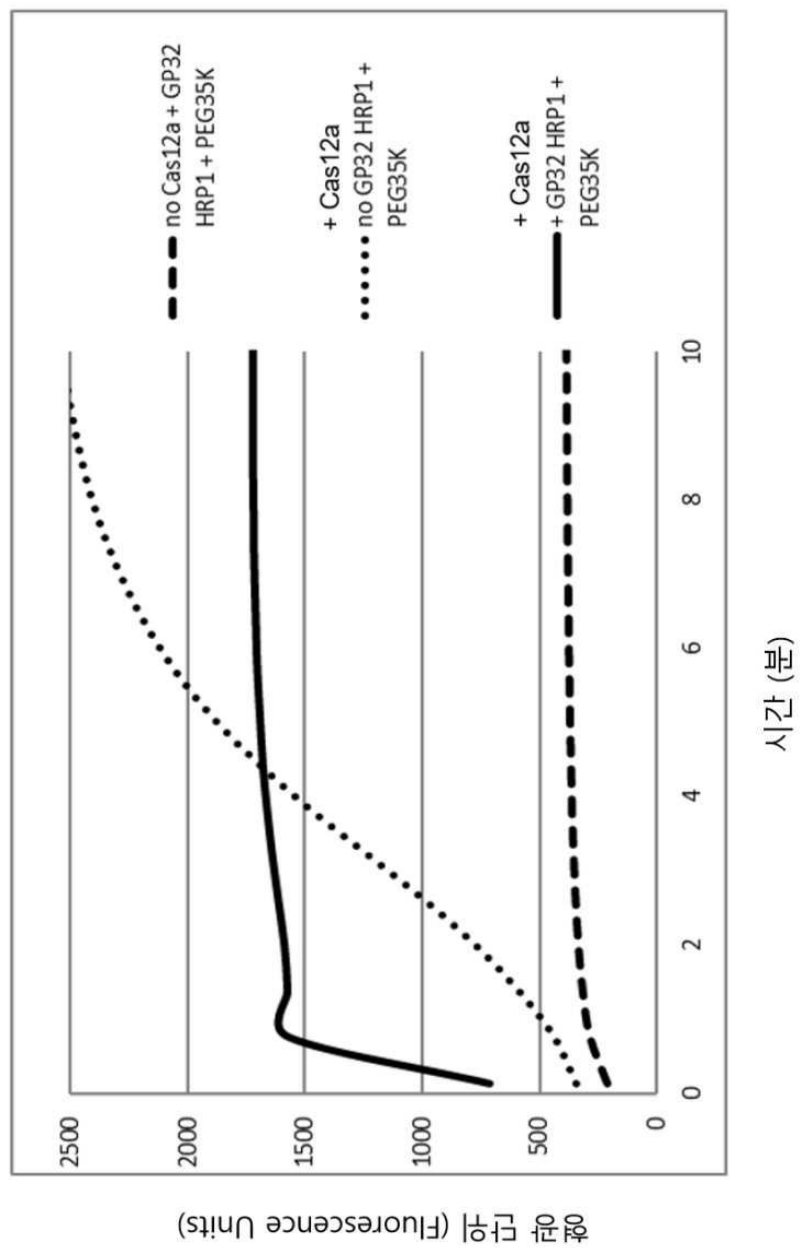
도면25b



도면26a



도면26b



서열 목록

SEQUENCE LISTING

<110> Biocrucible Limited

<120> BIOCHEMICAL REACTION METHODS AND REAGENTS COMPRISING
INTRINSICALLY DISORDERED REGIONS

<130> N417226WO

<150> GB 1916379.9

<151> 2019-11-11

<160> 137

<170> PatentIn version 3.5

<210> 1

<211> 43

<212> PRT

<213> Homo sapiens

<400> 1

Pro Gly Phe Ser Pro Arg Gly Gly Gly Phe Gly Gly Arg Gly Gly Phe

1 5 10 15

Gly Asp Arg Gly Gly Arg Gly Gly Arg Gly Gly Phe Gly Gly Gly Arg

20 25 30

Gly Arg Gly Gly Gly Phe Arg Gly Arg Gly Arg

35 40

<210> 2

<211> 86

<212> PRT

<213> Homo sapiens

<400> 2

Pro Gly Phe Ser Pro Arg Gly Gly Gly Phe Gly Gly Arg Gly Gly Phe

1 5 10 15

Gly Asp Arg Gly Gly Arg Gly Gly Arg Gly Gly Phe Gly Gly Gly Arg

20 25 30

Gly Arg Gly Gly Gly Phe Arg Gly Arg Gly Arg Pro Gly Phe Ser Pro

35 40 45

Arg Gly Gly Gly Phe Gly Gly Arg Gly Gly Phe Gly Asp Arg Gly Gly

50 55 60

Arg Gly Gly Arg Gly Gly Phe Gly Gly Gly Arg Gly Arg Gly Gly Gly

65 70 75 80

Phe Arg Gly Arg Gly Arg

85

<210> 3

<211> 25

<212> PRT

<213> Homo sapiens

<400> 3

Arg Gly Gly Gly Phe Gly Gly Arg Gly Gly Phe Gly Asp Arg Gly Gly

1 5 10 15

Arg Gly Gly Arg Gly Gly Phe Gly Gly

20 25

<210> 4

<211> 36

<212> PRT

<213> Artificial Sequence

<220><223> fib-2 (clipper1) - 7 amino acid deletion at the C-terminus of
fib-1

<400> 4

Pro Gly Phe Ser Pro Arg Gly Gly Gly Phe Gly Gly Arg Gly Gly Phe

1 5 10 15

Gly Asp Arg Gly Gly Arg Gly Gly Arg Gly Gly Phe Gly Gly Gly Arg

20 25 30

Gly Arg Gly Gly

35

<210> 5

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> fib-3 (clipper2) - 16 amino acids deletion at the C-terminus of
fib-1

<400> 5

Pro Gly Phe Ser Pro Arg Gly Gly Gly Phe Gly Gly Arg Gly Gly Phe

1 5 10 15

Gly Asp Arg Gly Gly Arg Gly Gly Arg Gly Gly

20 25

<210> 6

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> fib-4 (clipper3) - 28 amino acids deletion at the C-terminus of

fib-1

<400> 6

Pro Gly Phe Ser Pro Arg Gly Gly Gly Phe Gly Gly Arg Gly Gly

1 5 10 15

<210> 7

<211> 67

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 7

Gln Val Gln Met Gln Leu Arg Gln Val Phe Ser Gln Asp Gln Gln Val

1 5 10 15

Leu Gln Glu Arg Met Arg Tyr His Glu Leu Gln Gln Gln Gln Gln

20 25 30

Gln Tyr His Glu Thr Lys Asp Met Val Gly Ser Tyr Thr Gln Asn Ser

35 40 45

Asn Ser Ala Ile Pro Leu Phe Gly Asn Asn Ser Asp Thr Thr Asn Gln

50 55 60

Gln Asn Ser

65

<210> 8

<211> 59

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 8

Gln Asn Pro Thr Gly Ile Ser Tyr Ser Gln Pro Gln Gln Gln Gln

1 5 10 15

Pro Gln Gln Gln Pro Gln Tyr Met Gln Asn Phe Gln Gln Gln Gln Pro

20 25 30

Gln Tyr Ala Gln Asn Phe Gln Gln Gln Pro Gln Tyr Thr Gln Asn Tyr

35 40 45

Gln Gln Gln Pro Gln Tyr Ile Gln Pro His Gln

50 55

<210> 9

<211> 58

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 9

Gly Gly Asn Asn Gly Gly Asn Asn Met Asn Arg Arg Gly Gly Asn Phe

1 5 10 15

Gly Asn Gln Gly Asp Phe Asn Gln Met Tyr Gln Asn Pro Met Met Gly

20 25 30

Gly Tyr Asn Pro Met Met Asn Pro Gln Ala Met Thr Asp Tyr Tyr Gln

35 40 45

Lys Met Gln Glu Tyr Tyr Gln Gln Met Gln

50 55

<210> 10

<211> 116

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 10

Gly Gly Asn Asn Gly Gly Asn Asn Met Asn Arg Arg Gly Gly Asn Phe

1 5 10 15

Gly Asn Gln Gly Asp Phe Asn Gln Met Tyr Gln Asn Pro Met Met Gly

20 25 30

Gly Tyr Asn Pro Met Met Asn Pro Gln Ala Met Thr Asp Tyr Tyr Gln

35 40 45

Lys Met Gln Glu Tyr Tyr Gln Gln Met Gln Gly Gly Asn Asn Gly Gly

50 55 60

Asn Asn Met Asn Arg Arg Gly Gly Asn Phe Gly Asn Gln Gly Asp Phe

65 70 75 80

Asn Gln Met Tyr Gln Asn Pro Met Met Gly Gly Tyr Asn Pro Met Met

85 90 95

Asn Pro Gln Ala Met Thr Asp Tyr Tyr Gln Lys Met Gln Glu Tyr Tyr

100 105 110
 Gln Gln Met Gln
 115
 <210> 11
 <211> 240
 <212> PRT
 <213> Homo sapiens
 <400> 11
 Met Gly Asp Glu Asp Trp Glu Ala Glu Ile Asn Pro His Met Ser Ser
 1 5 10 15
 Tyr Val Pro Ile Phe Glu Lys Asp Arg Tyr Ser Gly Glu Asn Gly Asp
 20 25 30
 Asn Phe Asn Arg Thr Pro Ala Ser Ser Ser Glu Met Asp Asp Gly Pro
 35 40 45
 Ser Arg Arg Asp His Phe Met Lys Ser Gly Phe Ala Ser Gly Arg Asn
 50 55 60
 Phe Gly Asn Arg Asp Ala Gly Glu Cys Asn Lys Arg Asp Asn Thr Ser
 65 70 75 80
 Thr Met Gly Gly Phe Gly Val Gly Lys Ser Phe Gly Asn Arg Gly Phe
 85 90 95
 Ser Asn Ser Arg Phe Glu Asp Gly Asp Ser Ser Gly Phe Trp Arg Glu
 100 105 110
 Ser Ser Asn Asp Cys Glu Asp Asn Pro Thr Arg Asn Arg Gly Phe Ser
 115 120 125
 Lys Arg Gly Gly Tyr Arg Asp Gly Asn Asn Ser Glu Ala Ser Gly Pro
 130 135 140
 Tyr Arg Arg Gly Gly Arg Gly Ser Phe Arg Gly Cys Arg Gly Gly Phe
 145 150 155 160
 Gly Leu Gly Ser Pro Asn Asn Asp Leu Asp Pro Asp Glu Cys Met Gln
 165 170 175
 Arg Thr Gly Gly Leu Phe Gly Ser Arg Arg Pro Val Leu Ser Gly Thr
 180 185 190

Gly Asn Gly Asp Thr Ser Gln Ser Arg Ser Gly Ser Gly Ser Glu Arg

195 200 205

Gly Gly Tyr Lys Gly Leu Asn Glu Glu Val Ile Thr Gly Ser Gly Lys

210 215 220

Asn Ser Trp Lys Ser Glu Ala Glu Gly Gly Glu Ser Ser Asp Thr Gln

225 230 235 240

<210> 12

<211> 134

<212> PRT

<213> Homo sapiens

<400> 12

Ala Ser Ala Ser Ser Ser Gln Arg Gly Arg Ser Gly Ser Gly Asn Phe

1 5 10 15

Gly Gly Gly Arg Gly Gly Gly Phe Gly Gly Asn Asp Asn Phe Gly Arg

20 25 30

Gly Gly Asn Phe Ser Gly Arg Gly Gly Phe Gly Gly Ser Arg Gly Gly

35 40 45

Gly Gly Tyr Gly Gly Ser Gly Asp Gly Tyr Asn Gly Phe Gly Asn Asp

50 55 60

Gly Ser Asn Phe Gly Gly Gly Gly Ser Tyr Asn Asp Phe Gly Asn Tyr

65 70 75 80

Asn Asn Gln Ser Ser Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly

85 90 95

Gly Arg Ser Ser Gly Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys

100 105 110

Pro Gln Asn Gln Gly Gly Tyr Gly Val Ser Ser Ser Ser Ser Tyr

115 120 125

Gly Ser Gly Arg Arg Phe

130

<210> 13

<211> 204

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 13

Ser Pro Phe Gly Ala Tyr Gly Glu Ala Pro Thr Ser Pro Gly Phe Gly

1 5 10 15

Val Ser Ser Pro Gly Phe Ser Pro Thr Ser Pro Thr Tyr Ser Pro Thr

20 25 30

Ser Pro Ala Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro

35 40 45

Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr

50 55 60

Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro

65 70 75 80

Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser

85 90 95

Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser

100 105 110

Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser

115 120 125

Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ala Tyr Ser Pro Thr

130 135 140

Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro

145 150 155 160

Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Asn Tyr

165 170 175

Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Gly Tyr Ser Pro

180 185 190

Gly Ser Pro Ala Tyr Ser Pro Lys Gln Asp Glu Gln

195 200

<210> 14

<211> 42

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 14

Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser

1	5	10	15
Pro Gly Tyr Ser Pro Thr Ser Pro Ala Tyr Ser Pro Thr Ser Pro Thr			
	20	25	30
Tyr Ser Pro Thr Ser Pro Thr Tyr Ser Pro			
	35	40	

<210> 15

<211> 49

<212> PRT

<213> Artificial Sequence

<220><223> mimic1 - seven repeats of YDPTSPS motifs mimiking the c-terminus
of RNA polymerase II

<400> 15

Tyr Asp Pro Thr Ser Pro Ser Tyr Asp Pro Thr Ser Pro Ser Tyr Asp

1	5	10	15
Pro Thr Ser Pro Ser Tyr Asp Pro Thr Ser Pro Ser Tyr Asp Pro Thr			
	20	25	30
Ser Pro Ser Tyr Asp Pro Thr Ser Pro Ser Tyr Asp Pro Thr Ser Pro			
	35	40	45
Ser			

<210> 16

<211> 49

<212> PRT

<213> Artificial Sequence

<220><223> mimic2 - seven repeats of YSPTDPS motifs mimiking the c-terminus
of RNA polymerase II

<400> 16

Tyr Ser Pro Thr Asp Pro Ser Tyr Ser Pro Thr Asp Pro Ser Tyr Ser

1	5	10	15
Pro Thr Asp Pro Ser Tyr Ser Pro Thr Asp Pro Ser Tyr Ser Pro Thr			

20 25 30
 Asp Pro Ser Tyr Ser Pro Thr Asp Pro Ser Tyr Ser Pro Thr Asp Pro
 35 40 45

Ser

<210> 17

<211> 21

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 17

Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Tyr

1 5 10 15

Ser Pro Ala Tyr Ser

20

<210> 18

<211> 115

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 18

Met Ser Asp Ser Asn Gln Gly Asn Asn Gln Gln Asn Tyr Gln Gln Tyr

1 5 10 15

Ser Gln Asn Gly Asn Gln Gln Gln Gly Asn Asn Arg Tyr Gln Gly Tyr

20 25 30

Gln Ala Tyr Asn Ala Gln Ala Gln Pro Ala Gly Gly Tyr Tyr Gln Asn

35 40 45

Tyr Gln Gly Tyr Ser Gly Tyr Gln Gln Gly Gly Tyr Gln Gln Tyr Gln

50 55 60

Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr

65 70 75 80

Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Gln Phe Asn Pro Gln

85 90 95

Gly Gly Arg Gly Asn Tyr Lys Asn Phe Asn Tyr Asn Asn Asn Leu Gln

100 105 110

Gly Tyr Gln

115

<210> 19

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Sup1 - single repeat of YNPQGGYQQ

<400> 19

Tyr Asn Pro Gln Gly Gly Tyr Gln Gln

1 5

<210> 20

<211> 18

<212> PRT

<213> Artificial Sequence

<220><223> Sup2 - two repeat of YNPQGGYQQ

<400> 20

Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr

1 5 10 15

Gln Gln

<210> 21

<211> 27

<212> PRT

<213> Artificial Sequence

<220><223> Sup3 - three repeat of YNPQGGYQQ

<400> 21

Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr

1 5 10 15

Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln

20 25

<210> 22

<211> 36

<212> PRT

<213> Artificial Sequence

<220><223> Sup4 - four repeat of YNPQGGYQQ

<400> 22

Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr

1 5 10 15
Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro Gln Gly

20 25 30
Gly Tyr Gln Gln

35

<210> 23

<211> 31

<212> PRT

<213> Homo sapiens

<400> 23

Pro Gln Gln Asn Ala Leu His His His His Gly Asn Ser Ser His His

1 5 10 15
His His His His His His His His His His His Gly Gln Gln Ala
20 25 30

<210> 24

<211> 36

<212> PRT

<213> Homo sapiens

<400> 24

Ala Gly His His His His His Pro His Ala His His Pro Leu Ser Gln

1 5 10 15
Ser Ser Gly His His His His His His His His His His Gln Gly Tyr
20 25 30

Gly Gly Ser Gly

35

<210> 25

<211> 43

<212> PRT

<213> Homo sapiens

<400> 25

Ala Gly His His His His His His His His Leu Pro His Leu Pro Pro
1 5 10 15

Pro His Leu His His His His His Pro Gln His His Leu His Pro Gly
20 25 30

Ser Ala Ala Ala Val His Pro Val Gln Gln His
35 40

<210> 26

<211> 42

<212> PRT

<213> Homo sapiens

<400> 26

Asn His His Ala Ser His Gly His His Asn Ser His His Pro Gln His
1 5 10 15

His His His His His His His His His His Pro Pro Pro Pro Ala Pro
20 25 30

Gln Pro Pro Pro Pro Pro Gln Gln Gln Gln
35 40

<210> 27

<211> 76

<212> PRT

<213> Homo sapiens

<400> 27

Ser Gly His His Gly Ala His His Gly Ala His His Pro Ala Ala Ala
1 5 10 15

Ala Ala Tyr Glu Ala Phe Arg Gly Pro Gly Phe Ala Gly Gly Gly Gly
20 25 30

Ala Asp Asp Met Gly Ala Gly His His His Gly Ala His His Ala Ala
35 40 45

His His His His Ala Ala His His His His His His His His His His
50 55 60

Gly Gly Ala Gly His Gly Gly Gly Ala Gly His His

<210>	28
<211>	27
<212>	PRT
<213>	Homo sapiens
<400>	28
Gly Gly Gly Gly Asp His Gly Gly Gly Gly Gly His Gly His Ala Gly	
1 5 10 15	
His His His His His His His His His His His His	
	20 25
<210>	29
<211>	25
<212>	PRT
<213>	Homo sapiens
<400>	29
Pro Gly Ser Gly Gly Pro His Thr Ser His His His His His His His	
1 5 10 15	
His His His His His Gln Ser Arg His	
	20 25
<210>	30
<211>	25
<212>	PRT
<213>	Homo sapiens
<400>	30
Ala Pro Ser Tyr His His His His His His His His His His His His	
1 5 10 15	
His His Gln Gln Gln His Gln Gln Pro	
	20 25
<210>	31
<211>	24
<212>	PRT
<213>	Homo sapiens
<400>	31

Thr His Pro Pro Ser His His Pro His Pro His Pro His His His His

1 5 10 15

His His His His His His His His

20

<210> 32

<211> 38

<212> PRT

<213> Homo sapiens

<400> 32

Pro Pro Pro His His His His His His His His His His His His His

1 5 10 15

His His Arg Ala Gln Pro Pro Gln Gln Ser His His Pro Pro His His

20 25 30

His Arg Pro Gln Pro His

35

<210> 33

<211> 18

<212> PRT

<213> Homo sapiens

<400> 33

Asn His Ala His His His His His Pro Pro Gln His His His His His

1 5 10 15

His Gln

<210> 34

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> His12

<400> 34

Ala Gly His His His His His His His His His His Gly Gly Ala Gly

1 5 10 15

Gly Ala Gly Gly Ala Gly Gly Ala His His His His His His His His

20 25 30

His His Gly Gly Ala

35

<210> 35

<211> 37

<212> PRT

<213> Artificial Sequence

<220><223> His13

<400> 35

Ala Gly His His His His His His His His His Ser Ser Gly Ser

1 5 10 15

Gly Ser Ser Gly Ser Gly Ser Ser His His His His His His His His

20 25 30

His His Gly Gly Ala

35

<210> 36

<211> 17

<212> PRT

<213> Artificial Sequence

<220><223> His14

<400> 36

Ala Gly His His His His His His His His His His His His Gly Gly

1 5 10 15

Ala

<210> 37

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> His15

<400> 37

Ala Gly His His His His His His His His His His His His His His

1 5 10 15

His His His His His Gly Gly Ala

20

<210> 38

<211> 74

<212> PRT

<213> Escherichia phage

<400> 38

Lys Lys Val Ala Ser Lys Leu Lys Phe Lys Lys Lys Met Glu Ala Phe

1 5 10 15

Ser Ser Ala Lys Thr Lys Lys Lys Phe Met Ser Ser Ser Ser Ser Lys

20 25 30

Lys Ser Lys Leu Lys Lys Leu Leu Ala Gly Leu Met Glu Ala Phe Ser

35 40 45

Ser Ala Lys Thr Lys Lys Lys Phe Met Ser Ser Ser Ser Ser Lys Lys

50 55 60

Ser Lys Leu Lys Lys Leu Leu Ala Gly Leu

65 70

<210> 39

<211> 75

<212> PRT

<213> Escherichia phage

<400> 39

Asp Asp Val Ala Ser Asp Leu Asp Asp Phe Asp Asp Asp Met Glu Ala

1 5 10 15

Phe Ser Ser Ala Asp Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser

20 25 30

Asp Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Met Glu Ala Phe

35 40 45

Ser Ser Ala Asp Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp

50 55 60

Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu

65 70 75

<210> 40

<211> 90

<212> PRT

<213> Escherichia phage

<400> 40

Arg Ile Val Lys Ala Lys Val Lys Arg Leu Ile Asn Ser Lys Val Lys

1 5 10 15

Lys Phe Lys Ser Pro Lys Ser Lys Ser Lys Ser Ala Ala Lys Leu Lys

20 25 30

Thr Lys Leu Lys Gln Leu Ser Lys Met Lys Lys Phe Asn Lys Ile Val

35 40 45

Lys Ala Lys Val Lys Lys Leu Ile Asn Ser Lys Val Arg Lys Phe Lys

50 55 60

Ser Pro Lys Ser Lys Ser Lys Ser Ala Ala Lys Leu Lys Thr Lys Leu

65 70 75 80

Lys Gln Leu Ser Lys Met Lys Glu Phe Asn

85 90

<210> 41

<211> 90

<212> PRT

<213> Escherichia phage

<400> 41

Glu Ile Val Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Glu Val Glu

1 5 10 15

Glu Phe Asp Ser Pro Glu Ser Asp Ser Asp Glu Ala Ala Asp Leu Glu

20 25 30

Thr Asp Leu Glu Gln Leu Ser Asp Met Glu Glu Phe Asn Glu Ile Val

35 40 45

Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Glu Val Glu Asp Phe Asp

50 55 60

Ser Pro Glu Ser Asp Ser Asp Glu Ala Ala Asp Leu Glu Thr Asp Leu

65 70 75 80

Glu Gln Leu Ser Asp Met Glu Glu Phe Asn

85 90

<210> 42

<211> 90

<212> PRT

<213> Escherichia phage

<400> 42

Glu Ile Val Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu

1 5 10 15
Lys Phe Lys Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu
20 25 30
Thr Asp Leu Glu Gln Leu Ser Asp Met Glu Glu Phe Asn Glu Ile Val
35 40 45
Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys
50 55 60
Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu

65 70 75 80
Glu Gln Leu Ser Asp Met Glu Glu Phe Asn
85 90

<210> 43

<211> 135

<212> PRT

<213> Escherichia phage

<400> 43

Glu Ile Val Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu
1 5 10 15
Lys Phe Lys Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu
20 25 30
Thr Asp Leu Glu Gln Leu Ser Asp Met Glu Glu Phe Asn Glu Ile Val
35 40 45
Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys
50 55 60
Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu

Glu Gln Leu Ser Asp Met Glu Glu Phe Asn Glu Ile Val Glu Ala Glu
85 90 95
Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro Glu

100 105 110
Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln Leu
115 120 125
Ser Asp Met Glu Glu Phe Asn
130 135

<210> 44

<211> 400

<212> PRT

<213> Artificial Sequence

<220><223> UvsX (7His)

<400> 44

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser
1 5 10 15
Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20					25					30						
Val	Ile	Arg	Thr	Lys	Ile	Pro	Met	Leu	Asn	Ile	Ala	Ile	Ser	Gly	Ala	
35					40					45						
Ile	Asp	Gly	Gly	Met	Gln	Ser	Gly	Leu	Thr	Ile	Phe	Ala	Gly	Pro	Ser	
50					55					60						
Lys	Ser	Phe	Lys	Ser	Asn	Met	Ser	Leu	Thr	Met	Val	Ala	Ala	Tyr	Leu	
65					70					75					80	
Asn	Lys	Tyr	Pro	Asp	Ala	Ile	Cys	Leu	Phe	Tyr	Asp	Ser	Glu	Phe	Gly	

85					90					95					
Ile	Thr	Pro	Ala	Tyr	Leu	Lys	Ser	Met	Gly	Val	Asp	Pro	Glu	Arg	Val
100					105					110					
Ile	His	Thr	Pro	Ile	Gln	Ser	Val	Glu	Gln	Leu	Lys	Ile	Asp	Met	Val
115					120					125					
Asn	Gln	Leu	Glu	Thr	Ile	Glu	Arg	Gly	Glu	Lys	Val	Ile	Val	Phe	Ile

130 135 140
 Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

 145 150 155 160
 Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

 165 170 175
 Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

 180 185 190
 Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr

 195 200 205
 Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe

 210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser

 245 250 255
 Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser

 260 265 270
 Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys

 275 280 285
 Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
 290 295 300
 Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
 305 310 315 320
 Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

 325 330 335
 Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala

 340 345 350
 Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
 355 360 365
 Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln
 370 375 380

Leu Ser Asp Met Glu Glu Phe Asn Glu His His His His His His His

385 390 395 400

<210> 45

<211> 443

<212> PRT

<213> Artificial Sequence

<220><223> UvsX-fib-1

<400> 45

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50 55 60

Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu

65 70 75 80

Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly

85 90 95

Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val

100 105 110

Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

115 120 125

Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140

Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145 150 155 160

Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

165 170 175

Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

180 185 190
 Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
 195 200 205
 Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
 210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
 245 250 255

 Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
 260 265 270
 Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
 275 280 285
 Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
 290 295 300
 Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
 305 310 315 320

 Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
 325 330 335
 Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
 340 345 350
 Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
 355 360 365
 Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln
 370 375 380

 Leu Ser Asp Met Glu Glu Phe Asn Glu Pro Gly Phe Ser Pro Arg Gly
 385 390 395 400
 Gly Gly Phe Gly Gly Arg Gly Gly Phe Gly Asp Arg Gly Gly Arg Gly
 405 410 415
 Gly Arg Gly Gly Phe Gly Gly Gly Arg Gly Arg Gly Gly Gly Phe Arg
 420 425 430

Gly Arg Gly Arg His His His His His His His

435 440

<210> 46

<211> 437

<212> PRT

<213> Artificial Sequence

<220><223> UvsX-fib-2

<400> 46

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50 55 60

Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu

65 70 75 80

Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly

85 90 95

Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val

100 105 110

Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

115 120 125

Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140

Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145 150 155 160

Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

165 170 175

Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

180 185 190
 Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
 195 200 205
 Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
 210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
 245 250 255
 Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
 260 265 270
 Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
 275 280 285
 Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
 290 295 300
 Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
 305 310 315 320
 Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
 325 330 335
 Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
 340 345 350
 Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
 355 360 365
 Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln
 370 375 380
 Leu Ser Asp Met Glu Glu Phe Asn Glu Pro Gly Phe Ser Pro Arg Gly
 385 390 395 400
 Gly Gly Phe Gly Gly Arg Gly Gly Phe Gly Asp Arg Gly Gly Arg Gly
 405 410 415
 Gly Arg Gly Gly Phe Gly Gly Gly Arg Gly Arg Gly Gly Val Glu His
 420 425 430
 His His His His His

435

<210

> 47

<211> 428

<212> PRT

<213> Artificial Sequence

<220><223> UvsX-fib-3

<400> 47

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50 55 60

Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu

65 70 75 80

Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly

85 90 95

Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val

100 105 110

Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

115 120 125

Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140

Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145 150 155 160

Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

165 170 175

Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

180 185 190

Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
195 200 205

Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
210 215 220

Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
225 230 235 240

Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
245 250 255

Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
260 265 270

Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
275 280 285

Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
290 295 300

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
305 310 315 320

Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
325 330 335

Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
340 345 350

Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
355 360 365

Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln
370 375 380

Leu Ser Asp Met Glu Glu Phe Asn Glu Pro Gly Phe Ser Pro Arg Gly
385 390 395 400

Gly Gly Phe Gly Gly Arg Gly Gly Phe Gly Asp Arg Gly Gly Arg Gly
405 410 415

Gly Arg Gly Gly Val Glu His His His His His His
420 425

<210> 48

<211> 416

<212> PRT

<213> Artificial Sequence

<220><223> UvsX-fib-4

<400> 48

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50 55 60

Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu

65 70 75 80

Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly

85 90 95

Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val

100 105 110

Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

115 120 125

Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140

Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145 150 155 160

Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

165 170 175

Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

180 185 190

Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr

195 200 205

Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe

210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
 245 250 255

 Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
 260 265 270
 Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
 275 280 285
 Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
 290 295 300
 Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
 305 310 315 320

 Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
 325 330 335
 Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
 340 345 350
 Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
 355 360 365
 Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln
 370 375 380

 Leu Ser Asp Met Glu Glu Phe Asn Glu Pro Gly Phe Ser Pro Arg Gly
 385 390 395 400
 Gly Gly Phe Gly Gly Arg Gly Gly Val Glu His His His His His His
 405 410 415

 <210> 49
 <211> 534
 <212> PRT
 <213> Artificial Sequence
 <220><223> UvsX-HNRNPA1
 <400> 49
 Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15
 Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

 20 25 30
 Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
 35 40 45
 Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser
 50 55 60
 Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
 65 70 75 80
 Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly

 85 90 95
 Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val
 100 105 110
 Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val
 115 120 125
 Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile
 130 135 140
 Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

 145 150 155 160
 Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser
 165 170 175
 Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys
 180 185 190
 Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
 195 200 205
 Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe

 210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
 245 250 255

Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
 260 265 270
 Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
 275 280 285
 Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
 290 295 300
 Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
 305 310 315 320
 Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
 325 330 335
 Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
 340 345 350
 Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
 355 360 365
 Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln
 370 375 380
 Leu Ser Asp Met Glu Glu Phe Asn Glu Ala Ser Ala Ser Ser Ser Gln
 385 390 395 400
 Arg Gly Arg Ser Gly Ser Gly Asn Phe Gly Gly Gly Arg Gly Gly Gly
 405 410 415
 Phe Gly Gly Asn Asp Asn Phe Gly Arg Gly Gly Asn Phe Ser Gly Arg
 420 425 430
 Gly Gly Phe Gly Gly Ser Arg Gly Gly Gly Gly Tyr Gly Gly Ser Gly
 435 440 445
 Asp Gly Tyr Asn Gly Phe Gly Asn Asp Gly Ser Asn Phe Gly Gly Gly
 450 455 460
 Gly Ser Tyr Asn Asp Phe Gly Asn Tyr Asn Asn Gln Ser Ser Asn Phe
 465 470 475 480
 Gly Pro Met Lys Gly Gly Asn Phe Gly Gly Arg Ser Ser Gly Pro Tyr
 485 490 495
 Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro Gln Asn Gln Gly Gly Tyr

500 505 510
 Gly Val Ser Ser Ser Ser Ser Tyr Gly Ser Gly Arg Arg Phe His
 515 520 525
 His His His His His His
 530

<210> 50
 <211> 640
 <212> PRT
 <213> Artificial Sequence
 <220><223> UvsX-DDX
 <400> 50

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser
 1 5 10 15
 Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp
 20 25 30
 Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
 35 40 45
 Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50 55 60
 Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
 65 70 75 80
 Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly
 85 90 95
 Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val
 100 105 110
 Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

115 120 125
 Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile
 130 135 140
 Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu
 145 150 155 160
 Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

165 170 175
 Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

 180 185 190
 Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
 195 200 205
 Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
 210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser

 245 250 255
 Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
 260 265 270
 Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
 275 280 285
 Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
 290 295 300
 Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

 305 310 315 320
 Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
 325 330 335
 Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
 340 345 350
 Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
 355 360 365
 Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

 370 375 380
 Leu Ser Asp Met Glu Glu Phe Asn Glu Met Gly Asp Glu Asp Trp Glu
 385 390 395 400
 Ala Glu Ile Asn Pro His Met Ser Ser Tyr Val Pro Ile Phe Glu Lys
 405 410 415

Asp Arg Tyr Ser Gly Glu Asn Gly Asp Asn Phe Asn Arg Thr Pro Ala

420 425 430

Ser Ser Ser Glu Met Asp Asp Gly Pro Ser Arg Arg Asp His Phe Met

435 440 445

Lys Ser Gly Phe Ala Ser Gly Arg Asn Phe Gly Asn Arg Asp Ala Gly

450 455 460

Glu Cys Asn Lys Arg Asp Asn Thr Ser Thr Met Gly Gly Phe Gly Val

465 470 475 480

Gly Lys Ser Phe Gly Asn Arg Gly Phe Ser Asn Ser Arg Phe Glu Asp

485 490 495

Gly Asp Ser Ser Gly Phe Trp Arg Glu Ser Ser Asn Asp Cys Glu Asp

500 505 510

Asn Pro Thr Arg Asn Arg Gly Phe Ser Lys Arg Gly Gly Tyr Arg Asp

515 520 525

Gly Asn Asn Ser Glu Ala Ser Gly Pro Tyr Arg Arg Gly Gly Arg Gly

530 535 540

Ser Phe Arg Gly Cys Arg Gly Gly Phe Gly Leu Gly Ser Pro Asn Asn

545 550 555 560

Asp Leu Asp Pro Asp Glu Cys Met Gln Arg Thr Gly Gly Leu Phe Gly

565 570 575

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

580 585 590

Ser Arg Ser Gly Ser Gly Ser Glu Arg Gly Gly Tyr Lys Gly Leu Asn

595 600 605

Glu Glu Val Ile Thr Gly Ser Gly Lys Asn Ser Trp Lys Ser Glu Ala

610 615 620

Glu Gly Gly Glu Ser Ser Asp Thr Gln His His His His His His His

625 630 635 640

<210> 51

<211> 618

<212> PRT

<213> Artificial Sequence

<220><223> UvsX addPolCTD

<400> 51

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50 55 60

Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu

65 70 75 80

Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly

85 90 95

Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val

100 105 110

Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

115 120 125

Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140

Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145 150 155 160

Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

165 170 175

Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

180 185 190

Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr

195 200 205

Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe

210 215 220

Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr

225						230						235						240	
Gln	Phe	Val	Leu	Asn	Val	Glu	Lys	Ser	Arg	Thr	Val	Lys	Glu	Lys	Ser				
					245						250						255		
Lys	Phe	Phe	Ile	Asp	Val	Lys	Phe	Asp	Gly	Gly	Ile	Asp	Pro	Tyr	Ser				
					260						265						270		
Gly	Leu	Leu	Asp	Met	Ala	Leu	Glu	Leu	Gly	Phe	Val	Val	Lys	Pro	Lys				
					275						280						285		
Asn	Gly	Trp	Tyr	Ala	Arg	Glu	Phe	Leu	Asp	Glu	Glu	Thr	Gly	Glu	Met				
					290						295						300		
Ile	Arg	Glu	Glu	Lys	Ser	Trp	Arg	Ala	Lys	Asp	Thr	Asn	Cys	Thr	Thr				
					310						315						320		
Phe	Trp	Gly	Pro	Leu	Phe	Lys	His	Gln	Pro	Phe	Arg	Asp	Ala	Ile	Lys				
					325						330						335		
Arg	Ala	Tyr	Gln	Leu	Gly	Ala	Ile	Asp	Ser	Asn	Glu	Ile	Val	Glu	Ala				
					340						345						350		
Glu	Val	Asp	Glu	Leu	Ile	Asn	Ser	Lys	Val	Glu	Lys	Phe	Lys	Ser	Pro				
					355						360						365		
Glu	Ser	Lys	Ser	Lys	Ser	Ala	Ala	Asp	Leu	Glu	Thr	Asp	Leu	Glu	Gln				
					370						375						380		
Leu	Ser	Asp	Met	Glu	Glu	Phe	Asn	Glu	His	His	His	His	His	His	His				
					385						390						395	400	
Phe	Glu	Phe	Ser	Pro	Phe	Gly	Ala	Tyr	Gly	Glu	Ala	Pro	Thr	Ser	Pro				
					405						410						415		
Gly	Phe	Gly	Val	Ser	Ser	Pro	Gly	Phe	Ser	Pro	Thr	Ser	Pro	Thr	Tyr				
					420						425						430		
Ser	Pro	Thr	Ser	Pro	Ala	Tyr	Ser	Pro	Thr	Ser	Pro	Ser	Tyr	Ser	Pro				
					435						440						445		
Thr	Ser	Pro	Ser	Tyr	Ser	Pro	Thr	Ser	Pro	Ser	Tyr	Ser	Pro	Thr	Ser				
					450						455						460		
Pro	Ser	Tyr	Ser	Pro	Thr	Ser	Pro	Ser	Tyr	Ser	Pro	Thr	Ser	Pro	Ser				
					465						470						475	480	

Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser
485 490 495

Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr
500 505 510

Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro
515 520 525

Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ala Tyr
530 535 540

Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro
545 550 555 560

Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser
565 570 575

Pro Asn Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Gly
580 585 590

Tyr Ser Pro Gly Ser Pro Ala Tyr Ser Pro Lys Gln Asp Glu Gln Thr
595 600 605

Ala Ala Leu Glu His His His His His His
610 615

<210> 52

<211> 442

<212>

> PRT

<213> Artificial Sequence

<220><223> UvsX-fusPolII

<400> 52

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser
1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp
20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50				55				60							
Lys	Ser	Phe	Lys	Ser	Asn	Met	Ser	Leu	Thr	Met	Val	Ala	Ala	Tyr	Leu
65				70				75				80			
Asn	Lys	Tyr	Pro	Asp	Ala	Ile	Cys	Leu	Phe	Tyr	Asp	Ser	Glu	Phe	Gly
85				90				95							
Ile	Thr	Pro	Ala	Tyr	Leu	Lys	Ser	Met	Gly	Val	Asp	Pro	Glu	Arg	Val
100				105				110							
Ile	His	Thr	Pro	Ile	Gln	Ser	Val	Glu	Gln	Leu	Lys	Ile	Asp	Met	Val
115				120				125							
Asn	Gln	Leu	Glu	Thr	Ile	Glu	Arg	Gly	Glu	Lys	Val	Ile	Val	Phe	Ile
130				135				140							
Asp	Ser	Ile	Gly	Asn	Met	Ala	Ser	Lys	Lys	Glu	Thr	Glu	Asp	Ala	Leu
145				150				155				160			
Asn	Glu	Lys	Ser	Val	Ala	Asp	Met	Thr	Arg	Ala	Lys	Ser	Leu	Lys	Ser
165				170				175							
Leu	Phe	Arg	Ile	Val	Thr	Pro	Tyr	Phe	Ser	Ile	Lys	Asn	Ile	Pro	Cys
180				185				190							
Val	Ala	Val	Asn	His	Thr	Ile	Glu	Thr	Ile	Glu	Met	Phe	Ser	Lys	Thr
195				200				205							
Val	Met	Thr	Gly	Gly	Thr	Gly	Val	Met	Tyr	Ser	Ala	Asp	Thr	Val	Phe
210				215				220							
Ile	Ile	Gly	Lys	Arg	Gln	Ile	Lys	Asp	Gly	Ser	Asp	Leu	Gln	Gly	Tyr
225				230				235				240			
Gln	Phe	Val	Leu	Asn	Val	Glu	Lys	Ser	Arg	Thr	Val	Lys	Glu	Lys	Ser
245				250				255							
Lys	Phe	Phe	Ile	Asp	Val	Lys	Phe	Asp	Gly	Gly	Ile	Asp	Pro	Tyr	Ser
260				265				270							
Gly	Leu	Leu	Asp	Met	Ala	Leu	Glu	Leu	Gly	Phe	Val	Val	Lys	Pro	Lys
275				280				285							
Asn	Gly	Trp	Tyr	Ala	Arg	Glu	Phe	Leu	Asp	Glu	Glu	Thr	Gly	Glu	Met
290				295				300							

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

305 310 315 320

Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

325 330 335

Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala

340 345 350

Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro

355 360 365

Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

370 375 380

Leu Ser Asp Met Glu Glu Phe Asn Glu Thr Ser Pro Ser Tyr Ser Pro

385 390 395 400

Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Gly Tyr Ser Pro Thr Ser

405 410 415

Pro Ala Tyr Ser Pro Thr Ser Pro Thr Tyr Ser Pro Thr Ser Pro Thr

420 425 430

Tyr Ser Pro His His His His His His His

435 440

<210> 53

<211> 467

<212> PRT

<213> Artificial Sequence

<220><223> UvsX-PCF11

<400> 53

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50				55				60							
Lys	Ser	Phe	Lys	Ser	Asn	Met	Ser	Leu	Thr	Met	Val	Ala	Ala	Tyr	Leu
65				70				75				80			
Asn	Lys	Tyr	Pro	Asp	Ala	Ile	Cys	Leu	Phe	Tyr	Asp	Ser	Glu	Phe	Gly
85				90				95							
Ile	Thr	Pro	Ala	Tyr	Leu	Lys	Ser	Met	Gly	Val	Asp	Pro	Glu	Arg	Val
100				105				110							
Ile	His	Thr	Pro	Ile	Gln	Ser	Val	Glu	Gln	Leu	Lys	Ile	Asp	Met	Val
115				120				125							
Asn	Gln	Leu	Glu	Thr	Ile	Glu	Arg	Gly	Glu	Lys	Val	Ile	Val	Phe	Ile
130				135				140							
Asp	Ser	Ile	Gly	Asn	Met	Ala	Ser	Lys	Lys	Glu	Thr	Glu	Asp	Ala	Leu
145				150				155				160			
Asn	Glu	Lys	Ser	Val	Ala	Asp	Met	Thr	Arg	Ala	Lys	Ser	Leu	Lys	Ser
165				170				175							
Leu	Phe	Arg	Ile	Val	Thr	Pro	Tyr	Phe	Ser	Ile	Lys	Asn	Ile	Pro	Cys
180				185				190							
Val	Ala	Val	Asn	His	Thr	Ile	Glu	Thr	Ile	Glu	Met	Phe	Ser	Lys	Thr
195				200				205							
Val	Met	Thr	Gly	Gly	Thr	Gly	Val	Met	Tyr	Ser	Ala	Asp	Thr	Val	Phe
210				215				220							
Ile	Ile	Gly	Lys	Arg	Gln	Ile	Lys	Asp	Gly	Ser	Asp	Leu	Gln	Gly	Tyr
225				230				235				240			
Gln	Phe	Val	Leu	Asn	Val	Glu	Lys	Ser	Arg	Thr	Val	Lys	Glu	Lys	Ser
245				250				255							
Lys	Phe	Phe	Ile	Asp	Val	Lys	Phe	Asp	Gly	Gly	Ile	Asp	Pro	Tyr	Ser
260				265				270							
Gly	Leu	Leu	Asp	Met	Ala	Leu	Glu	Leu	Gly	Phe	Val	Val	Lys	Pro	Lys
275				280				285							
Asn	Gly	Trp	Tyr	Ala	Arg	Glu	Phe	Leu	Asp	Glu	Glu	Thr	Gly	Glu	Met
290				295				300							

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

305 310 315 320

Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

325 330 335

Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala

340 345 350

Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro

355 360 365

Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

370 375 380

Leu Ser Asp Met Glu Glu Phe Asn Glu Gln Val Gln Met Gln Leu Arg

385 390 395 400

Gln Val Phe Ser Gln Asp Gln Gln Val Leu Gln Glu Arg Met Arg Tyr

405 410 415

His Glu Leu Gln Gln Gln Gln Gln Gln Tyr His Glu Thr Lys Asp

420 425 430

Met Val Gly Ser Tyr Thr Gln Asn Ser Asn Ser Ala Ile Pro Leu Phe

435 440 445

Gly Asn Asn Ser Asp Thr Thr Asn Gln Gln Asn Ser His His His His

450 455 460

His His His

465

<210> 54

<211> 515

<212> PRT

<213> Artificial Sequence

<220><223> UvsX-Sup

<400> 54

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30
 Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
 35 40 45
 Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser
 50 55 60
 Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
 65 70 75 80
 Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly
 85 90 95
 Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val
 100 105 110
 Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val
 115 120 125
 Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile
 130 135 140
 Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu
 145 150 155 160
 Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser
 165 170 175
 Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys
 180 185 190
 Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
 195 200 205
 Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
 210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
 245 250 255
 Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
 260 265 270

Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys

275 280 285
Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met

290 295 300
Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

305 310 315 320
Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

325 330 335
Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala

340 345 350
Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro

355 360 365
Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

370 375 380
Leu Ser Asp Met Glu Glu Phe Asn Glu Met Ser Asp Ser Asn Gln Gly

385 390 395 400
Asn Asn Gln Gln Asn Tyr Gln Gln Tyr Ser Gln Asn Gly Asn Gln Gln

405 410 415
Gln Gly Asn Asn Arg Tyr Gln Gly Tyr Gln Ala Tyr Asn Ala Gln Ala

420 425 430
Gln Pro Ala Gly Gly Tyr Tyr Gln Asn Tyr Gln Gly Tyr Ser Gly Tyr

435 440 445
Gln Gln Gly Gly Tyr Gln Gln Tyr Gln Tyr Asn Pro Gln Gly Gly Tyr

450 455 460
Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro Gln Gly

465 470 475 480
Gly Tyr Gln Gln Gln Phe Asn Pro Gln Gly Gly Arg Gly Asn Tyr Lys

485 490 495
Asn Phe Asn Tyr Asn Asn Asn Leu Gln Gly Tyr Gln His His His His

500 505 510
His His His

515

<210> 55

<211> 499

<212> PRT

<213> Artificial Sequence

<220><223> UvsX-DoubleX

<400> 55

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15
Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp
20 25 30
Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
35 40 45
Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser
50 55 60
Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu

65 70 75 80
Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly
85 90 95
Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val
100 105 110
Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val
115 120 125
Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140
Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu
145 150 155 160
Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser
165 170 175
Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys
180 185 190

Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr

195 200 205

Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe

210 215 220

Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr

225 230 235 240

Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser

245 250 255

Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser

260 265 270

Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys

275 280 285

Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met

290 295 300

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

305 310 315 320

Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

325 330 335

Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala

340 345 350

Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro

355 360 365

Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

370 375 380

Leu Ser Asp Met Glu Glu Phe Asn Glu His His His His His His His

385 390 395 400

Phe Glu Phe Glu Ile Val Glu Ala Glu Val Asp Glu Leu Ile Asn Ser

405 410 415

Lys Val Glu Lys Phe Lys Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala

420 425 430

Asp Leu Glu Thr Asp Leu Glu Gln Leu Ser Asp Met Glu Glu Phe Asn

435 440 445
Glu Ile Val Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu

450 455 460
Lys Phe Lys Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu
465 470 475 480
Thr Asp Leu Glu Gln Leu Ser Asp Met Glu Glu Phe Asn His His His
485 490 495
His His His

<210> 56

<211> 444

<212> PRT

<213> Artificial Sequence

<220><223> Superpositive

<400> 56

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15
Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp
20 25 30
Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
35 40 45
Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser
50 55 60
Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu

65 70 75 80
Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly
85 90 95
Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val
100 105 110
Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val
115 120 125

Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140

Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145 150 155 160

Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

165 170 175

Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

180 185 190

Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr

195 200 205

Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe

210 215 220

Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr

225 230 235 240

Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser

245 250 255

Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser

260 265 270

Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys

275 280 285

Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met

290 295 300

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

305 310 315 320

Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

325 330 335

Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Arg Ile Val Lys Ala

340 345 350

Lys Val Lys Arg Leu Ile Asn Ser Lys Val Lys Lys Phe Lys Ser Pro

355 360 365

Lys Ser Lys Ser Lys Ser Ala Ala Lys Leu Lys Thr Lys Leu Lys Gln

370 375 380
 Leu Ser Lys Met Lys Lys Phe Asn Lys Ile Val Lys Ala Lys Val Lys

 385 390 395 400
 Lys Leu Ile Asn Ser Lys Val Arg Lys Phe Lys Ser Pro Lys Ser Lys

 405 410 415
 Ser Lys Ser Ala Ala Lys Leu Lys Thr Lys Leu Lys Gln Leu Ser Lys

 420 425 430
 Met Lys Glu Phe Asn His His His His His His His

 435 440
 <210> 57
 <211> 444
 <212> PRT
 <213> Artificial Sequence
 <220><223> Supernegative

 <400> 57
 Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser
 1 5 10 15
 Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

 20 25 30
 Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

 35 40 45
 Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

 50 55 60

 Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
 65 70 75 80
 Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly

 85 90 95
 Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val

 100 105 110
 Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

 115 120 125

Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile
 130 135 140
 Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu
 145 150 155 160
 Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser
 165 170 175
 Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys
 180 185 190

 Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
 195 200 205
 Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
 210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
 245 250 255

 Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
 260 265 270
 Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
 275 280 285
 Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
 290 295 300
 Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
 305 310 315 320

 Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
 325 330 335
 Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
 340 345 350
 Glu Val Asp Glu Leu Ile Asn Ser Glu Val Glu Glu Phe Asp Ser Pro
 355 360 365
 Glu Ser Asp Ser Asp Glu Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

370 375 380
 Leu Ser Asp Met Glu Glu Phe Asn Glu Ile Val Glu Ala Glu Val Asp
 385 390 395 400
 Glu Leu Ile Asn Ser Glu Val Glu Asp Phe Asp Ser Pro Glu Ser Asp
 405 410 415
 Ser Asp Glu Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln Leu Ser Asp
 420 425 430
 Met Glu Glu Phe Asn His His His His His His His
 435 440
 <210> 58
 <211
 > 429
 <212> PRT
 <213> Artificial Sequence
 <220><223> UvsX His2
 <400> 58
 Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser
 1 5 10 15
 Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp
 20 25 30
 Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
 35 40 45
 Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

 50 55 60
 Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
 65 70 75 80
 Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly
 85 90 95
 Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val
 100 105 110
 Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

 115 120 125

Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile
 130 135 140
 Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu
 145 150 155 160
 Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser
 165 170 175
 Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys
 180 185 190
 Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
 195 200 205
 Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
 210 215 220
 Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
 225 230 235 240
 Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
 245 250 255
 Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
 260 265 270
 Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
 275 280 285
 Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
 290 295 300
 Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
 305 310 315 320
 Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
 325 330 335
 Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
 340 345 350
 Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
 355 360 365
 Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

370 375 380
 Leu Ser Asp Met Glu Glu Phe Asn Glu Ala Gly His His His His His
 385 390 395 400
 Pro His Ala His His Pro Leu Ser Gln Ser Ser Gly His His His His
 405 410 415
 His His His His His His Gln Gly Tyr Gly Gly Ser Gly
 420 425
 <210> 59
 <211> 496
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> UvsX+PCF+His2
 <400> 59
 Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser
 1 5 10 15
 Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp
 20 25 30
 Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
 35 40 45
 Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser
 50 55 60

 Lys Ser Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
 65 70 75 80
 Asn Lys Tyr Pro Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly
 85 90 95
 Ile Thr Pro Ala Tyr Leu Lys Ser Met Gly Val Asp Pro Glu Arg Val
 100 105 110
 Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val
 115 120 125

 Asn Gln Leu Glu Thr Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile
 130 135 140
 Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145	150	155	160
Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser			
	165	170	175
Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys			
	180	185	190
Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr			
	195	200	205
Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe			
	210	215	220
Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr			
225	230	235	240
Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser			
	245	250	255
Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser			
	260	265	270
Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys			
	275	280	285
Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met			
	290	295	300
Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr			
305	310	315	320
Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys			
	325	330	335
Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala			
	340	345	350
Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro			
	355	360	365
Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln			
	370	375	380
Leu Ser Asp Met Glu Glu Phe Asn Glu Gln Val Gln Met Gln Leu Arg			
385	390	395	400

Gln Val Phe Ser Gln Asp Gln Gln Val Leu Gln Glu Arg Met Arg Tyr
405 410 415
His Glu Leu Gln Gln Gln Gln Gln Gln Tyr His Glu Thr Lys Asp
420 425 430
Met Val Gly Ser Tyr Thr Gln Asn Ser Asn Ser Ala Ile Pro Leu Phe
435 440 445

Gly Asn Asn Ser Asp Thr Thr Asn Gln Gln Asn Ser Ala Gly His His
450 455 460
His His His Pro His Ala His His Pro Leu Ser Gln Ser Ser Gly His
465 470 475 480
His His His His His His His His His Gln Gly Tyr Gly Gly Ser Gly
485 490 495

<210> 60

<211> 145

<212> PRT

<213> Artificial Sequence

<220><223> UvsY [Phage ST0]

<400> 60

Met His His His His His His His Met Lys Leu Glu Asp Leu Gln Glu

1 5 10 15
Glu Leu Asp Ala Asp Leu Ala Ile Asp Met Ser Lys Leu Gln Tyr Glu
20 25 30
Thr Ala Asn Asn Val Lys Leu Tyr Ser Lys Trp Leu Arg Lys His Ser
35 40 45
Phe Ile Arg Lys Glu Met Leu Arg Ile Glu Thr Gln Lys Lys Thr Ala
50 55 60
Leu Lys Ala Arg Leu Asp Tyr Tyr Ser Gly Arg Gly Asp Gly Asp Glu

65 70 75 80
Phe Ser Met Asp Arg Tyr Glu Lys Ser Glu Met Lys Thr Val Leu Ala
85 90 95
Ala Asp Lys Asp Val Leu Lys Ile Glu Thr Thr Leu Gln Tyr Trp Gly
100 105 110

Ile Leu Leu Glu Phe Cys Ser Gly Ala Leu Asp Ala Val Lys Ser Arg

115 120 125

Ser Phe Ala Leu Lys His Ile Gln Asp Met Arg Glu Phe Glu Ala Gly

130 135 140

Gln

145

<210> 61

<211> 172

<212> PRT

<213> Artificial Sequence

<220><223> UvsY C-Pol CTD

<400> 61

Met His His His His His His His His Lys Leu Glu Asp Leu Gln Glu

1 5 10 15

Glu Leu Asp Ala Asp Leu Ala Ile Asp Met Ser Lys Leu Gln Tyr Glu

20 25 30

Thr Ala Asn Asn Val Lys Leu Tyr Ser Lys Trp Leu Arg Lys His Ser

35 40 45

Phe Ile Arg Lys Glu Met Leu Arg Ile Glu Thr Gln Lys Lys Thr Ala

50 55 60

Leu Lys Ala Arg Leu Asp Tyr Tyr Ser Gly Arg Gly Asp Gly Asp Glu

65 70 75 80

Phe Ser Met Asp Arg Tyr Glu Lys Ser Glu Met Lys Thr Val Leu Ala

85 90 95

Ala Asp Lys Asp Val Leu Lys Ile Glu Thr Thr Leu Gln Tyr Trp Gly

100 105 110

Ile Leu Leu Glu Phe Cys Ser Gly Ala Leu Asp Ala Val Lys Ser Arg

115 120 125

Ser Phe Ala Leu Lys His Ile Gln Asp Met Arg Glu Phe Glu Ala Gly

130 135 140

Gln Ser Gly Ser Gly Ser Gly Pro Thr Ser Pro Ser Tyr Ser Pro Thr

145 150 155 160

Ser Pro Ser Tyr Ser Pro Tyr Ser Pro Ala Tyr Ser

165

170

<210> 62

<211> 176

<212> PRT

<213> Artificial Sequence

<220><223> UvsY fib[short]

<400> 62

Met His His His His His His His His Lys Leu Glu Asp Leu Gln Glu

1 5 10 15

Glu Leu Asp Ala Asp Leu Ala Ile Asp Met Ser Lys Leu Gln Tyr Glu

20 25 30

Thr Ala Asn Asn Val Lys Leu Tyr Ser Lys Trp Leu Arg Lys His Ser

35 40 45

Phe Ile Arg Lys Glu Met Leu Arg Ile Glu Thr Gln Lys Lys Thr Ala

50 55 60

Leu Lys Ala Arg Leu Asp Tyr Tyr Ser Gly Arg Gly Asp Gly Asp Glu

65 70 75 80

Phe Ser Met Asp Arg Tyr Glu Lys Ser Glu Met Lys Thr Val Leu Ala

85 90 95

Ala Asp Lys Asp Val Leu Lys Ile Glu Thr Thr Leu Gln Tyr Trp Gly

100 105 110

Ile Leu Leu Glu Phe Cys Ser Gly Ala Leu Asp Ala Val Lys Ser Arg

115 120 125

Ser Phe Ala Leu Lys His Ile Gln Asp Met Arg Glu Phe Glu Ala Gly

130 135 140

Gln Ser Gly Ser Gly Ser Gly Arg Gly Gly Gly Phe Gly Gly Arg Gly

145 150 155 160

Gly Phe Gly Asp Arg Gly Gly Arg Gly Gly Arg Gly Gly Phe Gly Gly

165 170 175

<210> 63

<211> 161

<212> PRT

<213> Artificial Sequence

<220><223> UvsY Sup1

<400> 63

Met His His His His His His His His Lys Leu Glu Asp Leu Gln Glu

1 5 10 15

Glu Leu Asp Ala Asp Leu Ala Ile Asp Met Ser Lys Leu Gln Tyr Glu

20 25 30

Thr Ala Asn Asn Val Lys Leu Tyr Ser Lys Trp Leu Arg Lys His Ser

35 40 45

Phe Ile Arg Lys Glu Met Leu Arg Ile Glu Thr Gln Lys Lys Thr Ala

50 55 60

Leu Lys Ala Arg Leu Asp Tyr Tyr Ser Gly Arg Gly Asp Gly Asp Glu

65 70 75 80

Phe Ser Met Asp Arg Tyr Glu Lys Ser Glu Met Lys Thr Val Leu Ala

85 90 95

Ala Asp Lys Asp Val Leu Lys Ile Glu Thr Thr Leu Gln Tyr Trp Gly

100 105 110

Ile Leu Leu Glu Phe Cys Ser Gly Ala Leu Asp Ala Val Lys Ser Arg

115 120 125

Ser Phe Ala Leu Lys His Ile Gln Asp Met Arg Glu Phe Glu Ala Gly

130 135 140

Gln Ser Gly Ser Gly Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Asn Asn

145 150 155 160

Leu

<210> 64

<211> 190

<212> PRT

<213> Artificial Sequence

<220><223> UvsY Sup1 HIS2

<400> 64

Met Ala Gly His His His His His Pro His Ala His His Pro Leu Ser

1 5 10 15
Gln Ser Ser Gly His His His His His His His His His His His Gln Gly
20 25 30
Tyr Gly Gly Ser Gly Lys Leu Glu Asp Leu Gln Glu Glu Leu Asp Ala
35 40 45
Asp Leu Ala Ile Asp Met Ser Lys Leu Gln Tyr Glu Thr Ala Asn Asn
50 55 60
Val Lys Leu Tyr Ser Lys Trp Leu Arg Lys His Ser Phe Ile Arg Lys

65 70 75 80
Glu Met Leu Arg Ile Glu Thr Gln Lys Lys Thr Ala Leu Lys Ala Arg
85 90 95
Leu Asp Tyr Tyr Ser Gly Arg Gly Asp Gly Asp Glu Phe Ser Met Asp
100 105 110
Arg Tyr Glu Lys Ser Glu Met Lys Thr Val Leu Ala Ala Asp Lys Asp
115 120 125
Val Leu Lys Ile Glu Thr Thr Leu Gln Tyr Trp Gly Ile Leu Leu Glu

130 135 140
Phe Cys Ser Gly Ala Leu Asp Ala Val Lys Ser Arg Ser Phe Ala Leu
145 150 155 160
Lys His Ile Gln Asp Met Arg Glu Phe Glu Ala Gly Gln Ser Gly Ser
165 170 175
Gly Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Asn Asn Leu Gln
180 185 190

<210> 65

<211> 306

<212> PRT

<213> Artificial Sequence

<220><223> Gp32 (7His) from phage vB

<400> 65

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45
 Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His

 50 55 60
 Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
 65 70 75 80
 His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
 85 90 95
 Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
 100 105 110
 Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

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

 180 185 190
 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp

 245 250 255

Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
260 265 270

Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
275 280 285

Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu His His His His His
290 295 300

His His
305

<210> 66

<211> 336

<

212> PRT

<213> Artificial Sequence

<220><223> Gp32Super +ve

<400> 66

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
 145 150 155 160
 Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
 165 170 175
 Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
 180 185 190
 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Lys
 245 250 255
 Lys Val Ala Ser Lys Leu Lys Phe Lys Lys Lys Met Glu Ala Phe Ser
 260 265 270
 Ser Ala Lys Thr Lys Lys Lys Phe Met Ser Ser Ser Ser Ser Lys Lys
 275 280 285
 Ser Lys Leu Lys Lys Leu Leu Ala Gly Leu Met Glu Ala Phe Ser Ser
 290 295 300
 Ala Lys Thr Lys Lys Lys Phe Met Ser Ser Ser Ser Ser Lys Lys Ser
 305 310 315 320
 Lys Leu Lys Lys Leu Leu Ala Gly Leu His His His His His His His
 325 330 335
 <210> 67
 <211> 337
 <212> PRT
 <213> Artificial Sequence
 <220><223> Gp32Super -ve
 <400> 67
 Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
 1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
 50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
 65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
 85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
 100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
 115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
 130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
 145 150 155 160

Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
 165 170 175

Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
 180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205

Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220

Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240

Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255

Asp Val Ala Ser Asp Leu Asp Asp Phe Asp Asp Asp Met Glu Ala Phe

260 265 270
 Ser Ser Ala Asp Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285

Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Met Glu Ala Phe Ser
 290 295 300
 Ser Ala Asp Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp Asp
 305 310 315 320
 Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu His His His His His His
 325 330 335
 His

<210> 68

<211> 403

<212> PRT

<213> Artificial Sequence

<220><223> Gp32-TripleXtail

<400

> 68

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
 1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45
 Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
 50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
 65 70 75 80
 His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
 85 90 95
 Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
 100 105 110
 Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

115 120 125
 Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
 130 135 140
 Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
 145 150 155 160
 Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
 165 170 175
 Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
 180 185 190

 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255

 Asp Val Ala Ser Glu Phe Glu Ile Val Glu Ala Glu Val Asp Glu Leu
 260 265 270
 Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro Glu Ser Lys Ser Lys
 275 280 285
 Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln Leu Ser Asp Met Glu
 290 295 300
 Glu Phe Asn Glu Ile Val Glu Ala Glu Val Asp Glu Leu Ile Asn Ser
 305 310 315 320

 Lys Val Glu Lys Phe Lys Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala
 325 330 335
 Asp Leu Glu Thr Asp Leu Glu Gln Leu Ser Asp Met Glu Glu Phe Asn
 340 345 350
 Glu Ile Val Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu
 355 360 365

Lys Phe Lys Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu
370 375 380

Thr Asp Leu Glu Gln Leu Ser Asp Met Glu Glu Phe Asn His His His
385 390 395 400
His His His

<210> 69

<211> 349

<212> PRT

<213> Artificial Sequence

<220><223> Gp32-fib

<400> 69

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
1 5 10 15
Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
20 25 30

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

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
100 105 110
Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
115 120 125
Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
130 135 140
Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr

145 150 155 160

 Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
 165 170 175
 Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
 180 185 190
 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220

 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255
 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285

 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Pro Gly Phe Ser Pro
 290 295 300
 Arg Gly Gly Gly Phe Gly Gly Arg Gly Gly Phe Gly Asp Arg Gly Gly
 305 310 315 320
 Arg Gly Gly Arg Gly Gly Phe Gly Gly Gly Arg Gly Arg Gly Gly Gly
 325 330 335
 Phe Arg Gly Arg Gly Arg His His His His His His His
 340 345

 <210> 70

 <211> 375
 <212> PRT
 <213> Artificial Sequence
 <220><223> Gp32-PCF11
 <400> 70

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
1 5 10 15
Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
20 25 30
Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
35 40 45
Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
50 55 60
Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
65 70 75 80
His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
85 90 95
Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
100 105 110
Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
115 120 125
Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
130 135 140
Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
145 150 155 160
Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
165 170 175
Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
180 185 190
Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
195 200 205
Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
210 215 220
Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
225 230 235 240
Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp

245 250 255
 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Asn Glu Gln Val Gln
 290 295 300
 Met Gln Leu Arg Gln Val Phe Ser Gln Asp Gln Gln Val Leu Gln Glu

305 310 315 320
 Arg Met Arg Tyr His Glu Leu Gln Gln Gln Gln Gln Gln Tyr His
 325 330 335
 Glu Thr Lys Asp Met Val Gly Ser Tyr Thr Gln Asn Ser Asn Ser Ala
 340 345 350
 Ile Pro Leu Phe Gly Asn Asn Ser Asp Thr Thr Asn Gln Gln Asn Ser
 355 360 365
 His His His His His His His
 370 375

<210> 71

<211> 421

<212> PRT

<213> Artificial Sequence

<220><223> Gp32-Sup

<400> 71

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
 1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45
 Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
 50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80
 His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
 85 90 95
 Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
 100 105 110
 Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

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

 180 185 190
 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp

 245 250 255
 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Met Ser Asp Ser Asn
 290 295 300
 Gln Gly Asn Asn Gln Gln Asn Tyr Gln Gln Tyr Ser Gln Asn Gly Asn

 305 310 315 320

Gln Gln Gln Gly Asn Asn Arg Tyr Gln Gly Tyr Gln Ala Tyr Asn Ala
 325 330 335
 Gln Ala Gln Pro Ala Gly Gly Tyr Tyr Gln Asn Tyr Gln Gly Tyr Ser
 340 345 350
 Gly Tyr Gln Gln Gly Gly Tyr Gln Gln Tyr Gln Tyr Asn Pro Gln Gly
 355 360 365
 Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro

370 375 380
 Gln Gly Gly Tyr Gln Gln Gln Phe Asn Pro Gln Gly Gly Arg Gly Asn
 385 390 395 400
 Tyr Lys Asn Phe Asn Tyr Asn Asn Asn Leu Gln Gly Tyr Gln His His
 405 410 415
 His His His His His
 420

<210> 72

<211> 315

<212> PRT

<213> Artificial Sequence

<220><223> Gp32-Sup1

<400> 72

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45
 Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
 50 55 60
 Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80
 His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
 85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
 100 105 110
 Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
 115 120 125
 Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
 130 135 140
 Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
 145 150 155 160
 Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
 165 170 175
 Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
 180 185 190
 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255
 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Asp
 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Tyr Asn Pro Gln Gly
 290 295 300
 Gly Tyr Gln Gln His His His His His His His
 305 310 315
 <210> 73
 <211> 324
 <212> PRT
 <213> Artificial Sequence

<220><223> Gp32-Sup2

<400> 73

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys

20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His

50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys

100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp

130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr

145 150 155 160

Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys

165 170 175

Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu

180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu

195 200 205

Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys

210 215 220

Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly

225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255

 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Tyr Asn Pro Gln Gly
 290 295 300
 Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln His His His
 305 310 315 320

His His His His

<210> 74

<211> 333

<212> PRT

<213> Artificial Sequence

<220><223> Gp32-Sup3

<400> 74

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
 1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45

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

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

 Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
 180 185 190
 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240

 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255
 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Tyr Asn Pro Gln Gly
 290 295 300

 Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro
 305 310 315 320
 Gln Gly Gly Tyr Gln Gln His His His His His His His
 325 330

 <210> 75
 <211> 342

<212> PRT

<213> Artificial Sequence

<220><223> Gp32-Sup4

<400> 75

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys

20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His

50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys

100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp

130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr

145 150 155 160

Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys

165 170 175

Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu

180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu

195 200 205

Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys

210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255
 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp

 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Tyr Asn Pro Gln Gly
 290 295 300
 Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln Tyr Asn Pro
 305 310 315 320
 Gln Gly Gly Tyr Gln Gln Tyr Asn Pro Gln Gly Gly Tyr Gln Gln His
 325 330 335
 His His His His His His
 340

<210> 76

<211> 557

<212> PRT

<213> Artificial Sequence

<220><223> Gp32-DDX

<400> 76

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
 1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45
 Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His

 50 55 60
 Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65						70						75						80						
His	Gly	Asp	Tyr	Asp	Ser	Cys	Pro	Val	Cys	Gln	Tyr	Ile	Ser	Lys	Asn									
					85					90					95									
Asp	Leu	Tyr	Asn	Thr	Asn	Lys	Thr	Glu	Tyr	Ser	Gln	Leu	Lys	Arg	Lys									
					100					105					110									
Thr	Ser	Tyr	Trp	Ala	Asn	Ile	Leu	Val	Val	Lys	Asp	Pro	Gln	Ala	Pro									
					115					120					125									
Asp	Asn	Glu	Gly	Lys	Val	Phe	Lys	Tyr	Arg	Phe	Gly	Lys	Lys	Ile	Trp									
130					135					140														
Asp	Lys	Ile	Asn	Ala	Met	Ile	Ala	Val	Asp	Thr	Glu	Met	Gly	Glu	Thr									
145					150					155					160									
Pro	Val	Asp	Val	Thr	Cys	Pro	Trp	Glu	Gly	Ala	Asn	Phe	Val	Leu	Lys									
					165					170					175									
Val	Lys	Gln	Val	Ser	Gly	Phe	Ser	Asn	Tyr	Asp	Glu	Ser	Lys	Phe	Leu									
					180					185					190									
Asn	Gln	Ser	Ala	Ile	Pro	Asn	Ile	Asp	Asp	Glu	Ser	Phe	Gln	Lys	Glu									
195					200					205														
Leu	Phe	Glu	Gln	Met	Val	Asp	Leu	Ser	Glu	Met	Thr	Ser	Lys	Asp	Lys									
210					215					220														
Phe	Lys	Ser	Phe	Glu	Glu	Leu	Asn	Thr	Lys	Phe	Asn	Gln	Val	Leu	Gly									
225					230					235					240									
Thr	Ala	Ala	Leu	Gly	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Val	Ala	Asp									
					245					250					255									
Lys	Val	Ala	Ser	Asp	Leu	Asp	Asp	Phe	Asp	Lys	Asp	Met	Glu	Ala	Phe									
260					265					270														
Ser	Ser	Ala	Lys	Thr	Glu	Asp	Asp	Phe	Met	Ser	Ser	Ser	Ser	Ser	Asp									
275					280					285														
Asp	Ser	Asp	Leu	Asp	Asp	Leu	Leu	Ala	Gly	Leu	His	His	His	His	His									
290					295					300														
His	His	Phe	Glu	Phe	Met	Gly	Asp	Glu	Asp	Trp	Glu	Ala	Glu	Ile	Asn									
					305					310					315					320				

Pro His Met Ser Ser Tyr Val Pro Ile Phe Glu Lys Asp Arg Tyr Ser
325 330 335

Gly Glu Asn Gly Asp Asn Phe Asn Arg Thr Pro Ala Ser Ser Ser Glu
340 345 350

Met Asp Asp Gly Pro Ser Arg Arg Asp His Phe Met Lys Ser Gly Phe
355 360 365

Ala Ser Gly Arg Asn Phe Gly Asn Arg Asp Ala Gly Glu Cys Asn Lys
370 375 380

Arg Asp Asn Thr Ser Thr Met Gly Gly Phe Gly Val Gly Lys Ser Phe
385 390 395 400

Gly Asn Arg Gly Phe Ser Asn Ser Arg Phe Glu Asp Gly Asp Ser Ser
405 410 415

Gly Phe Trp Arg Glu Ser Ser Asn Asp Cys Glu Asp Asn Pro Thr Arg
420 425 430

Asn Arg Gly Phe Ser Lys Arg Gly Gly Tyr Arg Asp Gly Asn Asn Ser
435 440 445

Glu Ala Ser Gly Pro Tyr Arg Arg Gly Gly Arg Gly Ser Phe Arg Gly
450 455 460

Cys Arg Gly Gly Phe Gly Leu Gly Ser Pro Asn Asn Asp Leu Asp Pro
465 470 475 480

Asp Glu Cys Met Gln Arg Thr Gly Gly Leu Phe Gly Ser Arg Arg Pro
485 490 495

Val Leu Ser Gly Thr Gly Asn Gly Asp Thr Ser Gln Ser Arg Ser Gly
500 505 510

Ser Gly Ser Glu Arg Gly Gly Tyr Lys Gly Leu Asn Glu Glu Val Ile
515 520 525

Thr Gly Ser Gly Lys Asn Ser Trp Lys Ser Glu Ala Glu Gly Gly Glu
530 535 540

Ser Ser Asp Thr Gln Leu Glu His His His His His His
545 550 555

<210> 77
<211> 521

<212> PRT

<213> Artificial Sequence

<220

><223> Gp32 PolCTD

<400> 77

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys

20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His

50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys

100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp

130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr

145 150 155 160

Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys

165 170 175

Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu

180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu

195 200 205

Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys

210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255

 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu His His His His His
 290 295 300
 His His Phe Glu Phe Ser Pro Phe Gly Ala Tyr Gly Glu Ala Pro Thr
 305 310 315 320

 Ser Pro Gly Phe Gly Val Ser Ser Pro Gly Phe Ser Pro Thr Ser Pro
 325 330 335
 Thr Tyr Ser Pro Thr Ser Pro Ala Tyr Ser Pro Thr Ser Pro Ser Tyr
 340 345 350
 Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro
 355 360 365
 Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser
 370 375 380

 Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser
 385 390 395 400
 Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser
 405 410 415
 Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr
 420 425 430
 Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro
 435 440 445

 Ala Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr
 450 455 460

Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro
 465 470 475 480
 Thr Ser Pro Asn Tyr Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser
 485 490 495
 Pro Gly Tyr Ser Pro Gly Ser Pro Ala Tyr Ser Pro Lys Gln Asp Glu
 500 505 510

Gln Leu Glu His His His His His His
 515 520

<210> 78

<211> 440

<212> PRT

<213> Artificial Sequence

<220><223> Gp32 HNRNPA1

<400> 78

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
 1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

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

100 105 110
 Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
 115 120 125
 Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
 130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
 145 150 155 160
 Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
 165 170 175
 Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
 180 185 190
 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255
 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270
 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Ala Ser Ala Ser Ser
 290 295 300
 Ser Gln Arg Gly Arg Ser Gly Ser Gly Asn Phe Gly Gly Gly Arg Gly
 305 310 315 320
 Gly Gly Phe Gly Gly Asn Asp Asn Phe Gly Arg Gly Gly Asn Phe Ser
 325 330 335
 Gly Arg Gly Gly Phe Gly Gly Ser Arg Gly Gly Gly Gly Tyr Gly Gly
 340 345 350
 Ser Gly Asp Gly Tyr Asn Gly Phe Gly Asn Asp Gly Ser Asn Phe Gly
 355 360 365
 Gly Gly Gly Ser Tyr Asn Asp Phe Gly Asn Tyr Asn Asn Gln Ser Ser
 370 375 380
 Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly Gly Arg Ser Ser Gly

385 390 395 400
 Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro Gln Asn Gln Gly
 405 410 415
 Gly Tyr Gly Val Ser Ser Ser Ser Ser Ser Tyr Gly Ser Gly Arg Arg

 420 425 430
 Phe His His His His His His His
 435 440
 <210> 79
 <211> 364
 <212> PRT
 <213> Artificial Sequence
 <220><223> Gp32 HRP1
 <400> 79
 Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
 1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

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

 100 105 110
 Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
 115 120 125
 Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
 130 135 140
 Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr

145 150 155 160
Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys

 165 170 175
Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu

 180 185 190
Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu

 195 200 205
Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys

 210 215 220
Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly

225 230 235 240
Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp

 245 250 255
Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe

 260 265 270
Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp

 275 280 285
Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Gly Gly Asn Asn Gly

 290 295 300
Gly Asn Asn Met Asn Arg Arg Gly Gly Asn Phe Gly Asn Gln Gly Asp

305 310 315 320
Phe Asn Gln Met Tyr Gln Asn Pro Met Met Gly Gly Tyr Asn Pro Met

 325 330 335
Met Asn Pro Gln Ala Met Thr Asp Tyr Tyr Gln Lys Met Gln Glu Tyr

 340 345 350
Tyr Gln Gln Met Gln His His His His His His His

 355 360

<210> 80

<211> 422

<212> PRT

<213> Artificial Sequence

<220><223> Gp32 HRP2

<400> 80

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys

20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His

50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys

100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp

130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr

145 150 155 160

Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys

165 170 175

Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu

180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu

195 200 205

Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys

210 215 220

Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly

225 230 235 240

Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp

245 250 255

Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe

260 265 270

Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp

275 280 285

Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Gly Gly Asn Asn Gly

290 295 300

Gly Asn Asn Met Asn Arg Arg Gly Gly Asn Phe Gly Asn Gln Gly Asp

305 310 315 320

Phe Asn Gln Met Tyr Gln Asn Pro Met Met Gly Gly Tyr Asn Pro Met

325 330 335

Met Asn Pro Gln Ala Met Thr Asp Tyr Tyr Gln Lys Met Gln Glu Tyr

340 345 350

Tyr Gln Gln Met Gln Gly Gly Asn Asn Gly Gly Asn Asn Met Asn Arg

355 360 365

Arg Gly Gly Asn Phe Gly Asn Gln Gly Asp Phe Asn Gln Met Tyr Gln

370 375 380

Asn Pro Met Met Gly Gly Tyr Asn Pro Met Met Asn Pro Gln Ala Met

385 390 395 400

Thr Asp Tyr Tyr Gln Lys Met Gln Glu Tyr Tyr Gln Gln Met Gln His

405 410 415

His His His His His His

420

<210> 81

<211> 330

<212> PRT

<213> Artificial Sequence

<220><223> Gp32 HIS1

<400> 81

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
1 5 10 15
Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
20 25 30
Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
35 40 45
Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
50 55 60
Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
65 70 75 80
His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
85 90 95
Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
100 105 110
Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
115 120 125
Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
130 135 140
Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
145 150 155 160
Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
165 170 175
Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
180 185 190
Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
195 200 205
Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
210 215 220
Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
225 230 235 240
Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp

245 250 255

Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe

260 265 270

Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp

275 280 285

Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Pro Gln Gln Asn Ala

290 295 300

Leu His His His His Gly Asn Ser Ser His His His His His His His

305 310 315 320

His His His His His His Gly Gln Gln Ala
325 330

<210> 82

<211> 335

<212> PRT

<213> Artificial Sequence

<220><223> GP32 HIS2

<400> 82

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys

20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

35 40 45
Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys

100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
115 120 125
Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
130 135 140
Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
145 150 155 160
Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
165 170 175
Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
180 185 190
Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
195 200 205
Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
210 215 220
Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
225 230 235 240
Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ser Val Ala Asp
245 250 255
Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
260 265 270
Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
275 280 285
Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Ala Gly His His His
290 295 300
His His Pro His Ala His His Pro Leu Ser Gln Ser Ser Gly His His
305 310 315 320
His His His His His His His His Gln Gly Tyr Gly Gly Ser Gly
325 330 335

<210> 83

<211> 342

<212> PRT

<213> Artificial Sequence

<220><223> GP32 HIS3

<400> 83

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
145 150 155 160

Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
165 170 175

Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
195 200 205

Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
210 215 220

Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly

225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255
 Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
 260 265 270

 Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
 275 280 285
 Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Ala Gly His His His
 290 295 300
 His His His His His Leu Pro His Leu Pro Pro Pro His Leu His His
 305 310 315 320
 His His His Pro Gln His His Leu His Pro Gly Ser Ala Ala Ala Val
 325 330 335

 His Pro Val Gln Gln His
 340

 <210> 84
 <211> 341
 <212> PRT
 <213> Artificial Sequence
 <220><223> GP32 HIS4
 <400> 84

 Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
 1 5 10 15
 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45

 Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
 50 55 60
 Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
 65 70 75 80
 His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

	85	90	95
Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys			
	100	105	110
Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro			
	115	120	125
Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp			
	130	135	140
Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr			
145	150	155	160
Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys			
	165	170	175
Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu			
	180	185	190
Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu			
	195	200	205
Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys			
	210	215	220
Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly			
225	230	235	240
Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ser Val Ala Asp			
	245	250	255
Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe			
	260	265	270
Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp			
	275	280	285
Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Asn His His Ala Ser			
	290	295	300
His Gly His His Asn Ser His His Pro Gln His His His His His His			
305	310	315	320
His His His His His Pro Pro Pro Pro Ala Pro Gln Pro Pro Pro Pro			
	325	330	335

Pro Gln Gln Gln Gln

340

<210> 85

<211> 375

<212> PRT

<213> Artificial Sequence

<220><223> GP32 HIS5

<400> 85

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys

20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His

50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys

100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp

130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr

145 150 155 160

Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys

165 170 175

Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu

180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
195 200 205

Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
210 215 220

Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
225 230 235 240

Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ser Val Ala Asp
245 250 255

Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
260 265 270

Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
275 280 285

Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Ser Gly His His Gly
290 295 300

Ala His His Gly Ala His His Pro Ala Ala Ala Ala Tyr Glu Ala
305 310 315 320

Phe Arg Gly Pro Gly Phe Ala Gly Gly Gly Gly Ala Asp Asp Met Gly
325 330 335

Ala Gly His His His Gly Ala His His Ala Ala His His His His Ala
340 345 350

Ala His His His His His His His His His His Gly Gly Ala Gly His
355 360 365

Gly Gly Gly Ala Gly His His
370 375

<210> 86

<211> 355

<212> PRT

<213> Artificial Sequence

<220><223> Gp32 Mimic1

<400> 86

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15

 Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
 20 25 30
 Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
 35 40 45
 Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
 50 55 60
 Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
 65 70 75 80

 His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn
 85 90 95
 Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys
 100 105 110
 Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro
 115 120 125
 Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp
 130 135 140

 Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr
 145 150 155 160
 Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys
 165 170 175
 Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu
 180 185 190
 Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu
 195 200 205

 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly
 225 230 235 240
 Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp
 245 250 255

Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe
260 265 270

Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp
275 280 285

Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Tyr Asp Pro Thr Ser
290 295 300

Pro Ser Tyr Asp Pro Thr Ser Pro Ser Tyr Asp Pro Thr Ser Pro Ser
305 310 315 320

Tyr Asp Pro Thr Ser Pro Ser Tyr Asp Pro Thr Ser Pro Ser Tyr Asp
325 330 335

Pro Thr Ser Pro Ser Tyr Asp Pro Thr Ser Pro Ser His His His His
340 345 350

His His His
355

<210> 87

<211> 355

<212> PRT

<213> Artificial Sequence

<220><223> Gp32 Mimic2

<400> 87

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys
1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys
20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu
35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His
50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr
65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

85	90	95
Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys		
100	105	110
Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro		
115	120	125
Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp		
130	135	140
Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr		
145	150	155
160		
Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys		
165	170	175
Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu		
180	185	190
Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu		
195	200	205
Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys		
210	215	220
Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly		
225	230	235
Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ser Val Ala Asp		
245	250	255
Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe		
260	265	270
Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp		
275	280	285
Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Tyr Ser Pro Thr Asp		
290	295	300
Pro Ser Tyr Ser Pro Thr Asp Pro Ser Tyr Ser Pro Thr Asp Pro Ser		
305	310	315
Tyr Ser Pro Thr Asp Pro Ser Tyr Ser Pro Thr Asp Pro Ser Tyr Ser		
325	330	335

Pro Thr Asp Pro Ser Tyr Ser Pro Thr Asp Pro Ser His His His His

340

345

350

His His His

355

<210> 88

<211> 348

<212> PRT

<213> Artificial Sequence

<220><223> Gp32 short pol ctd

<400> 88

Met Phe Lys Arg Lys Ser Thr Ala Asp Leu Ala Ala Gln Met Ala Lys

1 5 10 15

Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys

20 25 30

Leu Lys Leu Asp Ala Ser Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

35 40 45

Pro Ala Lys Thr Asp Asp Ala Leu Pro Phe Thr Ile Leu Val Asn His

50 55 60

Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80

His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

85 90 95

Asp Leu Tyr Asn Thr Asn Lys Thr Glu Tyr Ser Gln Leu Lys Arg Lys

100 105 110

Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Gln Ala Pro

115 120 125

Asp Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp

130 135 140

Asp Lys Ile Asn Ala Met Ile Ala Val Asp Thr Glu Met Gly Glu Thr

145 150 155 160

Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys

165 170 175

Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu

180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu

195 200 205

Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys

210 215 220

Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Asn Gln Val Leu Gly

225 230 235 240

Thr Ala Ala Leu Gly Gly Ala Ala Ala Ala Ala Ala Ser Val Ala Asp

245 250 255

Lys Val Ala Ser Asp Leu Asp Asp Phe Asp Lys Asp Met Glu Ala Phe

260 265 270

Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Ser Asp

275 280 285

Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu Thr Ser Pro Ser Tyr

290 295 300

Ser Pro Thr Ser Pro Ser Tyr Ser Pro Thr Ser Pro Gly Tyr Ser Pro

305 310 315 320

Thr Ser Pro Ala Tyr Ser Pro Thr Ser Pro Thr Tyr Ser Pro Thr Ser

325 330 335

Pro Thr Tyr Ser Pro His His His His His His His

340 345

<210> 89

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> 6-histidine

<400> 89

His His His His His His

1 5

<210> 90

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> C-myc epitope

<400> 90

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

1 5 10

<210> 91

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> FLAG octapeptide

<400> 91

Asp Tyr Lys Asp Asp Asp Asp Lys

1 5

<210> 92

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> Protein C

<400> 92

Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys

1 5 10

<210> 93

<211> 12

<212> PRT

<213> Artificial Sequence

<220><223> Tag-100

<400> 93

Glu Glu Thr Ala Arg Phe Gln Pro Gly Tyr Arg Ser

1 5 10

<210> 94

<211> 14

<212> PRT

<213> Artificial Sequence

<220><223> V5 epitope

<400> 94

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

1 5 10

<210> 95

<211> 11

<212> PRT

<213>

> Artificial Sequence

<220><223> VSV-G

<400> 95

Tyr Thr Asp Ile Glu Met Asn Arg Leu Gly Lys

1 5 10

<210> 96

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Xpress

<400> 96

Asp Leu Tyr Asp Asp Asp Asp Lys

1 5

<210> 97

<211> 9

<212> PRT

<213> Artificial Sequence

<220><223> Hemagglutinin

<400> 97

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala

1 5

<210> 98

<211>

> 36

<212> DNA

<213> Artificial Sequence

<220><223> Example 1 - Forward primer

<400> 98

cgccctgcaag tcctaagacg ccaatcgaaa agaaac 36

<210> 99

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> Example 1 - Reverse primer

<400> 99

ctgcatctcc gtgtataact aatacattgt tttta 35

<210> 100

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Example 1 - Probe

<220><221> MISC_FEATURE

<222> (32)..(32)

<223> FAM - fluorescein

<220><221> MISC_FEATURE

<222> (33)..(33)

<223> THF - tetrahydrofuran

<220><221> MISC_FEATURE

<222> (34)..(34)

<223> BHQ - Black Hole Quencher

<400> 100

cgaaaagaaa cagcggatg aaatcgataa gnnnatacaa ggattgga 48

<210> 101

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Example 5 - Forward primer

<400> 101

gcagctgtat agcaaattcc tgttgaaagc ag 32

<210> 102

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Example 5 - Reverse primer

<400> 102

tcctggctgt attcattgtt gttaaattgg

30

<210> 103

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Example 5 - probe

<220><221> MISC_FEATURE

<222> (30)..(30)

<223> FAM - fluorescein

<220><221> MISC_FEATURE

<222> (32)..(32)

<223> THF - tetrahydrofuran

<220><221> MISC_FEATURE

<222> (34)..(34)

<223> BHQ - Black Hole Quencher

<400> 103

cactgatgct tttcctagac acgagatgat gncctgtgga gcctttgt

48

<210> 104

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Example 12 - Probe

<220><221> MISC_FEATURE

<222> (1)..(37)

<223> FAM - Fluorescein

<400> 104

ccgcaatggg gcactctcag tacaatctgc tctgatg

37

<210> 105
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> Example 18 - PA30 forward primer
 <400> 105
 ccatctcatc cctgcgtgtc tccgactcag 30
 <210> 106
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> Example 18 - PB30 reverse primer
 <400> 106
 cctatcccct gtgtgccttg gcagtctcag 30
 <210> 107
 <211> 46
 <212> DNA
 <213> Artificial Sequence
 <220><223> Example 18 - Probe
 <220><221> MISC_FEATURE
 <222> (29)..(29)
 <223> FAM - fluorescein
 <220><221> MISC_FEATURE
 <222> (31)..(31)
 <223> THF - tetrahydrofuran
 <220><221> MISC_FEATURE
 <222> (33)..(33)
 <223> Quencher - Black Hole Quencher
 <400> 107
 agcagaagca ataccgccag caatagcatg ngtagagcga gctgcc 46
 <210> 108
 <211> 188
 <212> DNA

<213> Artificial Sequence

<220><223> Example 18 - TFIL template sequence

<400> 108

ccatctcatc cctgcgtgtc tccgactcag tgttttaggg tccccggggt taaaaggttt 60

cgaactcaac agctgtctgg cagctcgctc tacgcatgct attgctggcg gtattgcttc 120

tgctcttgct ggtggcgcca tgtctaaatt gtttggagct gagactgcca aggcacacag 180

gggatagg 188

<210> 109

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223>

> Example 18 - PB30' probe

<220><221> MISC_FEATURE

<222> (1)..(30)

<223> 5' - ROX (carboxyrhodamine)

<400> 109

ctgagactgc caaggcacac aggggatagg 30

<210> 110

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Example 18 - TFIL probe

<220><221> MISC_FEATURE

<222> (1)..(20)

<223> 5' FAM (fluorescein)

<400> 110

ggtttcgaac tcaacagctg 20

<210> 111

<211> 504

<212> PRT

<213> Artificial Sequence

<220><223> Example 20 - RB69 ligase

<400> 111

Met Ile Leu Asp Ile Leu Asn Gln Ile Ala Ala Ile Gly Ser Thr Lys
1 5 10 15
Thr Lys Gln Glu Ile Leu Lys Lys Asn Lys Asp Asn Lys Leu Leu Glu
20 25 30
Arg Val Tyr Arg Leu Thr Tyr Ala Arg Gly Ile Gln Tyr Tyr Ile Lys
35 40 45
Lys Trp Pro Gly Pro Gly Glu Arg Ser Gln Ala Tyr Gly Leu Leu Glu
50 55 60
Leu Asp Asp Met Leu Asp Phe Ile Glu Phe Thr Leu Ala Thr Arg Lys
65 70 75 80
Leu Thr Gly Asn Ala Ala Ile Lys Glu Leu Met Gly Tyr Ile Ala Asp
85 90 95
Gly Lys Pro Asp Asp Val Glu Val Leu Arg Arg Val Met Met Arg Asp
100 105 110
Leu Glu Val Gly Ala Ser Val Ser Ile Ala Asn Lys Val Trp Pro Gly
115 120 125
Leu Ile Gln Leu Gln Pro Gln Met Leu Ala Ser Ala Tyr Asp Glu Lys
130 135 140
Leu Ile Thr Lys Asn Ile Lys Trp Pro Ala Phe Ala Gln Leu Lys Ala
145 150 155 160
Asp Gly Ala Arg Cys Phe Ala Glu Val Arg Asp Asp Gly Val Gln Phe
165 170 175
Phe Ser Arg Ala Gly Asn Glu Tyr His Gly Leu Thr Leu Leu Ala Asp
180 185 190
Glu Leu Met Glu Met Thr Lys Glu Ala Arg Glu Arg His Pro Asn Gly
195 200 205
Val Leu Ile Asp Gly Glu Leu Val Tyr His Ser Phe Asp Ile Lys Lys
210 215 220
Ala Val Ser Ser Gly Asn Asp Leu Ser Phe Leu Phe Gly Asp Asn Glu
225 230 235 240
Glu Ser Glu Glu Val Gln Val Ala Asp Arg Ser Thr Ser Asn Gly Leu

245 250 255
 Ala Asn Lys Ser Leu Gln Gly Thr Ile Ser Pro Lys Glu Ala Glu Gly
 260 265 270
 Met Val Leu Gln Ala Trp Asp Tyr Val Pro Leu Asp Glu Val Tyr Ser
 275 280 285
 Asp Gly Lys Ile Lys Gly Gln Lys Tyr Asp Val Arg Phe Ala Ala Leu
 290 295 300
 Glu Asn Met Ala Glu Gly Phe Lys Arg Ile Glu Pro Ile Glu Asn Gln

 305 310 315 320
 Leu Val Arg Asn Leu Asp Glu Ala Lys Val Val Tyr Lys Lys Tyr Val
 325 330 335
 Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Arg Asp Ser Tyr Trp
 340 345 350
 Glu Asn Lys Arg Ser Lys Asn Leu Ile Lys Phe Lys Glu Val Ile Asp
 355 360 365
 Ile Ala Leu Glu Val Val Gly Tyr Tyr Glu His Ser Lys Asp Pro Asn

 370 375 380
 Lys Leu Gly Gly Val Glu Leu Val Ser Arg Cys Arg Arg Ile Thr Thr
 385 390 395 400
 Asp Cys Gly Ser Gly Phe Lys Asp Thr Thr His Lys Thr Val Asp Gly
 405 410 415
 Val Lys Val Leu Ile Pro Leu Asp Glu Arg His Asp Leu Asp Arg Glu
 420 425 430
 Arg Leu Met Ser Glu Ala Arg Glu Gly Lys Leu Ile Gly Arg Ile Ala

 435 440 445
 Asp Cys Glu Cys Asn Gly Trp Val His Ser Lys Gly Arg Glu Gly Thr
 450 455 460
 Val Gly Ile Phe Leu Pro Ile Ile Lys Gly Phe Arg Phe Asp Lys Thr
 465 470 475 480
 Glu Ala Asp Ser Phe Glu Asp Val Phe Gly Ser Trp Asp Gln Thr Gly
 485 490 495
 Val Ser Gly His His His His His

500

<210> 112

<211> 533

<212> PRT

<213> Artificial Sequence

<220><223> Example 20 - RB69 ligase-His2

<400> 112

Met Ile Leu Asp Ile Leu Asn Gln Ile Ala Ala Ile Gly Ser Thr Lys

1 5 10 15

Thr Lys Gln Glu Ile Leu Lys Lys Asn Lys Asp Asn Lys Leu Leu Glu

20 25 30

Arg Val Tyr Arg Leu Thr Tyr Ala Arg Gly Ile Gln Tyr Tyr Ile Lys

35 40 45

Lys Trp Pro Gly Pro Gly Glu Arg Ser Gln Ala Tyr Gly Leu Leu Glu

50 55 60

Leu Asp Asp Met Leu Asp Phe Ile Glu Phe Thr Leu Ala Thr Arg Lys

65 70 75 80

Leu Thr Gly Asn Ala Ala Ile Lys Glu Leu Met Gly Tyr Ile Ala Asp

85 90 95

Gly Lys Pro Asp Asp Val Glu Val Leu Arg Arg Val Met Met Arg Asp

100 105 110

Leu Glu Val Gly Ala Ser Val Ser Ile Ala Asn Lys Val Trp Pro Gly

115 120 125

Leu Ile Gln Leu Gln Pro Gln Met Leu Ala Ser Ala Tyr Asp Glu Lys

130 135 140

Leu Ile Thr Lys Asn Ile Lys Trp Pro Ala Phe Ala Gln Leu Lys Ala

145 150 155 160

Asp Gly Ala Arg Cys Phe Ala Glu Val Arg Asp Asp Gly Val Gln Phe

165 170 175

Phe Ser Arg Ala Gly Asn Glu Tyr His Gly Leu Thr Leu Leu Ala Asp

180 185 190

Glu Leu Met Glu Met Thr Lys Glu Ala Arg Glu Arg His Pro Asn Gly
 195 200 205
 Val Leu Ile Asp Gly Glu Leu Val Tyr His Ser Phe Asp Ile Lys Lys
 210 215 220
 Ala Val Ser Ser Gly Asn Asp Leu Ser Phe Leu Phe Gly Asp Asn Glu
 225 230 235 240

 Glu Ser Glu Glu Val Gln Val Ala Asp Arg Ser Thr Ser Asn Gly Leu
 245 250 255
 Ala Asn Lys Ser Leu Gln Gly Thr Ile Ser Pro Lys Glu Ala Glu Gly
 260 265 270
 Met Val Leu Gln Ala Trp Asp Tyr Val Pro Leu Asp Glu Val Tyr Ser
 275 280 285
 Asp Gly Lys Ile Lys Gly Gln Lys Tyr Asp Val Arg Phe Ala Ala Leu
 290 295 300

 Glu Asn Met Ala Glu Gly Phe Lys Arg Ile Glu Pro Ile Glu Asn Gln
 305 310 315 320
 Leu Val Arg Asn Leu Asp Glu Ala Lys Val Val Tyr Lys Lys Tyr Val
 325 330 335
 Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Arg Asp Ser Tyr Trp
 340 345 350
 Glu Asn Lys Arg Ser Lys Asn Leu Ile Lys Phe Lys Glu Val Ile Asp
 355 360 365

 Ile Ala Leu Glu Val Val Gly Tyr Tyr Glu His Ser Lys Asp Pro Asn
 370 375 380
 Lys Leu Gly Gly Val Glu Leu Val Ser Arg Cys Arg Arg Ile Thr Thr
 385 390 395 400
 Asp Cys Gly Ser Gly Phe Lys Asp Thr Thr His Lys Thr Val Asp Gly
 405 410 415
 Val Lys Val Leu Ile Pro Leu Asp Glu Arg His Asp Leu Asp Arg Glu
 420 425 430

 Arg Leu Met Ser Glu Ala Arg Glu Gly Lys Leu Ile Gly Arg Ile Ala

435 440 445
 Asp Cys Glu Cys Asn Gly Trp Val His Ser Lys Gly Arg Glu Gly Thr
 450 455 460
 Val Gly Ile Phe Leu Pro Ile Ile Lys Gly Phe Arg Phe Asp Lys Thr
 465 470 475 480
 Glu Ala Asp Ser Phe Glu Asp Val Phe Gly Ser Trp Asp Gln Thr Gly
 485 490 495

Val Ser Gly His His His His His Pro His Ala His His Pro Leu Ser
 500 505 510
 Gln Ser Ser Gly His His His His His His His His His Gln Gly
 515 520 525

Tyr Gly Gly Ser Gly

530

<210> 113

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> Example 21 - Lig170 Forward primer

<400> 113

gagcgcaacg caattaa

17

<210> 114

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Example 21 - Lig170_Rreverse primer

<400> 114

atccgctcac aattccacac

20

<210> 115

<211> 100

<212> DNA

<213> Artificial Sequence

<220><223> Example 21 - Lig170 template

<400> 115

gagcgcaacg caattaatgt gagttagctc actcattagg caccaccaggc tttaacttt 60
atgcttcgg ctcgtatgtt gtgtggaatt gtgagcgat 100
<210> 116
<211> 58
<212>
> DNA
<213> Artificial Sequence
<220><223> Example 21 - ILMN_AD_P5
<400> 116
aatgatacgg cgaccaccga gatctacact cttccctac acgacgtct tccgatct 58
<210> 117
<211> 63
<212> DNA
<213> Artificial Sequence
<220><223> Example 21 - ILMN_AD_P7rc_IDX01
<220><221> MISC_FEATURE
<222> (1)..(63)
<223> 5'-P04
<400> 117
gatcggaaga gcacacgtct gaactccagt cacatcacga tctcgatgc cgtcttctgc 60
ttg 63
<210>
118
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Example 22 - ILMN_P5
<400> 118
aatgatacgg cgaccaccga 20
<210> 119
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Example 22 - ILMN_P7
<400> 119

caagcagaag acggcatacg

20

<210> 120

<211> 366

<212> PRT

<213> Artificial Sequence

<220><223> Example 24 - T4-Gp32-HRP1

<400> 120

Met Phe Lys Arg Lys Ser Thr Ala Glu Leu Ala Ala Gln Met Ala Lys

1 5 10 15
Leu Asn Gly Asn Lys Gly Phe Ser Ser Glu Asp Lys Gly Glu Trp Lys

20 25 30
Leu Lys Leu Asp Asn Ala Gly Asn Gly Gln Ala Val Ile Arg Phe Leu

35 40 45
Pro Ser Lys Asn Asp Glu Gln Ala Pro Phe Ala Ile Leu Val Asn His

50 55 60
Gly Phe Lys Lys Asn Gly Lys Trp Tyr Ile Glu Thr Cys Ser Ser Thr

65 70 75 80
His Gly Asp Tyr Asp Ser Cys Pro Val Cys Gln Tyr Ile Ser Lys Asn

85 90 95
Asp Leu Tyr Asn Thr Asp Asn Lys Glu Tyr Ser Leu Val Lys Arg Lys

100 105 110
Thr Ser Tyr Trp Ala Asn Ile Leu Val Val Lys Asp Pro Ala Ala Pro

115 120 125
Glu Asn Glu Gly Lys Val Phe Lys Tyr Arg Phe Gly Lys Lys Ile Trp

130 135 140
Asp Lys Ile Asn Ala Met Ile Ala Val Asp Val Glu Met Gly Glu Thr

145 150 155 160
Pro Val Asp Val Thr Cys Pro Trp Glu Gly Ala Asn Phe Val Leu Lys

165 170 175
Val Lys Gln Val Ser Gly Phe Ser Asn Tyr Asp Glu Ser Lys Phe Leu

180 185 190

Asn Gln Ser Ala Ile Pro Asn Ile Asp Asp Glu Ser Phe Gln Lys Glu

195 200 205
 Leu Phe Glu Gln Met Val Asp Leu Ser Glu Met Thr Ser Lys Asp Lys
 210 215 220
 Phe Lys Ser Phe Glu Glu Leu Asn Thr Lys Phe Gly Gln Val Met Gly
 225 230 235 240
 Thr Ala Val Met Gly Gly Ala Ala Ala Thr Ala Ala Lys Lys Ala Asp
 245 250 255
 Lys Val Ala Asp Asp Leu Asp Ala Phe Asn Val Asp Asp Phe Asn Thr

260 265 270
 Lys Thr Glu Asp Asp Phe Met Ser Ser Ser Ser Gly Ser Ser Ser Ser
 275 280 285
 Ala Asp Asp Thr Asp Leu Asp Asp Leu Leu Asn Asp Leu Gly Gly Asn
 290 295 300
 Asn Gly Gly Asn Asn Met Asn Arg Arg Gly Gly Asn Phe Gly Asn Gln
 305 310 315 320
 Gly Asp Phe Asn Gln Met Tyr Gln Asn Pro Met Met Gly Gly Tyr Asn

325 330 335
 Pro Met Met Asn Pro Gln Ala Met Thr Asp Tyr Tyr Gln Lys Met Gln
 340 345 350
 Glu Tyr Tyr Gln Gln Met Gln His His His His His His His
 355 360 365

<210> 121

<211> 182

<212> DNA

<213> Artificial Sequence

<220><223> Example 24 - UP1-UP2'_TF1L template sequence

<400> 121

aatgatacgg cgaccaccgt gatctacact gttttacaac ctcagcatgg aaaaaggttt 60

cgaactcaac agctgtctgg cagctcgctc tacgcatgct attgctggcg gtattgcttc 120

tgctcttgct ggtggcgcca tgtctaaatt gtcgatacat ctcgtatgcc gtcttctgct 180

tg 182

<210> 122

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Example 24 - UP1 forward primer

<400> 122

aatgatacgg cgaccaccga gatctacac 29

<210> 123

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Example 24 - UP2-18 reverse primer

<400> 123

caagcagaag acggcata 18

<210> 124

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Amino acid sequence comprised in functional IDRs

<400> 124

Tyr Ser Pro Thr Ser Pro Ser

1 5

<210> 125

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Amino acid sequence comprised in functional IDRs

<400> 125

Phe Ser Pro Thr Ser Pro Thr

1 5

<210> 126

<211>

> 7

<212> PRT

<213> Artificial Sequence

<220><223> Amino acid sequence comprised in functional IDRs

<220><221> MISC_FEATURE

<222> (7)..(7)

<223> Xaa is A, N or G

<400> 126

Tyr Ser Pro Thr Ser Pro Xaa

1 5

<210> 127

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Amino acid sequence comprised in functional IDRs

<400> 127

Tyr Ser Pro Gly Ser Pro Ala

1 5

<210> 128

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> Amino acid sequence comprised in functional IDRs

<400> 128

Gln Gln Gln Pro Gln Tyr

1 5

<210> 129

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> YDPTSPS motif

<400> 129

Tyr Asp Pro Thr Ser Pro Ser

1 5

<210> 130

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> YSPTDPS motif

<400> 130

Tyr Ser Pro Thr Asp Pro Ser

1 5

<210> 131

<211> 31

<212> PRT

<213> Artificial Sequence

<220><223> Self-designed double sequence of MEAFSSAKTEDDFMSSSSDDSDLDDLGL

<400> 131

Met Glu Ala Phe Ser Ser Ala Lys Thr Glu Asp Asp Phe Met Ser Ser

1 5 10 15

Ser Ser Ser Asp Asp Ser Asp Leu Asp Asp Leu Leu Ala Gly Leu

20 25 30

<210> 132

<211> 45

<212> PRT

<213> Artificial Sequence

<220><223> Self-designed double sequence of

EIVEAEVDELINSKVEKFKSPESKSKSAADLETDLQLSDMEEFN

<400> 132

Glu Ile Val Glu Ala Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu

1 5 10 15

Lys Phe Lys Ser Pro Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu

20 25 30

Thr Asp Leu Glu Gln Leu Ser Asp Met Glu Glu Phe Asn

35 40 45

<210> 133

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Self designed linker of DDVASEF

<400> 133

Asp Asp Val Ala Ser Glu Phe

1 5

<210> 134

<211> 43

<212> RNA

<213> Artificial Sequence

<220><223> Example 26 - guide-RNA sequence

<400> 134

uaauuucuaac uguuguagau aaagugcuca ucauuggaaa acg 43

<210> 135

<211> 59

<212> DNA

<213> Artificial Sequence

<220><223> Example 26 - top oligo

<220><221> MISC_FEATURE

<222> (1)..(59)

<223> 5' FAM (fluorescein)

<400> 135

gaacgttttc caatgatgag cacttttaaa gttctatgta tcaaagcggc catttgagg 59

<210> 136

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> Example 26 - bottom oligo

<220><221> MISC_FEATURE

<222> (1)..(35)

<223> 3' - BHQ-1 (quencher)

<400> 136

agaactttaa aagtgtcat cattggaaaa cggtc 35

<210> 137

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> Hemagglutinin C terminus

<400> 137

Asp Val Pro Asp Tyr Ala

1

5