



(51) International Patent Classification:

C12Q 1/6848 (2018.01) *G01N 33/50* (2006.01)

(21) International Application Number:

PCT/GB2020/052866

(22) International Filing Date:

11 November 2020 (11.11.2020)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1916379.9 11 November 2019 (11.11.2019) GB

(71) Applicant: **BIOCRUCIBLE LIMITED** [GB/GB]; The Glenn Berge Building, Babraham Research Campus, Babraham, Cambridge, Cambridgeshire CB22 3FH (GB).

(72) Inventors: **ARMES, Niall**; c/o Biocrucible Limited, The Glenn Berge Building, Babraham Research Campus, Babraham, Cambridge, Cambridgeshire CB22 3FH (GB). **WILLIAMS, Hannah**; c/o Biocrucible Limited, The Glenn

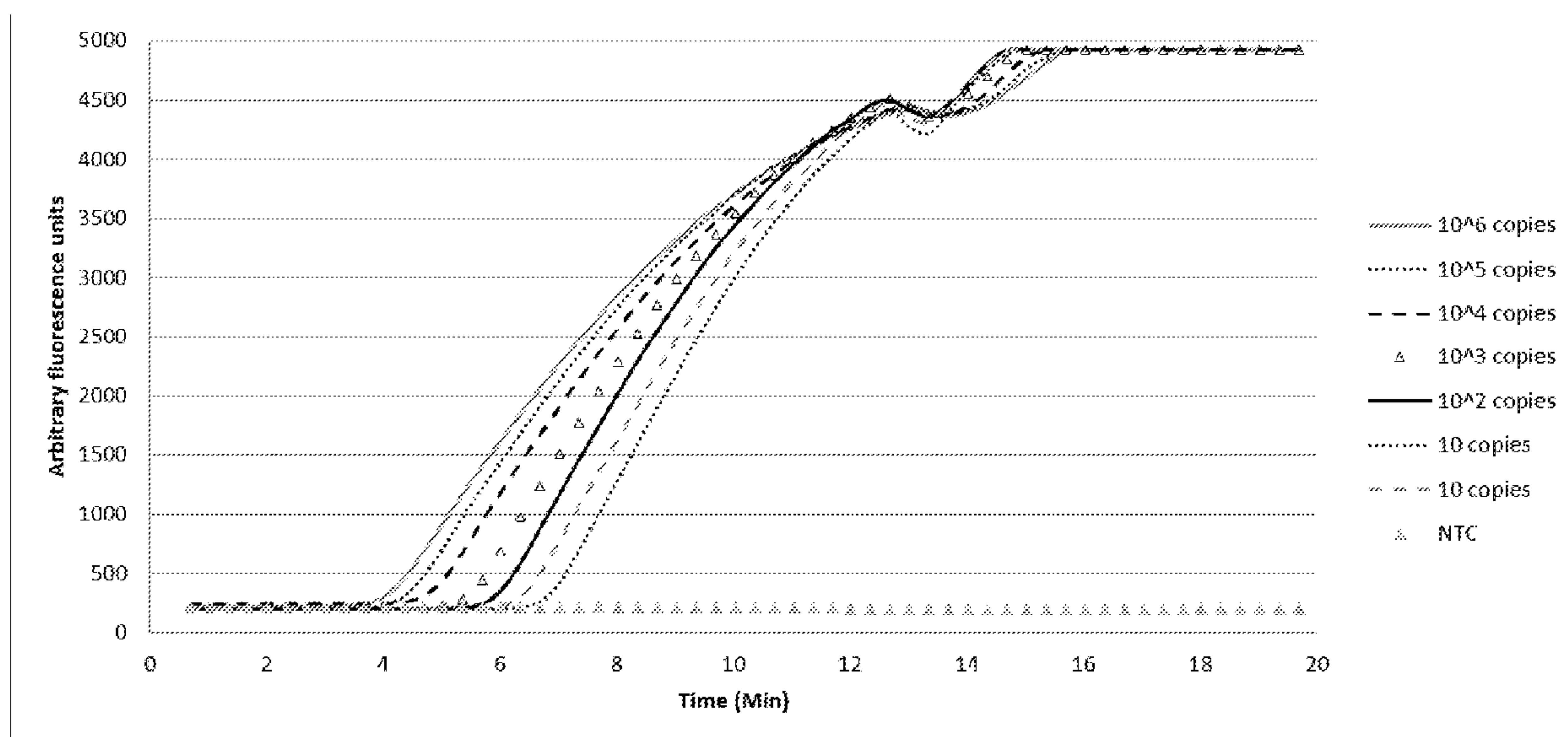
Berge Building, Babraham Research Campus, Babraham, Cambridge, Cambridgeshire CB22 3FH (GB). **FORREST, Matthew**; c/o Biocrucible Limited, The Glenn Berge Building, Babraham Research Campus, Babraham, Cambridge, Cambridgeshire CB22 3FH (GB). **PARKER, Mathew**; c/o Biocrucible Limited, The Glenn Berge Building, Babraham Research Campus, Babraham, Cambridge, Cambridgeshire CB22 3FH (GB). **LIU, Sidong**; c/o Biocrucible Limited, The Glenn Berge Building, Babraham Research Campus, Babraham, Cambridge, Cambridgeshire CB22 3FH (GB). **PARKER, Lauren**; c/o Biocrucible Limited, The Glenn Berge Building, Babraham Research Campus, Babraham, Cambridge, Cambridgeshire CB22 3FH (GB).

(74) Agent: **WILKINSON, Marc, George**; J A Kemp LLP, 14 South Square, Gray's Inn, London, Greater London WC1R 5JJ (GB).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,

(54) Title: BIOCHEMICAL REACTION METHODS AND REAGENTS COMPRISING INTRINSICALLY DISORDERED REGIONS

Figure 1



(57) Abstract: The invention relates to processes for performing biochemical reactions, such as in an aqueous *in vitro* reaction system. The processes involve macromolecules, particularly polypeptides, comprising one or more functional intrinsically disordered regions (IDRs). The invention also relates to IDR-macromolecules, including IDR-polypeptides, including macromolecules or polypeptides comprising a tagged amino acid sequence which comprises or consists of one or more functional IDRs. Such functional IDRs are capable of increasing the efficiency of the biochemical reaction. The invention relates to kits comprising any such macromolecules and polypeptides. The invention further relates to processes for stimulating or enhancing liquid-liquid demixing in a solution using any such macromolecules and polypeptides, including in combination with multivalent metal ions, thereby providing reagents capable of increasing the efficiency of a biochemical reaction.

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

**BIOCHEMICAL REACTION METHODS AND REAGENTS COMPRISING
INTRINSICALLY DISORDERED REGIONS**

FIELD OF THE INVENTION

5

The invention relates to processes for performing biochemical reactions, such as in an aqueous *in vitro* reaction system. The processes involve macromolecules, particularly polypeptides, comprising one or more functional intrinsically disordered regions (IDRs). The invention also relates to IDR-macromolecules, including IDR-
10 polypeptides, including macromolecules or polypeptides comprising a tagged amino acid sequence which comprises or consists of one or more functional IDRs. Such functional IDRs are capable of increasing the efficiency of the biochemical reaction. The invention relates to kits comprising any such macromolecules and polypeptides. The invention further relates to processes for stimulating or enhancing liquid-liquid
15 demixing in a solution using any such macromolecules and polypeptides, including in combination with multivalent metal ions, thereby providing reagents capable of increasing the efficiency of a biochemical reaction.

BACKGROUND TO THE INVENTION

20

The performance of biochemical reactions, and in particular *in vitro* biochemical reactions, is of fundamental importance in the biological sciences. Many biochemical reactions may need to be performed outside of the laboratory, such as at the point of care or in the field. In these settings it may not be possible to control biochemical
25 reactions in the precise manner afforded by the laboratory environment. Improving the efficiency of biochemical reactions performed in these settings would be of value. Indeed, it may be desirable to increase the efficiency of biochemical reactions, regardless of the exact setting, including *in vitro* and *in vivo* biochemical reactions. The present invention addresses these issues.

30

Many biochemical reactions require the use of co-factors to aid in driving performance efficiency. One particular example of such a co-factor is a

macromolecular crowding agent. Crowding agents are essential for the performance of many biochemical reactions. A notable example is the recombinase polymerase amplification (RPA) system for the amplification of nucleic acids. The use of a crowding agent has been considered essential in driving RPA performance efficiency.

5 However, crowding agents may have drawbacks. Accordingly, alternative means for driving performance efficiency of biochemical reactions, including RPA, and that obviate the need for added/exogenous crowding agents would be of use. In addition, reagents that add to or synergise with the functional effects of crowding agents in increasing the performance efficiency of biochemical reactions would be of use. The

10 present invention also addresses these issues.

SUMMARY OF THE INVENTION

The present invention provides a process of performing a biochemical reaction

15 in an aqueous *in vitro* reaction system, wherein the biochemical reaction is dependent on the function of at least one reaction macromolecule, optionally at least one reaction polypeptide, the process comprising: introducing at least one IDR-macromolecule into the *in vitro* reaction system under conditions suitable for performing the reaction, wherein the at least one IDR-macromolecule comprises one or more functional

20 intrinsically disordered regions (IDRs), wherein upon introduction of the at least one IDR-macromolecule into the *in vitro* reaction system the efficiency of the biochemical reaction is increased by the at least one IDR-macromolecule; preferably wherein the at least one IDR-macromolecule is at least one IDR-polypeptide.

In the above-described process, the biochemical reaction may be dependent on

25 the function of the at least one IDR-macromolecule, optionally the at least one IDR-polypeptide, wherein upon its introduction into the *in vitro* reaction system the at least one IDR-macromolecule or the at least one IDR-polypeptide performs its reaction function in the biochemical reaction and increases the efficiency of the reaction.

Any of the herein-described processes may further comprise maintaining the

30 IDR-macromolecule or the IDR-polypeptide in the system to cause liquid-liquid demixing and the formation of a plurality of phase-separated aqueous compartments

within the system by the IDR-macromolecule or the IDR-polypeptide, thereby increasing the efficiency of the biochemical reaction in the system.

Any of the herein-described processes may further comprise maintaining the IDR-macromolecule or the IDR-polypeptide in the system to cause molecules necessary
5 for the performance of the reaction to co-localise with the IDR-macromolecule or the IDR-polypeptide within the plurality of phase-separated aqueous compartments, or to further stimulate or enhance co-localisation of molecules necessary for the performance of the reaction with the IDR-macromolecule or the IDR-polypeptide within the plurality of phase-separated aqueous compartments, thereby increasing the efficiency of the
10 biochemical reaction in the system.

In any of the herein -described processes the plurality of phase-separated aqueous compartments may be a plurality of detectable phase-separated aqueous particles.

In an additional aspect the invention provides a process of performing a
15 biochemical reaction in an aqueous *in vitro* reaction system, wherein the biochemical reaction is dependent on the function of at least one reaction macromolecule, optionally at least one reaction polypeptide, the process comprising: introducing at least one polypeptide tagged with an amino acid sequence comprising or consisting of one or more functional intrinsically disordered regions (IDRs) (IDR-polypeptide) into the *in*
20 *vitro* reaction system under conditions suitable for performing the reaction, and maintaining the IDR-polypeptide in the system to cause liquid-liquid demixing and the formation of a plurality of phase-separated aqueous compartments, preferably detectable phase-separated aqueous particles, within the system by the IDR-polypeptide and to cause molecules necessary for the performance of the reaction to co-localise with the
25 IDR-polypeptide within the compartments thereby increasing the efficiency of the biochemical reaction in the system.

Optionally, in a process according to this additional aspect, the biochemical reaction is dependent on the function of at least one reaction polypeptide, wherein the reaction polypeptide is the at least one IDR-polypeptide, wherein upon introduction into
30 the system the at least one IDR-polypeptide performs its reaction function in the biochemical reaction and increases the efficiency of the reaction in the system.

In any of the processes according to this additional aspect, the conditions suitable for performing the reaction may further comprise providing multivalent metal ions to the IDR-polypeptide thereby further stimulating or enhancing the liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments caused by the IDR-polypeptide and thereby further increasing the efficiency of the biochemical reaction in the system; optionally wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. The multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

In any of the processes according to this additional aspect, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-polypeptide in the *in vitro* reaction system, thereby further stimulating or enhancing the liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments caused by the IDR-polypeptide and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

In any of the processes according to this additional aspect, the conditions suitable for performing the reaction may further comprise providing multivalent metal ions to the IDR-polypeptide thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-polypeptide within the plurality of phase-separated aqueous compartments thereby further increasing the efficiency of the biochemical reaction in the system; optionally wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. The multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

In any of the processes according to this additional aspect, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-

polypeptide in the *in vitro* reaction system, thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-polypeptide within the plurality of phase-separated aqueous compartments and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP
5 is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

In any of the processes according to this additional aspect, the efficiency of the reaction in the system may be increased by the IDR-polypeptide compared to the efficiency of the reaction in the system following introduction of the at least one
10 polypeptide under the same reaction conditions except that the at least one polypeptide has not been tagged with the amino acid sequence comprising or consisting of the one or more functional IDRs.

The invention also provides a process of performing a biochemical reaction in an aqueous *in vitro* reaction system, wherein the biochemical reaction is dependent on
15 the function of at least one reaction macromolecule, optionally at least one reaction polypeptide, the process comprising:

- i. introducing molecules comprising at least one IDR-macromolecule into the system under conditions suitable for performing the reaction, wherein the at
20 least one IDR-macromolecule comprises one or more functional intrinsically disordered regions (IDRs), preferably wherein the at least one IDR-macromolecule is at least one IDR-polypeptide;
- ii. maintaining the IDR-macromolecule or the IDR-polypeptide in the system to
25 cause liquid-liquid demixing in the system, wherein the liquid-liquid demixing is caused by the IDR-macromolecule or the IDR-polypeptide and forms a plurality of phase-separated aqueous compartments within the system;
- iii. maintaining the IDR-macromolecule or the IDR-polypeptide in the system to
30 cause molecules necessary for the performance of the reaction to co-localise

with the IDR-macromolecule or the IDR-polypeptide within the compartments;
and

- iv. allowing the biochemical reaction to proceed within the compartments; wherein
5 the efficiency of the biochemical reaction in the system is increased by the
presence of the at least one IDR-macromolecule.

In the above-described process, the biochemical reaction may be dependent on
the function of the at least one IDR-macromolecule, optionally the at least one IDR-
polypeptide, wherein upon its introduction into the *in vitro* reaction system the at least
10 one IDR-macromolecule or the at least one IDR-polypeptide performs its reaction
function in the biochemical reaction and increases the efficiency of the reaction. The
plurality of phase-separated aqueous compartments may be a plurality of detectable
phase-separated aqueous particles.

In a further aspect the invention provides a process of performing a biochemical
15 reaction in an aqueous *in vitro* reaction system, wherein the biochemical reaction is
dependent on the function of at least one reaction macromolecule, optionally at least
one reaction polypeptide, the process comprising:

- i. introducing molecules comprising at least one polypeptide tagged with an amino
20 acid sequence comprising or consisting of one or more functional intrinsically
disordered regions (IDRs) (IDR-polypeptide) into the system under conditions
suitable for performing the reaction;
- ii. maintaining the IDR-polypeptide in the system to cause liquid-liquid demixing
25 and the formation of a plurality of phase-separated aqueous compartments,
preferably detectable phase-separated aqueous particles, within the system,
wherein the liquid-liquid demixing is caused by the IDR-polypeptide;
- iii. maintaining the IDR-polypeptide in the system to cause molecules necessary for
30 the performance of the reaction to co-localise with the IDR-polypeptide within
the compartments; and

- iv. allowing the biochemical reaction to proceed within the compartments; wherein the efficiency of the biochemical reaction in the system is increased by the presence of the at least one IDR- polypeptide.

Optionally, in a process according to this further aspect, the biochemical
 5 reaction is dependent on the function of at least one reaction polypeptide, wherein the reaction polypeptide is the at least one IDR-polypeptide, wherein upon introduction into the system the at least one IDR-polypeptide performs its reaction function in the biochemical reaction and increases the efficiency of the reaction in the system.

In any of the processes according to this further aspect, the conditions suitable
 10 for performing the reaction may further comprise providing multivalent metal ions to the IDR-polypeptide thereby further stimulating or enhancing the liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments caused by the IDR-polypeptide and thereby further increasing the efficiency of the biochemical reaction in the system; optionally wherein the multivalent metal ions are provided at a
 15 concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. The multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

In any of the processes according to this further aspect, the conditions suitable
 20 for performing the reaction may further comprise providing ATP to the IDR-polypeptide in the *in vitro* reaction system, thereby further stimulating or enhancing the liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments caused by the IDR-polypeptide and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the
 25 system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

In any of the processes according to this further aspect, the conditions suitable
 for performing the reaction may further comprise providing multivalent metal ions to the IDR-polypeptide thereby further stimulating or enhancing molecules necessary for
 30 the performance of the reaction to co-localise with the IDR-polypeptide within the plurality of phase-separated aqueous compartments thereby further increasing the

efficiency of the biochemical reaction in the system; optionally wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. The multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

In any of the processes according to this further aspect, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-macromolecule or IDR-polypeptide in the *in vitro* reaction system, thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-polypeptide within the plurality of phase-separated aqueous compartments and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

In any of the processes according to this further aspect, the conditions suitable for performing the reaction may further comprise providing multivalent metal ions to the IDR-polypeptide thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-macromolecule or the IDR-polypeptide within the plurality of phase-separated aqueous compartments thereby further increasing the efficiency of the biochemical reaction in the system; optionally wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. The multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

In any of the processes according to this further aspect, the efficiency of the reaction in the system may be increased by the IDR-polypeptide compared to the efficiency of the reaction in the system following introduction of the at least one polypeptide under the same reaction conditions except that the at least one polypeptide has not been tagged with the amino acid sequence comprising or consisting of the one or more functional IDRs.

In any of the above-described processes, the process may be a biochemical reaction for synthesizing a nucleic acid molecule in an *in vitro* reaction system comprising:

- (a) providing at least one nucleic acid primer;
- 5 (b) providing a target nucleic acid molecule comprising at least one target strand, and contacting the at least one nucleic acid primer with the target strand, thereby forming a double stranded structure;
- (c) providing the IDR-macromolecule as an IDR-polypeptide, wherein the IDR-polypeptide is polymerase or one or more polypeptide co-factors; and
- 10 (d) allowing the reaction to proceed, thereby extending the 3' end of the at least one nucleic acid primer with polymerase and dNTPs, optionally in the presence of one or more polypeptide co-factors, to generate a double stranded nucleic acid, wherein the first strand comprises a sequence of the target strand and the second strand comprises a sequence which is
- 15 complementary thereto.

Alternatively, in any of the above-described processes, the process may be a biochemical for amplifying a single stranded target nucleic acid molecule or a double stranded target nucleic acid molecule in an *in vitro* reaction system, preferably wherein the target nucleic acid molecule is a DNA molecule.

- 20 The process may be a biochemical reaction for amplifying a double stranded target nucleic acid molecule in an *in vitro* reaction system comprising:

- (a) providing first and second nucleic acid primers;
- (b) providing a double stranded target nucleic acid molecule comprising a first strand and a second strand, and contacting the first and second nucleic acid
- 25 primers with the target nucleic acid molecule thereby forming a first double stranded structure with the first strand and a second double stranded structure with the second strand;
- (c) providing the IDR-macromolecule as an IDR-polypeptide, wherein the IDR-polypeptide is polymerase or one or more protein co-factors;
- 30 (d) allowing the reaction to proceed, thereby extending the 3' ends of the first and second nucleic acid primers with polymerase and dNTPs, optionally in

the presence of one or more protein co-factors, to generate first and second double stranded nucleic acids; and

(e) repeating steps (b) to (d) until a desired degree of amplification is reached.

In the above-described process for amplifying a double stranded target nucleic acid molecule in an *in vitro* reaction system, the process may be a recombinase polymerase amplification (RPA) process of amplifying the double stranded target nucleic acid molecule in the *in vitro* reaction system comprising:

- (a) providing a recombinase agent, optionally a recombinase loading protein, a single strand stabilizing agent, polymerase, first and a second nucleic acid primers, double stranded target nucleic acid comprising a first strand and a second strand, and optionally an exonuclease such as Exonuclease III;
- (b) contacting the recombinase agent with the first and second nucleic acid primers and optionally with the recombinase loading protein to form first and second nucleoprotein primers which comprise a single stranded region at their 3' ends;
- (c) contacting the first and second nucleoprotein primers with the target nucleic acid molecule thereby forming a first double stranded structure with the first strand and a second double stranded structure with the second strand;
- (d) allowing the reaction to proceed, thereby extending the 3' end of the first and second nucleoprotein primers with polymerase and dNTPs to generate first and second double stranded nucleic acids and first and a second displaced nucleic acid strands, wherein the single strand stabilizing agent stabilizes the first and second displaced strands; and
- (e) continuing the reaction by repeating steps (b) to (d) until a desired degree of amplification is reached;

wherein the recombinase agent, and/or the recombinase loading protein, and/or the single strand stabilizing agent, and/or the polymerase is provided as the IDR-polypeptide.

In the above-described RPA process of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, the recombinase agent may be selected from the group consisting of 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 or any functional analog, homolog or derivative thereof, and any
 5 combination thereof, preferably wherein the recombinase agent is UvsX, more preferably *Escherichia* phage vB_EcoM_DalCa UvsX.

In any one of the above-described RPA processes of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, the process may include a recombinase loading protein and wherein the recombinase loading protein is
 10 selected from the group consisting of UvsY, *E. coli* RecO, *E. coli* RecR or any functional analog, homolog or derivative thereof, and any combination thereof, preferably wherein the recombinase loading protein is UvsY, more preferably *Escherichia* phage STO UvsY.

In any one of the above-described RPA processes of amplifying a double
 15 stranded target nucleic acid molecule in the *in vitro* reaction system, the polymerase may be a eukaryotic polymerase selected from the group consisting of pol- α , pol- β , pol- δ , pol- ϵ or any functional analog, homolog or derivative thereof, and any combination thereof. The polymerase may be a prokaryotic polymerase selected from the group consisting of *Bacillus stearothermophilus* polymerase I large fragment, *Bacillus subtilis*
 20 Pol I large fragment (Bsu polymerase), *Listeria monocytogenes* DNA polymerase I, *S. aureus* DNA polymerase I (Sau polymerase), *E. coli* DNA polymerase I Klenow fragment, *E. coli* DNA polymerase I, *E. coli* DNA polymerase II, *E. coli* DNA polymerase III, *E. coli* DNA polymerase IV, *E. coli* DNA polymerase V, or any functional analog, homolog or derivative thereof, and any combination thereof,
 25 preferably wherein the polymerase is *S. aureus* DNA polymerase I (Sau polymerase) or *Bacillus subtilis* Pol I large fragment (Bsu polymerase). The polymerase may be a bacteriophage polymerase selected from the group consisting of bacteriophage T4 gp43 DNA polymerase, T7 DNA polymerase and Phi-29 DNA polymerase, or any functional analog, homolog or derivative thereof, and any combination thereof.

30 In any one of the above-described RPA processes of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, the single strand

stabilizing agent may be selected from the group consisting of Gp32, *E. coli* SSB protein, phage T4 Gp32 protein, phage Rb69 Gp32, phage vB_EcoM_NBG1 Gp32, or any functional analog, homolog or derivative thereof, and any combination thereof, preferably the single strand stabilizing agent is Gp32 or phage vB_EcoM_NBG1 Gp32.

5 In any one of the above-described RPA processes of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, only the recombinase agent may be provided as the IDR-polypeptide, or only the recombinase loading protein may be provided as the IDR-polypeptide, or only the single strand stabilizing agent may be provided as the IDR-polypeptide, or only the polymerase may
10 be provided as the IDR-polypeptide, or only the exonuclease may be provided as the IDR-polypeptide.

In any one of the above-described RPA processes of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, the one or more functional IDRs of the IDR-polypeptide may be tagged to the IDR-polypeptide as an
15 amino acid sequence comprising or consisting of the one or more IDRs such that the IDR-polypeptide is a genetically engineered fusion protein, wherein the one or more functional IDRs are located at the C-terminus of the IDR-polypeptide, at the N-terminus of the IDR-polypeptide, or at both the C-terminus of the IDR-polypeptide and the N-terminus of the IDR-polypeptide, or at any amino acid position along the length of the
20 polypeptide.

In any one of the above-described processes, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may be characterised as a sequence of amino acids which scores greater than 0.5 when analysed by the algorithm MetaDisorder.

25 In any one of the above-described processes, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise or consist of an amino acid sequence which includes one or more repeats of the tri-peptide sequence RGG. In any such process, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise or consist of an amino acid sequence which further
30 includes one or more repeats of the di-peptide sequence FG. In any such process, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may

comprise or consist of an amino acid sequence which further includes at least one aromatic amino acid residue consisting of tyrosine or phenylalanine.

In any one of the above-described processes, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise or consist of the amino acid sequence of

- i. (YNPQGGYQQ)_n (SEQ ID NO: 19), wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3; or
- ii. (YSPTSPTS)_n (SEQ ID NO: 124), wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3; or
- 10 iii. (FSPTSPT)_n (SEQ ID NO: 125), wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3; or
- iv. (YSPTSP-A/N/G)_n (SEQ ID NO: 126), wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3; or
- v. (YSPGSPA)_n (SEQ ID NO: 127), wherein n is a positive integer between 1 and 15 10, optionally wherein n = 1, 2, or 3.

In any one of the above-described processes, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise or consist of an amino acid sequence which is glutamine rich, optionally wherein the amino acid sequence comprises at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, or at least 10 consecutive glutamine residues. In any such process, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise or consist of an amino acid sequence comprising one or more repeats of the tri-peptide sequence QQQ. In any such process, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of the amino acid sequence of (QQQPQY)_n (SEQ ID NO: 128), wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3.

In any one of the above-described processes, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise or consist of a sequence of at least 5 consecutive amino acids of SEQ ID NO:1.

In any one of the above-described processes, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise or consist of an amino acid sequence of at least 5 consecutive amino acids of SEQ ID NO:9.

5 In any one of the above-described processes, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise an amino acid sequence containing one or more aromatic tyrosine residues and one or more phenylalanine residues which can engage in aromatic cation-pi interactions with multivalent metal ions, preferably divalent metal ions.

10 In any one of the above-described processes, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may comprise an amino acid sequence containing one or more arginine residues which can engage in guanidine-metal interactions with multivalent metal ions, preferably divalent metal ions.

15 In any one of the above-described processes, the IDR-macromolecule or the IDR-polypeptide may comprise or consist of the macromolecule or polypeptide tagged with an amino acid sequence which comprises or consists of the amino acid sequence of any one of SEQ ID NOs 1 to 43, or which comprises or consists of a functional variant amino acid sequence of SEQ ID NOs 1 to 43, *e.g.* which has 80% or more identity to any one of SEQ ID NOs 1 to 43.

20 In any one of the above-described RPA processes of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, the IDR-polypeptide may be a single strand stabilizing agent which is Gp32 and which has the amino acid sequence of any one of SEQ ID NOs 65 to 88, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 65 to 88.

25 In any one of the above-described RPA processes of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, the IDR-polypeptide may be a recombinase agent which is UvsX and which has the amino acid sequence of any one of SEQ ID NOs 44 to 59, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence
30 which has 80% or more identity to any one of SEQ ID NOs 44 to 59.

In any one of the above-described RPA processes of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, the IDR-polypeptide may be a recombinase loading protein which is UvsY and which has the amino acid sequence of any one of SEQ ID NOs 60 to 64, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 60 to 64.

In any one of the above-described processes, the process may further comprise providing multivalent metal ions to the IDR-macromolecule or IDR-polypeptide in the *in vitro* reaction system, thereby further simulating or enhancing the liquid-liquid demixing in the *in vitro* reaction system, thereby further increasing the efficiency of the biochemical reaction in the system, wherein the multivalent metal ions further stimulate or enhance the formation of the plurality of phase-separated aqueous compartments within the system thereby further increasing the efficiency of the biochemical reaction in the system, preferably wherein the multivalent metal ions further stimulate or enhance the formation of a plurality of detectable phase-separated aqueous particles; optionally wherein the multivalent metal ions are provided at a concentration of; optionally wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. In any such process, the multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

In any one of the above-described processes, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-macromolecule or the IDR-polypeptide in the *in vitro* reaction system, thereby further simulating or enhancing the liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments caused by the IDR-macromolecule or the IDR-polypeptide and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

In any one of the above-described processes, the conditions suitable for performing the reaction may further comprise providing multivalent metal ions to the

IDR-polypeptide thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-macromolecule or the IDR-polypeptide within the plurality of phase-separated aqueous compartments thereby further increasing the efficiency of the biochemical reaction in the system; optionally
 5 wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. The multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

10 In any one of the above-described processes, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-macromolecule or IDR-polypeptide in the *in vitro* reaction system, thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-polypeptide within the plurality of phase-separated aqueous
 15 compartments and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

In any one of the above-described processes, the biochemical reaction may be performed in a solid phase reaction system which comprises a surface. In any such
 20 process the biochemical reaction may be a process of amplifying a single stranded target nucleic acid molecule or a double stranded target nucleic acid molecule in the *in vitro* reaction system as described above, wherein the at least one nucleic acid primer and/or the IDR-macromolecule and/or the one or more polypeptide co-factors are attached to the surface.

25 In any one of the above-described RPA processes of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, the reaction may be performed in a solid phase reaction system which comprises a surface, and wherein the recombinase agent and/or the recombinase loading protein and/or the single strand stabilizing agent and/or the polymerase and/or the exonuclease and/or the first nucleic
 30 acid primer and/or the second nucleic acid primer is attached to the surface, preferably wherein: (i) the first nucleic acid primer or the second nucleic acid primer is attached to

the surface; or (ii) both the first and second nucleic acid primers are attached to the surface.

In any of the above described process performed in a solid phase reaction system which comprises a surface, the surface may be planar or may be a microbead, preferably the surface comprises silicon, glass, a gel-based material and/or a polymeric material such as polystyrene, more preferably the surface is a microbead comprising a polymeric material such as polystyrene. In any such process, the surface may be bound to a substrate, preferably the surface is planar and/or the substrate comprises glass. The surface, *e.g.* a planar surface and/or the substrate may be provided as a flow-cell.

10 The invention provides a processes for performing a biochemical reaction within a cell in culture by introducing at least one of any of the above-described IDR-macromolecules or at least one of any of the above-described IDR-polypeptides into a cultured host cell, or by expressing at least one of any of the above-described IDR-polypeptides in the cultured host cell, to increase the efficiency of the biochemical
15 reaction within the cultured host cell.

Any of the above-described processes for performing an *in vitro* biochemical reaction may comprise a biochemical reaction performed within a cell in culture, such as by introducing the at least one IDR-macromolecule or the at least one IDR-polypeptide into a cultured host cell, or by expressing the at least one IDR-polypeptide
20 in the cultured host cell, to increase the efficiency of the biochemical reaction within the cultured host cell.

The biochemical reaction may be any reaction which leads to the manipulation of a nucleic acid molecule within the cultured host cell, or which leads to the alteration of a nucleic acid molecule within the cultured host cell, such as a change in the structure
25 of a nucleic acid molecule, such as a change in the nucleotide sequence of a nucleic acid molecule. The biochemical reaction may be any reaction which leads to the synthesis of a nucleic acid molecule in the cultured host cell. The biochemical reaction may be any reaction which leads to the expression of a polypeptide from a nucleic acid molecule. The biochemical reaction may be any reaction which leads to the editing of a nucleic
30 acid sequence within the cultured host cell, *e.g.* wherein the IDR-polypeptide is a CRISPR polypeptide, such as a Cas polypeptide, including a Cas9 polypeptide). The

biochemical reaction may be any reaction which leads to the cleavage of a nucleic acid within the cultured host cell. The biochemical reaction may be any reaction which leads to the homologous recombination of nucleic acids within the cultured host cell. The biochemical reaction may be a metabolic reaction within the cultured host cell to
5 produce one or more biological products of interest within the cultured host cell, or to produce one or more biological products of interest which are secreted from the cultured host cell or otherwise released from the cultured host cell into the culture media.

In any one of the above-described processes, increasing the efficiency of the biochemical reaction may comprise increasing the efficiency of the reaction using the at
10 least one IDR-macromolecule or the at least one IDR-polypeptide compared to the efficiency of the reaction obtained by performing the reaction under the same conditions but wherein the relevant at least one macromolecule or the at least one polypeptide does not comprise or has not been tagged with one or more functional intrinsically disordered region polypeptide sequences, optionally wherein the reaction is performed in the
15 absence of an exogenously added crowding agent.

In any one of the above-described RPA processes, increasing or enhancing the efficiency or performance of an RPA biochemical reaction may comprise increasing the amount of amplified product obtained in the RPA reaction using the at least one IDR-polypeptide compared to the amount of amplified product obtained by performing the
20 reaction under the same conditions but wherein the relevant at least one polypeptide has not been tagged with one or more functional intrinsically disordered region polypeptide sequences, optionally wherein the reaction is performed in the absence of an exogenously added crowding agent.

In any one of the above-described processes involving introducing at least one
25 IDR-macromolecule or IDR-polypeptide into the *in vitro* reaction system, the efficiency of the reaction in the system is increased by the IDR-macromolecule or the IDR-polypeptide compared to the efficiency of the reaction in the system following the introduction of the at least one macromolecule or polypeptide under the same reaction conditions except that the at least one macromolecule or polypeptide does not comprise
30 one or more functional intrinsically disordered regions (IDRs).

In any one of the above-described processes involving introducing at least at least one polypeptide tagged with an amino acid sequence comprising or consisting of one or more functional intrinsically disordered regions (IDRs) (IDR-polypeptide) into the *in vitro* reaction system the efficiency of the reaction in the system is increased by the IDR-polypeptide compared to the efficiency of the reaction in the system following the introduction of the at least one polypeptide under the same reaction conditions except that the at least one polypeptide has not been tagged with the amino acid sequence comprising or consisting of the one or more functional IDRs.

The invention also provides a non-naturally occurring IDR-macromolecule comprising a macromolecule and a tag amino acid sequence, wherein the tag amino acid sequence comprises or consists of one or more functional intrinsically disordered regions (IDRs), wherein the IDR-macromolecule is capable of causing liquid-liquid demixing in an aqueous *in vitro* reaction system. Any such IDR-macromolecule may be capable of causing liquid-liquid demixing and the formation of a plurality of phase-separated aqueous compartments in the system, preferably a plurality of detectable phase-separated aqueous particles. Any such liquid-liquid demixing caused by any such non-naturally occurring IDR-macromolecule in the *in vitro* reaction system may thereby increase the efficiency of the biochemical reaction.

Any one of the above-described IDR-macromolecules may be a non-naturally occurring, artificial or genetically engineered IDR-macromolecule or IDR-polypeptide comprising a macromolecule or polypeptide and the tag amino acid sequence. In the case of an IDR-polypeptide, the tag amino acid sequence may be located at the C-terminus of the polypeptide, at the N-terminus of the polypeptide or at both the C-terminus of the polypeptide and the N-terminus of the polypeptide, or at any amino acid position along the length of the polypeptide.

In any one of the above-described IDR-macromolecules or IDR-polypeptides, the one or more functional IDRs of the tag amino acid sequence are functional IDRs as defined in any one of the above-described processes.

In any one of the above-described IDR-macromolecules or IDR-polypeptides the tag sequence comprises amino acid residues which are capable of engaging in aromatic cation- π interactions with multivalent metal cations, preferably divalent metal cations,

more preferably Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} ions, yet more preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , even more preferably Mg^{2+} .

In any one of the above-described IDR-macromolecules or IDR-polypeptides the IDR-macromolecule or the IDR-polypeptide comprises or consists of the
 5 macromolecule or polypeptide tagged with an amino acid sequence which comprises or consists of the amino acid sequence of any one of SEQ ID NOs 1 to 43, or which comprises or consists of a functional variant amino acid sequence of SEQ ID NOs 1 to 43, *e.g.* which has 80% or more identity to any one of SEQ ID NOs 1 to 43.

In any one of the above-described IDR-polypeptides, the polypeptide to which
 10 the sequence comprising or consisting of the one or more functional IDRs is tagged may be an enzyme, such as a helicase, a gyrase, a recombinase, *e.g.* an RPA recombinase agent, a nuclease, *e.g.* an exonuclease and an endonuclease, a ligase, a glycolyase, a methylase, a methyltransferase, a glucosyltransferase, a polymerase, a kinase, a phosphatase, a gene editing enzyme such as a CRISPR enzyme, *e.g.* a Cas9 enzyme; a
 15 co-factor, *e.g.* as an RPA recombinase loading protein and an RPA single strand stabilizing agent. The polypeptide to which the sequence comprising or consisting of the one or more functional IDRs is tagged may be a ligase, optionally an RB69 ligase, such as RB69 ligase-His2 (SEQ ID NO:112). The polypeptide to which the sequence comprising or consisting of the one or more functional IDRs is tagged may be an RPA
 20 single strand stabilizing agent, preferably Gp32; optionally wherein the IDR-polypeptide has the amino acid sequence of any one of SEQ ID NOs 65 to 88 and SEQ ID NO:120, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 65 to 88 and SEQ ID NO:120. The polypeptide to which the sequence
 25 comprising or consisting of the one or more functional IDRs is tagged may be an RPA recombinase agent, preferably UvsX; optionally wherein the IDR-polypeptide has the amino acid sequence of any one of SEQ ID NOs 44 to 59, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 44 to 59.
 30 The polypeptide to which the sequence comprising or consisting of the one or more functional IDRs is tagged may be an RPA recombinase loading protein, preferably

UvsY; optionally wherein the IDR-polypeptide has the amino acid sequence of any one of SEQ ID NOs 60 to 64, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 60 to 64.

5 The invention also provides an isolated nucleic acid molecule comprising a first nucleic acid sequence encoding any of the above-described IDR-polypeptides; optionally including a second nucleic acid sequence encoding a promoter, wherein the first nucleic acid sequence is operably linked to the second nucleic acid sequence. The invention also provides a recombinant polynucleotide expression vector comprising any
10 such nucleic acid molecule. The invention also provides a host cell comprising any such nucleic acid molecule, or any such recombinant polynucleotide expression vector. The invention also provides a cell culture comprising a growth medium and a population of any such host cells.

 The invention also provides a kit comprising any of the above-described non-
15 naturally occurring IDR-macromolecules or IDR-polypeptides. Any such kit may further comprising additional RPA components comprising an RPA recombinase agent, and/or an RPA recombinase loading protein, and/or polymerase, and/or first and second nucleic acid primers, and/or an exonuclease, and/or a buffer, and/or a source of multivalent metal ions, preferably divalent metal cations. In any such kit, all
20 components may be provided in lyophilized form.

 The invention also provides a process of stimulating or enhancing liquid-liquid demixing in a solution, the process comprising providing a solution comprising any of the above-described IDR-macromolecules or any of the above-described IDR-polypeptides, and contacting the IDR-macromolecule or IDR-polypeptide in solution
25 with multivalent metal ions whereupon liquid-liquid demixing in the solution is stimulated or enhanced. The invention also provides a further process of stimulating or enhancing in an aqueous *in vitro* reaction system liquid-liquid demixing caused by an IDR-macromolecule or an IDR-polypeptide, the process comprising providing any one of the above-described IDR-macromolecules or any one of the above-described IDR-
30 polypeptides into the system, providing multivalent metal ions into the system and allowing the IDR-macromolecule or IDR-polypeptide to contact the multivalent metal

ions whereupon liquid-liquid demixing caused by the IDR-macromolecule or the IDR-polypeptide in the solution is stimulated or enhanced. In any such processes the liquid-liquid demixing may result in the formation of a plurality of phase-separated aqueous compartments, preferably a plurality of detectable phase-separated aqueous particles in the solution. In any such process the multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} . In any such further process, the multivalent metal ions may engage in aromatic cation-pi interactions with amino acid residues in the one or more functional IDR amino acid sequences, thereby promoting liquid-liquid demixing.

10 In any such further process, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-macromolecule or the IDR-polypeptide in the *in vitro* reaction system, thereby further simulating or enhancing the liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments caused by the IDR-macromolecule or the IDR-polypeptide and thereby
15 further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

In any such further process, the conditions suitable for performing the reaction may further comprise providing multivalent metal ions to the IDR-polypeptide thereby
20 further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-macromolecule or the IDR-polypeptide within the plurality of phase-separated aqueous compartments thereby further increasing the efficiency of the biochemical reaction in the system; optionally wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein
25 the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. The multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

In any such further process, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-macromolecule or IDR-polypeptide in
30 the *in vitro* reaction system, thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-polypeptide

within the plurality of phase-separated aqueous compartments and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

5 The invention also provides the use of multivalent metal ions in stimulating or enhancing liquid-liquid demixing in a solution, said demixing being mediated by any one of the above-described IDR-macromolecules or any one of the above-described IDR-polypeptides. The invention also provides the use of multivalent metal ions in stimulating or enhancing in an aqueous *in vitro* reaction system liquid-liquid demixing
10 caused by an IDR-macromolecule or an IDR-polypeptide which has been introduced into the system, wherein said IDR-macromolecule or IDR-polypeptide is any one of the above-described IDR-macromolecules or any one of the above-described IDR-polypeptides. In any such use, said liquid-liquid demixing may result in the formation of a plurality of phase-separated aqueous compartments caused by the IDR-
15 macromolecule or IDR-polypeptide, preferably a plurality of detectable phase-separated aqueous particles in the solution. In any such use, the multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} . In any such use, the multivalent metal ions may engage in aromatic cation-pi interactions with amino acid residues in the one or more functional
20 IDR amino acid sequences, thereby promoting liquid-liquid demixing.

 In any such use, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-macromolecule or the IDR-polypeptide in the *in vitro* reaction system, thereby further stimulating or enhancing the liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments
25 caused by the IDR-macromolecule or the IDR-polypeptide and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

 In any such use, the conditions suitable for performing the reaction may further
30 comprise providing multivalent metal ions to the IDR-polypeptide thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-

localise with the IDR-macromolecule or the IDR-polypeptide within the plurality of phase-separated aqueous compartments thereby further increasing the efficiency of the biochemical reaction in the system; optionally wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM. The multivalent metal ions may be divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

In any such use, the conditions suitable for performing the reaction may further comprise providing ATP to the IDR-macromolecule or IDR-polypeptide in the *in vitro* reaction system, thereby further stimulating or enhancing molecules necessary for the performance of the reaction to co-localise with the IDR-polypeptide within the plurality of phase-separated aqueous compartments and thereby further increasing the efficiency of the biochemical reaction in the system, wherein ATP is provided in the system at a concentration of 1mM to 3.5mM, optionally 1mM to 2mM, preferably 1mM.

The invention also provides any one of the above-described non-naturally occurring IDR-macromolecules or any one of the above-described IDR-polypeptides for use in therapy, for use as a medicament, for use as a pharmaceutical, for use in a diagnostic method or for use as a diagnostic agent.

The invention also provides a process for making any one of the above-described non-naturally occurring IDR-macromolecules or any one of the above-described IDR-polypeptides, the process comprising providing a macromolecule, optionally a polypeptide, and tagging the macromolecule or polypeptide with one or more functional intrinsically disordered region polypeptide sequences. Said tagging may be performed by any means described and defined herein. Said one or more functional intrinsically disordered region polypeptide sequences may be any of the same as described and defined herein. Said macromolecule or polypeptide may be any suitable macromolecule or polypeptide, including any macromolecule or polypeptide described and defined herein.

Any one of the above-described IDR-macromolecules or IDR-polypeptides may increase the efficiency of the biochemical reaction. Increasing the efficiency of the biochemical reaction may comprise increasing the efficiency of the reaction using the

IDR-macromolecule or the IDR-polypeptide compared to the efficiency of the reaction obtained by performing the reaction under the same conditions but wherein the relevant macromolecule or the relevant polypeptide does not comprise or has not been tagged with one or more functional intrinsically disordered region polypeptide sequences, optionally wherein the reaction is performed in the absence of an exogenously added crowding agent.

Any one of the above-described IDR-macromolecules or IDR-polypeptides may increase the efficiency of a biochemical reaction, wherein the reaction is a recombinase polymerase amplification (RPA) reaction. Increasing the efficiency or performance of an RPA biochemical reaction may comprise increasing the amount of amplified product obtained in the RPA reaction using the IDR-polypeptide compared to the amount of amplified product obtained by performing the reaction under the same conditions but wherein the relevant polypeptide does not comprise or has not been tagged with one or more functional intrinsically disordered region polypeptide sequences, optionally wherein the RPA reaction is performed in the absence of an exogenously added crowding agent.

The invention also provides a method for determining the nucleotide sequence of one or more target polynucleotide molecules, said method comprising the steps of:

- (i) performing an above-described process to amplify the one or more target polynucleotide molecules thereby obtaining a population comprising a plurality of copies of the one or more target polynucleotide molecules; and
- (ii) performing one or more nucleic acid sequencing reactions on the population comprising the plurality of copies of the target polynucleotide molecules, preferably wherein the method is performed in a solid phase reaction system which comprises a surface.

The invention also provides the use of any one of the above-described IDR-macromolecules or any one of the above-described IDR-polypeptides in a method for determining the nucleotide sequence of one or more target polynucleotide molecules, preferably wherein the method is as described above.

The invention also provides a polypeptide or an isolated polypeptide which comprises or consists of the amino acid sequence of any one of SEQ ID NOs 1 to 43, or

which comprises or consists of a functional variant amino acid sequence of any one of SEQ ID NOs 1 to 43, *e.g.* which has 80% or more identity to any one of SEQ ID NOs 1 to 43. Any such polypeptide can be attached to/tagged to a macromolecule or a polypeptide to form an IDR-tagged macromolecule or an IDR-tagged polypeptide as described further herein. The macromolecule or polypeptide which is tagged may be a macromolecule or polypeptide required for the performance of a biochemical reaction in an aqueous reaction system. When maintained within the aqueous reaction system under conditions for performing the biochemical reaction, any such IDR-tagged macromolecule or IDR-tagged polypeptide can cause liquid-liquid demixing caused by the amino acid sequence of any one of SEQ ID NOs 1 to 43 or any functional variant thereof thereof and the formation of a plurality of phase-separated aqueous compartments, preferably a plurality of detectable phase-separated aqueous particles, in the system thereby increasing the efficiency of the biochemical reaction in the system. When maintained within the aqueous reaction system under conditions for performing the biochemical reaction, any such IDR-tagged macromolecule or IDR-tagged polypeptide causes molecules necessary for the performance of the reaction to co-localise with the IDR-macromolecule or the IDR-polypeptide within the plurality of phase-separated aqueous compartments, thereby increasing the efficiency of the biochemical reaction in the system.

The aqueous reaction system may be an aqueous *in vitro* reaction system.

The invention also provides an isolated nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide which comprises or consists of the amino acid sequence of any one of SEQ ID NOs 1 to 43, or which comprises or consists of a functional variant amino acid sequence of SEQ ID NOs 1 to 43, *e.g.* which has 80% or more identity to any one of SEQ ID NOs 1 to 43.

The invention also provides the use of an IDR moiety, which is a polypeptide which comprises or consists of one or more functional intrinsically disordered regions (IDRs), in producing an IDR-tagged macromolecule or IDR-tagged polypeptide wherein the IDR moiety is attached to/tagged to a macromolecule or a polypeptide; wherein the macromolecule or polypeptide which is tagged is a macromolecule or polypeptide required for the performance of a biochemical reaction in an aqueous

reaction system, and wherein when maintained within the aqueous reaction system under conditions for performing the biochemical reaction, the IDR-tagged macromolecule or IDR-tagged polypeptide causes liquid-liquid demixing caused by the IDR moiety and causes the formation of a plurality of phase-separated aqueous compartments, preferably a plurality of detectable phase-separated aqueous particles, in the system thereby increasing the efficiency of the biochemical reaction in the system. When maintained within the aqueous reaction system under conditions for performing the biochemical reaction, any such IDR-tagged macromolecule or IDR-tagged polypeptide causes molecules necessary for the performance of the reaction to co-localise with the IDR-tagged macromolecule or the IDR-tagged polypeptide within the plurality of phase-separated aqueous compartments, thereby increasing the efficiency of the biochemical reaction in the system.

Preferably, the IDR moiety is attached to/tagged to a polypeptide, thereby producing an IDR-tagged polypeptide, preferably produced as a recombinant genetic fusion protein.

Preferably, the IDR moiety is a polypeptide which comprises or consists of the amino acid sequence of any one of SEQ ID NOs 1 to 43, or a polypeptide which comprises or consists of a functional variant amino acid sequence of any one of SEQ ID NOs 1 to 43, *e.g.* which has 80% or more identity to any one of SEQ ID NOs 1 to 43.

Any of the above-described IDR-tagged macromolecules or IDR-tagged polypeptides may be defined as non-naturally occurring, artificial or genetically engineered macromolecules or polypeptides.

Any of the above-described IDR-tagged macromolecules or IDR-tagged polypeptides may further possess the characteristic features of any one or more of the IDR-macromolecules or IDR-polypeptides described and defined herein.

The aqueous reaction system may be an aqueous *in vitro* reaction system.

The invention further provides an IDR-tagged macromolecule or an IDR-tagged polypeptide obtained in accordance with any of the above-described uses.

The invention also provides a method for producing an IDR-tagged macromolecule or an IDR-tagged polypeptide comprising providing a macromolecule or a polypeptide and attaching/tagging an IDR moiety thereto, wherein the IDR moiety

is a polypeptide which comprises or consists of one or more functional intrinsically disordered regions (IDRs); wherein the macromolecule or polypeptide which is tagged is a macromolecule or polypeptide required for the performance of a biochemical reaction in an aqueous reaction system, and wherein when maintained within the aqueous reaction system under conditions for performing the biochemical reaction, the IDR-tagged macromolecule or IDR-tagged polypeptide causes liquid-liquid demixing caused by the IDR moiety and causes the formation of a plurality of phase-separated aqueous compartments, preferably a plurality of detectable phase-separated aqueous particles, in the system thereby increasing the efficiency of the biochemical reaction in the system. When maintained within the aqueous reaction system under conditions for performing the biochemical reaction, any such IDR-tagged macromolecule or IDR-tagged polypeptide causes molecules necessary for the performance of the reaction to co-localise with the IDR-tagged macromolecule or the IDR-tagged polypeptide within the plurality of phase-separated aqueous compartments, thereby increasing the efficiency of the biochemical reaction in the system.

Preferably, the method comprises providing a polypeptide and attaching/tagging an IDR moiety thereto to produce an IDR-tagged polypeptide, preferably produced as a recombinant genetic fusion protein.

Preferably, the IDR moiety is a polypeptide which comprises or consists of the amino acid sequence of any one of SEQ ID NOs 1 to 43, or a polypeptide which comprises or consists of a functional variant amino acid sequence of any one of SEQ ID NOs 1 to 43, *e.g.* which has 80% or more identity to any one of SEQ ID NOs 1 to 43.

Any of the above-described IDR-tagged macromolecules or IDR-tagged polypeptides may be defined as non-naturally occurring, artificial or genetically engineered macromolecules or polypeptides.

Any of the above-described IDR-tagged macromolecules or IDR-tagged polypeptides may further possess the characteristic features of any one or more of the IDR-macromolecules or IDR-polypeptides described and defined herein.

The aqueous reaction system may be an aqueous *in vitro* reaction system.

The invention further provides an IDR-tagged macromolecule or an IDR-tagged polypeptide obtained by any of the above-described methods.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-HIS2) at varying template nucleic acid concentrations.

Figure 2 shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-HIS5) at varying template nucleic acid concentrations.

Figure 3 shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-HRP1) at varying template nucleic acid concentrations.

Figure 4 shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-Sup1) at varying template nucleic acid concentrations.

Figure 5 shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-Sup2) at varying template nucleic acid concentrations. The experiments shown in Figures 5A, B, C and D use respectively a Gp32 fusion protein with one, two, three and four Sup2 IDR repeats.

Figure 6 shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-HIS5) at varying MgOAc concentrations.

Figure 7 shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-HIS2) at varying phosphocreatine concentrations.

Figure 8 shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-HRP1) at varying KOAc concentrations.

Figure 9 shows real-time recombinase polymerase amplification traces using Gp32 tagged with seven histidine residues (for protein purification purpose, *i.e.* no IDR tag) compared with an IDR-tagged Gp32 fusion protein (Gp32-Sup1) either in the presence or absence of a crowding agent (PEG).

Figure 10 shows the effect of multivalent metal cations on the promotion of phase separation (particle formation) mediated by IDR amino acid sequence tags in the

absence of a crowding agent. The IDR amino acid sequences were tagged to the Gp32 protein to create the Gp32-HIS2 fusion protein (Figure 10A), the Gp32-HRP1 fusion protein (Figure 10B), the Gp32-Sup1 fusion protein (Figure 10C) and the Gp32-Fib fusion protein (Figure 10D). In each case the effect of representative concentrations of
 5 divalent metal cations, *i.e.* magnesium (MgOAc), manganese (MgCl₂) and calcium (CaCl₂) was tested.

Figure 11 shows the effect of multivalent metal cations on the promotion of phase separation (particle formation) mediated by IDR amino acid sequence tags in the absence of a crowding agent. The IDR amino acid sequences were tagged to the Gp32
 10 protein to create the Gp32-Fib fusion protein (Figure 11A), the Gp32-Sup1 fusion protein (Figure 11B), the Gp32-HIS2 fusion protein (Figure 11C), the Gp32-HRP1 fusion protein (Figure 11D), the Gp32-HIS5 fusion protein (Figure 11E). In each case the effect of representative concentrations of divalent metal cations, *i.e.* magnesium (MgOAc), manganese (MgCl₂) and calcium (CaCl₂) was tested.

Figure 12 shows the effect of a divalent metal cation, *i.e.* magnesium (MgOAc),
 15 on the capability of an IDR-tagged Gp32 fusion protein (Gp32-HRP1) to promote phase separation (particle formation) in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

Figure 13 shows the effect of a divalent metal cation, *i.e.* magnesium (MgOAc),
 20 on the capability of an IDR-tagged Gp32 fusion protein (Gp32-HIS2) to promote phase separation (particle formation) in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

Figure 14 shows the effect of varying concentrations of a divalent metal cation, *i.e.* magnesium (MgOAc), on the capability of an IDR-tagged Gp32 fusion protein
 25 (Gp32-HRP1) to promote phase separation (particle formation) in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

Figure 15 shows the effect of varying concentrations of a divalent metal cation, *i.e.* magnesium (MgOAc), on the capability of an IDR-tagged Gp32 fusion protein
 (Gp32-HIS2) to promote phase separation (particle formation) in an exemplary *in vitro*
 30 biochemical reaction environment in the absence of a crowding agent.

Figure 16 shows the effect of addition of a divalent metal cation, *i.e.* magnesium (MgOAc), on the capability of an IDR-tagged Gp32 fusion protein (Gp32-HRP1) to promote phase separation in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent. Phase separation is demonstrated by the formation of an opaque solution following the addition of MgOAc due to particle formation (Figure 16A) and particle formation is further demonstrated by pelleting of the particles (Figure 16B). RPA protein components are demonstrated to associate with particles as revealed by SDS-PAGE analysis of pelleted material.

Figure 17 shows real-time recombinase polymerase amplification traces using a native Gp32 fusion protein either in the presence of absence of crowding agent. The experiment reveals that Gp32 which is not tagged with an amino acid sequence comprising an intrinsically disordered region (IDR) is unable to mediate amplification in the absence of a crowding agent.

Figure 18. Figure 18A is a cartoon depicting a reaction mixture set up for real-time amplification using dual-primer beads. Figure 18B is a cartoon depicting amplified products in the real-time reaction. Figure 18C is a cartoon depicting amplification characterisation in the end-point reaction. Figure 18D shows real-time recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-HIS2) using primers attached to a solid surface or using primers free in solution. Figures 18E and F show end-point recombinase polymerase amplification traces using an IDR-tagged Gp32 fusion protein (Gp32-HIS2) using primers attached to a solid surface.

Figure 19 shows disorder profiles generated using the MetaDisorder software program for Gp32 (Figure 19A), UvsY (Figure 19B) and UvsX (Figure 19C).

Figure 20 shows the effect of varying concentrations of a divalent metal cation, *i.e.* magnesium (MgCl₂), on the capability of an IDR-tagged RB69 ligase fusion protein (RB69 ligase-HIS2) to promote phase separation (particle formation) in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

Figure 21 shows the ligase activity performance of an IDR-tagged RB69 ligase fusion protein (RB69 ligase-HIS2) in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

Figure 22 shows the ligase activity performance of an IDR-tagged RB69 ligase fusion protein (RB69 ligase-HIS2) in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent compared with untagged RB69 ligase and T4 DNA ligase.

5 **Figure 23** shows the ligase activity performance of an IDR-tagged RB69 ligase fusion protein (RB69 ligase-HIS2) in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent compared with NEBNext Ultra II ligation master mix.

10 **Figure 24** shows the effect of ATP on the capability of an IDR-tagged RB69 ligase fusion protein (RB69 ligase-HIS2) to promote phase separation (particle formation) in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

15 **Figure 25** shows brightfield and fluorescent images of representative sections of FlexWell™ chambers that had 0, 5, 10, 20, 40 or 80 copies of single stranded UP1-UP2'-TF1L template per bead annealed at 50°C for 1 hour and then amplified by recombinase polymerase amplification using an IDR-tagged Gp32 fusion protein (Gp32-Hrp1) using primers attached to a solid surface in the absence of a crowding agent such as PEG. Amplification was detected by nicking amplicons with Nt. BbvCI and extending the nick with Aminoallyl-dUTP-XX-ATTO-594. No fluorescence was
20 observed on beads where template had not been added and an increasing amount of fluorescence was observed on beads where increasing amounts of template had been annealed. This shows that amplification occurred on the solid surface of the beads in the absence of a crowding agent.

25 **Figure 26 A** shows brightfield and fluorescent images demonstrating the formation of phase-separated aqueous particles mediated by an IDR-tagged Gp32 fusion protein (Gp32-Hrp1). Figure 26 B shows a plot demonstrating increased efficiency of a reaction (rate of nucleic acid cutting by Cas12a) upon the formation of phase-separated aqueous particles mediated by Gp32-Hrp1.

30

DETAILED DESCRIPTION OF THE INVENTION

Recombinase polymerase amplification is a technique for amplifying nucleic acid molecules. The system utilises, *inter alia*, a recombinase enzyme and preferably a recombinase loading protein. These protein components form a complex with amplification primers. Following binding to the target nucleic acid molecule to be amplified, the complex “scans” the target nucleic acid molecule and “searches” for regions of complementarity between the target and primer sequences. Once a complementary region is found the complex facilitates the binding of the primer to the target sequence. A polymerase enzyme can then extend the primer to generate a copy of the target sequence. The use of the recombinase complex provides a key difference over other nucleic acid amplification methods such as PCR. In RPA there is no requirement for melting and annealing steps driven by thermal cycling, as the recombinase complex provides an entirely enzymatic-based solution to the issue of primer binding. RPA is thus an isothermal technique. The absence of a requirement for extreme thermal cycling means that RPA has many obvious advantages over techniques such as PCR.

A well-documented requirement in RPA methods is the presence of a “crowding agent”, also commonly referred to in the technical field as a “macromolecular crowding agent”. These agents are well known in the technical field and have an art-understood meaning. Crowding agents are discussed in more detail herein. One of the most commonly used crowding agents in RPA methods is polyethylene glycol (PEG), although other crowding agents can be used. Prior to the present invention the use of a crowding agent was considered to be an essential requirement in RPA methods.

The present inventors have surprisingly discovered that it is possible to bypass the hitherto-perceived critical requirement for a crowding agent in RPA methods. The present invention is founded on this discovery.

The inventors have surprisingly discovered that by “tagging” a macromolecule, such as a protein component required in RPA methods, with amino acid sequences which comprise or consist of one or more functional “intrinsically disordered regions” (IDRs), the IDR amino acid sequence tag is capable of facilitating efficient RPA in the

complete absence of a crowding agent. As such, efficient amplification can be achieved in an RPA system without reliance on a crowding agent, thus reducing the complexity of RPA reactions.

The inventors have also surprisingly discovered that the efficiency of
5 amplification in RPA methods involving an IDR-tagged macromolecular component in the absence of a crowding agent can be correlated with the functional capability of the IDR amino acid tag sequence to promote liquid-liquid demixing leading to phase separation in the biochemical reaction system. Phase separation may be assessed by the formation in the biochemical reaction milieu of phase-separated aqueous compartments,
10 particularly spherical-like aqueous globular foci or phase-separated aqueous particles which are amenable to detection by standard methods, including by microscopic observation, as described further herein.

Furthermore, the inventors have also surprisingly discovered that the provision of an IDR-tagged macromolecular component and a crowding agent can provide
15 additive and even synergistic effects with respect to the efficiency of amplification in RPA methods.

Yet further, the inventors have surprisingly discovered that the efficiency of amplification in RPA methods involving an IDR amino acid-tagged macromolecular component in the absence of a crowding agent can be correlated with the concentration
20 of multivalent metal cations introduced into the biochemical reaction milieu. Thus multivalent metal cations can further stimulate or enhance liquid-liquid demixing caused by IDR-macromolecules or IDR-polypeptides and thereby further increase reaction efficiency.

The inventors have also surprisingly discovered that certain concentrations of
25 ATP, as described further herein, can further stimulate or enhance liquid-liquid demixing caused by IDR-macromolecules or IDR-polypeptides and thereby further increase reaction efficiency.

Based on these surprising discoveries the invention provides methods and reagents which increase the efficiency of enzymatic-based *in vitro* biochemical
30 reactions, including RPA reactions, as further described herein.

The IDR amino acid sequences and IDR reagents described and defined herein have broad applicability as useful reagents to be applied to any suitable macromolecular component of a biochemical reaction, such as a polypeptide, thereby promoting liquid-liquid demixing and including promoting phase separation in the biochemical reaction milieu without reliance on macromolecular crowding agents, particularly when the IDR amino acid sequences are used in concert with a multivalent metal cation. The invention further embraces the use of a multivalent metal cation, such as a divalent metal cation, or any functional equivalent thereof, in promoting IDR amino acid sequence-mediated phase separation in the biochemical reaction milieu, without reliance on macromolecular crowding agents.

The invention thus provides IDR-based processes, macromolecules, polypeptides, nucleic acids, vectors, host cells and uses, as described and defined further herein.

Elements of the invention are described in turn below.

Biochemical reactions

As explained above, the inventors have surprisingly discovered that it is possible to bypass the requirement for a crowding agent, a previously-considered essential component of RPA and other reactions. As described in detail herein, this may be achieved by attaching/tethering/tagging an amino acid sequence comprising one or more functional intrinsically disordered regions (IDRs) to a protein component required in the RPA reaction. The inventors have also surprisingly shown that a functional intrinsically disordered region attached to a ligase enzyme is capable of increasing the efficiency of a ligase reaction. The inventors have shown that the degree of phase separation induced by the IDR amino acid sequence can be correlated with the efficiency of the reaction, *e.g.* amplification, in the absence of a crowding agent, and can be enhanced with a multivalent metal cation. Based on these surprising observations it is plausibly expected that such IDR amino acid sequences associated with a macromolecule or protein component of a biochemical reaction will improve the

efficiency of the reaction in the *in vitro* or *in vivo* biochemical reaction environment, particularly in the absence of an added/exogenous crowding agent.

Accordingly, the invention embraces the use of any of the IDR amino acid sequences described and defined herein to be applied to any suitable macromolecular or polypeptide component of an *in vitro* or *in vivo* biochemical reaction, thus providing IDR reagents capable of promoting liquid-liquid demixing in the biochemical reaction milieu and increasing the efficiency of the biochemical reaction. Such liquid-liquid demixing in the biochemical reaction milieu may result in phase separation the biochemical reaction milieu. Such liquid-liquid demixing in the biochemical reaction milieu may result in phase separation leading to, causing or promoting the formation of phase-separated aqueous compartments, including detectable phase-separated aqueous particles in the biochemical reaction milieu, as described further herein. Such IDR reagents, or IDR-based reagents, as described and defined further herein may be referred to interchangeably to describe any one or more of an IDR-macromolecule or IDR-tagged macromolecule, or an IDR-polypeptide or IDR-tagged polypeptide.

In any one of the methods, processes and uses, or in any one of the non-naturally occurring IDR-macromolecules, IDR-fusion macromolecules or isolated nucleic acid molecules encoding the same, recombinant polynucleotide expression vectors or host cells described and defined herein, increasing or enhancing the efficiency or performance of a biochemical reaction may comprise increasing the efficiency of the reaction using any one or more of the IDR-based macromolecules or polypeptides described herein compared to the efficiency obtained by performing the reaction under the same conditions but wherein the relevant macromolecule or polypeptide does not comprise or has not been tagged with one or more functional intrinsically disordered region polypeptide sequences, optionally wherein the reaction is performed in the absence of an exogenously added crowding agent.

Increasing or enhancing the efficiency or performance of a biochemical reaction is to be understood according to generally accepted concepts. For example, reaction efficiency in an RPA reaction, or in any other nucleic acid amplification reaction, may be understood as providing an equivalent total population of amplicons using comparatively less starting target nucleic acid, or providing a comparatively faster time

to detection or a comparatively faster rate of amplification using the same amount of starting target nucleic acid.

Increasing or enhancing the efficiency or performance of an RPA biochemical reaction may comprise increasing the amount of amplified product obtained in the RPA reaction using any one or more of the IDR-based macromolecules or polypeptides described herein compared to the amount of amplified product obtained by performing the reaction under the same conditions but wherein the relevant macromolecule or polypeptide does not comprise or has not been tagged with one or more functional intrinsically disordered region polypeptide sequences, and optionally wherein the reaction is performed in the absence of an exogenously added crowding agent.

Increasing the efficiency of a biochemical reaction in a reaction system, such as an *in vitro* reaction system, may comprise increasing any measurable parameter of the reaction in the system over a specified time period, such as the rate of the reaction over a time period, the amount of a substrate consumed over a time period, the amount of a product produced over a time period etc.

Increasing the efficiency of a biochemical reaction in a reaction system, such as an *in vitro* reaction system, may comprise increasing the parameter of the reaction within detectable phase-separated compartments, such as detectable phase-separated aqueous particles. This can, for example, be inferred indirectly by measuring the parameter of the reaction and correlating the increase with formation of detectable phase-separated aqueous particles and/or the detectable co-localisation of reaction molecules into detectable phase-separated aqueous particles.

Described herein are straightforward bioinformatics methods and phase separation assays which can be used to establish whether any IDR amino acid sequence is capable of functioning in the required manner of promoting liquid-liquid demixing and phase separation in the desired biochemical reaction milieu when used with a given macromolecule or protein and included in the desired *in vitro* biochemical reaction environment. Moreover the suitability of any given co-factor, in particular a multivalent, *e.g.* divalent, metal cation, may be established in these assays in a very straightforward way.

Accordingly, the invention provides IDR reagents as described and defined herein that may usefully be applied in any given desired *in vitro* or *in vivo* biochemical reaction environment.

Any of the IDR amino acid sequences as described and defined herein may be
5 used with any macromolecular or protein component required for the performance of an *in vitro* or *in vivo* biochemical reaction, such as any of the reactions described herein.

Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of a nucleic acid synthesis reaction.

10 Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of a nucleic acid synthesis reaction wherein a polymerase is used to synthesize a new nucleic acid molecule by extending a primer nucleic acid molecule.

Any of the IDR amino acid sequences as described and defined herein may be
15 used with any macromolecular or protein component required for the performance of a nucleic acid amplification reaction. The nucleic acid amplification reaction may be a reaction which involves thermal cycling. The nucleic acid amplification reaction may be an isothermal amplification reaction. The nucleic acid amplification reaction may be polymerase chain reaction (PCR), polymerase spiral reaction (PSR), loop mediated
20 isothermal amplification (LAMP), nucleic acid sequence based amplification (NASBA), self-sustained sequence replication (3SR), rolling circle amplification (RCA), strand displacement amplification (SDA), multiple displacement amplification (MDA), ligase chain reaction (LCR), helicase dependant amplification (HDA), ramification amplification method (RAM), recombinase polymerase amplification (RPA),
25 transcription-mediated amplification (TMA) or nicking enzyme amplification reaction (NEAR).

Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of a gene editing reaction.

Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of a CRISPR reaction.

Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of a prime editing gene editing reaction, wherein a CRISPR enzyme, such as a Cas enzyme, *e.g.* Cas9, is provided in a complex with at least a reverse transcriptase enzyme, optionally further with a prime editing guide RNA (pegRNA), and wherein any component of the prime editing complex may be provided tagged with one or more functional intrinsically disordered region (IDR) polypeptide sequences, *e.g.* wherein the CRISPR enzyme is tagged with the one or more functional IDR polypeptide sequences or wherein the reverse transcriptase enzyme is tagged with the one or more functional IDR polypeptide sequences.

Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of a ligation reaction.

Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of an exonuclease reaction.

Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of an endonuclease reaction, a transcription reaction, a DNA methylation reaction, a DNA glycosylation reaction, an antibody-antigen reaction, a drug-target reaction.

Any of the IDR amino acid sequences as described and defined herein may be used with any macromolecular or protein component required for the performance of a reaction involving protein:protein interactions.

A process for performing an *in vitro* biochemical reaction as used herein is intended to embrace biochemical reactions performed directly in solution in a reaction vessel, such as an RPA reaction as described further herein.

A process for performing an *in vitro* biochemical reaction as used herein also includes biochemical reactions performed within cells in culture, such as by expressing

an IDR reagent as defined herein in a cultured host cell to increase the efficiency of a biochemical reaction within the cultured host cell.

A process for performing an *in vitro* biochemical reaction as used herein includes a biochemical reaction performed within a host cell in culture by introducing
5 an IDR reagent as defined herein into a cultured host cell or expressing an IDR reagent as defined herein in a cultured host cell to increase the efficiency of a biochemical reaction within the cultured host cell, wherein the biochemical reaction is any reaction which leads to the manipulation of a nucleic acid molecule within the cultured host cell, or which leads to the alteration of a nucleic acid molecule within the cultured host cell,
10 such as a change in the structure of a nucleic acid molecule, such as a change in the nucleotide sequence of a nucleic acid molecule.

A process for performing an *in vitro* biochemical reaction as used herein includes a biochemical reaction performed within cells in culture by introducing an IDR reagent as defined herein into a cultured host cell or expressing an IDR reagent as
15 defined herein in a cultured host cell to increase the efficiency of a biochemical reaction within the cultured host cell, wherein the biochemical reaction is any reaction which leads to the synthesis of a nucleic acid molecule in the cultured host cell.

A process for performing an *in vitro* biochemical reaction as used herein includes biochemical reactions performed within cells in culture by introducing an IDR
20 reagent as defined herein into a cultured host cell or expressing an IDR reagent as defined herein in a cultured host cell to increase the efficiency of a biochemical reaction within the cultured host cell, wherein the biochemical reaction is any reaction which leads to the expression of a polypeptide from a nucleic acid molecule.

A process for performing an *in vitro* biochemical reaction as used herein
25 includes biochemical reactions performed within cells in culture by introducing an IDR reagent as defined herein into a cultured host cell or expressing an IDR reagent as defined herein in a cultured host cell to increase the efficiency of a biochemical reaction within the cultured host cell, wherein the biochemical reaction is any reaction which leads to the editing of a nucleic acid sequence within the cultured host cell (*e.g.* wherein
30 the IDR-polypeptide is a CRISPR polypeptide, such as a Cas polypeptide, including a Cas9 polypeptide, or wherein the IDR-polypeptide is a polypeptide in complex with a

CRISPR polypeptide, such as wherein the IDR-polypeptide is a reverse transcriptase enzyme), the cleavage of a nucleic acid within the cultured host cell and homologous recombination of nucleic acids within the cultured host cell.

A process for performing an *in vitro* biochemical reaction as used herein
5 includes biochemical reactions performed within cells in culture by introducing an IDR reagent as defined herein into a cultured host cell or expressing an IDR reagent as defined herein in a cultured host cell to increase the efficiency of a biochemical reaction within the cultured host cell, wherein the biochemical reaction is a metabolic reaction within the cultured host cell to produce one or more biological products of interest
10 within the cultured host cell, or to produce one or more biological products of interest which are secreted from the cultured host cell or otherwise released from the cultured host cell into the culture media.

The invention is also intended to embrace biochemical reactions performed *ex vivo*, for example by expressing the IDR reagents defined herein in cells of *e.g.* a tissue
15 culture or any other suitable complex biological system developed external to the body. Accordingly, any reference to a process for performing a biochemical reaction in an aqueous *in vitro* reaction system as used herein using any of the IDR reagents as defined herein may alternatively be defined as a process for performing a biochemical reaction in an aqueous *ex vivo* reaction system using any of the IDR reagents as defined herein.

20 The invention also provides processes, reagents and methods for increasing the efficiency of a biochemical reaction *in vivo*. Accordingly, any reference to a process for performing a biochemical reaction in an aqueous *in vitro* reaction system as used herein using any of the IDR reagents as defined herein may alternatively be defined as a process for performing a biochemical reaction in an aqueous *in vivo* reaction system
25 using any of the IDR reagents as defined herein.

The invention provides any non-naturally occurring IDR-macromolecule or IDR-polypeptide described or defined herein for use in therapy, for use as a therapeutic, for use as a medicament, for use as a pharmaceutical agent or for use as a diagnostic agent.

The invention provides any non-naturally occurring IDR-macromolecule or IDR-polypeptide described or defined herein for use in a method for the treatment of the human or animal body by therapy.

5 The invention provides any non-naturally occurring IDR-macromolecule or IDR-polypeptide described or defined herein for use in a diagnostic method practised on the human or animal body.

The invention provides any non-naturally occurring IDR-macromolecule or IDR-polypeptide described or defined herein for use in the manufacture of a medicament for the treatment of the human or animal body by therapy.

10 The invention provides any non-naturally occurring IDR-macromolecule or IDR-polypeptide described or defined herein for use in the manufacture of a diagnostic agent for a diagnostic method practised on the human or animal body.

The invention provides a method of treatment of a human or animal comprising administering a therapeutically effective amount of any non-naturally occurring IDR-
15 macromolecule or IDR-polypeptide described or defined herein to a human or animal in need thereof.

In any one of the above-described processes, reagents and methods for increasing the efficiency of a biochemical reaction, the non-naturally occurring IDR-macromolecule or IDR-polypeptide is capable of promoting liquid-liquid demixing.
20 Said liquid-liquid demixing may be capable of promoting the formation of phase-separated aqueous compartments in solution, including detectable phase-separated aqueous particles in solution. Said liquid-liquid demixing or said formation of detectable phase-separated compartments or particles thereby increases the efficiency of the biochemical reaction caused by the IDR-macromolecule or the IDR-polypeptide.

25 A process for performing an *in vitro* biochemical reaction as used herein includes any biochemical reaction performed *in vitro* in solution in a reaction vessel or within a host cell in culture by introducing an IDR reagent as defined herein into the solution or introducing or expressing an IDR reagent in a cultured host cell to promote liquid-liquid demixing in the solution or in the cultured host cell. In any such
30 biochemical reaction, the liquid-liquid demixing in the solution or in the cultured host

cell promotes phase separation in the solution or in the cultured host cell, as described and defined herein.

Any such biochemical reaction may be performed in order to assess the efficacy of any IDR amino acid sequence as described and defined herein in promoting liquid-liquid demixing in the solution or in the cultured host cell and/or in promoting phase separation in the solution or in the cultured host cell.

Any such biochemical reaction may be performed in order to assess the efficacy of a test agent, such as a drug, a polypeptide or any other molecule, in stimulating or enhancing liquid-liquid demixing mediated by the IDR amino acid sequence in the solution or in the cultured host cell and/or in stimulating or enhancing phase separation mediated by the IDR amino acid sequence in the solution or in the cultured host cell, preferably wherein the test agent interacts with the IDR amino acid sequence.

Any such biochemical reaction may be performed in order to assess the efficacy of a test agent, such as a drug, a polypeptide or any other molecule, in inhibiting liquid-liquid demixing mediated by the IDR amino acid sequence in the solution or in the cultured host cell and/or in inhibiting phase separation mediated by the IDR amino acid sequence in the solution or in the cultured host cell, preferably wherein the test agent interacts with the IDR amino acid sequence.

Any of the processes described herein for performing an *in vitro*, *in vivo* or *ex vivo* biochemical reaction may exclude a process for cloning a human being.

Any of the processes described herein for performing an *in vitro*, *in vivo* or *ex vivo* biochemical reaction may exclude a process for modifying the germ line genetic identity of a human being.

Any of the processes described herein for performing an *in vitro*, *in vivo* or *ex vivo* biochemical reaction may exclude a process involving the use of a human embryo, or the use of a totipotent human cell.

Any host cell described herein may exclude a human embryo, or a totipotent human cell, or a human germ line cell.

Whilst embracing *in vivo* use in some aspects, the invention embraces the exclusion of *in vivo* use in some aspects. Accordingly, any of the processes, uses or

methods etc. described herein for performing a biochemical reaction in an aqueous reaction system may exclude an *in vivo* aqueous reaction system.

Whilst embracing *ex vivo* use in some aspects, the invention embraces the exclusion of *ex vivo* use in some aspects. Accordingly, any of the processes, uses or
5 methods etc. described herein for performing a biochemical reaction in an aqueous reaction system may exclude an *ex vivo* aqueous reaction system.

In any of the methods, processes, uses or IDR reagents described herein, the efficiency of the reaction in the system may be increased by the IDR-macromolecule or the IDR-polypeptide compared to the efficiency of the reaction in the system following
10 the introduction of the at least one macromolecule or polypeptide under the same reaction conditions except that the at least one macromolecule or polypeptide does not comprise one or more functional intrinsically disordered regions (IDRs).

In any of the methods, processes, uses or IDR reagents described herein involving at least one macromolecule or at least one polypeptide tagged with an amino
15 acid sequence comprising or consisting of one or more functional intrinsically disordered regions (IDRs) (IDR-tagged macromolecule or IDR-tagged polypeptide) the efficiency of the reaction in the system may be increased by the IDR-tagged macromolecule or IDR-tagged polypeptide compared to the efficiency of the reaction in the system following the introduction of the at least one macromolecule or
20 polypeptide under the same reaction conditions except that the at least one macromolecule or polypeptide has not been tagged with the amino acid sequence comprising or consisting of the one or more functional IDRs.

Accordingly, whether an IDR-macromolecule or IDR-polypeptide, or an IDR-tagged macromolecule or IDR-tagged polypeptide, is capable of increasing the
25 efficiency of the reaction in the system can be established very simply by comparing the reaction efficiency of the macromolecule or polypeptide with or without the one or more functional IDRs. The skilled person is able to carry out simple comparative tests to establish the relevant functional capability. Exemplary test assays are described further herein.

30 Similarly, whether an IDR-macromolecule or IDR-polypeptide, or an IDR-tagged macromolecule or IDR-tagged polypeptide, is capable of causing molecules

necessary for the performance of the reaction to co-localise with the IDR-macromolecule or the IDR-polypeptide, or with the IDR-tagged macromolecule or IDR-tagged polypeptide, within the plurality of phase-separated aqueous compartments, or to further stimulate or enhance co-localisation of molecules necessary for the performance of the reaction within the plurality of phase-separated aqueous compartments, thereby increasing the efficiency of the biochemical reaction in the system, can also be established very simply by comparing co-localisation with or without the one or more functional IDRs. Again, the skilled person is able to carry out simple comparative tests to establish the relevant functional capability. Exemplary test assays are described further herein.

Similarly whether providing multivalent metal ions to an IDR-macromolecule or an IDR-polypeptide, or to an IDR-tagged macromolecule or IDR-tagged polypeptide thereby further stimulating or enhancing liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments and thereby further increasing the efficiency of the biochemical reaction in the system can also be established very simply by comparing liquid-liquid demixing with or without providing multivalent metal ions. Again, the skilled person is able to carry out simple comparative tests to establish the relevant functional capability. Exemplary test assays are described further herein.

Similarly whether providing ATP to an IDR-macromolecule or an IDR-polypeptide, or to an IDR-tagged macromolecule or IDR-tagged polypeptide may further simulate or enhance liquid-liquid demixing and the formation of the plurality of phase-separated aqueous compartments and thereby further increase the efficiency of the biochemical reaction in the system can also be established very simply by comparing liquid-liquid demixing with or without providing ATP. Whether providing ATP to an IDR-macromolecule or an IDR-polypeptide, or to an IDR-tagged macromolecule or IDR-tagged polypeptide to further stimulate or enhance co-localisation of molecules necessary for the performance of the reaction within the plurality of phase-separated aqueous compartments, thereby increasing the efficiency of the biochemical reaction in the system, can also be established very simply by comparing co-localisation with or without providing ATP. Again, the skilled person is

able to carry out simple comparative tests to establish the relevant functional capability. Exemplary test assays are described further herein.

Assays are described herein to establish the capability to cause liquid-liquid demixing by reference to the capability to cause the formation of phase-separated aqueous particles (see *e.g.* “phase separation assay method” as described herein). The same assay can be used to establish the capability to cause co-localisation of molecules necessary for the performance of the reaction within the phase-separated aqueous compartments (particles). Assays are described herein to establish the capability to increase the efficiency of a reaction by reference to the capability to increase the efficiency of an RPA method (see *e.g.* “RPA assay method” as described herein). Such an assay can be used to evaluate the capability of an amino acid sequence consisting of or comprising one or more functional intrinsically disordered regions (IDRs) to increase the efficiency of a reaction, and/or to evaluate the capability of divalent metal ions to further increase the efficiency of a reaction and/or to evaluate the capability of ATP to further increase the efficiency of a reaction.

Using simple assays as described herein the skilled person is able to determine an increase in the efficiency of a reaction of 5% or more, the increase in the efficiency of the reaction may be 10% or more, 15% or more, 20% or more, 25% or more, 30% or more, 35% or more, 40% or more, 45% or more, 50% or more, 55% or more, 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more or 100% or more.

Crowding agent

A crowding agent is typically a high molecular weight macromolecule, such as a protein or a synthetic block polymer. A crowding agent is considered to be essentially biochemically inert, *i.e.* it does not contribute to a specific interaction or catalysis.

It is widely postulated that a crowding agent exerts an influence on a biological/biochemical system, be it an *in vitro* or an *in vivo* system, through the effects of its physical occupation of volume in a solution thus causing steric obstruction and a reduction in available open solvent space. By this excluded volume mechanism

crowding agents appear to increase the effective concentration of other macromolecules, with particular influence on altering dissociation constants and favouring the association of interacting macromolecules such as multiple proteins which come together into specific organised complexes. The size of the crowding effect depends in particular on
5 the molecular mass of the molecules involved, being generally much stronger with larger molecules. Thus, as a general rule, macromolecular crowding is an effect exerted by large molecules on the properties of other large molecules.

Furthermore, crowding agents have been widely described as being capable of facilitating the formation of biological/biochemical systems with preferred phases in
10 which reactants segregate themselves into micron-sized phase-separated particles. This effect arises substantially from the effects of volume exclusion on the dissociation constant of macromolecular complexes which become relatively “confined” due to the inability to readily diffuse into the largely volume-occupied bulk solvent. Additionally and/or alternatively, some crowding agents such as block chain polymers like
15 polyethylene glycol may exhibit kosmotropic character which leads them to exert an overall alteration in the structure of bulk water, typically lowering water density. Such changes in bulk solvent characteristics may also exert complex effects on other macromolecules and their assemblies whose surfaces must interact with water. This in turn may also promote segregation of those other macromolecules into an alternate
20 phase, significantly enriched in the biological components and concomitantly depleted in the crowding agent which principally occupies the bulk solvent phase.

In either scenario, either by simple volume occupation or solvent modification, the effects individually or in combination of crowding agents in stimulating the condensation of macromolecules into phase-distinct condensates appears to operate by a
25 “repulsive” rather than by an “attractive” mechanism from the perspective of the condensate components. In other words, from the perspective of the components that are highly enriched in the condensate the crowding agent acts to create a bulk phase environment that cannot be readily penetrated by diffusion, and/or whose bulk solvent character is modified in a manner that it presents a net enthalpic disadvantage to
30 disperse into. It is in this manner that the effects of high concentrations of crowding agents, typically greater than 1% w/v, are referred to herein in stimulating phase

separation by functioning through an “obstructive” or “repulsive” mechanism insofar as the phenomenon arises due to the condensate components being unable to readily disperse as they would in the absence of the crowding agent. However, at the same time, in view of its generally inert properties, the crowding agent has little or no direct
5 debilitating effects on other specific molecules in the system, for example because the crowding agent does not significantly interact with or exert effects on specific molecular side-chains in a direct fashion.

In standard RPA reactions polyethylene glycols (PEGs) can have profound effects on recombination/DNA synthesis. PEGs can influence the number of multiple
10 invasion/extension cycles that occur, for example, when RecA is combined with Gp32. PEGs can stimulate amplification reactions configured in several different ways. PEGs and other similar crowding agents may affect the cooperativity of Gp32 and recombinases, they may affect polymerase processivity and they may affect the hybridisation rate and behaviour of oligonucleotides in solution. The chain length of
15 the polyethylene glycol can influence results. PEGs may also increase the stability of recombinase-loaded filaments and the increased persistence may increase RPA efficacy.

To exert its effects in an *in vitro* biochemical reaction milieu an added crowding agent is typically present at concentrations at which steric exclusion/confinement effects are predicted to occur, typically above approximately 1% by volume of the reaction, or
20 by weight of the reaction.

In standard RPA reactions a crowding agent is present at concentrations of approximately 1% to 12% by volume of the reaction or by weight of the reaction.

The terms “macromolecular crowding agent” or more simply “crowding agent” are very well recognised and art-understood terms. This is apparent from the literature
25 in which the terms are used extensively. For example Kuznetsova, I., M. *et al.*, (What Macromolecular Crowding Can Do to a Protein, 2014, Int. J. Mol. Sci., 15, pp 23090-23140) provide a review which purports to cover more than 320 papers and is suggested to represent one of the most comprehensive compendia of the current knowledge in the field. The term “crowding agent” is used extensively throughout the text highlighting
30 its ubiquitous use (see also Mixed Macromolecular Crowding: A Protein and Solvent Perspective, Biswas, S. *et al.*, 2018, ACS Omega, 3(4), pp4316-4330 and Common

Crowding Agents Have Only a Small Effect on Protein-Protein Interactions, Phillip Y. *et al.*, 2009, Biophysical Journal, 97 pp875-885 875).

A compound or macromolecule can be identified as a crowding agent by means known in the art. For example a crowding agent can be identified as such via its
 5 experimentally determined and calculated hydrodynamic radius (Kuznetsova *et al.*, *supra*). A crowding agent can be identified as such via sol-gel glass encapsulation analysis (Kuznetsova *et al.*, *supra*).

The following compounds are examples of known crowding agents. A synthetic block polymer, a polyethylene glycol (PEG), PEG 1450, PEG 2050, PEG3000, PEG
 10 4600, PEG 6000, PEG 8000, PEG 10000, PEG 20000, PEG 35000, PEG compound molecular weight 15,000 to 20,000 (also known as Carbowax 20M), a dextran, Dextran 6, Dextran 40, Dextran 70, Dextran 670, Dextran sulfate 10, Dextran sulfate 500, a ficoll, Ficoll 70, Ficoll 400, Poly(sodium 4-styrene sulfonate) (PSS), Bovine pancreatic trypsin inhibitor (BPTI), Ribonuclease A, Lysozyme, β -Lactoglobulin, Hemoglobin,
 15 Bovine serum albumin (BSA).

In any one of the methods, processes and uses of the present invention, including in any one of the RPA methods, processes and uses of the present invention, the methods, processes and uses may be performed in the absence of a crowding agent.

In any one of the methods, processes and uses of the present invention, including
 20 in any one of the RPA methods, processes and uses of the present invention, the methods, processes and uses may be performed in the presence of a crowding agent.

In any one of the methods, processes and uses of the present invention, including in any one of the RPA methods, processes and uses of the present invention, the methods, processes and uses may be performed in the presence of a crowding agent,
 25 wherein the crowding agent is provided at a concentration which provides an enhancement of the increase in the efficiency of the biochemical reaction which is provided by the IDR-macromolecule or the IDR-polypeptide.

In any one of the methods, processes and uses of the present invention, including in RPA methods, processes and uses of the present invention, the methods, processes
 30 and uses may be performed in the presence of a crowding agent, wherein the crowding agent is provided at a concentration which provides an additive effect on the efficiency

of the biochemical reaction which is provided by the IDR-macromolecule or the IDR-polypeptide.

In any one of the methods, processes and uses of the present invention, including in any one of the RPA methods, processes and uses of the present invention, the methods, processes and uses may be performed in the presence of a crowding agent, wherein the crowding agent is provided at a concentration which provides a synergistic effect on the efficiency of the biochemical reaction which is provided by the IDR-macromolecule or the IDR-polypeptide.

In any one of the methods, processes and uses of the present invention, including in any one of the RPA methods, processes and uses of the present invention, the methods, processes and uses may be performed in the presence of a crowding agent, wherein the introduction of the IDR-macromolecule or the IDR-polypeptide into the biochemical reaction system lowers the concentration of crowding agent which would be required to achieve the same increase in the efficiency of the biochemical reaction in the absence of the introduction of the IDR-macromolecule or the IDR-polypeptide into the biochemical reaction system.

In any one of the above-described methods, processes and uses which may be performed in the presence of a crowding agent, the crowding agent may be present at a concentration below that at which its normal biological effects (steric exclusion/confinement effects) occur.

In any one of the above-described methods, processes and uses which may be performed in the presence of a crowding agent, the crowding agent may be present at concentrations below approximately 3% by volume of the reaction or by weight of the reaction, approximately 2% by volume of the reaction or by weight of the reaction, approximately 1% by volume of the reaction or by weight of the reaction or approximately 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2% or 0.1% by volume of the reaction or by weight of the reaction.

If used in any one of the methods, processes and uses of the invention, including RPA reaction methods, any suitable crowding agent may be used. Examples of suitable crowding agents are provided herein.

Macromolecule or polypeptide comprising an intrinsically disordered region (IDR)

The methods, processes and reagents of the present invention involve “IDR-macromolecules”, including “IDR-tagged macromolecules”, as described herein. The methods, processes and reagents of the present invention involve “IDR-polypeptides”, including “IDR-tagged polypeptides”, as described herein. Any such IDR-macromolecules, IDR-tagged macromolecules, IDR-polypeptides or IDR-tagged polypeptides may be referred to interchangeably herein as IDR reagents or IDR-based reagents.

10 An IDR-macromolecule or IDR-polypeptide or an IDR-tagged macromolecule or IDR-tagged polypeptide as used herein is any macromolecule, or polypeptide or protein, which comprises one or more functional intrinsically disordered regions (IDRs).

15 An IDR-macromolecule or IDR-polypeptide or an IDR-tagged macromolecule or IDR-tagged polypeptide as used herein is any macromolecule, or polypeptide or protein, which comprises an amino acid sequence consisting of or comprising one or more functional intrinsically disordered regions (IDRs).

20 Accordingly, an IDR-macromolecule or IDR-polypeptide, as referred to herein, may thus refer to: a macromolecule or polypeptide which comprises an amino acid sequence consisting of one or more functional intrinsically disordered regions; or a macromolecule or polypeptide which comprises an amino acid sequence comprising one or more functional intrinsically disordered regions.

In addition, an IDR-macromolecule comprising one or more functional intrinsically disordered regions (IDRs), as referred to herein may be a macromolecule of interest tagged with an amino acid sequence consisting of or comprising one or more functional intrinsically disordered regions (IDRs). Such an IDR-tagged macromolecule is also an IDR reagent as defined herein. An IDR-tagged polypeptide comprising one or more functional intrinsically disordered regions (IDRs), as referred to herein, may be a polypeptide of interest tagged with an amino acid sequence consisting of or comprising one or more functional intrinsically disordered regions (IDRs). Such an IDR-tagged polypeptide is also an IDR reagent as defined herein.

25

30

An IDR-tagged macromolecule or IDR-tagged polypeptide, as used herein, is any macromolecule, or polypeptide or protein, which is “tagged” with an amino acid sequence consisting of or comprising one or more functional intrinsically disordered regions (IDRs).

5 Accordingly, an IDR-tagged macromolecule or IDR-tagged polypeptide, as referred to herein, may thus refer to: a macromolecule or polypeptide which is tagged with an amino acid sequence consisting of one or more functional intrinsically disordered regions; or a macromolecule or polypeptide which is tagged with an amino acid sequence comprising one or more functional intrinsically disordered regions.

10 A tagged amino acid sequence consisting of or comprising one or more functional intrinsically disordered regions (IDRs) is not found naturally or ordinarily in the macromolecule or polypeptide or protein to which it is tagged at the tagged position. Accordingly, a tagged amino acid sequence consisting of or comprising one or more functional intrinsically disordered regions (IDRs) can be considered to be an exogenous
15 amino acid sequence compared to the macromolecule or polypeptide or protein to which it is tagged. The tagged macromolecule or polypeptide or protein can therefore be considered to be a non-naturally occurring, artificial or genetically engineered macromolecule, polypeptide or protein.

 Mechanisms by which amino acid sequences may be “tagged” to polypeptides
20 and other macromolecules are explained further herein.

Any one or more functional intrinsically disordered regions (IDRs) may be tagged to a macromolecule or polypeptide or protein, including any one or more of the specific IDR amino acid tag sequences disclosed herein, or any one or more functional variants, analogs, homologs or derivatives thereof.

25 For use in the present invention both an intrinsically disordered region polypeptide sequence and a domain thereof should be “functional”. The term “functional” means that any IDR amino acid sequence must possess one of the functional properties outlined further herein.

 The term “intrinsically disordered region” is an art-understood term used
30 commonly in the technical field. For a comprehensive review see: Classification of

Intrinsically Disordered Regions and Proteins, van der Lee et al., 2014, Chem. Rev. 114, pp 6589–6631.

The invention provides, *inter alia*, a process of performing a biochemical reaction in an aqueous *in vitro* reaction system, wherein the biochemical reaction is dependent on the function of at least one reaction macromolecule, optionally at least one reaction polypeptide, the process comprising: introducing at least one IDR-macromolecule into the *in vitro* reaction system under conditions suitable for performing the reaction, wherein the at least one IDR-macromolecule comprises one or more functional intrinsically disordered regions (IDRs), wherein upon introduction of the at least one IDR-macromolecule into the *in vitro* reaction system the efficiency of the biochemical reaction is increased by the at least one IDR-macromolecule. The efficiency of the biochemical reaction is increased by the one or more functional IDRs of the IDR-macromolecule. In any such process the at least one IDR-macromolecule may be at least one IDR-polypeptide. In any such process, the IDR-macromolecule or IDR-polypeptide comprising one or more functional intrinsically disordered regions (IDRs) may not be the “reaction macromolecule” or “reaction polypeptide” upon whose function the biochemical reaction depends. Thus in any such process the IDR-macromolecule or IDR-polypeptide may have no intrinsic biochemical role in the biochemical reaction *per se*. Nevertheless, its introduction into the reaction system leads to an increase in the efficiency of the biochemical reaction.

The process of performing a biochemical reaction in an *in vitro* reaction system may be a process wherein the biochemical reaction is dependent on the function of the at least one IDR-macromolecule, wherein upon its introduction into the *in vitro* reaction system the at least one IDR-macromolecule performs its reaction function in the biochemical reaction and increases the efficiency of the reaction. In any such process the at least one IDR-macromolecule may be at least one IDR-polypeptide. In any such process, the IDR-macromolecule or IDR-polypeptide does have an intrinsic biochemical role in the biochemical reaction *per se*. Thus the at least one IDR-macromolecule or IDR-polypeptide comprising one or more functional intrinsically disordered regions (IDRs) is the “reaction macromolecule” or “reaction polypeptide” upon whose function the biochemical reaction depends.

The at least one IDR-macromolecule or IDR-polypeptide comprises an amino acid sequence comprising or consisting of one or more functional intrinsically disordered regions. The IDR-macromolecule or IDR-polypeptide is introduced into the biochemical reaction system under conditions suitable for performing the biochemical reaction. Because of the presence of the one or more functional intrinsically disordered regions, the IDR-macromolecule or IDR-polypeptide increases the efficiency of the reaction.

By increasing the efficiency of the reaction it is meant that the efficiency of the reaction is improved compared to the efficiency of the reaction which would be observed if the IDR-macromolecule or IDR-polypeptide was provided without an amino acid sequence comprising or consisting of the one or more functional intrinsically disordered regions. Such an improvement can readily be established by comparative testing of the reaction macromolecule or polypeptide with and without the IDR amino acid sequence.

In any one of the methods of the invention the one or more functional intrinsically disordered regions of the IDR-macromolecule or IDR-polypeptide promotes/causes liquid-liquid demixing in the reaction system leading to phase separation. The functional capability of promoting liquid-liquid demixing leading to phase separation in a reaction system can readily be established, for example by performing a phase separation assay as described herein. Such liquid-liquid demixing promotes phase separation and this may lead to the formation of phase separated compartments in the reaction system, such as particles which are detectable, *e.g.* under microscopic observation, as detailed further herein.

An IDR-macromolecule or IDR-polypeptide may or may not possess catalytic activity. For example, an IDR-polypeptide may have catalytic activity, such as a polymerase enzyme used in a recombinase polymerase amplification reaction, as described further herein. An IDR-polypeptide may not have catalytic activity, such as a single strand stabilizing agent used in a recombinase polymerase amplification reaction, *e.g.* Gp32 as described further herein.

As discussed further herein, whether an IDR-macromolecule or IDR-polypeptide has catalytic activity or not, the IDR-macromolecule or IDR-polypeptide may possess a

function which is required for or influences the biochemical reaction, such that in the absence of the IDR-macromolecule or IDR-polypeptide in the biochemical reaction system the biochemical reaction cannot proceed or proceeds with reduced efficiency. Alternatively, as discussed further herein, the IDR-macromolecule or IDR-polypeptide may not possess any function which is required for or which influences the biochemical reaction *per se*. Nevertheless, because of the IDR amino acid sequences, the introduction of the IDR-macromolecule or IDR-polypeptide into the biochemical reaction system leads to an increase in the efficiency of the biochemical reaction compared to the efficiency observed in the absence of the IDR-macromolecule or IDR-polypeptide, or in the presence of the same macromolecule or polypeptide without the IDR amino acid sequences.

Structural properties of an IDR polypeptide

The presence of IDRs in amino acid sequences may readily be determined by structural analysis. A large number of bioinformatics-based platforms are available for the prediction of the presence of IDRs within polypeptides and proteins. These include ELM, MiniMotif, SLiMPrints, phylo-HMM, DiliMot, SLiMFinder, Phospho.ELM, PhosphoSite, PHOSIDA, ScanSite, NetPhorest, NetworKIN, PhosphoNET, IDEAL, MoRFpred, ANCHOR, Pfam, FFPred, DisProt, D²P², and MetaDisorder. Any of these methods may be used to identify IDR amino acid sequences. If necessary, such IDR amino acid sequences can be tested to assess their functional properties, as described further herein.

A preferred bioinformatics-based platform for IDR amino acid sequence identification is the MetaDisorder software program (MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. Kozlowski, L. P., et al., BMC Bioinformatics, 2012, 13(1): 111).

The MetaDisorder program is available freely online (<http://genesilico.pl/metadisorder/>). Using this program an amino acid sequence of interest is simply pasted into an internet browser window and the program is initiated.

As the online documentation explains, any amino acid region which scores >0.5 in the software package is considered to comprise an intrinsically disordered region.

Using the MetaDisorder software platform the inventors have identified a number of amino acid sequences which comprise one or more intrinsically disordered regions. These are set out in Table 1.

Accordingly, in any one of the methods, processes and uses, or in any one of the non-naturally occurring IDR-macromolecules, IDR-fusion macromolecules or isolated nucleic acid molecules encoding the same, recombinant polynucleotide expression vectors or host cells, the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide may be characterised as a sequence of amino acids which scores greater than 0.5 when analysed by the algorithm MetaDisorder. The sequence of amino acids may be a sequence of amino acids which scores greater than 0.5 when analysed by the algorithm MetaDisorder according to the methods of Kozlowski, L. P., et al., BMC Bioinformatics, 2012, 13(1): 111.

The invention provides and relates to preferred IDR amino acid sequences which comprise or consist of the amino acid sequence of any one of SEQ ID NOs 1 to 43 (Table 1) and variants thereof. In all cases a variant of an amino acid sequence of any one of SEQ ID NOs 1 to 43 is a functional variant which retains the IDR functional properties, as described further herein.

Furthermore, as described further herein, the IDR-macromolecule or IDR-polypeptide may comprise or consist of a macromolecule or polypeptide tagged with an amino acid sequence which comprises or consists of the amino acid sequence of any one of SEQ ID NOs 1 to 43, or which comprises or consists of a functional variant amino acid sequence of SEQ ID NOs 1 to 43.

The functional variant may have at least 80% sequence identity compared to the herein-described IDR amino acid sequence (Table 1). The functional variant may have at least 81% sequence identity compared to the herein-described IDR amino acid sequence (Table 1), or 82% sequence identity, or 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity.

For the purpose of this invention, in order to determine the percent identity *e.g.* between an amino acid sequence of any one of SEQ ID NOs 1 to 43 and a functional

variant of an amino acid sequence of any one of SEQ ID NOs 1 to 43 the two respective amino acid sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in a first sequence for optimal alignment with a second sequence). The nucleotide residues at nucleotide positions are then compared. When a position in the first sequence is occupied by the same nucleotide residue as the corresponding position in the second sequence, then the nucleotides are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = number of identical positions /total number of positions in the reference sequence x 100).

Typically the sequence comparison is carried out over the full length of the reference sequence. For example, if the skilled person wished to determine whether a given (“variant”) sequence is 80% identical to SEQ ID NO. 2, SEQ ID NO. 2 would be the reference sequence. For example, to assess whether a variant sequence is at least 80% identical to SEQ ID NO. 2 (an example of a reference sequence), the skilled person would carry out an alignment over the length of SEQ ID NO. 2, and identify how many positions in the test sequence were identical to those of SEQ ID NO. 2. If at least 80% of the positions are identical, the test sequence is at least 80% identical to SEQ ID NO. 2. If the sequence is shorter than SEQ ID NO. 2, the gaps or missing positions should be considered to be non-identical positions.

The skilled person is aware of different computer programs that are available to determine the homology or identity between two sequences. For instance, a comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. The percent identity between two amino acid or nucleic acid sequences may be determined, for example, using the Needleman and Wunsch (1970) algorithm which has been incorporated into the GAP program in the Accelrys GCG software package (available at <http://www.accelrys.com/products/gcg/>), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6.

A functional variant of an amino acid sequence of any one of SEQ ID NOs 1 to 43 may be an amino acid sequence that differs by having a number of amino acids less compared to the amino acid sequence of any one of SEQ ID NOs 1 to 43 respectively

- (*i.e.* the functional variant is shorter), or that differs by having a number of amino acids more compared to the amino acid sequence of any one of SEQ ID NOs 1 to 43 respectively (*i.e.* the functional variant is longer). The functional variant may thus contain one or more amino acid deletions and/or one or more insertions compared to the
- 5 reference amino acid sequence. The number of amino acids in the functional variant amino acid sequence by which the variant differs from the reference sequence can be 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, 15 or more, 16 or more, 17 or more, 18 or more, 19 or more or 20 or more.
- 10 A functional variant of an amino acid sequence of any one of SEQ ID NOs 1 to 43 may for example comprise conservative amino acid substitutions of amino acid residues indicated in the sequences listed in Table 1. Conservative substitutions may be made, for example according to the table below which describes a generally accepted grouping of amino acids. The functional variant may thus contain conservative amino
- 15 acid substitutions compared to the reference amino acid sequence. The number of amino acids in the functional variant amino acid sequence which are conservative amino acid substitutions compared to the reference sequence can be 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, 15 or more, 16 or more, 17 or
- 20 more, 18 or more, 19 or more or 20 or more.

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

Whether a given variant retains the IDR functional properties may readily be established, for example by methods described further herein.

Table 1 - amino acid sequences comprising intrinsically disordered regions

[illegible]

										deletion at the C-terminus of fib (fib-1)				
fib-4 (clipper3)	N/A	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	GlnValGlnMetGlnLeuArgGlnValPheSerGlnAspGlnGlnValLeuGlnGluArgMetArgTyrHisGluLeuGlnGlnGlnGlnGlnGlnTyrHisGluThrLysAspMetValGlySerTyrThrGlnAsnSerAsnSerAlaIleProLeuPheGlyAsnAsnSerAspThrThrAsnGlnGlnAsnSer	7					
ENT1	Ent2p	Uniprot: C7GIT6	Saccharomyces cerevisiae	530-588	59		N/A	GlnAsnProThrGlyIleSerTyrSerGlnProGlnGlnGlnGlnGlnProGlnGlnGlnProGlnTyrMetGlnAsnPheGlnGlnGlnGlnProGlnTyrAlaGlnAsnPheGlnGlnGlnProGlnTyrThrGlnAsnTyrGlnGlnGlnProGlnTyrIleGlnProHisGln	8					
HRP1	Nuclear polyadenylated RNA-binding protein 4	Uniprot: Q99383	Saccharomyces cerevisiae	331-388	58		N/A	GlyGlyAsnAsnGlyGlyAsnAsnMetAsnArgArgGlyGlyAsnPheGlyAsnGlnGlyAspPheAsnGlnMetTyrGlnAsnProMetMetGlyGlyTyrAsnProMetMetAsnProGlnAlaMetThrAspTyrTyrGlnLysMetGlnGluTyrTyrGlnGlnMetGln	9					

HRP2	Nuclear polyadenylated RNA-binding protein 4	Uniprot: Q99383	Saccharomyces cerevisiae	331-388	116	double of HRP1 tag	GlyGlyAsnAsnGlyGlyAsnAsnMetAsnArgArgGlyGlyAsnPh eGlyAsnGlnGlyAspPheAsnGlnMetTyrGlnAsnProMetMe tGlyGlyTyrAsnProMetMetAsnProGlnAlaMetThrAspTyr TyrGlnLysMetGlnGluTyrTyrGlnGlnMetGlnGlyGlyAsnA snGlyGlyAsnAsnMetAsnArgArgGlyGlyAsnPheGlyAsnGl nGlyAspPheAsnGlnMetTyrGlnAsnProMetMetGlyGlyTyr AsnProMetMetAsnProGlnAlaMetThrAspTyrTyrGlnLysM etGlnGluTyrTyrGlnGlnMetGln	10
DDX	Probable ATP-dependent RNA helicase DDX4	Uniprot: Q9NQI0	Homo sapiens	4-240	240	N/A	MetGlyAspGluAspTrpGluAlaGluIleAsnProHisMetSerS erTyrValProIlePheGluLysAspArgTyrSerSerGlyGluAsnGl yAspAsnPheAsnArgThrProAlaSerSerSerGluMetAspAsp GlyProSerArgArgAspHisPheMetLysSerGlyPheAlaSerG lyArgAsnPheGlyAsnArgAspAlaGlyGluCysAsnLysArgAs pAsnThrSerThrMetGlyGlyPheGlyValGlyLysSerPheGly AsnArgGlyPheSerAsnSerArgPheGluAspGlyAspSerSerG lyPheTrpArgGluSerSerAsnAspCysGluAspAsnProThrAr gAsnArgGlyPheSerLysArgGlyGlyTyrArgAspGlyAsnAsn SerGluAlaSerGlyProTyrArgArgGlyGlyArgGlySerPhea rgGlyCysArgGlyGlyPheGlyLeuGlySerProAsnAsnAspLe uAspProAspGluCysMetGlnArgThrGlyGlyLeuPheGlySer ArgArgProValLeuSerGlyThrGlyAsnGlyAspThrSerGlnS erArgSerGlySerGlySerGluArgGlyGlyTyrLysGlyLeuAs nGluGluValIleThrGlySerGlyLysAsnSerTrpLysSerGlu AlaGluGlyGlyGluSerSerAspThrGln	11
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1	Uniprot: Q0VAC0	Homo sapiens	187-320	134	N/A	AlaSerAlaSerSerSerGlnArgGlyArgSerGlySerGlyAsnP heGlyGlyGlyArgGlyGlyGlyPheGlyGlyAsnAspAsnPheGl yArgGlyGlyAsnPheSerGlyArgGlyGlyPheGlyGlySerArg GlyGlyGlyGlyTyrGlySerGlyAspGlyTyrAsnGlyPheG lyAsnAspGlySerAsnPheGlyGlyGlySerTyrAsnAspPh eGlyAsnTyrAsnAsnGlnSerSerAsnPheGlyProMetLysGly GlyAsnPheGlyGlyArgSerSerGlyProTyrGlyGlyGlyG lnTyrPheAlaLysProGlnAsnGlnGlyGlyTyrGlyValSerSe rSerSerSerSerTyrGlySerGlyArgArgPhe	12

PoliICTD	DNA-directed RNA polymerase II subunit RPB1	UniProt: P04050	Saccharomyces cerevisiae	1521-1724	204	N/A	SerProPheGlyAlaTyrGlyGluAlaProThrSerProGlyPheGlyValSerSerProGlyPheSerProThrSerProThrTyrSerPr oThrSerProAlaTyrSerProThrSerProSerTyrSerProThr SerProSerTyrSerProThrSerProSerTyrSerProThrSerP roSerTyrSerProThrSerProSerTyrSerProThrSerProSe rTyrSerProThrSerProSerTyrSerProThrSerProSerTyr SerProThrSerProSerTyrSerProThrSerProSerTyrSerP roThrSerProSerTyrSerProThrSerProSerTyrSerProTh rSerProSerTyrSerProThrSerProSerTyrSerProThrSer ProAlaTyrSerProThrSerProSerTyrSerProThrSerPros erTyrSerProThrSerProSerTyrSerProThrSerProSerTy rSerProThrSerProAsnTyrSerProThrSerProSerTyrSer ProThrSerProGlyTyrSerProGlySerProAlaTyrSerProL ysGlnAspGluGln
shortPolCTD	DNA-directed RNA polymerase II subunit RPB1	UniProt: P04050	Saccharomyces cerevisiae	1636-1677	42	modifications: S1653G S1667T S1674T	ThrSerProSerTyrSerProThrSerProSerProSerTyrSerProThrSerProGlyTyrSerProThrSerProAlaTyrSerProThrSerPr oThrTyrSerProThrSerProThrTyrSerPro
mimic1	N/A	N/A	N/A	N/A	49	self-designed: seven repeats of YDPTSPS (SEQ ID NO: 129) motifs mimicking the c-terminus of RNA polymerase II	TyrAspProThrSerProSerTyrAspProThrSerProSerTyrA spProThrSerProSerTyrAspProThrSerProSerTyrAspPr oThrSerProSerTyrAspProThrSerProSerTyrAspProThr SerProSer

64

[illegible]

[illegible]

[illegible]

							of EIVEAEVDE LINSKVEKF KSPESKSKS AADLETDL E QLSDMEEFN SEQ ID NO: 132) and change amino acids D/E to R/K			ysPheLysSerProLysSerLysSerLysSerAlaAlaLysLeuLys sThrLysLeuLysGlnLeuSerLysMetLysGluPheAsn	
supernegative	RecA-like recombination protein	GenBank: AYP69747.1	Escherichia phage vB_EcoM_Dal Ca	348-392	90		self- designed: double sequence of EIVEAEVDE LINSKVEKF KSPESKSKS AADLETDL E QLSDMEEFN SEQ ID NO: 132) and change amino acids R/K to D/E			GluIleValGluAlaGluValValAspGluLeuIleAsnSerGluValG luGluPheAspSerProGluSerAspSerAspGluAlaAlaAspLe uGluThrAspLeuGluGlnLeuSerAspMetGluGluPheAsnGlu IleValGluAlaGluValValAspGluLeuIleAsnSerGluValGluA spPheAspSerProGluSerAspSerAspGluAlaAlaAspLeuGl uThrAspLeuGluGlnLeuSerAspMetGluGluPheAsn	41
DoubleX	RecA-like recombination protein	GenBank: AYP69747.1	Escherichia phage vB_EcoM_Dal Ca	348-392	90		self- designed: a linker of FEF and double sequence of EIVEAEVDE LINSKVEKF KSPESKSKS AADLETDL E QLSDMEEFN			GluIleValGluAlaGluValValAspGluLeuIleAsnSerLysValG luLysPheLysSerProGluSerLysSerLysSerAlaAlaAspLe uGluThrAspLeuGluGlnLeuSerAspMetGluGluPheAsnGlu IleValGluAlaGluValValAspGluLeuIleAsnSerLysValGluL ysPheLysSerProGluSerLysSerLysSerAlaAlaAspLeuGl uThrAspLeuGluGlnLeuSerAspMetGluGluPheAsn	42

TripleXtail	RecA-like recombination protein	GenBank: AYP69747.1	Escherichia phage vB_EcoM_Dal Ca	348-392	135	SEQ ID NO: 132)	self-designed: a linker of DDVASEF (SEQ ID NO: 133) and triple sequence of EIVEAEVDE LINSKVEKF KSPESKSKS AADLETDL EQLSDMEEFN (SEQ ID NO: 132) GluIleValGluAlaGluValAspGluLeuIleAsnSerLysValGluLysPheLysSerProGluSerLysSerLysSerAlaAlaAspLeuGluThrAspLeuGluGlnLeuSerAspMetGluGluPheAsnGluIleValGluAlaGluValAspGluLeuIleAsnSerLysValGluLysPheLysSerProGluSerLysSerLysSerAlaAlaAspLeuGluThrAspLeuGluGlnLeuSerAspMetGluGluPheAsnGluIleValGluAlaGluValAspGluLeuIleAsnSerLysValGluLysPheLysSerProGluSerLysSerAlaAlaAspLeuGluThrAspLeuGluGlnLeuSerAspMetGluGluPheAsn	43	
-------------	---------------------------------	---------------------	----------------------------------	---------	-----	-----------------	---	----	--

Like-for-like substitutions may be made, in the case of amino acids such as basic for basic, acidic for acidic, polar for polar, etc. Non-homologous substitution may also occur, *i.e.* from one class of residue to another, or alternatively involving the inclusion of unnatural amino acids such as ornithine, diaminobutyric acid ornithine, norleucine
5 ornithine, pyriylalanine, thienylalanine, naphthylalanine and phenylglycine.

The specific IDR amino acid sequences disclosed herein (see Table 1) can broadly be classified into four groups. Some IDR sequences can be classified into more than one group. The RGG/RG group includes IDR sequences which are FG/YG rich. This group includes fib, hnrpnA1, DDX, HRP1 and Sup. The Poly Q group includes
10 IDR sequences which are Q/N rich. This group includes PCF11, Ent-1, HRP1, Sup, His4, His8 and His10. The Poly P group includes sequences which are P rich. This group includes His4, His9 and His10. The Poly H group includes sequences which are H rich. This group includes His1-11. Some key features of the IDR amino acid
15 sequences are that they demonstrate cation–pi interactions and pi-pi interactions, and are capable of forming amide bridges and salt bridges. Key features and key inter/intra molecular interactions of preferred IDR amino acid sequences are presented in Tables 2 to 20 below.

Table 2

IDR	fib
Sequence	PGFSPRGGGGFGGRGGFGDRGGRGGRGGFGGGGRGRGGGFRGRGR (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)

Table 3

IDR	HNRPNA1
Sequence	ASASSSQ <u>RGR</u> SGSGN <u>FGG</u> <u>RGG</u> G <u>FG</u> GNDN <u>FGR</u> GGNFSG <u>RGG</u> <u>FG</u> GS RGGGGYGGSGDGYNGFGNDGSNFGGGGSYNDFGNYNQSSN <u>F</u> GP MKGGN <u>F</u> GGRSSGPY <u>G</u> GGGQYFAKPQNQGGY <u>G</u> VSSSSSSSYGSGRRF (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)

Table 4

IDR	DDX
Sequence	MGDEDWEAEINPHMSSYVPIFEKDRYSGENGDNFNRTPASSEMDD GPSRRDHFMKSGFASGRN <u>F</u> GNRDAGECNKRDNTSTMGG <u>F</u> GVGK <u>S</u> <u>F</u> <u>G</u> N <u>R</u> GFSNSRFEDGDSSGFWRESSNDCEDNPTRN <u>R</u> GFSK <u>R</u> GGYRDGN NSEASGPYRR <u>G</u> GRGS <u>F</u> <u>R</u> G <u>C</u> RGG <u>F</u> GLGSPNNDLDPDECMQRTGG <u>L</u> <u>F</u> G SRRPVLSGTGNGDTSQSRSGSGSER <u>G</u> GYKGLNEEVITGSGKNSWKS EAEGGESDQTQ (SEQ ID NO: 11)
Key feature	RGG/RG, FG
Key inter/intra	R/K-Y/F/H (cation-pi interaction) Y/F/H-Y/F/H (pi-pi interaction) R/K-D/E (salt bridge)

molecular interactions	Q/N-Q/N (amide interaction)
------------------------	-----------------------------

Table 5

IDR	HRP1
Sequence	GGNNGGNNMNR <u>RG</u> <u>GN</u> <u>FG</u> NQGD FNQMYQNPMMGGYNPMMNPQA MTD <u>YYQ</u> <u>KMQ</u> <u>EYYQ</u> <u>QMQ</u> (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)

Table 6

IDR	Sup
Sequence	MSDSNQGN NQQNYQQYSQNGNQQQGN NRYQGYQAYNAQAQPAG GYYQNYQGYSGYQQGGYQQYQ <u>YNPQGGYQQYNPQGGYQQYNPQ</u> <u>GGYQQQFNPQGG</u> <u>RG</u> NYKNFNYN NNLQGYQ (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)

Table 7

IDR	PolCTD
Sequence	<u>SPFGAYGEAPTSPGFGVSSPGFSPTSPTYSPTSPAYSPTSPSYSPTSPSY</u> <u>SPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPT</u> <u>SPSYSPTSPSYSPTSPSYSPTSPSYSPTSPSYSPTSPAYSPTSPSYSPTSPS</u> <u>YSPTSPSYSPTSPSYSPTPNYSPTSPSYSPTSPGYSPGSPAYSPKQDEQ</u> (SEQ ID NO: 13)
Key feature	repeats of YSPTSPS (SEQ ID NO: 124) and its variants (FSPTSPT (SEQ ID NO: 125), YSPTSP-A/N/G (SEQ ID NO: 126), YSPGSPA (SEQ ID NO: 127))
Key inter/intra molecular interactions	Q/N-Q/N(amide bridge) K-Y/F/H (cation-pi interaction) Y/F/H-Y/F/H (pi-pi interaction) K-D/E (salt bridge)

Table 8

IDR	PCF11
Sequence	<u>Q</u> <u>V</u> <u>Q</u> <u>M</u> <u>Q</u> <u>L</u> <u>R</u> <u>Q</u> <u>V</u> <u>F</u> <u>S</u> <u>Q</u> <u>D</u> <u>Q</u> <u>Q</u> <u>V</u> <u>L</u> <u>Q</u> <u>E</u> <u>R</u> <u>M</u> <u>R</u> <u>Y</u> <u>H</u> <u>E</u> <u>L</u> <u>Q</u> <u>Q</u> <u>Q</u> <u>Q</u> <u>Q</u> <u>Q</u> <u>Q</u> <u>Y</u> <u>H</u> <u>E</u> <u>T</u> <u>K</u> <u>D</u> <u>M</u> <u>V</u> <u>G</u> <u>S</u> Y <u>T</u> <u>Q</u> <u>N</u> <u>S</u> <u>N</u> <u>S</u> <u>A</u> <u>I</u> <u>P</u> <u>L</u> <u>F</u> <u>G</u> <u>N</u> <u>N</u> <u>S</u> <u>D</u> <u>T</u> <u>T</u> <u>N</u> <u>Q</u> <u>Q</u> <u>N</u> <u>S</u> (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)

Table 9

IDR	ENT1
Sequence	<u>QNPTGISYSQPQQQQQQPQQQPQYMQNFQQQQPQYAQN</u> <u>FQQQPQY</u> <u>TQNYQQQPQYIQPHQ</u> (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)

Table 10

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

Table 11

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

Table 12

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

Table 13

IDR	HIS4
Sequence	NHHASHGHHNSHHPQH HHHHHHHHHHHHPPPPAPQPPPPPQQQQ (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)

Table 14

IDR	HIS5
Sequence	SGHHGAHHGAHHPAAAAAYEAFRGPGFAGGGGADDMGAGHHH GAHHAAHHHHAHHHHHHHHHHHGGAGHGGG 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)

Table 15

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

Table 16

IDR	HIS7
Sequence	PGSGGPHTSHHHHHHHHHHHHQSRRH (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)

Table 17

IDR	HIS8
Sequence	APSYHHHHHHHHHHHHHHHHHHHHQQQHQQP (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)

Table 18

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

Table 19

IDR	HIS10
Sequence	PPPHHHHHHHHHHHHHHHRAQPPQQSHHPPHHHRPQPH (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)

Table 20

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

5 **Functional properties of an IDR-macromolecule or IDR-polypeptide**

An IDR-tagged macromolecule or polypeptide, or IDR-macromolecule or IDR-polypeptide, as defined herein, must possess an amino acid sequence consisting of or

comprising one or more functional IDRs to be of use in the processes of the invention. Whether or not such an IDR amino acid sequence or domain is functional can be established by routine methods, such as those described herein.

5 Particle formation

The inventors have surprisingly discovered that IDR-tagged polypeptides or IDR-polypeptides are capable of forming particles in an appropriate solution. This is believed to occur by liquid-liquid demixing leading to phase separation of fluids within
10 the solution mixture mediated by the IDR amino acid sequence.

The formation of particles mediated by IDR amino acid sequences is described further in the examples below. The particles demonstrate a spheroidal appearance and can be described as “globules”, “globular foci” or “particles”.

The terms “particle”, “globules” or “globular foci” as referred to herein are
15 intended to be synonymous and can be used interchangeably. Conditions and methods which permit the observation and detection of particles are set out herein including in the examples below.

In the examples described herein, particle formation was observed to occur in a simple system comprising merely a solution of IDR-tagged polypeptide and divalent
20 metal cation. Particle formation was also found to occur in more complex mixtures including those comprising components required for RPA where one of the RPA protein components (Gp32) was IDR-tagged. In these situations, reaction components were seen to strongly co-localise with the particles, for example particles were seen to be dense in oligonucleotide as detected by the fluorescent labelled probe attached thereto,
25 as well comprising all other RPA reaction protein components.

Detection and monitoring of the particles can be performed using any suitable method, as well as the methods set out in the examples below. Exemplary methods include microscopy, light scattering, flow cytometry, and microfluidic methods.

The particles can be detected using microscopy, *e.g.*, differential interference
30 contrast or fluorescence microscopy, to directly observe the particles at high magnification. With the aid of a computer, microscope images can be automatically

obtained and analyzed. Additionally, microscopy can allow for continual or frequent monitoring of at least a portion of a mixture containing particles.

The particles can be detected using flow cytometry. In flow cytometry, one or more beams of light, *e.g.*, each of a single wavelength, are directed onto a hydrodynamically-focused stream of fluid. Suspended particles passing through the beams scatter the light, and fluorescent chemicals found in the particles or attached to the particles may be excited. The scattered and/or fluorescent light is analysed by detectors within the device, from which information about particle size and fluorescence can be determined. Modern flow cytometers can analyze several thousand particles every second, in "real time", and can actively separate and isolate particles having specified properties.

The particles can be detected using cytometry methods, devices, and systems as disclosed *e.g.* in US patent application publication nos. US2009/0079963 and US2010/0179068, and international patent application publication no. WO2009/112594.

The particles can be detected using microfluidic methods, devices, and systems. For example, the particles can be detected using a lab-on-a-chip device or system, or the like (see, *e.g.*, US patent application publication nos. US2009/0326903 and US2009/0297733).

Particles may be about 0.5-20 μm in size, *e.g.*, between about any two sizes selected from 0.5, 1, 1.5, 2, 2.5, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 18, and 20 μm (*e.g.*, about 1-10 μm in size).

The concentration of particles may be approximately 10 to 5000 particles/nl, *e.g.*, between any two numbers of particles selected from 10, 20, 50, 100, 200, 500, 1000, 2000, and 5000 particles per nl, may be detected (*e.g.* about 100-500 particles per nl). The concentration of particles may be approximately 200-400 particles per nanolitre.

Such phase-separated particles may be smaller than about 0.5 μm in size. Phase-separated particles, including those that are smaller than about 0.5 μm in size, may be detected by changes in the turbidity of the solution. Changes in the turbidity of solutions can be measured by standard means and can be quantified typically according to the Formazin Turbidity Unit (FTU) or Formazin Nephelometric Unit (FNU). Other

methods include size exclusion chromatography, including multiangle light scattering (SEC-MALS).

Experimental determination of IDR function

5

An IDR-macromolecule or IDR-polypeptide, or an IDR-tagged macromolecule or IDR-tagged polypeptide, as defined herein, can be determined to possess a functional intrinsically disordered region (IDR) amino acid sequence and/or domain thereof, and thus to be of use in the processes and reagents of the invention, by using *e.g.* a phase separation assay method or a RPA assay method as described below.

Accordingly, an IDR-macromolecule, IDR-polypeptide or an IDR-tagged macromolecule or IDR-tagged polypeptide is a macromolecule or polypeptide which comprises, or which is tagged with, an amino acid sequence consisting of one or more functional intrinsically disordered regions; or a macromolecule or polypeptide which is tagged with an amino acid sequence comprising one or more functional intrinsically disordered regions.. In all cases the functional intrinsically disordered region is one which may be determined to be functional in the phase separation assay method described below and/or in the RPA assay method as described below.

20 Phase separation assay method

The phase separation assay method is a method comprising:

1. tagging the one or more intrinsically disordered region amino acid sequences to a polypeptide to create an IDR-polypeptide fusion protein, preferably tagged to the recombinant phage vB EcoM NBG1 Gp32 protein to create a Gp32-IDR fusion protein, and providing purified IDR-polypeptide fusion protein;
2. adding the IDR-polypeptide fusion protein to a volume of water to a final concentration of 1000 ng/ μ l, wherein the final volume of the mixture is 50 μ l, and preferably adding a divalent metal cation to a final concentration of 2 mM or more, more preferably wherein the divalent metal cation is Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} ;

3. vortexing the mixture followed by pulse centrifuging the mixture;
4. transferring a 10 μ l sample of the mixture supernatant to a haemocytometer slide;
5. viewing the haemocytometer slide under a microscope at x400 magnification;
- 5 6. observing the formation of particles in the mixture;
7. in the absence of particles in the mixture, repeating steps 1 to 6 and incrementally increasing the concentration of the divalent metal cation until the formation of particles in the mixture is observed;
8. counting the number of particles which form in a magnification area of 218 μ m x 175 μ m at x400 magnification; and
- 10 9. (i) establishing that the amino acid sequence consisting of or comprising one or more intrinsically disordered regions (IDRs) is functional when 10 or more particles are counted in the magnification area, preferably when 50 or more particles are counted in the magnification area, more preferably when 100 or more particles are counted in the magnification area; or
- 15 (ii) establishing that the amino acid sequence consisting of or comprising one or more intrinsically disordered regions (IDRs) is non-functional if the concentration of the divalent metal cation is increased to 100 mM or more and no particles are observed to form in the magnification area, or less than 10
- 20 particles are counted in the magnification area.

In the above method, should it be desirable to examine the effect of providing a divalent metal cation on particle formation, step 2 may comprise adding a divalent metal cation to any desired final concentration. The effects of different concentrations of divalent metal cations may thus be examined.

- 25 In the above method, should it be desirable to examine the effect of providing ATP on particle formation, step 2 may comprise adding ATP to any desired final concentration. The effects of different concentrations of ATP may thus be examined. ATP may be provided for example at a concentration of 1mM to 3.5mM, *e.g.* 1mM to 2mM.

- 30 Step 2 may comprise adding a detectable nucleic acid molecule and wherein step 8 comprises counting the number of particles by the detection means. For example, step

2 may comprise adding a probe having the nucleic acid sequence set for the in SEQ ID NO:104, which is labelled with FAM (fluorescein) and step 8 may comprise detecting particles by fluorescence. The detectable nucleic acid molecule may be added to any suitable final concentration, such as 0.5 μ M.

5 Accordingly, the above assay may be used to examine reaction efficiency, the capability of causing liquid-liquid demixing and the capability of causing molecules to co-localise within a plurality of phase-separated aqueous compartments (particles).

In the above method, if the divalent metal cation is Mg^{2+} the source of the cation is preferably MgOAc. If the divalent metal cation is Ca^{2+} , the source of the cation is preferably $CaCl_2$. If the divalent metal cation is Mn^{2+} , the source of the cation is preferably $MnCl_2$.

RPA assay method

15 The RPA assay method is a method comprising:

1. tagging the one or more intrinsically disordered region polypeptide sequences to a Gp32 protein, preferably the recombinant phage vB EcoM NBG1 Gp32 protein, to create a Gp32-IDR fusion protein and providing purified Gp32-IDR fusion protein;
- 20 2. creating a reaction mixture comprising
 - a. Tris HCl pH 8.3, 25 mM;
 - b. KOAc, 7.5 mM;
 - c. DTT, 1 mM;
 - d. ATP, 2.5 mM;
 - 25 e. phosphocreatine, 20 mM;
 - f. creatine kinase, 1 μ M;
 - g. dNTPs, 1 mM;
 - h. purified Gp32-IDR fusion protein, 20 μ M;
 - i. purified UvsX, 4.8 μ M
 - 30 j. purified UvsY, 8.6 μ M
 - k. *S. aureus* DNA polymerase 1 (Sau), 0.135 μ M

- l. Exonuclease III, 0.27 μ M
 - m. forward primer, 0.4 μ M
 - n. reverse primer, 0.4 μ M
 - o. probe, 0.12 μ M
3. initiating a recombinase polymerase amplification reaction by adding to the reaction mixture 33 mM MgOAc and 10 copies of template nucleic acid;
 4. incubating the reaction mixture at 39°C in a fluorometer with magnetic mixing using a bearing ball;
 5. (i) establishing that the one or more intrinsically disordered region polypeptide sequence is functional when a 2-fold or more more increase in amplified product is detectable within 15 minutes by a measurable increase in fluorescence compared to baseline in a template-dependent fashion; or
 - (ii) establishing that the one or more intrinsically disordered region polypeptide sequence is non-functional when no amplified product is detected by fluorescence after 15 minutes.

In the above method, the forward primer sequence is:

CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98); the reverse primer sequence is: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99); the probe sequence is:

CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM] [THF][BHQ-1]ATACAAGGATTGGA (SEQ ID NO:100), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher; and the template is *Listeria monocytogenes* genomic DNA.

Step 5 of the above RPA assay method may alternatively comprise establishing that the one or more intrinsically disordered region polypeptide sequence is functional when a 5-fold or more more increase in amplified product is detectable within 15 minutes by a measurable increase in fluorescence compared to baseline in a template-dependent fashion, or when a 10-fold or more more increase is detectable, or a 20-fold or more more increase, or a 30-fold or more more increase, or a 40-fold or more more increase, or a 50-fold or more more increase, or a 100-fold or more more increase, or a 150-fold or more more increase, or a 200-fold or more more increase, or a 250-fold or

more more increase, or a 300-fold or more more increase, or a 350-fold or more more increase, or a 400-fold or more more increase, or a 450-fold or more more increase, or a 500-fold or more, 1000-fold or more, 2000-fold or more, 3000-fold or more, 4000-fold or more or 5000-fold or more more increase. By an increase in amplified product over
5 baseline it is meant the increase in amplified product compared to the amount of amplified product obtained by performing the reaction under the same conditions except wherein the Gp32 protein has not been tagged with the one or more intrinsically disordered region polypeptide sequences.

In the above method, should it be desirable to examine the effect of providing a
10 divalent metal cation on reaction efficiency, step 3 may comprise adding a divalent metal cation to any desired final concentration. The effects of different concentrations of divalent metal cations may thus be examined.

In the above method, should it be desirable to examine the effect of providing ATP on reaction efficiency, step 2 may comprise adding ATP to any desired final
15 concentration. The effects of different concentrations of ATP may thus be examined. ATP may be provided for example at a concentration of 1mM to 3.5mM, *e.g.* 1mM to 2mM.

Tagging of IDR amino acid sequences to macromolecule and polypeptides

20

The methods, processes and reagents of the present invention involve, *inter alia*, IDR-tagged macromolecules and IDR-tagged polypeptides, wherein an IDR-tagged macromolecule or an IDR-tagged polypeptide is a macromolecule or polypeptide of interest tagged with an amino acid sequence consisting of or comprising one or more
25 intrinsically disordered regions (IDRs) (which may be referred to herein as an IDR moiety).

The term “tag” or “tagging” is to be understood in its broadest sense. The terms are to be understood to mean that an IDR moiety, *i.e.* an amino acid sequence which consists of or comprises one or more functional IDRs, is attached to, tethered to, bound
30 to or otherwise associated with a macromolecule or a polypeptide of interest in any suitable way.

The most preferred means by which an IDR moiety is tagged to a polypeptide of interest is by creating a recombinant genetic fusion protein, wherein the polypeptide of interest is genetically engineered at the nucleotide level such that when transcribed and translated the expressed protein comprises the polypeptide of interest together with the
5 IDR moiety.

If desired, linkers may be placed between the polypeptide of interest and the IDR moiety. For example, flexible, rigid and cleavable linkers are well known in the art and are widely used in the manufacture of fusion proteins (see, *e.g.*: Fusion Protein Linkers: Property, Design and Functionality, Chen, X., *et al.* 2013, Adv. Drug Deliv.
10 Rev., 15, 65(10), pp1357–1369).

Standard methods for genetic engineering are well known in the art (see for example, Sambrook *et al.*, 2001, Molecular Cloning: a Laboratory Manual, 3rd edition, Cold Spring Harbour Laboratory Press; and Current Protocols in Molecular Biology, Greene Publishing and Wiley-Interscience, New York (1995)), as are methods for
15 protein expression and purification.

Other means by which an IDR moiety may be tagged to a macromolecule or polypeptide of interest is by one or more covalent bonds or by an affinity interaction.

The IDR moiety can be tagged to a polypeptide in any suitable orientation, such as at the N-terminus of the polypeptide of interest, at the C-terminus of the polypeptide
20 of interest, or the polypeptide of interest may comprise an IDR moiety at both its N-terminus and C-terminus, or at any amino acid position along the length of the polypeptide.

Peptides/oligopeptides/polypeptides/proteins may be attached/tethered conjugated to other macromolecules, including other
25 peptides/oligopeptides/polypeptides/proteins, by the use of methods very known in this technical field.

One such method is “click chemistry”. The term “click chemistry” is typically used to describe the reaction of an azide with an alkyne resulting in a 1,5-disubstituted 1,2,3-triazole in the presence of a copper catalyst. Click chemistry allows
30 peptides/oligopeptides/polypeptides/proteins to be conjugated to a vast range of other macromolecules including other peptides/oligopeptides/polypeptides/proteins, as well as

e.g. carbohydrates, nucleic acids, polymers, drugs, aptamers, hydrogels etc. The method is also referred to as “CuAAC” (Cu catalyzed alkyne azide cycloaddition) (see *e.g.* “Click” reactions: a versatile toolbox for the synthesis of peptide-conjugates. Tang, W. *et al.*, 2014, Chem. Soc. Rev., 43, pp7013-7039).

5 Many other linker/crosslinker chemistries are available to conjugate peptides/oligopeptides/polypeptides/proteins to other macromolecules, such as crosslinkers which contain maleimide, sulfhydryl reactive groups or succinimidyl esters (often referred to as NHS esters), which react with amines. For example, succinimides can be used to form covalent bonds between proteins or peptides and plastics materials.

10 Standard chemistries can be used which are commonly employed to create conjugates between polypeptides and non-polypeptide molecules, such as chemistries to create antibody-drug conjugates. Many such techniques are well known in this technical field.

Affinity-based interactions can also be employed. For example, an amino acid
15 sequence consisting of or comprising one or more functional intrinsically disordered regions may be attached/tethered to a macromolecule or polypeptide of interest by an affinity-based interaction, such as streptavidin-biotin, receptor-ligand interaction etc.

Multivalent metal cations for IDR amino acid sequence function

20

When IDR-macromolecules or IDR-polypeptides, as described and defined herein, are used in *in vitro* biochemical reactions, the *in vitro* biochemical reaction buffer preferably contains multivalent metal cations, preferably divalent metal cations.

The presence of multivalent/divalent metal cations in the reaction buffer helps to
25 facilitate and enhance liquid-liquid demixing leading to phase separation in the *in vitro* biochemical reaction milieu mediated/caused by an IDR-macromolecule or an IDR-polypeptide.

The functional capability of a divalent metal cation to enhance phase separation in an *in vitro* biochemical reaction milieu mediated/caused by an IDR-macromolecule
30 or IDR-polypeptide can readily be established, such as by techniques disclosed and defined herein. In particular, such functional capability can be established by the ability

of the multivalent/divalent metal cation to induce the formation of globular foci or particles, as further described and defined herein, in the *in vitro* biochemical reaction milieu in an IDR-dependent manner, for example as determined by assays described herein.

5 The use of a divalent metal cation in promoting/enhancing IDR-dependent liquid-liquid demixing leading to phase separation is preferred. However, functional equivalents of any multivalent or any divalent metal cation are envisaged. A functional equivalent of a multivalent/divalent metal cation as described herein is any agent that may substitute for a divalent metal cation in promoting IDR-dependent liquid-liquid
10 demixing leading to phase separation in an *in vitro* biochemical reaction milieu, for example as determined by assays described herein.

Any suitable multivalent/divalent metal cation may be used, either as a single agent or combination of agents and optionally in the presence of a chelating agent such as ethylenediaminetetraacetic acid (EDTA), ethylene glycol-bis(β -aminoethyl ether)-
15 N,N,N',N'-tetraacetic acid (EGTA) or nitriloacetic acid (NTA).

The divalent metal cation may be Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} , Ni^{2+} or Cu^{2+} . Any of these cations may be used as a single agent, or any combination of cations may be used. Preferably they are used as single agents. Preferred divalent metal cations are Mg^{2+} , Mn^{2+} and Ca^{2+} .

20 The specific multivalent/divalent metal cation which achieves optimal results in promoting IDR-mediated phase separation in an *in vitro* biochemical reaction milieu, as well as the specific concentration of the multivalent/divalent metal cation used may depend upon the specific intrinsically disordered region amino acid sequence which is used to tag the macromolecule or polypeptide of interest. The optimal
25 multivalent/divalent metal cation and optimal concentration can be established empirically using routine tests. Phase separation assays as described further herein may be used for this purpose.

Preferred concentrations ranges of the multivalent/divalent metal cations are from about 300 μ M to about 100 mM, from about 300 μ M to about 50 mM, from about
30 400 μ M to about 50 mM, from about 400 μ M to about 20 mM, from about 400 μ M to

about 30 mM, from about 500 μ M to about 10 mM, from about 500 μ M to about 25 mM and from about 1 mM to about 35 mM.

An *in vitro* biochemical reaction buffer may contain Mg^{2+} ions. A preferred concentration range is from about 300 μ M to about 100 mM, more preferably from about 400 μ M to about 50 mM, yet more preferably from about 500 μ M to about 40 mM, even more preferably from about 25 mM to about 35 mM, such as 33 mM. Preferably the buffer contains MgOAc at the indicated concentrations.

An *in vitro* biochemical reaction buffer may contain Ca^{2+} ions. A preferred concentration range is from about 300 μ M to about 100 mM, more preferably from about 400 μ M to about 50 mM, yet more preferably from about 1 mM to about 40 mM, even more preferably from about 25 mM to about 35 mM, such as 33 mM. Preferably the buffer contains $CaCl_2$ at the indicated concentrations.

An *in vitro* biochemical reaction buffer buffer may contain Mn^{2+} ions. A preferred concentration range is from about 300 μ M to about 50 mM, more preferably from about 400 μ M to about 50 mM, yet more preferably from about 500 μ M to about 40 mM, even more preferably from about 25 mM to about 35 mM, such as 33 mM. Preferably the buffer contains $MnCl_2$ at the indicated concentrations.

Recombinase polymerase amplification (RPA)

20

Recombinase polymerase amplification (RPA) is a method for isothermal amplification of nucleic acids. In general, in a first step of RPA, a recombinase agent is contacted with first and second nucleic acid primers and a recombinase loading protein to form first and second nucleoprotein primers. In general, in a second step, the first and second nucleoprotein primers are contacted with a double stranded template nucleic acid to form a first double stranded structure at a first portion of the first strand of the template nucleic acid, and a second double stranded structure at a second portion of the second strand of the template nucleic acid such that the 3' ends of the first nucleic acid primer and the second nucleic acid primer are orientated towards each other on a given nucleic acid molecule. In general, in a third step, the 3' end of the first and the second nucleoprotein primers are extended by polymerase to generate first and second double

stranded nucleic acids, and first and second displaced single strands of nucleic acid. A single stranded stabilizing agent is employed to stabilize the first and second displaced single strands of nucleic acid. Generally, the second and third steps can be repeated until a desired degree of amplification is reached.

5 RPA methods are disclosed extensively, *e.g.*, in US Patent No. 7,270,981; US Patent No. 7,399,590; US Patent No. 7,666,598; US Patent No. 7,435,561 and International Patent Application Publication No. WO2010/141940. In addition, for a comprehensive recent review see: Review: a comprehensive summary of a decade development of the recombinase polymerase amplification, Li, J. *et al.*, 2019, Analyst,
10 144, pp31-67).

Recombinase agent

RPA methods, including those of the present invention, use a recombinase
15 agent.

Any of the one or more IDR-polypeptides of the present invention may be attached/tethered/tagged to any recombinase agent.

A recombinase agent is a molecule, typically an enzyme, that can coat a single stranded nucleic acid, typically DNA (ssDNA) to form a nucleoprotein filament. Such
20 filaments can then “scan” a double stranded nucleic acid molecule, typically DNA (dsDNA) for regions of sequence homology/complementarity. When complementary sequences are located, the nucleoprotein filament (comprising the recombinase agent) strand invades the double stranded nucleic acid molecule creating a short hybrid and a displaced strand bubble known as a D-loop.

25 Any suitable recombinase agent may be used in the RPA methods described herein, and may be tagged with any of the IDR amino acid sequences as described herein.

The recombinase agent may originate from a prokaryotic, eukaryotic or viral organism.

30 The recombinase agent may be RecA, UvsX, RadA, RadB, Rad 51 or any functional variant, analogue, homologue or derivative of any of these proteins.

Any combination of these proteins may be used.

Suitable recombinase agents include the *E. coli* RecA protein, the T4 UvsX protein, or any homologous protein or protein complex from any phyla.

Eukaryotic RecA homologues are generally named Rad51 after the first member of this group to be identified. Other non-homologous recombinase agents may be utilized in place of RecA, for example RecT or RecO.

Exemplary recombinase agents include RecA and UvsX, and fragments or mutants thereof and combinations thereof. The RecA and UvsX proteins can be obtained from any species. RecA and UvsX fragments or mutant proteins can also be produced using the available RecA and UvsS protein and nucleic acids sequences, and molecular biology techniques. Exemplary UvsX proteins include those derived from myoviridae phages, such as T4, T2, T6, Rb69, Aeh1, KVP40, Acinetobacter phage 133, Aeromonas phage 65, cyanophage P-SSM2, cyanophage PSSM4, cyanophage S-PM2, Rb14, Rb32, Aeromonas phage 25, Vibrio phage nt-1, phi-1, Rb16, Rb43, Phage 31, phage 44RR2.81, Rb49, phage Rb3, and phage LZ2. Additional exemplary recombinase agents include archaeobacterial RADA and RADB proteins and eukaryotic (*e.g.*, plant, mammal, and fungal) Rad51 proteins (*e.g.*, RAD51, RAD51B, RAD51C, RAD51D, DMC1, XRCC2, XRCC3, and recA).

The recombinase agent is preferably UvsX, T4 UvsX, T6 UvsX, RB18 UvsX, *E. coli* phage wV7 UvsX, *Shigella* phage CB8 UvsX, *Shigella* phage Shfl2 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 or any functional variant, analog, homolog or derivative thereof or any combination thereof. A particularly preferred recombinase agent is *Escherichia* phage vB_EcoM_DalCa UvsX.

The recombinase agent may also comprise a C-terminal deletion of acidic residues to improve its activity.

Any functional variants, analogs, homologs or derivatives of the recombinase agent above may also function itself as a recombinase agent and these functional variants, analogs, homologs or derivatives are also contemplated as a recombinase agent to be used in the processes described and defined herein.

For example, a small peptide from RecA, has been shown to retain some aspects of the recombination properties of RecA. This peptide comprises residues 193 to 212 of *E. coli* RecA and can mediate pairing of single stranded oligonucleotides.

The recombinase agent (*e.g.*, UvsX) may be a mutant or hybrid recombinase agent. Mutant forms of UvsX are described in U.S. Patent No. 8,071,308. The mutant UvsX may be an Rb69 UvsX that includes at least one mutation in the Rb69 UvsX amino acid sequence, wherein the mutation is selected from the group consisting of (a) an amino acid which is not histidine at position 64, a serine at position 64, the addition of one or more glutamic acid residues at the C-terminus, the addition of one or more aspartic acid residues at the C-terminus, and a combination thereof.

The mutant UvsX may be a T6 UvsX having at least one mutation in the T6 UvsX amino acid sequence, wherein the mutation is selected from the group consisting of (a) an amino acid which is not histidine at position 66; (b) a serine at position 66; (c) the addition of one or more glutamic acid residues at the C-terminus; (d) the addition of one or more aspartic acid residues at the C-terminus; and (e) a combination thereof. Where a hybrid recombinase agent is used, the hybrid protein may, for example, be a UvsX protein that includes at least one region that includes an amino acid sequence derived from a different UvsX species. The region may be, for example, the DNA-binding loop-2 region of UvsX.

If desired, the recombinase agent may be a temperature-sensitive (referred to herein as "ts") recombinase agent. If a ts recombinase agent is used, the RPA reaction may be started at one temperature (the permissive temperature) and terminated at another temperature (the non-permissive temperature). Combinations of permissive temperatures may be, for example 25°C/30°C, 30°C/37°C, 37°C/42°C and the like. The ts protein may be reversible. A reversible ts protein's activity is restored when it is shifted from the non-permissive temperature to the permissive temperature.

While any recombinase agent concentration may be used, preferred recombinase concentrations may be, for example, in the range of 0.2-12 μ M, 6-12 μ M, 4-12 μ M and 4-6 μ M, preferably about 5 μ M, more preferably about 4.8 μ M.

Recombinase agents generally require the presence of ATP, ATP γ S, or other nucleoside triphosphates or their analogs. It is preferred that recombinase agents are

used in a reaction environment in which regeneration of targeting sites can occur shortly following a round of D-loop stimulated synthesis. Completed recombination events involving recombinase disassembly will avoid a stalling of amplification or very inefficient linear amplification of ssDNA caused by oscillating single sided synthesis

5 from one end to the other.

Exemplary UvsX recombinase agents tagged with amino acid tag sequences comprising intrinsically disordered regions are set out in Table 21 below.

Table 21

Name of protein	SEQ ID NO.	Amino acid sequence
UvsX (7His)	44	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELG FVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPF RDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDL EQLSDMEEFNEHHHHHHH
UvsX-fib-1	45	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF

		DGGIDPYSGLLDMALELG FVVKPKN GWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPF RDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNEPGFSPRGGGGFGGRGGFGDR GGRGGRGGFGGGGRGRGGGGFRGRGRHHHHHHH
UvsX-fib-2	46	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELG FVVKPKN GWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPF RDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNEPGFSPRGGGGFGGRGGFGDR GGRGGRGGFGGGGRGRGGGVEHHHHHHH
UvsX-fib-3	47	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELG FVVKPKN GWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPF RDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNEPGFSPRGGGGFGGRGGFGDR GGRGGRGGGVEHHHHHHH

UvsX-fib-4	48	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPFRAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNPGFSPRGGGGFGGRGGVEH HHHHH
UvsX- HNRNPA1	49	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPFRAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNESASSSQRGRSGSGNFGGG RGGGFGGNDNFGRGGNFSGRGGFGGSRGGGGYGGG GDGYNGFGNDGSNFGGGGSGYNDFGNYNQSSNFGP MKGGNFGGRSSGPYGGGGQYFAKPQNQGGYGVSSSS SSYGSGRRFHHHHHHH

UvsX-DDX	50	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELG FVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWG PLFKHQPF RDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDL EQLSDMEEFNEMGDEDWEAEINPHMSSYV PIFEKDRYSGENGDNFNRT PASSEMDDGPSRRDHFM KSGFASGRNFGNRDAGECNKRDNTSTMGGFGVGKSF GNRGFSNSRFEDGDSSGFWRESSND CEDNPTRNRGFS KRGGYRDGNNSEASGPYRRGGRGSFRGCRGGFGLGS PNNDLDPDECMQRTGGLFGSRRPVLSGTGNGDTSQS RSGSGSERGGYKGLNEEVITGSGKNSWKSEAEGGESS DTQH HHHHHH
UvsX addPolCTD	51	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELG FVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWG PLFKHQPF RDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDL EQLSDMEEFN EHHHHHHHHFEFSPFGAYGE APTSPGFGVSSPGFSPTSPTYSPTSPAYSPTSPSYSPTSP

		SYSPTSPSYSPTSYSPTSPSYSPTSYSPTSPSYSPTS PSYSPTSPSYSPTSYSPTSPSYSPTSYSPTSPSYSPTS SPSYSPTSPAYSPTSYSPTSPSYSPTSYSPTSPSYSPTS TSPNYSPTSPSYSPTS PGYSPGSPAYSPKQDEQTAAL HHHHHH
UvsX-fusPolII	52	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELG FVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWG PLFKHQPF RDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNETS PSYSPTSPSYSPTS PGYSP TSPAYSPTSPTYSPTSPTYSPHHHHHHH
UvsX-PCF11	53	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELG FVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWG PLFKHQPF RDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFN EQVQMQLRQVFSQDQQVLQ

		ERMRYHELQQQQQQQYHETKDMVGSYTQNSNSAIPL FGNNSDTTNQQNSHHHHHHH
UvsX-Sup	54	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNEMSDSNQGNNQQNYQQYSQ NGNQQQGNNRYQGYQAYNAQAQPAGGYQNYQGY SGYQQGGYQQYQYNPQGGYQQYNPQGGYQQYNPQ GGYQQQFNPQGGRGNYKNFNYNNNLQGYQHHHHH HH
UvsX- DoubleX	55	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNE HHHHHHHFEFEIVEAEVDELINSKVEKFKSPESKSKSA

		ADLETDLSQLSDMEEFNEIVEAEVDELINSKVEKFKSP ESKSKSAADLETDLSQLSDMEEFNHHHHHHH
Superpositive	56	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKR AYQLGAIDSNRIVKAKVKRLINSKVKKFKSPKSKSKS AAKLKTKLKQLSKMKKFNKIVKAKVKKLINSKVRKF KSPKSKSKSAAKLKTKLKQLSKMKEFNHHHHHHH
Supernegative	57	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKR AYQLGAIDSNEIVEAEVDELINSEVEEFDSPESDSDEA ADLETDLSQLSDMEEFNEIVEAEVDELINSEVEDFDSP ESDSDEAADLETDLSQLSDMEEFNHHHHHHH
UvsX His2	58	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV

		AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNEAGHHHHHPHAHHPLSQSS GHHHHHHHHHHQGYGGSG
UvsX+PCF+His2	59	MSIADLKSRLIKASTSKMTAELTTSKFFNEKDVIRTKIP MLNIAISGAIDGGMQSGLTIFAGPSKSFKSNMSLTMV AAYLNKYPDAICLFYDSEFGITPAYLKSMGVDPERVI HTPIQSVEQLKIDMVNQLETIERGEKVIVFIDSIGNMAS KKETEDALNEKSVADMTRAKSLKSLFRIVTPYFSIKNI PCVAVNHTIETIEMFSKTVMTGGTGVMYSADTVFIIG KRQIKDGSDLQGYQFVLNVEKSRTVKEKSKFFIDVKF DGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEE TGEMIREEKSWRAKDTNCTTFWGPLFKHQPFDAIKR AYQLGAIDSNEIVEAEVDELINSKVEKFKSPESKSKSA ADLETDLQLSDMEEFNEQVQMQLRQVFSQDQQVLQ ERMRYHELQQQQQQQYHETKDMVGSYTQNSNSAIPL FGNNSDTTNQQNSAGHHHHHPHAHHPLSQSSGHHHH HHHHHHQGYGGSG

Recombinase loading protein

RPA methods, including those of the present invention, may additionally include/use a recombinase loading protein.

5 Any suitable recombinase loading protein may be used in the RPA methods described herein.

Any of the one or more IDR-polypeptides of the present invention may be attached/tethered/tagged to any recombinase loading protein.

The recombinase loading protein may originate from a prokaryotic, viral or
 10 eukaryotic organism. Exemplary recombinase loading proteins include *E. coli* RecO, *E. coli* RecR, UvsY, and mutants or fragments thereof, or combinations thereof. Exemplary UvsY proteins include those derived from myoviridae phages, such as T4, T2, T6, Rb69, Aeh1, KVP40, Acinetobacter phage 133, Aeromonas phage 65, cyanophage P-SSM2, cyanophage PSSM4, cyanophage S-PM2, Rb14, Rb32,
 15 Aeromonas phage 25, Vibrio phage nt-1, phi-1, Rb16, Rb43, Phage 31, phage 44RR2.8t, Rb49, phage Rb3, and phage LZ2.

Preferred recombinase loading proteins are UvsY, *E. coli* RecO, *E. coli* RecR or any functional variant, analogue, homologue or derivative of any of these proteins. A particularly preferred UvsY recombinase loading protein is *Escherichia* phage STO
 20 UvsY.

Any combination of any of these proteins may be used.

Preferred concentrations of these proteins is between 0.1-24 μ M, 6-24 μ M, 4-24 μ M and 4-12 μ M, preferably about 10 μ M, more preferably about 8.6 μ M. The recombinase loading protein may be present at between about 0.5 to about 2 times the
 25 micromolar concentration of the recombinase agent.

Exemplary UvsY recombinase loading proteins tagged with amino acid tag sequences comprising intrinsically disordered regions are set out in Table 22 below.

Table 22

Name of protein	SEQ ID NO.	Amino acid sequence
UvsY [Phage ST0]	60	MHHHHHHHMKLEDLQEELDADLAIDMSKLQYETAN NVKLYSKWLRKHSFIRKEMLRJETQKKTALKARLDY YSGRGDGDEFSDRYEKSEMKTVLAADKDVVKIETT LQYWGILLEFCSGALDAVKSRSFALKHIQDMREFEAG Q
UvsY C-Pol CTD	61	MHHHHHHHHKLEDLQEELDADLAIDMSKLQYETAN NVKLYSKWLRKHSFIRKEMLRJETQKKTALKARLDY YSGRGDGDEFSDRYEKSEMKTVLAADKDVVKIETT LQYWGILLEFCSGALDAVKSRSFALKHIQDMREFEAG QSGSGSGPTSPSYSPSPSYSPYSPAYS
UvsY fib[short]	62	MHHHHHHHHKLEDLQEELDADLAIDMSKLQYETAN NVKLYSKWLRKHSFIRKEMLRJETQKKTALKARLDY YSGRGDGDEFSDRYEKSEMKTVLAADKDVVKIETT LQYWGILLEFCSGALDAVKSRSFALKHIQDMREFEAG QSGSGSGRGGGFGGRGGFGDRGGRGGRGGFGG
UvsY Sup1	63	MHHHHHHHHKLEDLQEELDADLAIDMSKLQYETAN NVKLYSKWLRKHSFIRKEMLRJETQKKTALKARLDY YSGRGDGDEFSDRYEKSEMKTVLAADKDVVKIETT LQYWGILLEFCSGALDAVKSRSFALKHIQDMREFEAG QSGSGYNPQGGYQQNNL
UvsY Sup1 HIS2	64	MAGHHHHHPAHHPHSQSSGHHHHHHHHHHHQGYG GSGKLEDLQEELDADLAIDMSKLQYETANNVKLYSK

		WLRKHSFIRKEMLRIETQKKTALKARLDYYSGRGDG DEFSMDRYEKSEMKTVLAADKDVVKIETTLQYWGIL LEFCSGALDAVKSRSFALKHIQDMREFEAGQSGSGYN PQGGYQQNNLQ
--	--	--

Single strand stabilizing agent

RPA methods, including those of the present invention, use a single strand
5 stabilizing agent.

Any suitable single strand stabilizing agent (single stranded DNA binding protein) may be used in the RPA methods described herein.

Any of the one or more IDR-polypeptides thereof of the present invention may be attached/tethered/tagged to any single strand stabilizing agent.

10 A single strand stabilizing agent is used to stabilize nucleic acids during the various exchange reactions that occur during the RPA reaction. In particular a single strand stabilizing agent is used to stabilize recombinase/ssDNA nucleoprotein filaments.

A single strand stabilizing agent can be derived or obtained from any species, *e.g.*, from a prokaryotic, viral or eukaryotic species.

15 Single strand stabilizing agents include single stranded DNA binding proteins from *E. coli* and those derived from myoviridae phages, such as T4, T2, T6, Rb69, Aeh1, KVP40, Acinetobacter phage 133, Aeromonas phage 65, cyanophage P-SSM2, cyanophage PSSM4, cyanophage S-PM2, Rb14, Rb32, Aeromonas phage 25, Vibrio phage nt-1, phi-1, Rb16, Rb43, Phage 31, phage 44RR2.81, Rb49, phage Rb3, and
20 phage LZ2. Additional examples of single strand stabilizing agents include *A. denitrificans* Alide_2047, *Burkholderia thailandensis* BthaB_33951, *Prevotella pollens* HMPREF9144_0124, and eukaryotic single stranded DNA binding protein replication protein A.

Preferred single strand stabilizing agents are selected from the group consisting
25 of Gp32, *E. coli* SSB protein, phage T4 Gp32 protein, phage Rb69 Gp32, phage vB_EcoM_NBG1 Gp32, and derivatives thereof and any combination thereof.

Particularly preferred single strand stabilizing agents are Gp32 and in particular phage vB_EcoM_NBG1 Gp32.

Any combination of any of these proteins may be used.

One preferred concentration of the single strand stabilizing agent is between approximately 5-30 μ M, such as approximately 8.6 μ M, preferably between approximately 15-25 μ M, more preferably approximately 20 μ M.

Exemplary Gp32 single strand stabilizing agents tagged with amino acid tag sequences comprising intrinsically disordered regions are set out in Table 23 below.

10 Table 23

Name of protein	SEQ ID NO.	Amino acid sequence
Gp32 (7His) from phage vB	65	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLHHHHHHH
Gp32Super +ve	66	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVAKKVASKLKFKKKMEAFSSAKTKKKFMSS

		SSSKKSKLKKLLAGLMEAFSSAKTKKKFMSSSSSKK SKLKKLLAGLHHHHHHH
Gp32Super -ve	67	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADDVASDLDDFDDDMEAFSSADTEDDFMS SSSSDDSDLDDLLAGLMEAFSSADTEDDFMSSSSSD DSDLDDLLAGLHHHHHHH
Gp32- TripleXtail	68	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADDVASEFEIVEAEVDELINSKVEKFKSPES KSKSAADLETDLQLSDMEEFNEIVEAEVDELINSKV EKFKSPESKSKSAADLETDLQLSDMEEFNEIVEAEV DELINSKVEKFKSPESKSKSAADLETDLQLSDMEEF NHHHHHH
Gp32-fib	69	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK

		KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFOKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLPGFSPRGGGGFGGRGGFGDRG GRGGRGGFGGGGRGRGGGGFRGRGRHHHHHHH
Gp32-PCF11	70	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFOKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLNEQVQMQLRQVFSQDQQVL QERMRYHELQQQQQQQYHETKDMVGSYTQNSNSA IPLFGNNSDTTNQQNSHHHHHHH
Gp32-Sup	71	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFOKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLMSDSNQGNNQQNYQQYSQN GNQQQGNNRYQGYQAYNAQAQPAGGYQNYQGY SGYQQGGYQQYQYNPQGGYQQYNPQGGYQQYNPQ

		GGYQQQFNPQGGRGNYKNFNYNNNLQGYQH HHHH HH
Gp32-Sup1	72	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLYNPQGGGYQQHHHHHHH
Gp32-Sup2	73	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLYNPQGGGYQQYNPQGGGYQQH HHHHHH
Gp32-Sup3	74	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA

		AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLYNPQGGYQQYNPQGGYQQY NPQGGYQQHHHHHHH
Gp32-Sup4	75	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLYNPQGGYQQYNPQGGYQQY NPQGGYQQYNPQGGYQQHHHHHHH
Gp32-DDX	76	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLHHHHHHHFEFMGDEDWEAE INPHMSSYVPIFEKDRYSGENGDNFNRTPASSEMDD GPSRRDHFMKSGFASGRNFGNRDAGECNKRDNTST MGGFGVGKSFGNRGFSNSRFEDGDSSGFWRESSND CEDNPTRNRGFSKRGGYRDGNNSEASGPYRRGGRG SFRGCRGGFGLGSPNNDLDPDECMQRTGGLFGSRRP VLSGTGNGDTSQSRSGSGSERGGYKGLNEEVITGSG KNSWKSEAEGGESDQTLEHHHHHH

Gp32 PolCTD	77	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLHHHHHHHHFEFSPFGAYGEAP TSPGFGVSSPGFSPTSPTYSP TSPA YSPTSPSYSP TSPS YSPTSPSYSP TSPSYSP TSPSYSP TSPSYSP TSPSYSP TSPSYSP PSYSP TSPSYSP TSPSYSP TSPSYSP TSPSYSP TSPSYSP TSPSYSP TSPSYSP TSPA YSPTSPSYSP TSPSYSP TSPSYSP TSPSYSP TSPSY SPTSPNYSPTSPSYSP TSPGYSPGSPAYSPKQDEQLEH HHHHH
Gp32 HNRNPA1	78	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLASASSSQRGRSGSGNFGGGR GGGFGGNDNFGRGGNFSGRGGFGGSRGGGGYGG S GDGYNGFGNDGSNFGGGGSYNDFGNYNQSSNFGP MKGGNFGGRSSGPYGGGGQYFAKPQNQGGYGVSSS SSSYGSGRRFH HHHHHH

Gp32 HRP1	79	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLGGNNGGNNMNR RGGNFGNQ GDFNQMYQNPMMSGGYNPMMNPQAMTDYYQKMQ EYYQQMQHHHHHHH
Gp32 HRP2	80	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLGGNNGGNNMNR RGGNFGNQ GDFNQMYQNPMMSGGYNPMMNPQAMTDYYQKMQ EYYQQMQGGNNGGNNMNR RGGNFGNQGDFNQMY QNPMMSGGYNPMMNPQAMTDYYQKMQEYYQQMQ HHHHHHH
Gp32 HIS1	81	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK

		VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLPQQNALHHHHGNSSHHHHH HHHHHHHHHGQQA
GP32 HIS2	82	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLAGHHHHHPHAHHPLSQSSGH HHHHHHHHHQGYGGSG
GP32 HIS3	83	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLAGHHHHHHHHHLPPLPPHLH HHHHPQHHLHPGSAAAVHPVQQH
GP32 HIS4	84	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY

		SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLNHHASHGHHNSHHPQH HHH HHHHHHHPPPPAPQPPPPPQQQQ
GP32 HIS5	85	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLSGHHGAHHGAHHPA AAAAY EAFRGPGFAGGGGADDMGAGHHHGAHHA AHHHH AAHHHHHHHHHHHGGAGHG GGGAGHH
Gp32 Mimic1	86	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDESFKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLYDPTSPSYDPTSPSYDPTSPSY DPTSPSYDPTSPSYDPTSPSYDPTSPSHHHHHHH

Gp32 Mimic2	87	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLYSPTDPSYSPTDPSYSPTDPSY SPTDPSYSPTDPSYSPTDPSYSPTDPSHHHHHHH
Gp32 short pol ctd	88	MFKRKSTADLAAQMAKLNGNKGFSSEDKGEWKLK LDASGNGQAVIRFLPAKTDDALPFTILVNHGFKKNG KWYIETCSSTHGDYDSCPVCQYISKNDLYNTNKTEY SQLKRKTSYWANILVVKDPQAPDNEGKVFKYRFGK KIWDKINAMIAVDTEMGETPVDVTCPWEGANFVLK VKQVSGFSNYDESKFLNQSAIPNIDDES FQKELFEQM VDLSEMTSKDKFKSFEELNTKFNQVLGTAALGGAA AAAASVADKVASDLDDFDKDMEAFSSAKTEDDFMS SSSSDDSDLDDLLAGLTSPSYSPTSPSYSPTSPGYST SPAYSPTSPTYSPTSPTYSPHHHHHHH

Polymerase

- RPA methods, including those of the present invention, use a polymerase.
- 5 Any suitable polymerase may be used in the methods described herein.
- Any of the one or more IDR-polypeptides of the present invention may be attached/tethered/tagged to any suitable polymerase.

For the synthesis or amplification of DNA, DNA polymerase are preferably used.

One advantage of the RPA reaction is that there is no limit on the type of polymerase that can be used. For example, eukaryotic, prokaryotic and bacteriophage
5 polymerases can be used.

The DNA polymerase may be a eukaryotic polymerase. Examples of eukaryotic polymerases that may be used include pol- α , pol- β , pol- δ , pol- ϵ or any functional variant, analogue, homologue or derivative thereof and any combination thereof.

The DNA polymerase may be a prokaryotic polymerase. Examples of
10 prokaryotic polymerases that may be used include *E. coli* DNA polymerase I Klenow fragment, *E. coli* DNA polymerase I, *E. coli* DNA polymerase II, *E. coli* DNA polymerase III, *E. coli* DNA polymerase IV, *E. coli* DNA polymerase V, *Bacillus stearothermophilus* polymerase I large fragment, *Bacillus subtilis* Pol I large fragment (Bsu polymerase), *Listeria monocytogenes* DNA polymerase I, *Staphylococcus aureus*
15 DNA polymerase 1 (Sau) or any functional variant, analogue, homologue or derivative thereof and any combination thereof.

The DNA polymerase may be a bacteriophage polymerase. Examples of bacteriophage polymerases that may be used in the methods described herein include Phi-29 DNA polymerase, T7 DNA polymerase, bacteriophage T4 gp43 DNA
20 polymerase, or any functional variant, analogue, homologue or derivative thereof and any combination thereof.

The DNA polymerase typically contains strand displacing properties.

DNA polymerases can use the free 3'-hydroxyl of the invading strand to catalyze DNA synthesis by incorporation of new nucleotides. A number of polymerases
25 can use the 3'-hydroxyl of the invading strand to catalyze synthesis and simultaneously displace the other strand as synthesis occurs. For example *E. coli* polymerase II or III can be used to extend invaded D-loops. In addition, *E. coli* polymerase V normally used in SOS-lesion-targeted mutations in *E. coli* can be used. All of these polymerases can be rendered highly processive through their interactions and co-operation with the
30 β -dimer clamp, as well as single stranded DNA binding protein (SSB) and other

components. Other polymerases from prokaryotes, viruses, and eukaryotes can also be used to extend the invading strand.

Many DNA polymerases possess 3'-5' exonuclease activity, and some also possess 5'-3' exonuclease activity, which is undesirable in RPA reactions as it results in digestion of one DNA strand progressively as the polymerase moves forward, rather than displacement.

The 3'-5' exonuclease has potential advantages as well as its obvious disadvantages. On the one hand 3'-5' exonuclease activity increases the fidelity of the replication reaction, and can also prevent stalling of polymerases at points of misincorporation. High fidelity amplification is desirable for many DNA applications. The 3'-5' exonuclease activity may also be appropriate for amplification of larger DNA fragments where stalling due to misincorporation could inhibit effective amplification.

Despite these clear advantages of 3'-5' exonuclease activity there are some disadvantages. The free oligonucleotides can be subject to end-dependant degradation when polymerases possessing 3'-5' exonuclease are employed.

Reaction noise can be reduced by utilising polymerases lacking 3'-5' exonuclease activity. This suggests mispriming may result from oligonucleotides that have been shortened by the 3'-5' exonuclease activity of polymerases. Consequently 3'-5' exonuclease editing activity, pyrophosphorylysis, or any other similar editing activity can be a source of noise. This can be suppressed to a large extent by using saturating amounts of relatively cooperative Gp32 protein with some polymerases such as the Klenow fragment. Nevertheless, polymerases for use in the methods described herein might be provided which lack 3'-5' exonuclease activity.

The DNA polymerase may be present at a concentration of between 10,000 units/ml to 10 units/ml, such as between 5000 units/ml to 500 units/ml.

Accessory agents

RPA reactions, including those of the present invention, may further utilize accessory agents.

Any of the one or more IDR-polypeptides of the present invention may be attached/tethered/tagged to any accessory agent.

These accessory agents include a single strand binding protein, a helicase, a topoisomerase, a resolvase and any combination thereof. Such agents may possess
5 unwinding, relaxing, and resolving activities respectively on nucleic acids.

The accessory agents may also include RuvA, RuvB, RuvC, RecG, PriA, PriB, PriC, DnaT, DnaB, DnaC, DnaG, DnaX clamp loader, polymerase core complex, DNA ligase and a sliding clamp and any combination thereof. The sliding clamp may be *E. coli* β -dimer sliding clamp, the eukaryotic PCNA sliding clamp, or the T4 sliding clamp
10 gp45 and a combination thereof. The accessory agents may include, in addition, DNA Polymerase III holoenzyme complex consisting of β -Clamp, DnaX Clamp Loader, and the Polymerase Core Complex. These latter accessory agents would allow the performance of leading and lagging RPA.

RPA reactions may be performed with one or more additional enzymes that can
15 promote efficient disassembly of recombinase agent/dsDNA complexes after initiation of DNA synthesis. These enzymes include those that are capable of stimulating 3' to 5' disassembly and those capable of supporting 5' to 3' disassembly.

Such additional enzymes include several polymerases that can displace RecA in the 3' to 5' direction and can stimulate 3' to 5' disassembly of recombinase
20 agent/dsDNA complexes. These DNA polymerases include *E. coli* PolV and homologous polymerase of other species. Inclusion of *E. coli* PolV or any functional variant, analogue, homologue or derivative thereof may improve the amplification efficiency.

Other enzymes include a class of enzymes called helicases that can be used to
25 promote the disassembly of RecA from dsDNA. These promote disassembly in both the 5' to 3' and 3' to 5' directions. An ideal helicase complex for stimulating disassembly of RecA from intermediates consists of the *E. coli* proteins RuvA and RuvB. The RuvAB complex promotes branch migration, and dissociates the RecA protein, allowing RecA to be recycled. Incorporation of RuvAB into the RPA mixture
30 can promote the dissociation of RecA from dsDNA following strand exchange and displacement, allowing renewed synthesis of the duplicated template from the same site.

Additionally, the RuvAB complex can act in concert with RuvC, which finally cuts and resolves Holliday junctions. With RuvC added to the RPA reaction mixture, complicated structures such as Holliday junctions formed at invasion sites, can be resolved.

5 Still other enzymes include the *E. coli* RecG protein. RecG can stimulate disassembly of branch structures.

Other enzymes useful in an RPA reaction mixture are those that allow continual generation of RecA nucleoprotein filaments in the presence of ATP and the single strand stabilizing agent. Accordingly, RecO and RecR, and optionally RecF proteins
10 may be used.

Exonuclease enzymes are often included in RPA reaction mixtures. These are included for the efficient operation of cleavable probes. One example of an exonuclease enzyme commonly used is Exonuclease III. Any of the IDR polypeptides of the present invention may be attached/tethered/tagged to any exonuclease.

15

Primers

RPA methods employ polymerases to generate copies of template nucleic acid molecules. RPA methods, including those of the present invention, therefore use
20 primers to initiate extension by polymerases.

It is a necessity of most nucleic acid polymerases that incorporation requires a free 3'-hydroxyl moiety on the terminal sugar of a short stretch of double stranded nucleic acid adjacent to the site of new synthesis. This stretch of double stranded nucleic acid is typically formed on a template by a short oligonucleotide typically
25 having a complementary sequence, called a primer, which serves as an initiation site for the polymerase synthesis reaction. In some cases a 3' modification, such as a sulfhydryl, may be utilized to prime the synthesis reaction. The primer nucleic acid, which is base-paired with the template and extended by the polymerase, can be RNA or DNA. Typically, for *in vitro* reactions the primer is supplied as a short, often chemically
30 synthesized, single stranded DNA (or modified DNA or RNA), and is usually referred to as an oligonucleotide primer. The primer is often of a specific sequence, although

random primers can also be used. The primer is targeted to complementary sequences by virtue of its specific base-pairing capacity. Formation of hybrids between the oligonucleotide primer and target nucleic acid are typically formed by incubation of the two in solution under conditions of salt, pH, and temperature that allow spontaneous
5 annealing.

The primers used in RPA may have a single stranded region for hybridization to the target DNA in the presence of a recombinase agent. The single stranded region may be, for example, about 10 bases, about 15 bases, about 20 bases, about 25 bases, about 30 bases, about 40 bases, and about 50 bases. Even longer regions such as about 75
10 bases, about 100 bases, about 150 bases or more may in theory be used. The choice of single stranded regions will depend on the complexity of the starting nucleic acid so that for example, a human genome may require a longer primer while a plasmid may require a much shorter primer.

A preferred primer length is between about 30 to about 50 bases. For example,
15 between 30 to 45 bases, between 30 to 40 bases, between 30 to 35 bases, between 35 to 40 bases, between 40 to 45 bases, and between 45 to 50 bases. While the above-referenced primer lengths are indicated, a recombinase and/or single stranded binding protein with an optimum primer length of less than 30 bases is also possible and envisioned.

20 The primers used in RPA are preferably DNA although PNA, and RNA are also suitable for use as primers. It is noted that in fact, in natural DNA replication, DNA polymerases elongate genomic DNA by extension from RNA primers.

Primers may be synthesized according to standard techniques. Modified bases and/or linker backbone chemistries may be desirable and functional in some cases.
25 Additionally oligonucleotides may be modified at their ends, either 5' or 3', with groups that serve various purposes *e.g.* fluorescent groups, quenchers, protecting (blocking) groups (reversible or not), magnetic tags, proteins etc. In some cases single stranded oligonucleotides may be used for strand invasion, in others only partly single stranded nucleic acids may be used, the 5' stretch of sequence of an invading nucleic acid being
30 already hybridized to an oligonucleotide.

The primers may comprise a 5' region that is not homologous to the target nucleic acid. It should be noted that amplification may be achieved even if the primers are not completely complementary to the target nucleic acid. The primers may be non-complementary by having additional sequences at their 5' ends. These additional
5 sequences may be, for example, the sequence for a restriction endonuclease recognition site or the sequence that is complementary to a sequencing primer. The restriction endonuclease recognition site may be useful for subsequent cleavage of the amplified sequence. The use of restriction endonuclease that cleaves nucleic acid outside the restriction endonuclease recognition site is also contemplated. The sequence that is
10 complementary for a sequencing primer may allow rapid DNA sequencing of the amplified product using commercially available primers or commercially available sequencing apparatus.

Software to design oligonucleotides for use in *in vitro* DNA synthesis reactions is well established, particularly for use in PCR. The considerations for the RPA method
15 are similar and include the optimisation of the melting temperature of the oligonucleotide, avoidance of hairpin formation within an oligonucleotide and selection against complementarity with other oligonucleotides present in a given reaction. It is therefore important to design oligonucleotide primer pairs to avoid undesirable side reactions.

20 Besides optimising oligonucleotide sequence design there are additional approaches to reduce or eliminate primer dimer formation. As noted elsewhere herein, reaction noise can be reduced by utilising polymerases lacking 3'-5' exonuclease activity. This suggests mispriming may result from oligonucleotides that have been shortened by the 3'-5' exonuclease activity of polymerases. Consequently 3'-5'
25 exonuclease editing activity, pyrophosphorylysis, or any other similar editing activity can be a source of noise. In addition to using polymerases lacking exonuclease activity and the removal of pyrophosphate with pyrophosphatase, use of synthetic oligonucleotides with a non-hydrolysable backbone at the ultimate and/or penultimate link may be beneficial to reduce reaction noise. Alternative backbones could be
30 selected from the considerable range of chemistries available such as phosphorothiorate, morpholino, locked nucleic acid, or peptide nucleic acid.

Reagents for use in RPA reactions

Reagents for use in RPA methods, including those of the present invention, are
5 outlined below.

dNTPs

dNTPs, for example dATP, dGTP, dCTP, and dTTP, and derivatives and
10 analogs thereof, may be added to an RPA reaction. In leading and lagging strand RPA, ATP, GTP, CTP, and UTP may also be included for synthesis of RNA primers. In addition, ddNTPs (ddATP, ddTTP, ddGTP and ddGTP and derivatives and analogs thereof) may be used to generate fragment ladders.

The dNTP may be used at a concentration of between 1 mM to 200 mM of each
15 NTP species.

A mixture of dNTP and ddNTP may be used with ddNTP concentrations at 1/100 to 1/1000 of that of the dNTP (1 mM to 200 mM).

The RPA may be performed in the presence of ATP, a hydrolysable ATP
analog, or another nucleoside triphosphate. The ATP analog may be, for example,
20 dATP, ddATP, or another nucleoside triphosphate analog such as UTP.

Reducing agents

Reducing agents which may be used in the RPA reaction include DTT. The
25 DTT concentration may be between 1 mM and 10 mM, preferably 1 mM.

ATP

ATP or an ATP analog may be used in the RPA reaction.

The ATP or ATP analog may be any of ATP, ATP- γ -S, ATP- β -S, ddATP or a combination thereof. A preferred ATP or ATP analog concentration is between 1 mM and 10 mM, preferably 2.5 mM.

5 System for ATP regeneration

Other components of the RPA reaction may include a system for ATP regeneration (*i.e.* a system to convert ADP to ATP). Such a system may be, for example, phosphocreatine and creatine kinase.

10 An ATP regeneration system permits persistent recombination reactions, as recombinases have an extremely high rate of ATP hydrolysis when bound to nucleic acids. In particular, the UvsX protein has a hydrolysis rate 10-20 times higher than RecA and can consume 200 molecules of ATP per minute per monomer. A number of systems are available. The creatine kinase/phosphocreatine system is preferred. When
15 UvsX is employed the AMP that is produced may be converted into ATP. Chicken myokinase may additionally be used, which converts a molecule of AMP and one of ATP to two molecules of ADP. ADP can then be converted to ATP using the creatine kinase/phosphocreatine system. Poor regeneration of ATP can reduce the reaction rate.

In the RPA methods described herein phosphocreatine is preferably used at a
20 concentration of between 15-25 mM, more preferably 20 mM. Creatine kinase is preferably used at a concentration of between about 0.25-5.0 μ M, more preferably 1 μ M.

Multivalent metal cations

25

The buffer solution in an RPA reaction preferably contains multivalent metal cations. The buffer may contain a functional equivalent of a multivalent metal cation.

The buffer solution in an RPA reaction more preferably contains divalent metal cations. The buffer may contain a functional equivalent of a divalent metal cation.

30

Any suitable multivalent or divalent metal cation or functional equivalent thereof may be used, either as a single agent or combination of agents.

The specific multivalent or divalent metal cation or functional equivalent thereof which achieves optimal results in promoting/enhancing IDR-mediated phase separation in an RPA reaction, as well as the specific concentration of the multivalent/divalent metal cation used may depend upon the specific IDR polypeptide which is used. The optimal multivalent/divalent metal cation or functional equivalent thereof, and the optimal concentration thereof, can be established empirically using routine tests, including RPA reactions themselves and/or the phase separation assays which are described further herein.

The divalent metal cation may be Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} , Ni^{2+} or Cu^{2+} . Any of these cations may be used as a single agent, or any combination of cations may be used. Preferably they are used as single agents. Preferred divalent metal cations are Mg^{2+} , Mn^{2+} and Ca^{2+} . A particularly preferred divalent metal cation is Mg^{2+} .

A preferred concentration range is from 30 to 40 mM, more preferably from 33 to 39 mM.

The buffer may contain Mg^{2+} ions, preferably at the indicated concentrations. More preferably the buffer contains MgOAc at the indicated concentrations.

The buffer may contain Ca^{2+} ions, preferably at the indicated concentrations. More preferably the buffer contains CaCl_2 at the indicated concentrations.

The buffer may contain Mn^{2+} ions, preferably at the indicated concentrations. More preferably the buffer contains MnCl_2 at the indicated concentrations.

Buffers

The buffer solution in an RPA reaction may be a Tris-HCl buffer, a Tris-Acetate buffer, or a combination thereof. The buffers may be present at a concentration of between about 10 mM to about 100 mM. A preferred buffer is a Tris-HCl buffer used at a concentration of between about 20 mM to about 30 mM, most preferably 25 mM. The buffered pH may be between 6.5 to 9.0, preferably pH 8.3.

The buffer may contain potassium acetate between about 5 mM to about 50 mM, preferably between about 10 mM to about 40 mM.

Reaction components

A preferred but non-limiting set of reaction components for an RPA reaction is as follows below.

5		
	Tris HCl pH 8.3	25 mM
	KOAc	7.5 mM
	DTT	1 mM
	ATP	2.5 mM
10	Phosphocreatine	20 mM
	Creatine kinase	1 μM
	dNTPs	1 mM
	Gp32	20 μM
	UvsX	4.8 μM
15	UvsY	8.6 μM
	<i>S. aureus</i> DNA polymerase 1 (Sau)	0.135 μM
	Or <i>B. subtilis</i> DNA polymerase 1 (Bsu)	
	Exonuclease III	0.27 μM
	MgOAc	33 mM
20	Forward primer	0.4 μM
	Reverse primer	0.4 μM
	Exo probe	0.12 μM

RPA reaction conditions

25

RPA reactions, including those of the present invention, may incubated for any suitable length of time.

Any of the RPA reactions may be incubated for between 5 minutes and 16 hours or more, such as between 15 minutes and 3 hours or between 30 minutes and 2 hours.

The incubation may be performed until a desired degree of amplification is achieved. The desired degree of amplification may be 10 fold, 100 fold, 1000 fold, 10,000 fold, 100,000 fold or 1,000,000 fold amplification.

One benefit of RPA is that the reaction may be performed at reduced
 5 temperatures compared to techniques which require thermal cycling, such as PCR. A further advantage of RPA is that the temperature is not critical and precise control, while preferred, is not absolutely necessary. For example, in a field environment, it is sufficient to incubate the RPA reaction at room temperature, or close to body temperature (35°C to 38°C), for example by placing the sample in a body crevice.
 10 Furthermore, the RPA reaction may be performed without temperature induced melting of the template nucleic acid.

Thus any of the RPA reactions may be performed at any suitable temperature.

The RPA reactions may be performed at less than 45°C. The RPA reactions may be performed at less than 40°C. The RPA reactions may be performed at less than
 15 35°C. The RPA reactions may be performed at less than 30°C.

The RPA reactions may be performed at between 20°C and 50°C, between 20°C and 40°C, such as between 20°C and 30°C.

Freeze drying of RPA reaction components

20

One advantage of RPA reactions is that the reagents, with the possible exception of the crowding agent (if used) and buffer, may be freeze dried (*i.e.*, lyophilized) before use. Freeze dried reagents offer the advantage of not requiring refrigeration to maintain activity. For example, a tube of RPA reagents may be stored at room temperature. This
 25 advantage is especially useful in field conditions where access to refrigeration is limited.

RPA reagents may be freeze dried onto the bottom of a tube, or on a bead or any other suitable type of solid support. To perform an RPA reaction the freeze dried reagents are reconstituted in a buffer solution and with a crowding agent (if used), or
 30 simply a buffered solution or water, depending on the composition of the freeze-dried reagents. Then a target nucleic acid, or a sample suspected to contain a target nucleic

acid is added. The reconstitution liquid may also contain the sample nucleic acid. The reconstituted reaction is incubated for a period of time and the amplified nucleic acid, if present, is detected.

In any one of the RPA methods described herein, the reagents that can be freeze dried before use include, at least, the recombinase agent, the recombinase loading protein, the single strand stabilizing agent, the DNA polymerase, the dNTPs or the mixture of dNTPs and ddNTPs, the reducing agent, the ATP or ATP analog, primers and probe.

Stabilizing agents such as trehalose sugar may be included in the freeze dried mixture, for example at 20 mM to 200 mM and most optimally 40 mM to 80 mM in the reconstituted reaction, in order to improve freeze-drying performance and shelf life. If desired, the freeze dried reagents may be stored for 1 day, 1 week, 1 month or 1 year or more before use.

Biochemical reaction reagents, such as RPA reagents, may be freeze dried together with a crowding agent. However, complex inter-related issues may exist which may justify the omission of a crowding agent in a lyophilized mixture. For example, the user may experience difficulty in the effective rehydration of the freeze dried crowding agent, or the user may experience other detrimental effects, including the need for larger lyophilized pellets. Accordingly, there may be advantages in being able to exclude some or all of a crowding agent from freeze dried materials which include among other things reduction of pellet size, shorter cycle times, and easier rehydration. However this has the consequential disadvantage that a crowding agent, if used, would need to be added fresh prior to use after the biochemical reaction mixture was rehydrated and prepared for use. This could be problematic in certain situations, such as for point-of-care use or field use. An advantage of the IDR-based reagents of the present invention is that they would not be expected to exhibit the same drawbacks as crowding agents in a lyophilized setting, and could therefore readily be freeze dried with other biochemical reaction components, thus obviating the need to add fresh additional reagents prior to use.

30

Detection of RPA reaction products

Detection of RPA reaction products may be performed using any suitable method.

5 For example, detection may be performed using electrophoresis on an agarose or PAGE gel followed by ethidium bromide staining.

 Monitoring a RPA reaction may involve, for example, removing a fraction of an RPA, reaction, isolating the unincorporated fraction, and detecting the unincorporated primer. Since the size of an unincorporated primer may be less than 50 bp, less than 40
10 bp, less than 30 bp or less than 25 bp, and the size of the amplified product may be greater than 1Kb, greater than 2 Kb, greater than 5 Kb, or greater than 10 Kb, there is a great size difference between the incorporated and unincorporated primer. The isolation of the unincorporated primer may be performed rapidly using size exclusion chromatography such as, for example, a spin column. If a primer is labeled, a monitor
15 procedure comprising a spin column and a measurement (*e.g.*, fluorescence or radioactivity) can be performed in less than one minute.

 Another alternative for separating elongated primers from unelongated primers involve the use of PAGE. For example, the elongated primer may be separated from the unelongated primer by gel electrophoresis in less than 5 minutes.

20 Yet another alternative for separating elongated primers involves the use of immobilized oligonucleotides. For example oligonucleotides homologous to sequences found uniquely within the amplified DNA sequence can be used to capture nucleic acids produced by primer elongation specifically. These capturing oligonucleotides can be immobilized on a chip, or other substrate. Capture of the elongated oligonucleotides by
25 the capturing oligonucleotides can be performed by RecA protein mediated methods, or by traditional solution hybridizations if necessary.

 The use of fluorescent probes is most commonly used and preferred for the detection of RPA amplification products and have the advantage of providing for real-time detection.

30 These probes are labelled with a fluorophore, such as fluorescein (FAM) and a quencher, such as a Black Hole Quencher, in close proximity to the fluorophore. The

probes have a blocking group at the 3' end to prevent extension from the probe by polymerase. A fluorescent signal is detected when the probe is cleaved and when the quencher and fluorophore are separated, allowing for real-time detection. The probe contains an abasic site, typically tetrahydrofuran (THF) or dR group, and cleavage
 5 occurs at the abasic site, typically by *E. coli* Exonuclease III (cleaving at THF) or *E. coli* fpg (glycolyase/lyase) (cleaving at the dR group).

Kits comprising RPA reaction components

10 The invention also provides a kit for performing an RPA reaction.

The kit may comprise any of the reagents described herein for RPA in any one of the concentrations described above.

The kit may comprise any of the IDR-tagged macromolecules and/or IDR-tagged polypeptides described and defined herein. Preferably the kit further comprises
 15 additional RPA components selected from an RPA recombinase agent, and/or an RPA recombinase loading protein, and/or polymerase, and/or first and second nucleic acid primers, and/or an exonuclease, and/or a buffer, and/or a source of multivalent metal ions, preferably divalent metal cations such as Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} ,.

The reagents of the kit may be freeze dried, in which case the reagents may be
 20 provided in any suitable amount such that when reconstituted the appropriate reagent concentration is achieved.

Polymerase

25 As discussed above, any of the IDR amino acid sequences as described and defined herein may be tagged to any protein component required for the performance of a nucleic acid synthesis reaction.

Any of the IDR amino acid sequences as described and defined herein may be tagged to any protein component required for the performance of a nucleic acid
 30 synthesis reaction wherein a polymerase is used to synthesize a new nucleic acid molecule by extending a primer nucleic acid molecule.

Accordingly, any suitable polymerase may be tagged with an IDR amino acid sequence as described and defined herein. The polymerase may be one which is compatible with and may be used in any reaction used to synthesize a new nucleic acid molecule by extending a primer nucleic acid molecule.

5 The polymerase may be one which is compatible with and may be used in any nucleic acid amplification reaction. The nucleic acid amplification reaction may be a reaction which involves thermal cycling. The nucleic acid amplification reaction may be an isothermal amplification reaction. The nucleic acid amplification reaction may be polymerase chain reaction (PCR), polymerase spiral reaction (PSR), loop mediated
10 isothermal amplification (LAMP), nucleic acid sequence based amplification (NASBA), self-sustained sequence replication (3SR), rolling circle amplification (RCA), strand displacement amplification (SDA), multiple displacement amplification (MDA), ligase chain reaction (LCR), helicase dependant amplification (HDA), ramification amplification method (RAM), recombinase polymerase amplification (RPA),
15 transcription-mediated amplification (TMA) or nicking enzyme amplification reaction (NEAR).

Sequence tags

20 Any of the IDR-macromolecules or any of the IDR-polypeptides involved in biochemical reactions as described herein, including those involved in RPA reactions, may comprise one or more sequence tags. If used, any such sequence tags are preferably attached to a polypeptide as a fusion protein, as described herein. Sequence tags and means for attaching sequence tags to polypeptides are well known in the art.

25 Sequence tags may be short amino acid sequences or larger polypeptides including proteins.

Sequence tags may be attached to the C-terminus of a polypeptide, to the N-terminus of a polypeptide tag, to both the C- and N-terminus of a polypeptide or at any amino acid position along the length of a polypeptide in any combination.

30 Non-limiting examples of suitable amino acid sequence tags include 6-histidine (6X-His; HHHHHH; SEQ ID NO:89), c-myc epitope (EQKLISEEDL; SEQ ID NO:90),

FLAG® octapeptide (DYKDDDDK; SEQ ID NO:91), Protein C (EDQVDPRLIDGK; SEQ ID NO:92), Tag-100 (EETARFQPGYRS; SEQ ID NO:93), V5 epitope (GKPIPNNLLGLDST; SEQ ID NO:94), VSV-G (YTDIEMNRLGK; SEQ ID NO:95), Xpress (DLYDDDDK; SEQ ID NO:96), and hemagglutinin (YPY-DVPDYA; SEQ ID NO:97).

Non-limiting examples of suitable protein tags include β -galactosidase, thioredoxin, His-patch thioredoxin, IgG-binding domain, intein chitin binding domain, T7 gene 10, glutathione-S-transferase (GST), green fluorescent protein (GFP), and maltose binding protein (MBP).

It will be understood by those in the art that sequence tags and protein tags can be used interchangeably, *e.g.*, for purification and/or identification purposes.

Solid phase biochemical reactions

Biochemical reactions performed in processes according to the invention may be performed using solid phase or reversible solid phase techniques. Solid phase reaction systems suitable for performing the processes, uses and methods described herein may comprise a surface. Any suitable surface may be used.

Data described herein demonstrate that biochemical reactions may be performed with IDR-based reagents according to the invention using solid phase techniques in the absence of a crowding agent. One particular example is recombinase polymerase amplification of nucleic acids wherein primers are attached to a solid surface. Any suitable biochemical reaction amenable to performance using solid phase methods may be performed using such methods using processes according to the invention involving any of the IDR-based reagents described and defined herein.

A variety of such solid phase techniques are known in the art and may be used.

Macromolecules, such as polynucleotides, including nucleic acid amplification primers, peptides, haptens, hormones, drugs, etc. may be immobilized to a surface.

Any suitable macromolecular component of a biochemical reaction may be immobilized to a surface, including the IDR-based reagents described and defined herein.

Macromolecules, such as polynucleotides, *e.g.* primers for use in amplification reactions, may be immobilized to surfaces directly or indirectly. For example they may be attached directly to surfaces by chemical bonding. They may be indirectly attached to surfaces via an intermediate surface.

5 A surface may be *e.g.* a planar surface such as glass, a gel-based material, or the surface of a microparticle such as a bead or functionalised quantum dot. The material comprising the surface may itself be bound to a substrate. The substrate may be comprise any suitable material such as glass, plastics or polymeric material.

10 Macromolecules involved in biochemical reactions according to processes of the invention may be immobilized to a gel-based material such as *e.g.* polyacrylamide or a hydrogel, and wherein the gel-based material is itself bound to a supporting substrate such as glass or a plastics or polymeric material.

 Pre-formed polynucleotides, for example, can be immobilized to surfaces by methods commonly employed to create nucleic acid microarrays. For example,
15 polynucleotides may be synthesized and then spotted or printed onto a surface, typically a planar surface. Polynucleotides may be deposited onto surfaces using contact printing techniques. For example, solid or hollow tips or pins may be dipped into solutions comprising pre-formed polynucleotides and contacted with the surface. Alternatively, polynucleotides may be adsorbed onto micro-stamps and then transferred to a surface
20 by physical contact. Non-contact printing techniques include thermic printing or piezoelectric printing wherein sub-nanolitre size microdroplets comprising pre-formed polynucleotides may be ejected from a printing tip using methods similar to those used in inkjet and bubblejet printing.

 Polynucleotides may be synthesised directly on surfaces such as using so-called
25 “on-chip” methods employed to create nucleic acid microarrays. On-chip techniques for generating polynucleotides include photolithography which involves the use of UV light directed through a photolithographic mask to selectively activate a protected nucleotide allowing for the subsequent incorporation of a new protected nucleotide. Cycles of UV-mediated deprotection and coupling of pre-determined nucleotides allows
30 the in situ generation of a polynucleotide having a desired sequence. As an alternative to the use of a photolithographic mask, polynucleotides may be created on surfaces by

the sequential deposition of nucleobases using inkjet printing technology and the use of cycles of coupling, oxidation and deprotection to generate an oligonucleotide having a desired sequence (for a review see Kosuri and Church, Nature Methods, 2014, 11, 499-507).

- 5 Surfaces for the attachment of macromolecules, including polynucleotides, peptides, haptens, hormones, drugs, etc. can be made of any suitable material. Typically a surface may comprise silicon, glass or any suitable polymeric material, such as polystyrene. A surface may comprise a gel surface, such as a polyacrylamide surface or a hydrogel surface. The gel surface may in turn be coupled to or bound to a solid
- 10 support or a substrate, said support or substrate may comprise any suitable material such as silicon, glass or any suitable polymeric material. A surface may comprise a hydrogel material coupled to a polystyrene material.

A surface may be the surface of a microparticle, often referred to as a microsphere or microbead, or simply a bead.

- 15 A surface may comprise a hydrogel material coupled to a polystyrene material in the form of a microbead.

A variety of surface attachment methods and chemistries are available for the immobilization of macromolecules, such as polynucleotides, to surfaces such as microbeads. Surfaces may be functionalised or derivatized to facilitate attachment.

- 20 Such functionalisations are known in the art. For example, a surface may be functionalised with a polyhistidine-tag (hexa histidine-tag, 6xHis-tag, His6 tag or His-tag®), Ni-NTA, streptavidin, biotin, an oligonucleotide, a polynucleotide (such as DNA, RNA, PNA, GNA, TNA or LNA), carboxyl groups, quaternary amine groups, thiol groups, azide groups, alkyne groups, DIBO, lipid, FLAG-tag (FLAG octapeptide),
- 25 polynucleotide binding proteins, peptides, proteins, antibodies or antibody fragments. The surface may be functionalised with a molecule or group which specifically binds to the macromolecule to be immobilized or to another moiety attached to the macromolecule to be immobilized. Covalent immobilization of macromolecules to surfaces is commonly employed. Purely by way of example, a carboxylate-modified
- 30 polystyrene latex surface is suitable for the covalent attachment of *e.g.* amine-terminated proteins, DNA, or other molecules for example via EDAC-mediated

coupling. Other techniques are available. Although macromolecules will typically be attached chemically, they may also be attached to surfaces by indirect means such as via affinity interactions. For example, macromolecules to be immobilized may be functionalised with biotin and bound to surfaces coated with avidin or streptavidin, or
 5 vice-versa.

In any of the processes, uses and methods described and defined herein, a macromolecule may be attached to a surface via one or more covalent bonds. The one or more covalent bonds may be formed between a functional group on the surface and a functional group on the macromolecule. The functional group on the macromolecule
 10 may be *e.g.* an amine group, a thiol group, a thiophosphate group or a thioamide group. The functional group on the surface may be *e.g.* a bromoacetyl group, optionally wherein the bromoacetyl group is provided on a polyacrylamide surface derived using N- (5- bromoacetamidylpentyl) acrylamide (BRAPA).

In any of the processes, uses and methods described and defined herein a
 15 macromolecule may be attached to a surface, either directly or indirectly, via a linker. Any suitable linker which is biocompatible in nature may be used.

A linker may be a linear linker or a branched linker.

A linker may comprise a hydrocarbon chain. A hydrocarbon chain may comprise from 2 to about 2000 or more carbon atoms. The hydrocarbon chain may
 20 comprise an alkylene group, *e.g.* C2 to about 2000 or more alkylene groups. The hydrocarbon chain may have a general formula of $-(CH_2)_n-$ wherein n is from 2 to about 2000 or more. The hydrocarbon chain may be optionally interrupted by one or more ester groups (*i.e.* $-C(O)-O-$) or one or more amide groups (*i.e.* $-C(O)-N(H)-$).

Any linker may be selected from the group comprising polyacrylamide, poly(2-
 25 hydroxyethyl methacrylate), Poly-2-methyl-2-oxazoline (PMOXA), zwitterionic polymers, *e.g.* poly(carboxybetaine methacrylate) (PCBMA), poly[N -(3-sulfopropyl)-N-methacryloxyethyl- N, N dimethyl ammonium betaine] (PSBMA), glycopolymers, and polypeptides.

A linker may comprise oligoethylene glycol-phosphate units having a general
 30 formula of $-[(CH_2-CH_2-O)_n-PO_2^- -O]_m-$ where n is from 1 to about 600 or more and m could be 1-200 or more.

Any of the above-described linkers may be attached at one end of the linker to a macromolecule described herein, and at the other end of the linker to a first functional group wherein the first functional group may provide a covalent attachment to a surface. The first functional group may be *e.g.* an amine group, a thiol group, a thiophosphate group or a thioamide group as further described herein. The surface may be functionalised with a further functional group to provide a covalent bond with the first functional group. The further functional group may be *e.g.* a 2-bromoacetamido group as further described herein. Optionally a bromoacetyl group is provided on a polyacrylamide surface derived using N-(5-bromoacetamidylpentyl) acrylamide (BRAPA). The further functional group on the surface may be a bromoacetyl group, optionally wherein the bromoacetyl group is provided on a polyacrylamide surface derived using N-(5-bromoacetamidylpentyl) acrylamide (BRAPA) and the first functional group may be *e.g.* an amine group, a thiol group, a thiophosphate group or a thioamide group as appropriate. The surface to which polynucleotides are attached may comprise a gel. The surface may comprise a polyacrylamide surface, such as about 2% polyacrylamide, preferably the polyacrylamide surface is coupled to a solid support such as glass.

Microparticles and beads which facilitate reversible immobilization may be used. Solid phase reversible immobilization (SPRI) methods or modified methods are known in the art and may be employed (*e.g.* see DeAngelis M. M. et al. (1995) Solid-Phase Reversible Immobilization for the Isolation of PCR Products, Nucleic Acids Research, 23(22): 4742–4743.).

Surfaces can be provided in the form of *e.g.* paramagnetic beads. Paramagnetic beads can agglomerate under the influence of a magnetic field. For example, paramagnetic surfaces can be provided with chemical groups, *e.g.* carboxyl groups, which in appropriate attachment conditions will act as binding moieties for macromolecules including nucleic acids. Macromolecules can be eluted from such surfaces in appropriate elution conditions. Surfaces of microparticles and beads can be provided with UV-sensitive polycarbonate. Nucleic acids, for example, can be bound to the activated surface in the presence of a suitable immobilization buffer.

5 Microparticles and beads may be allowed to move freely within a reaction solution and then reversibly immobilized, *e.g.* by holding the bead within a microwell or pit etched into a surface. A bead can be localised as part of an array *e.g.* by the use of a unique nucleic acid “barcode” attached to the bead or by the use of colour-coding.

10 The surface may be part of an electrowetting-on-dielectric system (EWOD). EWOD systems provide a dielectric-coated surface which facilitates microfluidic manipulation of very small liquid volumes in the form of microdroplets (*e.g.* see Chou, W-L., et al. (2015) Recent Advances in Applications of Droplet Microfluidics, Micromachines, 6: 1249-1271.). Droplet volumes can programmably be created,
15 moved, partitioned and combined on-chip by electrowetting techniques. Thus electrowetting systems provide alternative means to reversibly immobilize macromolecules to surfaces and/or to manipulate macromolecules immobilized to surfaces.

20 Accordingly, in any one of the processes or uses according to the invention as described or defined herein, the biochemical reaction may be performed in a solid phase reaction system which comprises a surface.

25 In any one of the processes or uses according to the invention as described or defined herein wherein the biochemical reaction is performed in a solid phase reaction system which comprises a surface, any macromolecule required for the performance of the reaction may be attached to the surface. For example, in one such process wherein the biochemical reaction is a process of amplifying a single stranded target nucleic acid molecule or a double stranded target nucleic acid molecule in the *in vitro* reaction system as described herein, the at least one nucleic acid primer, and/or the reaction macromolecule, and/or the IDR-macromolecule and/or the one or more polypeptide co-
30 factors may be attached to the surface.

 In any one of the processes or uses according to the invention as described or defined herein wherein the biochemical reaction is performed in a solid phase reaction system which comprises a surface, the IDR-macromolecule which is required for the performance of the reaction may be attached to the surface.

35 In any one of the processes or uses according to the invention as described or defined herein wherein the biochemical reaction is a recombinase polymerase

amplification process of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, wherein the reaction is performed in a solid phase reaction system which comprises a surface, and wherein the recombinase agent and/or the recombinase loading protein and/or the single strand stabilizing agent and/or the
5 polymerase and/or the exonuclease and/or the first nucleic acid primer and/or the second nucleic acid primer may be attached to the surface. In one such process or use, the first nucleic acid primer or the second nucleic acid primer may be attached to the surface. Alternatively, in other such processes or uses, both the first nucleic acid primer and the second nucleic acid primer may be attached to the surface.

10 In any one of the processes or uses according to the invention as described or defined herein wherein the biochemical reaction is performed in a solid phase reaction system, the surface to which a macromolecule is attached may be a microbead, preferably wherein the microbead comprises a silicon, glass, gel or a polymeric material, such as polystyrene, or any combination thereof.

15 In any one of the processes or uses described herein wherein the biochemical reaction is performed in a solid phase reaction system which comprises a surface and/or a substrate, the surface and/or the substrate may be provided as a flow-cell. Any suitable flow-cell compatible with the biochemical reaction being performed may be used. A suitable flow-cell may comprise a plurality of fluidic channels through which
20 reagents used to perform the biochemical reaction may flow. Any one or more macromolecules used to perform the biochemical reaction may be attached to a surface lining a fluidic channel. A suitable flow-cell may be used to perform a biochemical reaction for amplification of a single stranded or a double stranded target nucleic acid molecule. Sequencing reactions performed using the processes, uses and methods
25 described herein may also be performed using a suitable flow-cell.

EXAMPLES

The following Examples are provided to illustrate the invention but not to limit the invention.

5

Example 1. Recombinase polymerase amplification of *Listeria monocytogenes* gene hly using Gp32 having an IDR tag derived from human Otx1.

Purpose and summary of experiment

10

This experiment was performed to assess the performance of a Gp32 fusion protein preparation containing a tag comprising a histidine-rich amino acid domain sequence found in the intrinsically disordered region (IDR) of the human homeobox protein Otx1.

15

The example demonstrates recombinase polymerase amplification (RPA) of the *Listeria monocytogenes* gene hly across a range of template concentrations using Gp32 C-terminally tagged with the histidine rich intrinsically disordered region (IDR) domain (Otx1) in the absence of a crowding agent.

20 Materials and methods

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG (SEQ ID NO:24). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The fusion protein was designated Gp32-HIS2. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 82 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent, using the indicated

30

copies of a DNA template derived from *Listeria monocytogenes* genomic DNA. The test template was titrated in copy number as indicated in Figure 1.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 5 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M Gp32 fusion, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M Exonuclease III. Reactions were initiated by the addition of template at the given concentration and with 33 mM MgOAc.

The relevant primers and probe are indicated below.

10 Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM]
15 [THF][BHQ-1]ATACAAGGATTGGA (SEQ ID NO:100), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

20 Results and conclusion

As shown in Figure 1 the test template was readily detected at high sensitivity within 7 minutes of initiation of the RPA reaction. Amplicons were detected with as little as 10 copies of target.

25 It was therefore found that amplification in the absence of a crowding agent, such as PEG, occurred efficiently using this Gp32 IDR-tagged fusion protein.

Example 2. Recombinase polymerase amplification of the *Listeria monocytogenes* gene hly using Gp32 having an IDR tag derived from human MafA.

Purpose and summary of experiment

5

This experiment was performed to assess the performance of a Gp32 fusion protein preparation containing a tag comprising a histidine-rich domain sequence found in the intrinsically disordered region (IDR) of the human transcription factor MafA.

10 The example demonstrates recombinase polymerase amplification (RPA) of the *Listeria monocytogenes* gene hly across a range of template concentrations using Gp32 C-terminally tagged with the histidine-rich intrinsically disordered region (IDR) domain (MafA) in the absence of a crowding agent.

Materials and methods

15

The specific amino acid sequence of the IDR domain tag used was SGHHGAHHGAHHHPAAAAAYEAFRGPGFAGGGGADDMGAGHHHGAHHAAHH HHAHHHHHHHHHHHGGAGHGGGAGHH (SEQ ID NO:27). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was
20 purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The fusion protein was designated Gp32-HIS5. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 85 (Table 23).

25 The recombinant phage vB EcoM NBG1 Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent, using the indicated copies of a DNA template derived from *Listeria monocytogenes* genomic DNA. The test template was titrated in copy number as indicated in Figure 2.

30 Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M Gp32 fusion, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M

Exonuclease III. Reactions were initiated by the addition of template at the given concentration and with 33 mM MgOAc.

The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC
5 (SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM]
[THF][BHQ-1]ATACAAGGATTGGA (SEQ ID NO:100), where FAM is fluorescein,
10 THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

Results and conclusion

15 As shown in Figure 2 the test template was readily detected at high sensitivity within 10 minutes of initiation of the RPA reaction. Amplicons were detected with as little as 10 copies of target.

It was therefore found that amplification in the absence of a crowding agent,
20 such as PEG, occurred efficiently using this Gp32 IDR-tagged fusion protein.

Example 3. Recombinase polymerase amplification of the *Listeria monocytogenes* gene hly using Gp32 having an IDR tag derived from *Saccharomyces cerevisiae* Hrp1.

25

Purpose and summary of experiment

This experiment was performed to assess the performance of a Gp32 fusion protein preparation containing a tag comprising an intrinsically disordered region (IDR)
30 of the *Saccharomyces cerevisiae* Hrp1 protein.

The example demonstrates recombinase polymerase amplification (RPA) of the *Listeria monocytogenes* gene hly across a range of template concentrations using Gp32 C-terminally tagged with the sequence comprising an intrinsically disordered region (IDR) of the yeast Hrp1 protein in the absence of a crowding agent.

5

Materials and methods

The specific amino acid sequence of the IDR domain tag used was GGNNGGNNMNRERGNGFNQGDFNQMYQNPMMGGINPMMNPQAMTDYYQ
 10 KMQEYYQQMQ (SEQ ID NO:9). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR tag at the C-terminus of the fusion protein. The fusion protein was
 15 designated Gp32-HRP1. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 79 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent, using the indicated copies of a DNA template derived from *Listeria monocytogenes* genomic DNA. The
 20 test template was titrated in copy number as indicated in Figure 3.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M Gp32 fusion, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M
 25 Exonuclease III. Reactions were initiated by the addition of template at the given concentration and with 33 mM MgOAc.

The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98).

30 Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM]
 [THF][BHQ-1]ATACAAGGATTGGA (SEQ ID NO:100), where FAM is fluorescein,
 THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with
 5 magnetic mixing using a bearing ball.

Results and conclusion

As shown in Figure 3 the test template was readily detected at high sensitivity
 10 within 7 minutes of initiation of the RPA reaction. Amplicons were detected with as
 little as 10 copies of target.

It was therefore found that amplification in the absence of a crowding agent,
 such as PEG, occurred efficiently using this Gp32 IDR-tagged fusion protein.

15 **Example 4. Recombinase polymerase amplification of the *Listeria monocytogenes***
gene hly using Gp32 having an IDR tag derived from *Saccharomyces cerevisiae*
Sup2.

Purpose and summary of experiment

20

This experiment was performed to assess the performance of a Gp32 fusion
 protein preparation containing a tag comprising an intrinsically disordered region (IDR)
 domain of the *Saccharomyces cerevisiae* Sup2 protein.

The example demonstrates recombinase polymerase amplification (RPA) of the
 25 *Listeria monocytogenes* gene hly across a range of template concentrations using Gp32
 C-terminally tagged with the sequence comprising an intrinsically disordered region
 (IDR) domain of the yeast Sup2 protein in the absence of a crowding agent.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was YNPQGGYQQ (SEQ ID NO:19). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR domain tag at the C-terminus of the fusion protein. The fusion protein was designated Gp32-Sup1. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 72 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent, using the indicated copies of a DNA template derived from *Listeria monocytogenes* genomic DNA. The test template was titrated in copy number as indicated in Figure C.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M Gp32 fusion, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M Exonuclease III. Reactions were initiated by the addition of template at the given concentration and with 33 mM MgOAc.

The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM] [THF][BHQ-1]ATACAAGGATTGGA (SEQ ID NO:100), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

Results and conclusion

As shown in Figure 4 the test template was readily detected at high sensitivity within 7 minutes of initiation of the RPA reaction. Amplicons were detected with as little as 10 copies of target.

It was therefore found that amplification in the absence of a crowding agent, such as PEG, occurred efficiently using this Gp32 IDR-tagged fusion protein.

Example 5. Recombinase polymerase amplification of the human apoB gene using Gp32 having IDR tags derived from *Saccharomyces cerevisiae* Sup2.

Purpose and summary of experiment

This experiment was performed to assess the performances of a number of Gp32 fusion protein preparations containing tags comprising an intrinsically disordered region (IDR) domain amino acid sequence of the *Saccharomyces cerevisiae* Sup2 protein. Variable numbers of an IDR domain repeat unit were assessed and a range of concentrations of the fusion proteins were examined.

The example demonstrates recombinase polymerase amplification (RPA) of the human apolipoprotein B (apoB) gene using Gp32 C-terminally tagged with the sequence comprising an intrinsically disordered region (IDR) domain of the yeast Sup2 protein in the absence of a crowding agent.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was YNPQGGYQQ (SEQ ID NO:19). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. Either a single YNPQGGYQQ unit was attached, or two, three or four repeats were attached. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.*

placed after the IDR domain tag at the C-terminus of the fusion protein. The fusion proteins were designated Gp32-Sup2 (two repeats; SEQ ID NO:20), Gp32-Sup3 (three repeats; SEQ ID NO:21) and Gp32-Sup4 (four repeats; SEQ ID NO:22). The full amino acid sequences of the fusion proteins are presented respectively as SEQ ID NO: 73, SEQ ID NO: 74 and SEQ ID NO: 75 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion proteins were then tested, along with Gp32-Sup1, in PEG-free amplification, *i.e.* in the absence of a crowding agent, using a DNA template derived from human genomic DNA.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, Gp32 fusion protein at the concentrations indicated in Figures 5 A to D, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M Exonuclease III. Reactions were initiated by the addition of template and with 33 mM MgOAc. The test template copy number used in each case was 10,000.

The relevant primers and probe are indicated below.

Forward primer: GCAGCTGTATAGCAAATTCCTGTTGAAAGCAG (SEQ ID NO:101).

Reverse primer: TCCTGGCTGTATTCATTGTTGTTAAATTGG (SEQ ID NO:102).

Probe: CACTGATGCT TTTCCTAGACACGAGATGA[FAM-dT]G[THF]C[BHQ1-dT]TGTGGAGCCTTTGT (SEQ ID NO:103), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

Results and conclusion

The results are shown in Figure 5. Figure 5A shows the results using a single IDR domain tag unit. Figures 5 B to D show the results using respectively two, three

and four of the IDR domain tag unit repeats. The test template was detected after about 10 minutes of initiation of the RPA reaction.

It was found that amplification in the absence of a crowding agent, such as PEG, occurred efficiently using these Gp32-IDR tagged fusion proteins. The best
 5 performance was seen with a single IDR domain tag unit and with two IDR domain tag units. Three IDR domain tag units also gave good performance.

**Example 6. Recombinase polymerase amplification of the *Listeria monocytogenes* gene hly using Gp32 having an IDR tag derived from human MafA – comparison
 10 of magnesium ion concentrations.**

Purpose of experiment

This experiment was performed to assess the performance of a Gp32 fusion
 15 protein preparation containing a tag comprising a histidine-rich domain sequence found in the intrinsically disordered region (IDR) of the human transcription factor MafA. The experiment assessed performance across a range of magnesium concentrations.

The example demonstrates recombinase polymerase amplification (RPA) of the *Listeria monocytogenes* gene hly across a range of magnesium concentrations using
 20 Gp32 C-terminally tagged with the histidine-rich intrinsically disordered region (IDR) domain (MafA) in the absence of a crowding agent.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was
 25 SGHHGAHHGAHHHPAAAAAYEAFRGPGFAGGGGADDMGAGHHHGAHHAAHH
 HHAHHHHHHHHHHHGGAGHGGGAGHH (SEQ ID NO:27). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography
 30 relying on the histidines naturally present in the IDR domain tag of the fusion protein

under test. The fusion protein was designated Gp32-HIS5. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 85 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent, using the indicated
5 copies of a DNA template derived from *Listeria monocytogenes* genomic DNA. The test template was provided at 10,000 copies per reaction and the magnesium ion concentration was varied from 5.6 mM up to 44.8 mM.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs,
10 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M Gp32 fusion, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M Exonuclease III. Reactions were initiated by the addition of template and the indicated concentration of MgOAc from 5.6 mM to 44.8 mM MgOAc.

The relevant primers and probe are indicated below.

15 Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM]
20 [THF][BHQ-1]ATACAAGGATTGGA (SEQ ID NO:100), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

25 Results and conclusion

It was found that amplification in the absence of a crowding agent such as PEG occurred efficiently using this Gp32 IDR-tagged fusion protein.

As shown in Figure 6, it was found that good amplification occurred using this
30 Gp32 IDR-tagged fusion protein when 28 mM or more magnesium was present. The

optimum concentration in this experiment appeared to be 33.6 mM, and further increases up to 44.8 mM yielded similar time to detection.

Example 7. Effects of phosphocreatine levels on recombinase polymerase amplification of a human ApoB gene fragment.

Purpose and summary of experiment

This experiment was performed to assess the effects of varying the phosphocreatine levels on the performance of a Gp32 fusion protein preparation containing a tag comprising a histidine-rich domain sequence found in the intrinsically disordered region (IDR) of the human homeobox protein Otx1.

The example demonstrates recombinase polymerase amplification (RPA) of a fragment of the human apolipoprotein (apoB) gene using Gp32 C-terminally tagged with the histidine-rich intrinsically disordered region (IDR) domain (Otx1) in the absence of a crowding agent.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG (SEQ ID NO:24). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The fusion protein was designated Gp32-HIS2. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 82 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent. A phosphocreatine titration was carried out using the human apoB assay. The test template was provided at a concentration of 10^4 copies.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, levels of phosphocreatine indicated in the figures, 1 μ M creatine kinase, 1 mM dNTPs, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M Gp32 fusion, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M Exonuclease III. Reactions were initiated by the addition of 10^4 copies template per reaction and with 33 mM MgOAc.

Forward primer: GCAGCTGTATAGCAAATTCCTGTTGAAAGCAG (SEQ ID NO:101).

Reverse primer: TCCTGGCTGTATTCATTGTTGTTAAATTGG (SEQ ID NO:102).

Probe: CACTGATGCT TTTCCTAGACACGAGATGA[FAM-dT]G[THF]C[BHQ1-dT]TGTGGAGCCTTTGT (SEQ ID NO:103), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

Results and conclusion

It was found that amplification occurred in the absence of a crowding agent such as PEG using this Gp32 IDR-tagged fusion protein. As shown in Figures 7 A, B and C, at the standard phosphocreatine concentration used in PEG based RPA (50 mM), little amplification activity was seen within 20 minutes. Reducing phosphocreatine to 20 mM resulted in optimum performance, but good performance was also observed between 15-25 mM, and lower levels of amplification within 20 minutes was also observed at 30-35 mM.

Example 8. Recombinase polymerase amplification of the *Listeria monocytogenes* gene hly using Gp32 having an IDR tag derived from *Saccharomyces cerevisiae* Hrp1 – comparison of salt concentrations.

5

Purpose and summary of experiment

This experiment was performed to assess the performance of a Gp32 fusion protein preparation containing a tag comprising an intrinsically disordered region (IDR) of the *Saccharomyces cerevisiae* Hrp1 protein. The experiment assessed performance across a range of salt concentrations, in this case using potassium acetate.

The example demonstrates recombinase polymerase amplification (RPA) of the *Listeria monocytogenes* gene hly can be optimised across a range of salt concentrations using Gp32 C-terminally tagged with an intrinsically disordered region (IDR) of the *Saccharomyces cerevisiae* Hrp1 protein in the absence of a crowding agent.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was GGNNGGNNMNRGNGNFGNQGFNQMYQNPMMGGYNPMMNPQAMTDYYQ KMQEYYQQMQ (SEQ ID NO:9). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR tag at the C-terminus of the fusion protein. The fusion protein was designated Gp32-HRP1. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 79 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent, using 100 copies of a DNA template derived from *Listeria monocytogenes* genomic DNA. The potassium acetate concentration was varied from 10 mM up to 100 mM.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M Gp32 fusion, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M Exonuclease III. Reactions were initiated by the addition of template and 33 mM MgOAc.

The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM] [THF][BHQ-1]ATACAAGGATTGGA (SEQ ID NO:100), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

Results and conclusion

It was found that amplification in the absence of a crowding agent such as PEG occurred efficiently using this Gp32 IDR-tagged fusion protein.

It was also found that amplification in the absence of a crowding agent using this Gp32 IDR-tagged fusion protein can be optimised across a range of salt concentrations, of which potassium acetate is a representative example.

As shown in Figure 8, it was found that good amplification occurred using this Gp32 IDR-tagged fusion protein when 10 mM or more potassium acetate was present. The optimum concentration range in this experiment appeared to be between 10 to 40 mM. At concentrations above 40 mM less efficient amplification was observed.

Example 9. Recombinase polymerase amplification of a human ApoB gene fragment using Gp32 having an IDR tag derived from *Saccharomyces cerevisiae* Sup2 – synergistic effect with crowding agent.

5

Purpose and summary of experiment

This experiment was performed to assess the effect of low concentrations of a crowding agent, in this case PEG, on the reaction efficiency of a Gp32 fusion protein preparation containing a histidine-rich sequence found in the intrinsically disordered region of the yeast SUP2 gene, specifically the Sup1 sequence YNPQGGYQQ (SEQ ID NO:19), attached to the C-terminus of phage vB EcoM NBG1 Gp32. The performance of this fusion protein was compared to a Gp32 protein lacking the Sup1 IDR tag in recombinase polymerase amplification of a fragment of the human apolipoprotein (apoB) gene.

It was found that low concentrations of a crowding agent can enhance the reaction efficiency of the Sup1 IDR-tagged Gp32 and that conditions can be achieved where a synergistic effect can be observed.

20 Materials and methods

Gp32-Sup1

The specific amino acid sequence of the IDR domain tag used was YNPQGGYQQ (SEQ ID NO:19). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR domain tag at the C-terminus of the fusion protein. The fusion protein was designated Gp32-Sup1. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 72 (Table 23).

Gp32 (7His)

Phage vB EcoM NBG1 Gp32 was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on a heptahistidine tag placed at the very C-terminus of the protein under test. The fusion protein was designated Gp32 (7His). The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 65 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion proteins were tested in RPA reactions, either in the presence or absence of a crowding agent using a DNA template comprising a fragment of the human apolipoprotein (apoB) gene.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 50 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M Gp32 fusion protein, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase, and 0.27 μ M Exonuclease III. Reactions were initiated by the addition of template and with 33 mM MgOAc. The test template copy number used in each case was 10,000. PEG was added to the final concentrations noted in the relevant figure. The species of PEG used was PEG molecular weight 35,000.

The relevant primers and probe are indicated below.

Forward primer: GCAGCTGTATAGCAAATTCCTGTTGAAAGCAG (SEQ ID NO:101).

Reverse primer: TCCTGGCTGTATTCATTGTTGTTAAATTGG (SEQ ID NO:102).

Probe: CACTGATGCT TTTCCTAGACACGAGATGA[FAM-dT]G[THF]C[BHQ1-dT]TGTGGAGCCTTTGT (SEQ ID NO:103), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher.

Reactions were then incubated at 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

Results and conclusion

The results are shown in Figure 9. Figure 9 shows that when the Gp32-Sup1
5 IDR-tagged fusion protein was tested in the absence of the crowding agent PEG the test template was detected efficiently.

When the Gp32-7His fusion protein without the Sup1 IDR tag was tested in the presence of the crowding agent PEG, between 0.5% to 2%, a small but nevertheless detectable amount of amplified product was observed.

10 When the Gp32-Sup1 IDR-tagged fusion protein was tested in the presence of the crowding agent PEG the test template was detected efficiently. In this case a synergistic effect could be observed, with the amount of amplified product exceeding the combined amount when comparing the amount observed with: (i) the Gp32-Sup1 IDR-tagged fusion protein in the absence of PEG and (ii) the Gp32-7His fusion protein
15 without the Sup1 IDR tag in the absence of PEG (see for example Figure 9 and compare Sup1 1% PEG with Sup1 0% PEG + normal GP32 1% PEG).

These results demonstrate that an enhanced effect on the performance efficiency of a biochemical reaction can be observed when combining an IDR-tagged macromolecular component of the reaction with low concentrations of a crowding
20 agent, and that conditions can be achieved that promote a synergistic effect on reaction efficiency when combining an IDR-tagged macromolecular component of the reaction with low concentrations of a crowding agent.

Example 10. Promotion of phase separation by IDR tags in the presence of 25 multivalent metal cations.

Purpose and summary of experiment

This experiment was performed to assess, in an aqueous *in vitro* biochemical
30 system, the effects of multivalent metal cations on the promotion of phase separation

driven/caused by several Gp32 fusion proteins each having a tag comprising an intrinsically disordered region (IDR) domain amino acid sequence.

The example demonstrates that tags comprising IDR domain amino acid sequence are surprisingly capable of promoting phase separation, and more surprisingly
5 that this effect is enhanced by the presence of multivalent metal cations.

Materials and methods

Gp32-HIS2 fusion protein

10

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG (SEQ ID NO:24). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity
15 chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The full amino acid sequence of the Gp32-HIS2 fusion protein is provided as SEQ ID NO:82 (Table 23).

Gp32-HRP1 fusion protein

20

The specific amino acid sequence of the IDR domain tag used was GGNNGGNNMNRGNGNFGNQGDFNQMYQNPMMGGYNPMMNPQAMTDYYQ
KMQEYYQQMQ (SEQ ID NO:9). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step
25 immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR tag at the C-terminus of the fusion protein. The full amino acid sequence of the Gp32-HRP1 fusion protein is provided as SEQ ID NO:79 (Table 23).

Gp32-Sup1 fusion protein

The specific amino acid sequence of the IDR domain tag used was
 5 YNPQGGYQQ (SEQ ID NO:19). This was attached to the C-terminus of phage vB
 EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step
 immobilised metal (nickel) affinity chromatography relying on an additional
 heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.*
 placed after the IDR domain tag at the C-terminus of the fusion protein. The full amino
 10 acid sequence of the Gp32-Sup1 fusion protein is provided as SEQ ID NO:72 (Table
 23).

Gp32-Fib fusion protein

15 The specific amino acid sequence of the IDR domain tag used was
 PGFSPRGGGFGGRGGFGDRGGRGGRGGFGGGRGRGGGFRGRGR (SEQ ID
 NO:1). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The
 recombinant fusion protein was purified using standard 1-step immobilised metal
 (nickel) affinity chromatography relying on an additional heptahistidine tag placed at
 20 the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR domain tag
 at the C-terminus of the fusion protein. The full amino acid sequence of the Gp32-Fib
 fusion protein is provided as SEQ ID NO:69 (Table 23).

Phase separation assay

25

The methods outlined below apply for all the fusion proteins tested. The
 volumes of the fusion protein solutions used depended upon the protein concentrations
 after purification.

A 50 µl solution was made up in each case, comprising the tagged fusion protein
 30 at a final concentration of 1000 ng/µl and metal ion, in either the acetate or the chloride

form, at the target concentration indicated below and in the relevant figures presented herein.

For the Gp32-HIS2 fusion, protein concentration after purification was 48 mg/ml. 1.04 μ l of this fusion protein was used in each 50 μ l reaction to achieve a final concentration of 1000 ng/ μ l in solution. For the Gp32-HRP1 fusion, protein concentration after purification was 39 mg/ml. 1.28 μ l of this fusion protein was used in each 50 μ l reaction to achieve a final concentration of 1000 ng/ μ l in solution. For the Gp32-Sup1 fusion, protein concentration after purification was 36 mg/ml. 1.4 μ l of this fusion protein was used in each 50 μ l reaction to achieve a final concentration of 1000 ng/ μ l in solution. For the Gp32-Fib fusion, protein concentration after purification was 20.2 mg/ml. 2.48 μ l of this fusion protein was used in each 50 μ l reaction to achieve a final concentration of 1000 ng/ μ l in solution.

In these experiments the divalent metal cation concentration required for detectible phase separation enhancement to occur was tested with representative divalent metal cations: magnesium (MgOAc), manganese (MgCl₂) and calcium (CaCl₂). The acetate forms of manganese and calcium were not used simply due to their known instability in solution. Manganese will oxidise over time in the acetate solution, and calcium acetate seems to support the growth of some bacteria in solution whereas calcium chloride does not.

Following constitution of a mixture comprising water, the IDR-tagged protein and divalent metal cations, the mixture was vortexed, spun down and a 10 μ l sample of the mixture was transferred to a DHC-B01 C-Chip haemocytometer slide. The slide was then imaged under a microscope at x400 magnification. Detectible phase separation was assessed by the formation of spherical-like globular foci/particles which can be identified visually via magnification and counted using the haemocytometer. A globular foci count per unit volume can then be performed. Globular foci counts were performed by counting the number of globular foci which formed in a magnification area of 218 μ m x 175 μ m at x400 magnification. This was done by splitting the magnified image into 20 square segments (4 x 5 of the image), counting the globular foci in one of these segments and then multiplying this number by 20.

Results and conclusions

It was observed that the transition that occurs between just below the minimum
5 detectible phase separating concentration (MPSC) in this assay and just above the
MPSC occurred very suddenly in all of the reactions performed. At just below the
MPSC no detectible phase-separated aqueous particles were observed at all, and the
solution was found to be empty of visually detectible particles (globular foci). Above
the MPSC the transition was very apparent with hundreds of visually detectible
10 particles (globular foci) suddenly forming.

The size of the globular foci varied and was found to be corellated with the IDR
tag and the divalent metal cation used. It was determined that the specific sizes of the
globular foci was unimportant.

Globular foci were present as particle-like structures being broadly spherical in
15 shape. For any given IDR tag and any given divalent metal ion combination an average
diameter of a population of globular foci can readily be determined using standard
methods.

Results using individual fusion proteins are outlined below.

20 Gp32-HIS2 fusion protein

The minimum concentration of magnesium required to enhance the formation of
detectible phase-separated aqueous particles in these conditions was determined to be
10 mM - approximately 600 particles (globular foci) were counted within the field.

25 The minimum concentration of calcium ions required to enhance the formation
of detectible phase-separated aqueous particles in these conditions was determined to be
12 mM – approximately 500 particles (globular foci) were counted within the field.

The minimum concentration of manganese ions required to enhance the
formation of detectible phase-separated aqueous particles in these conditions was
30 determined to be 2 mM - approximately 180 particles (globular foci) were counted
within the field.

Representative magnified images are shown in Figure 10A.

Gp32-HRP1 fusion protein

5 The minimum concentration of magnesium ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was determined to be 16 mM. At this concentration approximately 580 particles (globular foci) were counted within the field.

10 The minimum concentration of calcium ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was determined to be 24 mM. At this concentration approximately 240 particles (globular foci) were counted within the field.

15 The minimum concentration of manganese ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was determined to be 6 mM. At this concentration approximately 260 particles (globular foci) were counted within the field.

Representative magnified images are shown in Figure 10B.

Gp32-Sup1 fusion protein

20

 The minimum concentration of magnesium ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was determined to be 24 mM. At this concentration approximately 280 particles (globular foci) were counted within the field.

25 The minimum concentration of calcium ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was determined to be 32 mM. At this concentration approximately 460 particles (globular foci) were counted within the field.

30 The minimum concentration of manganese ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was

determined to be 4 mM. At this concentration approximately 220 particles (globular foci) were counted within the field.

Representative magnified images are shown in Figure 10C.

5 Gp32-Fib fusion protein

The minimum concentration of magnesium ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was determined to be 500 μ M. At this concentration approximately 340 particles (globular foci) were counted within the field.

The minimum concentration of calcium ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was determined to be 1 mM. At this concentration approximately 500 particles (globular foci) were counted within the field.

15 The minimum concentration of manganese ions required to enhance the formation of detectible phase-separated aqueous particles in these conditions was determined to be 500 μ M. At this concentration approximately 360 particles (globular foci) were counted within the field.

Representative magnified images are shown in Figure 10D.

20 Using these assays it was determined that the functional capability of an IDR or IDR domain to enhance the formation of detectible phase-separated aqueous particles in an *in vitro* biochemical environment when tagged to a protein can be established when 10 or more particles (globular foci) were are formed in a magnification area of 218 μ m x 175 μ m at x400 magnification. The functional capability of an IDR or IDR domain to induce phase separation in an *in vitro* biochemical environment when tagged to a protein can be established when preferably 50 or more particles (globular foci) were are formed in a magnification area of 218 μ m x 175 μ m at x400 magnification, more preferably when 100 or more particles (globular foci) were are formed.

The term “globular foci” as used herein is synonymous with “globule”,
30 “particle” or “globular particle” and these terms can be used interchangeably.

Example 11. Formation of globular foci by IDR tags in the presence of multivalent metal cations.

Purpose and summary of experiment

5

This experiment was performed to assess, in an *in vitro* biochemical reaction system, the effects of multivalent metal cations on the promotion of phase separation driven/caused by several Gp32 fusion proteins each having a tag comprising an intrinsically disordered region (IDR) domain amino acid sequence.

10

The example demonstrates that tags comprising IDR domain amino acid sequences are capable of promoting/enhancing phase separation, and that this effect occurs in the presence of various multivalent metal cations.

Materials and methods

15

Gp32-Fib fusion protein

The specific amino acid sequence of the IDR domain tag used was PGFSPRGGGFGGRGGFGDRGGRGGRGGFGGGRGRGGGFRGRGR (SEQ ID NO:1). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR domain tag at the C-terminus of the fusion protein. The full amino acid sequence of the Gp32-Fib fusion protein is provided as SEQ ID NO:69 (Table 23).

25

Gp32-Sup1 fusion protein

The specific amino acid sequence of the IDR domain tag used was YNPQGGYQQ (SEQ ID NO:19). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step

30

immobilised metal (nickel) affinity chromatography relying on an additional
 heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.*
 placed after the IDR domain tag at the C-terminus of the fusion protein. The full amino
 acid sequence of the Gp32-Sup1 fusion protein is provided as SEQ ID NO:72 (Table
 5 23).

Gp32-HIS2 fusion protein

The specific amino acid sequence of the IDR domain tag used was
 10 AGHHHHHPHAHHPLSQSSGHHHHHHHHHHHQQGYGGSG (SEQ ID NO:24). This
 was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant
 fusion protein was purified using standard 1-step immobilised metal (nickel) affinity
 chromatography relying on the histidines naturally present in the IDR domain tag of the
 fusion protein under test. The full amino acid sequence of the Gp32-HIS2 fusion
 15 protein is provided as SEQ ID NO:82 (Table 23).

Gp32-HRP1 fusion protein

The specific amino acid sequence of the IDR domain tag used was
 20 GGNNGGNNMNRGNGNFGNQGFNQMYQNPMMGYNPMMNPQAMTDYYQ
 KMQEYYQQMQ (SEQ ID NO:9). This was attached to the C-terminus of phage vB
 EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step
 immobilised metal (nickel) affinity chromatography relying on an additional
 heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.*
 25 placed after the IDR tag at the C-terminus of the fusion protein. The full amino acid
 sequence of the Gp32-HRP1 fusion protein is provided as SEQ ID NO:79 (Table 23).

Gp32-HIS5 fusion protein

30 The specific amino acid sequence of the IDR domain tag used was
 SGHHGAHHGAHHHPAAAAAYEAFRGPGFAGGGGADDMGAGHHHGAHHAAHH

HHAAHHHHHHHHHHGGAGHGGGAGHH (SEQ ID NO:27). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The full amino acid sequence of the Gp32-HIS5 fusion protein is presented as SEQ ID NO: 85.

Phase separation assay

The methods outlined below apply for all the fusion proteins tested. The volumes of the fusion protein solutions used depended upon the protein concentrations after purification.

A 50 μ l solution was made up in each case comprising the tagged fusion protein at a final concentration of 1000 ng/ μ l (29.4 μ M) and divalent metal cation. The metal ions tested were Mg^{2+} ($MgOAc$), Mn^{2+} ($MnCl_2$) and Ca^{2+} ($CaCl_2$) and in each case these were used at a final concentration of 20 mM.

Following constitution of a mixture comprising water, the IDR-tagged protein and multivalent metal cations, the mixture was vortexed, spun down and a 10 μ l sample of the mixture was transferred to a DHC-B01 C-Chip haemocytometer slide. The slide was then imaged using bright field microscopy at x400 magnification. Phase separation was assessed by the formation of spherical-like globular foci (particles) which can be identified visually via magnification and counted using the haemocytometer.

Results and conclusions

Representative magnified images are shown in Figures 11 A to E.

For each of the IDR-tagged Gp32 fusion proteins tested, detectible phase separation was observed as determined by the formation of detectible spherical-like phase-separated particles (globular foci). In each case the effect was observed in the presence of Mg^{2+} , Mn^{2+} and Ca^{2+} divalent metal ions.

Thus the ability of multivalent metal ions to induce/enhance phase separation appears to be a general property applicable to a broad range of different IDR tags having quite different amino acid sequences.

5 **Example 12. Formation of globular foci by Gp32 having an IDR tag derived from *Saccharomyces cerevisiae* Hrp1.**

Purpose and summary of experiment

10 This experiment was performed to assess the capability of a Gp32 fusion protein preparation containing a tag comprising an intrinsically disordered region (IDR) of the *Saccharomyces cerevisiae* Hrp1 protein in forming globular foci in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

15 The example demonstrates that the tag comprising the IDR domain amino acid sequences was capable of promoting/enhancing phase separation, as determined by the formation of detectible phase-separated aqueous particles, in the exemplary *in vitro* biochemical reaction environment and in the absence of a crowding agent.

Materials and methods

20 The specific amino acid sequence of the IDR domain tag used was GGNNGGNNMNRGNGNFGNQGDFNQMYQNPMMGGYNPMMNPQAMTDYYQ KMQEYYQQMQ (SEQ ID NO:9). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step
25 immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR tag at the C-terminus of the fusion protein. The fusion protein was designated Gp32-HRP1. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 79 (Table 23).

An exemplary *in vitro* biochemical reaction environment was created to test the effects of the IDR domain sequence tag. In this case, the environment was one which characterises recombinase polymerase amplification reactions.

Reactions were set up according to the following protocol. A reaction mixture
 5 was created with the following components: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.2 μ M forward primer, 0.2 μ M reverse primer, 0.516 μ M probe, 22.6 μ M Gp32-HRP fusion, 8.4 μ M UvsX, 15.3 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase (large subunit), and 0.27 μ M Exonuclease III. Gp32, UvsX, UvsY, polymerase and
 10 Exonuclease III were prepared as a premix before being added in one step to the mixture of primer, buffer, nucleotides and creatine kinase. The total volume was 44 μ l. Once combined, 6 μ l of 280 mM MgOAc was added to the mixture to achieve a final concentration of 33 mM. 10 μ l of the reaction mix was then transferred to a C-Chip haemocytometer slide which was placed on a heated stage set to 39°C before being
 15 observed under the microscope where images were taken under bright field light conditions and fluorescence conditions.

The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC
 (SEQ ID NO:98).

20 Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATG (SEQ ID NO:104) labelled with FAM (fluorescein).

25 Results and conclusion

As shown in Figure 12, the HRP tag attached to Gp32 promoted the formation of many detectible phase-separated aqueous particles (globular foci) which were seen to be dense in oligonucleotide (as detected by the fluorescent labelled probe).

30 A separate experiment was conducted with identical materials and conditions except that the Gp32 protein was tagged only with the heptahistidine sequence and not

with the HRP IDR tag. In these experiments globular foci did not form (data not shown) indicating that the formation of globular foci was driven specifically by the IDR tag, and also that consequently the heptahistidine sequence is not a functional IDR as defined herein.

5 The results demonstrate the functional capability of an IDR domain tag, in this case represented by the *Saccharomyces cerevisiae* HRP1 amino acid sequence tag indicated above, of promoting detectable phase separation in an *in vitro* biochemical reaction environment, in this case represented by the reaction mixture environment which characterises recombinase polymerase amplification reactions, and in the absence
10 of a crowding agent.

Example 13. Formation of globular foci by Gp32 having an IDR tag derived from human Otx1.

15 **Purpose and summary of experiment**

This experiment was performed to assess the capability of a Gp32 fusion protein preparation containing a tag comprising an intrinsically disordered region (IDR) of the human Otx1 protein in forming globular foci in an exemplary *in vitro* biochemical
20 reaction environment in the absence of a crowding agent.

The example demonstrates that the tag comprising the IDR domain amino acid sequences was capable of promoting detectable phase separation in the exemplary *in vitro* biochemical reaction environment and in the absence of a crowding agent.

25 **Materials and methods**

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG (SEQ ID NO:24). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant
30 fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the

fusion protein under test. The fusion protein was designated Gp32-HIS2. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 82 (Table 23).

An exemplary *in vitro* biochemical reaction environment was created to test the effects of the IDR domain sequence tag. In this case the environment was one which characterises recombinase polymerase amplification reactions.

Reactions were set up according to the following protocol. A reaction mixture was created with the following components: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.2 μ M forward primer, 0.2 μ M reverse primer, 0.516 μ M probe, 22.6 μ M Gp32-HIS2 fusion, 8.4 μ M UvsX, 15.3 μ M UvsY, 0.135 μ M *S. aureus* DNA polymerase (large subunit), and 0.27 μ M Exonuclease III. Gp32-His2, UvsX, UvsY, polymerase and Exonuclease III were prepared as a premix before being added in one step to the mixture of primer, buffer, nucleotides and creatine kinase. The total volume was 44 μ l. Once combined, 6 μ l of 280 mM MgOAc was added to the mixture to achieve a final concentration of 33 mM. 10 μ l of the reaction mix was then transferred to a C-Chip haemocytometer slide which was placed on a heated stage set to 39°C before being observed under the microscope where images were taken under bright field light conditions and fluorescence conditions.

The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATG (SEQ ID NO:104) labelled with FAM (fluorescein).

Results and conclusion

As shown in Figure 13, the HIS2 IDR tag attached to Gp32 promoted the formation of many detectible phase-separated aqueous particles (globular foci) which were seen to be dense in oligonucleotide (as detected by the fluorescent labelled probe).

It was noted that the globules appeared smaller in size compared to those which formed when the HRP IDR tag was attached to Gp32 as described further herein.

The results demonstrate the functional capability of an IDR domain tag, in this case represented by the HIS2 amino acid sequence tag indicated above, of promoting
 5 detectable phase separation in an *in vitro* biochemical reaction environment, in this case represented by the reaction mixture environment which characterises recombinase polymerase amplification reactions, and in the absence of a crowding agent.

**Example 14. Effect of multivalent metal cations on the formation of globular foci
 10 by Gp32 having an IDR tag derived from *Saccharomyces cerevisiae* Hrp1.**

Purpose and summary of experiment

This experiment was performed to assess the effects of multivalent metal cations
 15 on the ability of a Gp32 fusion protein preparation containing a tag comprising an intrinsically disordered region (IDR) of the *Saccharomyces cerevisiae* Hrp1 protein in forming globular foci in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

The example demonstrates that the tag comprising the IDR domain amino acid
 20 sequences was capable of promoting/enhancing phase separation, as determined by the formation of detectible phase-separated aqueous particles, in the absence of a crowding agent, that phase separation is enhanced upon the presence of multivalent metal cations and optimised concentrations for promoting phase separation can be determined.

25 Materials and methods

The specific amino acid sequence of the IDR domain tag used was
 GGNNGGNNMNRGNGNFGNQGDFNQMYQNPMMGGYNPMMNPQAMTDYYQ
 KMQEYYQQMQ (SEQ ID NO:9). This was attached to the C-terminus of phage vB
 30 EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (Nickel) affinity chromatography relying on an additional

heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR tag at the C-terminus of the fusion protein. The fusion protein was designated Gp32-HRP1. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 79 (Table 23).

5 An exemplary *in vitro* biochemical reaction environment was created to test the effects of the IDR domain sequence tag in the presence of varying concentrations of divalent metal cation. In this case, the environment was one which characterises recombinase polymerase amplification reactions.

Reactions were set up according to the following protocol. A reaction mixture
10 was created with the following components: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.26 μ M forward primer, 0.26 μ M reverse primer, 0.4 μ M probe, 22.6 μ M Gp32-HRP fusion, 8.4 μ M UvsX and 15.3 μ M UvsY. Gp32, UvsX and UvsY were prepared as a premix before being added in one step to the mixture of primer, buffer, nucleotides and
15 creatine kinase. MgOAc was added to the mixture to achieve the final concentrations as shown in the relevant figures. 10 μ l of the reaction mix was then transferred to a C-Chip haemocytometer slide which was placed on a heated stage set to 39°C before being observed under the microscope where images were taken under bright field light conditions and fluorescence conditions.

20 The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC
(SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ
ID NO:99).

25 Probe: CCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATG (SEQ ID NO:104) labelled with FAM (fluorescein).

Results and conclusion

As shown in Figures 14 A & B, the HRP IDR tag attached to Gp32 promoted the formation of many detectible phase-separated aqueous particles (globular foci) which were seen to be dense in oligonucleotide (as detected by the fluorescent labelled probe).

Globular foci were clearly visible at 22.4 mM MgOAc. Optimum formation of globular foci occurred at 33 mM MgOAc. Some clumping of globules began to be observed at concentrations above 33 mM.

Remarkably, 33 mM MgOAc is the concentration of magnesium at which optimum amplification efficiency is observed in recombinase polymerase amplification (RPA) reactions using IDR-tagged Gp32 in the absence of a crowding agent, as described herein. Thus the efficiency of IDR tag-mediated formation of globular foci surprisingly correlates with the efficiency of an exemplary biochemical reaction in an *in vitro* system in the absence of a crowding agent, in this case amplification in RPA reactions using an IDR-tagged protein as an example test biochemical system.

The results support the surprising conclusion that the performance of an IDR domain sequence tag in driving/increasing the efficiency of a biochemical reaction in the absence of a crowding agent can correlate with the efficiency of phase separation, and this in turn appears to be enhanced by the concentration of multivalent metal cations, or a functional equivalent thereof, included in the system to affect the function of the intrinsically disordered region or domain.

Example 15. Effect of multivalent metal cations on the formation of globular foci by Gp32 having an IDR tag derived from human Otx1.

Purpose and summary of experiment

This experiment was performed to assess the effects of multivalent metal cations on the ability of a Gp32 fusion protein having a tag comprising an intrinsically disordered region (IDR) of the human Otx1 protein in forming particles/globular foci in

an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

The example demonstrates that the tag comprising the IDR domain amino acid sequences was capable of promoting detectible phase separation in the exemplary *in vitro* biochemical reaction environment and in the absence of a crowding agent, that detectible phase separation is enhanced by the presence of multivalent metal cations and optimised concentrations for promoting detectible phase separation can be determined.

Materials and methods

10

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG (SEQ ID NO:24). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The fusion protein was designated Gp32-HIS2. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 82 (Table 23).

An exemplary *in vitro* biochemical reaction environment was created to test the effects of the IDR domain sequence tag in the presence of varying concentrations of divalent metal cation. In this case, the environment was one which characterises recombinase polymerase amplification reactions.

Reactions were set up according to the following protocol. A reaction mixture was created with the following components: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.26 μ M forward primer, 0.26 μ M reverse primer, 0.4 μ M probe, 20 μ M Gp32-HIS2 fusion, 8.4 μ M UvsX and 8.6 μ M UvsY. Gp32, UvsX and UvsY, were prepared as a premix before being added in one step to the mixture of primer, buffer, nucleotides and creatine kinase. MgOAc was added to the mixture to achieve the final concentrations as shown in the relevant figures. 10 μ l of the reaction mix was then transferred to a C-Chip haemocytometer slide which was placed on a heated stage set to 39°C before being

observed under the microscope where images were taken under bright field light conditions and fluorescence conditions.

The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC
5 (SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATG (SEQ ID NO:104) labelled with FAM (fluorescein).

10

Results and conclusion

As shown in Figures 15 A and B, the HIS2 tag attached to Gp32 promoted the formation of many globular foci which were seen to be dense in oligonucleotide (as
15 detected by the fluorescent labelled probe).

Globular foci were clearly visible at 22.4 mM MgOAc. Optimum formation of globular foci occurred between 33 and 39 mM MgOAc. Some clumping of globules began to be observed at concentrations above 39mM.

Remarkably, 33 mM to 39 mM MgOAc is the concentration of magnesium at
20 which optimum amplification efficiency is observed in recombinase polymerase amplification (RPA) reactions using IDR-tagged Gp32 in the absence of a crowding agent, as described herein. Thus the efficiency of IDR tag-mediated formation of globular foci surprisingly correlates with the efficiency of an exemplary biochemical reaction in an *in vitro* system in the absence of a crowding agent, in this case
25 amplification in RPA reactions using an IDR-tagged protein as an example test biochemical system.

The results support the surprising conclusion that the performance of an IDR domain sequence tag in driving/increasing the efficiency of a biochemical reaction in the absence of a crowding agent can correlate with the efficiency of phase separation,
30 and this in turn appears to be enhanced by the concentration of multivalent metal

cations, or a functional equivalent thereof, included in the system to affect the function of the intrinsically disordered region or domain.

Example 16. Effect of magnesium concentration on the formation of globular foci by Gp32 having an IDR tag derived from *Saccharomyces cerevisiae* Hrp1.

Purpose and summary of experiment

This experiment was performed to assess the effects of magnesium ions on the ability of a Gp32 fusion protein preparation containing a tag comprising an intrinsically disordered region (IDR) of the *Saccharomyces cerevisiae* Hrp1 protein in forming globular foci in an exemplary *in vitro* biochemical reaction environment in the absence of a crowding agent.

The example demonstrates that the tag comprising the IDR domain amino acid sequences was capable of promoting/enhancing phase separation, as determined by the formation of detectible phase-separated aqueous particles, in the exemplary *in vitro* biochemical reaction environment and in the absence of a crowding agent, that phase separation is dependent upon the presence of magnesium ions and that all protein components of the reaction mixture were found to be associated with the phase-separated particles and not with the bulk phase.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was
 GGNNGGNNMNRGNGNFGNQGFNQMYQNPMMGGYNPMMNPQAMTDYYQ
 KMQEYYQQMQ (SEQ ID NO:9). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.*
 placed after the IDR tag at the C-terminus of the fusion protein. The fusion protein was

designated Gp32-HRP1. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 79 (Table 23).

5 An exemplary *in vitro* biochemical reaction environment was created to test the effects of the IDR domain sequence tag either in the presence or absence of magnesium ions.

Reactions were set up according to the following protocol. A 1ml reaction mixture was created with the following components: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.4 μ M probe, 20.26 μ M Gp32-HRP
10 fusion, 5 μ M UvsX, 8.67 μ M UvsY, 0.127 μ M *S. aureus* DNA polymerase (large subunit) and either 0 mM or 33.6 mM MgOAc.

The relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98).

15 Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

Probe: CCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATG (SEQ ID NO:104) labelled with FAM (fluorescein).

Photographs were taken of the completed mixtures.

20 Mixtures were spun at 2,000 rcf for 1 minute. Supernatants were removed from the MgOAc mixture. In the mixture with 33.6 mM MgOAc a small pellet was left behind which was assumed to be composed of phase-separated globules/particles. No similar pellet was seen in mixtures without MgOAc. 10 μ l of 1% SDS solution was added to the pellet for solubilisation. The pellet was estimated to be 4.5 μ l in volume
25 making an estimated total volume of 14.5 μ l. 1 μ l of each sample was analysed by SDS-PAGE.

Results and conclusion

30 As shown in Figure 16A, addition of magnesium acetate to the 1ml of RPA mixture caused the mixture to become opaque. This was not observed in the absence of

magnesium acetate. This opaque effect was the same effect as seen in equivalent smaller reactions, and when globular foci/phase-separated particles had been observed to form with typical diameters estimated to be in the range of 2-3 microns by microscopy, and typically at about 200-400 particles per nanolitre. When subject to centrifugation these opaque mixtures cleared and a pellet or lower phase was seen to form at the bottom of the tube which was assumed to be a mass of particles forced together into a single volume (Figure 16B). The estimated volume of this pelleted fraction was approximately 4 μ l, which is roughly the predicted total volume of particles anticipated to form assuming an average particle diameter of 3 μ m (hence a volume of approximately 13 femtoliters) and an abundance of about 400 particles per nanolitre, 400,000 particles per microliter based on haemocytometer/microscope field calculations, which generates an estimated volume of around 5nl of particles per microliter of mixture, and thus about 5 microliters per ml of mixture.

Analysis of one microliter of the bulk mixture (or cleared phase) before and after the addition of magnesium acetate shows that before addition the various proteins can be identified as expected in the clear liquid – Gp32 being the most prominent protein by mass. Following condensation and clearing only trace amounts of protein can be found in the supernatant, while the pellet is vastly enriched in all proteins added to the RPA mixture (Figure 16C). By deduction it can be assumed that roughly a 200-fold concentration of reactants had been achieved and that total protein was at a concentration of approximately 200 μ g/ μ l.

The results demonstrate that all protein components of the RPA reaction mixture, *i.e.* creatine kinase, Gp32-HRP fusion, UvsX, UvsY and polymerase are associated with the phase-separated particles and not with the bulk phase.

25

Example 17. Demonstration of the essential nature of an amino acid sequence comprising an intrinsically disordered region in increasing the efficiency of a biochemical reaction.

5 Purpose and summary of experiment

This experiment was performed to assess the performance of a Gp32 protein lacking a tag comprising an amino acid sequence comprising an intrinsically disordered region (IDR) in an exemplary *in vitro* biochemical reaction environment either in the
10 presence or absence of a crowding agent.

The example demonstrates that in the absence of a tag comprising the IDR domain amino acid sequences, Gp32 was unable in the absence of a crowding agent to mediate recombinase polymerase amplification efficiently and to a point at which detection was made in this assay system within the period of analysis. By comparison
15 with other Examples described above, such as Examples 1 to 5, these data establish that the tag comprising the IDR domain amino acid sequences is essential in increasing the efficiency of the biochemical reaction in the absence of a crowding agent.

Materials and methods

20

The phage vB EcoM NBG1 Gp32 protein was purified in its native form lacking any form of exogenous IDR tag or histidine tag. The protein was purified using heparin resin and eluted with a NaCl step gradient. Native Gp32 protein from the 400mM NaCl fraction was subjected to testing.

25 An exemplary *in vitro* biochemical reaction environment was created to test the effects of the Native Gp32 protein either in the presence or absence of crowding agent.

Reactions were set up according to the following protocols.

PEG-free reaction mixtures were created with the following components: 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM
30 phosphocreatine, 1 μM creatine kinase, 1 mM dNTPs, 0.4 μM forward primer, 0.4 μM reverse primer, 0.12 μM probe, 20 μM native Gp32 protein, 4.8 μM UvsX, 8.6 μM

UvsY, 0.135 μ M *S. aureus* DNA polymerase (large subunit) and 0.27 μ M Exonuclease III.

PEG-based reaction mixtures were created with the following components: 50 mM Tris HCl pH 8.3, 100 mM KOAc, 1 mM DTT, 2.5 mM ATP, 50 mM
 5 phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 0.4 μ M forward primer, 0.4 μ M reverse primer, 0.12 μ M probe, 20 μ M native Gp32 protein, 4.8 μ M UvsX, 8.6 μ M UvsY, 0.27 μ M *S. aureus* DNA polymerase (large subunit), 0.27 μ M Exonuclease III and PEG at a final concentration as shown in the relevant figure. The species of PEG used was PEG molecular weight 35,000.

10 In all reactions the relevant primers and probe are indicated below.

Forward primer: CGCCTGCAAGTCCTAAGACGCCAATCGAAAAGAAAC (SEQ ID NO:98).

Reverse primer: CTGCATCTCCGTGGTATACTAATACATTGTTTTTA (SEQ ID NO:99).

15 Probe: CGAAAAGAAACACGCGGATGAAATCGATAAG[FAM]
 [THF][BHQ-1]ATACAAGGATTGGA (SEQ ID NO:100), where FAM is fluorescein, THF is tetrahydrofuran and BHQ is Black Hole Quencher.

All reactions were initiated by the addition of 33mM MgOAc and 100 copies of a DNA template derived from *Listeria* genomic DNA. Reactions were then incubated at
 20 39°C and placed in a fluorometer with magnetic mixing using a bearing ball.

Results and conclusion

As shown in Figure 17, rapid amplification with native Gp32 protein was
 25 observed in the presence of 5.5% PEG. However, no amplification was observed with native Gp32 protein in the absence of PEG.

In other Examples described above, such as Examples 1 to 5, Gp32-mediated amplification was observed in the absence of PEG only when the Gp32 protein was tagged with an amino acid sequence comprising an intrinsically disordered region
 30 (IDR).

Accordingly, and taken together with data presented in other Examples described herein, these data establish that the tag comprising the amino acid sequence comprising the intrinsically disordered region (IDR) applied to a protein component essential for the function of an *in vitro* biochemical reaction is able to bypass the requirement for a crowding agent in the reaction and increases the efficiency of the biochemical reaction compared to the efficiency observed in the absence of the IDR tag sequence.

Example 18. Recombinase polymerase amplification on a solid surface using Gp32 having an IDR tag derived from human Otx1.

Purpose and summary of experiment

This experiment was performed to assess the performance of a Gp32 fusion protein preparation containing a tag comprising a histidine-rich amino acid domain sequence found in the intrinsically disordered region (IDR) of the human homeobox protein Otx1.

The example demonstrates recombinase polymerase amplification (RPA) of an artificial nucleic acid template on a solid surface using Gp32 C-terminally tagged with the histidine rich intrinsically disordered region (IDR) domain (Otx1) in the absence of a crowding agent both in real-time and end-point assays.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHHQGYGGSG (SEQ ID NO:24). This was attached to the C-terminus of phage vB EcoM NBG1 Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The fusion protein was designated Gp32-HIS2. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 82 (Table 23).

The recombinant phage vB EcoM NBG1 Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent, on a solid surface. The tests were performed using two oligonucleotide primers attached to the surface of beads in differing proportions. Amplification was detected by fluorescence, either in real-time using a cleavable quenched fluorescent probe or by end-point detection of annealed fluorescent probes. In both real-time and end-point RPA reactions the beads were the same. Beads were sourced from Bangs Laboratories, Inc. (<https://www.bangslabs.com/>) and had a polystyrene core that was carboxylated and had a hydrogel grown on it to which oligonucleotides were covalently attached.

Real-time RPA reactions

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 μ M creatine kinase, 1 mM dNTPs, 120nM probe, 20 μ M Gp32 fusion, 4.9 μ M UvsX, 7.6 μ M UvsY, 0.146 μ M *S. aureus* DNA polymerase and 0.34 μ M Exonuclease III. The reaction mixture also included 800,000 beads/ μ l, each bead having approximately 750,000 oligonucleotide primers per bead consisting of a mix of PA30 forward primer and PB30 reverse primer.

Reactions were initiated by the addition of 33.6 mM MgOAc and an artificial DNA template called TF1L at a final concentration of 800,000 template copies per μ l reaction mixture.

The relevant primers, probe and template are indicated below.

PA30 forward primer: CCATCTCATCCCTGCGTGTCTCCGACTCAG (SEQ ID NO:105).

PB30 reverse primer: CCTATCCCCTGTGTGCCTTGGCAGTCTCAG (SEQ ID NO:106).

Probe: AGCAGAAGCAATACCGCCAGCAATAGCA[dT-FAM]G[THF]G[dT-Quencher]AGAGCGAGCTGCC (SEQ ID NO:107), where FAM is fluorescein, THF is tetrahydrofuran and Quencher is Black Hole Quencher.

TF1L template sequence:
CCATCTCATCCCTGCGTGTCTCCGACTCAGTGTTTTAGGGTCCCCGGGGTTA

AAAGGTTTCGAACTCAACAGCTGTCTGGCAGCTCGCTCTACGCATGCTATTG
CTGGCGGTATTGCTTCTGCTCTTGCTGGTGGCGCCATGTCTAAATTGTTTGG
AGCTGAGACTGCCAAGGCACACAGGGGATAGG (SEQ ID NO:108).

Reactions were then incubated for 30 minutes at 39°C in a T8 fluorometer and
5 fluorescence in the FAM channel recorded.

Figure 18A is a cartoon depicting a reaction mixture set up for real-time
amplification using dual-primer beads. Figure 18B is a cartoon depicting amplified
products in the real-time reaction.

10 ***End-point RPA reactions***

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1
mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1 µM creatine kinase, 1 mM dNTPs,
20 µM Gp32 fusion, 4.9 µM UvsX, 7.6 µM UvsY and 0.146 µM *S. aureus* DNA
15 polymerase. The reaction mixture also included 800,000 beads/µl, each bead having
approximately 750,000 oligonucleotide primers per bead consisting of a mix of PA30
forward primer and PB30 reverse primer.

Reactions were initiated by the addition of 33.6 mM MgOAc and an artificial
DNA template called TF1L at a final concentration of 800,000 template copies per µl
20 reaction mixture.

The relevant primers and template are indicated below.

PA30 forward primer: CCATCTCATCCCTGCGTGTCTCCGACTCAG (SEQ
ID NO:105).

PB30 reverse primer: CCTATCCCCTGTGTGCCTTGGCAGTCTCAG (SEQ
25 ID NO:106).

TF1L template sequence:

CCATCTCATCCCTGCGTGTCTCCGACTCAGTGTTTTAGGGTCCCCGGGGTTA
AAAGGTTTCGAACTCAACAGCTGTCTGGCAGCTCGCTCTACGCATGCTATTG
CTGGCGGTATTGCTTCTGCTCTTGCTGGTGGCGCCATGTCTAAATTGTTTGG
30 AGCTGAGACTGCCAAGGCACACAGGGGATAGG (SEQ ID NO:108).

Reactions were then incubated for 30 minutes at 39°C and then stopped by the addition of Sodium Dodecyl Sulfate (SDS) to 1% final concentration and heating to 65°C for 10 minutes to denature the proteins.

SDS was removed by diluting tenfold with water, vortexing, centrifuging for 15 minutes at ~18,000g and then removing the supernatant. Beads were resuspended in TE pH 8.0, 0.05% Triton X-100 buffer to give approximately 800,000 beads/μl.

Two fluorescent oligonucleotide probes (PB30' probe (ROX-5'-CTGAGACTGCCAAGGCACACAGGGGATAGG; SEQ ID NO:109) and TF1L probe (FAM-5'-GGTTTCGAACTCAACAGCTG; SEQ ID NO:110), where ROX is carboxyrhodamine and FAM is fluorescein) were hybridised to beads in TE pH 8.0, 0.05% Triton X-100, 100 mM NaCl buffer, with both probes at a final concentration of 1 μM and 80,000 beads/μl. Hybridisation was performed by heating to 95°C for 2 minutes, then cooling to 25°C at 0.1°C/second. A positive control was run using beads that already had TF1L amplicon attached. Beads were then cleaned to remove unhybridized probes by diluting the hybridisation mixture sixfold in TE pH 8.0, 0.05% Triton X-100 buffer and centrifuging for 15 minutes at approximately 18,000g and then as much supernatant as possible was removed. Beads were resuspended in TE pH 8.0, 0.05% Triton X-100 buffer. Reactions were then incubated for 5 minutes at 39°C in a T8 fluorometer (FAM level set to 17%, ROX level set to 8%) and fluorescence in the FAM and ROX channels were recorded.

Figure 18C is a cartoon depicting amplification characterisation in the end-point reaction.

Results

Real-time RPA reactions

Figure 18D shows real-time fluorescence detection of TF1L amplicon using the specific exonuclease cleaved probe. PA30 primer percentages specified in the figure denote the percentage of bead-bound oligonucleotide that are PA30 oligonucleotide, with the remainder of the bead-bound oligonucleotide being PB30 oligonucleotide.

Amplification is detected when all PA30 and PB30 oligonucleotide primers are bead-bound, as well as when PA30 and PB30 oligonucleotide primers are in liquid phase or when PB30 is bead-bound and PA30 is in liquid phase. No amplification was detected when only PB30 was present on the beads and when there was no PA30.

5

End-point RPA reactions

End-point fluorescence detection of TF1L amplicon was observed using probes specific to the PB30 oligonucleotide primer (ROX-labelled, Figure 18E) and TF1L amplicon (FAM-labelled, Figure 18F). Percentages specified in the figure denote the percentage of bead-bound oligonucleotide that are PA30 oligonucleotide, with the remainder of the bead-bound oligonucleotide being PB30 oligonucleotide. The table below shows the levels of fluorescence for each bead type, the ratio of TF1L probe fluorescence to PB30' probe fluorescence and the same ratio normalised to unamplified control beads with TF1L amplicon attached directly to account for background fluorescence caused by imperfect washing.

10
15

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

Conclusion

20

It was found that nucleic acid amplification in the absence of a crowding agent, such as PEG, occurred efficiently using the Gp32-HIS2 fusion protein both in real-time and end-point assays.

Example 19. Identification of amino acid sequences comprising intrinsically disordered regions.

5 The amino acid sequences of phage vB EcoM NBG1 Gp32, T4 UvsY and T4 UvsX were examined via the MetaDisorder software program (MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. Kozlowshi, L. P., *et al.*, BMC Bioinformatics, 2012, 13(1): 111).

10 As shown in Figures 19 A, B and C respectively, the full length amino acid sequences of phage vB EcoM NBG1 Gp32, T4 UvsY and T4 UvsX contain amino acid sequence stretches which score greater than 0.5 when analysed by the algorithm, and thus comprise intrinsically disordered region sequences.

15 The example demonstrates that intrinsically disordered region sequences or domains thereof can readily be identified using standard analytical methods.

Example 20. Comparison of the phase separation promoting activity of RB69 ligase and RB69 ligase having an IDR tag derived from human Otx1.

Purpose and summary of experiment

20

 This experiment was performed to assess the phase separation promoting activity of a ligase enzyme fusion protein preparation containing a tag comprising a histidine-rich amino acid domain sequence found in the intrinsically disordered region (IDR) of the human homeobox protein Otx1 (His2 tag).

25

 The experiment demonstrated that the formation of phase-separated aqueous particles (globular foci) by RB69 ligase-His2 in the absence of a crowding agent was enhanced by Mg^{2+} concentration, whereas the formation of globular foci by RB69 ligase correlated poorly or not at all with Mg^{2+} concentration.

30

Materials and methods

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHHQQGYGGSG (SEQ ID NO:24; Table 1).

- 5
- This was attached to the C-terminus of RB69 DNA ligase. The recombinant fusion protein and IDR-free protein were purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test and a poly-histidine tag at the C-terminus of IDR-free protein. The fusion protein was designated RB69 ligase-His2 and the IDR-
- 10
- free protein was designated RB69 ligase. The full amino acid sequence of the proteins are presented as SEQ ID NO: 111 and SEQ ID NO: 112 respectively in Table 24 below.

Table 24

Name of protein	SEQ ID NO.	Amino acid sequence
RB69 ligase	111	MILDILNQIAAIGSTKTKQEILKKNKDNKLLERVYRLTYARGIQY YIKKWPGPGERSQAYGLLELDDMLDFIEFTLATRKL TGNAAIKEL MGYIADGKPDDVEVLRRVMMRDLEVGASVSIANKVWPGLIQLQ PQMLASAYDEKLITKNIKWPAFAQLKADGARCFAEVRDDGVQF FSRAGNEYHGLTLLADELMEMTKEARERHPNGVLIDGELVYHSF DIKKAVSSGNDLSFLFGDNEESEEVQVADRSTSNGLANKSLQGTI SPKEAEGMVLQAWDYVPLDEVYSDGKIKGQKYDVRFAALENM AEGFKRIEPIENQLVRNLDEAKVVYKKYVDQGLEG IILKNRDSY WENKRSKNLIK FKEVIDIALEVVGYYEHSKDPNKLGGVELV SRC RRITDCGSGFKDTTHKTVDGVKVLIP LDERHDLDRERLMSEAR EGKLIGRIADCECNGWVHSKGREGTVGIFLPIIKGFRFDKTEADSF EDVFGSWDQTGVSGHHHHH
RB69 ligase-His2	112	MILDILNQIAAIGSTKTKQEILKKNKDNKLLERVYRLTYARGIQY YIKKWPGPGERSQAYGLLELDDMLDFIEFTLATRKL TGNAAIKEL MGYIADGKPDDVEVLRRVMMRDLEVGASVSIANKVWPGLIQLQ

		PQMLASAYDEKLITKNIKWPAFAQLKADGARCFAEVRDDGVQF FSRAGNEYHGLTLLADELMEMTKEARERHPNGVLIDGELVYHSF DIKKAVSSGNDLSFLFGDNEESEEVQVADRSTSNGLANKSLQGTI SPKEAEGMVLQAWDYVPLDEVYSDGKIKGQKYDVRFAALENM AEGFKRIEPIENQLVRNLDEAKVVYKKYVDQGLEGILKNRDSY WENKRSKNLIKFKEVIDIALEVVGYYEHSKDPNKLGGVELVSR RRITDCGSGFKDTTHKTVDGVKVLPLDERHDLDRERLMSEAR EGKLIGRIADCECNGWVHSGREGTVGIFLPIIKGFRFDKTEADSF EDVFGSWDQTGVSGHHHHHPHAHHPLSQSSGHHHHHHHHHHHQ GYGGSG
--	--	--

The probe oligo used in the experiments was:
CCGCAATGGTGC ACTCTCAGTACAATCTGCTCTGATG (SEQ ID NO:104)
labelled with FAM (fluorescein).

5 A 50µl solution was made up comprising the ligase at a final concentration of
1mg/ml, 50mM NaCl, 0.4µM FAM-oligo and MgCl₂ at the target concentration
indicated in the relevant figures. 10µl of the reaction mix was then transferred to a C-
Chip haemocytometer slide and images were taken under bright field light conditions
and fluorescence conditions.

10

Results and conclusion

As shown in Figure 20, RB69 ligase-His2 enhanced the formation of many
phase-separated aqueous particles (globular foci) which were seen to be dense in the
15 oligonucleotide probe (as detected by the fluorescent label) in the presence of Mg²⁺.
Untagged RB69 ligase had very little effect in enhancing globular foci formation, even
at 20mM Mg²⁺.

Example 21. Assessment of ligase activity performance of RB69 ligase having an IDR tag derived from human Otx1.

Purpose and summary of experiment

5

This experiment was performed to assess the ligase activity performance of a ligase enzyme fusion protein preparation containing a tag comprising a histidine-rich amino acid domain sequence found in the intrinsically disordered region (IDR) of the human homeobox protein Otx1 (His2 tag).

10

The experiment demonstrated that double-ligated products increased when the concentration of RB69 ligase-His2 was increased.

Materials and methods

15

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG (SEQ ID NO:24; Table 1). This was attached to the C-terminus of RB69 DNA ligase. The recombinant IDR fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The fusion protein was designated RB69 ligase-His2. The full amino acid sequence of the protein is presented in Table 24 above.

20

25

30

The ligation template was a 170bp fragment (Lig170) amplified from a pUC19 vector (New England Biolabs). A 50µl amplification reaction was set up by mixing 25µl DreamTaq Green Master Mix (Thermo Fisher Scientific), 0.2µM Lig170_FW primer, 0.2µM Lig170_RV primer, 1pg pUC19. PCR reactions were performed as follows: 95°C for 2 minutes; 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute; followed by a final extension for 5 minutes at 72°C. Amplification product was run in a 2% agarose gel. The band of the target DNA was excised and purified by a Monarch DNA Gel Extraction Kit (New England Biolabs). DNA was further phosphorylated at the 5' end by T4 polynucleotide kinase (T4 PNK, Thermo Fisher Scientific). A 50µl phosphorylation reaction was set up by mixing 1x reaction

buffer A, 1mM ATP, 1U T4 PNK and DNA from the previous step. The phosphorylation reaction was incubated at 37°C for 30 minutes. 5'-phosphorylated double-stranded DNA was purified by a Monarch PCR & DNA Cleanup Kit (New England Biolabs) and quantified by a Qubit dsDNA HS assay kit (Thermo Fisher Scientific).

The relevant primers and template sequences are indicated below.

Lig170_FW primer: 5'-GAGCGCAACGCAATTAA-3' (SEQ ID NO: 113).

Lig170_RV primer: 5'-ATCCGCTCACAATTCCACAC-3' (SEQ ID NO: 114).

Lig170 template: 5'-

10 GAGCGCAACGCAATTAAATGTGAGTTAGCTCACTCATT
AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG
GAATTGTGAGCGGAT-3' (SEQ ID NO: 115).

Illumina adaptor was prepared by slow annealing two oligos, 1.5µM ILMN_AD_P5 and 1.5µM ILMN_AD_P7rc_IDX01. The annealing process was heating the oligo mixture to 95°C and cooled to 14°C at a rate of 0.1°C/min.

ILMN_AD_P5: 5'-

AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCT-3' (SEQ ID NO: 116).

ILMN_AD_P7rc_IDX01: 5'-PO₄-

20 GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCC
GTCTTCTGCTTG-3' (SEQ ID NO: 117).

RB69 ligase-His2 was 35mg/ml and it was diluted to 1mg/ml as a working stock. A 20µl solution was made up comprising T4 PNK treated Lig170 at a final concentration of 1ng/µl, 187.5nM Illumina adaptor, 5% PEG 35000, 1x T4 DNA Ligase Reaction Buffer (New England Biolabs), and RB69 ligase-His2 at a final concentration of 0.1/0.2/0.3/0.4 mg/ml. Ligation reactions were performed at 16°C for 20 minutes and 65°C for 15 minutes. To be visualised on an agarose gel, 8 parallel reactions were set up for each reaction condition and combined before loading to a 2% agarose gel. Gel image was analysed by ImageJ (National Institutes of Health) and optical densities of bands were plotted by Excel (Microsoft).

Results and conclusion

As shown in Figure 21, double-ligated products increased when concentrations of RB69 ligase-His2 increased. When 0.4mg/ml of RB69 ligase-His2 was used in a 20µl reaction (0.8µg in the figure), 93.5% of template could be ligated at both ends by Illumina adaptors.

Example 22. Comparison of the ligase activity performance of RB69 ligase and RB69 ligase having an IDR tag derived from human Otx1.

Purpose and summary of experiment

This experiment was performed to assess the activity performance of a ligase enzyme fusion protein preparation containing a tag comprising a histidine-rich amino acid domain sequence found in the intrinsically disordered region (IDR) of the human homeobox protein Otx1 (His2 tag).

The experiment demonstrated that the His2 tag could significantly increase the TA ligation efficiency of RB69 ligase compared to the efficiency of untagged RB69 ligase.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG (SEQ ID NO:24; Table 1). This was attached to the C-terminus of RB69 DNA ligase. The recombinant fusion protein and IDR-free protein were purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test and a poly-histidine tag at the C-terminus of the IDR-free protein. The fusion protein was designated RB69 ligase-His2 and the IDR-free protein was designated RB69 ligase. The full amino acid sequence of proteins are presented in Table 24 above.

The ligation template was a 170bp fragment (Lig170) amplified from a pUC19 vector (New England Biolabs). A 50µl amplification reaction was set up by mixing 25µl DreamTaq Green Master Mix (Thermo Fisher Scientific), 0.2µM Lig170_FW primer, 0.2µM Lig170_RV primer, 1pg pUC19. PCR reaction was performed as follows: 95°C for 2 minutes; 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute; followed by a final extension for 5 minutes at 72°C. Amplification product was run in a 2% agarose gel. The band of the target DNA was excised and purified by a Monarch DNA Gel Extraction Kit (New England Biolabs). DNA was further phosphorylated at the 5' end by T4 polynucleotide kinase (T4 PNK, Thermo Fisher Scientific). A 50µl phosphorylation reaction was set up by mixing 1x reaction buffer A, 1mM ATP, 1U T4 PNK and DNA from the previous step. The phosphorylation reaction was incubated at 37°C for 30 minutes. 5'-phosphorated double-stranded DNA was purified by a Monarch PCR & DNA Cleanup Kit (New England Biolabs) and quantified by a Qubit dsDNA HS assay kit (Thermo Fisher Scientific).

The relevant primers and template sequences are indicated below.

Lig170_FW primer: 5'-GAGCGCAACGCAATTAA-3' (SEQ ID NO: 113).

Lig170_RV primer: 5'-ATCCGCTCACAATTCCACAC-3' (SEQ ID NO: 114).

Lig170 template: 5'-

GAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT
AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG
GAATTGTGAGCGGAT-3' (SEQ ID NO: 115).

Illumina adaptor was prepared by slow annealing two oligos, 1.5µM ILMN_AD_P5 and 1.5µM ILMN_AD_P7rc_IDX01. The annealing process was heating the oligo mixture to 95°C and cooled to 14°C at a rate of 0.1°C/min.

ILMN_AD_P5: 5'-

AATGATACGGCGACCAACGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCT-3' (SEQ ID NO: 116).

ILMN_AD_P7rc_IDX01: 5'-PO₄-

GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCC
GTCTTCTGCTTG-3' (SEQ ID NO: 117).

RB69 ligase-His2 was 35mg/ml and RB69 ligase was 27.75mg/ml. They were diluted to 1mg/ml as working stocks. T4 DNA ligase was quantified by Pierce BCA protein assay kit (Thermo Fisher Scientific) and diluted to 1mg/ml as a working stock. A 20µl solution was made up comprising T4 PNK treated Lig170 at a final

5 concentration of 1ng/ul, 187.5nM Illumina adaptor, 5% PEG 35000, 1x T4 DNA Ligase Reaction Buffer (New England Biolabs), and T4 DNA ligase/RB69 ligase/RB69 ligase-His2 at a final concentration of 0.075 mg/ml. Ligation reactions were performed at 16°C for 20 minutes and 65°C for 15 minutes. DNA was purified by 0.8x AMPure XP for PCR beads (Beckman Coulter) according to the manufacturer's instructions.

10 Purified DNA was mixed with 25µl DreamTaq Green Master Mix (Thermo Fisher Scientific), 0.2µM ILMN_P5 primer, 0.2µM ILMN_P7 primer. PCR reactions were performed as follows: 95°C for 2 minutes; 10 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute; followed by a final extension for 5 minutes at 72°C. Amplification product was purified by 1x AMPure XP for PCR beads. Purified DNA

15 was quantified by Qubit dsDNA HS assay kit and amount of DNA was plotted by Excel (Microsoft).

ILMN_P5: 5'-AATGATACGGCGACCACCGA-3' (SEQ ID NO: 118)

ILMN_P7: 5'-CAAGCAGAAGACGGCATACG-3' (SEQ ID NO: 119)

20 Results and conclusion

As shown in Figure 22, only double-ligated products could be amplified. Both RB69 ligase and RB69 ligase-His2 demonstrated significantly increased ligation efficiencies. Especially RB69 ligase-His2 could have up to 3.1 fold enhancement of ligation

25 efficiency.

Example 23. Comparison of the ligase activity performance of NEBNext Ultra II ligase and RB69 ligase having an IDR tag derived from human Otx1.

Purpose and summary of experiment

5

This experiment was performed to assess the activity performance of NEBNext Ultra II Ligation Master Mix compared with a ligase enzyme fusion protein preparation containing a tag comprising a histidine-rich amino acid domain sequence found in the intrinsically disordered region (IDR) of the human homeobox protein Otx1 (His2 tag).

10

The experiment demonstrated that RB69 ligase-His2 had a significantly enhanced ligation efficiency compared to NEBNext Ultra II Ligation Master Mix.

Materials and methods

15

The specific amino acid sequence of the IDR domain tag used was AGHHHHHPHAHHPLSQSSGHHHHHHHHHHQGYGGSG (SEQ ID NO:24). This was attached to the C-terminus of RB69 DNA ligase. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on the histidines naturally present in the IDR domain tag of the fusion protein under test. The fusion protein was designated RB69 ligase-His2. The full amino acid sequence of the protein is presented in Table 24 above.

20

25

The ligation template was a 170bp fragment (Lig170) amplified from a pUC19 vector (New England Biolabs). A 50µl amplification reaction was set up by mixing 25µl DreamTaq Green Master Mix (Thermo Fisher Scientific), 0.2µM Lig170_FW primer, 0.2µM Lig170_RV primer, 1pg pUC19. PCR reaction was performed as follows: 95°C for 2 minutes; 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute; followed by a final extension for 5 minutes at 72°C. Amplification product was run in a 2% agarose gel. The band of the target DNA was excised and purified by a Monarch DNA Gel Extraction Kit (New England Biolabs). DNA was further phosphorylated at the 5' end by T4 polynucleotide kinase (T4 PNK, Thermo Fisher Scientific). A 50µl phosphorylation reaction was set up by mixing 1x reaction

30

buffer A, 1mM ATP, 1U T4 PNK and DNA from the previous step. The phosphorylation reaction was incubated at 37°C for 30 minutes. 5'-phosphorated double-stranded DNA was purified by a Monarch PCR & DNA Cleanup Kit (New England Biolabs) and quantified by a Qubit dsDNA HS assay kit (Thermo Fisher Scientific).

The relevant primers and template sequences are indicated below.

Lig170_FW primer: 5'-GAGCGCAACGCAATTAA-3' (SEQ ID NO: 113).

Lig170_RV primer: 5'-ATCCGCTCACAATTCCACAC-3' (SEQ ID NO: 114).

Lig170 template: 5'-

10 GAGCGCAACGCAATTAAATGTGAGTTAGCTCACTCATT
AGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG
GAATTGTGAGCGGAT-3' (SEQ ID NO: 115).

Illumina adaptor was prepared by slow annealing two oligos, 1.5µM ILMN_AD_P5 and 1.5µM ILMN_AD_P7rc_IDX01. The annealing process was heating the oligo mixture to 95°C and cooled to 14°C at a rate of 0.1°C/min.

ILMN_AD_P5: 5'-

AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCT-3' (SEQ ID NO: 116).

ILMN_AD_P7rc_IDX01: 5'-PO₄-

20 GATCGGAAGAGCACACGTCTGAACTCCAGTCACATCACGATCTCGTATGCC
GTCTTCTGCTTG-3' (SEQ ID NO: 117).

RB69 ligase-His2 was 35mg/ml and was diluted to 1mg/ml as a working stock.

A 93.5µl solution was made up comprising 10ng T4 PNK treated Lig170, 187.5nM Illumina adaptor, 30µl NEBNext Ultra II Ligation Master Mix, and 1µl NEBNext Ligation Enhancer. Ligation reactions were performed at 20°C for 15 minutes. Ligation reactions were terminated by adding 6.5µl of 0.5M EDTA and DNA was purified by 0.8x AMPure XP for PCR beads (Beckman Coulter) according to the manufacturer's instructions. A 93.5µl solution was made up comprising 10ng T4 PNK treated Lig170, 187.5nM Illumina adaptor, 5%/7% PEG 35000, 1x T4 DNA Ligase Reaction Buffer (New England Biolabs), and RB69 ligase-His2 at a final concentration of 0.2/0.3/0.4mg/ml. Ligation reactions were performed at 16°C for 20 minutes.

Ligation reactions were terminated by adding 4.5µl of 0.5M EDTA and 2ul Protease K (New England Biolabs) and incubated at 40°C for 30 minutes. DNA was purified by 0.8x AMPure XP for PCR beads (Beckman Coulter) according to the manufacturer's instructions. Purified DNA was mixed with 25µl DreamTaq Green Master Mix (Thermo Fisher Scientific), 0.2µM ILMN_P5 primer, 0.2µM ILMN_P7 primer. PCR reactions were performed as follows: 95°C for 2 minutes; 10 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 minute; followed by a final extension for 5 minutes at 72°C. Amplification product was purified by 1x AMPure XP for PCR beads. Purified DNA was quantified by Qubit dsDNA HS assay kit and amount of DNA were plotted by Excel (Microsoft). DNA was also analysed by a 2% agarose gel.

ILMN_P5: 5'-AATGATACGGCGACCACCGA-3' (SEQ ID NO: 118).

ILMN_P7: 5'-CAAGCAGAAGACGGCATACG-3' (SEQ ID NO: 119).

Results and conclusion

15

Only double-ligated products could be amplified. As shown in Figure 23, RB69 ligase-His2 could have up to 2.8 fold enhancement of ligation efficiency compared to NEBNext Ultra II Ligation Master Mix.

20 **Example 24. Analysis of the effect of ATP on the phase separation performance of RB69 ligase having an IDR tag derived from human Otx1.**

Purpose and summary of experiment

25

The purpose of this experiment was to analyse the effect of ATP on the ability of a ligase enzyme fusion protein preparation to cause phase separation. The ligase enzyme fusion protein has a tag comprising a histidine-rich amino acid domain sequence found in the intrinsically disordered region (IDR) of the human homeobox protein Otx1 (His2 tag).

30

The experiment demonstrated that ATP significantly enhanced phase separation mediated by the His2 tag.

Materials and methods

The specific amino acid sequence of the IDR domain tag used was
 5 AGHHHHHPHAHHPLSQSSGHHHHHHHHHHHQQGYGGSG (SEQ ID NO:24). This
 was attached to the C-terminus of RB69 DNA ligase. The recombinant IDR fusion
 protein and IDR-free protein were purified using standard 1-step immobilised metal
 (nickel) affinity chromatography relying on the histidines naturally present in the IDR
 domain tag of the fusion protein under test and a poly-histidine tag at the C-terminus of
 10 IDR-free protein. The fusion protein was designated RB69 ligase-His2 and the IDR
 free protein was designated RB69 ligase. The full amino acid sequences of the proteins
 are presented in Table 24 above. FAM-oligo is
 CCGCAATGGTGCACTCTCAGTACAATCTGCTCTGATG (SEQ ID NO:104)
 labelled with FAM (fluorescein).

15 A 50µl solution was made up comprising the ligase at a final concentration of
 1mg/ml, 0.4µM FAM-oligo and 0/20mM MgCl₂ and 0/1mM ATP. 10µl of the reaction
 mix was then transferred to a C-Chip haemocytometer slide and images were taken
 under bright field light conditions and fluorescence conditions.

20 Results and conclusion

As shown in Figure 24, very few phase-separated aqueous particles (globular
 foci) were observed using the IDR-free RB69 ligase. RB69 ligase-His2 significantly
 promoted the formation of many globular foci in the presence of 1mM ATP. After
 25 addition of 20mM Mg²⁺, brighter fluorescence of globular foci was observed indicating
 a further enhancement of globular foci formation and/or that more DNA was forced to
 co-localise in the globular foci.

Example 25. Recombinase polymerase amplification on a solid surface using Gp32 having an IDR tag derived from *Saccharomyces cerevisiae* Hrp1.

Purpose and summary of experiment

5

This experiment was performed to assess the capability of a Gp32 fusion protein preparation containing a tag comprising an intrinsically disordered region (IDR) of the *Saccharomyces cerevisiae* Hrp1 protein in amplification on a solid surface.

10 The example demonstrates recombinase polymerase amplification (RPA) of an artificial nucleic acid template on a solid surface using Gp32 C-terminally tagged with the intrinsically disordered region (IDR) of the *Saccharomyces cerevisiae* Hrp1 protein in the absence of a crowding agent both in real-time and end-point assays.

Materials and methods

15

The specific amino acid sequence of the IDR domain tag used was GGNNGGNNMNRGNGNFGNQGDFNQMYQNPMMGGYNPMMPQAMTDYYQ KMQEYYQQMQ (SEQ ID NO:9). This was attached to the C-terminus of T4 phage Gp32. The recombinant fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on an additional heptahistidine tag placed at the very C-terminus of the fusion protein under test, *i.e.* placed after the IDR tag at the C-terminus of the fusion protein. The fusion protein was designated T4-Gp32-HRP1. The full amino acid sequence of the fusion protein is presented as SEQ ID NO: 120 below

25

MFKRKSTAELAAQMAKLNGNKGFSSEDKGEWKLKLDNAGNGQAVIRFLPSKN
DEQAPFAILVNHGFKKNGKWKYIETCSSTHGDYDSCPVCQYISKNDLYNTDNKE
YSLVKRKTSYWANILVVKDPAAPENEGKVFKYRFGKKIWDKINAMIAVDVEM
GETPVDVTCPWEGANFVLKVQVSGFSNYDESKFLNQSAIPNIDDESFKELFE
30 QMVDLSEMTSKDKFKSFEELNTKFGQVMGTAVMGGAATAAKKADKVADDL
DAFNVDDEFNTKTEDDFMSSSSGSSSSADDTDLDDLNDLGGNNGGNNMNRG

GNFGNQGDFNQMYQNPMMGGYNPMMNPQAMTDYYQKMQEYYQQMQHHH
HHHH

(SEQ ID NO: 120)

- 5 The recombinant T4 phage Gp32 fusion protein was then tested in PEG-free amplification, *i.e.* in the absence of a crowding agent, on a solid surface. The tests were performed using two oligonucleotide primers attached to the surface of beads. Amplification was detected by the incorporation of a fluorescent nucleotide at a nickase site found in the amplicon using an end-point reaction. Beads were sourced from Bangs
- 10 Laboratories, Inc. (<https://www.bangslabs.com/>) and had a polystyrene core that was carboxylated and was grafted with a copolymer to which oligonucleotides were covalently attached. Beads were deposited onto a glass substrate that was patterned into discrete regions using standard microfabrication technologies like photolithography, soft lithography, etching, etc. The resulting regions have characteristics such as being
- 15 hydrophobic or hydrophilic and can attract or repel samples to be analyzed. Grace Bio-Labs FlexWell™ removable incubation chambers were used to divide a single piece of patterned glass into eight reaction chambers measuring 6.5mm x 6.5mm each estimated to contain 12.25 million beads in an ordered array on the surface. Differing amounts of amplification template were added to different reaction chambers such that if all
- 20 template hybridised to primers on beads a reaction chamber would have 0, 5, 10, 20, 40 or 80 copies of template per bead, with the assumption that hybridisation would be much less than 100% efficient. Single stranded DNA template (UP1-UP2' _TF1L template sequence:
- 25 AATGATACGGCGACCACCGTGATCTACACTGTTTTACAACCTCAGCATGGA
AAAAGGTTTCGAACTCAACAGCTGTCTGGCAGCTCGCTCTACGCATGCTATT
GCTGGCGGTATTGCTTCTGCTCTTGCTGGTGGCGCCATGTCTAAATTGTCGA
TACATCTCGTATGCCGTCTTCTGCTTG (SEQ ID NO: 121) was added to reaction chambers in 50 µl buffer (10mM Tris HCl pH 8.0, 1mM EDTA, 0.05% Triton X-100, 100mM NaCl), covered with FlexWell™ SealStrips™ and heated to 50°C for 1 hour to
- 30 allow template to anneal to complementary oligonucleotides on the beads. Excess buffer was then removed and the beads washed twice with TTM buffer (10mM Tris

HCl pH 8.0, 10mM MgCl₂, 0.05% Triton X-100) and then twice with reaction buffer (25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT) to remove any template that had not annealed.

Reactions were set up by mixing 25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT, 2.5 mM ATP, 20 mM phosphocreatine, 1.7μM creatine kinase, 1 mM dNTPs, 6.6 μM Gp32 fusion, 2.7 μM UvsX, 2.7 μM UvsY, 0.22 μM *S. aureus* DNA polymerase and 23mM MgOAc. The reaction mixture also included 290,000 beads/mm², with 0.59 μl reaction mix/mm² each bead having approximately 600,000 oligonucleotide primers per bead consisting of a mix of UP1 forward primer and UP2-18 reverse primer having the sequences below:

UP1 forward primer: AATGATACGGCGACCACCGAGATCTACAC (SEQ ID NO: 122).

UP2-18 reverse primer: CAAGCAGAAGACGGCATA (SEQ ID NO: 123).

Reactions were then incubated for 60 minutes at 43°C and then stopped by washing (adding/removing) twice with STTM buffer (10mM Tris HCl pH 8.0, 10mM MgCl₂, 0.05% Triton X-100 and Sodium Dodecyl Sulfate (SDS) to 1% final concentration) to denature the proteins. The beads were then washed twice by adding/removing TTM buffer (10mM Tris HCl pH 8.0, 1mM EDTA, 0.05% Triton X-100) to remove the SDS.

The beads were then covered in 25 μl 1x CutSmart buffer (NEB – 50mM KOAc, 20mM Tris-acetate, 10mM MgOAc, 100 μg/ml BSA, pH 7.9) with 2.5U nickase Nt. BbvCI. The UP1-UP2' _TF1L template includes a single copy of the recognition site of Nt. BbvCI (CC/TCAGC) which introduces a nick in one strand of the DNA. Beads were heated to 37°C for 45 minutes to ensure any amplicons were nicked. Nicking was stopped by washing (adding/removing) twice with STTM buffer (10mM Tris HCl pH 8.0, 1mM EDTA, 0.05% Triton X-100 and Sodium Dodecyl Sulfate (SDS) to 1% final concentration) to denature the nickase. The beads were then washed twice by adding/removing TTM buffer (10mM Tris HCl pH 8.0, 1mM EDTA, 0.05% Triton X-100) to remove the SDS and twice with reaction buffer (25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT).

The final step in the protocol was to incorporate fluorescently labelled dUTP into amplicons. This was done by submerging the beads in reaction buffer (25 mM Tris HCl pH 8.3, 7.5 mM KOAc, 1 mM DTT) with 0.11 μ M *S. aureus* DNA polymerase, 160pM Aminoallyl-dUTP-XX-ATTO-594 (Jena Bioscience) and 23mM MgOAc and heating to 43°C for 45 minutes. Extension was stopped by washing (adding/removing) twice with STTM buffer (10mM Tris HCl pH 8.0, 1mM EDTA, 0.05% Triton X-100 and Sodium Dodecyl Sulfate (SDS) to 1% final concentration) to denature the nickase. The beads were then washed twice by adding/removing TTM buffer (10mM Tris HCl pH 8.0, 1mM EDTA, 0.05% Triton X-100) to remove the SDS. The Flexwells were then removed and a glass coverslip and small volume of TTM buffer were placed on the glass wafer. The wafer was examined using a fluorescent microscope, with brightfield and fluorescent pictures taken of the same locations.

Results

End-point fluorescence detection of UP1-UP2'-TF1L amplicon was observed using the incorporation of ATTO-594 labelled dUTP into the Nt. Bbv CI nicking site of the amplicon. Figure 25 shows that when no template was added, beads remained dark, but as template increased an increasing proportion of beads were fluorescent and the level of fluorescence increased.

Conclusion

It was found that nucleic acid amplification in the absence of a crowding agent, such as PEG, occurred efficiently using the T4 Gp32-Hrp1 fusion protein on surfaces.

Example 26. Analysis of the effect of Gp32-HRP1 ssDNA binding protein with associated phase separation on Cas12a protein performance.

Purpose and summary of experiment

5

The purpose of this experiment was to analyse the effect of an IDR-tagged Gp32 ssDNA binding protein in the presence of Polyethylene Glycol (PEG) 35K under conditions that promote phase separation on the activity of the Cas12a nuclease protein, in association with a guide RNA, to bind and cut a double-stranded DNA target monitored by fluorescence readout. The Gp32 ssDNA binding protein used has a tag comprising an amino acid domain sequence found in the intrinsically disordered region (IDR) of the yeast HRP protein (Gp32-HRP1). In the presence of this tag and PEG, phase separation occurs substantially in the absence of other factors even at low concentrations of the protein.

15

The duplex nucleic acid target for Cas12a in this case has a 6-FAM/BHQ1 pairing which when cleaved generates a nucleotide fragment containing the 6-FAM label which should melt essentially immediately from the annealed hybrid leading to lead to a measurable fluorescence increase. This template was additionally engineered to interact with Gp32-HRP1 by virtue of an additional single-stranded region to hook.

20

The experiment demonstrates that using Gp32-HRP1 ssDNA binding protein in the presence of PEG35K leads to the formation of phase-separated aqueous particles (globules or globular foci) and concomitantly significantly enhanced the rate that Cas12a cuts its DNA target in an *in vitro* system.

25

Materials and methods

30

The specific amino acid sequence of the IDR domain tag used was GGNNGGNNMNRRRGGNFGNQGDFNQMYQNPMMGGYNPMMNPQAMTDYYQ KMQEYYQQMQ (SEQ ID NO: 9). This was attached to the C-terminus of T4 Gp32 ssDNA binding protein. The recombinant IDR fusion protein was purified using standard 1-step immobilised metal (nickel) affinity chromatography relying on 7

additional histidines appended to the C-terminus of the IDR tag. The fusion protein was designated T4 GP32-HRP1. The full amino acid sequence of the protein is presented as SEQ ID NO: 120.

The guide-RNA sequence is 5'-

- 5 UAAUUUCUACUGUUGUAGAUAAAGUGCUCAUCAUUGGAAAACG-3' (SEQ ID NO: 134).

The double-stranded/single-stranded DNA target was prepared by annealing two oligos, top oligo 5'-

- GAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTATGTATCAAAGCGGCCA
 10 TTTGCGG-3' labelled with FAM (fluorescein) at 5' end (SEQ ID NO: 135) and bottom oligo 5'-AGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC-3' labelled with BHQ-1 (quencher) at 3' end (SEQ ID NO: 136). The annealing process was heating the 1µM oligo mixture to 95°C and cooled to 14°C at a rate of 0.1°C/min. This provides a duplex target site for the cas12a nuclease with the supplied guide RNA but also an
 15 additional 24 single-stranded residues which may interact with Gp32-HRP1 with an anticipated binding of about 3 monomers of protein. In this fashion it was anticipated that much of the annealed target will be forced to locate in phase-separated Gp32-HRP1 globules should they arise. Furthermore the presence of a fluorophore and quencher on either strand which should be in close proximity when annealed, but disperse following
 20 a cut (as the resulting hybrid is only a few nucleotides long) provides a convenient mechanism to assess the rate of cutting. As expected, in a Cas12a -dependent fashion fluorescence changes from generally low levels and increases over time.

EnGen Lba Cas12a protein was purchased from New England Biolabs.

- Solutions were made up which did or did not include Cas12a protein, PEG35K
 25 or T4 GP32 HRP1 protein. The solutions were comprised of: 30mM NaCl, 10mM Tris Acetate pH8.3, 20mM Mg Acetate, 0.1mg/ml BSA, 33.3nM guide RNA, 50nM dsDNA, 5% PEG35K. When included in the reactions, the following components were present at the following concentrations: 33.3nM Cas12a protein, 333ng/µl T4 GP32 HRP1 protein. To assess reaction rate behavior, 30µl of reaction solution was transferred to
 30 0.2ml tubes and assayed using an Axxin T16 fluorescence reader, using a run temperature of 42°C. Independently, 20µl of reaction solution was warmed at 42°C for

approximately 1 minute and then transferred to a C-Chip haemocytometer slide and images were taken under bright field light conditions and fluorescence conditions. These images therefore represented a snapshot of the microscopic state of the system within the first few minutes of the reaction.

5

Results and conclusion

As shown in Figure 26A, without Cas12a protein but in the presence of Gp32-HRP1 and PEG, phase-separated aqueous particles (globular foci) were observed but demonstrated only weak fluorescence when the labelled, but largely quenched, target DNA was present in the presence of Gp32-HRP1 and PEG (at the early timepoint at which the reaction was analysed microscopically). Naturally, no cutting is anticipated absent Cas12a protein as supported by minimal change in fluorescence change over time as shown in the flat character of the fluorescence plot (Figure 26B) albeit some fluorescence is observed, possibly due to incomplete annealing of the fluorophore/quencher probe and consequent background.

When Cas12a protein was present, but in the absence of T4 Gp32 HRP1, no globular foci were observed, indicating the as-expected requirement for T4 Gp32 HRP1 to enable globule formation. The kinetic analysis indicated that target cutting increased steadily over time as assessed up to 10 minutes. Microscopically the overall fluorescence appeared slightly higher than in the Cas12a-minus sample indicating that within a few minutes some annealed probe had been processed consistent with the kinetic study.

However in stark and remarkable contrast, in the presence of Cas12a and T4 Gp32-HRP1 (and PEG), many globular foci were observed and generally much stronger fluorescence was observed throughout the microscope image, suggesting both the need for T4 Gp32-HRP1 for globule formation, but in addition that this lead to more processing (note once processed the small released products would not necessarily be expected to localize any longer to the globules). The kinetic graph was markedly and stunningly different under these conditions as well, and showed very rapid fluorescence accumulation to a peak, around, or before 1 minute (just after the sample was placed

into the reader), and then a plateau for the remainder of the analysis time. This significant enhancement in DNA cutting rate observed in the presence of T4 Gp32-HRP1 is we propose consistent with the notion that the phase-separated particles markedly promoted specific cutting, presumably caused by a co-localisation of the Cas12a protein and its dsDNA target inside the globular foci enabling a greatly increased rate of reaction. In a fashion similar to the amplification system demonstrated herein this indicates that even when only a single system component acts to drive phase separation, other participants may be drawn to that phase leading local high concentrations and massively accelerated kinetics.

It is to be understood that different applications of the disclosed IDR-based methods, processes, macromolecules, polypeptides and uses may be tailored to the specific needs in the art. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments of the invention only, and is not intended to be limiting.

As used in this specification and the appended claims, the singular forms “a”, “an”, and “the” include plural referents unless the content clearly dictates otherwise. Thus, for example, references to an entity such as “a macromolecule”, “a polypeptide”, “a polynucleotide”, “a cell”, “a host cell” and so on, includes two or more such entities.

Terms such as “about” and “approximately” are to be understood as encompassing the relevant figure +/- 10% of the value of the figure, or +/- 5% of the value of the figure unless the content clearly dictates otherwise.

Where a range of numbers is presented as being “between” a lower value and an upper value, the range is to be interpreted as including the upper and lower values. For example, a range of between 22 mM to 50 mM, or between about 22 mM to about 50 mM, should be interpreted as including the values of 22 mM and 50 mM or the values of about 22 mM and about 50 mM.

All publications, patents and patent applications cited herein are hereby incorporated by reference in their entirety.

CLAIMS

1. A process of performing a biochemical reaction in an aqueous *in vitro* reaction system, wherein the biochemical reaction is dependent on the function of at least one reaction macromolecule, optionally at least one reaction polypeptide, the process comprising: introducing at least one IDR-macromolecule into the *in vitro* reaction system under conditions suitable for performing the reaction, wherein the at least one IDR-macromolecule comprises one or more functional intrinsically disordered regions (IDRs), wherein upon introduction of the at least one IDR-macromolecule into the *in vitro* reaction system the efficiency of the biochemical reaction is increased by the at least one IDR-macromolecule; preferably wherein the at least one IDR-macromolecule is at least one IDR-polypeptide.
2. A process according to claim 1, wherein the biochemical reaction is dependent on the function of the at least one IDR-macromolecule, optionally the at least one IDR-polypeptide, wherein upon introduction into the system the at least one IDR-macromolecule or the at least one IDR-polypeptide performs its reaction function in the biochemical reaction and increases the efficiency of the reaction in the system.
3. A process according to claim 1 or claim 2, further comprising maintaining the IDR-macromolecule or the IDR-polypeptide in the system to cause liquid-liquid demixing and the formation of a plurality of phase-separated aqueous compartments within the system by the IDR-macromolecule or the IDR-polypeptide, thereby increasing the efficiency of the biochemical reaction in the system.
4. A process according to claim 3, further comprising maintaining the IDR-macromolecule or the IDR-polypeptide in the system to cause molecules necessary for the performance of the reaction to co-localise with the IDR-macromolecule or the IDR-polypeptide within the plurality of phase-separated aqueous compartments, thereby increasing the efficiency of the biochemical reaction in the system.

5. A process according to claim 3 or claim 4, wherein the plurality of phase-separated aqueous compartments is a plurality of detectable phase-separated aqueous particles.
- 5 6. A process according to any one of the preceding claims, wherein the process is a biochemical reaction for synthesizing a nucleic acid molecule in an *in vitro* reaction system comprising:
- (a) providing at least one nucleic acid primer;
 - (b) providing a target nucleic acid molecule comprising at least one target
10 strand, and contacting the at least one nucleic acid primer with the target strand, thereby forming a double stranded structure;
 - (c) providing the IDR-macromolecule as an IDR-polypeptide, wherein the IDR-polypeptide is polymerase or one or more polypeptide co-factors; and
 - (d) allowing the reaction to proceed, thereby extending the 3' end of the at least
15 one nucleic acid primer with polymerase and dNTPs, optionally in the presence of one or more polypeptide co-factors, to generate a double stranded nucleic acid, wherein the first strand comprises a sequence of the target strand and the second strand comprises a sequence which is complementary thereto.
- 20
7. A process according to claim 6, wherein the process is a biochemical reaction for amplifying a single stranded target nucleic acid molecule or a double stranded target nucleic acid molecule in an *in vitro* reaction system, preferably wherein the target nucleic acid molecule is a DNA molecule.
- 25
8. A process according to any one of claims 1 to 5, wherein the process is a biochemical reaction for amplifying a double stranded target nucleic acid molecule in an *in vitro* reaction system comprising:
- (a) providing first and second nucleic acid primers;
 - (b) providing a double stranded target nucleic acid molecule comprising a first
30 strand and a second strand, and contacting the first and second nucleic acid

primers with the target nucleic acid molecule thereby forming a first double stranded structure with the first strand and a second double stranded structure with the second strand;

- (c) providing the IDR-macromolecule as an IDR-polypeptide, wherein the IDR-polypeptide is polymerase or one or more protein co-factors;
- (d) allowing the reaction to proceed, thereby extending the 3' ends of the first and second nucleic acid primers with polymerase and dNTPs, optionally in the presence of one or more protein co-factors, to generate first and second double stranded nucleic acids; and
- (e) repeating steps (b) to (d) until a desired degree of amplification is reached.

9. A process according to claim 8, wherein the process is a recombinase polymerase amplification process of amplifying the double stranded target nucleic acid molecule in the *in vitro* reaction system comprising:

- (a) providing a recombinase agent, optionally a recombinase loading protein, a single strand stabilizing agent, polymerase, first and a second nucleic acid primers, double stranded target nucleic acid comprising a first strand and a second strand, and optionally an exonuclease such as Exonuclease III;
- (b) contacting the recombinase agent with the first and second nucleic acid primers and optionally with the recombinase loading protein to form first and second nucleoprotein primers which comprise a single stranded region at their 3' ends;
- (c) contacting the first and second nucleoprotein primers with the target nucleic acid molecule thereby forming a first double stranded structure with the first strand and a second double stranded structure with the second strand;
- (d) allowing the reaction to proceed, thereby extending the 3' end of the first and second nucleoprotein primers with polymerase and dNTPs to generate first and second double stranded nucleic acids and first and a second displaced nucleic acid strands, wherein the single strand stabilizing agent stabilizes the first and second displaced strands; and

(e) continuing the reaction by repeating steps (b) to (d) until a desired degree of amplification is reached;

wherein the recombinase agent, and/or the recombinase loading protein, and/or the single strand stabilizing agent, and/or the polymerase is provided as the IDR-

5 polypeptide.

10. A process according to claim 9, wherein the recombinase agent is selected from the group consisting of UvsX, T4 UvsX, T6 UvsX, RB18 UvsX, *E.coli* phage wV7 UvsX, *Shigella* phage CB8 UvsX, *Shigella* phage Shfl2 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 or any functional analog, homolog or derivative thereof, and any combination thereof, preferably wherein the recombinase agent is UvsX, more preferably *Escherichia* phage vB_EcoM_DalCa UvsX.

15

11. A process according to claim 9 or claim 10, wherein the process includes a recombinase loading protein and wherein the recombinase loading protein is selected from the group consisting of UvsY, *E. coli* RecO, *E. coli* RecR or any functional analog, homolog or derivative thereof, and any combination thereof, preferably wherein the recombinase loading protein is UvsY, more preferably *Escherichia* phage STO UvsY.

20

12. A process according to any one of claims 6 to 11, wherein the polymerase is a eukaryotic polymerase selected from the group consisting of pol- α , pol- β , pol- δ , pol- ϵ or any functional analog, homolog or derivative thereof, and any combination thereof.

25

13. A process according to any one of claims 6 to 11, wherein the polymerase is a prokaryotic polymerase selected from the group consisting of *Bacillus stearothermophilus* polymerase I large fragment, *Bacillus subtilis* Pol I large fragment (Bsu polymerase), *Listeria monocytogenes* DNA polymerase I, *S. aureus* DNA polymerase I (Sau polymerase), *E. coli* DNA polymerase I Klenow fragment, *E. coli*

30

DNA polymerase I, *E. coli* DNA polymerase II, *E. coli* DNA polymerase III, *E. coli* DNA polymerase IV, *E. coli* DNA polymerase V, or any functional analog, homolog or derivative thereof, and any combination thereof, preferably wherein the polymerase is *S. aureus* DNA polymerase I (Sau polymerase) or *Bacillus subtilis* Pol I large fragment (Bsu polymerase).

14. A process according to any one of claims 6 to 11, wherein the polymerase is a bacteriophage polymerase selected from the group consisting of bacteriophage T4 gp43 DNA polymerase, T7 DNA polymerase and Phi-29 DNA polymerase, or any functional analog, homolog or derivative thereof, and any combination thereof.

15. A process according to any one of claims 9 to 14, wherein the single strand stabilizing agent is selected from the group consisting of Gp32, *E. coli* SSB protein, phage T4 Gp32 protein, phage Rb69 Gp32, phage vB_EcoM_NBG1 Gp32, or any functional analog, homolog or derivative thereof, and any combination thereof, preferably the single strand stabilizing agent is Gp32 or phage vB_EcoM_NBG1 Gp32.

16. A process according to any one of claims 9 to 15, wherein only the recombinase agent is provided as the IDR-polypeptide, or wherein only the recombinase loading protein is provided as the IDR-polypeptide, or wherein only the single strand stabilizing agent is provided as the IDR-polypeptide, or wherein only the polymerase is provided as the IDR-polypeptide, or wherein only the exonuclease is provided as the IDR-polypeptide.

17. A process according to any one of the preceding claims, wherein the one or more functional IDRs of the IDR-polypeptide are tagged to the IDR-polypeptide as an amino acid sequence comprising or consisting of the one or more IDRs such that the IDR-polypeptide is a genetically engineered fusion protein, wherein the one or more functional IDRs are located at the C-terminus of the IDR-polypeptide, at the N-terminus of the IDR-polypeptide, or at both the C-terminus of the IDR-polypeptide and the N-

terminus of the IDR-polypeptide, or at any amino acid position along the length of the polypeptide.

18. A process according to any one of the preceding claims, wherein the one or
5 more functional IDRs of the IDR-macromolecule or the IDR-polypeptide is characterised as a sequence of amino acids which scores greater than 0.5 when analysed by the algorithm MetaDisorder.
19. A process according to any one of the preceding claims, wherein the one or
10 more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of an amino acid sequence which includes one or more repeats of the tri-peptide sequence RGG.
20. A process according to claim 19, wherein the one or more functional IDRs of the
15 IDR-macromolecule or the IDR-polypeptide comprises or consists of an amino acid sequence which further includes one or more repeats of the di-peptide sequence FG.
21. A process according to claim 19 or claim 20, wherein the one or more functional
20 IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of an amino acid sequence which further includes at least one aromatic amino acid residue consisting of tyrosine or phenylalanine.
22. A process according to any one of the preceding claims, wherein the one or
25 more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of the amino acid sequence of
- i. (YNPQGGYQQ)_n, wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3; or
 - ii. (YSPTSPTS)_n, wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3; or
 - 30 iii. (FSPTSPT)_n, wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3; or

- iv. (YSPTSP-A/N/G)_n, wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3; or
- v. (YSPGSPA)_n, wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3.

5

23. A process according to any one of the preceding claims, wherein the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of an amino acid sequence which is glutamine rich, optionally wherein the amino acid sequence comprises at least 4, at least 5, at least 6, at least 7, at least 8, at
10 least 9, or at least 10 consecutive glutamine residues.

24. A process according to claim 23, wherein the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of an amino acid sequence comprising one or more repeats of the tri-peptide sequence QQQ.

15

25. A process according to claim 23 or claim 24, wherein the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of the amino acid sequence of (QQQPQY)_n, wherein n is a positive integer between 1 and 10, optionally wherein n = 1, 2, or 3.

20

26. A process according to any one of the preceding claims, wherein the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of a sequence of at least 5 consecutive amino acids of SEQ ID NO:1.

25 27. A process according to any one of the preceding claims, wherein the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises or consists of an amino acid sequence of at least 5 consecutive amino acids of SEQ ID NO:9.

30 28. A process according to any one of the preceding claims, wherein the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises an

amino acid sequence containing one or more aromatic tyrosine residues and one or more phenylalanine residues which can engage in aromatic cation- π interactions with multivalent metal ions, preferably divalent metal ions.

- 5 29. A process according to any one of the preceding claims, wherein the one or more functional IDRs of the IDR-macromolecule or the IDR-polypeptide comprises an amino acid sequence containing one or more arginine residues which can engage in guanidine-metal interactions with multivalent metal ions, preferably divalent metal ions.
- 10 30. A process according to any one of the preceding claims, wherein the IDR-polypeptide comprises or consists of the polypeptide tagged with an amino acid sequence which comprises or consists of the amino acid sequence of any one of SEQ ID NOs 1 to 43, or which comprises or consists of a functional variant amino acid sequence of SEQ ID NOs 1 to 43, *e.g.* which has 80% or more identity to any one of SEQ ID NOs
15 1 to 43.
31. A process according to any one of claims 9 to 30, wherein the IDR-polypeptide is a single strand stabilizing agent which is Gp32 and which has the amino acid sequence of any one of SEQ ID NOs 65 to 88, or wherein the IDR-polypeptide is a
20 functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 65 to 88.
32. A process according to any one of claims 9 to 30, wherein the IDR-polypeptide is a recombinase agent which is UvsX and which has the amino acid sequence of any
25 one of SEQ ID NOs 44 to 59, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 44 to 59.
33. A process according to any one of claims 9 to 30, wherein the IDR-polypeptide
30 is a recombinase loading protein which is UvsY and which has the amino acid sequence of any one of SEQ ID NOs 60 to 64, or wherein the IDR-polypeptide is a functional

variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 60 to 64.

34. A process according to any one of the preceding claims, further comprising
5 providing multivalent metal ions to the IDR-macromolecule or IDR-polypeptide in the *in vitro* reaction system, thereby simulating or enhancing the liquid-liquid demixing in the *in vitro* reaction system, thereby increasing the efficiency of the biochemical reaction, preferably wherein said liquid-liquid demixing promotes the formation of detectable phase-separated particles in the *in vitro* reaction system thereby increasing
10 the efficiency of the biochemical reaction; optionally wherein the multivalent metal ions are provided at a concentration of about 22 mM or more, preferably wherein the multivalent metal ions are provided at a concentration of between about 22 mM to 50 mM.

15 35. A process according to claim 34, wherein the multivalent metal ions are divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

20 36. A process according to any one of the preceding claims, wherein the biochemical reaction is performed in a solid phase reaction system which comprises a surface.

25 37. A process according to claim 36, wherein the biochemical reaction is a process of amplifying a single stranded target nucleic acid molecule or a double stranded target nucleic acid molecule in the *in vitro* reaction system as defined in claim 3 or claim 4, wherein the at least one nucleic acid primer and/or the IDR-macromolecule and/or the one or more polypeptide co-factors are attached to the surface.

30 38. A process according to any one of claims 9 to 34, wherein the biochemical reaction is a recombinase polymerase amplification process of amplifying a double stranded target nucleic acid molecule in the *in vitro* reaction system, wherein the

reaction is performed in a solid phase reaction system which comprises a surface, and wherein the recombinase agent and/or the recombinase loading protein and/or the single strand stabilizing agent and/or the polymerase and/or the exonuclease and/or the first nucleic acid primer and/or the second nucleic acid primer is/are attached to the surface, preferably wherein: (i) the first nucleic acid primer or the second nucleic acid primer is attached to the surface; or (ii) both the first and second nucleic acid primers are attached to the surface.

39. A process according to any one of claims 36 to 38, wherein the surface is planar or is a microbead, preferably wherein the surface comprises silicon, glass, a gel-based material and/or a polymeric material such as polystyrene, more preferably wherein the surface is a microbead comprising a polymeric material such as polystyrene.

40. A process according to claim 39, wherein the surface is bound to a substrate, preferably wherein the surface is planar and/or the substrate comprises glass, optionally wherein the surface and/or the substrate are provided as a flow-cell.

41. A non-naturally occurring IDR-macromolecule comprising a macromolecule and a tag amino acid sequence, wherein the tag sequence comprises or consists of one or more functional intrinsically disordered regions (IDRs), wherein the IDR-macromolecule is capable of promoting liquid-liquid demixing in solution.

42. An IDR-macromolecule according to claim 41, wherein said liquid-liquid demixing promotes the formation of detectable phase-separated particles in the *in vitro* reaction system.

43. An IDR-macromolecule according to claim 41 or claim 42, wherein said liquid-liquid demixing and formation of phase-separated aqueous compartments thereby increases the efficiency of the biochemical reaction in the system, or wherein the formation of the detectable phase-separated aqueous particles thereby increases the efficiency of the biochemical reaction in the system.

44. An IDR-macromolecule according to any one of claims 41 to 43, which is a non-naturally occurring, artificial or genetically engineered IDR-polypeptide comprising a polypeptide and the tag amino acid sequence.
- 5
45. An IDR-polypeptide according to claim 44, wherein the tag amino acid sequence is located at the C-terminus of the polypeptide, at the N-terminus of the polypeptide or at both the C-terminus of the polypeptide and the N-terminus of the polypeptide, or at any amino acid position along the length of the polypeptide.
- 10
46. An IDR-polypeptide according to claim 45, wherein the one or more functional IDRs of the tag amino acid sequence are functional IDRs as defined in any one of claims 14 to 28.
- 15
47. An IDR-macromolecule or polypeptide according to any one of claims 41 to 46, wherein the tag sequence comprises amino acid residues which are capable of engaging in aromatic cation-pi interactions with multivalent metal cations, preferably divalent metal cations, more preferably Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} ions, yet more preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , even more preferably Mg^{2+} .
- 20
48. An IDR-polypeptide according to any one of claims 44 to 47, wherein the polypeptide to which the sequence comprising or consisting of the one or more functional IDRs is tagged is an enzyme, such as a helicase, a gyrase, a recombinase, *e.g.* an RPA recombinase agent, a nuclease, *e.g.* an exonuclease and an endonuclease, a
- 25
- ligase, a glycolyase, a methylase, a methyltransferase, a glucosyltransferase, a polymerase, a kinase, a phosphatase, a gene editing enzyme such as a CRISPR enzyme, *e.g.* a Cas9 enzyme; a co-factor, *e.g.* as an RPA recombinase loading protein and an RPA single strand stabilizing agent.
- 30
49. An IDR-polypeptide according to claim 48, wherein the polypeptide to which the sequence comprising or consisting of the one or more functional IDRs is tagged is

an RPA single strand stabilizing agent, preferably Gp32; optionally wherein the IDR-polypeptide has the amino acid sequence of any one of SEQ ID NOs 65 to 88, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID
5 NOs 65 to 88.

50. An IDR-polypeptide according to claim 48, wherein the polypeptide to which the sequence comprising or consisting of the one or more functional IDRs is tagged is an RPA recombinase agent, preferably UvsX; optionally wherein the IDR-polypeptide
10 has the amino acid sequence of any one of SEQ ID NOs 44 to 59, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID NOs 44 to 59.

51. An IDR-polypeptide according to claim 48, wherein the polypeptide to which the sequence comprising or consisting of the one or more functional IDRs is tagged is an RPA recombinase loading protein, preferably UvsY; optionally wherein the IDR-polypeptide has the amino acid sequence of any one of SEQ ID NOs 60 to 64, or wherein the IDR-polypeptide is a functional variant thereof, *e.g.* an IDR-polypeptide having an amino acid sequence which has 80% or more identity to any one of SEQ ID
15
20 NOs 60 to 64.

52. An isolated nucleic acid molecule comprising a first nucleic acid sequence encoding the IDR-polypeptide according to any one of claims 41 to 51; optionally including a second nucleic acid sequence encoding a promoter, wherein the first nucleic
25 acid sequence is operably linked to the second nucleic acid sequence.

53. A recombinant polynucleotide expression vector comprising the nucleic acid molecule according to claim 52.

30 54. A host cell comprising the nucleic acid molecule defined in claim 52, or the recombinant polynucleotide expression vector defined in claim 53.

55. A cell culture comprising a growth medium and a population of host cells, wherein the population comprises host cells as defined according to claim 54.
- 5 56. A kit comprising a non-naturally occurring IDR-macromolecule or an IDR-polypeptide according to any one of claims 41 to 51.
57. A kit according to claim 56, further comprising additional RPA components comprising an RPA recombinase agent, and/or an RPA recombinase loading protein,
10 and/or polymerase, and/or first and second nucleic acid primers, and/or an exonuclease, and/or a buffer, and/or a source of multivalent metal ions, preferably divalent metal cations.
58. A kit according to claim 56 or claim 57, wherein all components are provided in
15 lyophilized form.
59. A process of stimulating or enhancing liquid-liquid demixing in a solution, the process comprising providing a solution comprising an IDR-macromolecule or an IDR-polypeptide according to any one of claims 41 to 51, and contacting the IDR-
20 macromolecule or IDR-polypeptide in solution with multivalent metal ions whereupon liquid-liquid demixing in the solution is stimulated or enhanced.
60. A process according to claim 59, wherein liquid-liquid demixing results in the formation of detectable phase-separated particles, in the solution.
25
61. A process according to claim 59 or claim 60, wherein the multivalent metal ions are divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} , Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .

62. The use of multivalent metal ions in stimulating or enhancing liquid-liquid demixing in a solution, said demixing being mediated by an IDR-macromolecule or an IDR-polypeptide according to any one of claims 41 to 51.
- 5 63. The use according to claim 62, wherein liquid-liquid demixing results in the formation of detectable phase-separated particles in the solution.
64. The use according to claim 62 or claim 63, wherein the multivalent metal ions are divalent metal ions, optionally Mg^{2+} , Mn^{2+} , Ca^{2+} , Co^{2+} or Ni^{2+} , preferably Mg^{2+} ,
10 Mn^{2+} or Ca^{2+} , more preferably Mg^{2+} .
65. The use according to any one of claims 62 to 64, wherein the multivalent metal ions engage in aromatic cation- π interactions with amino acid residues in the one or more functional IDR amino acid sequences, thereby promoting liquid-liquid demixing.
15
66. A non-naturally occurring IDR-macromolecule or an IDR-polypeptide according to any one of claims 41 to 48 for use in therapy or for use as a diagnostic agent.
67. A method for determining the nucleotide sequence of one or more target
20 polynucleotide molecules, said method comprising the steps of:
- (i) performing a process according to any one of claims 1 to 40 to amplify the one or more target polynucleotide molecules thereby obtaining a population comprising a plurality of copies of the one or more target polynucleotide molecules; and
 - (ii) performing one or more nucleic acid sequencing reactions on the
25 population comprising the plurality of copies of the target polynucleotide molecules, preferably wherein the method is performed in a solid phase reaction system which comprises a surface.
68. The use of an IDR-macromolecule or an IDR-polypeptide according to any one
30 of claims 41 to 51 in a method for determining the nucleotide sequence of one or more

target polynucleotide molecules, preferably DNA molecules, preferably wherein the method is as defined in claim 67.

Figure 1

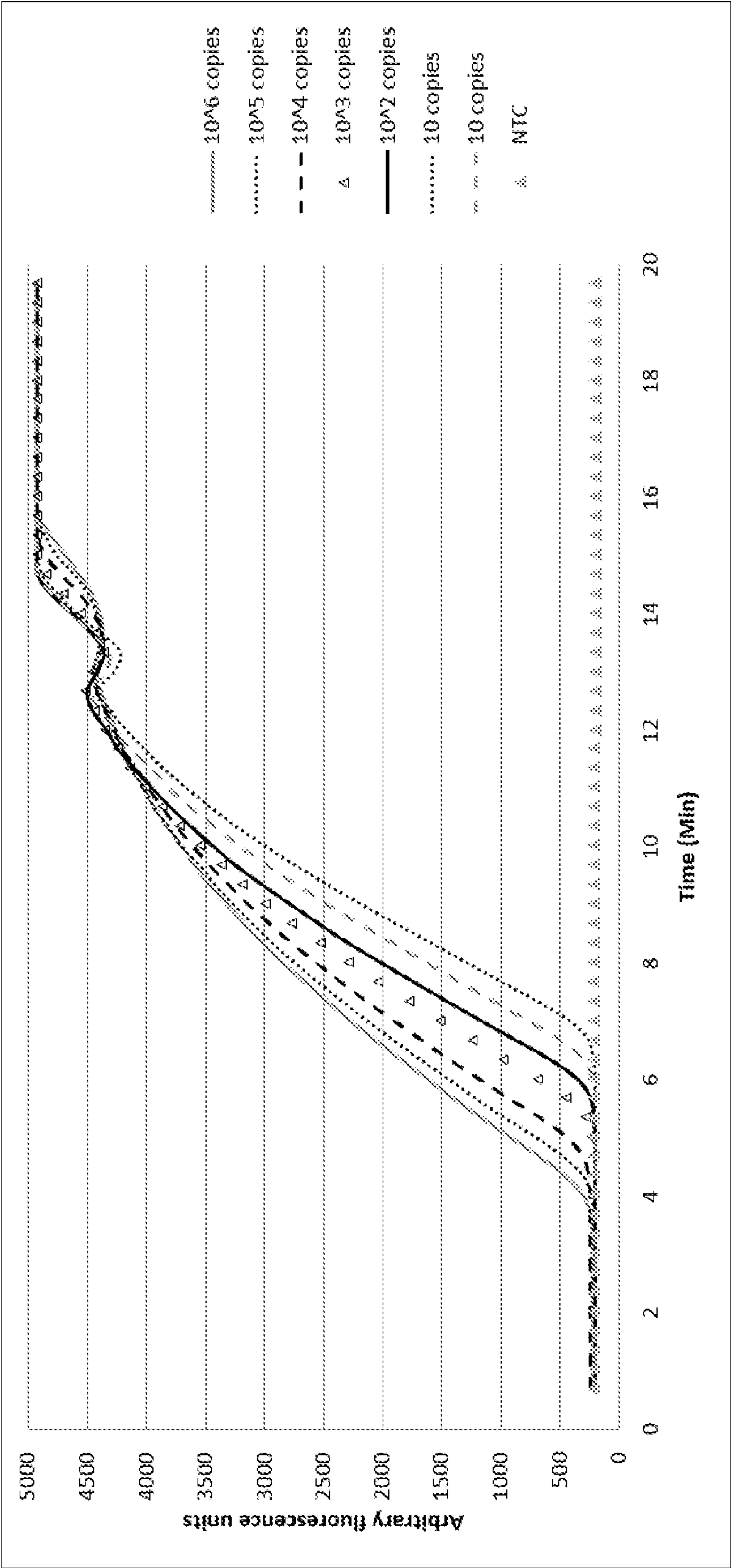


Figure 2

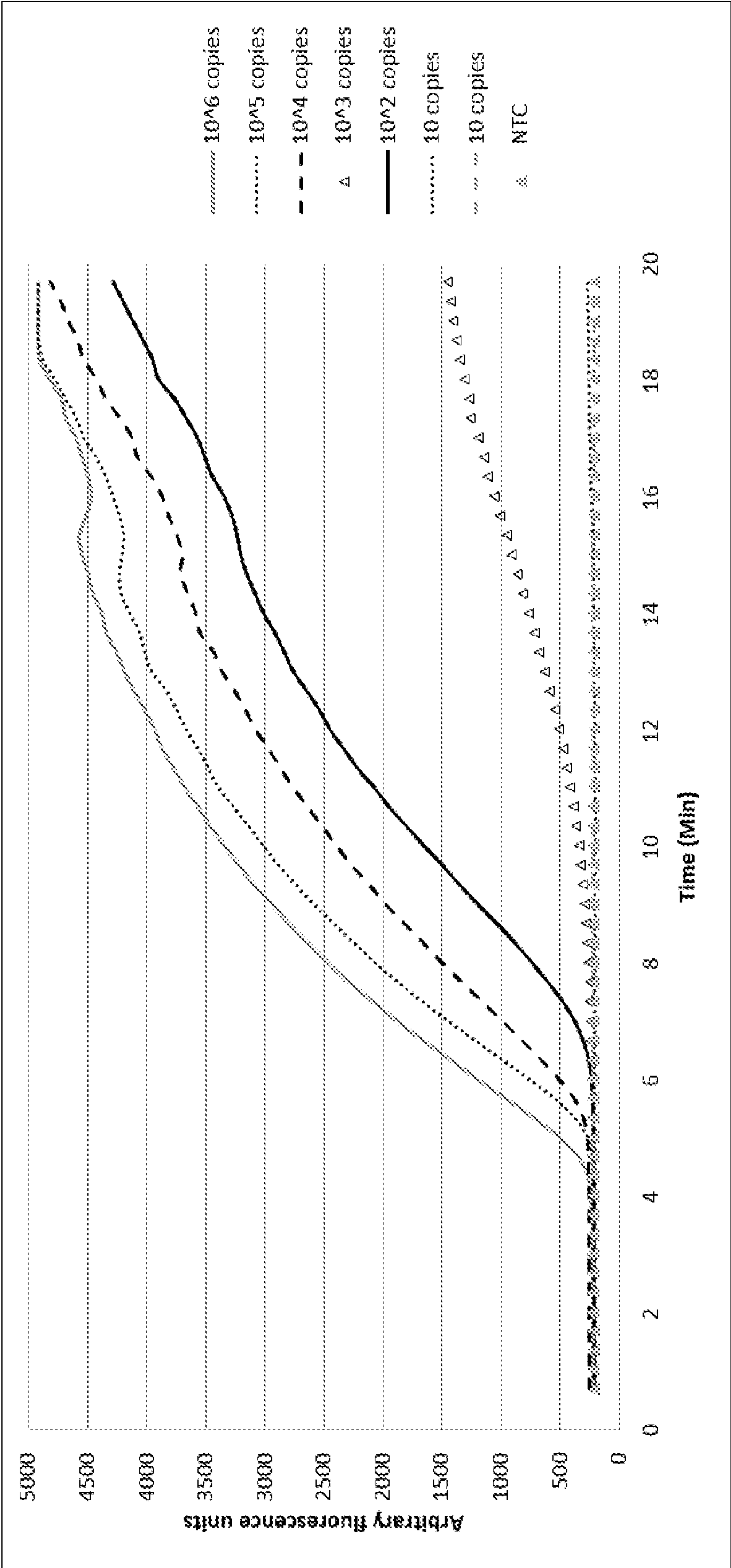


Figure 3

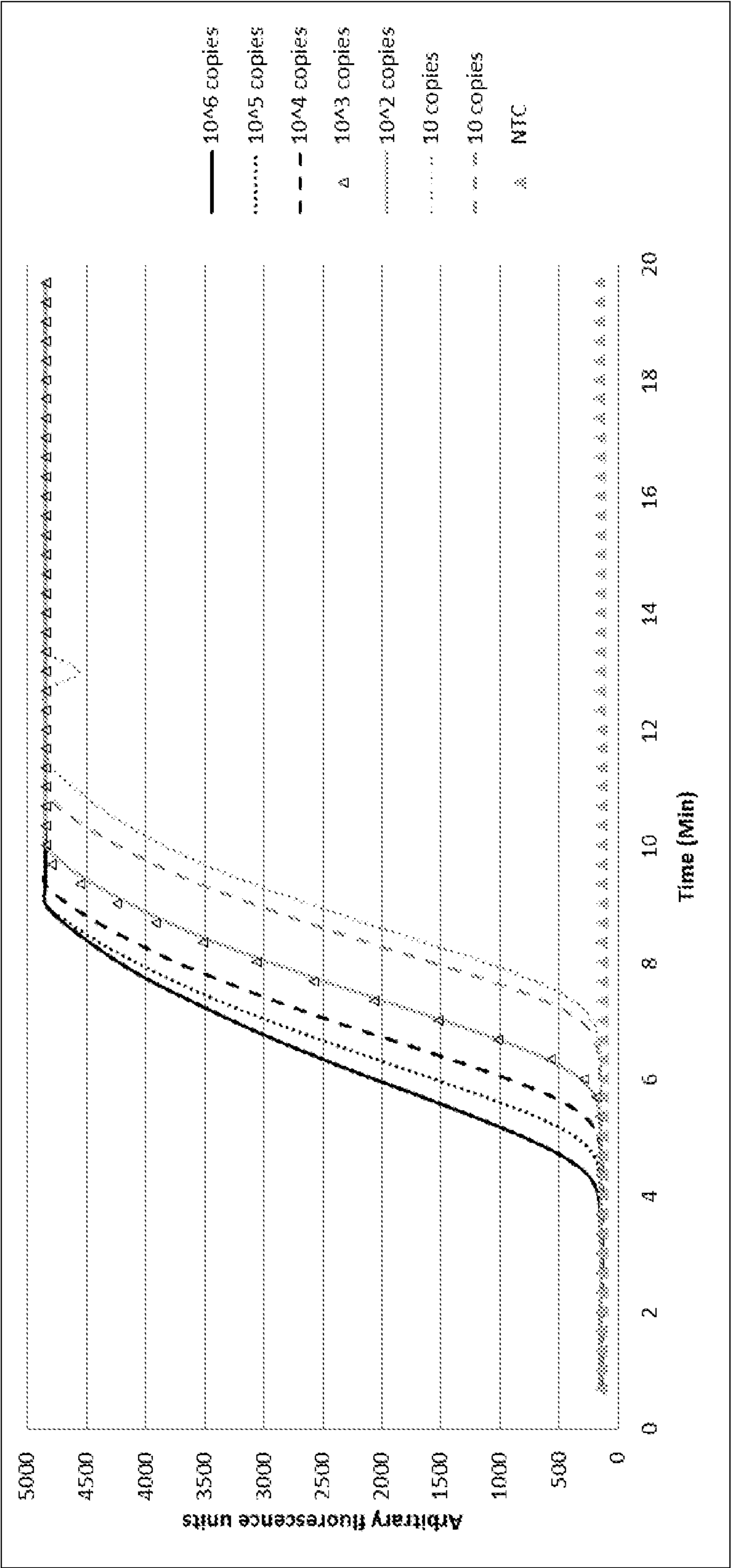


Figure 4

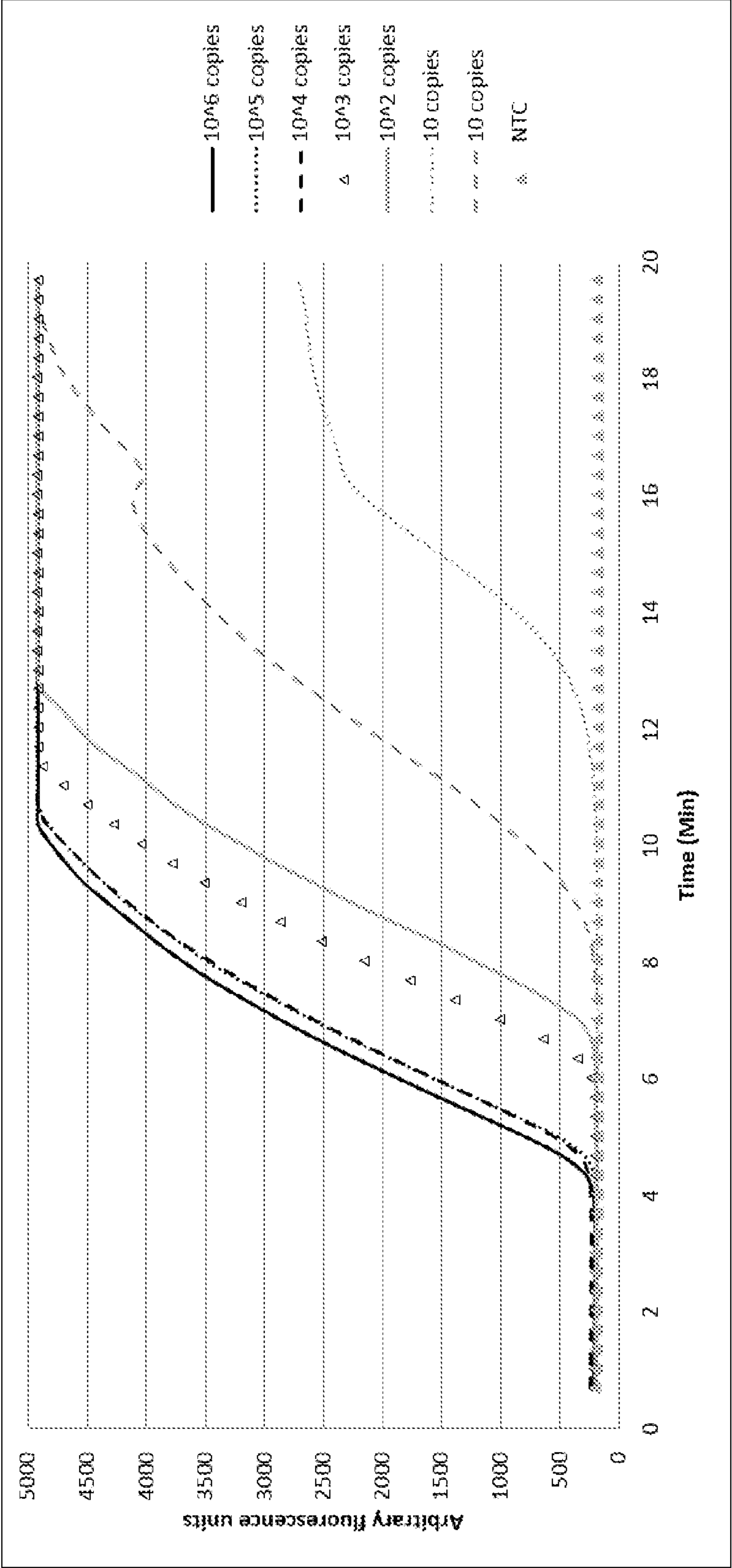


Figure 5A

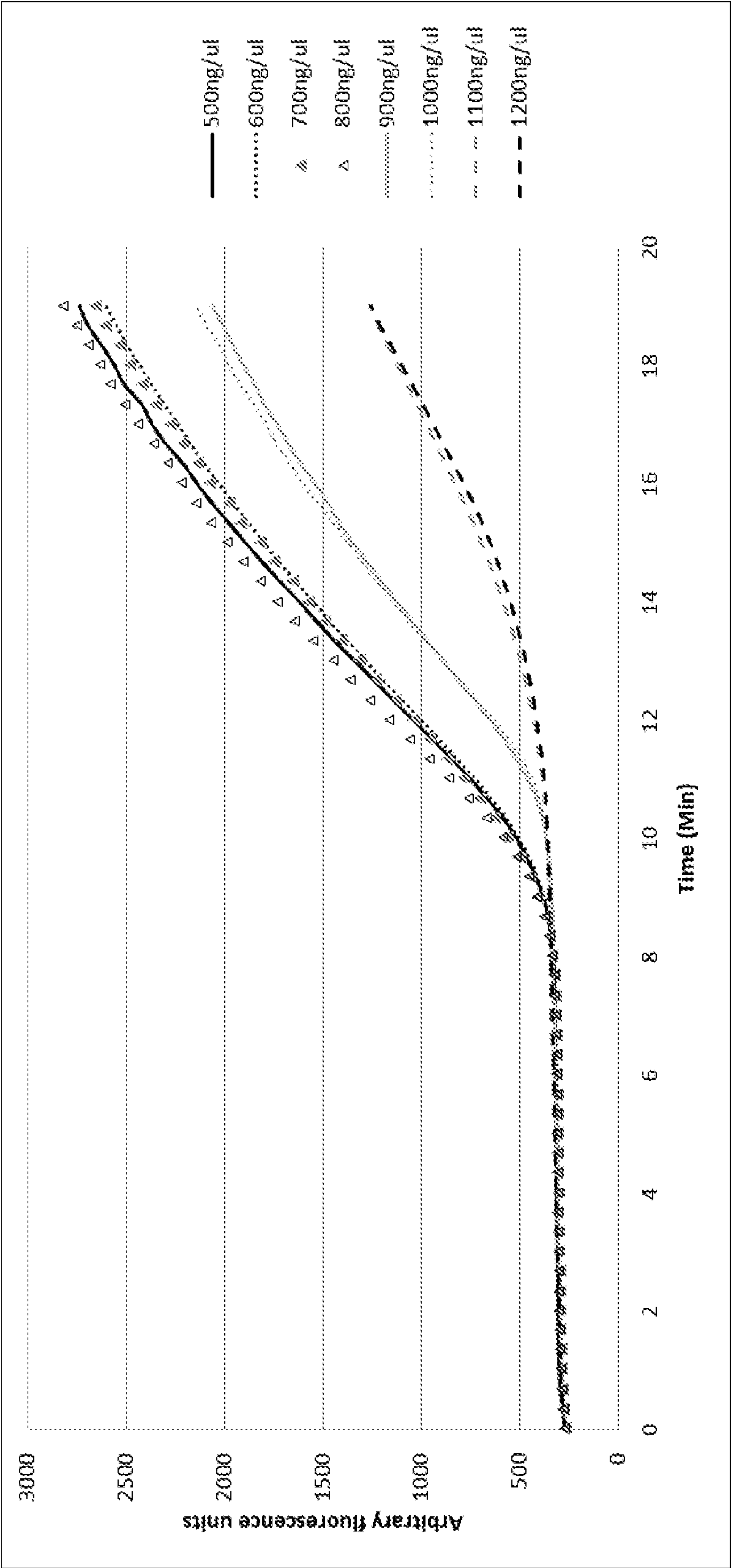


Figure 5B

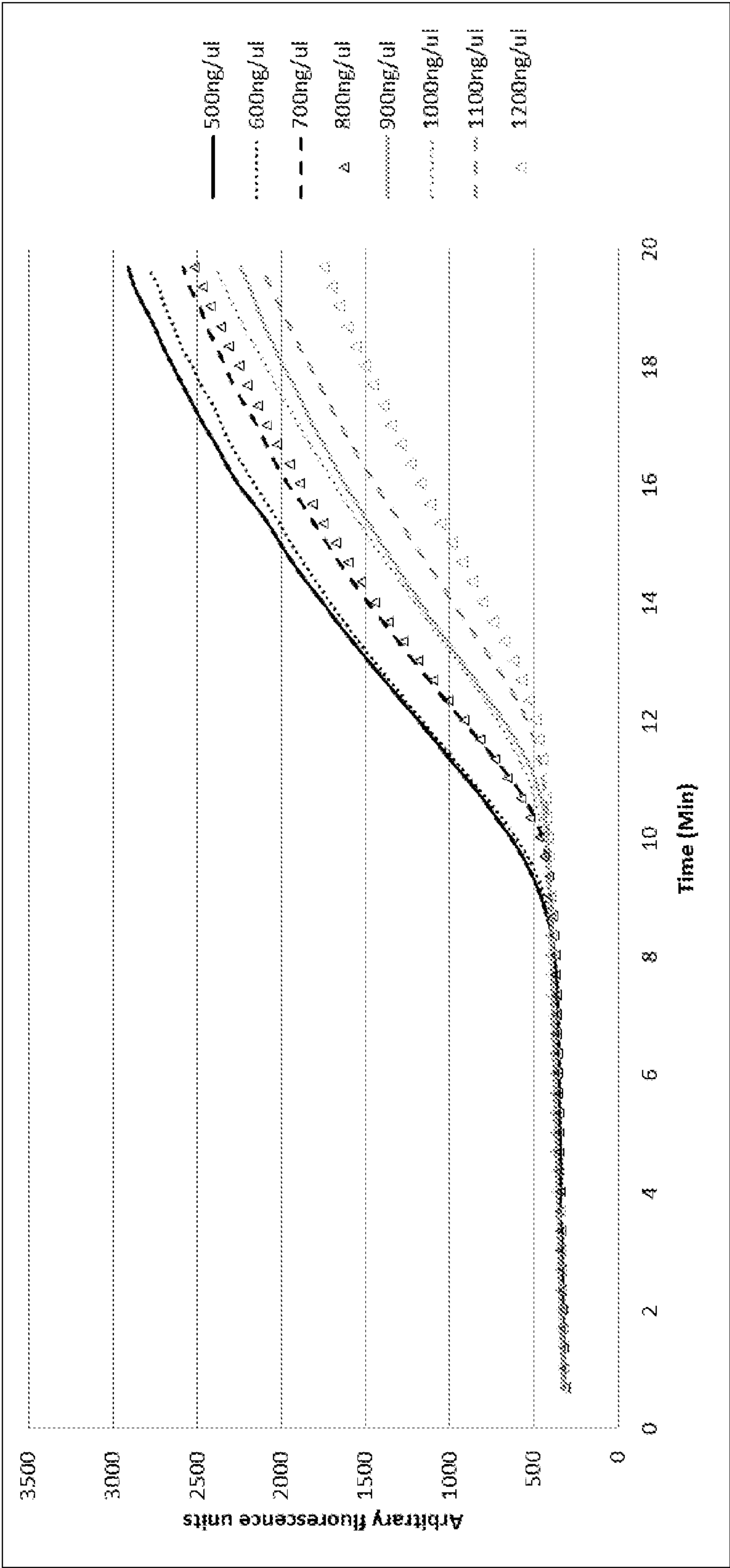


Figure 5C

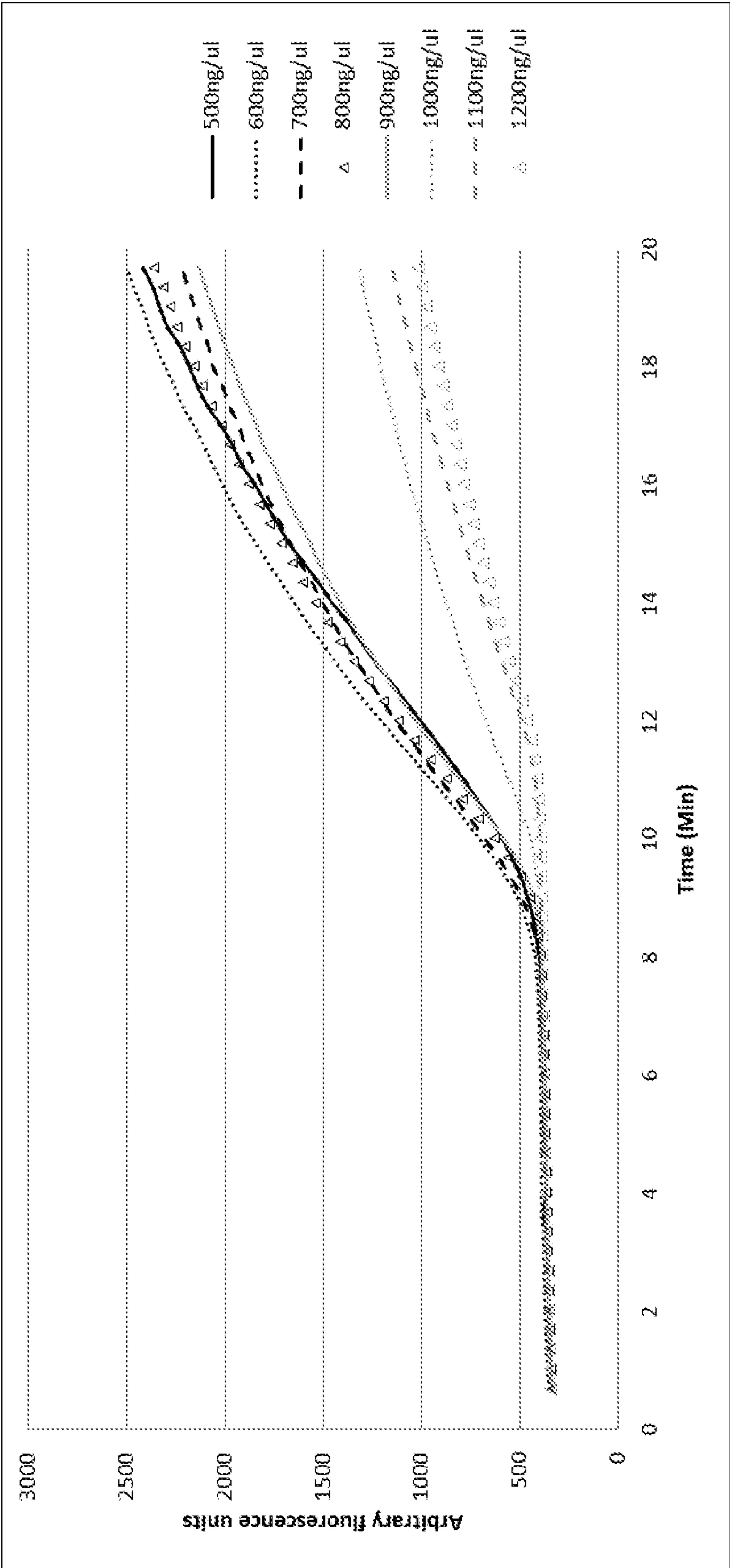


Figure 5D

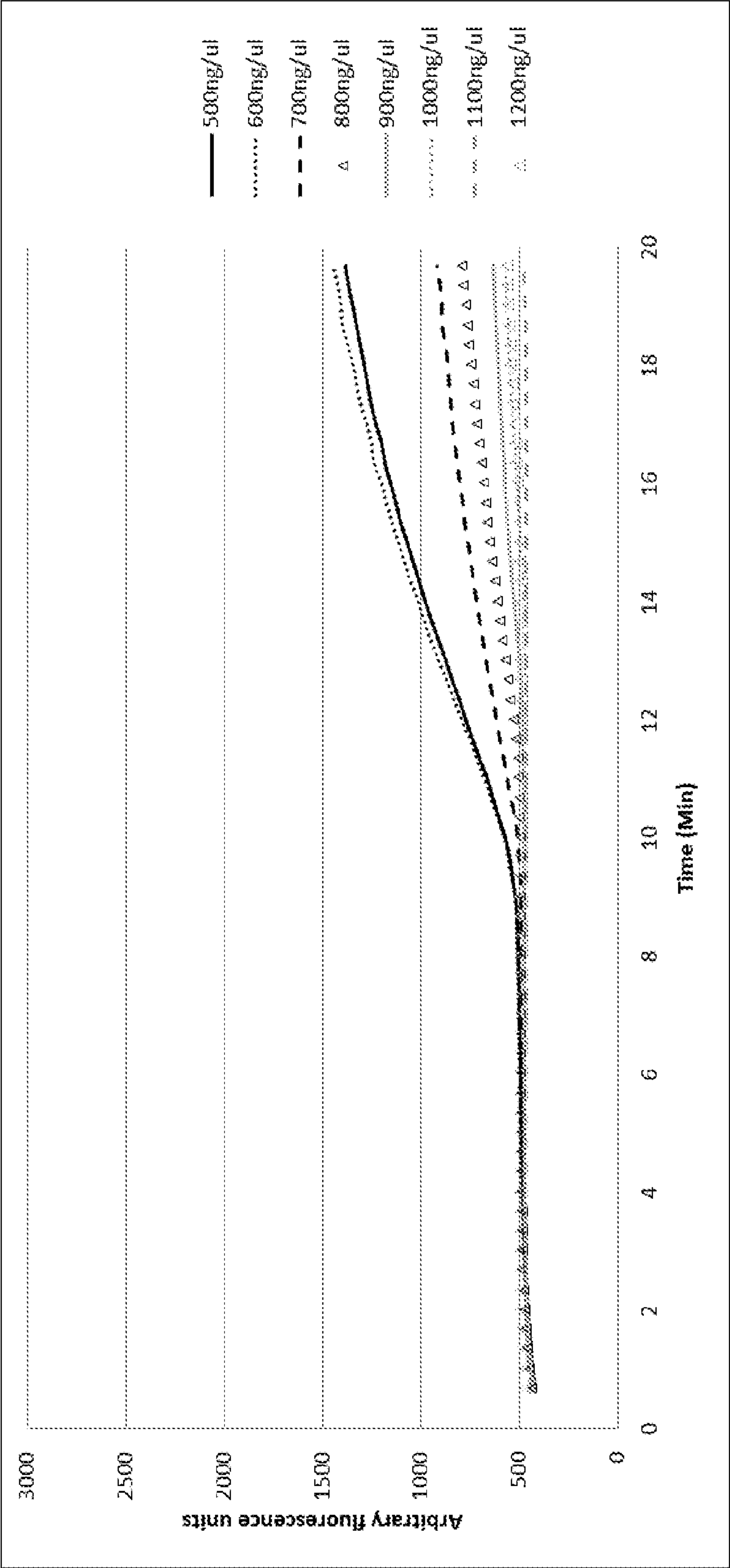


Figure 6

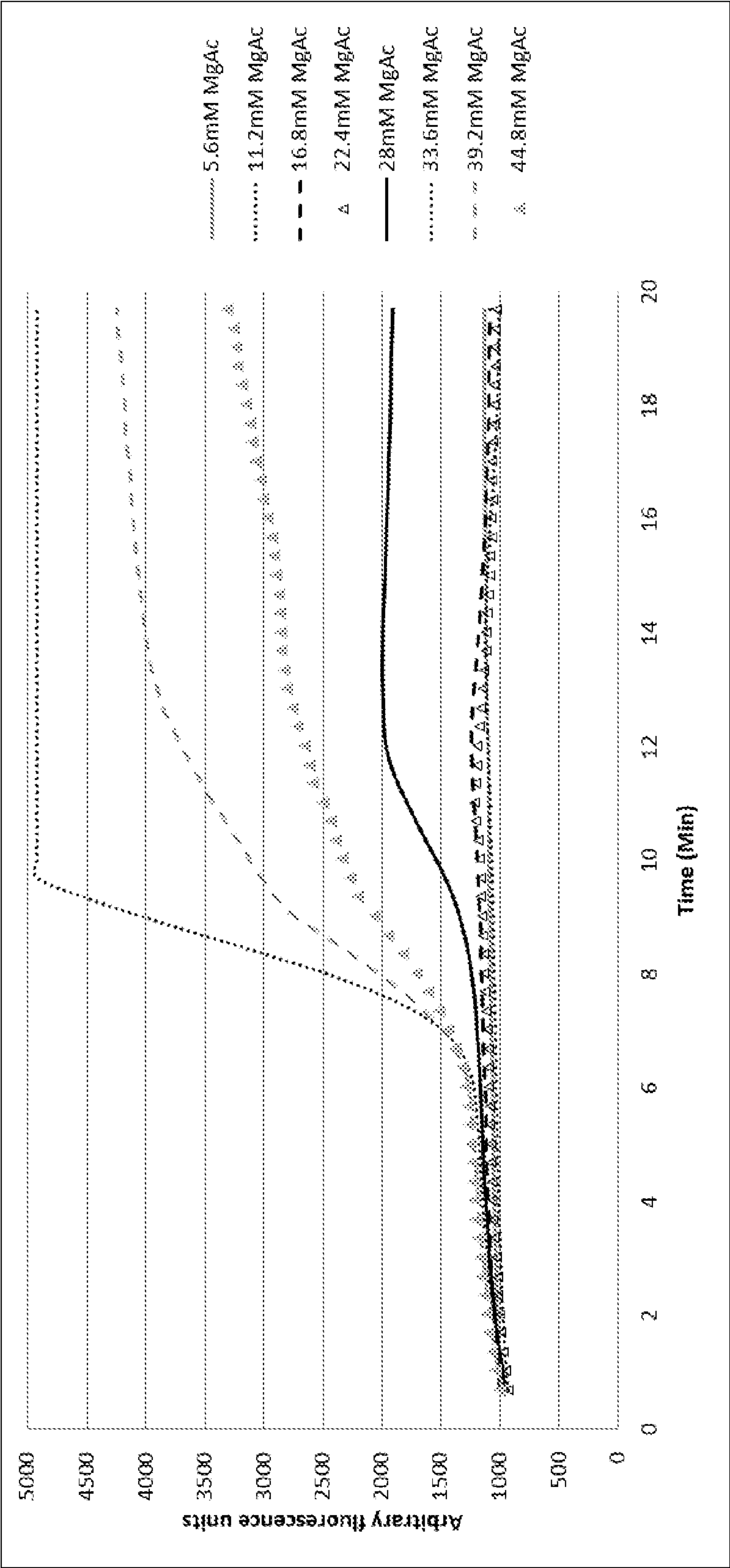


Figure 7A

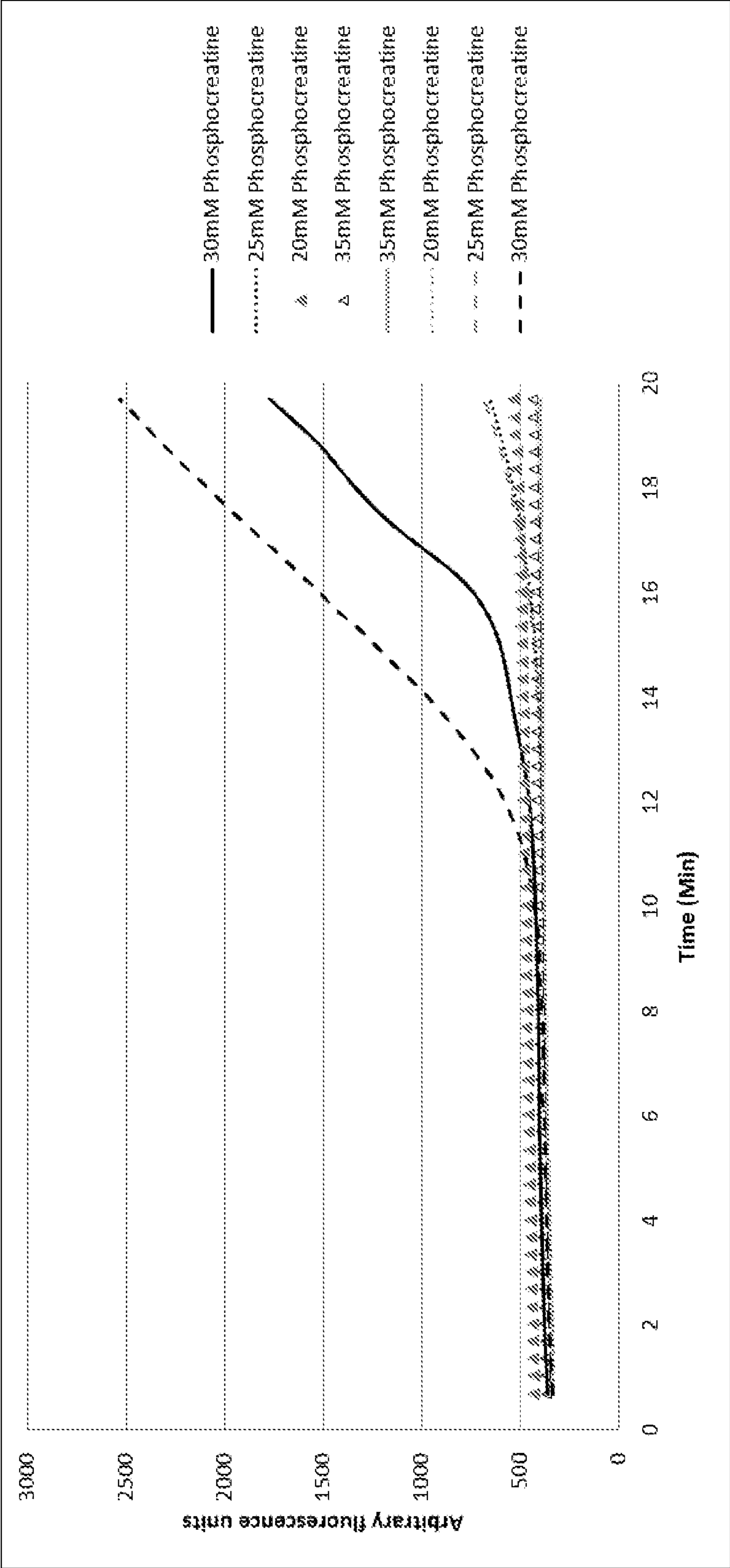


Figure 7B

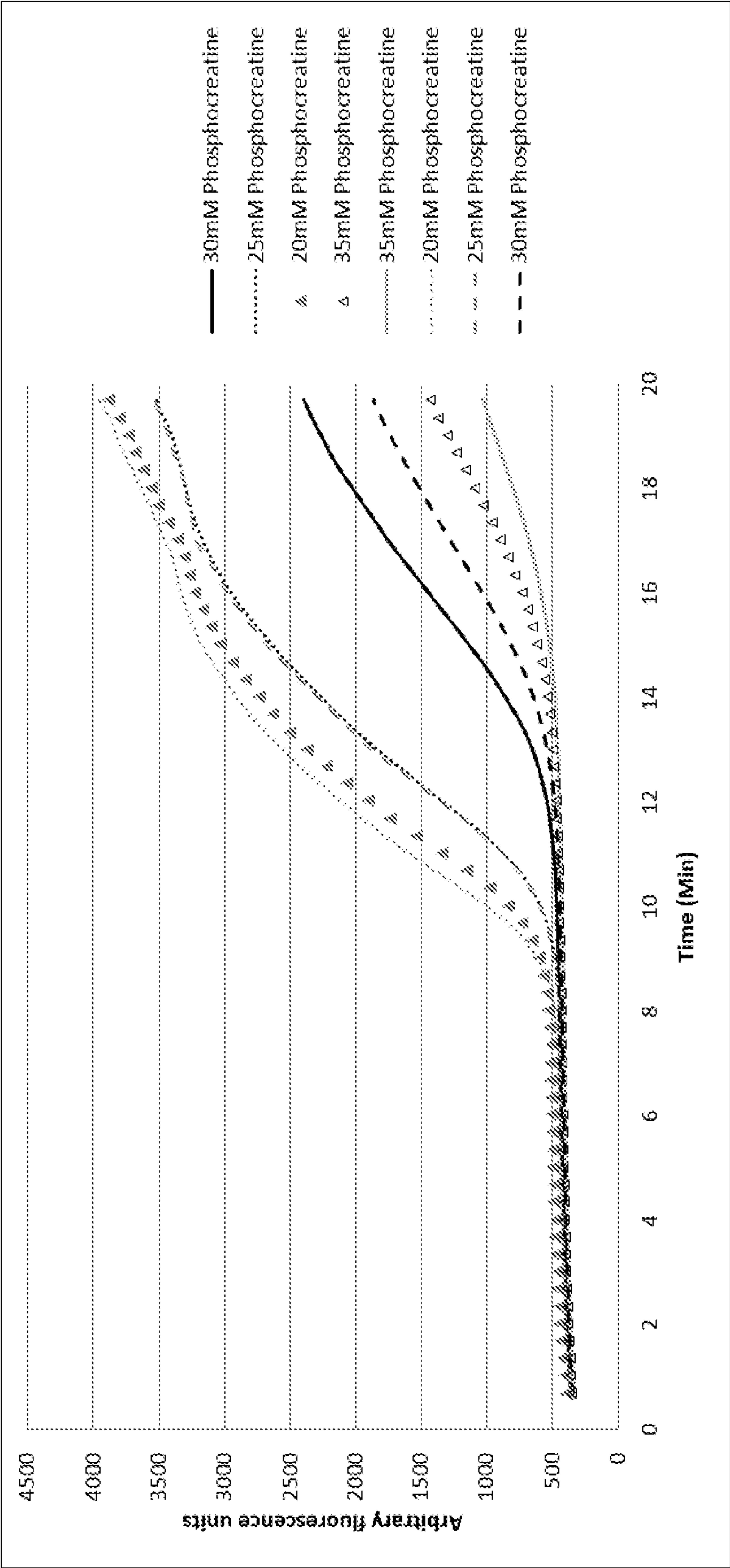


Figure 7C

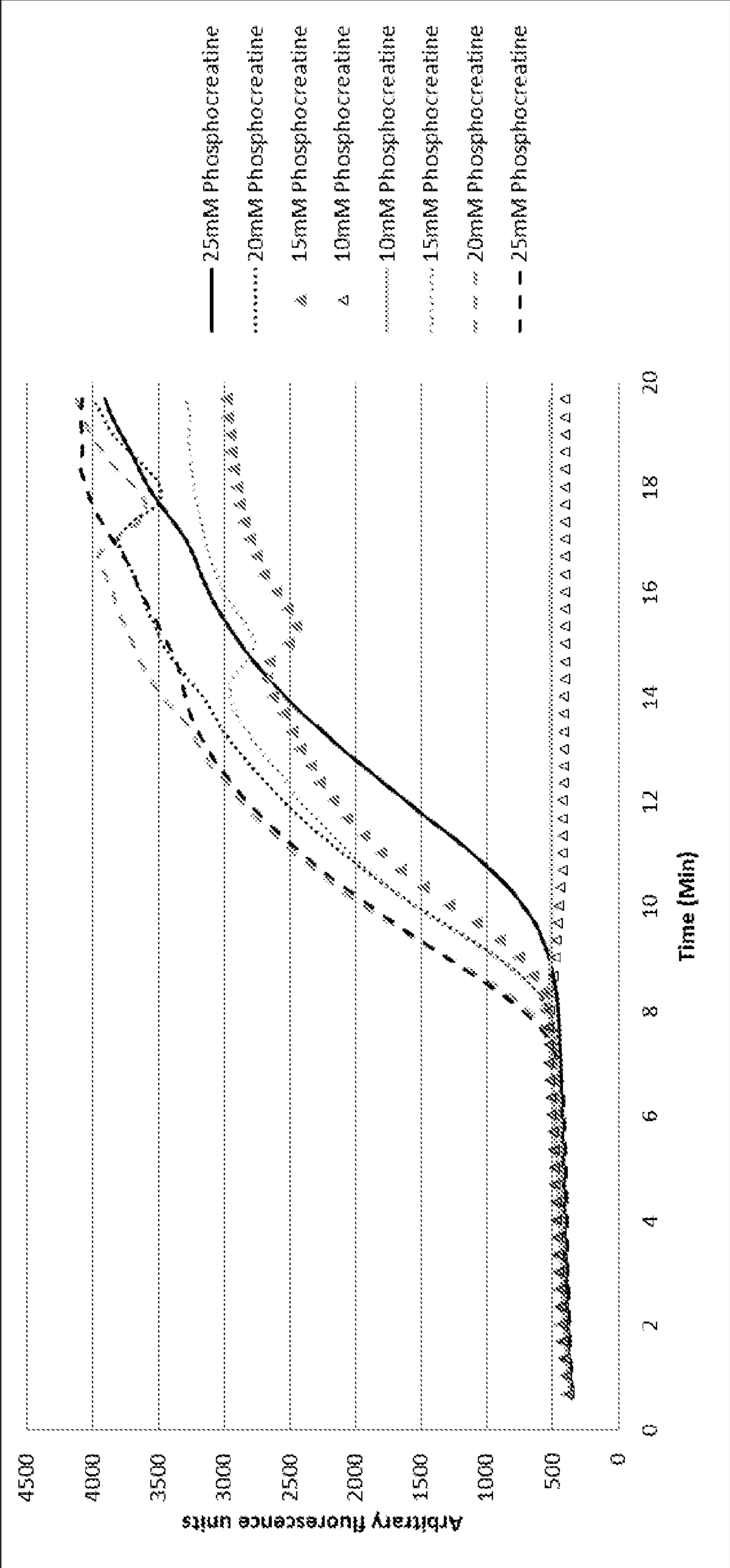


Figure 8

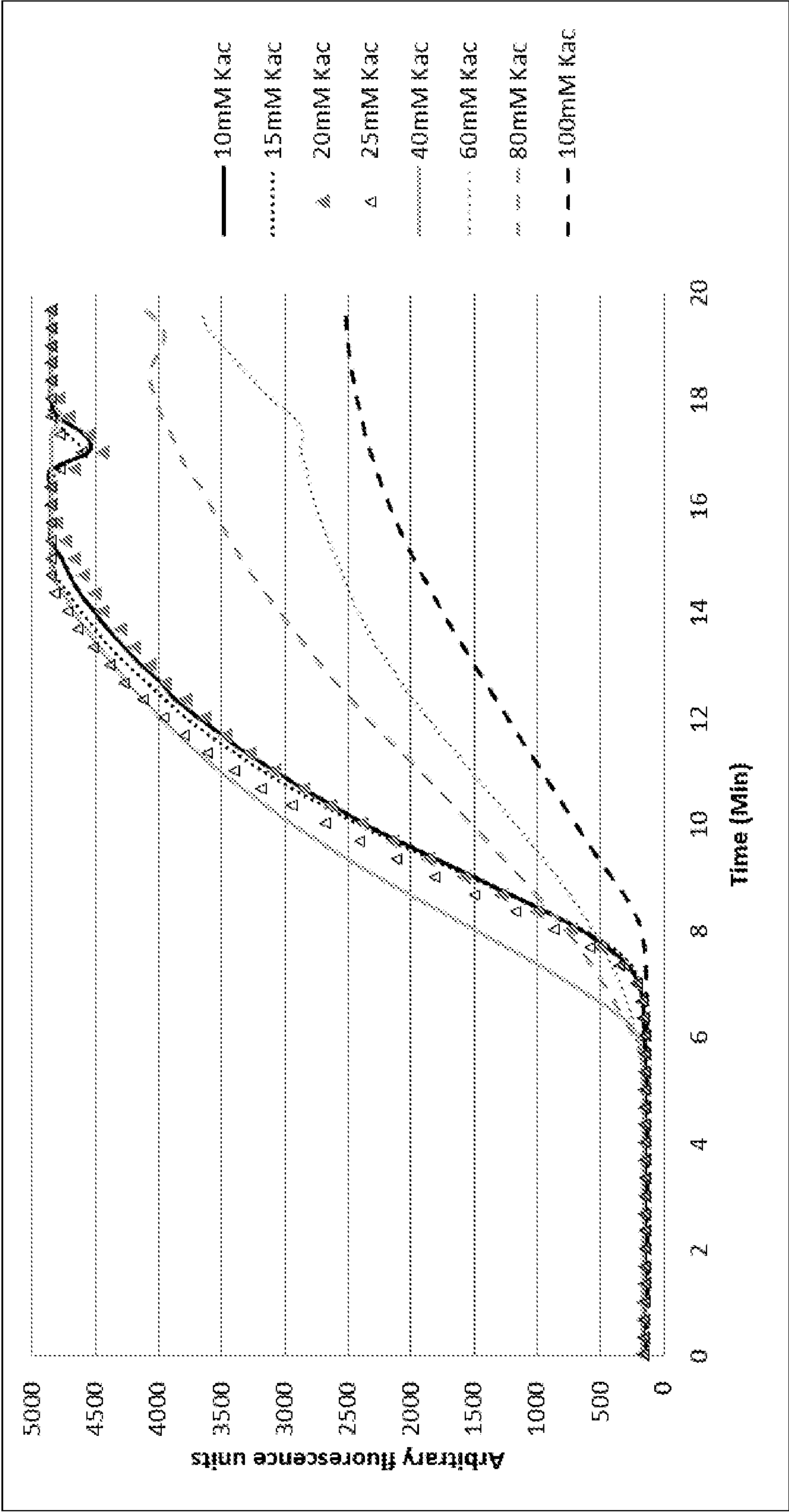


Figure 9

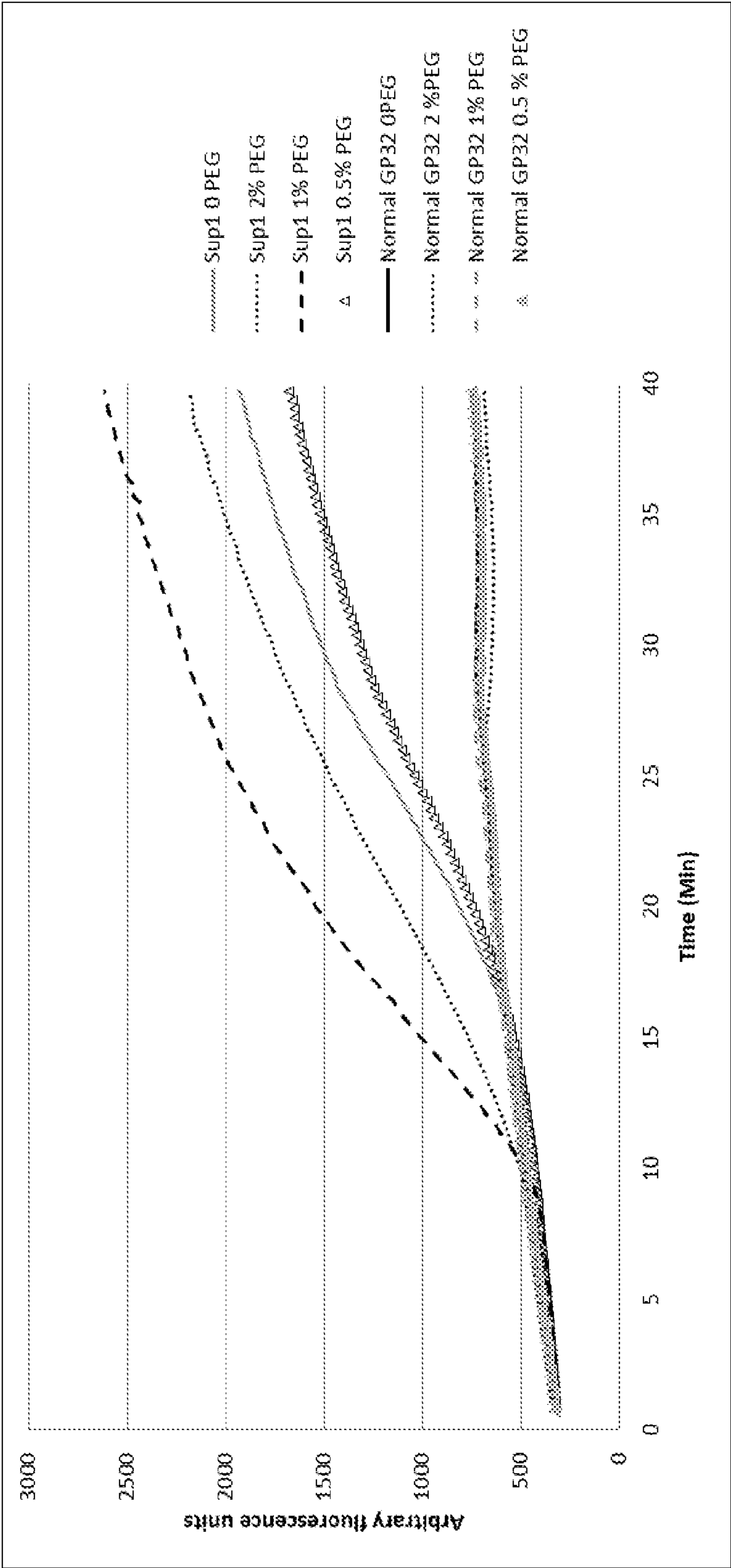


Figure 10A

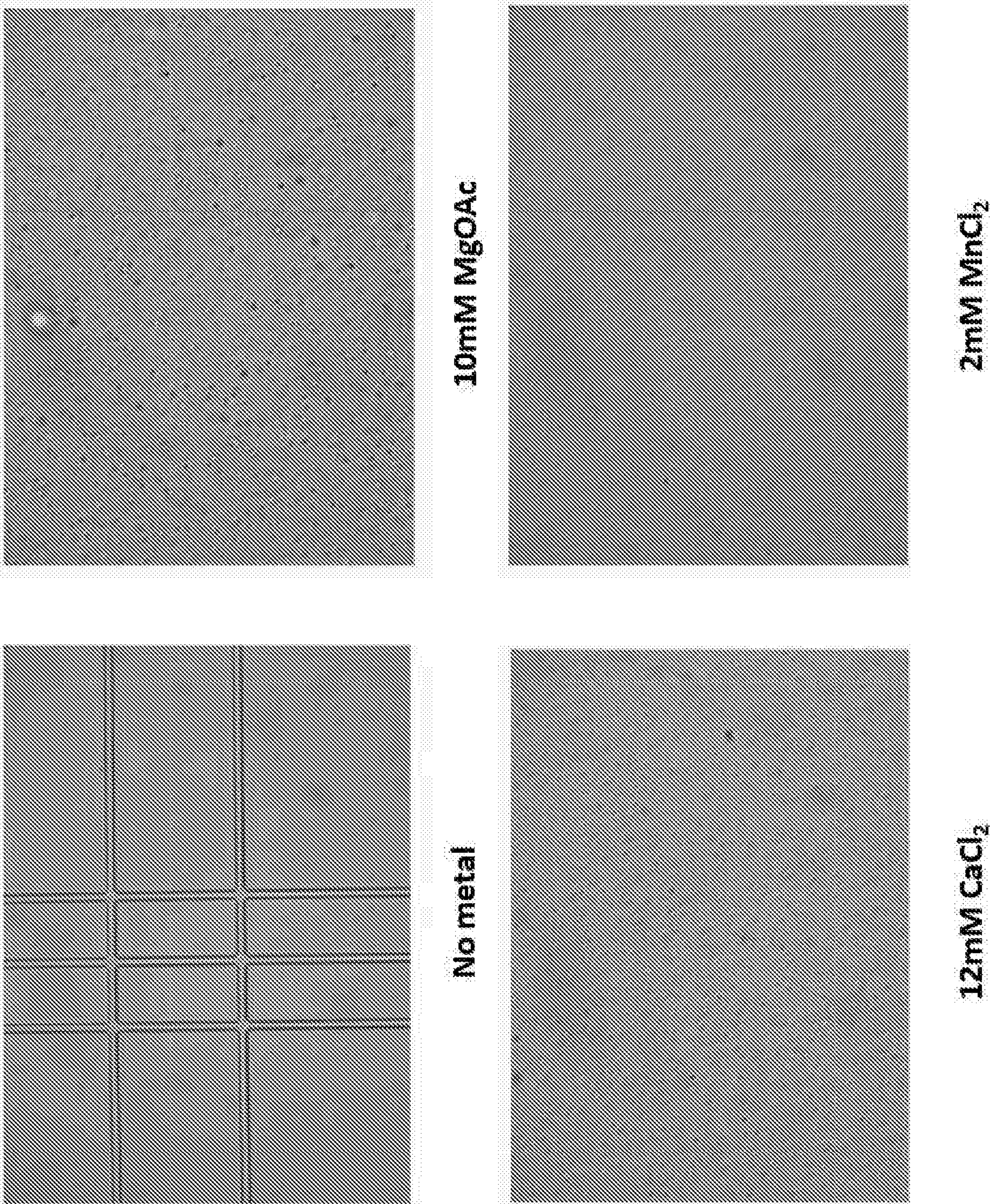


Figure 10B

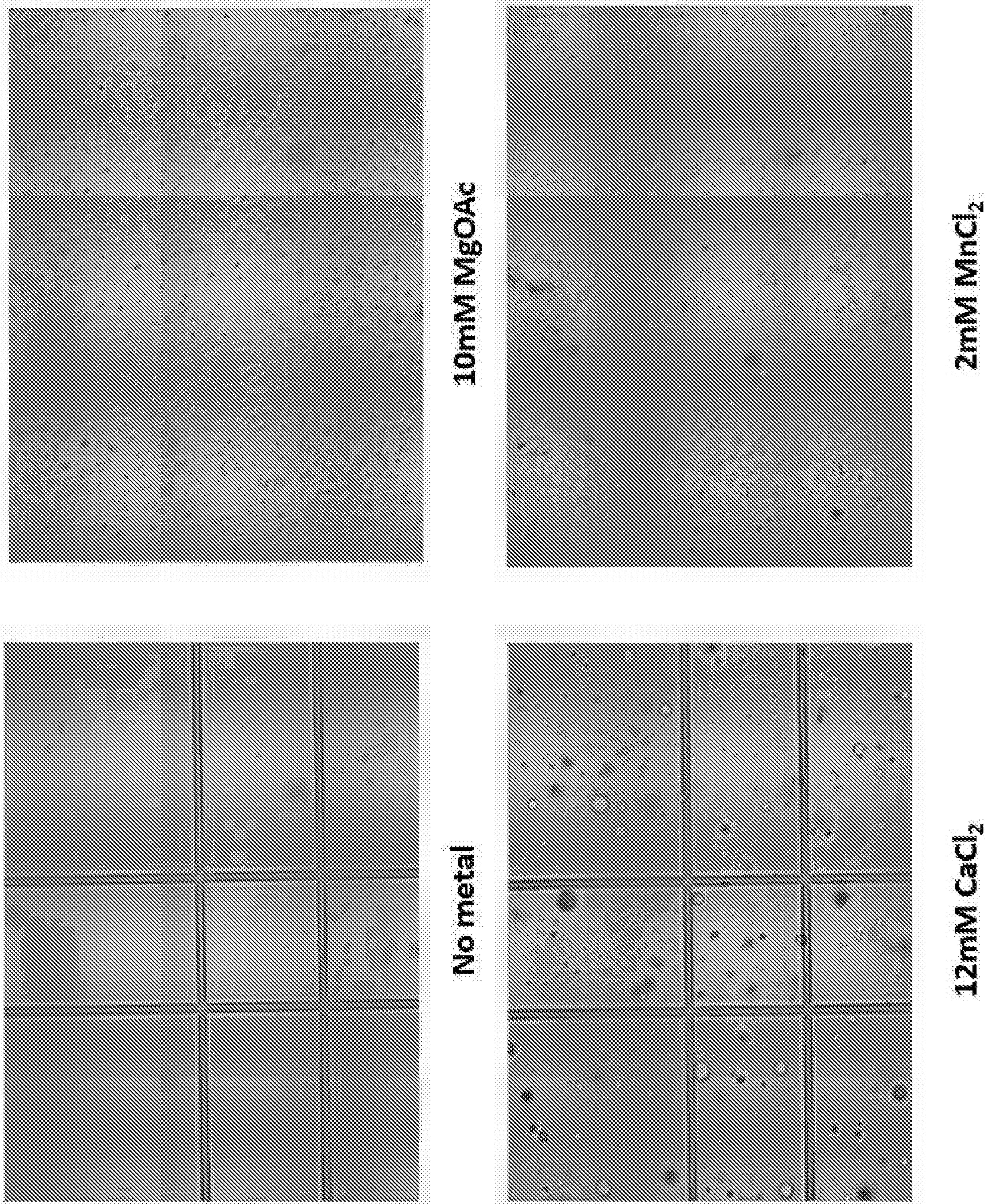
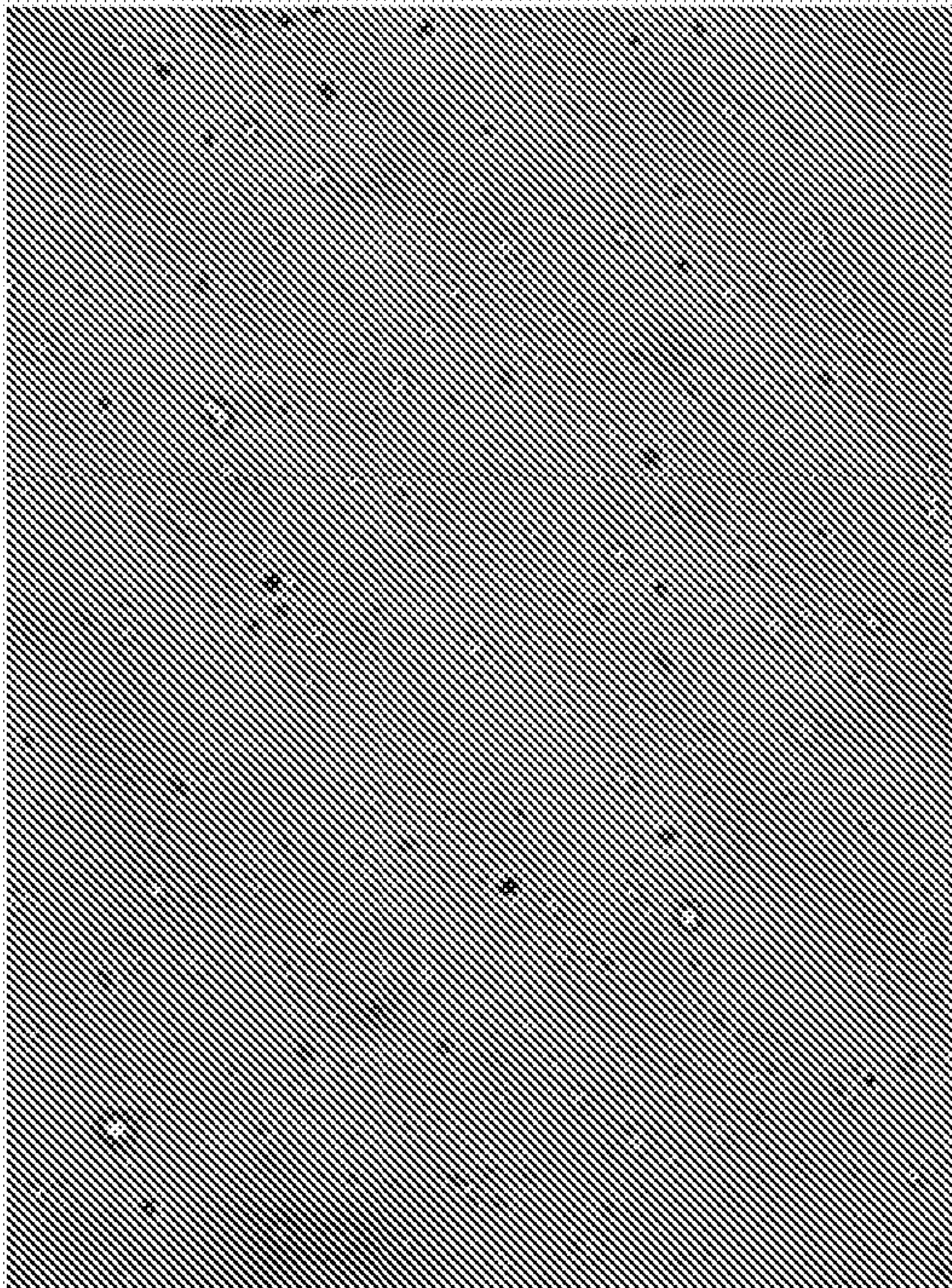
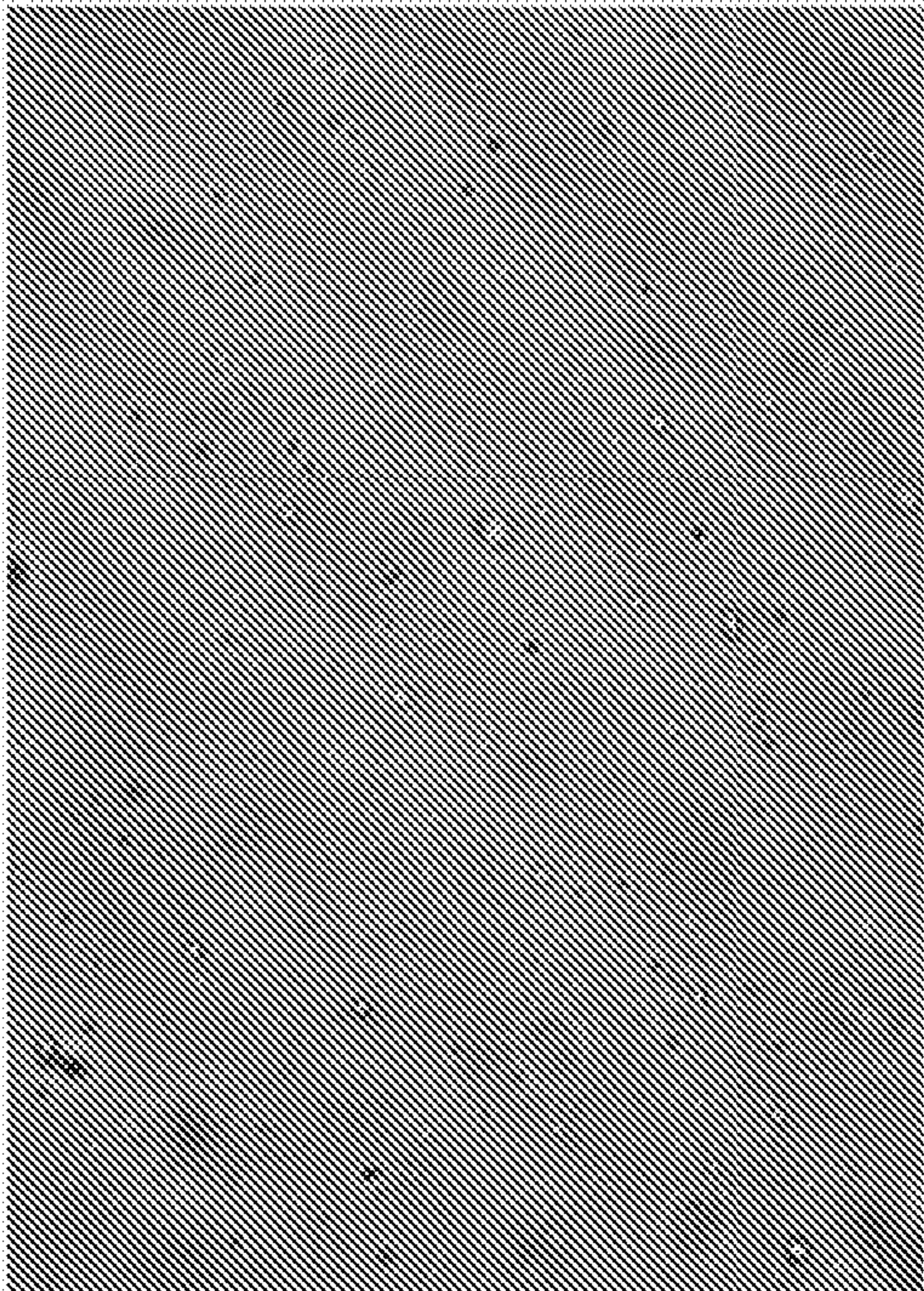


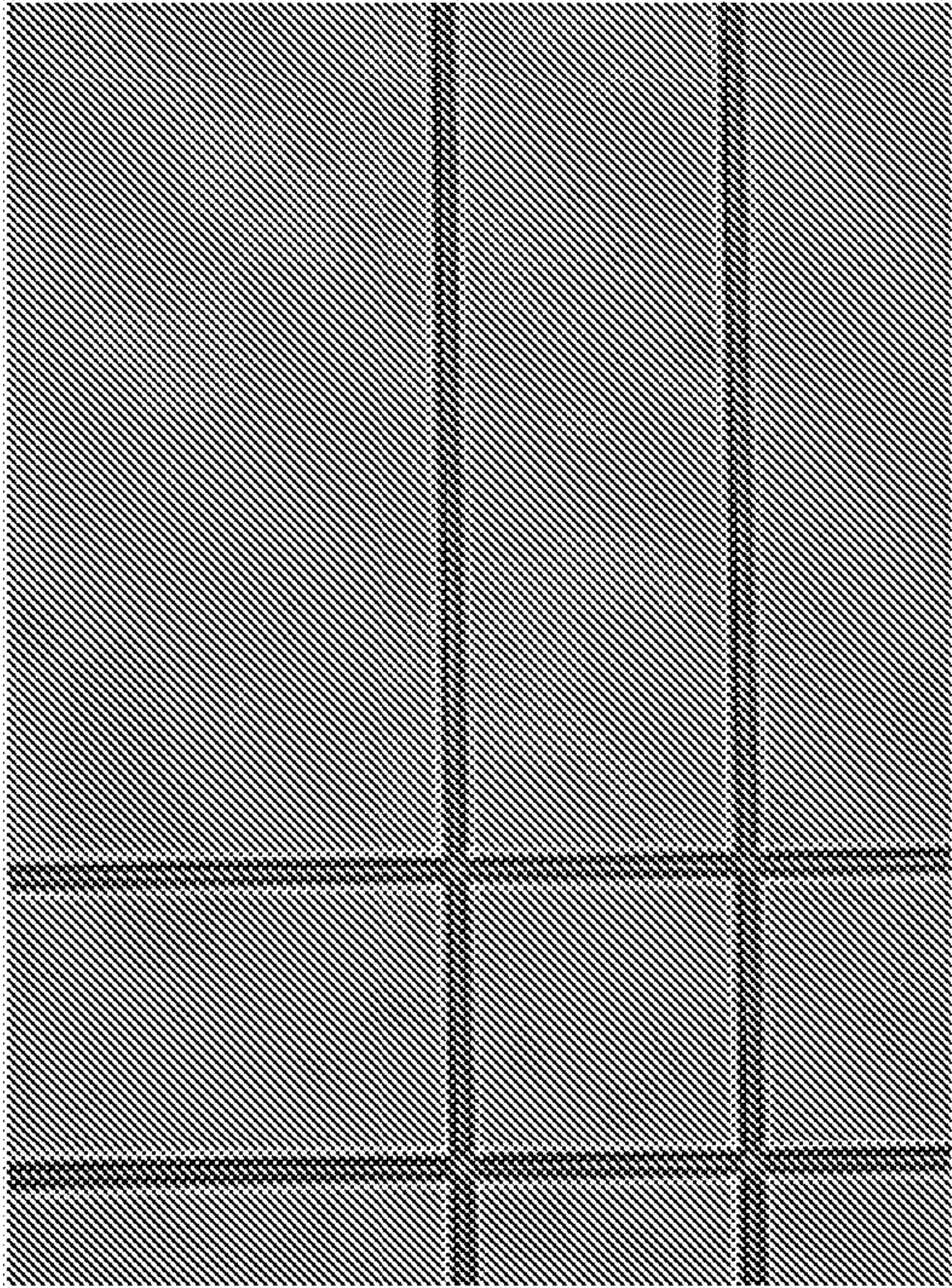
Figure 10C



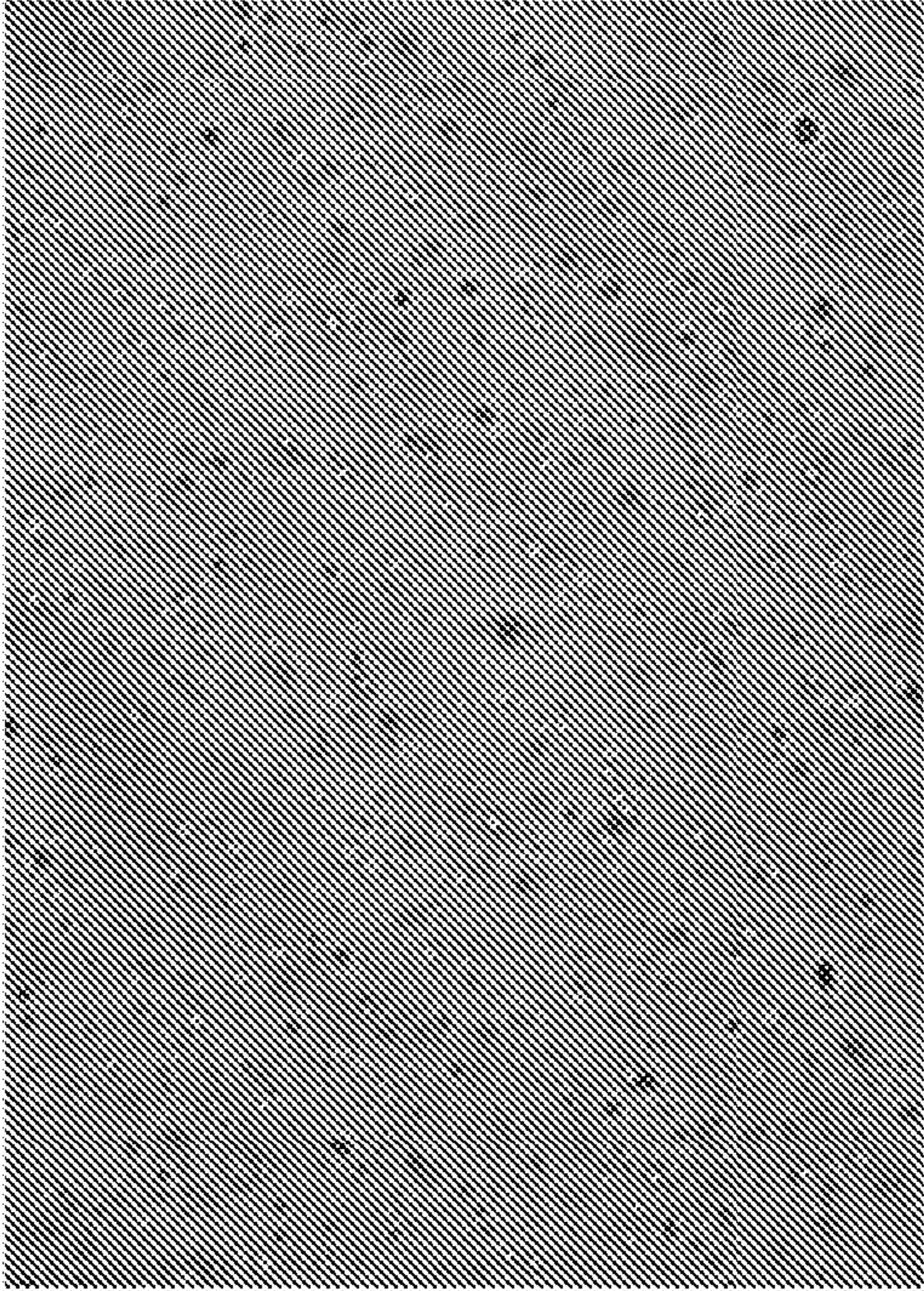
10mM MgOAc



2mM MnCl₂



No metal



12mM CaCl₂

Figure 10D

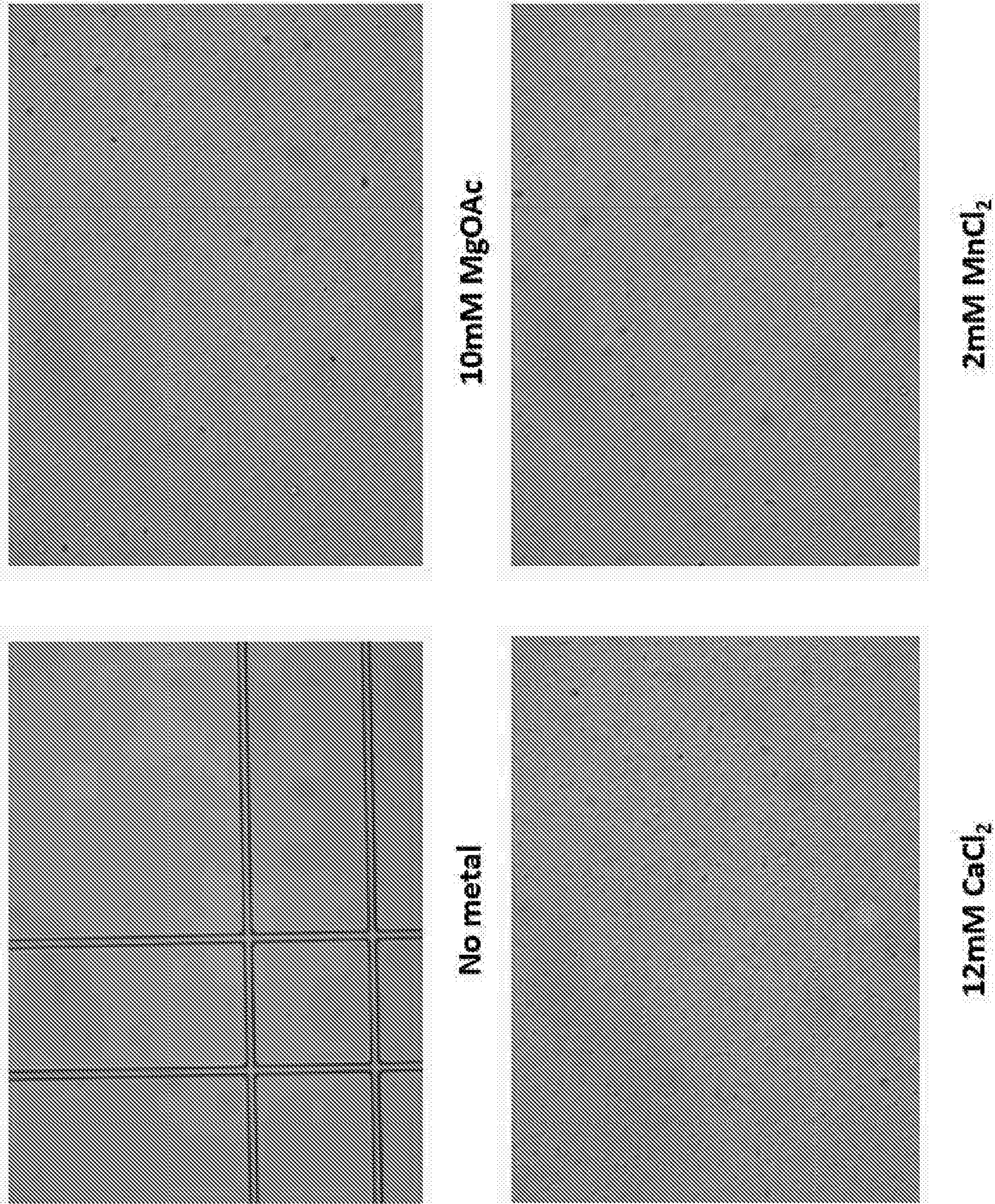


Figure 11A

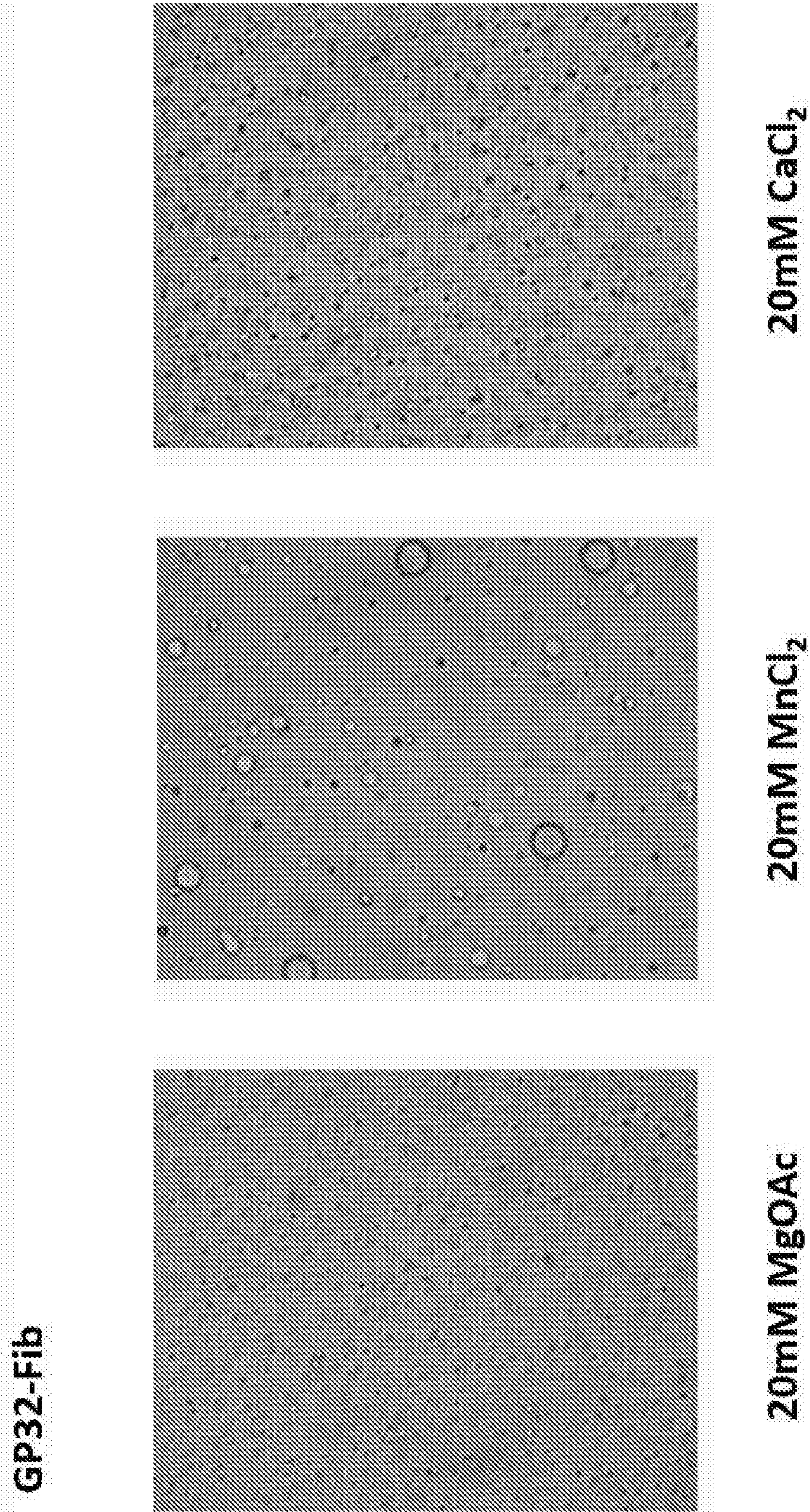


Figure 11B

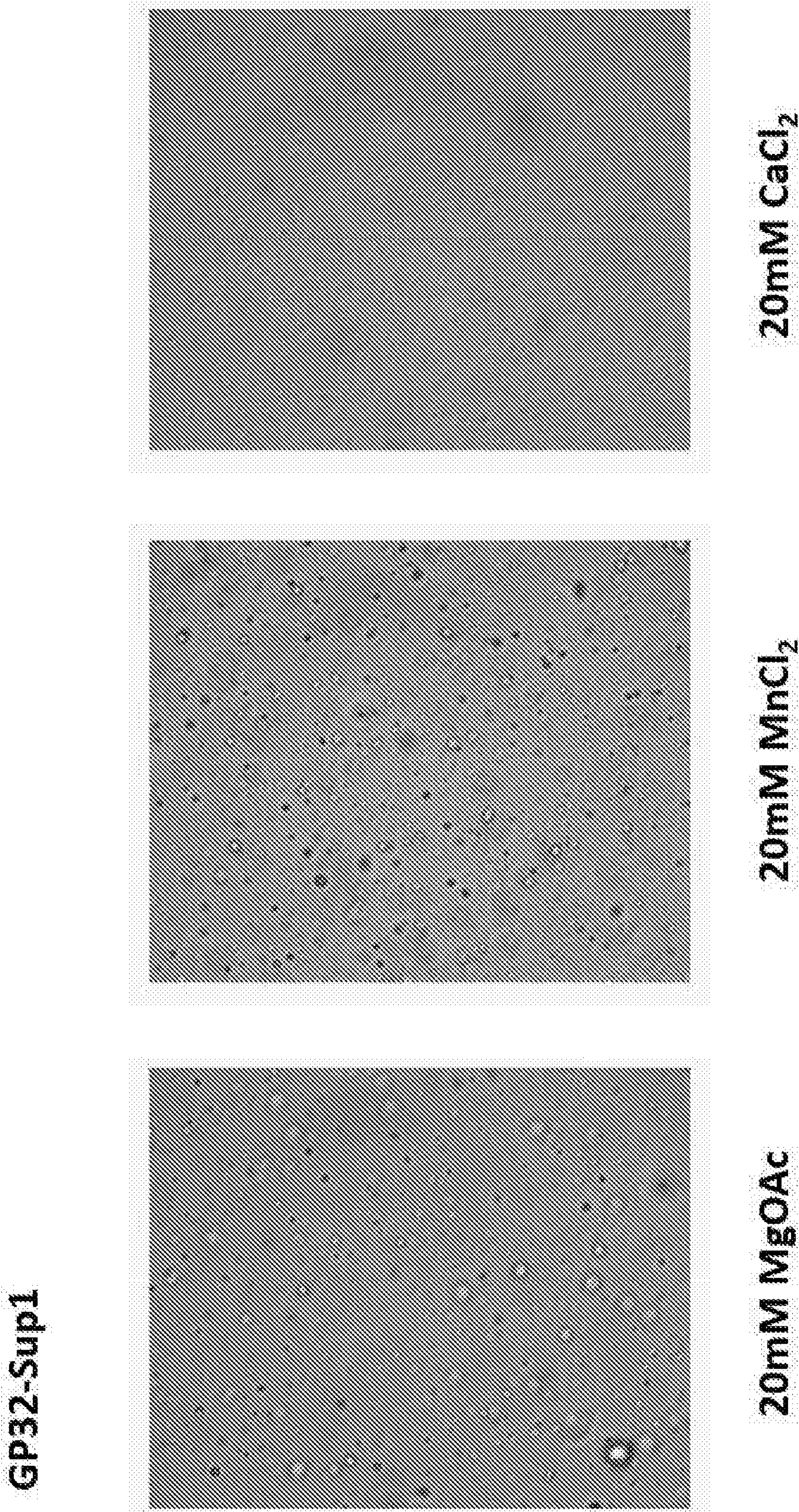
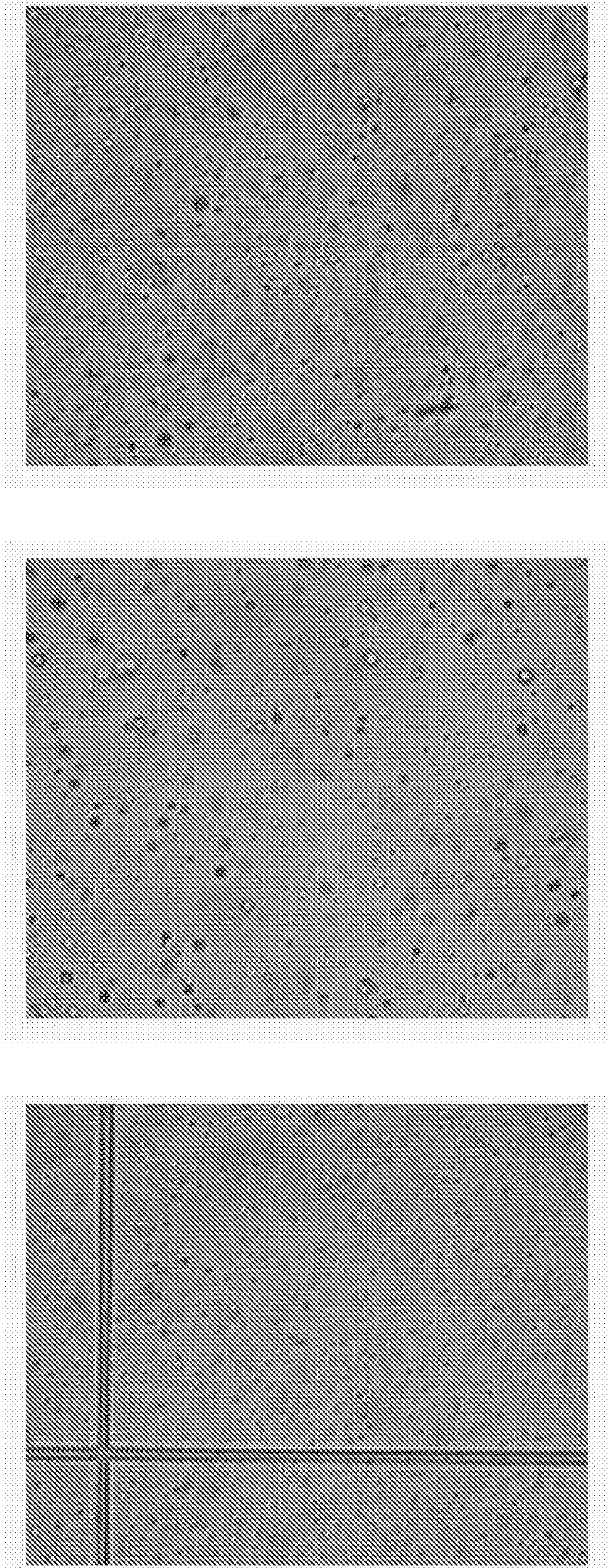


Figure 11C

GP32-his2



20mM MgOAc

20mM MnCl₂

20mM CaCl₂

Figure 11D

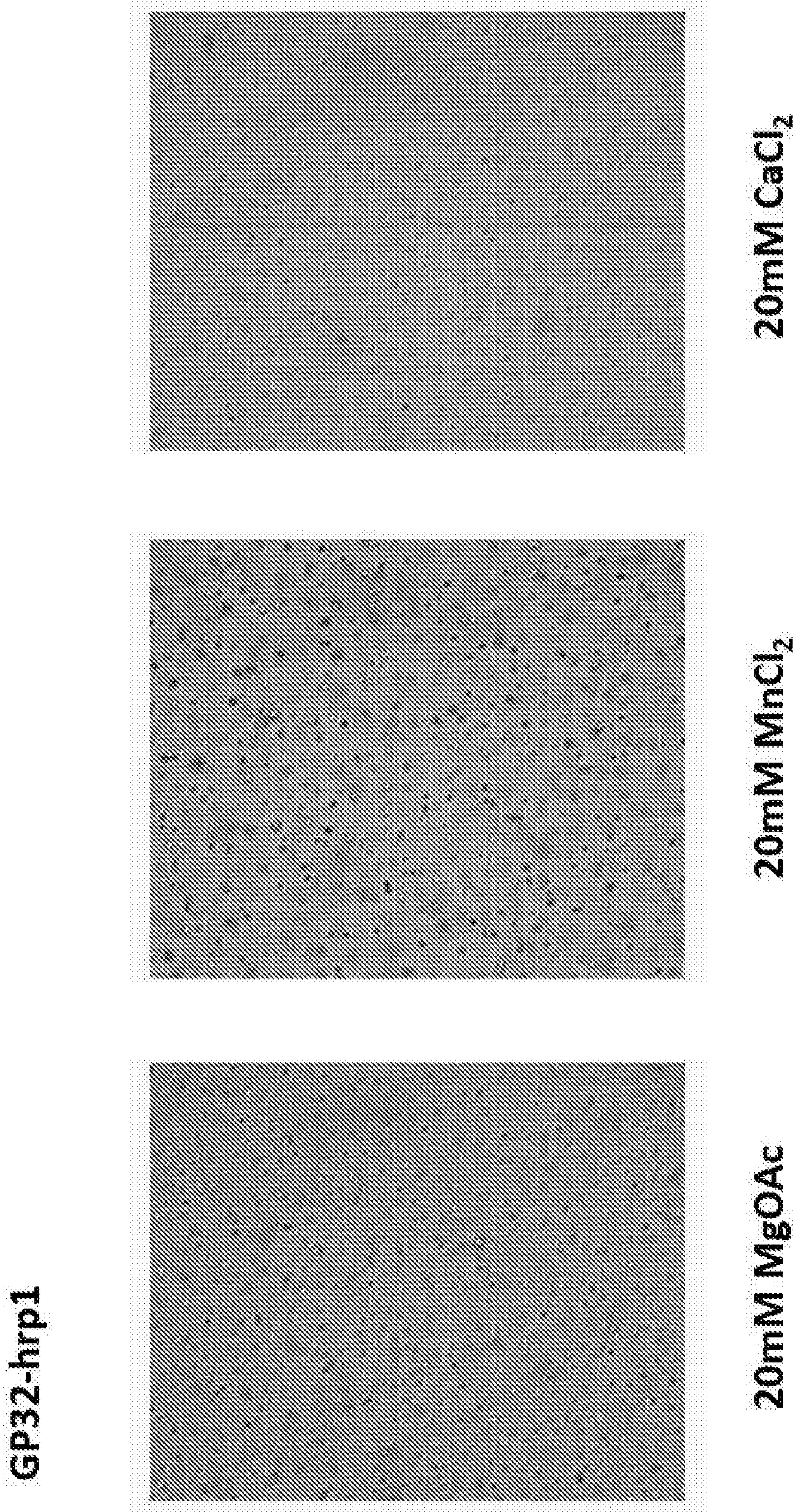
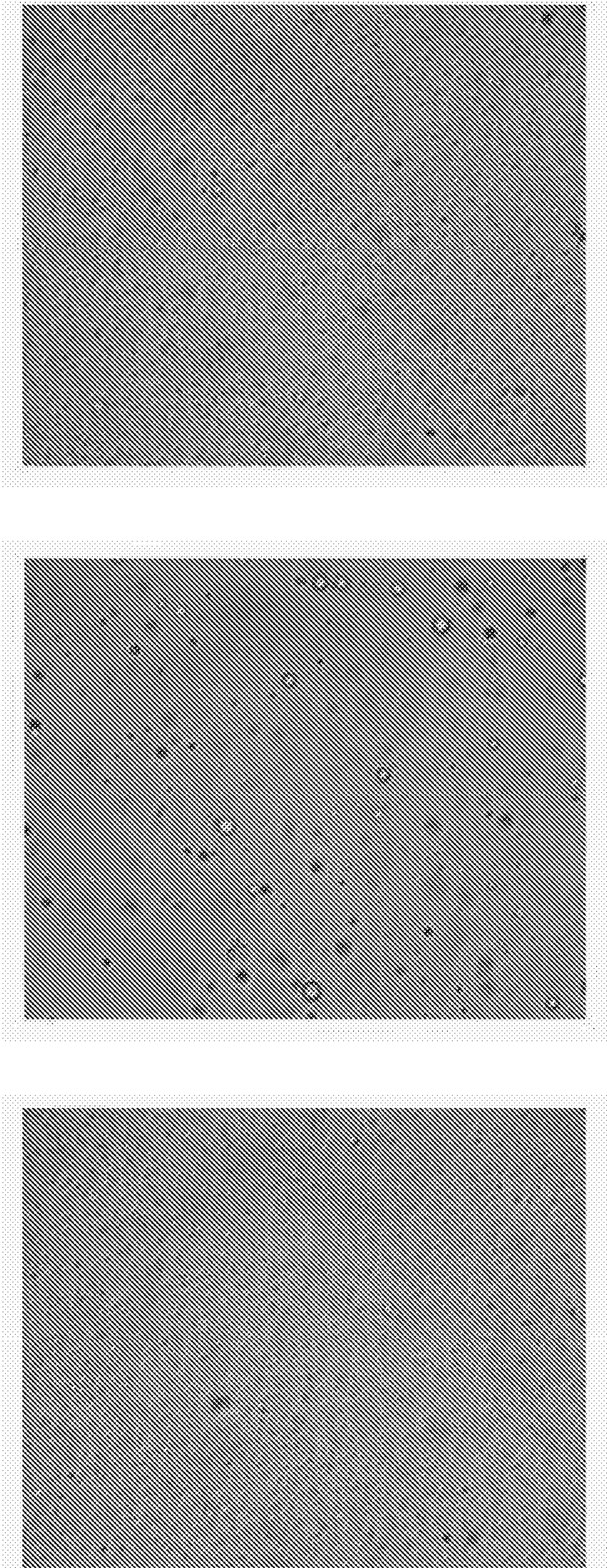


Figure 11E

GP32-his5



20mM MgOAc

20mM MnCl₂

20mM CaCl₂

Figure 12

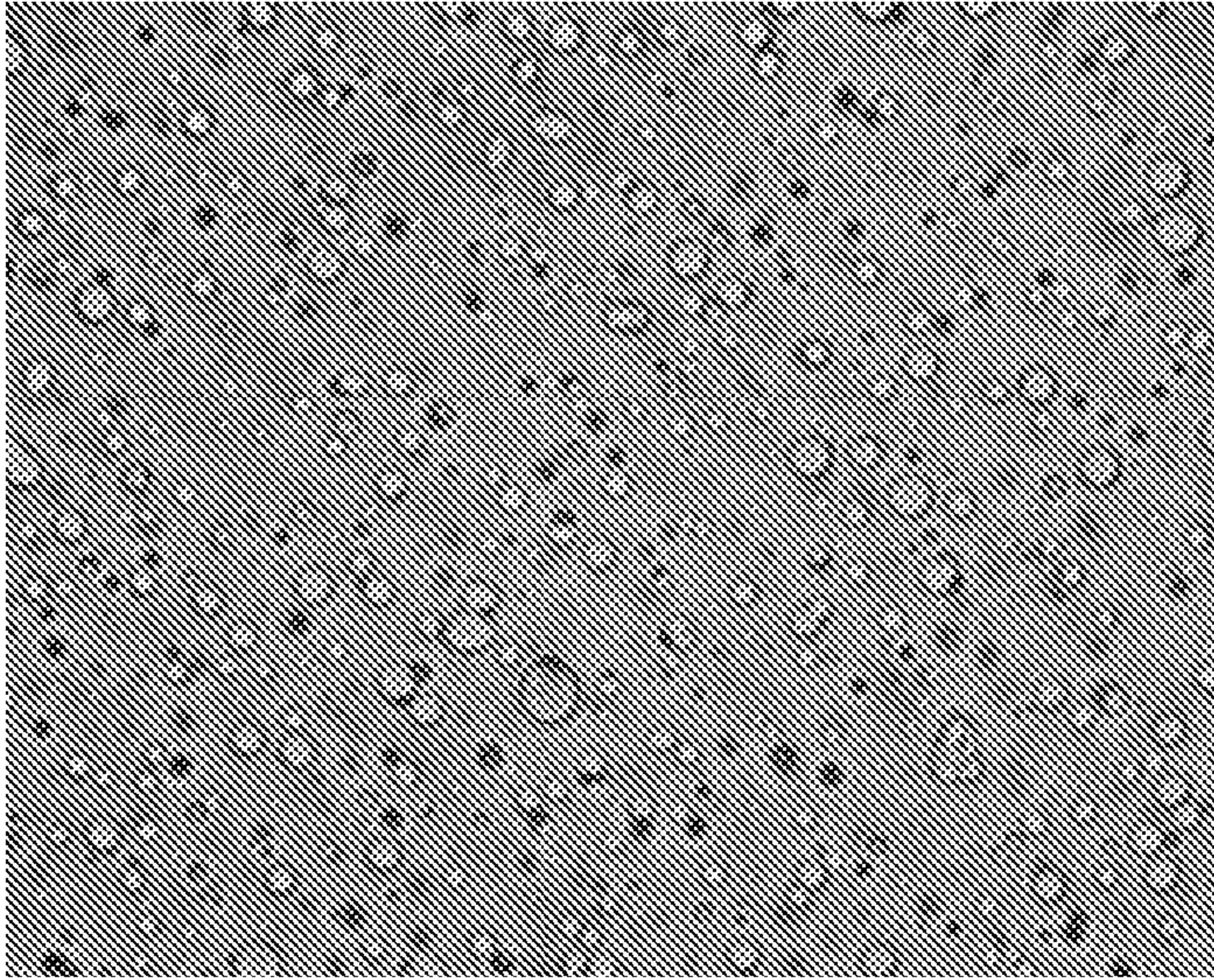


Image 1 – Bright field on glass cover slip

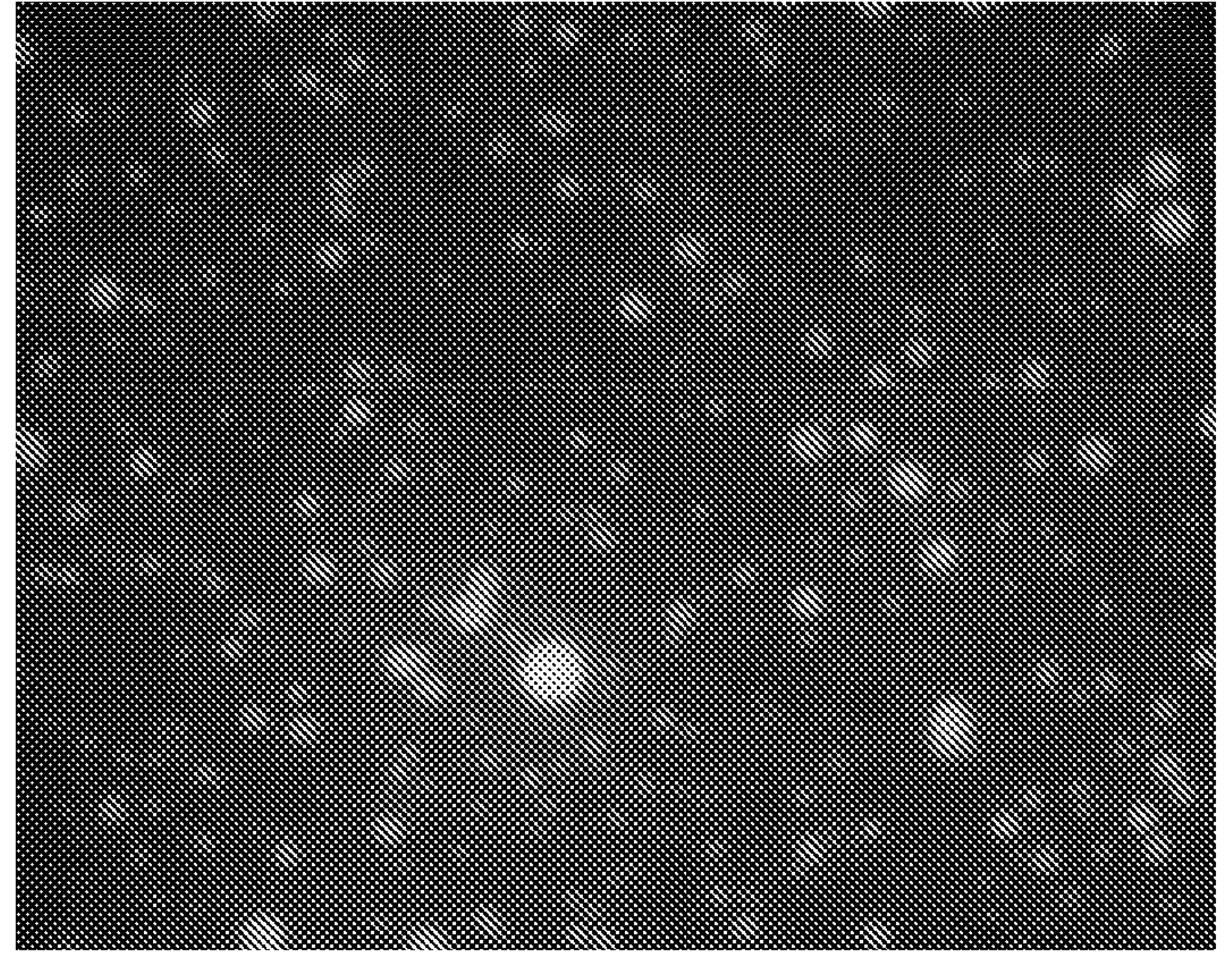


Image 1 – Fluorescence on glass cover slip

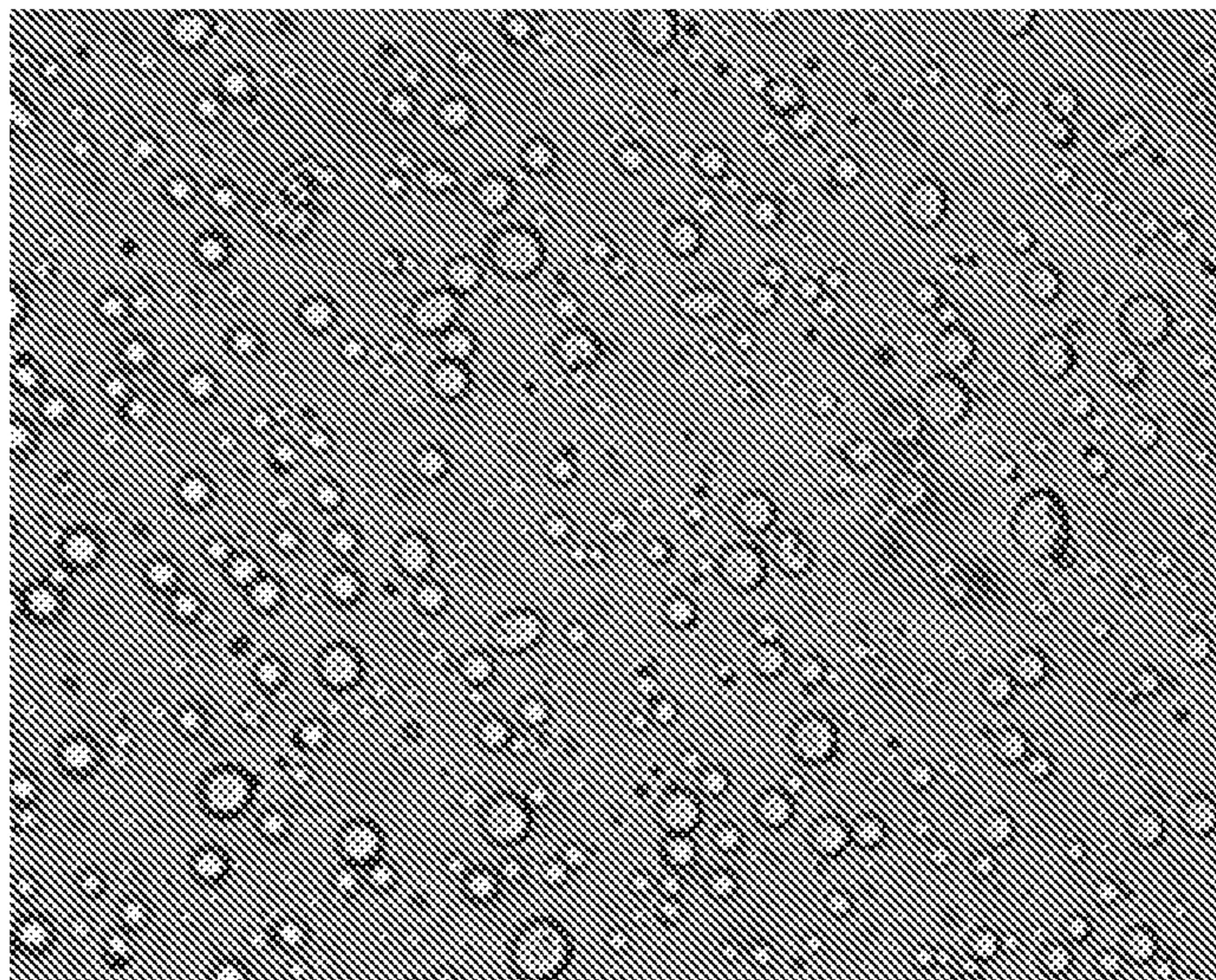


Image 2 – Bright field on C-Chip haemocytometer

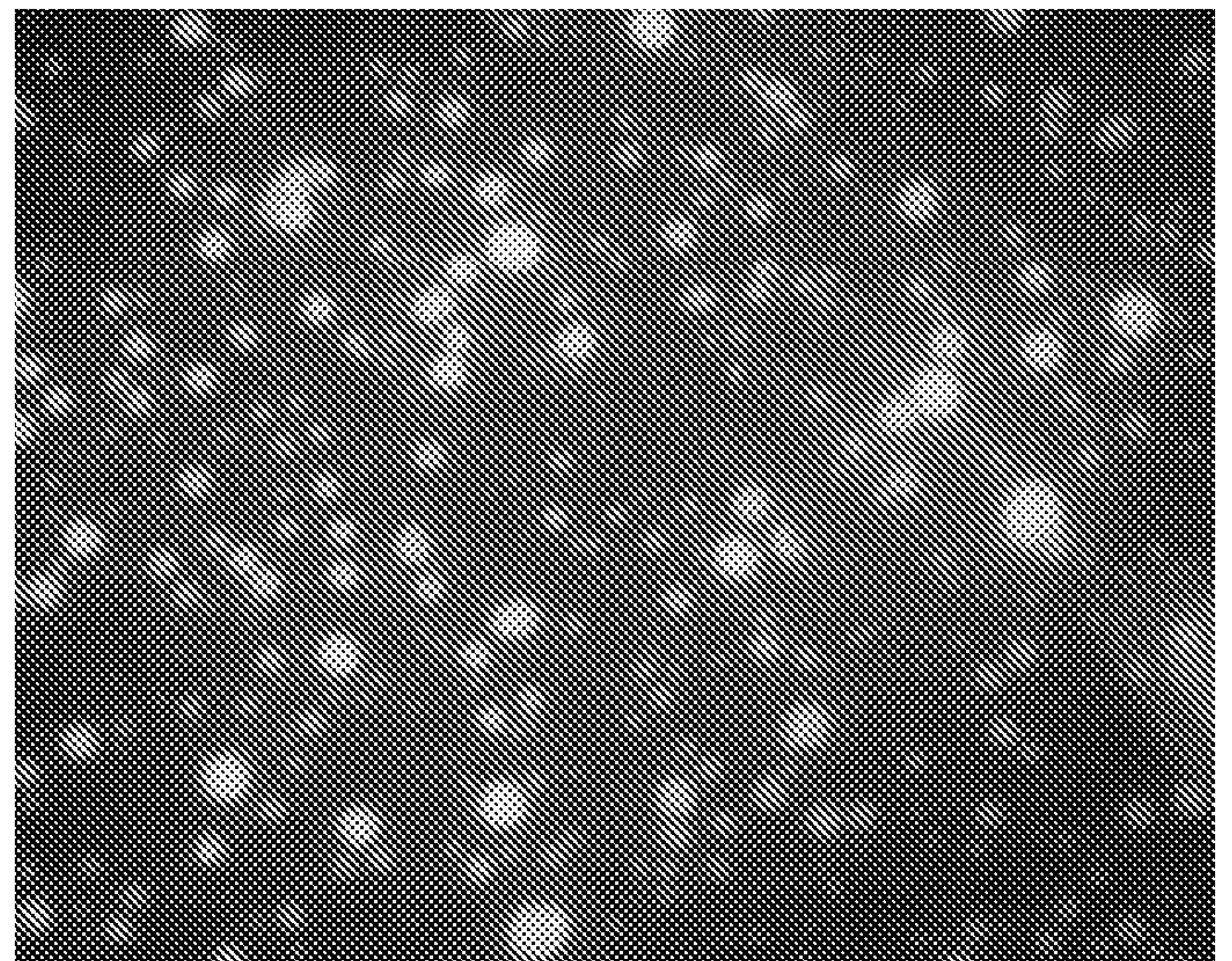


Image 2 – Fluorescence on C-Chip haemocytometer

Figure 13

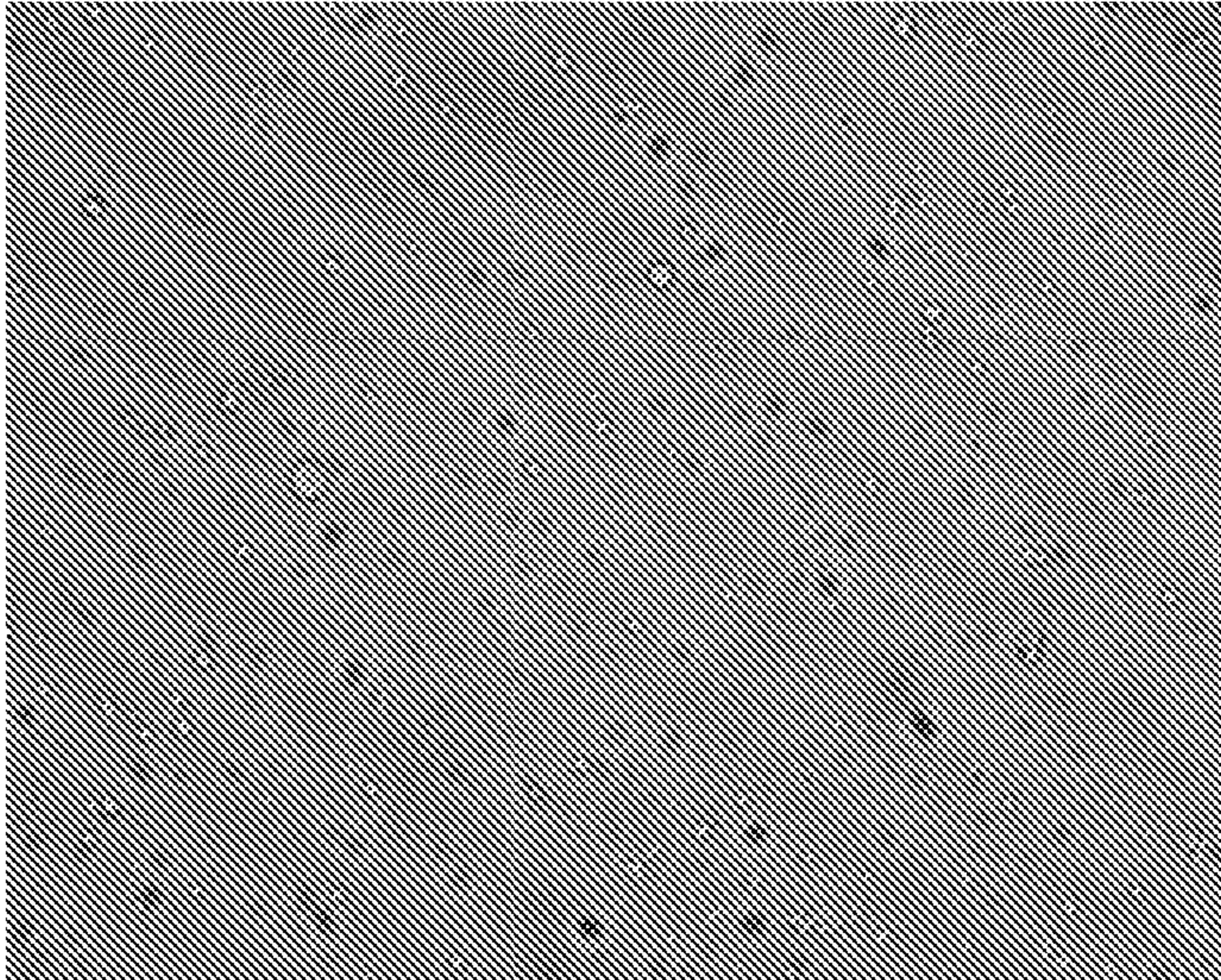


Image 1 – Bright field on glass cover slip

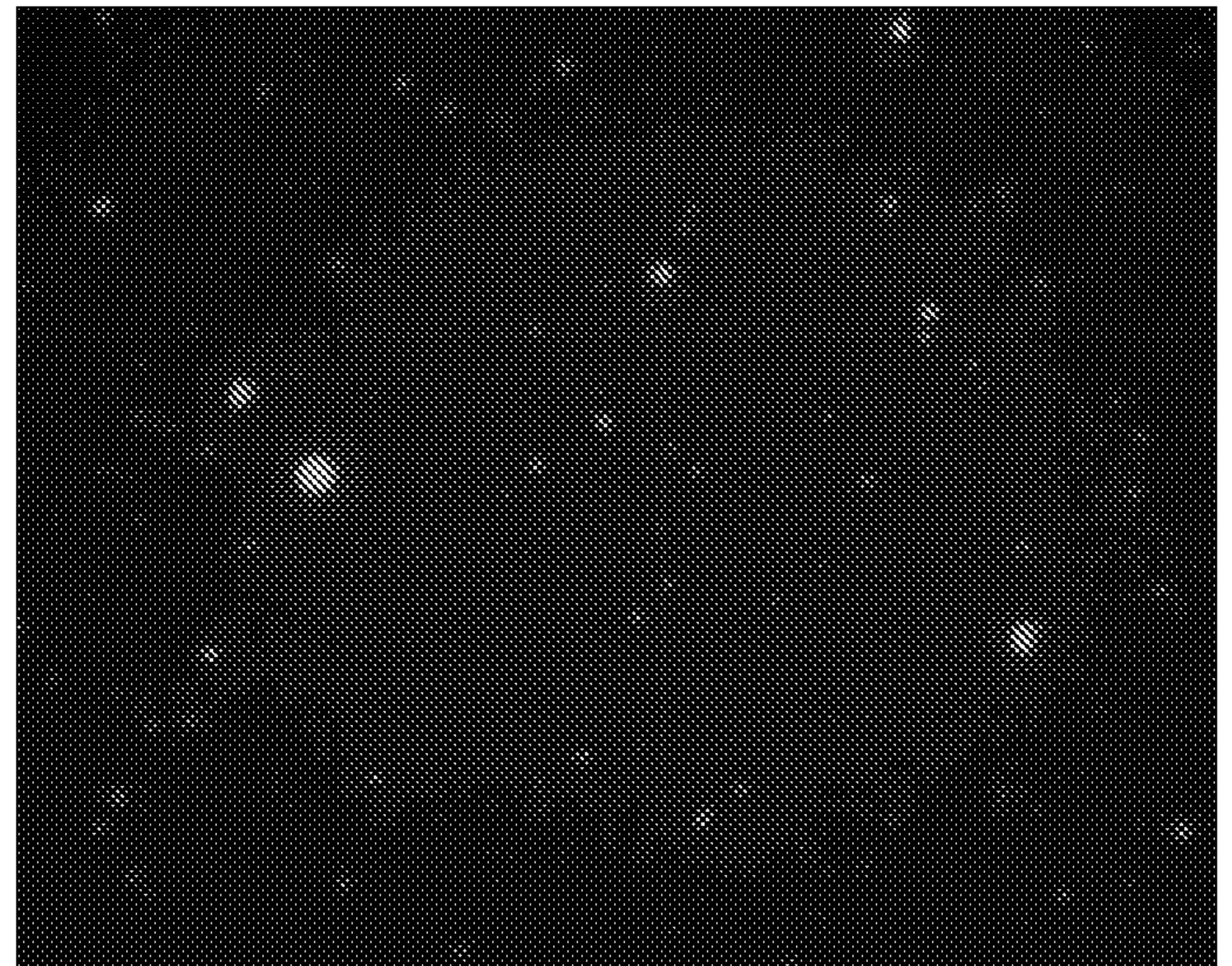


Image 1 – Fluorescence on glass cover slip

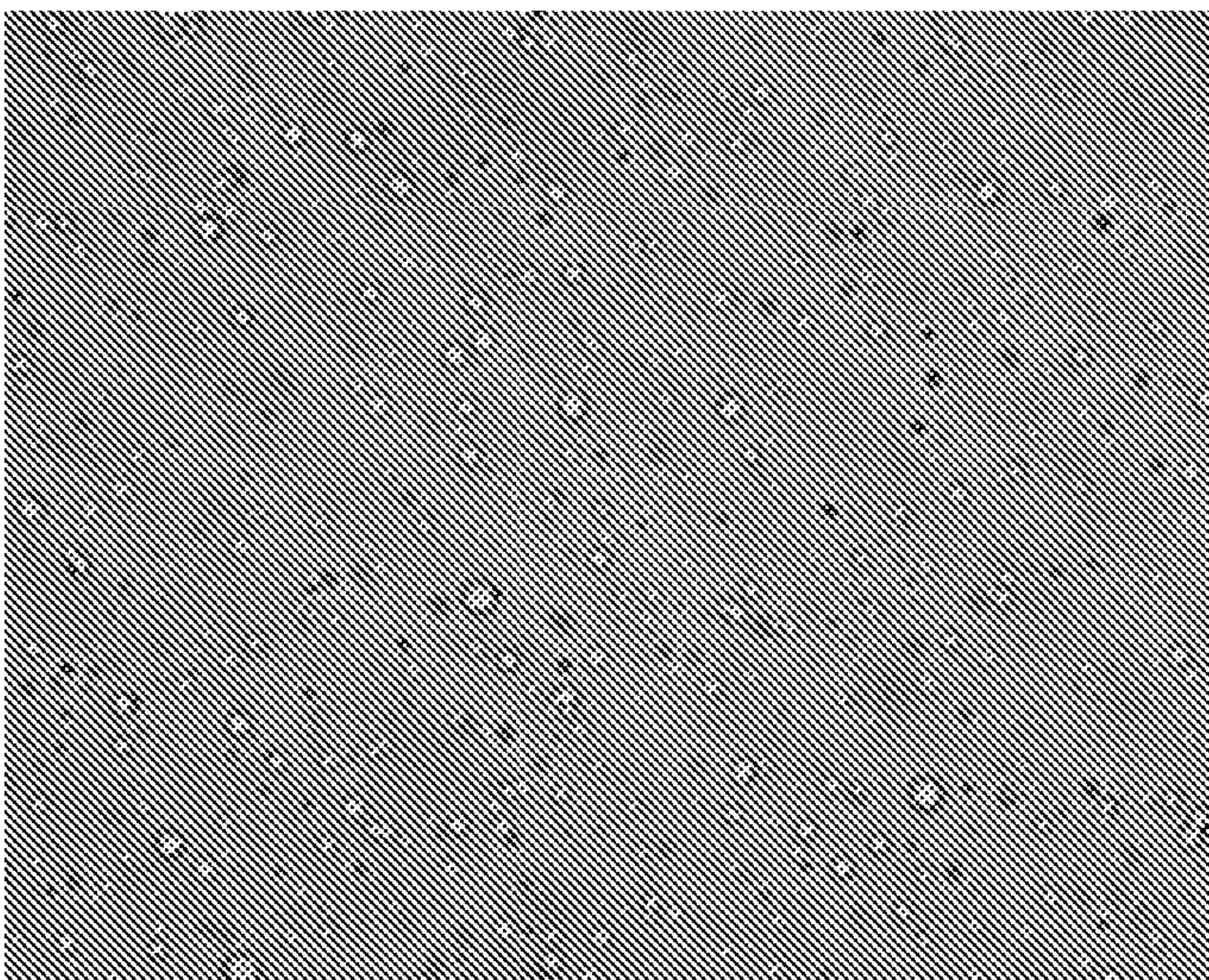


**Image 2 – Bright field on C-Chip
haemocytometer**

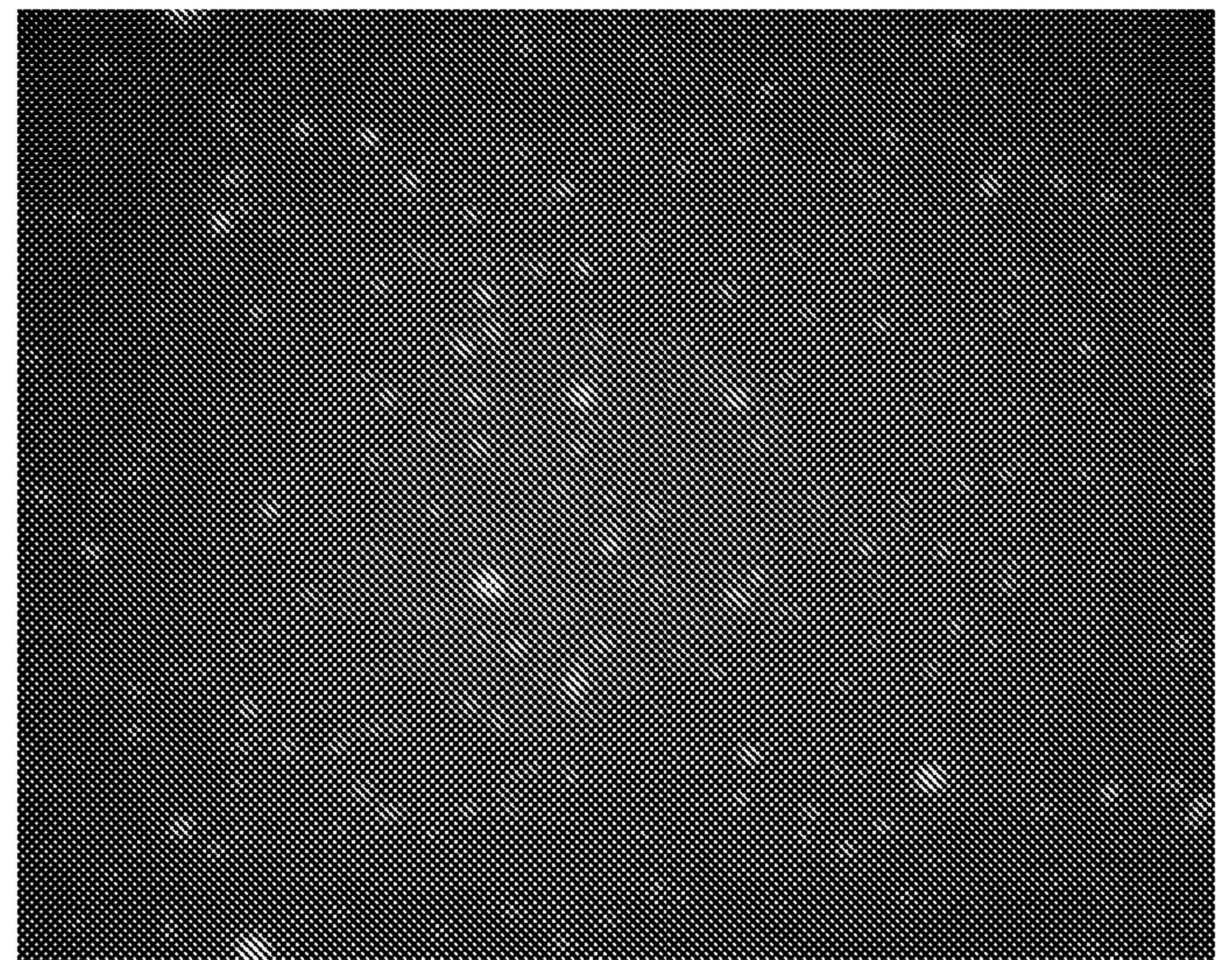


**Image 2 – Fluorescence on C-
Chip haemocytometer**

Figure 14A

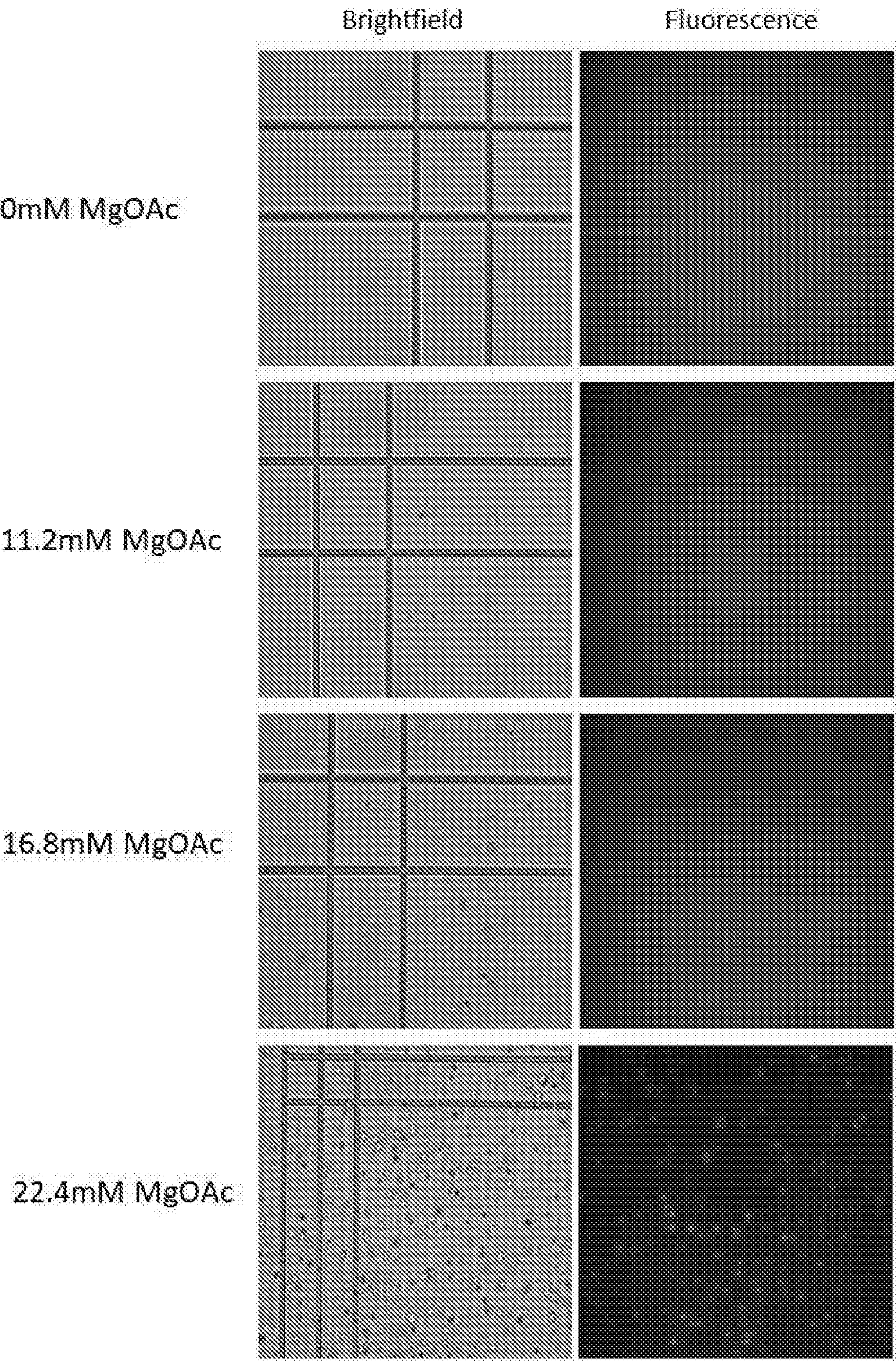


Figure 14B

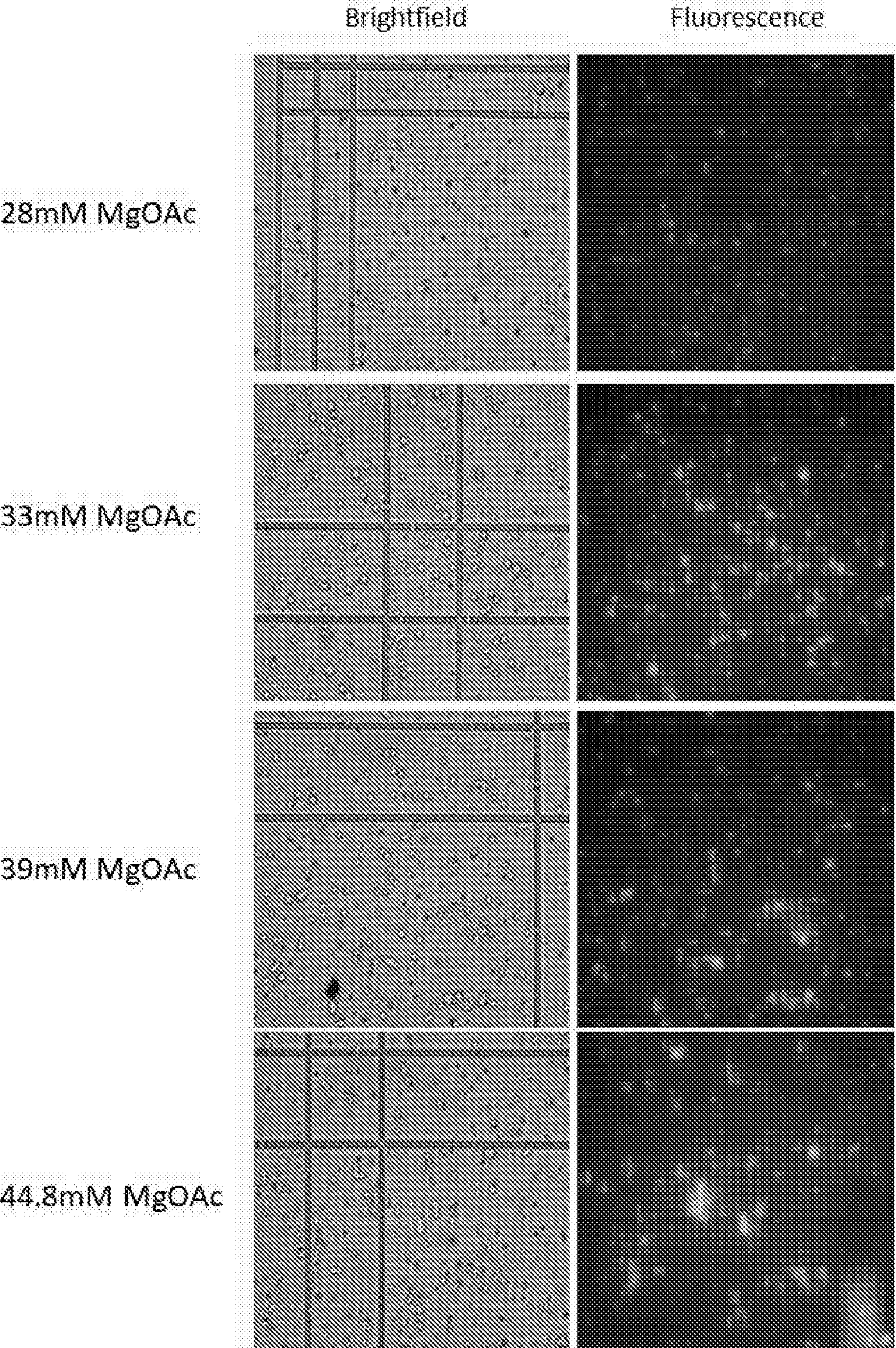


Figure 15A

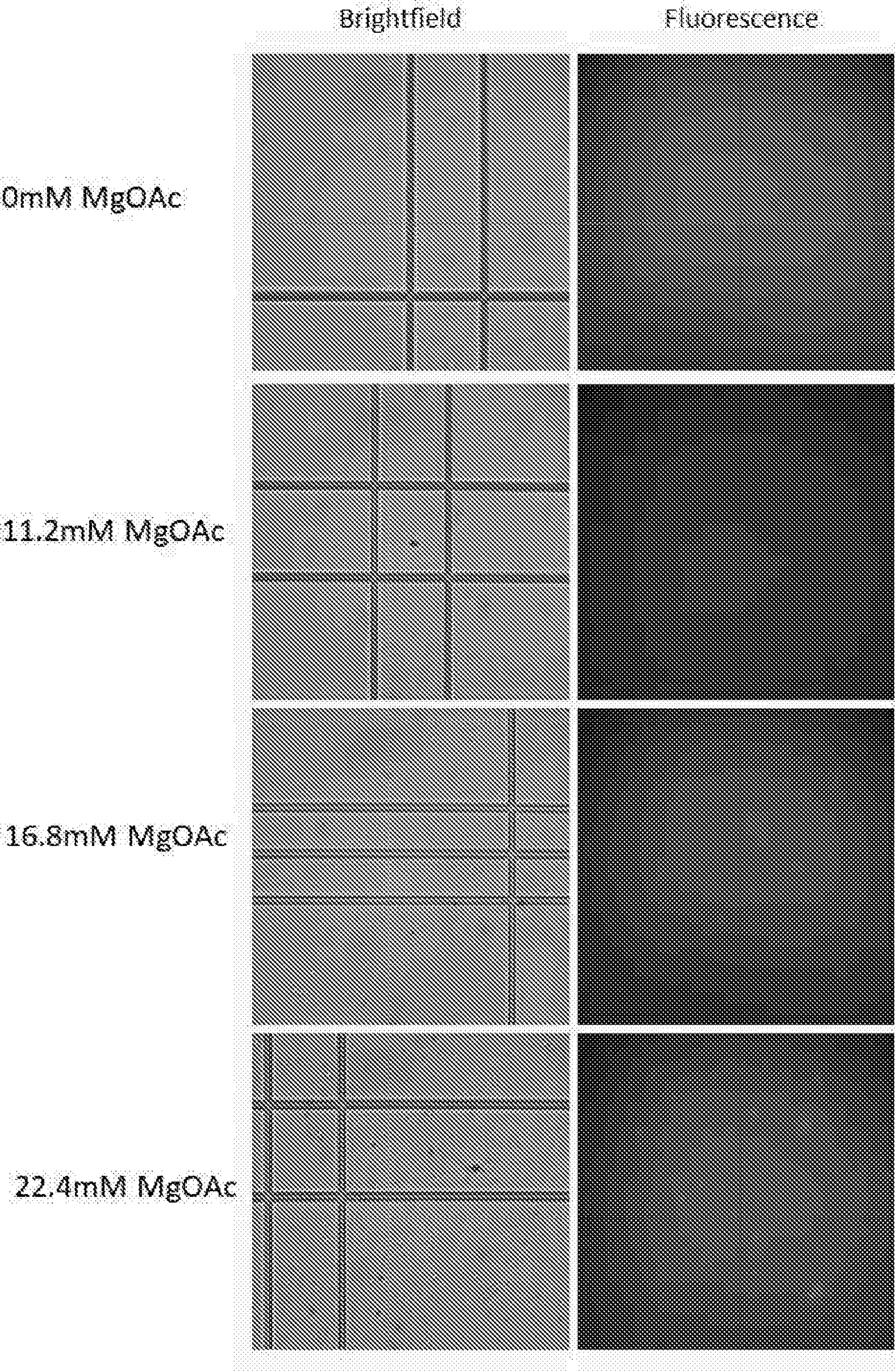


Figure 15B

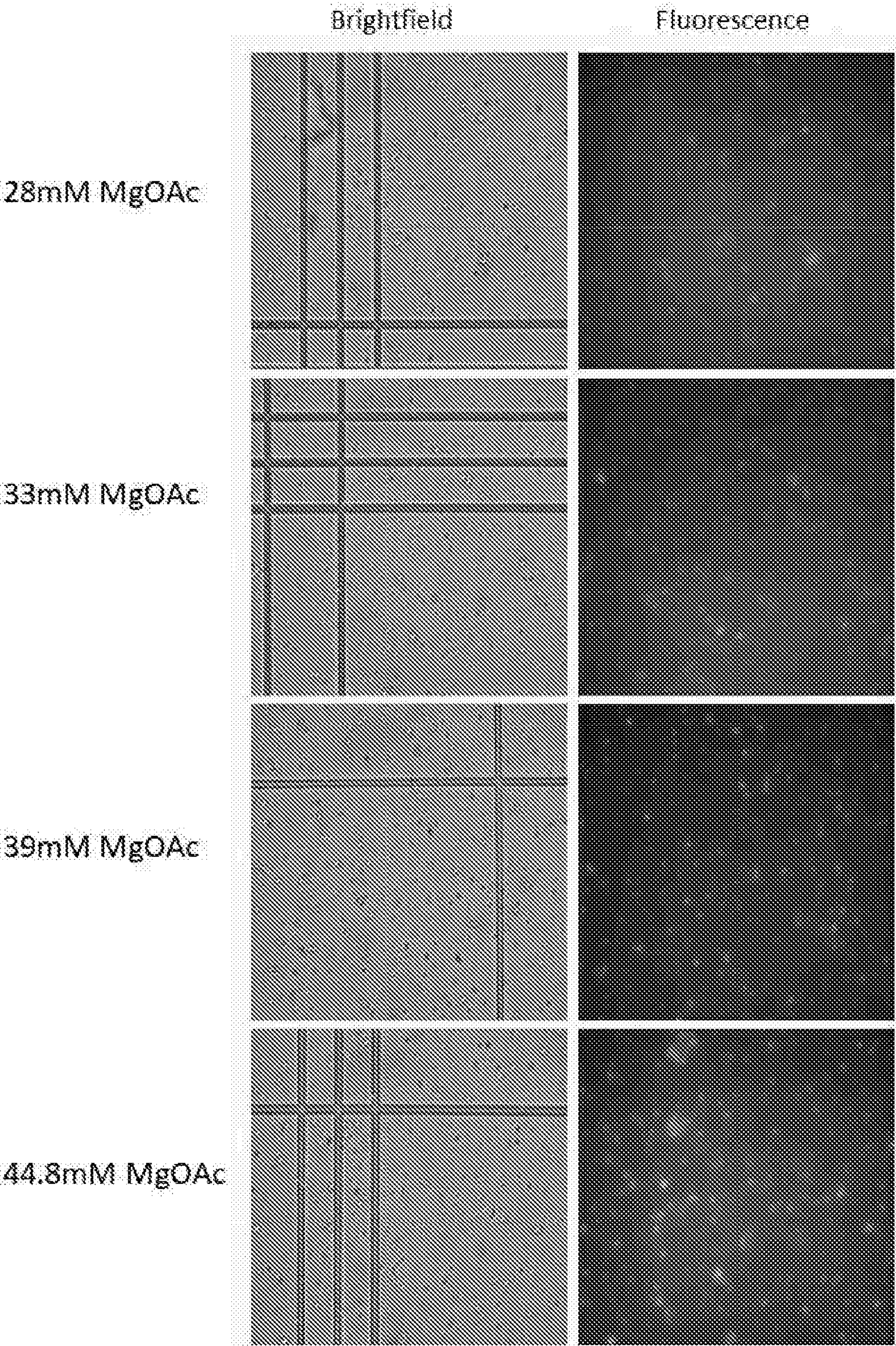


Figure 16A

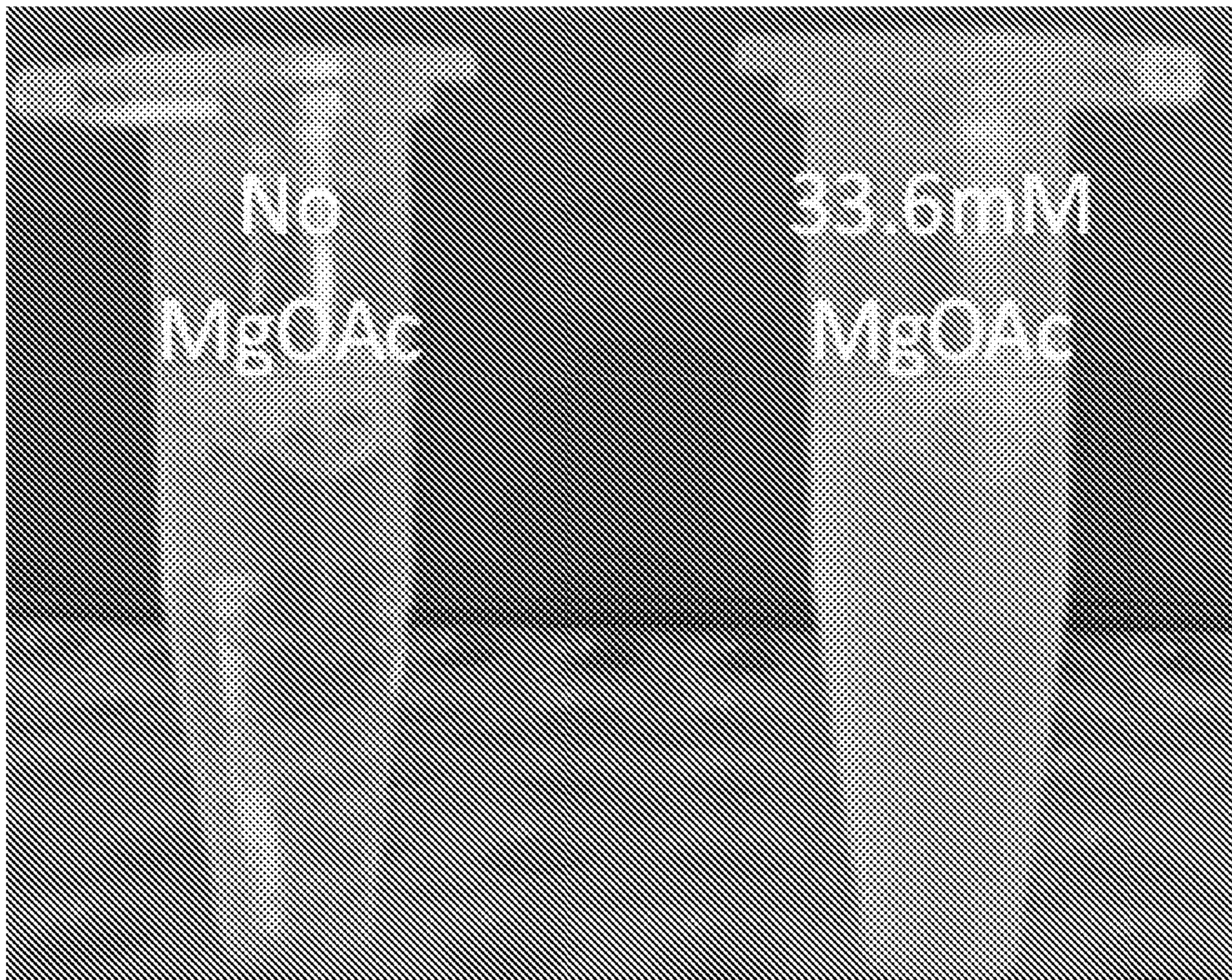


Figure 16B

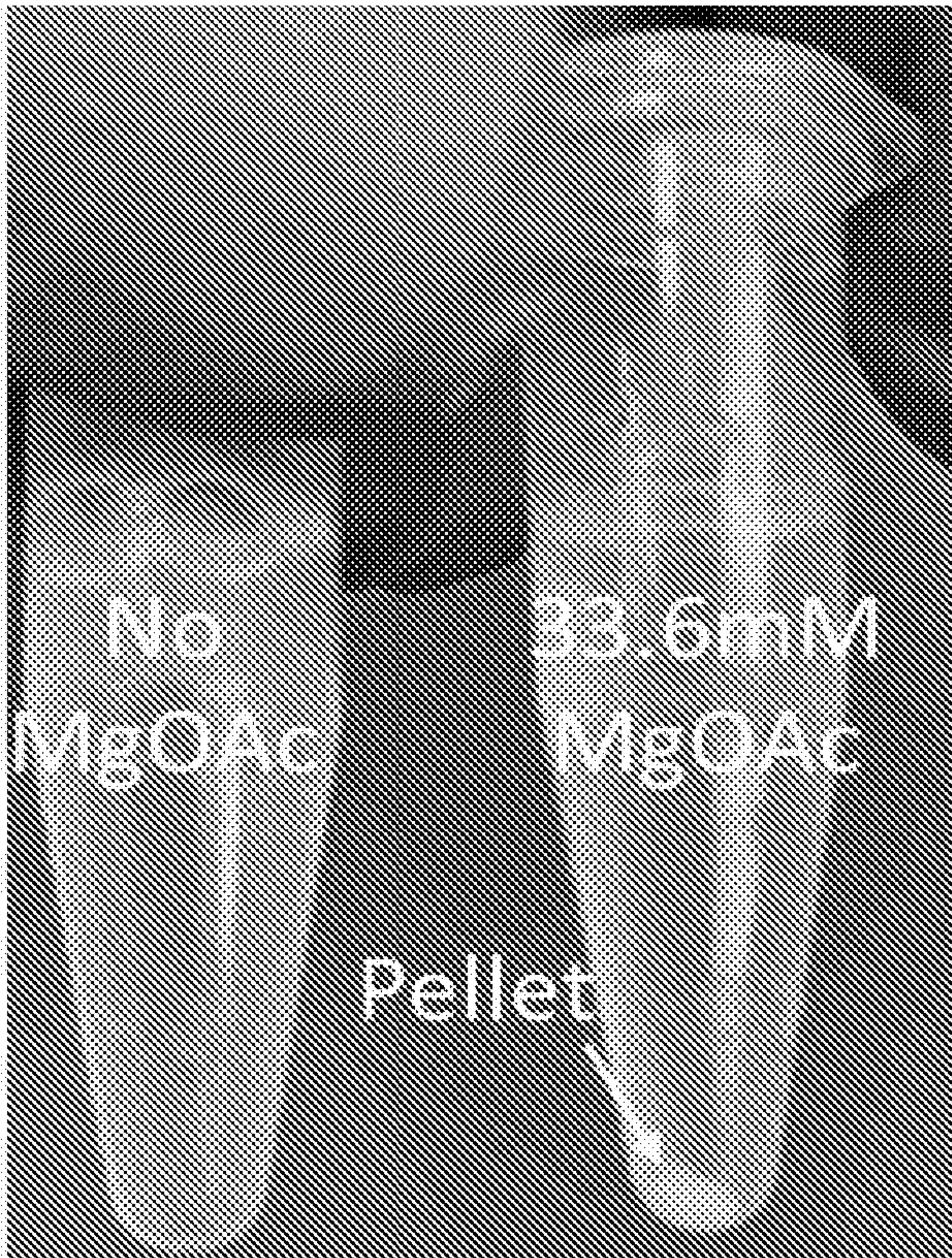


Figure 16C

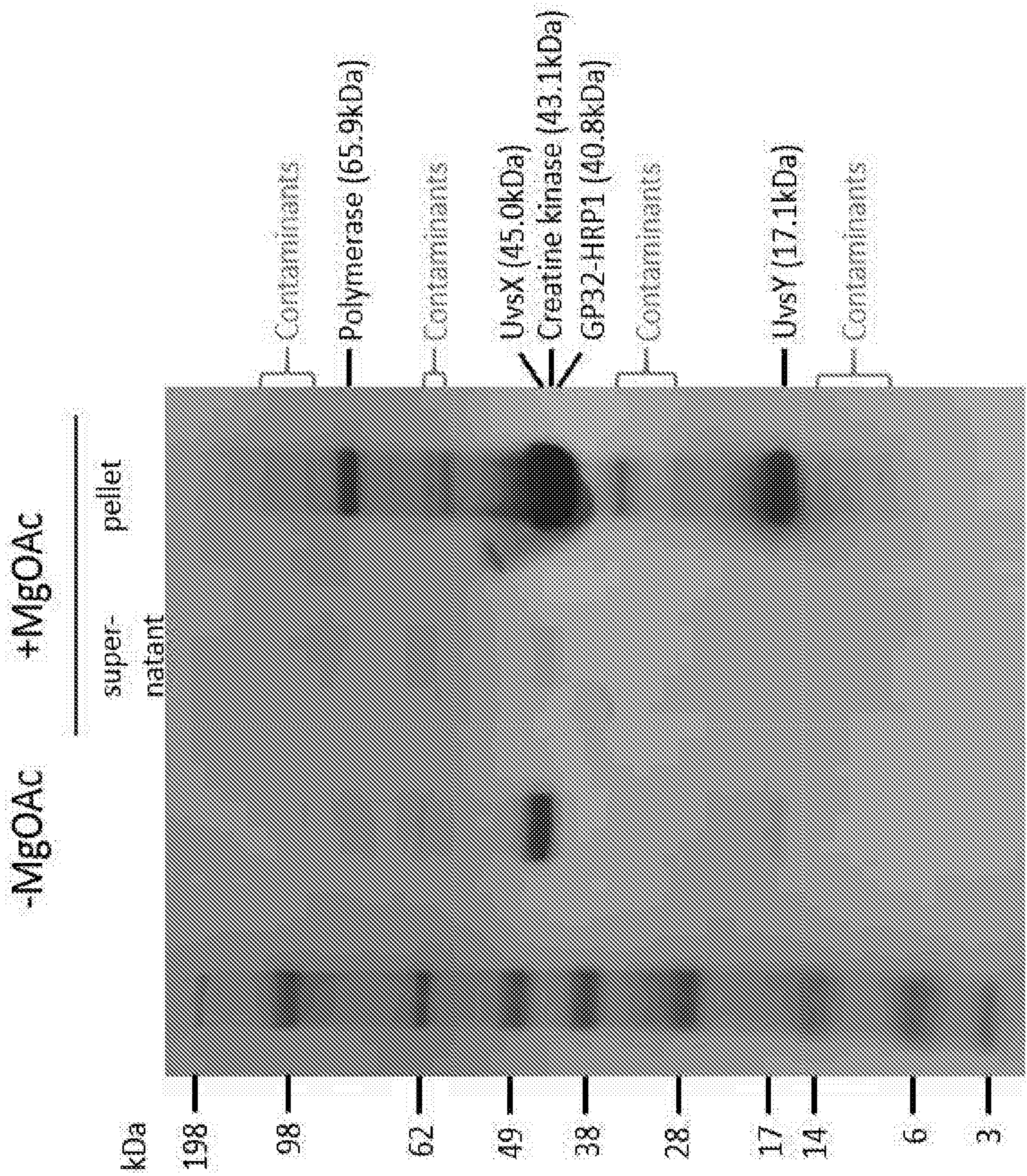


Figure 17

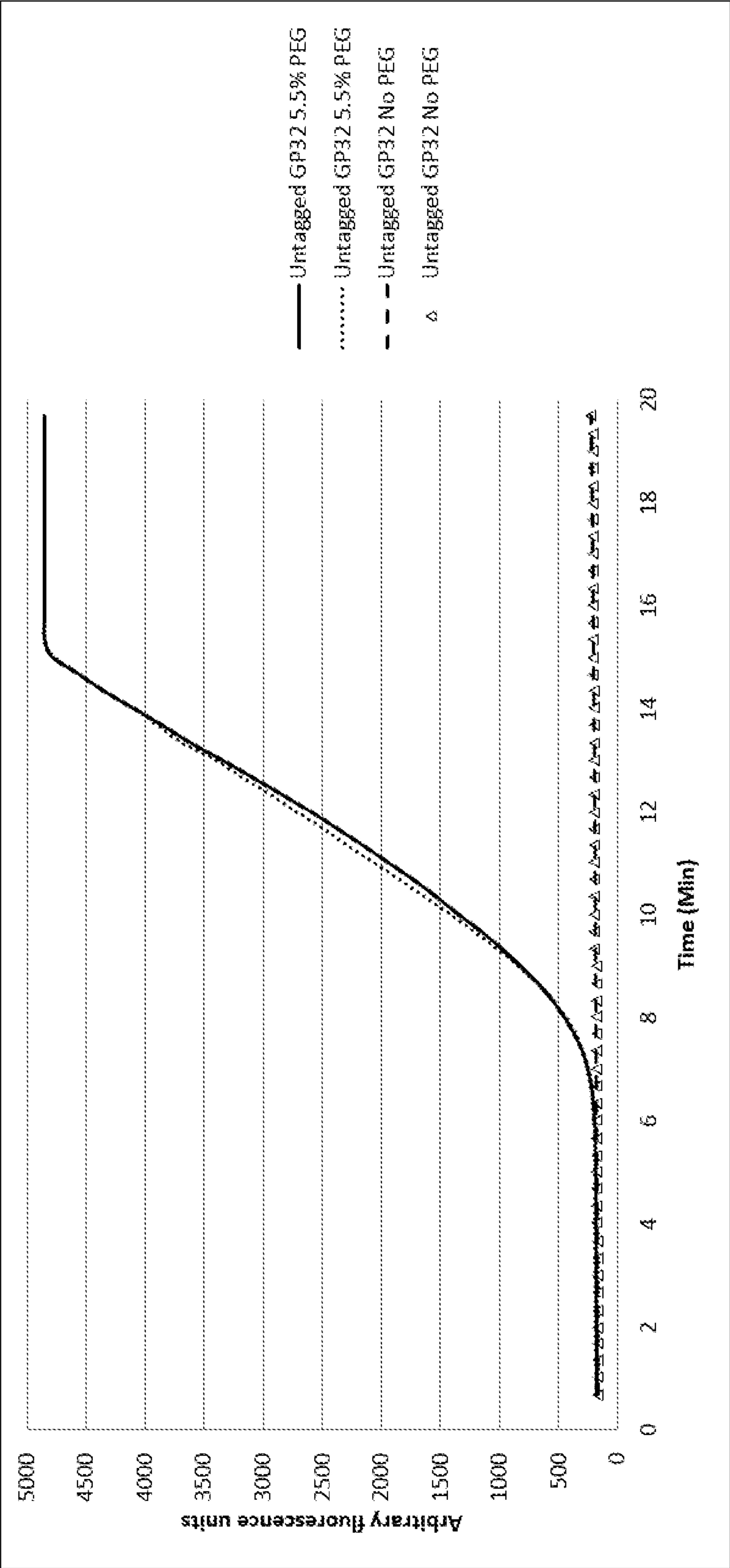


Figure 18A

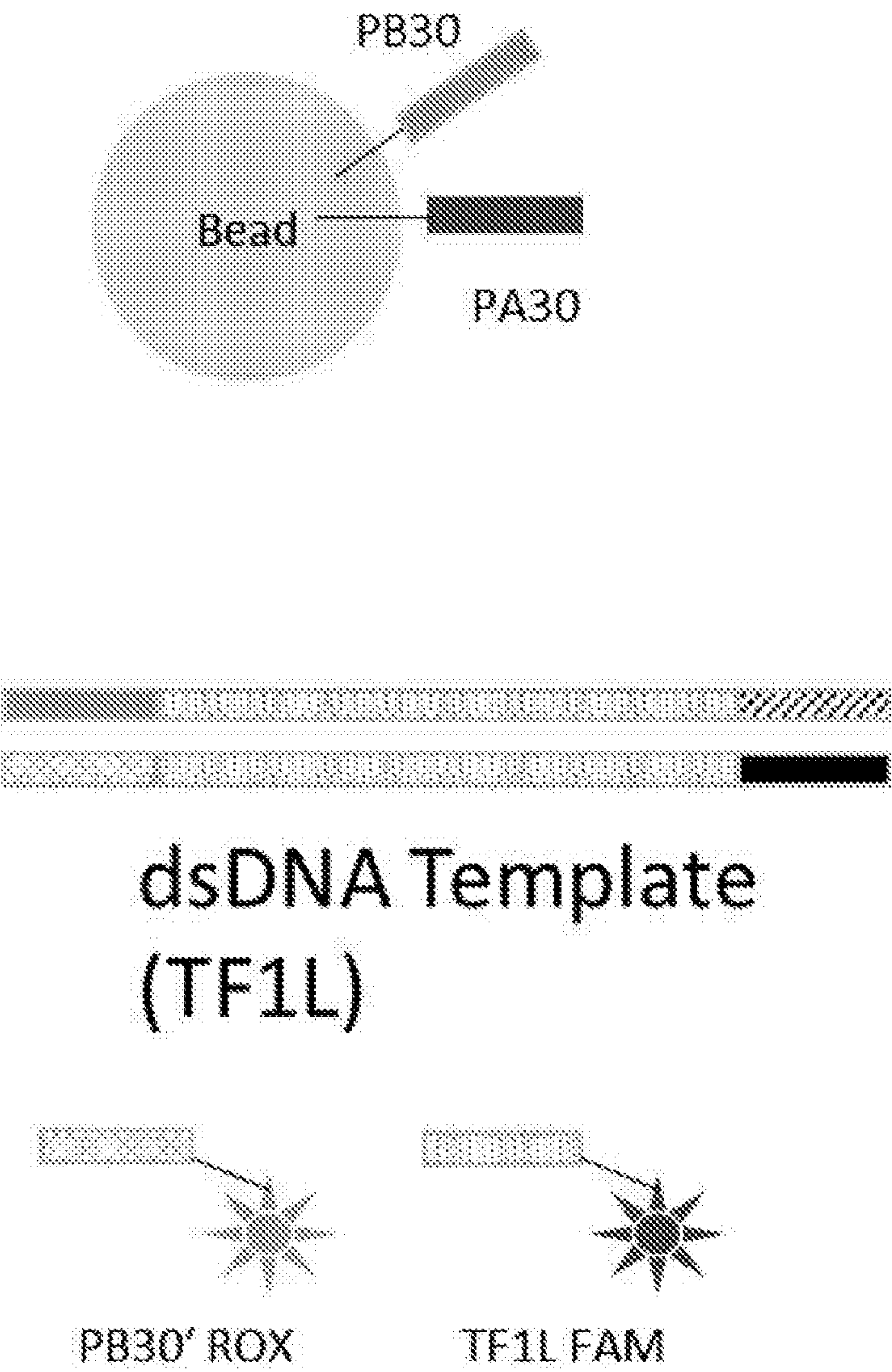


Figure 18B

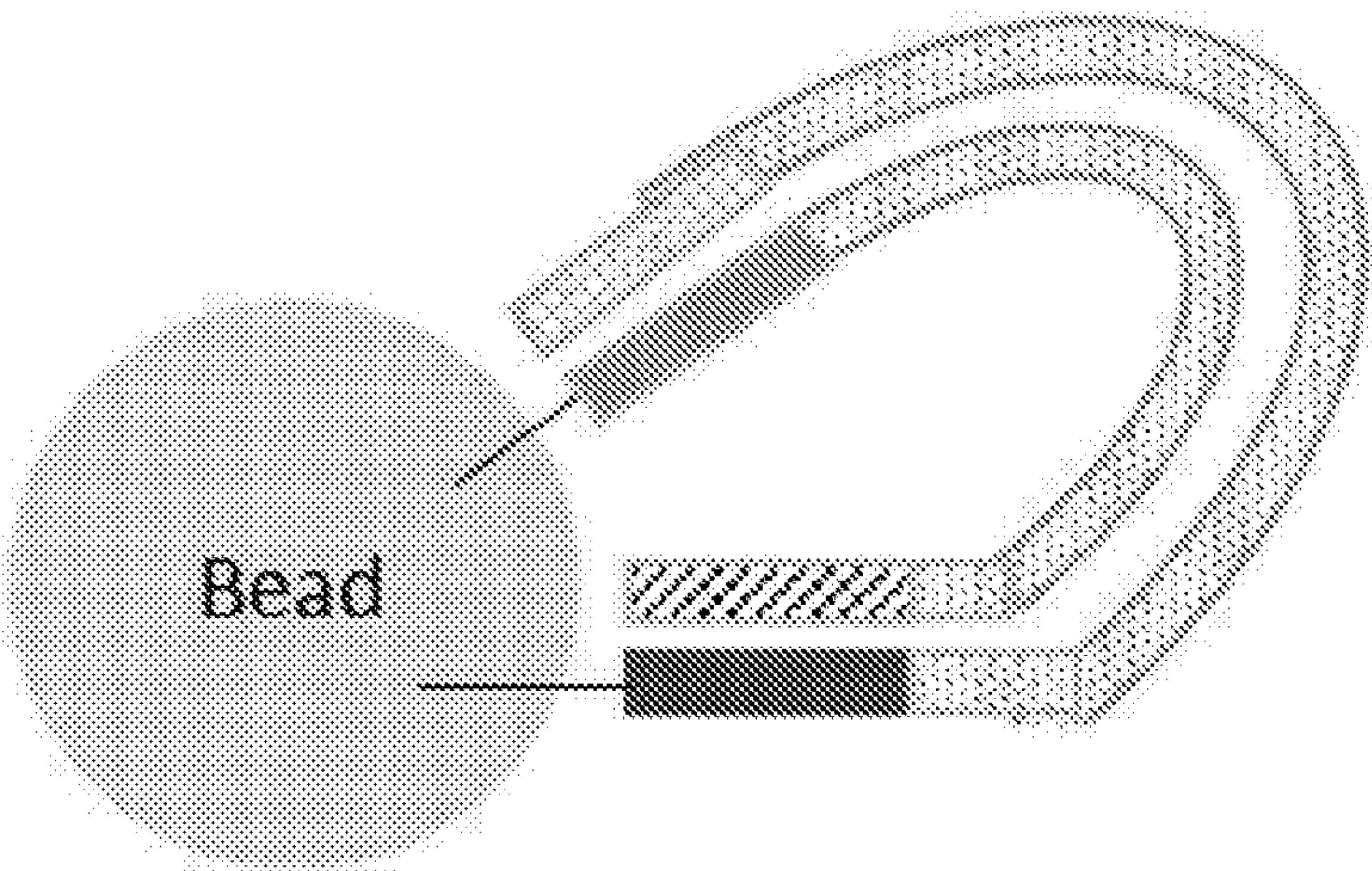


Figure 18C

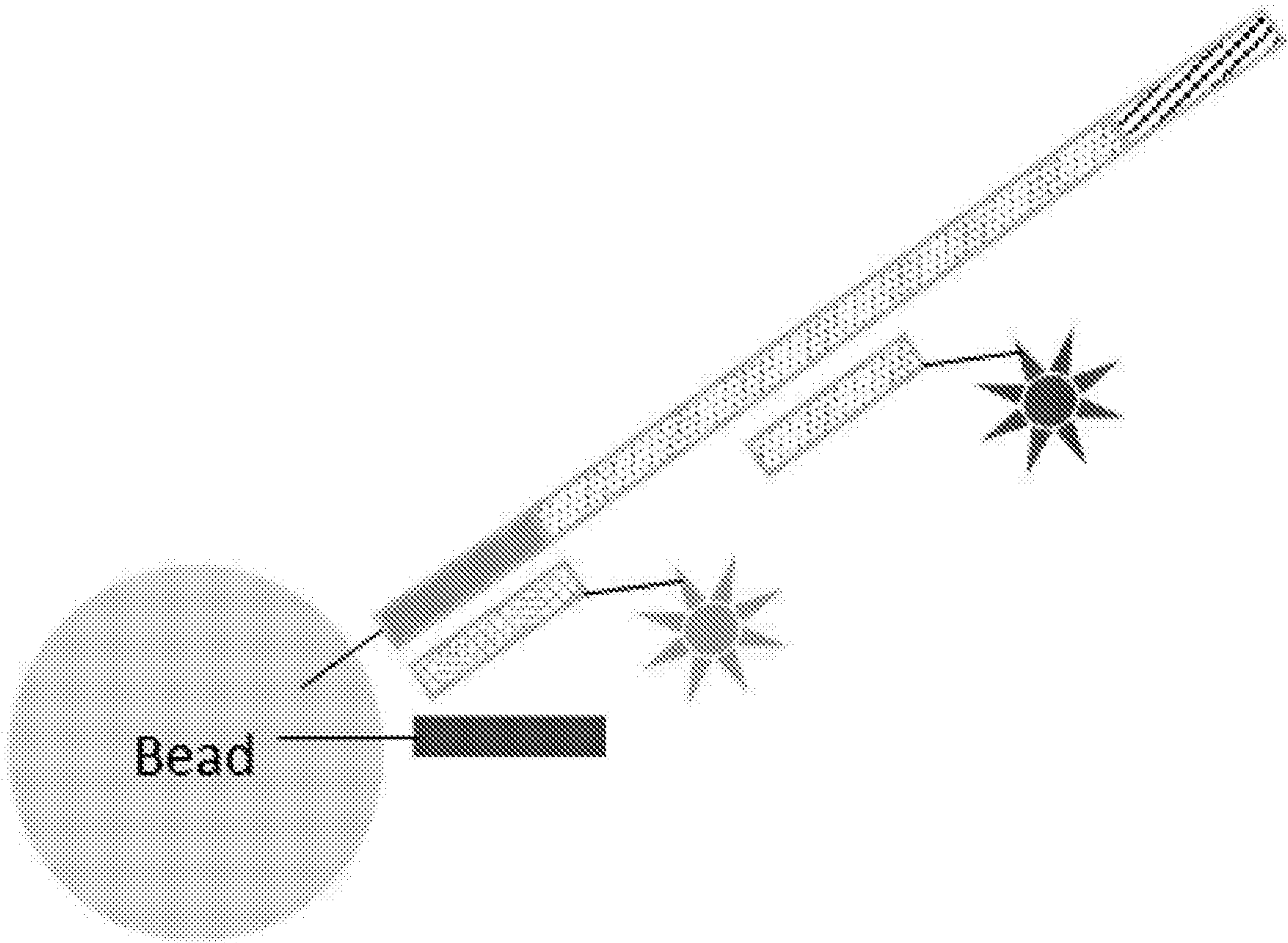


Figure 18D

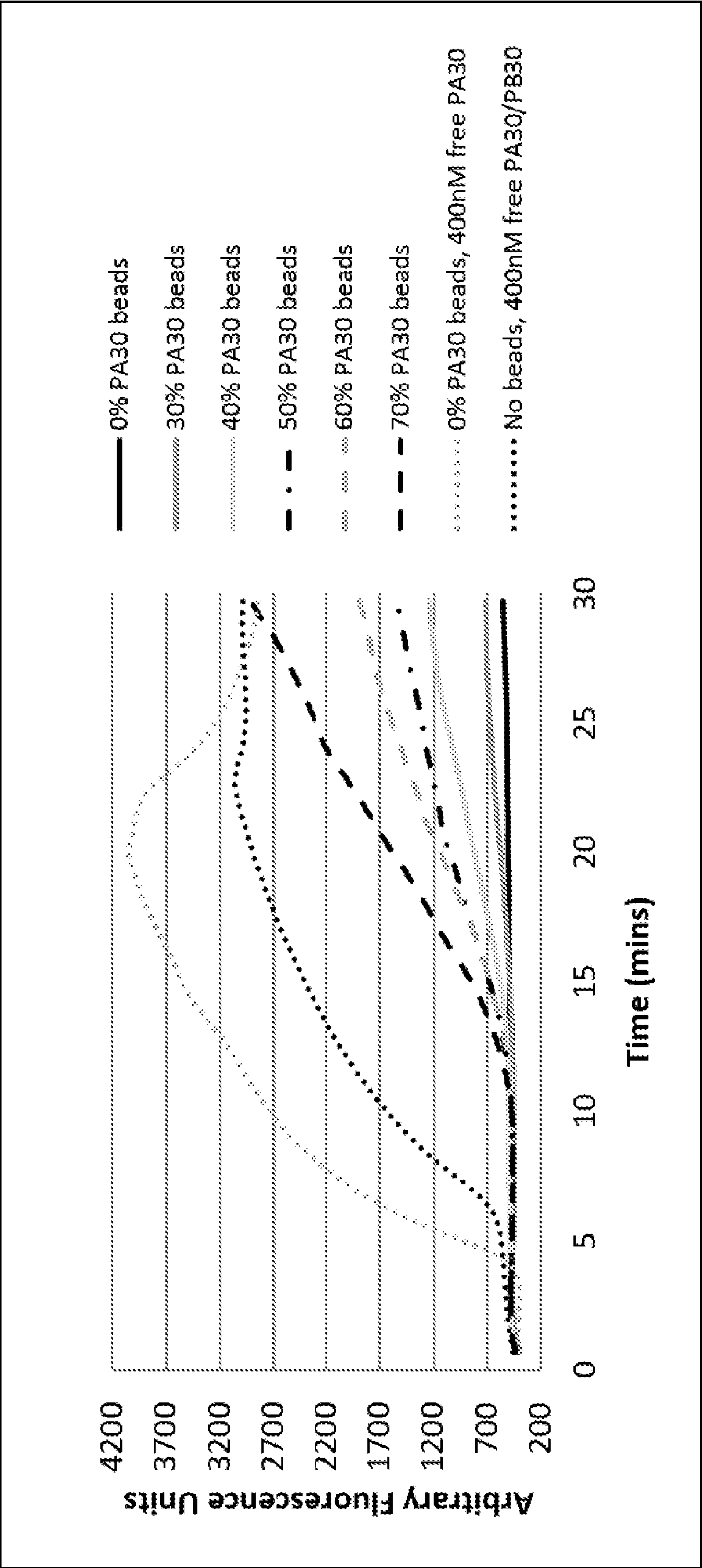


Figure 18E

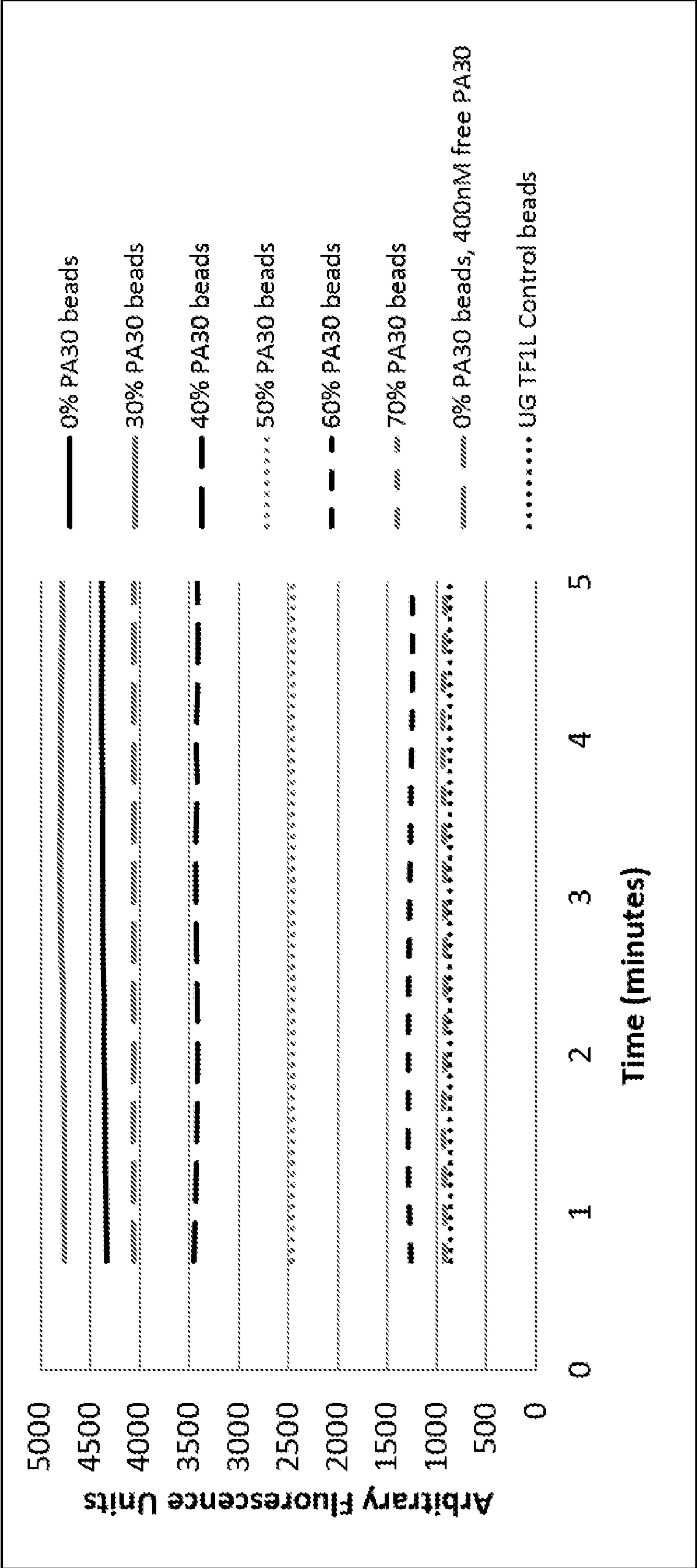


Figure 18F

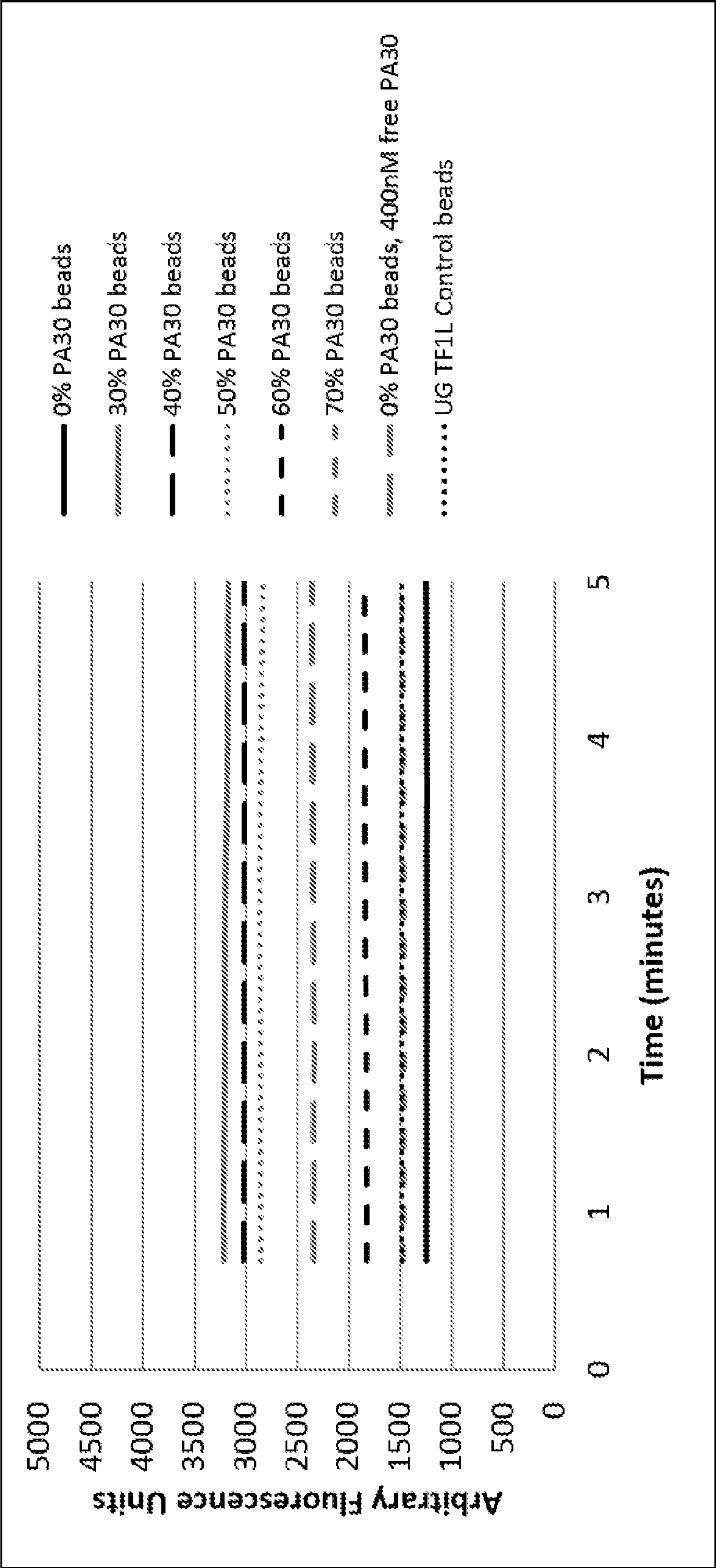


Figure 19A

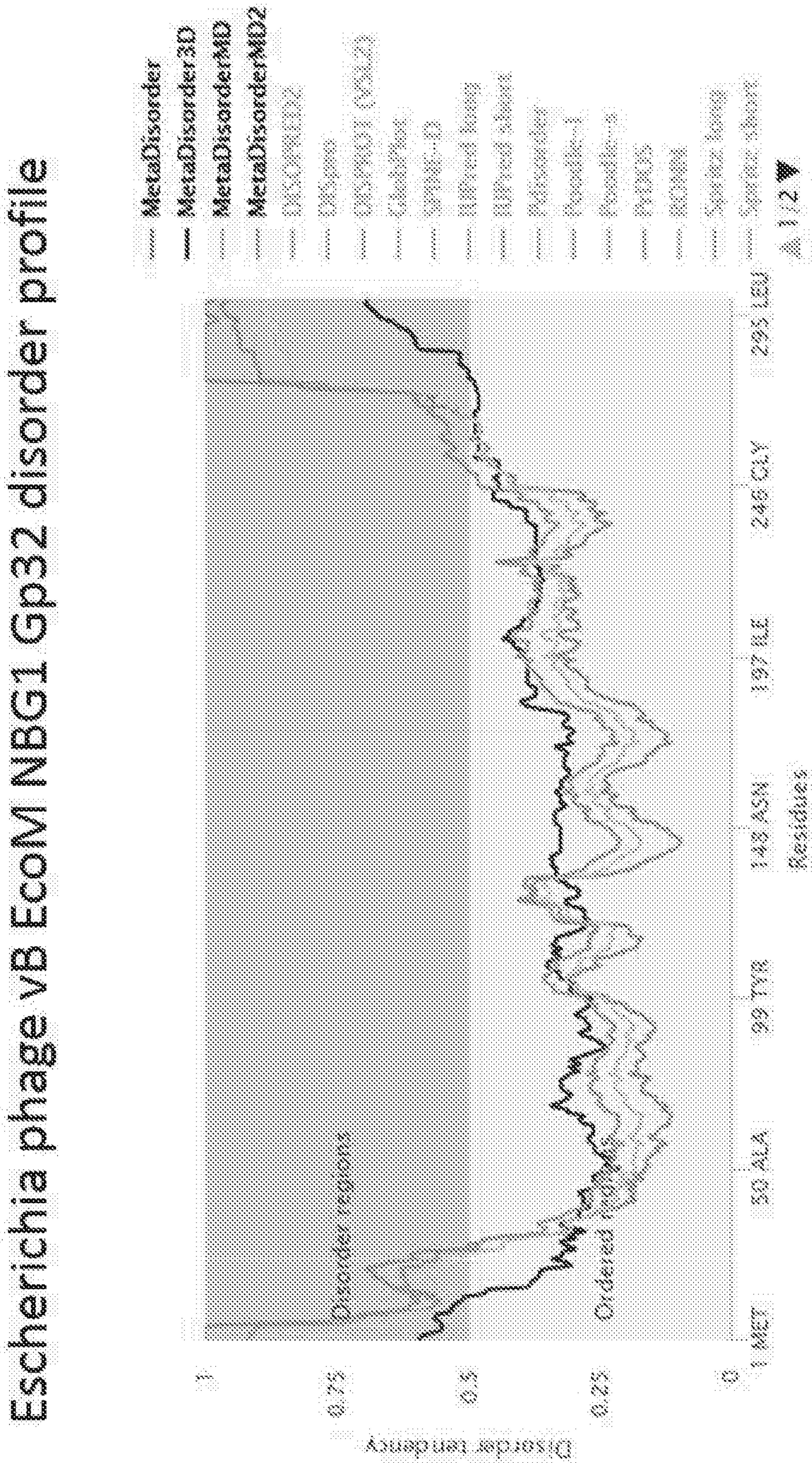


Figure 19B

T4 UvsY disorder profile

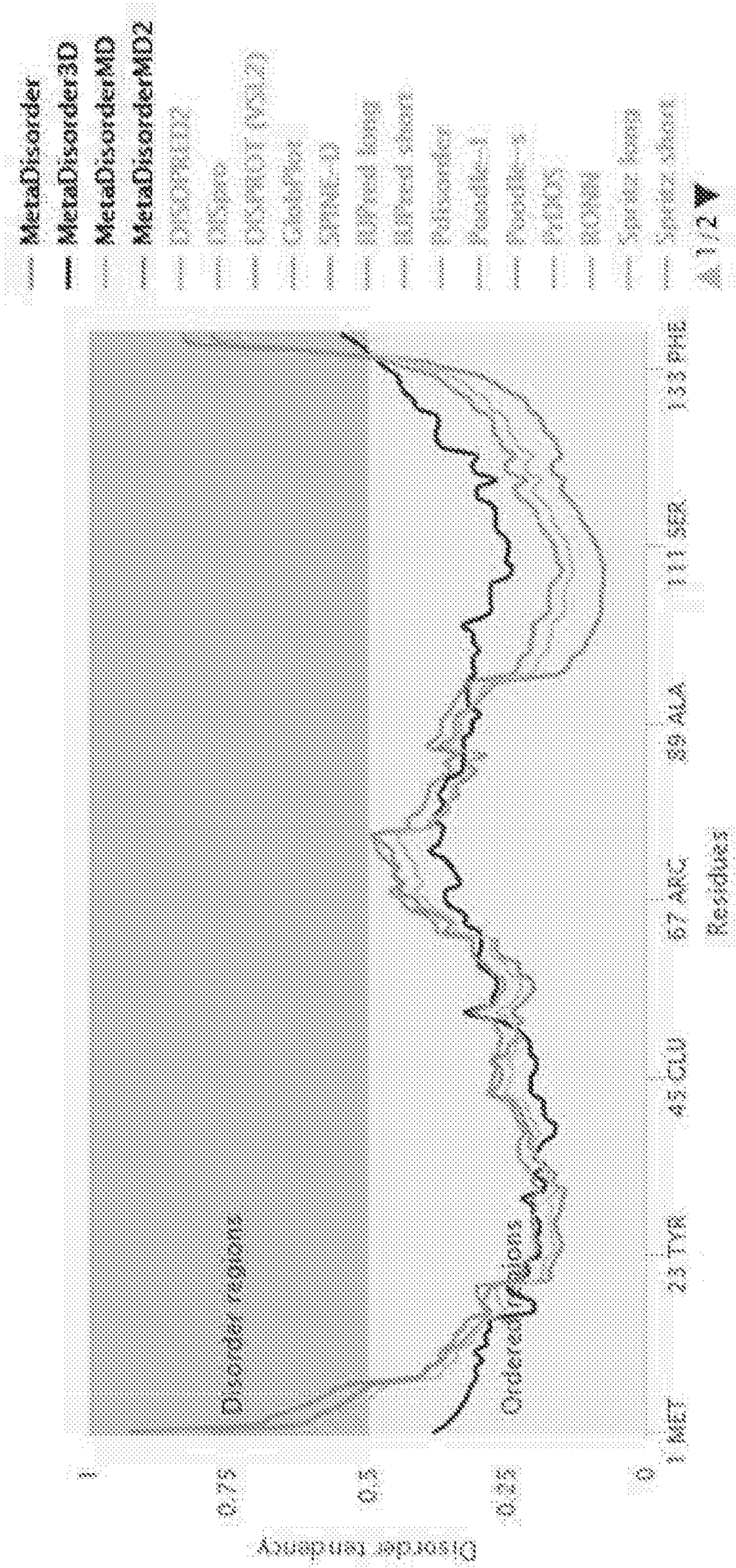


Figure 19C

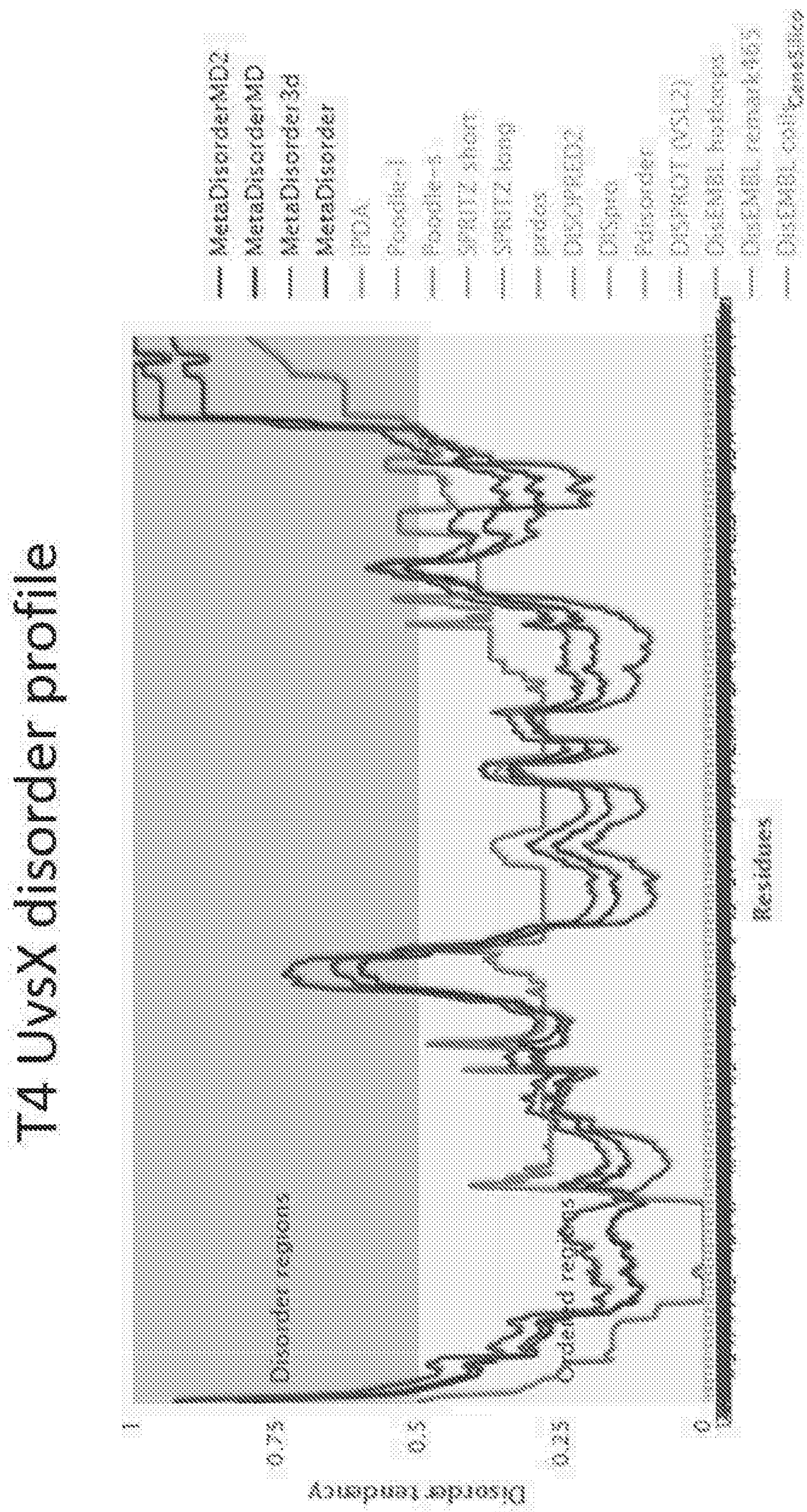


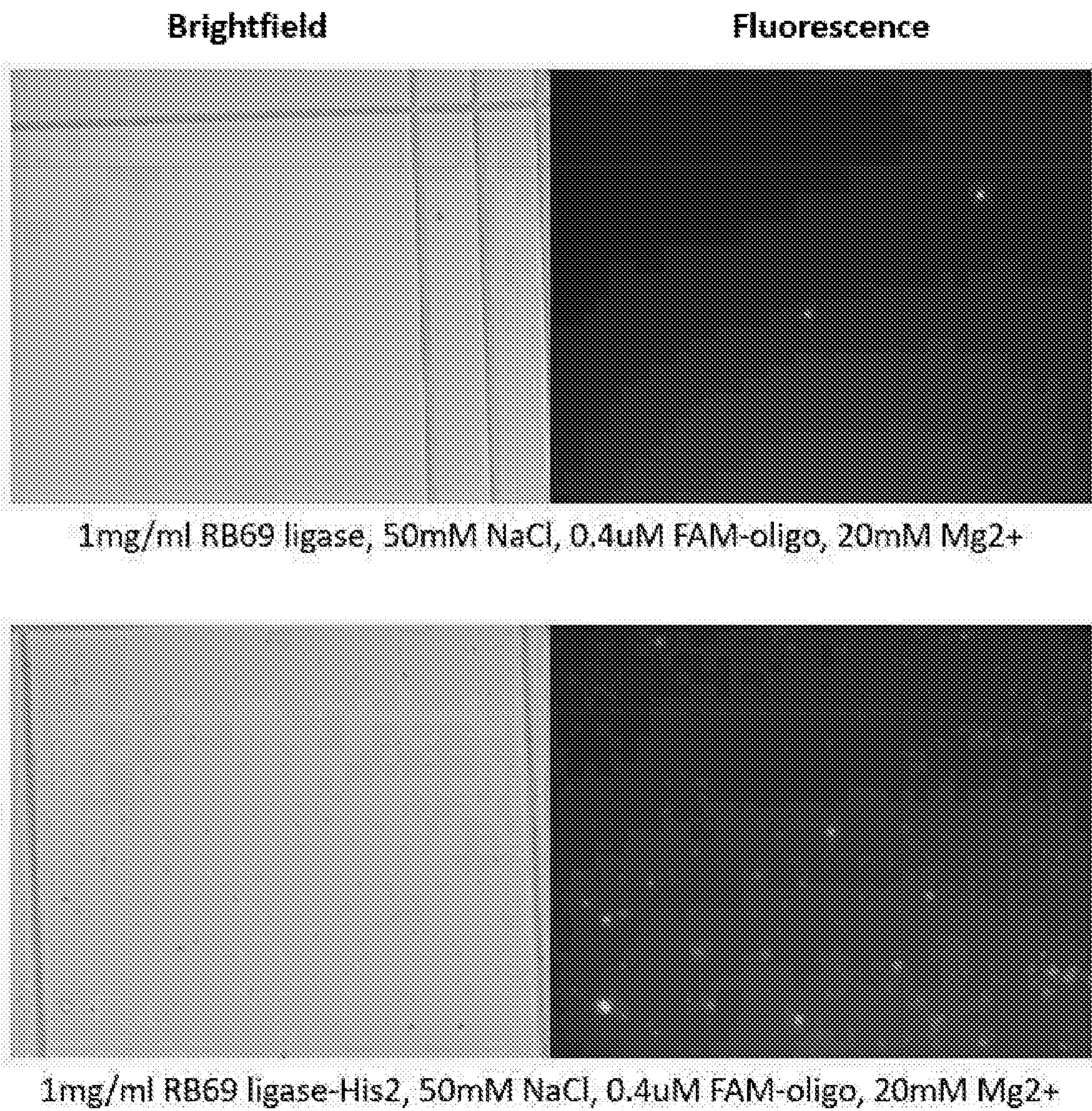
Figure 20A

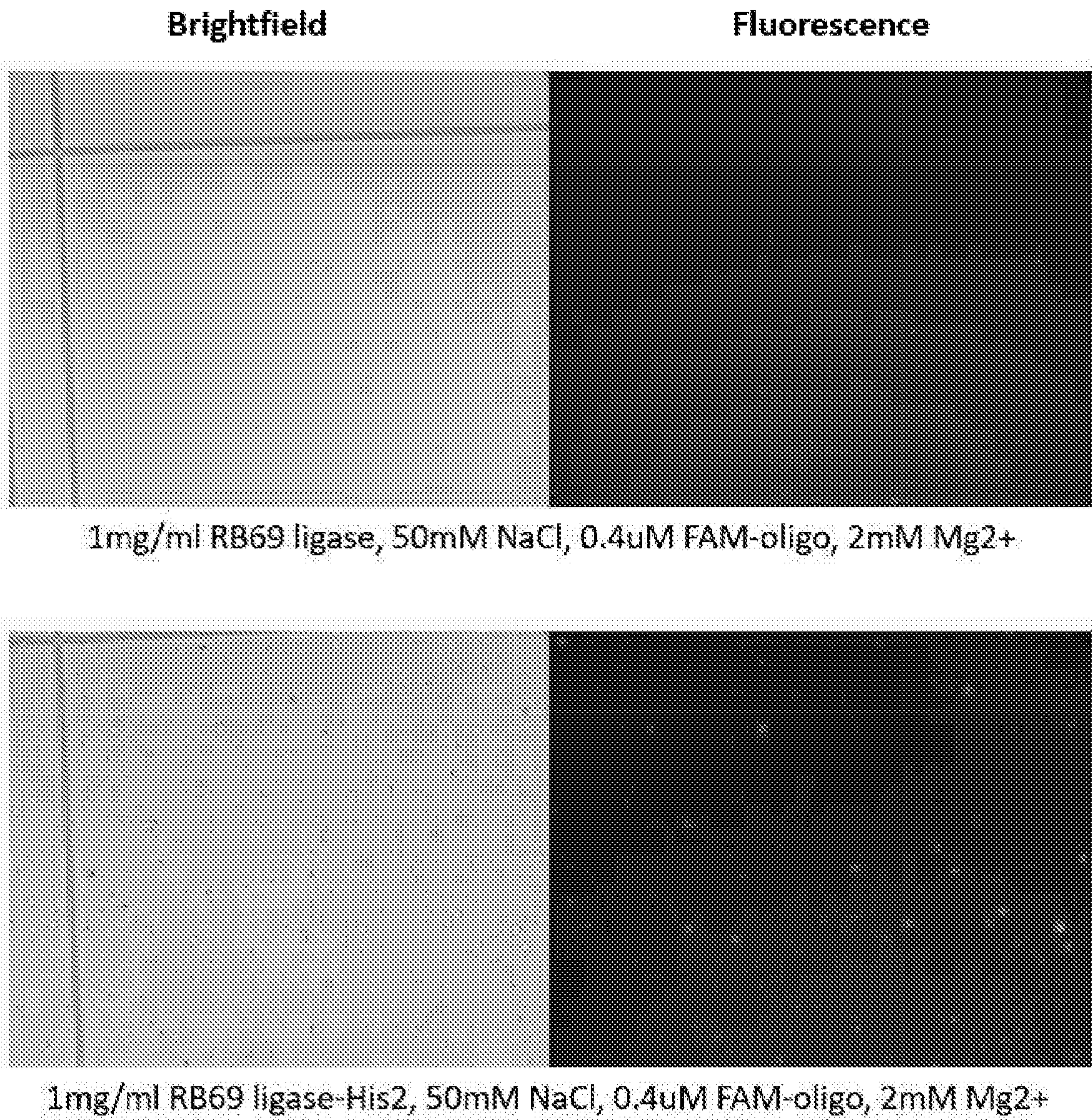
Figure 20B

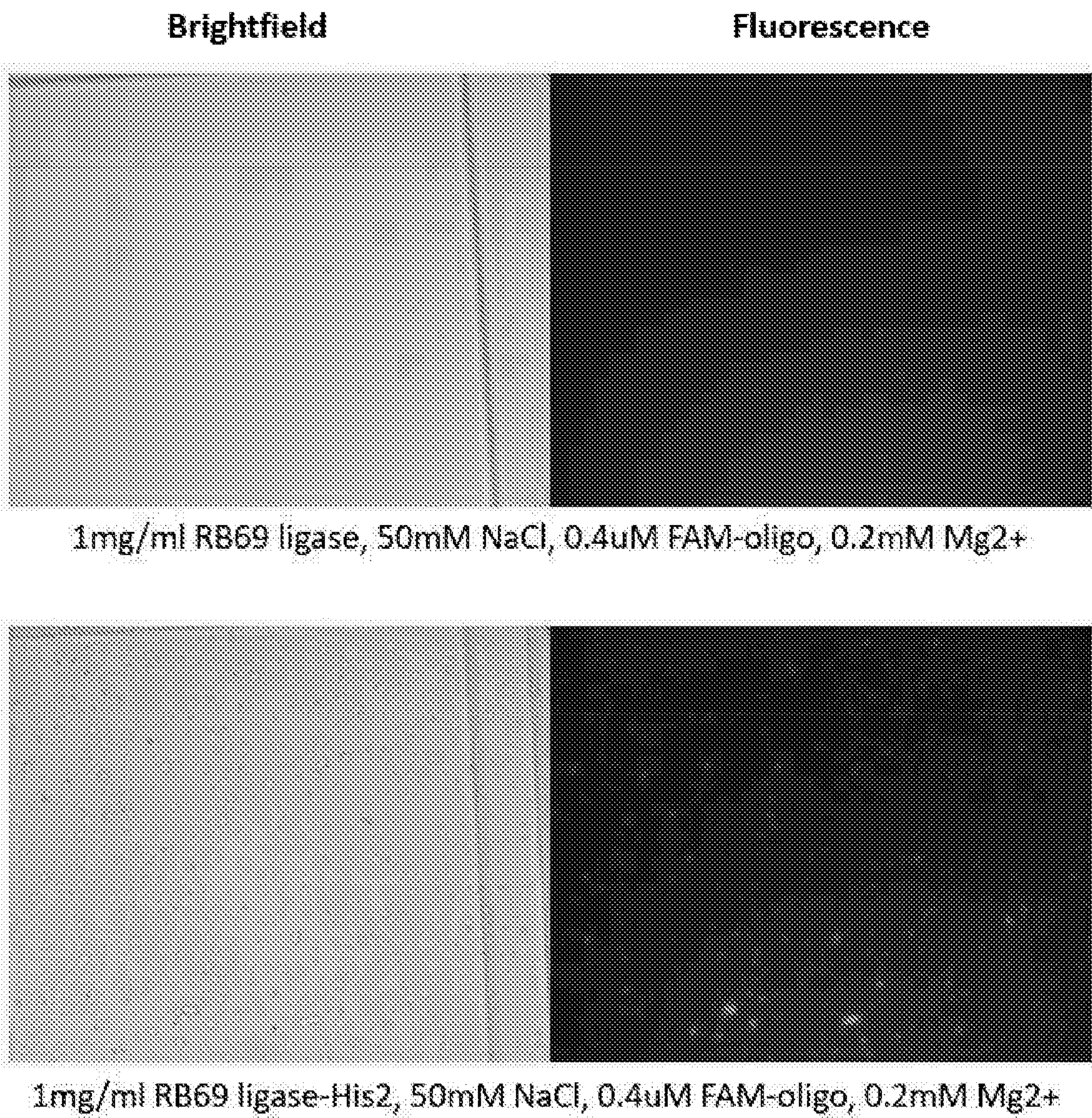
Figure 20C

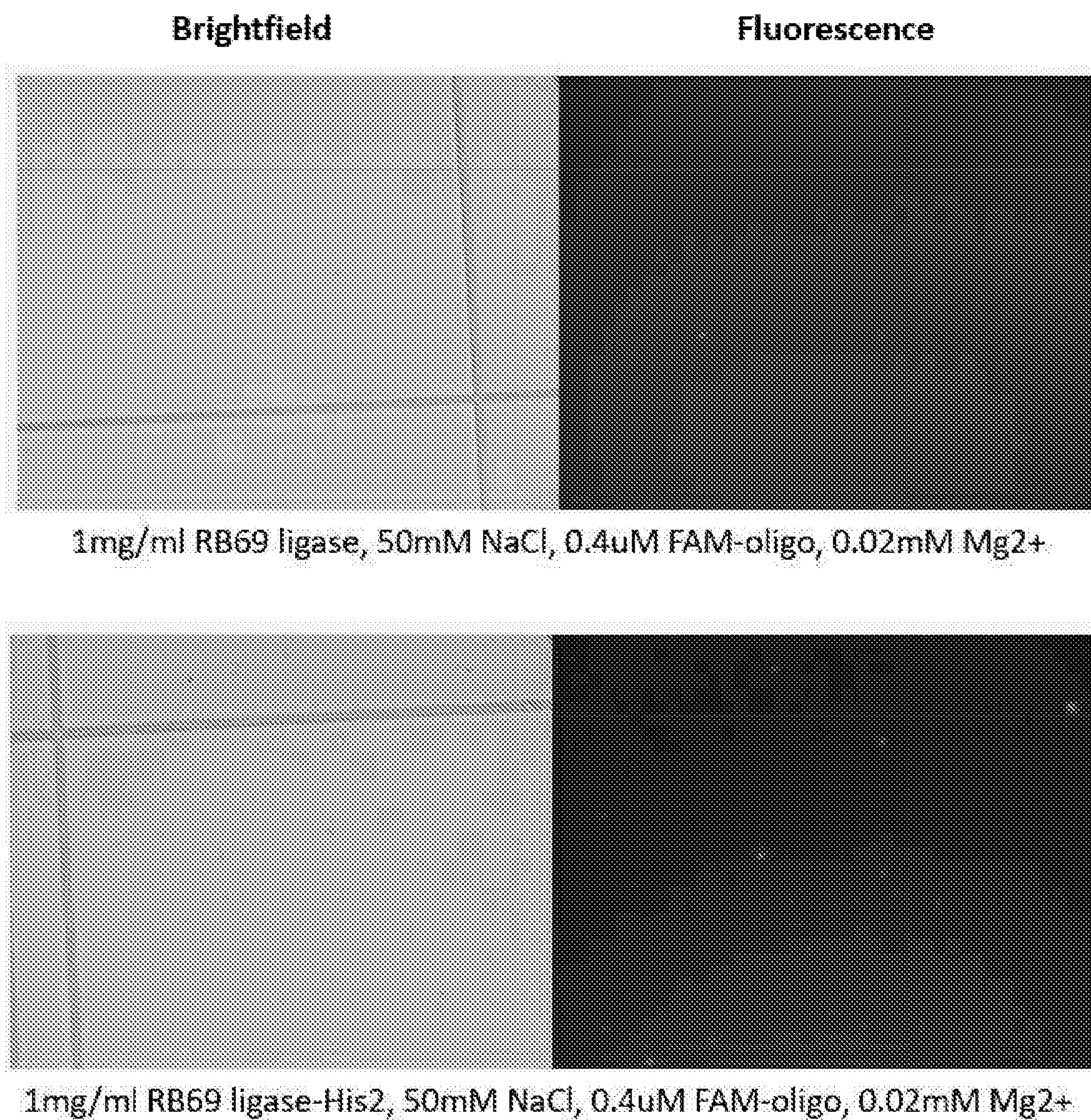
Figure 20D

Figure 21

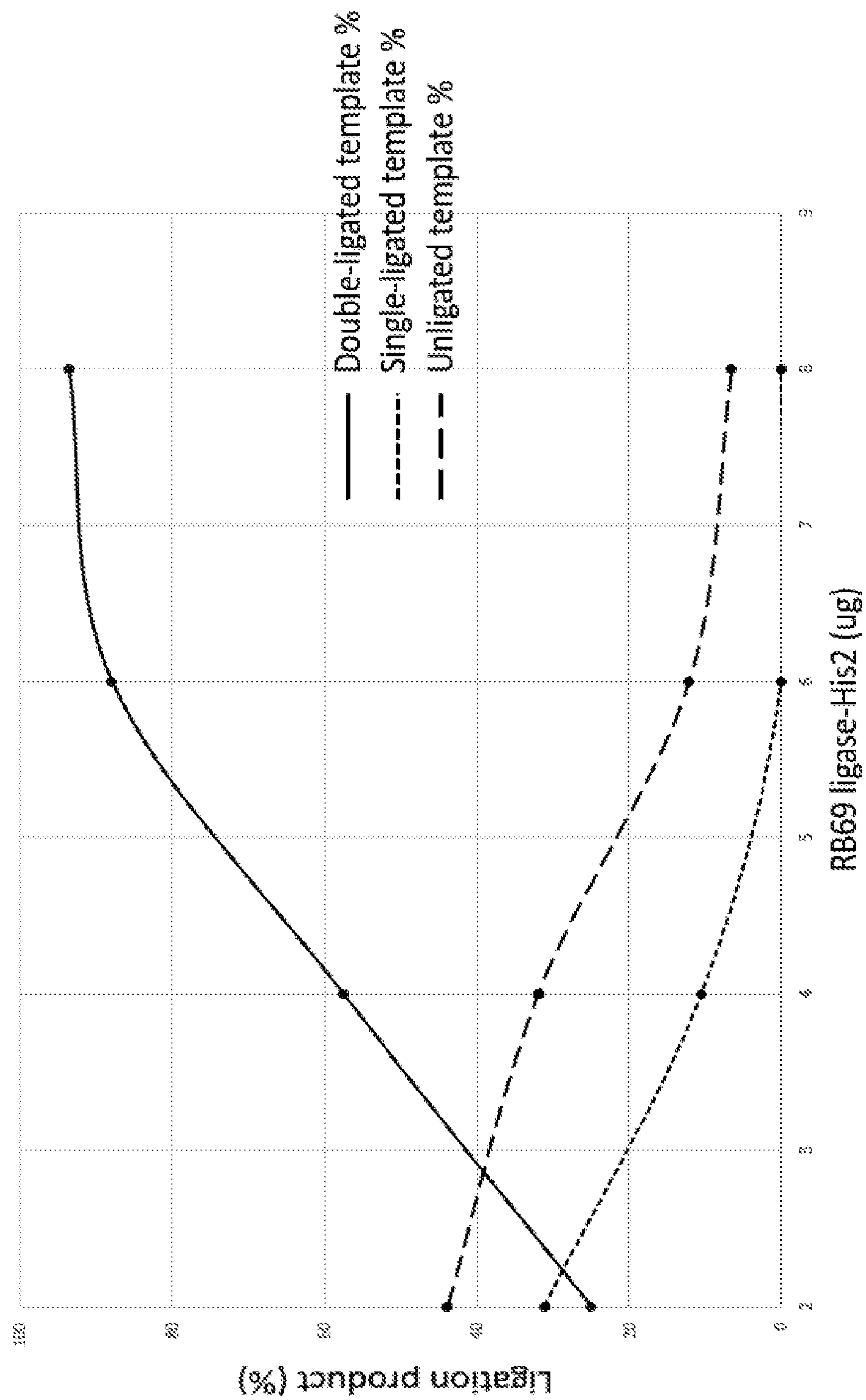


Figure 22

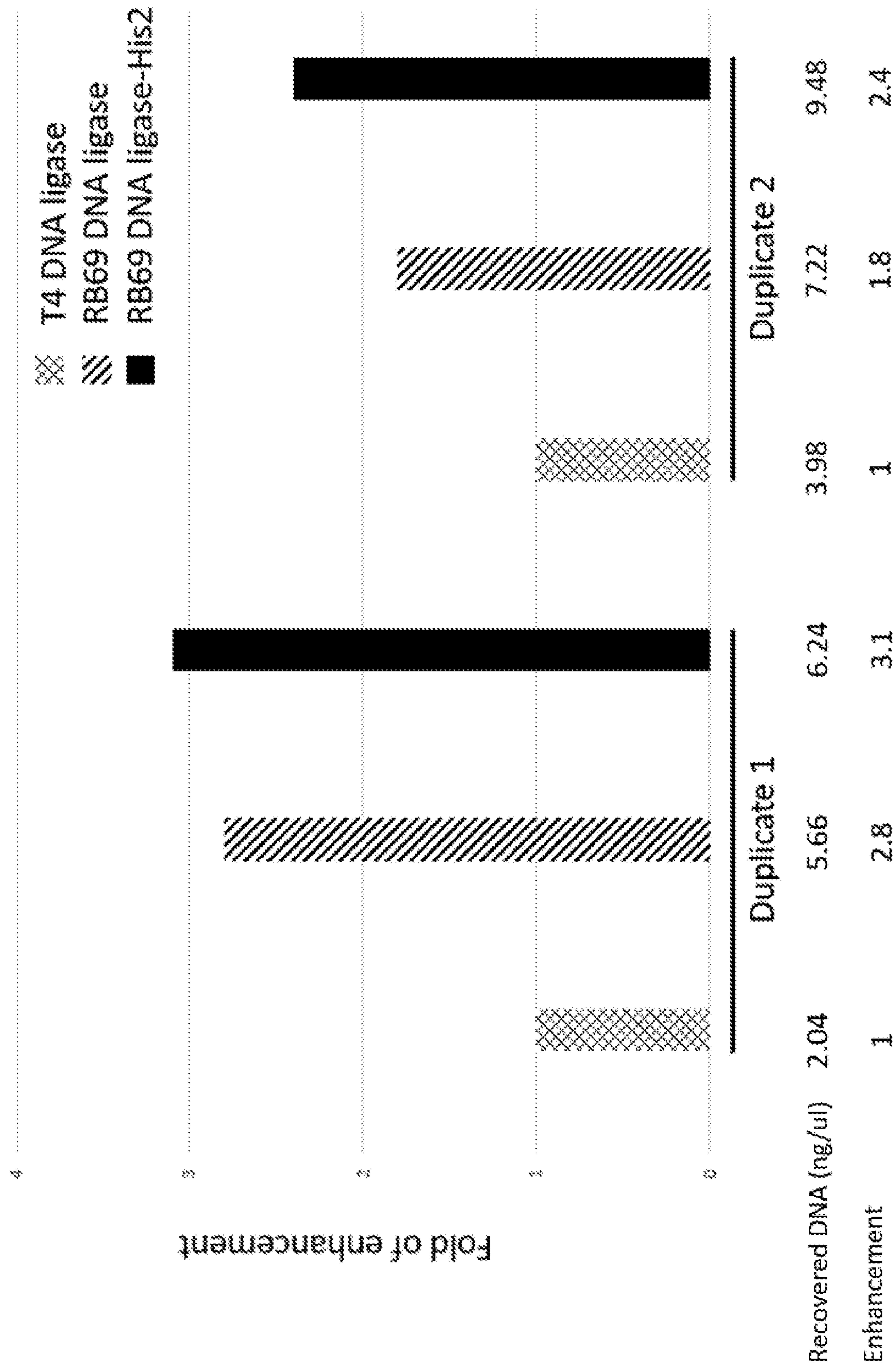


Figure 23

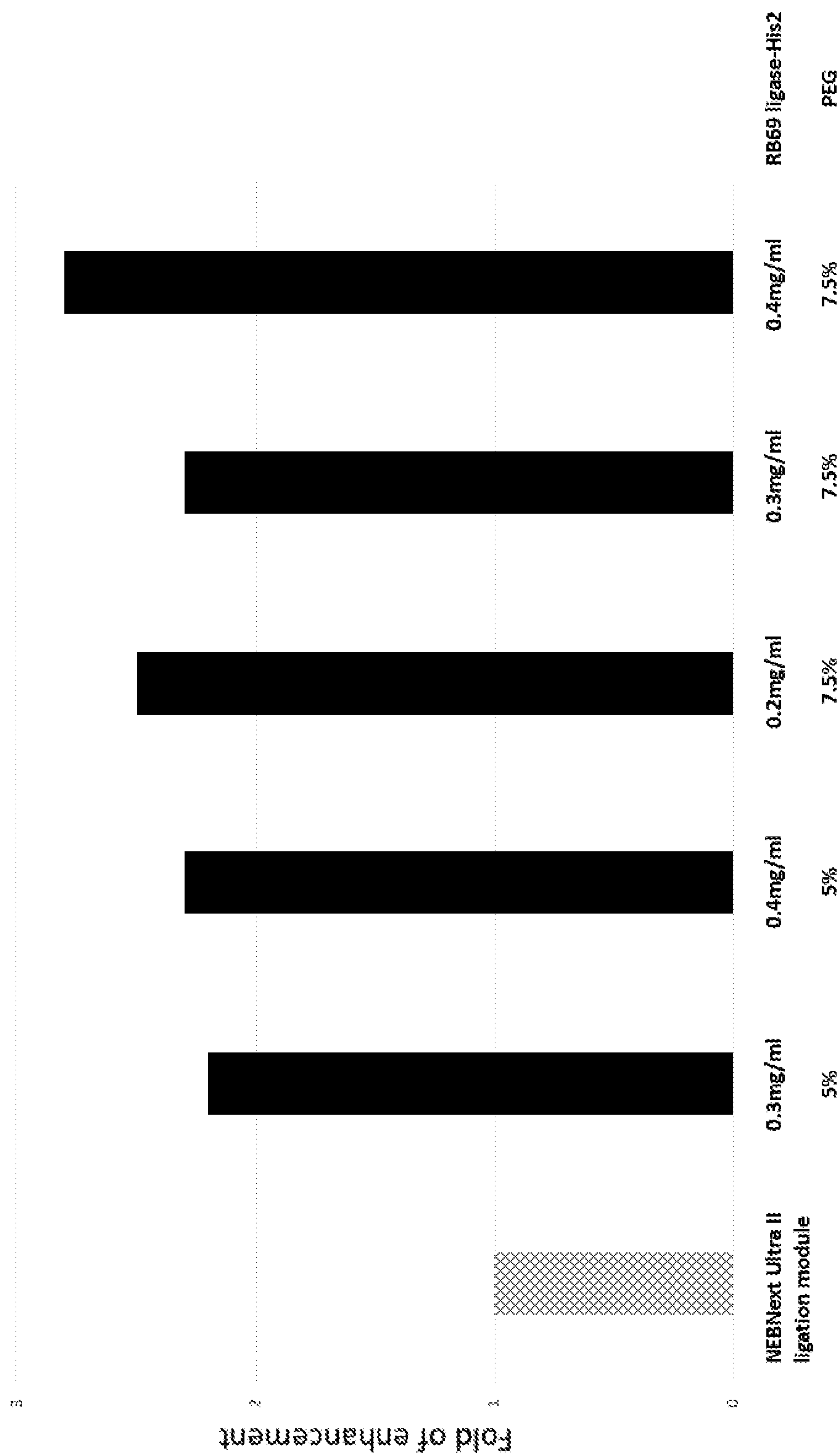


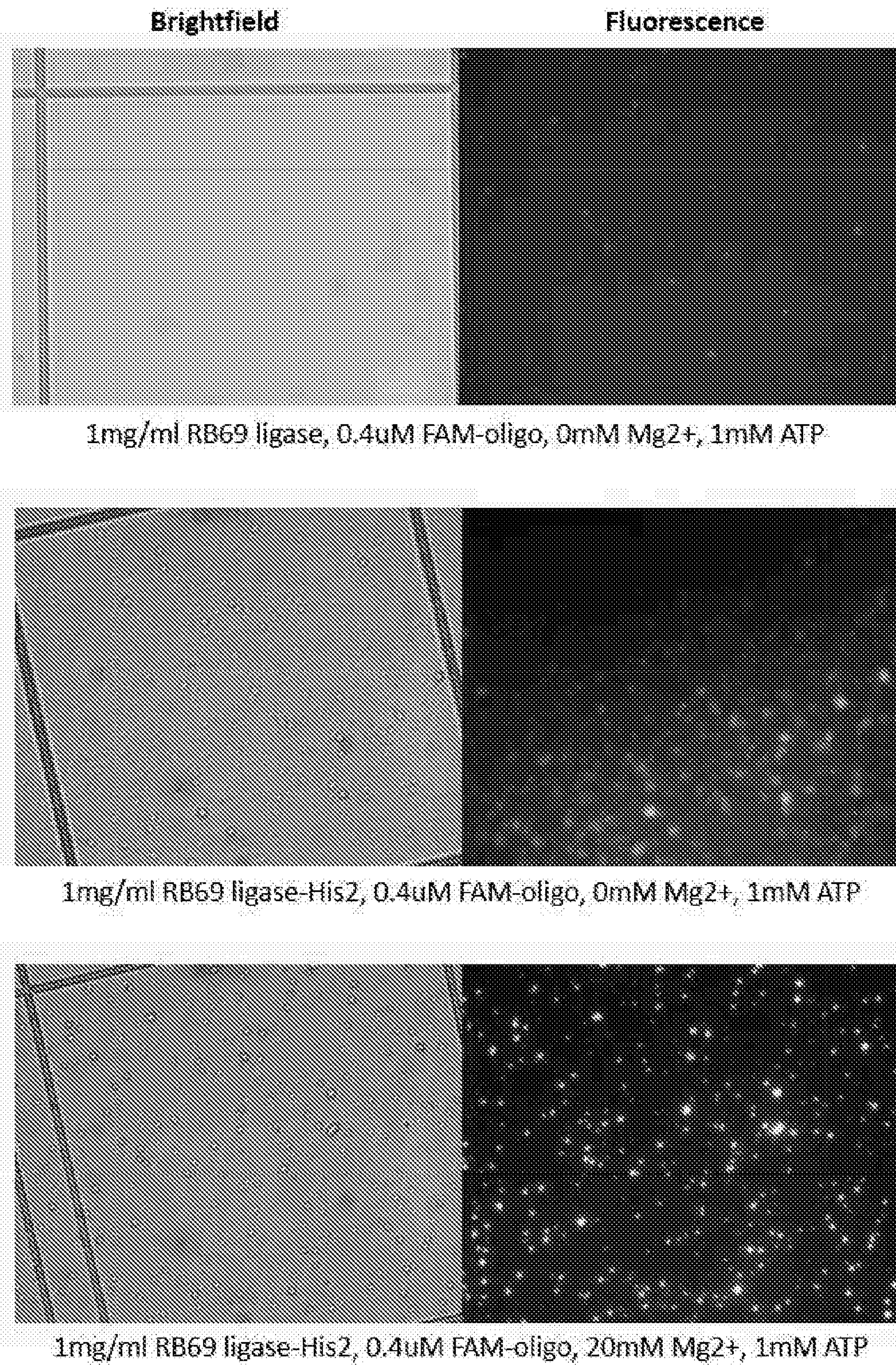
Figure 24

Figure 25A

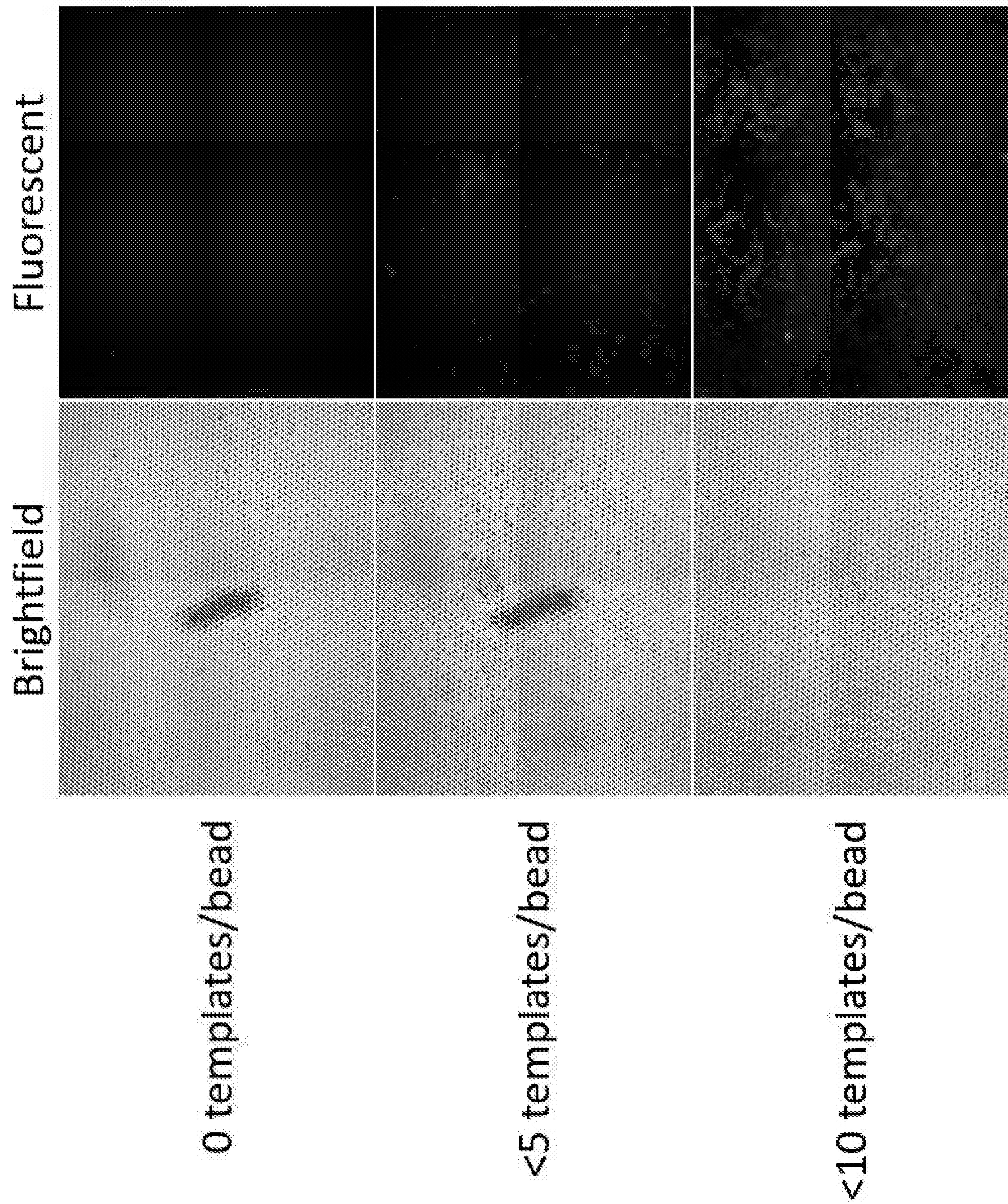


Figure 25B

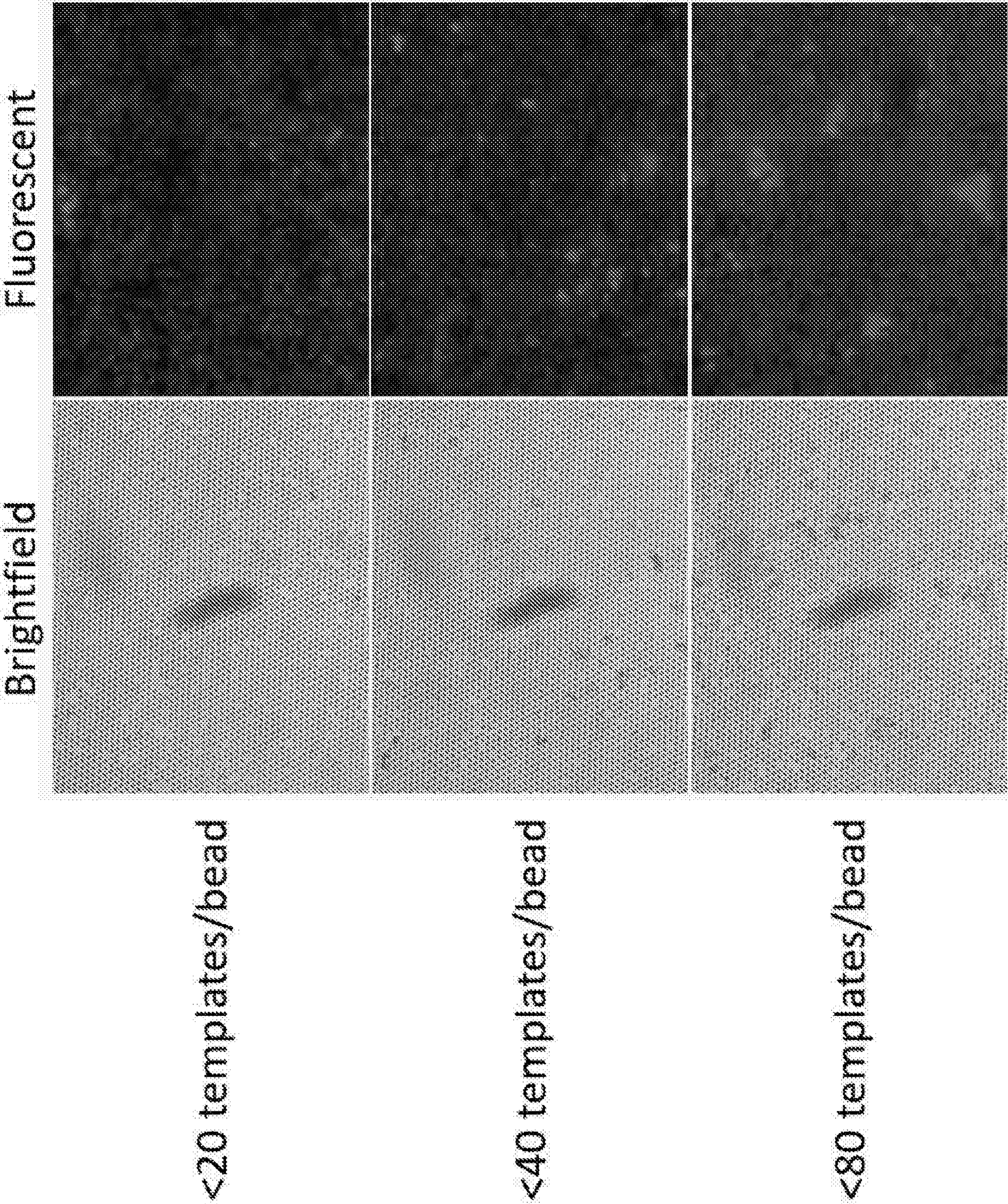


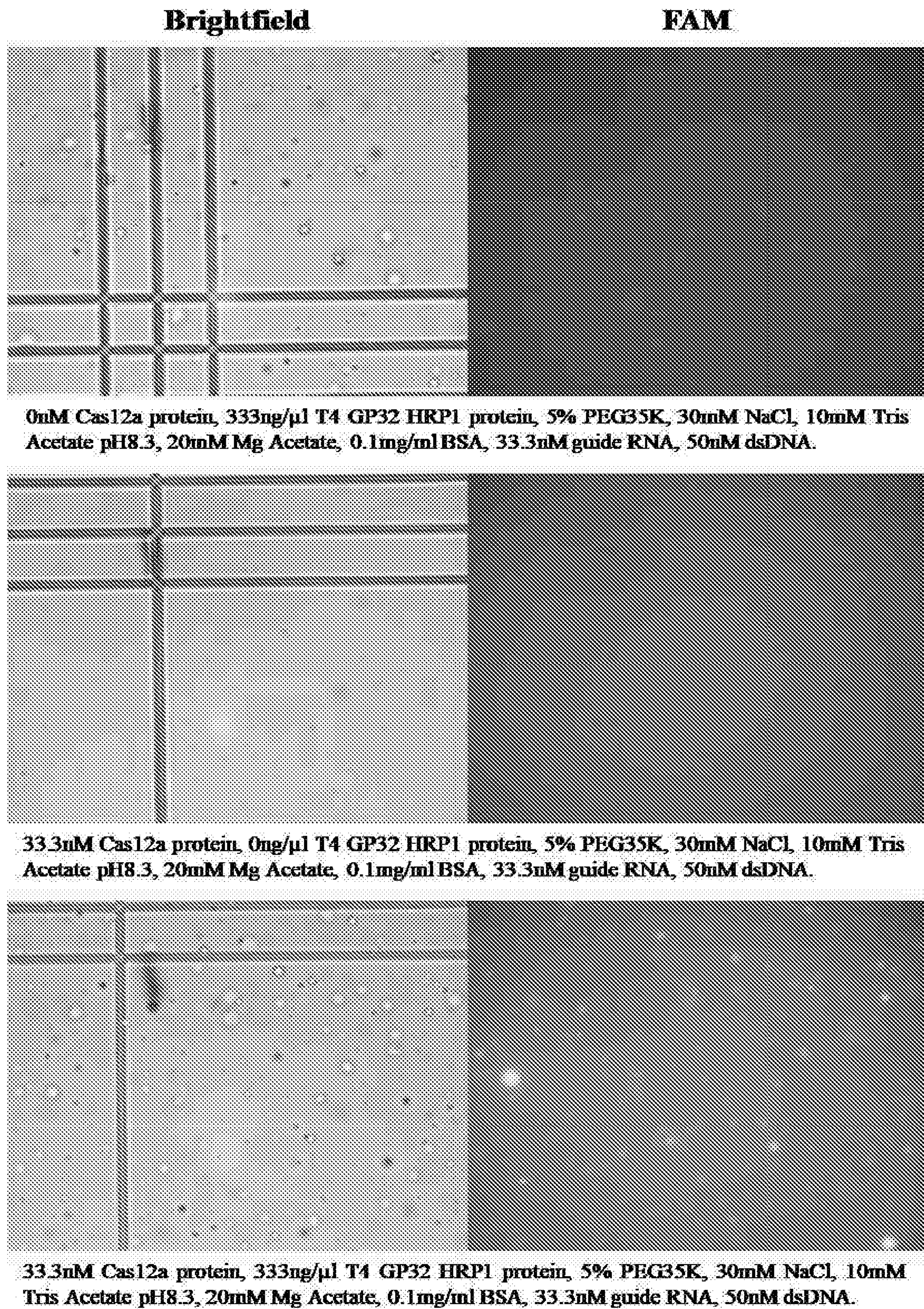
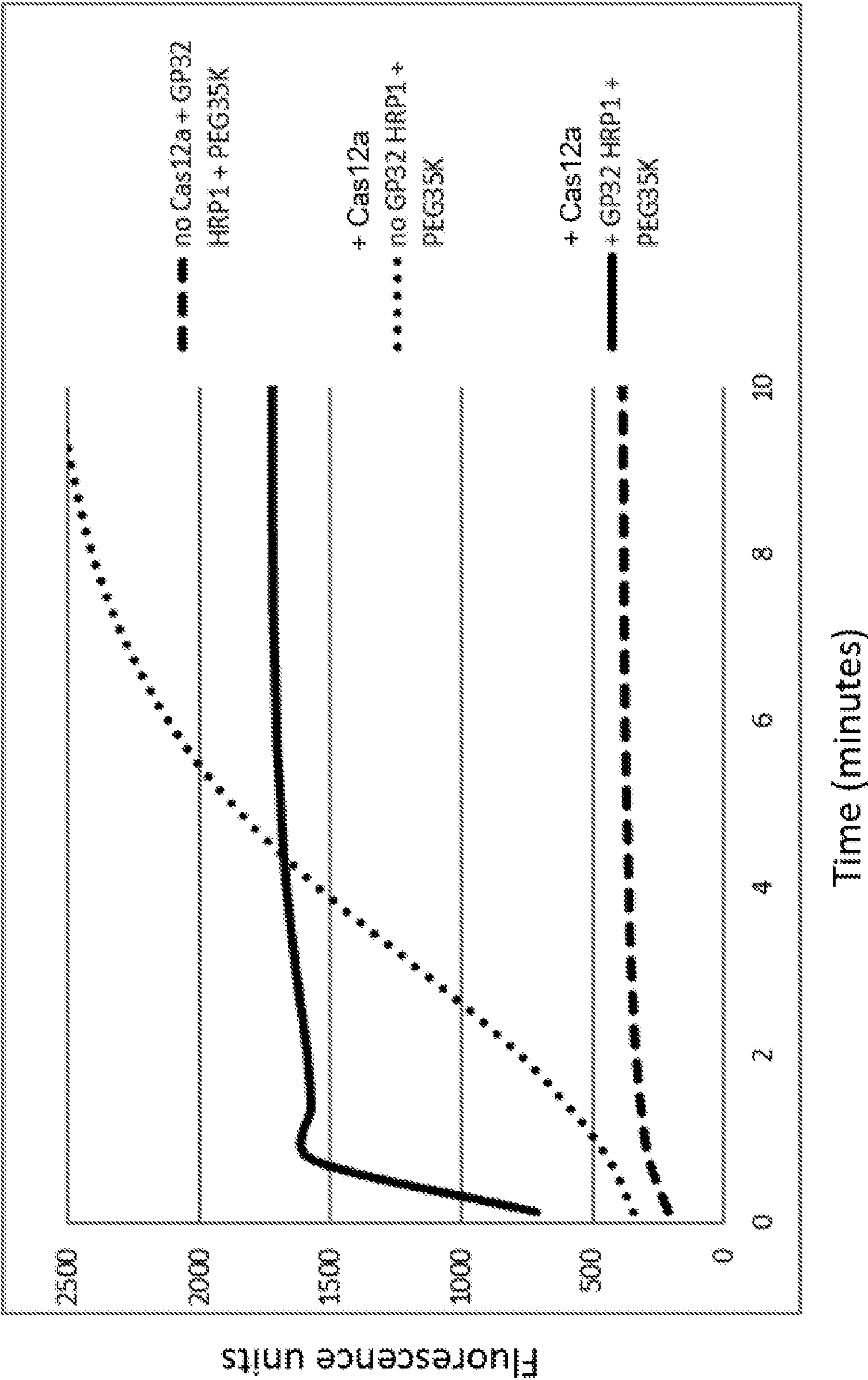
Figure 26A

Figure 26B



INTERNATIONAL SEARCH REPORT

International application No

PCT/GB2020/052866

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/6848 G01N33/50
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, CHEM ABS Data, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HUI YIN TAN ET AL: "The intrinsically disordered linker of E . coli SSB is critical for the release from single-stranded DNA : SSB IDL is Required for DNA Release", PROTEIN SCIENCE, vol. 26, no. 4, 1 April 2017 (2017-04-01), pages 700-717, XP055770349, US	1-5
Y	ISSN: 0961-8368, DOI: 10.1002/pro.3115 abstract; figure 5B page 705, column 1, paragraph first full page 706, column 1, last paragraph - column 2, paragraph first full ----- -/-	1-40,67



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

2 February 2021

Date of mailing of the international search report

07/04/2021

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Celler, Jakub

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2020/052866

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/112631 A1 (PIEPENBURG OLAF [GB] ET AL) 26 May 2005 (2005-05-26)	1-40,67
Y	abstract; claims 1,22,111,120,124,125,141,142 paragraphs [0030], [0217] -----	1-40,67
A	CHONG P ANDREW ET AL: "Liquid-liquid phase separation in cellular signaling systems", CURRENT OPINION IN STRUCTURAL BIOLOGY, ELSEVIER LTD, GB, vol. 41, 20 August 2016 (2016-08-20), pages 180-186, XP029835291, ISSN: 0959-440X, DOI: 10.1016/J.SBI.2016.08.001 the whole document -----	1-40,67

INTERNATIONAL SEARCH REPORT

International application No.

PCT/GB2020/052866

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13~~ter~~.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13~~ter~~.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13~~ter~~.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2020/052866

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-40, 67

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-40, 67

Subject matter of the above claims as characterised by the constellation of technical features of claim 1.

2. claims: 41-66, 68

Subject matter of the claims as characterised by the constellation of technical features of claim 41.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2020/052866

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005112631 A1	26-05-2005	US 2005112631 A1	26-05-2005
		US 2008293045 A1	27-11-2008
		US 2009017462 A1	15-01-2009
		US 2011059506 A1	10-03-2011
		US 2012082990 A1	05-04-2012
		US 2014234846 A1	21-08-2014
		US 2016032374 A1	04-02-2016
		US 2017342473 A1	30-11-2017
