



(51) International Patent Classification:

C07H 19/10 (2006.01) C12Q 1/68 (2018.01)
C07H 19/20 (2006.01)

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))

(21) International Application Number:

PCT/US2024/032844

(22) International Filing Date:

06 June 2024 (06.06.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/471,425 06 June 2023 (06.06.2023) US

(71) Applicant: ILLUMINA, INC. [US/US]; 5200 Illumina Way, San Diego, CA 92122 (US).

(72) Inventors: PANTOJA, Rigoberto; 5200 Illumina Way, San Diego, CA 92122 (US). YANG, Xiangyuan; 5200 Illumina Way, San Diego, 92122 (US). VERMAAS, Eric; 5200 Illumina Way, San Diego, CA 92122 (US). TEO, Yin Nah; 5200 Illumina Way, San Diego, CA 92122 (US). DI, Liu; 5200 Illumina Way, San Diego, California 92122 (US).

(74) Agent: BURNETTE, Jennifer; MARSHALL, GERSTEIN & BORUN LLP, 233 S. Wacker Drive, 6300 Willis Tower, Chicago, IL 60606-6357 (US).

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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).

(54) Title: MULTIVALENT BINDING UNIT MODIFIED DIDEOXYNUCLEOTIDE TRIPHOSPHATES AND METHODS OF MAKING AND USING THE SAME

(57) Abstract: Compounds of the disclosure include modified dideoxynucleotide triphosphates (ddNPT) including multiple binding unit containing arms to provide multivalent binding unit molecules.

WO 2024/254329 A1

MULTIVALENT BINDING UNIT MODIFIED DIDEOXYNUCLEOTIDE TRIPHOSPHATES AND METHODS OF MAKING AND USING THE SAME

CROSS-REFERENCE TO RELATED APPLICATION

[0001] The benefit of priority to U.S. Provisional Patent Application No. 63/471,425 filed June 6, 2023, is hereby claimed and the disclosure is incorporated herein by reference in its entirety.

FIELD

[0002] The disclosure relates to multivalent binding unit modified dideoxynucleotide triphosphates (ddNTP) and methods of making and using the same.

BACKGROUND

[0003] Single nucleotide polymorphisms (SNPs) are a common form of polymorphisms present in the human genome. In genotyping assays, allele specificity can be conferred by a single base extension that incorporates one or four labeled nucleotides. When excited by a laser, the nucleotide label emits a signal. The intensity of that signal conveys information about allelic ratio at that locus.

SUMMARY

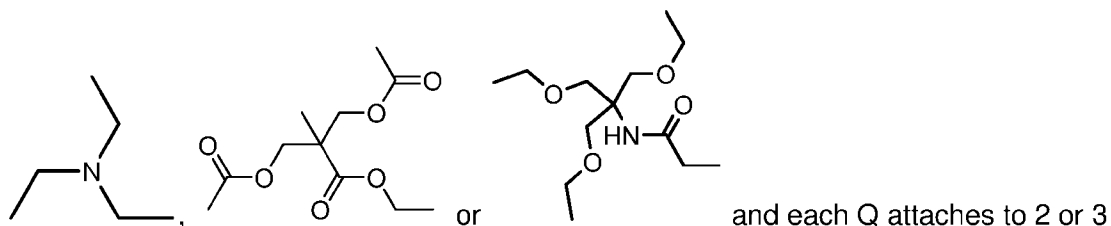
[0004] Signal detection deficiencies have been observed in red channel (ddATP-DNP and ddUTP-DNP) detection. Lower signal in the red channel after 3 rounds of staining has been observed as compared to green channel detection. Without intending to be bound by theory, it is believed that some of the difference in signal intensity between red and green channels may be due to higher non-specific anti-DNP antibody-dye conjugate binding. This was observed through higher red signal in homozygous green SNPs.

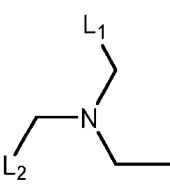
[0005] The compounds and methods of the disclosure utilize multivalent binding unit modified ddNTP to enhance binding unit binding partner binding affinity and thereby improve signal detection. This can beneficially result in decreased nonspecific binding that is mediated by both binding unit binding partner surface and dyes covalently attached to the binding unit binding partner for fluorescence detection. This can also beneficially reduce the amount of binding unit binding partner required in current reagent formulations, which in turn can minimize nonspecific binding and improve COGS. Finally, the compounds and method of the disclosure can beneficially decrease the number of staining rounds required for signal amplification. For example, the binding unit molecule can be haptens and the binding unit binding partner can be hapten binding partners.

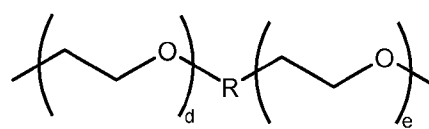
[0006] In accordance with the disclosure, a multivalent binding unit modified ddNTP can be a compound of formula I:

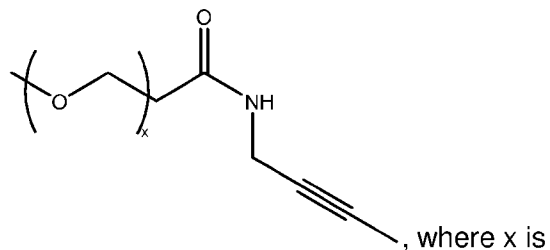


wherein a is at least 2, b is 1 to 3, c is 0 or 1, each ARM comprising a binding unit and a PEO-containing linker, each Q is independently selected from,



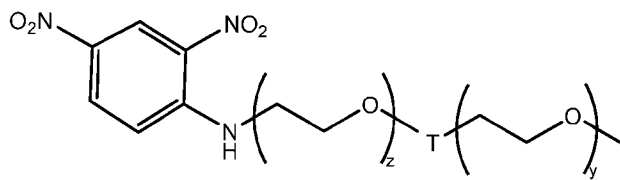
ARMS, and each M is , each L₁ and L₂ are independently selected from

[0007] , wherein for each L₁ and L₂, d is independently selected from 1 to 10, and e is independently selected from 1 to 10, R is independently selected from a functionality resulting from a coupling reaction between one of the following pairs: amine-imidoester, amine-pentafluorophenyl ester, amine-N-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-



cyclooctyne, azide-norbornene, or azide-alkyne, J is 1 to 48.

In accordance with the disclosure, each ARM can include



, wherein for each ARM z is independently selected from 1 to 48 and y is independently selected from 1 to 48 T is independently selected from a functionality resulting from a coupling reaction between one of the following pairs: amine-imidoester, amine-pentafluorophenyl ester, amine-~~N~~-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-cyclooctyne, azide-norbornene, or azide-alkyne.

[0008] In accordance with the disclosure, a kit can include a bead array comprising a plurality of beads, each bead having at least one capture oligonucleotide probe attached to a surface thereof; anti-DNP antibodies, each comprising a first fluorescence label; at least one compound of formula I; and instructions for use such that upon contact, one or more anti-DNP antibodies bind to each ARM of the at least one compound of formula I.

[0009] In accordance with the disclosure, a process for genotyping can include contacting a bead array with a sample DNA, each bead of the bead array comprising one or more capture oligonucleotide probes that each bind to a target complementary sequence in the sample, wherein the target complementary sequence binds to the one or more capture oligonucleotide probes stopping at least one base before a locus of interest thereby providing a bead array having captured target complementary sequence; admixing the bead array having the captured target complementary sequences with a composition comprising at least a first hapten modified ddNTPs and a second hapten modified ddNTP under conditions sufficient to hybridize the first and second hapten modified ddNTPs to respective ones of the captured target complementary sequences as a single base extension of an associated capture oligonucleotide probe, wherein the first and second haptens are different and the first hapten modified ddNTP is a compound of formula I; contacting the bead array comprising the hybridized first and second hapten modified ddNTPs with a labeling composition comprising a first hapten binding partner and a second hapten binding partner under conditions to bind the hapten binding partner to the first hapten and the second hapten binding partner to the second hapten, wherein the first hapten binding partner comprises a first fluorescence label, and the hapten binding partner comprises a second fluorescence label and upon binding of the first and/or second hapten with the respective first

and/or second hapten binding partner, the bead array is fluorescently labeled; and analyzing a fluorescence signal of the bead array after binding of the first and second hapten binding partners.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] Figure 1 is a schematic illustration of a method of making a multivalent hapten-modified ddNTP in accordance with the disclosure;

[0011] Figure 2 is a schematic illustration of a method of making a multivalent hapten-modified ddNTP in accordance with the disclosure;

[0012] Figure 3 is a graph showing fluorescence analysis results comparing use of a multivalent hapten modified ddNTP to a commercially available monovalent DNP-ddNTP and a lab-produced monovalent DNP-ddNTP.

[0013] Figure 4 is a graph showing fluorescence analysis results comparing use of a multivalent hapten modified ddNTP to a commercially available monovalent DNP-ddNTP and a lab-produced monovalent DNP-ddNTP after multiple staining rounds.

[0014] Figure 5 is a graph showing fluorescence analysis in the green channel.

[0015] Figure 6 is a schematic illustration of an example assay in accordance with the disclosure.

[0016] Figure 7A is a graph showing fluorescence analysis in the red channel comparing extension efficiency using a multivalent hapten modified ddNTP to a commercially available monovalent DNP-ddNTP and a lab-produced monovalent DNP-ddNTP.

[0017] Figure 7B is a graph showing fluorescence analysis in the green channel comparing extension efficiency using a multivalent hapten modified ddNTP to a commercially available monovalent DNP-ddNTP and a lab-produced monovalent DNP-ddNTP.

[0018] Figure 8 is a graph showing the results in both extension and non-extension control sets comparing extension efficiency using a multivalent hapten modified ddNTP to a commercially available monovalent DNP-ddNTP and a lab-produced monovalent DNP-ddNTP.

[0019] Figure 9A is a graph showing the fluorescence analysis results of two rounds of staining using increased hapten-modified ddNTP concentrations.

[0020] Figure 9B is a graph showing the fluorescence analysis results of three rounds of staining using increased hapten-modified ddNTP concentrations.

[0021] Figure 10 is a graph showing the signal-to-noise ratio for performance testing on the Infinium EX microarray platform for compounds in accordance with the disclosure as compared to standard design ddNTP.

[0022] Figure 11 is a graph showing call rate and logR deviation metrics for performance testing of compounds in accordance with the disclosure as compared to standard design ddNTPs on the Infinium EX microarray platform.

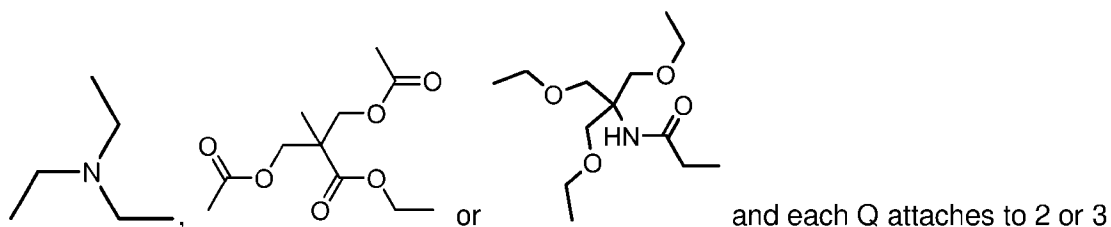
DETAILED DESCRIPTION

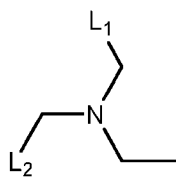
[0023] Compounds of the disclosure are multivalent binding unit modified dideoxynucleotide triphosphates (ddNTP). The compounds of the disclosure can be useful in a variety of applications, including genotyping and multi-omics applications. The binding unit modified ddNTP in accordance with the disclosure include multiple binding units bound or attached to the ddNTP, which has been observed to enhance binding affinity of the fluorescently labeled binding unit binding partner. The binding unit modified ddNTPs in accordance with the disclosure can decrease nonspecific binding. This can beneficially allow for a reduced amount of binding unit binding partner to be used and/or reduce the number of staining rounds required for signal amplification.

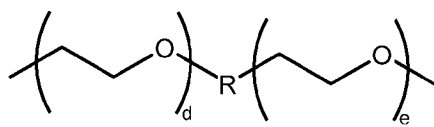
[0024] The binding unit modified ddNTP can be a compound of formula I:



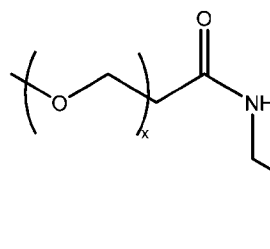
wherein a is at least 2, b is 1 to 3, c is 0 or 1, each ARM comprising a binding unit and a PEO-containing linker, each Q is independently selected from,



ARMS, and each M is , each L₁ and L₂ are independently selected from



, wherein for each L_1 and L_2 , d is independently selected from 1 to 10, and e is independently selected from 1 to 10, R is independently selected from a functionality resulting from a coupling reaction between one of the following pairs: amine-imidoester, amine-pentafluorophenyl ester, amine-*N*-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-



cyclooctyne, azide-norbornene, or azide-alkyne, J is 1 to 48.

[0025] In the compound of formula I, a can be at least 2 and up to 9. For example, a can be 2, 3, 4, 5, 6, 7, 8, or 9.

[0026] The binding unit in each ARM can be the same, for example. The binding unit in two or more of the ARMS can be different, alternatively.

[0027] The binding unit can be any one or more of antibodies, proteins, oligonucleotides, DNA, RNA, drugs, small molecules, dyes, peptides, and the like.

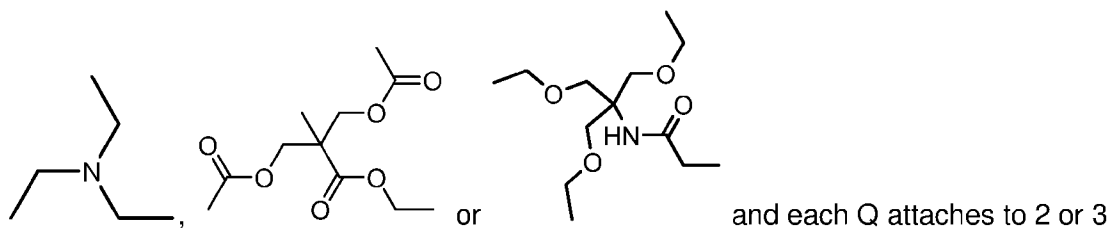
Hapten-Modified Multivalent DDNTP

[0028] In aspects, the compounds of the disclosure are hapten modified dideoxynucleotide triphosphates (ddNTP), which can be useful in genotyping applications to improve signal detection. The hapten modified ddNTP in accordance with the disclosure include multivalent hapten molecules, which has been observed to enhance binding affinity of the fluorescently labeled hapten binding partner. The hapten modified ddNTPs in accordance with the disclosure can decrease nonspecific binding. This can beneficially allow for a reduced amount of hapten binding partner to be used and/or reduce the number of staining rounds required for signal amplification.

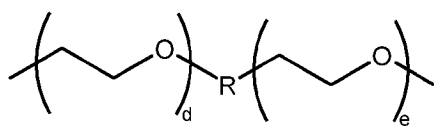
[0029] The hapten modified ddNTP can be a compound of formula IA:



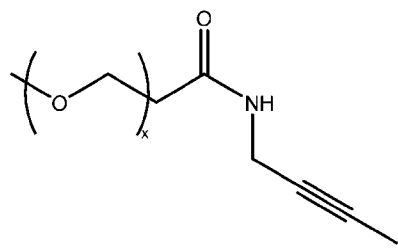
wherein a is at least 2, b is 1 to 3, c is 0 or 1, each ARM comprising a hapten and a PEO-containing linker, each Q is independently selected from,



ARMS, and each M is , each L₁ and L₂ are independently selected from



, wherein for each L₁ and L₂, d is independently selected from 1 to 10, and e is independently selected from 1 to 10, R is independently selected from a functionality resulting from a coupling reaction between one of the following pairs: amine-imidoester, amine-pentafluorophenyl ester, amine-*N*-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-



cyclooctyne, azide-norbornene, or azide-alkyne, J is 1 to 48.

[0030] In the compound of formula IA, a can be at least 2 and up to 9. For example, a can be 2, 3, 4, 5, 6, 7, 8, or 9.

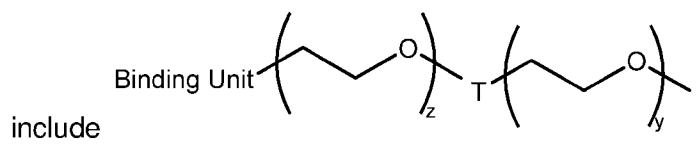
[0031] The hapten in each ARM can be the same, for example. The hapten in two or more of the ARMS can be different, alternatively.

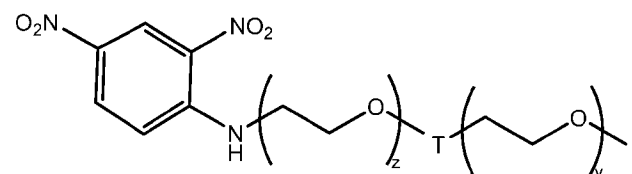
[0032] The hapten in each ARM can be independently selected from, for example, digoxigenin (DIG), 3-nitrotyrosine, biotin, alkyl chloride, O-GlcNAc, alkyl guanine, Ni-NTA and

2,4-dinitrophenol (DNP). The table below shows binding partners for the various haptens that can be included as part of the ARM of compounds of formula (I).

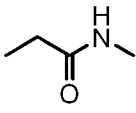
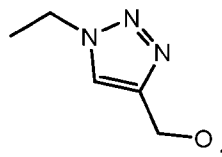
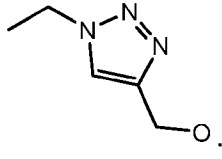
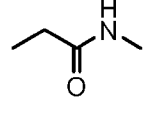
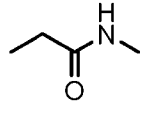
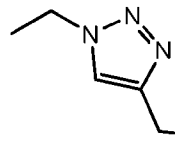
Hapten	Hapten Binding Partner
DNP	Anti-DNP antibody
Biotin	Avidin
DIG	Anti-DIG antibody
O-GlcNAc	WGA
Alkyl guanine	SNAP tag
3-nitrotyrosine	Anti-nitrotyrosine antibody
Ni-NTA	His-Tag

[0033] In either the compounds of formula I or formula IA, for example, each ARM can

include  . For example, each ARM can include

 . In each such ARM, z can be independently selected from 1 to 48, 1 to 12, 1 to 10, 13 to 24, 35 to 36, or 37 to 48, and any values therebetween and ranges defined by such values. In each ARM y can be independently selected from 1 to 48, 1 to 12, 1 to 10, 13 to 24, 35 to 36, or 37 to 48.

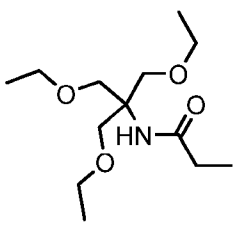
[0034] In each ARM, T is independently selected from a functionality that results from a coupling reaction between one of the following pairs of compounds: amine-imidoester, amine-pentafluorophenyl ester, amine-N-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-cyclooctyne, azide-norbornene, or azide-alkyne.

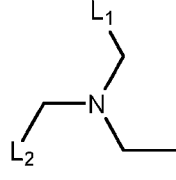
[0035] For example, T can be  or . For example, in each ARM, T can be . For example, in each ARM T can be . For example, T can be  in at least one arm and  in at least one other arm.

[0036] In each ARM, z and y can be same and the compound can be a symmetrical compound. Alternatively, in two or more of the ARM, y and z can be different and the compound can be asymmetrical.

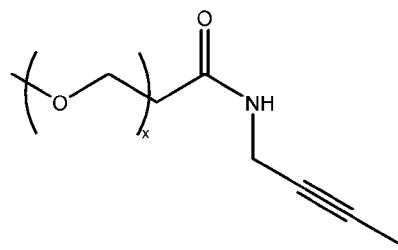
[0037] In the compound of formula I or formula IA, b can be 1, 2 or 3.

[0038] In the compound of formula I or formula IA, c can be 0 or 1. For example, when c is 0,

Q can be . Other selections for Q when c is 0 are also contemplated herein.

M, when present is, . Each L₁ and L₂ are independently selected from $\left(\text{CH}_2 \right)_d \text{O} \left(\text{CH}_2 \right)_e$, wherein for each L₁ and L₂, d is independently selected from 1 to 48, 1 to 12, 1 to 10, 13 to 24, 35 to 36, or 37 to 48, and e is independently selected from 1 to 48, 1 to 12, 1 to 10, 13 to 24, 35 to 36, or 37 to 48, and R is independently selected from the functionality resulting from a coupling reaction between one of the following pairs amine-imidoester, amine-pentafluorophenyl ester, amine-N-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl

disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-cyclooctyne, azide-norbornene, or azide-alkyne.

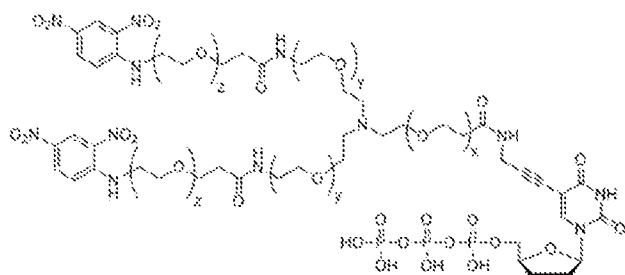


[0039] In the compound of formula I or formula IA, J is wherein x is from 1 to 48, 1 to 12, 1 to 10, 13 to 24, 35 to 36, or 37 to 48.

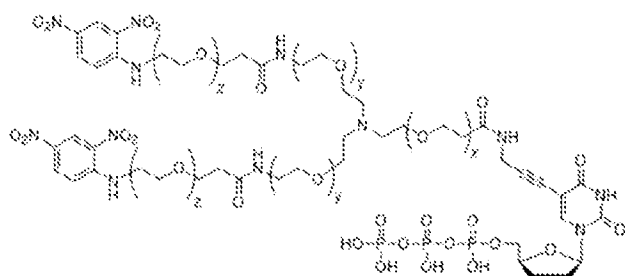
[0040] The ddNTP in the compound of formula I or formula IA can be any one of ddATP, ddTTP, ddUTP, ddGTP, or ddCTP.

[0041] In various cases, the compound of formula I or formula IA is a structure as shown in Table A

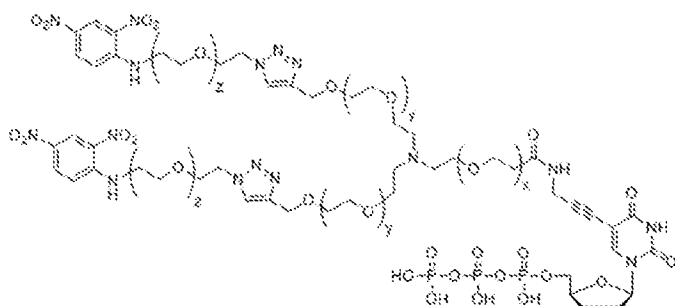
Table A



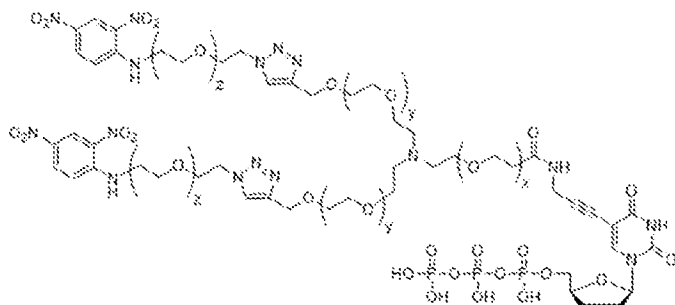
where z and y have the same value on each ARM, and x is 1 to 48.



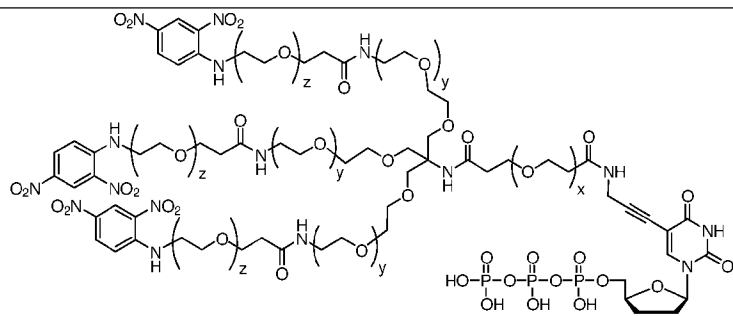
where at least one of z and y have different values on each ARM, and x is 1 to 48.



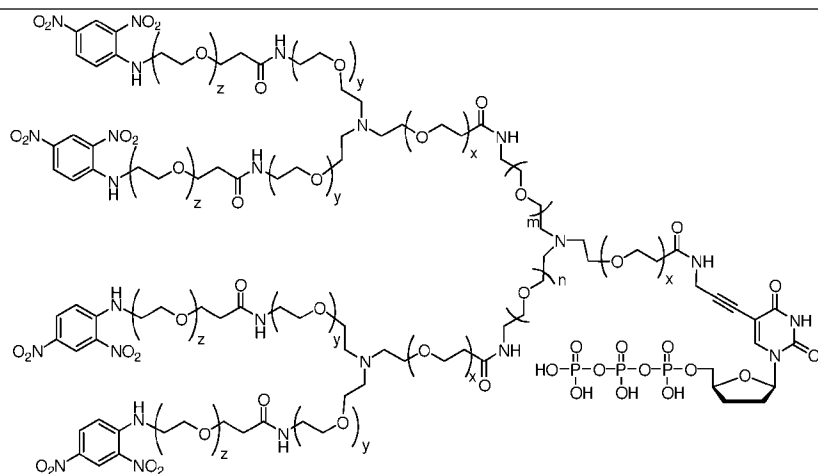
where z and y have the same value on each ARM, and x is 1 to 48.



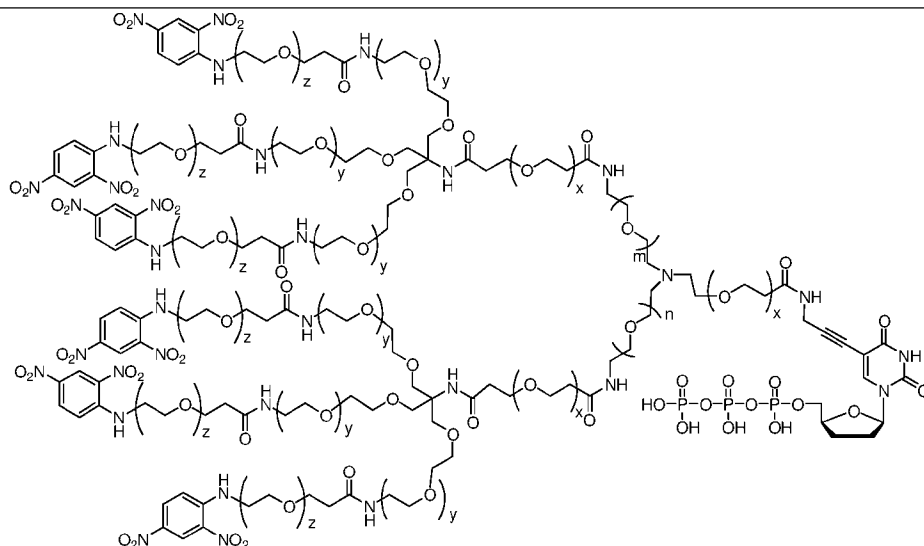
where at least one of z and y have different values on each ARM, and x is 1 to 48.



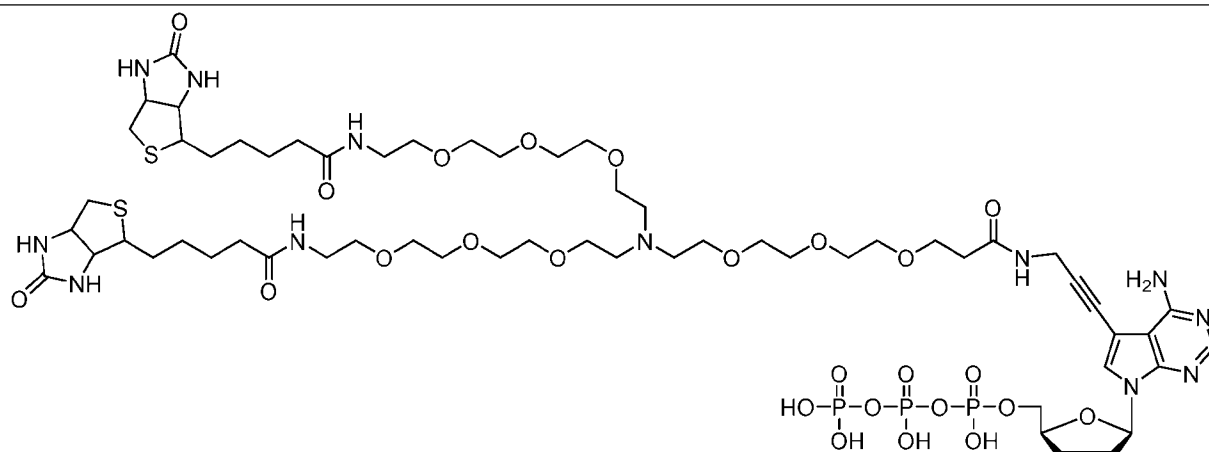
, wherein z , and y are each independently selected on each arm to be from 1 to 48, and each x is independently selected to be from 1 to 48.

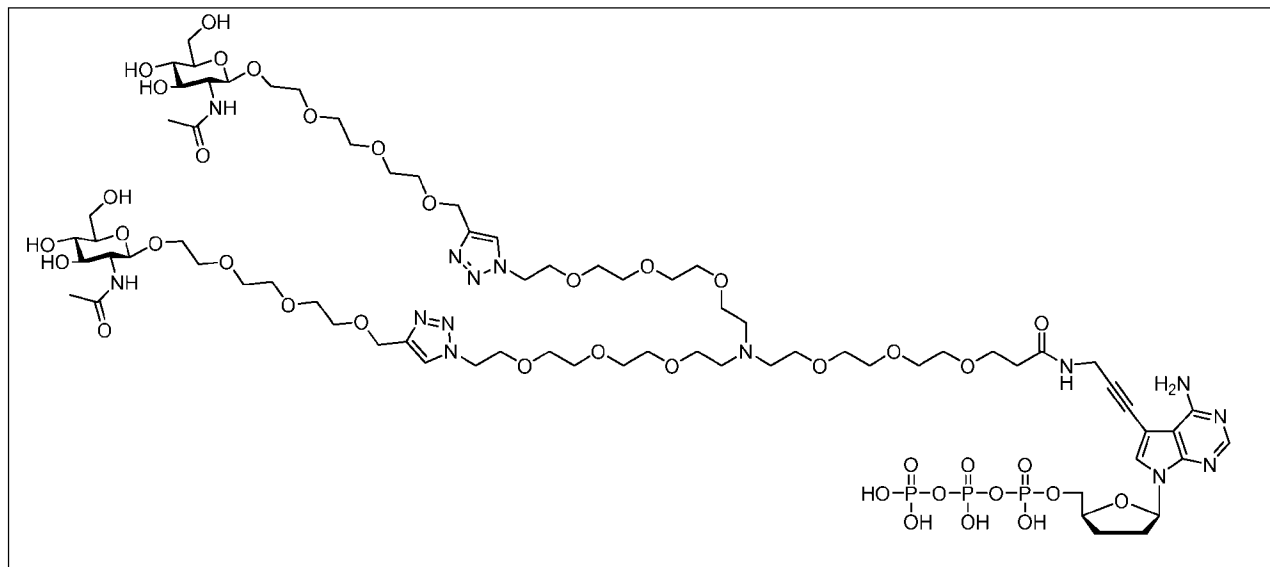


, wherein z and y are each independently selected on each arm to be from 1 to 48, and each x is independently select to be from 1 to 48.



, wherein z and y are each independently selected on each arm to be 1 to 48, and each x is independently selected to be from 1 to 48.





Kits

[0042] A kit can include, for example, a bead array comprising a plurality of beads, each bead having at least one capture oligonucleotide probe attached to a surface thereof, a binding unit binding partner comprising a first label, at least one compound of formula I and/or formula IA, and instructions for use. The binding unit binding partner containing the first label selectively binds to the binding unit present on the ARM(s) of the compound of formula I and/or formula IA. For example, where the hapten is DNP, the hapten binding partner can be an anti-DNP antibody comprising a first fluorescence label. Upon contact with the anti-DNP antibody with the compound of formula I, the anti-DNP antibody hybridizes to the DNP, thereby labeling the bead with the first label. The first label can be, for example, a fluorescence label. The label or interaction between the binding unit and the binding unit binding pair can results in a detectable signal such as fluorescence, bioluminescent, or electrical detection. The kit can be useful for screening and/or identifying binding pairs between binding units and binding unit binding partners. For example, a plurality of binding unit bindings partners and/or a plurality of binding units can be included in the kit to screen for binding interactions through use of the kit.

[0043] A kit for genotyping can include, for example, a bead array comprising a plurality of beads, each bead having at least one capture oligonucleotide probe attached to a surface thereof, a hapten binding partner comprising a first fluorescence label, at least one compound of formula I, and instructions for use. The hapten binding partner containing the first fluorescence label selectively binds to the hapten present on the ARM(s) of the compound of formula I. For example, where the hapten is DNP, the hapten binding partner can be an anti-DNP antibody

comprising a first fluorescence label. Upon contact with the anti-DNP antibody with the compound of formula I, the anti-DNP antibody hybridizes to the DNP, thereby labeling the bead with the first fluorescence label.

[0044] Binding unit binding partners can be independently selected from antibodies, proteins, oligonucleotides, DNA, RNA, drugs, small molecules, dyes, peptides, and the like. For example, the hapten binding partners can be independently selected from anti-DNP antibodies, avidin, anti-DIG antibodies, WGA, SNAP tags, anti-nitrotyrosine antibodies, and His-Tag. Hapten/Hapten binding partner pairs are identified above, and suitable binding unit binding partners can be selected based on the binding unit present on the compounds of formula I. The binding unit on the compound of formula I can be, for example, a small molecule, peptide, or antibody, while the binding unit binder partner can be a protein, oligo, RNA or DNA. Alternatively, the binding unit can be a protein and the binding unit binding partner can be a small molecule, peptide, or antibody to be screened for binding to the protein, oligo, RNA, or DNA. The binding unit binding partner can include a label, such as a fluorescence label, bioluminescence label, or charge (for electrical detection), to allow for indication of binding pairs that occur. The compounds and methods of the disclosure can be used, for example, to screen for interactions that may be useful, for example, in drug development.

[0045] Compounds and methods of the disclosure can utilize binding unit/binding unit partners for various applications. For example, the compounds and methods of the disclosure can be used in genotyping applications, binding pair screening and/or identification, and multi-omics. For example, counting multi-omics can be performed using the compounds of the disclosure in which the binding sets provided by the compounds of the disclosure and the binding unit partner can allow for detection of multiple analytes such as DNA, RNA, and protein binding sites.

The kit can include multiple different compounds of formula I and/or formula IA, for example, with different ddNTP portions. For example, the compounds of formula I can include a binding unit-modified ddATP and a binding unit-modified ddUTP, such that a single fluorescence signal from the first fluorescence label will be generated when both A and T are present on the bead. Single detection of A and T as a single channel is commonly used in various genotyping process. Any known bead array technology can be used with the kits of the disclosure. For example, the beads can be an Infinium bead array (Illumina).

[0046] For example, the kit can include two sets of compounds of formula I and/or formula IA, each set having different binding units for binding to different fluorescently labeled binding unit

binding partners. For example, a first set of the compound of formula I can include DNP as the hapten present on the arm, and an anti-DNP antibody with a first fluorescent label can be provided. This first set of compounds can include a two subset of compounds, one subset including ddATP as the ddNTP and the other subset including ddUTP as the ddNTP. The first set of compounds of formula I would result in fluorescently labeling A and T detection with a first fluorescent label. The kit can further include a second set of compounds of formula I, for example, where the hapten present on the arm is biotin and the hapten binding partner is avidin with a second fluorescent label, different from the first fluorescent label. The second set of compounds of formula I can include two subsets, one in which the ddNTP is ddGTP and the other in which the ddNTP is ddCTP. The second set of compounds of formula I would result in fluorescently labeling G and C detection with the second fluorescent label. This can result in two channels of detection, for example, a red (A and T) and a green channel (G and C).

[0047] It is also contemplated herein that kits of the disclosure can include other platforms, such as non-bead platforms, in which the capture oligonucleotide probe is attached to a different substrate. The methods and kits of the disclosure can be useful, for example, with DNA, RNA, and protein sequencing technologies.

[0048] Process of Using the Compounds of the Disclosure Methods of the disclosure can utilize the compounds for various application, such as identifying binding pairs, phenotyping, genotyping, and in multi-omics applications.

[0049] A method of using the compounds of the disclosure can include capturing a target oligonucleotide on a substrate or bead array having a capture oligonucleotide. Upon contact with the sample, the capture oligonucleotide binds the target oligonucleotide from the sample, thereby capturing the target oligonucleotide. The substrate having the captured target oligonucleotides are then contacted with one or more compounds of formula I and/or formula IA, wherein the ddNTP portion of the compound of formula I and/or formula IA binds to the target capture oligonucleotide as a single base extension thereof. The bead array or substrate is then contacted with one or more binding unit partners. The binding unit partners can be labeled such that binding between the binding unit present on the compound of formula I and/or formula IA and the binding unit partner can be identified. Selection of the binding pairs (binding unit and corresponding binding unit binding partners) can be made depending on the information desired to obtain from the process. For example, binding unit pairs can be used in methods of drug development and discovery to identifying potential candidates with a targeted activity.

[0050] Compounds of the disclosure can be used, for example, in processes of genotyping using generally the same procedures as conventional ddNTP-hapten molecules. In general, the beads array or other substrate includes a capture oligonucleotide that, upon contact with a sample, captures through complementary binding target oligonucleotides from the sample. The capture oligonucleotide binds to the complementary sequence in the sample, stopping one base before the locus of interest, thereby providing a bead having a captured target complementary sequence. Allele specificity is conferred by a single base extension that includes one of four labeled nucleotides. The labeled nucleotides hybridized to the captured target oligonucleotide having a complementary base thereto, thereby extending the capture oligonucleotide by a single base. At least one of the label nucleotides can be a compound of formula I and/or formula IA. After single base extension using the labeled nucleotides, the sample is contacted with one or more fluorescently labeled hapten binding partners. For single signal detecting, a single hapten binding partner that specifically binds to the hapten present on the compound of formula I and/or formula IA is used. For multiple signal detection, multiple hapten binding partners, each having distinct fluorescent labels can be contacted with the sample to bind to the respective hapten present on the labeled nucleotides.

[0051] In some cases, the labeled nucleotides can include the compound of formula I for detection of a single nucleotide, while conventional labeled nucleotides can be used for detection of the other nucleotides. In some cases, the detection can be two-channel detection, whereby A and T bases are detected by a first signal and G and C bases are detected by a second single. In such cases, the labeled nucleotides comprising ddATP and ddTTP or ddUTP can be the compounds of formula I and/or formula IA.

[0052] For example, a process of genotyping using a two-signal detection, with a first signal detecting A and T and a second signal detecting G and C can include contacting a bead array or other substrate with a sample DNA, each bead or surface region of the bead array or other substrate, respectively, includes one or more capture oligonucleotide probes that each bind to a target complementary sequence in the sample, wherein the target complementary sequence binds to the one or more capture oligonucleotide probes stopping at least one base before a locus of interest thereby providing a bead array having captured target complementary sequence bound thereto. The bead array or other substrate having the captured target complementary sequences bound to the capture oligonucleotide probe is then contacted with a composition comprising at least a first hapten modified ddNTPs and a second hapten modified ddNTP under conditions sufficient to hybridize the first and second hapten modified ddNTP to

respective ones of the captured target complementary sequences as a single base extension of an associated capture oligonucleotide probe. The first and second haptens are different. One or both of the first and second hapten modified ddNTPs can be a compound of formula I. The method further includes contacting the bead array comprising the hybridized first and second hapten modified ddNTPs with a labeling composition that includes a first hapten binding partner and a second hapten binding partner under conditions sufficient to bind the first hapten binding partner to the first hapten and the second hapten binding partner to the second hapten. The first hapten binding partner comprises a first fluorescence label, and the second hapten binding partner comprises a second fluorescence label. Upon binding of the first and/or second hapten modified ddNTPs with the respective first and/or second hapten binding partner, the bead array or other substrate is fluorescently labeled. The fluorescent signal resulting from the bead array can then be analyzed for genotyping the sample. For example, in the process, the first hapten modified ddNTP can include a first hapten modified ddATP and a first hapten modified ddUTP and the second hapten modified ddNTP comprises a second hapten modified ddGTP and a second hapten modified ddCTP.

[0053] In some instance, for example, when using DNP as the first hapten and biotin as the second hapten, it can be advantageous to have the first hapten modified ddNTP (DNP modified ddNTP) to be a compound of formula I and the biotin modified ddNTP to be monovalent. This can balance the signal in the red and green channels, for example. Conventionally, ddATP-DNP and ddUTP-DNP are used in genotyping applications to generate a red channel signal. It has been observed that when monovalent ddNTP are utilized, the red channel signal is weaker than a green channel signal resulting from ddGTP-biotin and ddCTP-biotin labels. .

[0054] In some cases, all four labeled nucleotides can be compounds of formula I. For example, for two channel detection, a first set of compounds of formula I can include a first hapten and a second set of compounds of formula I can include a second hapten. For example, the first hapten can be DNP and the second hapten can be biotin. Within the first set, two subsets of compounds of formula I can be further provided, with one set having ddATP as the nucleotide and the other set having ddTTP or ddUTP as the nucleotide. Similarly, within the second set, two subsets of compounds of formula I can be provided, with one set having ddGTP and the other subset having ddCTP. As a result of the distinct sets of compounds of formula I, detection of A or T can be made by the presence of a first signal resulting from the first fluorescent label, while detection of G or C can be made by the presence of a second signal resulting the second fluorescent label/

[0055] In still further embodiments, the compound of formula I can be provided as four distinct sets, each set having a different nucleotide and different hapten. This can allow, for example, for four distinct fluorescent labels to be applied by a corresponding hapten binding partner.

[0056] In any of the methods disclosed herein, the contacting of the bead array or other substrate with the hapten binding partner can be performed one or more times for each of the distinct hapten binding partners. For example, the staining by contacting can be performed 2 or more times, for example 3 or more times. For example, when the hapten-labeled nucleotide is a compound of formula I, the contacting step for staining the bead array or other substrate can be performed a reduced number of times as compared when a monovalent hapten labeled nucleotide while maintaining the same or even having improved signal intensity.

[0057] Use of the multi-valent hapten modified ddNTPs of the disclosure can improve hapten binding partner binding affinity to the hapten, thereby allowing for a reduced amount of hapten binding partner to be used and/or reduce a number of staining or contacting rounds needed for labeling the ddNTP with the fluorescent labeled hapten binding partner.

[0058] When using the compounds of the disclosure with bead arrays in particular, the improved binding efficiency and increased signal to noise ratio resulting therefrom, can also allow the bead array to be formed using smaller beads and/or with reduced bead spacing. This can beneficially increase the bead density in the array, allowing for an increased amount of content that can be assayed by the array.

Chemical Definitions

[0059] As used herein, the term "alkyl" refers to straight chained and branched saturated hydrocarbon groups containing one to thirty carbon atoms, for example, one to twenty carbon atoms, or one to ten carbon atoms. The term C_n means the alkyl group has "n" carbon atoms. For example, C_4 alkyl refers to an alkyl group that has 4 carbon atoms. C_{1-6} alkyl refers to an alkyl group having a number of carbon atoms encompassing the entire range (i.e., 1 to 6 carbon atoms), as well as all subgroups (e.g., 1-5, 2-5, 1-4, 2-5, 1, 2, 3, 4, 5, and 6 carbon atoms). Nonlimiting examples of alkyl groups include, methyl, ethyl, n-propyl, isopropyl, n-butyl, sec-butyl (2-methylpropyl), and t-butyl (1,1-dimethylethyl). Unless otherwise indicated, an alkyl group can be an unsubstituted alkyl group or a substituted alkyl group.

[0060] As used herein, the term "alkylene" refers to a bivalent saturated aliphatic radical. The term C_n means the alkylene group has "n" carbon atoms. For example, C_{1-6} alkylene refers to an

alkylene group having a number of carbon atoms encompassing the entire range, as well as all subgroups, as previously described for "alkyl" groups.

[0061] As used herein, the term "alkene" or "alkenyl" is defined identically as "alkyl" except for containing at least one carbon-carbon double bond, and having two to thirty carbon atoms, for example, two to twenty carbon atoms, or two to ten carbon atoms. The term C_n means the alkenyl group has "n" carbon atoms. For example, C_4 alkenyl refers to an alkenyl group that has 4 carbon atoms. C_{2-7} alkenyl refers to an alkenyl group having a number of carbon atoms encompassing the entire range (i.e., 2 to 7 carbon atoms), as well as all subgroups (e.g., 2-6, 2-5, 3-6, 2, 3, 4, 5, 6, and 7 carbon atoms). Specifically contemplated alkenyl groups include ethenyl, 1-propenyl, 2-propenyl, and butenyl. Unless otherwise indicated, an alkenyl group can be an unsubstituted alkenyl group or a substituted alkenyl group. Unless otherwise indicated, an alkenyl group can be a cis-alkenyl or trans-alkenyl.

[0062] As used herein, the term "alkyne" or "alkynyl" is defined identically as "alkyl" except for containing at least one carbon-carbon triple bond, and having two to thirty carbon atoms, for example, two to twenty carbon atoms, or two to ten carbon atoms. The term C_n means the alkynyl group has "n" carbon atoms. For example, C_4 alkynyl refers to an alkynyl group that has 4 carbon atoms. C_{2-7} alkynyl refers to an alkynyl group having a number of carbon atoms encompassing the entire range (i.e., 2 to 7 carbon atoms), as well as all subgroups (e.g., 2-6, 2-5, 3-6, 2, 3, 4, 5, 6, and 7 carbon atoms). Specifically contemplated alkynyl groups include ethynyl, 1-propynyl, 2-propynyl, and butynyl. Unless otherwise indicated, an alkynyl group can be an unsubstituted alkynyl group or a substituted alkynyl group.

[0063] As used herein, the term "carbocycle" refers to an aromatic or nonaromatic (i.e., fully or partially saturated) ring in which each atom of the ring is carbon. A carbocycle can include, for example, from three to ten carbon atoms, four to eight carbon atoms, or five to six carbon atoms. As used herein, the term "carbocycle" also includes polycyclic ring systems having two or more cyclic rings in which two or more carbons are common to two adjoining rings wherein at least one of the rings is carbocyclic, e.g., the other cyclic rings can be cycloalkyls, cycloalkenyls, aryls, heteroaryls, and/or heterocycles.

[0064] As used herein, the term "cycloalkyl" specifically refers to a non-aromatic carbocycle. The term C_n means the cycloalkyl group has "n" carbon atoms. For example, C_5 cycloalkyl refers to a cycloalkyl group that has 5 carbon atoms in the ring. C_{5-8} cycloalkyl refers to cycloalkyl groups having a number of carbon atoms encompassing the entire range (i.e., 5 to 10 carbon atoms), as well as all subgroups (e.g., 5-10, 5-9, 5-8, 5-6, 6-8, 7-8, 5-7, 5, 6, 7, 8, 9 and

10 carbon atoms). Nonlimiting examples of cycloalkyl groups include cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, and cyclooctyl. Unless otherwise indicated, a cycloalkyl group can be an unsubstituted cycloalkyl group or a substituted cycloalkyl group.

[0065] As used herein, the term “aryl” refers to an aromatic carbocycle, and can be monocyclic or polycyclic (e.g., fused bicyclic and fused tricyclic) carbocyclic aromatic ring systems. Examples of aryl groups include, but are not limited to, phenyl, naphthyl, tetrahydronaphthyl, phenanthrenyl, biphenylenyl, indanyl, indenyl, anthracenyl, fluorenyl, tetralinyl. Unless otherwise indicated, an aryl group can be an unsubstituted aryl group or a substituted aryl group.

[0066] As used herein, the term “heterocycle” is defined similarly as carbocycle, except the ring contains one to four heteroatoms independently selected from oxygen, nitrogen, and sulfur. For example, a heterocycle can be a 3-10 membered aromatic or non-aromatic ring having 1 or 2 heteroatoms selected from N, O, and S. As another example, a heterocycle can be a 5-6 membered ring having 1 or 2 ring heteroatoms selected from N, O, and S. Nonlimiting examples of heterocycle groups include piperidine, tetrahydrofuran, tetrahydropyran, dihydrofuran, morpholine, oxazepaneryl, thiazole, pyrrole, and pyridine.

[0067] Carbocyclic and heterocyclic groups can be saturated or partially unsaturated ring systems optionally substituted with, for example, one to three groups, independently selected alkyl, alkoxy, alkylene, OH, C(O)NH₂, NH₂, oxo (=O), aryl, haloalkyl, haloalkoxy, C(O)-alkyl, SO₂alkyl, halo, OH, NHC₁₋₃alkylene-aryl, OC₁₋₃alkylene-aryl, C₁₋₃alkylene-aryl, and C₃₋₆heterocycloalkyl having 1-3 heteroatoms selected from N, O, and S. Heterocyclic groups optionally can be further N-substituted as described herein. Other substituents contemplated for the disclosed rings is provided elsewhere in this disclosure.

[0068] As used herein, the term “heteroaryl” refers to an aromatic heterocycle, and can be monocyclic or polycyclic (e.g., fused bicyclic and fused tricyclic) aromatic ring systems, wherein one to four ring atoms are selected from oxygen, nitrogen, or sulfur, and the remaining ring atoms are carbon, said ring system being joined to the remainder of the molecule by any of the ring atoms. Nonlimiting examples of heteroaryl groups include, but are not limited to, pyridyl, pyridazinyl, pyrazinyl, pyrimidinyl, pyrrolyl, pyrazolyl, imidazolyl, thiazolyl, tetrazolyl, oxazolyl, isooxazolyl, thiadiazolyl, oxadiazolyl, furanyl, thienyl, quinolinyl, isoquinolinyl, benzoxazolyl, benzimidazolyl, benzofuranyl, benzothiazolyl, triazinyl, triazolyl, purinyl, pyrazinyl, purinyl, indolinyl, phthalizinyl, indazolyl, quinolinyl, isoquinolinyl, cinnolinyl, quinazolinyl, naphthyridinyl,

pyridopyridinyl, indolyl, 3H-indolyl, pteridinyl, and quinoxaliny. Unless otherwise indicated, a heteroaryl group can be an unsubstituted heteroaryl group or a substituted heteroaryl group.

[0069] As used herein, the term “hydroxy” or “hydroxyl” as used herein refers to an “—OH” group. Accordingly, a “hydroxyalkyl” refers to an alkyl group substituted with one or more —OH groups.

[0070] As used herein, the term “alkoxy” or “alkoxyl” refers to a “—O-alkyl” group.

[0071] As used herein, the term “halo” is defined as fluoro, chloro, bromo, and iodo. Accordingly, a “haloalkyl” refers to an alkyl group substituted with one or more halo atoms. A “haloalkoxy” refers to an alkoxy group that is substituted with one or more halo atoms.

[0072] A “substituted” functional group (e.g., a substituted alkyl, cycloalkyl, aryl, or heteroaryl) is a functional group having at least one hydrogen radical that is substituted with a non-hydrogen radical (i.e., a substituent). Examples of non-hydrogen radicals (or substituents) include, but are not limited to, alkyl, cycloalkyl, alkenyl, cycloalkenyl, alkynyl, ether, aryl, O-alkylene aryl, N-alkylene aryl, alkylene aryl, heteroaryl, heterocycloalkyl, hydroxy, hydroxyalkyl, haloalkoxy, amido, oxy (or oxo), alkoxy, ester, thioester, acyl, carboxyl, cyano, nitro, amino, sulfhydryl, and halo. When a substituted alkyl group includes more than one non-hydrogen radical, the substituents can be bound to the same carbon or two or more different carbon atoms.

Synthesis of Compounds of the Disclosure

[0073] The compounds provided herein can be synthesized using conventional techniques readily available starting materials known to those skilled in the art. In general, the compounds provided herein are conveniently obtained via standard organic chemistry synthesis methods.

[0074] Although not limited to any one or several sources, classic texts such as Smith, M. B., March, J., March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure, 5th edition, John Wiley & Sons: New York, 2001 ; and Greene, T.W., Wuts, P.G. M., Protective Groups in Organic Synthesis, 3rd edition, John Wiley & Sons: New York, 1999, are useful and recognized reference textbooks of organic synthesis known to those in the art. The following descriptions of synthetic methods are designed to illustrate, but not to limit, general procedures for the preparation of compounds of the present disclosure.

[0075] The synthetic processes disclosed herein can tolerate a wide variety of functional groups; therefore, various substituted starting materials can be used. The processes generally provide the desired final compound at or near the end of the overall process, although it may be

desirable in certain instances to further convert the compound to a pharmaceutically acceptable salt thereof.

[0076] In general, the compounds of the disclosure can be synthesized in line with the examples shown below. For example, the compounds can be prepared by alkylation of the appropriate amine having a carboxyl group, with appropriate protecting groups as necessary. The intermediate can be saponified, for example, to expose a reactive carboxylate. Then, amide coupling between the appropriate amine and the free carboxylate can occur.

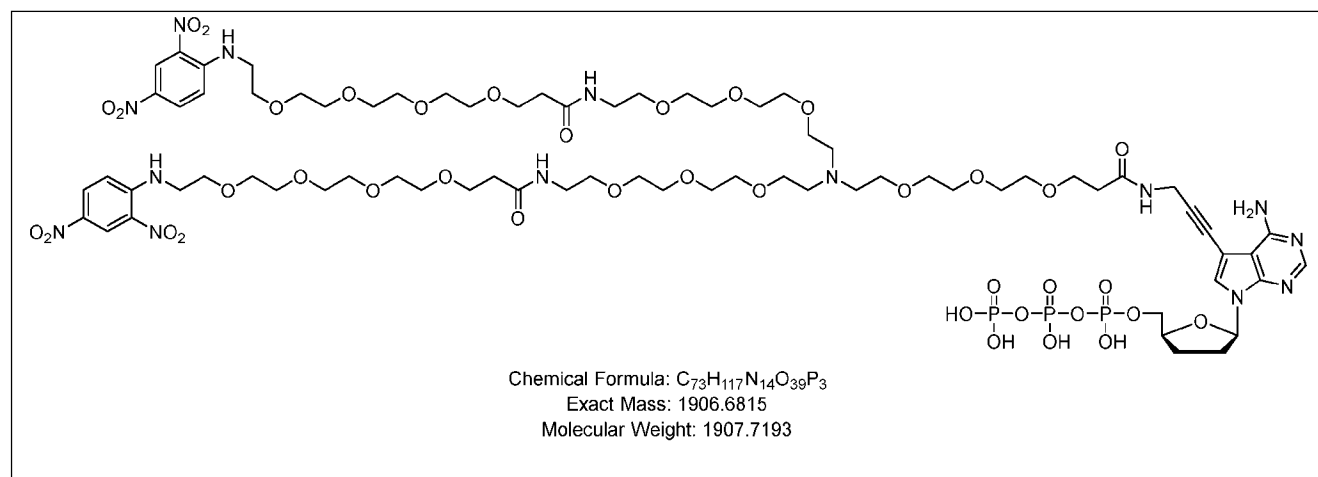
[0077] The amine for the amide coupling noted above can be prepared via known synthetic techniques using appropriate starting materials and protecting groups, as necessary.

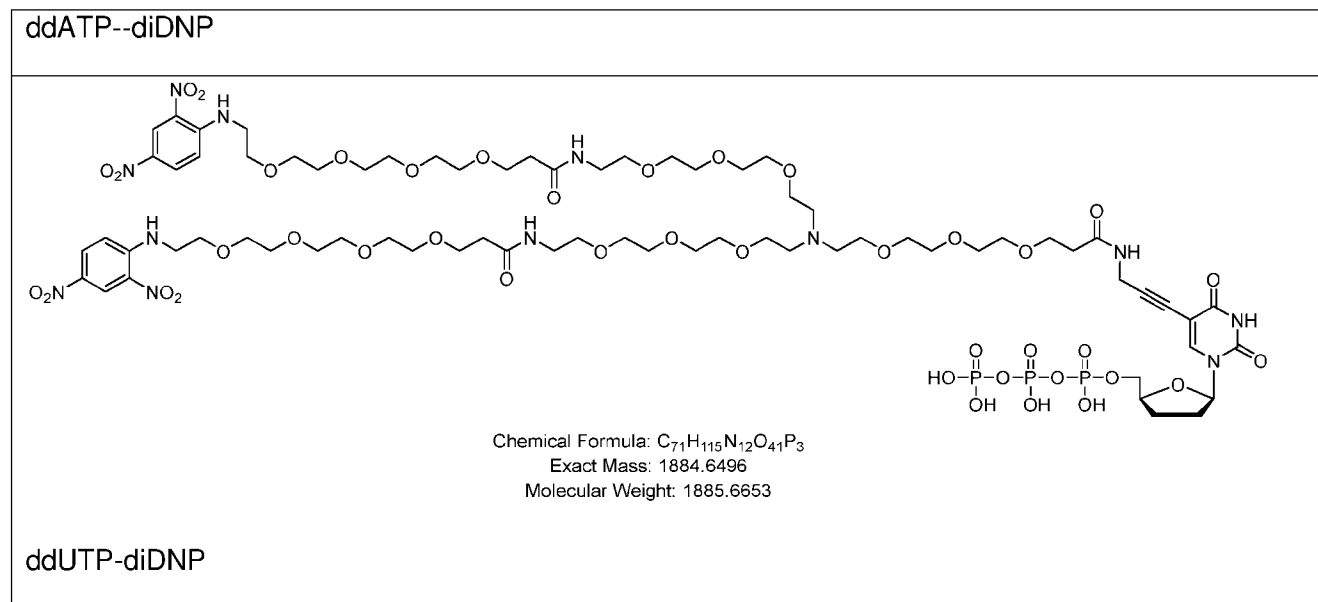
[0078] Further modifications can be performed, e.g., to introduce additional substituents such as halo groups or alkyl groups.

EXAMPLES

Example 1

[0079] Single base extension reagent was formulated with compounds in accordance with the disclosure, as well as standard monovalent compounds as controls. A single base extension reagent having a compound in accordance with formula I was prepared by admixing single base extension premix (EML premix), SSBs, tTAQ608, and the labeled nucleotides. For the G and C (green) channel detection, conventional monovalent ddGTP-biotin and ddCTP-biotin labels were used. For the A and T (red) channel detection following compounds in accordance with the disclosure were used. In the data set, the samples run using the extension reagent formulated with the compounds of formula I are referred to as diEML

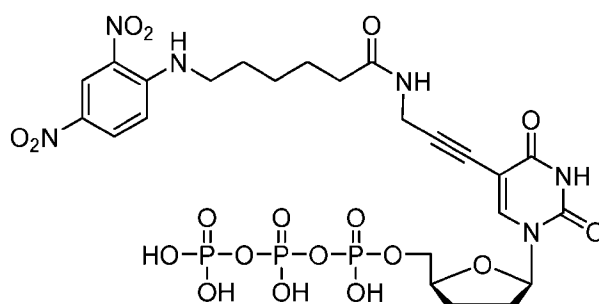




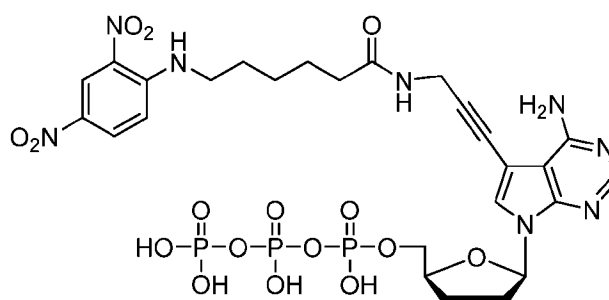
[0080] Figure 1 shows a synthesis scheme for making the ddATP-diDNP and Figure 2 shows a synthesis scheme for making ddUTP-diDNP. The ddATP-diDNP and ddUTP-diDNP were prepared by admixing 1 equiv bis-amine linker with 2.2 equiv DNP-NHS dissolved in 0.2M $NaHCO_3$:DMSO (1.4 v/v). The admixture was stirred for 8 hours at room temperature. The mixture was purified by HPLC using water and ACN as eluant. The di-DNP-COOH product was then lyophilized. The lyophilized product was then reconstituted in DMF. 1.5 equiv TSTU was added to the mixture, followed by 20 equiv DIPEA and then stirred for 30 min at room temperature. 1 equiv PA-ddNTP was then added and stirred overnight at room temperature. The resulting product was purified by HPLC using 50 mM TEAA (pH 7.2) and ACN as eluant. The product was then lyophilized.

[0081] Controls were run using a convention, commercially available single base extension premix having monovalent labeled nucleotides (referred to herein as pEML), as well as a single base extension premix formulated with monovalent ddNTPs (referred to herein as devEML). In particular, the formulated control included the same EML premix, SSBs, and tTAQ608 as used in the diEML, but included as the labeled nucleotides DNP-11-ddATP, DNP-11-ddUTP, as well as the same ddGTP-biotin and ddCTP-biotin used in diEML. Structures for DNP-11-ddATP and DNP-11-ddUTP are shown below.

DNP-11-ddUTP



DNP-11-ddATP



[0082] Each of the reagents included a nominal ddNTP concentration of 0.125 μ M. The reagents were separately tested using an Infinium bead array (Illumina) with a standard MSA7 workflow (Infinium XT-ST workflow) with DPB purity instead of precipitate. Amplification was preformed using RAM 10 nm input with 18-hour incubation. The sample tested for this genotype testing included a single DNA type – NA11993.

[0083] Staining with the fluorescent labeled hapten binding partner was performed using the Xstain protocol in Tecan. Three staining rounds were performed and thee signal strength was evaluated with an iScan N392 after each round.

[0084] Figure 3 illustrates the signal in the red channel for each of the reagent samples tested. The diEML having the compounds in accordance with the disclosure provided a detectible signal even after a single staining round. Figure 3 further illustrates that ddNTPs coupled to PEG linkers with multiple DNP moieties can be incorporated by the tTaq608 polymerase using the same biochemical conditions as the monovalent ddNTP linker DNP species.

[0085] Referring to Figure 4, it can be seen that the increased signal resulting in a third round of staining as compared to the second round of staining was comparable between the diEML and pEML control, and that the diEML had a greater improvement in signal as compared to the devEML control.

[0086] Finally, Figure 5 illustrates that detection in the green channel with the biotin-modified ddGTP and ddCTP was unaffected by the use of the compounds in accordance with the claims.

Example 2

[0087] The compounds of the disclosure were also evaluated using synthetic probes. The same single base extension reagents as described in Example 1 were used, including the controls. The example evaluated the extension efficiency of A, T, C, and G nucleotides from a hairpin probe. Both red (A, T), and green (C, G) channels were monitored. Figure 6 is a schematic illustration of the process.

[0088] Referring to Figures 7A and 7B, it was observed that the compounds in accordance with the disclosure (divalent DNP formulations) showed a higher signal. Nonspecific extension was observed by the signal in extension C and G. Nucleotide incorporation efficiency and specificity is impacted by templating base, pH, and the concentration and stoichiometry of the components: ddNTP-diDNP, polymerase, and catalytic metal (Mg^{2+}). It is contemplated that further improvement of efficiency and specificity can be achieved with the compounds of the disclosure with further optimization of these parameters.

[0089] Figure 8 illustrates that increased signal was achieved with the A and T extension, in the diEML, utilizing the compound of the disclosure, as compared to the controls. The dashed box in Figure 7 is the non-extension control set and no increase in signal was anticipated in that non-extension control set.

Example 3

[0090] The effect of ddNTP concentration on incorporation rates was evaluated by comparing the performance of reagents having 0.125 μ M ddNTP to reagents having 0.625 μ M ddNTP. Other than concentrations, the reagents were prepared as described in Example 1.

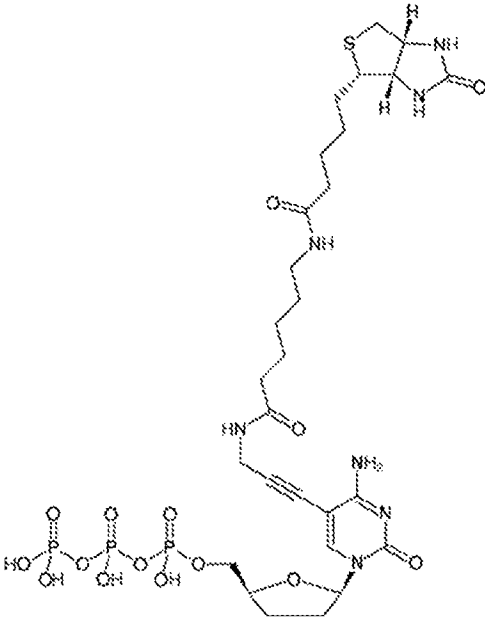
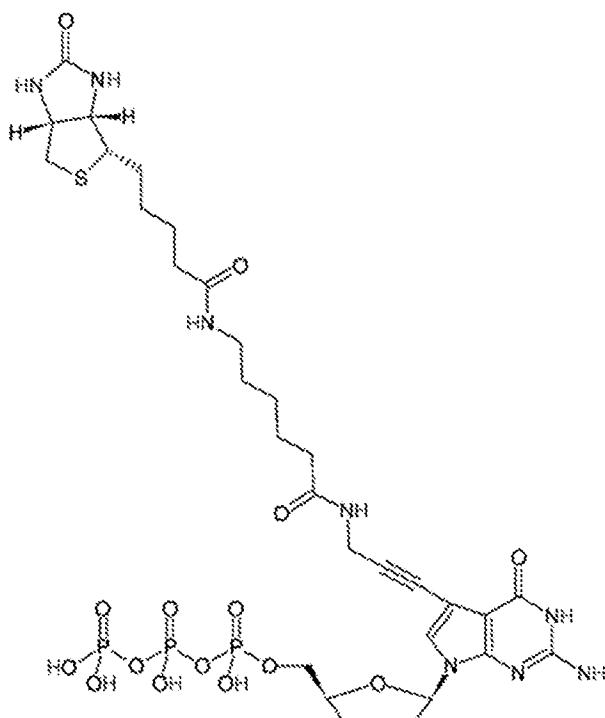
[0091] Referring to Figure 9A, it was observed that after 2 rounds of staining, the diEML having an increased concentration of the compound in accordance with the disclosure of 0.625 μ M showed significantly higher intensity in the red channel as compared to the production EML, but not significantly higher than the devEML also formulated with the higher (0.625 μ M)

concentration of monovalent ddNTP. However, as shown in Figure 9B, increased signal of the higher concentration diEML sample was observed.

Example 4

[0092] Compounds of the disclosure were evaluated using the Infinium EX[®] platform. Performance using compounds in accordance with the disclosure (ddATP-diDNP and ddUTP-diDNP) were compared to standard ddNTPs as shown in the table below.

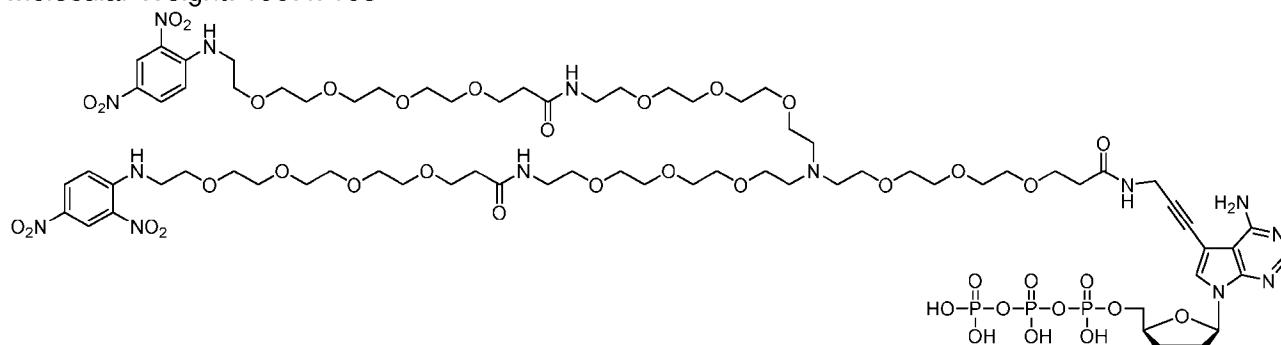
Standard Compounds Used for Comparison

Bio-11-ddCTP (Standard Design) Mol Formula: C₂₈H₄₄N₇O₁₅P₃S (free acid) MW: 843.67 g/mol (free acid) Exact Mass: 843.18 g/mol (free acid)  Expected m/z M-H: 842,175079	Bio-11-ddGTP (Standard Design) Mol Formula: C₃₀H₄₅N₈O₁₅P₃S (free acid) MW: 882.71 g/mol (free acid) Exact Mass: 882.19 g/mol (free acid)  Expected m/z M-H: 881.185978
---	--

Compounds in accordance with the disclosure:**ddATP-diDNP****Divalent DNP**Mol. Formula $C_{73}H_{117}N_{14}O_{39}P_3$

Exact Mass: 1906.6815

Molecular Weight: 1907.7193

Chemical Formula: $C_{73}H_{117}N_{14}O_{39}P_3$

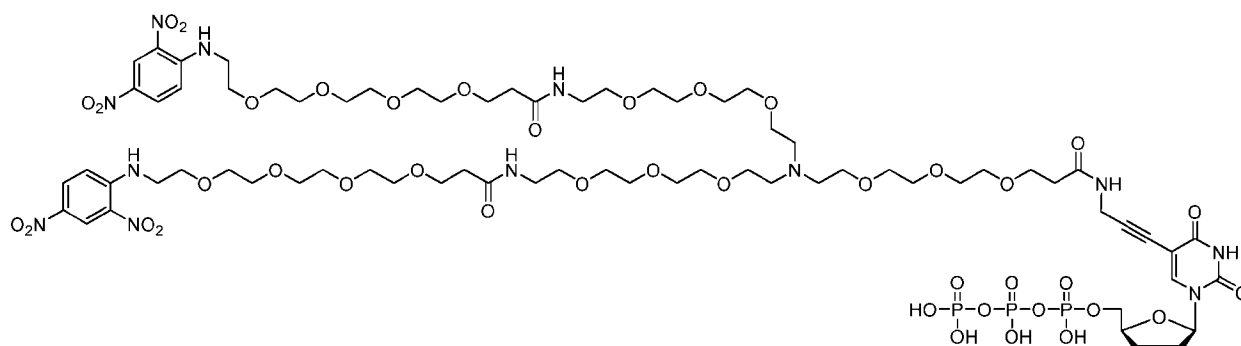
Exact Mass: 1906.6815

Molecular Weight: 1907.7193

ddUTP-diDNP**Divalent DNP**Mol. Formula: $C_{71}H_{115}N_{12}O_{41}P_3$

Exact Mass: 1884.6496

Molecular Weight: 1885.6653

Chemical Formula: $C_{71}H_{115}N_{12}O_{41}P_3$

Exact Mass: 1884.6496

Molecular Weight: 1885.6653

[0093] The purity of the hapten labeled ddNTPs was $\geq 95\%$ (HPLC). The ddNTPs (standard and in accordance with the disclosure) were prepared in a 1 mM concentration in a formulation that included 10 mM Tris-HCl, 1 mM EDTA and had a pH of 7.6 ± 0.2 .

[0094] EML performance was measured using standard EML reagents in an Infinium EX microarray platform, which included a beadchip with 1 μm beads. Testing was done with decreasing antibody concentration in STX and ATX. Standard EML reagents for testing the standard design ddNTPs included equimolar amounts of the standard design ddNTPS. Testing was performed with increasing concentrations of divalent ddNTPs in accordance with the disclosure. Testing conditions used are shown in the table below:

	Bio-11- ddCTP Standard Design	Bio-11- ddGTP Standard Design	DNP-11- ddATP Standard Design	DNP-11- ddUTP Standard Design	ddATP- diDNP Divalent DNP	ddUTP- diDNP Divalent DNP
Standard EML	0.125 nM	0.125 nM	0.125 nM	0.125 nM	0	0
EML + divalent Test Condition 1	0.125 nM	0.125 nM	0	0	0.0625 nM	0.0625 nM
EML + divalent Test Condition 2	0.125 nM	0.125 nM	0	0	0.125 nM	0.125 nM
EML + divalent Test Condition 3	0.125 nM	0.125 nM	0	0	0.250 nM	0.250 nM

[0095] Referring to Figure 10, while there was an observed high signal-to-noise ratio at p95 for the compounds in accordance with the disclosure, a lower signal-to-noise ratio at p50 was observed. It is believed that modification of the testing conditions can be made to reduce the signal-to-noise ratio at p95. Signal to noise ratio of the EML reagents is included in Figure 10 as a control. Referring to Figure 11, an increasing call rate was observed with increasing divalent ddNTP concentration and was independent of antibody concentration. Log R deviation metric was observed to decrease with increasing divalent ddNTP concentration independent of antibody concentration. Low Log R deviation is beneficial for optimal genotyping on the Infinium EX microarray platform.

[0096] The foregoing description is given for clearness of understanding only, and no unnecessary limitations should be understood therefrom, as modifications within the scope of the disclosure may be apparent to those having ordinary skill in the art.

[0097] All patents, patent applications, government publications, government regulations, and literature references cited in this specification are hereby incorporated herein by reference in their entirety. In the case of conflict, the present description, including definitions, will control.

[0098] Throughout the specification, where the compounds, compositions, methods, and/or processes are described as including components, steps, or materials, it is contemplated that the compounds, compositions, methods, and/or processes can also comprise, consist essentially of, or consist of any combination of the recited components or materials, unless described otherwise. Component concentrations can be expressed in terms of weight concentrations, unless specifically indicated otherwise. Combinations of components are contemplated to include homogeneous and/or heterogeneous mixtures, as would be understood by a person of ordinary skill in the art in view of the foregoing disclosure.

What is claimed is:

1. A compound of formula I,

$\text{ARM}_a - \text{Q}_b - \text{M}_c - \text{J} - \text{ddNTP}$,

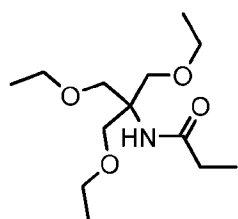
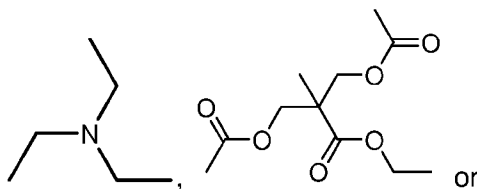
wherein:

a is at least 2,

each ARM comprises a binding unit and a PEO containing linker,

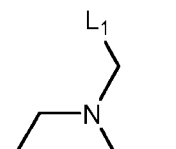
b is 1 to 3,

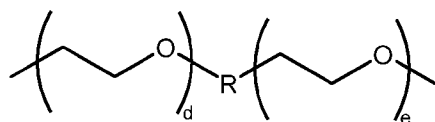
each Q is independently selected from



, each Q attaches to 2 or 3 ARMS

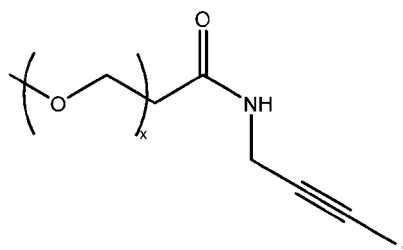
c is 0 or 1,

M is , each L₁ and L₂ are independently selected from



, wherein for each L₁ and L₂, d is independently

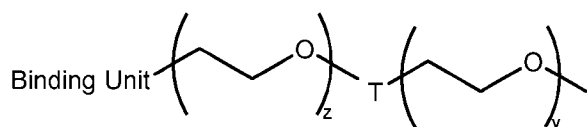
selected from 1 to 10, and e is independently selected from 1 to 10, R is independently selected from a functionality resulting from a coupling reaction between one of the following pairs: amine-imidoester, amine-pentafluorophenyl ester, amine-*N*-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-cyclooctyne, azide-norbornene, or azide-alkyne, and



J is

2. The compound of claim 1, wherein the binding unit is one or more of antibodies, proteins, oligonucleotides, DNA, RNA, drugs, small molecules, dyes, and peptides.

3. The compound of claim 1 or 2, wherein each ARM comprises



selected from 1 to 48 and y is independently selected from 1 to 48 T is independently selected from a functionality resulting from a coupling reaction between one of the following pairs: amine-imidoester, amine-pentafluorophenyl ester, amine-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-cyclooctyne, azide-norbornene, or azide-alkyne

4. A compound of formula IA,



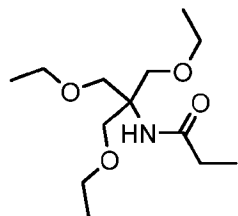
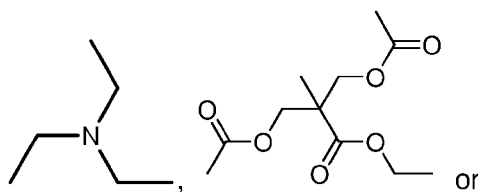
wherein:

a is at least 2,

each ARM comprises a hapten and a PEO containing linker,

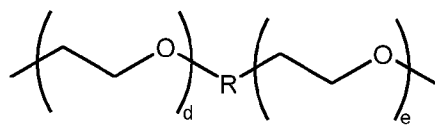
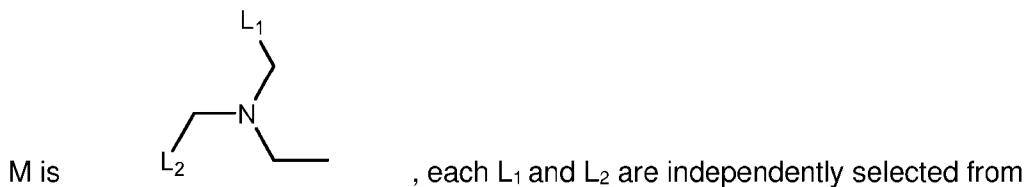
b is 1 to 3,

each Q is independently selected from

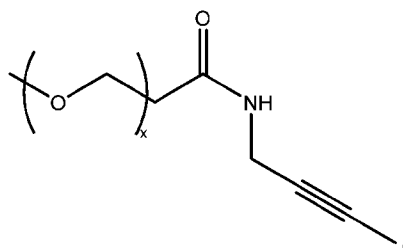


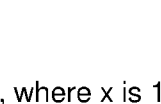
, each Q attaches to 2 or 3 ARMS

c is 0 or 1,



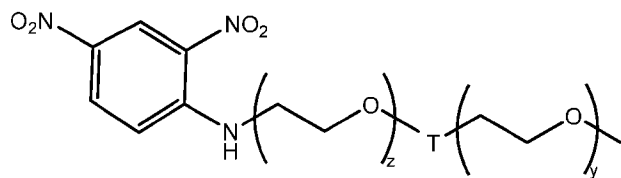
, wherein for each L₁ and L₂, d is independently selected from 1 to 10, and e is independently selected from 1 to 10, R is independently selected from a functionality resulting from a coupling reaction between one of the following pairs: amine-imidoester, amine-pentafluorophenyl ester, amine-*N*-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine, hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-cyclooctyne, azide-norbornene, or azide-alkyne, and



J is , where x is 1 to 48.

5. The compound of claim 4, wherein the hapten for each ARM is independently selected from digoxigenin (DIG), 3-nitrotyrosine, biotin, alkyl chloride, O-GlcNAc, alkyl guanine, Ni-NTA and 2,4-dinitrophenol (DNP).

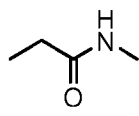
6. The compound of 4 or 5, wherein each ARM comprises

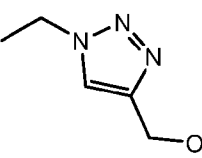


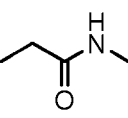
, wherein for each ARM z is independently selected from 1 to 48 and y is independently selected from 1 to 48 T is independently selected from a functionality resulting from a coupling reaction between one of the following pairs: amine-imidoester, amine-pentafluorophenyl ester, amine-*N*-hydroxysuccinimide ester, amine-hydroxymethyl phosphine, carboxyl-carbodiimide, thiol-maleimide, thiol-haloacetyl, thiol-pyridyl disulfide, thiol-thiosulfonate, thiol-vinyl sulfone, aldehyde-hydrazide, aldehyde-alkoxyamine,

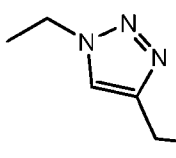
hydroxy-isocyanate, azide-phosphine, transcyclooctene-tetrazine, norbornene-tetrazine, azide-cyclooctyne, azide-norbornene, or azide-alkyne.

7. The compound of claim 3 or 5, wherein z and y are the same in each ARM.
8. The compound of claim 3 or 5, z and y are different in each least two of the ARM.

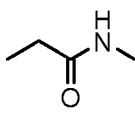
9. The compound of claim 3 or 5, wherein T is  in each of ARM.

10. The compound of claim 3 or 5, wherein T is  in each ARM.

11. The compound of claim 3 or 5, where T is  in at least one ARM and T

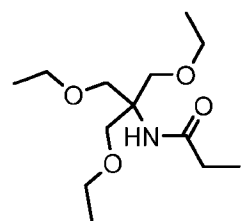
is  in at least one other ARM

12. The compound of claim 3 or 5, wherein T and

13. The compound of any one of the preceding claims, wherein R .

14. The compound of any one of the preceding claims, wherein c is 1.

15. The compound of any one of claims 1 to 13, wherein c is 0 and Q is

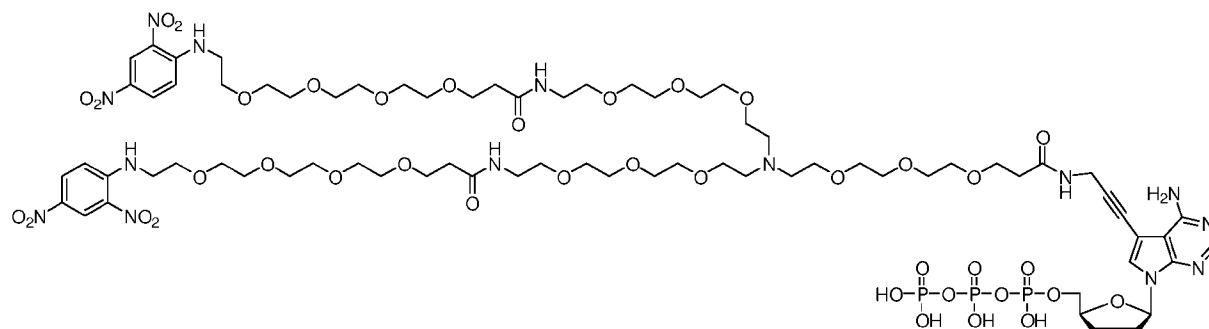


16. The compound of any one of the preceding claims, wherein the ddNTP is ddATP or ddUTP.

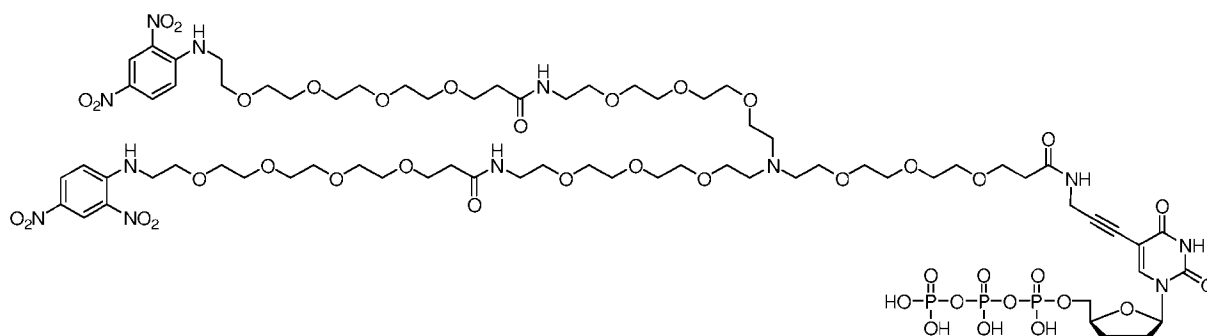
17. The compound of any one of the preceding claims, wherein the ddNTP is ddGTP or ddCTP.

18. The compound of any one of the preceding claims, wherein a is 2 to 9.

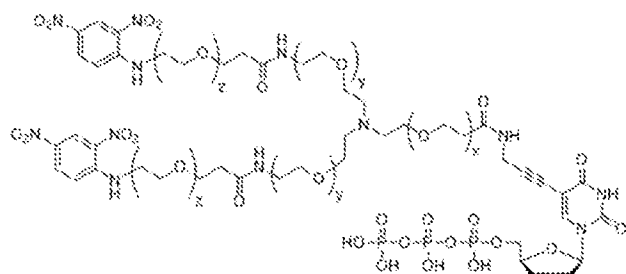
19. The compound of claim 4, wherein the compound is



20. The compound of claim 4, wherein the compound is ,

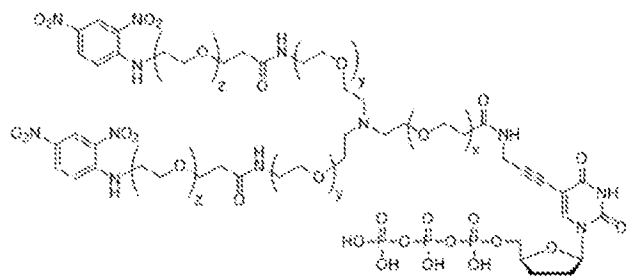


21. The compound of claim 4, wherein the compound is one or more of:



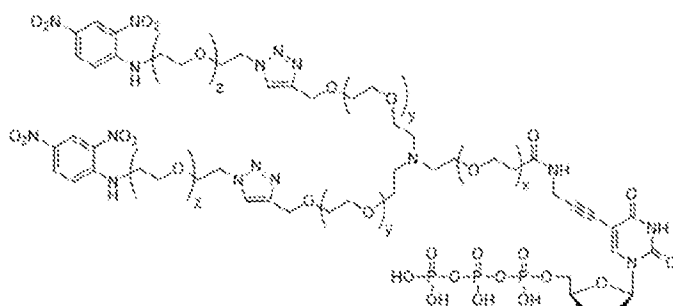
wherein z and y have the same

value on each ARM,



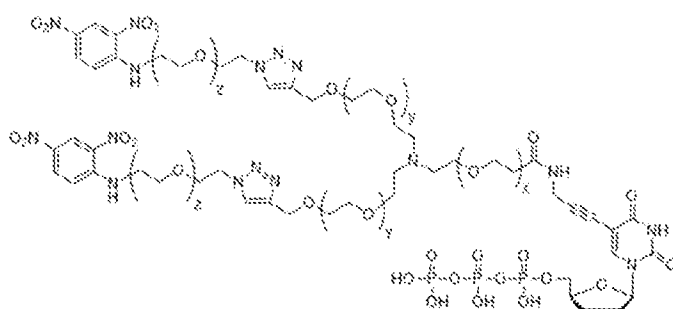
wherein at least one of z and y

has different values on each ARM,



wherein z and y have the same

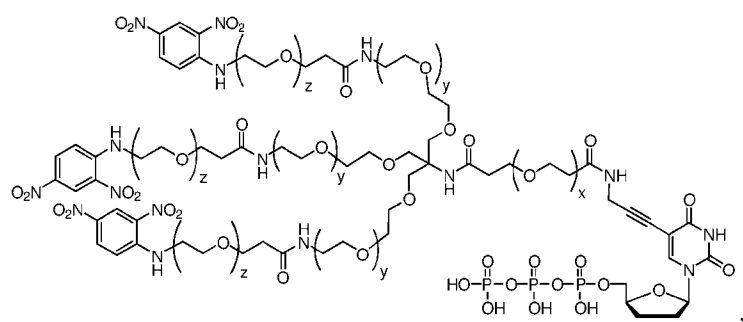
value on each ARM, or

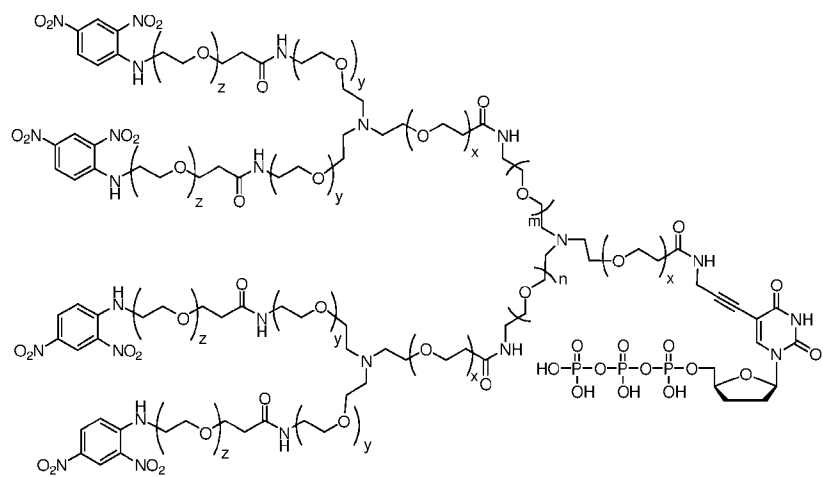


wherein at least one of z and y

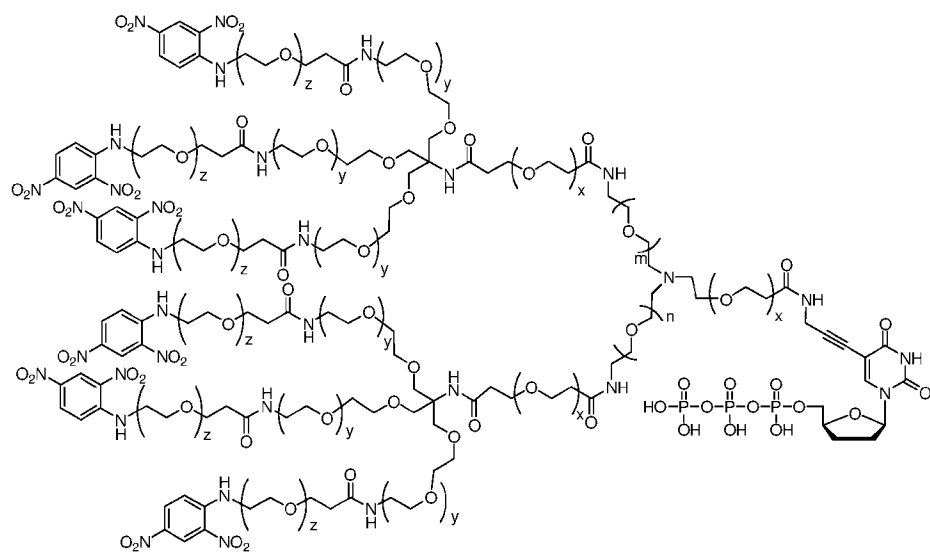
has different values on each ARM.

22. The compound of formula I, selected from

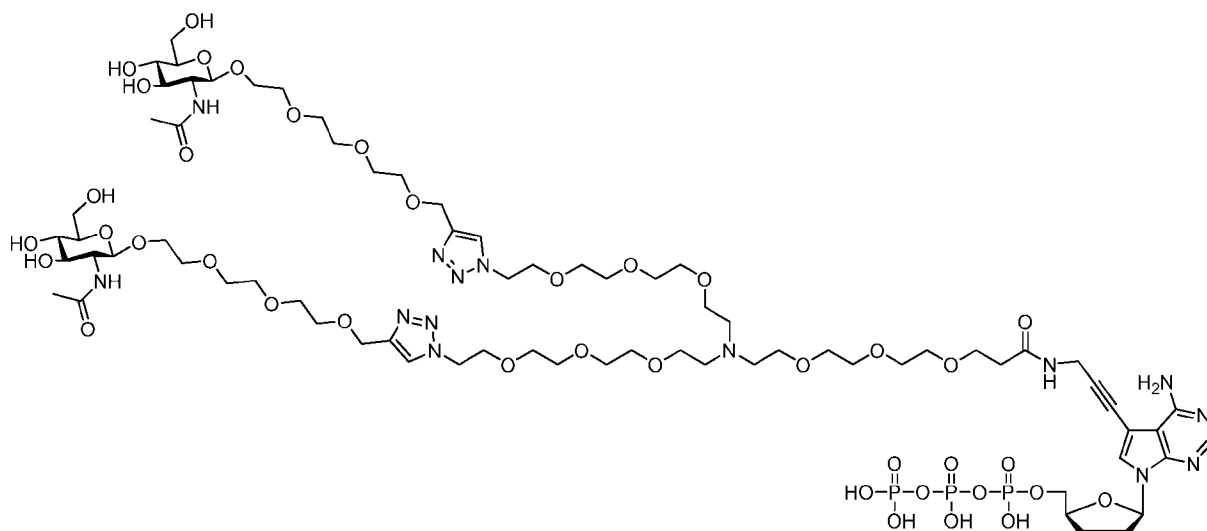




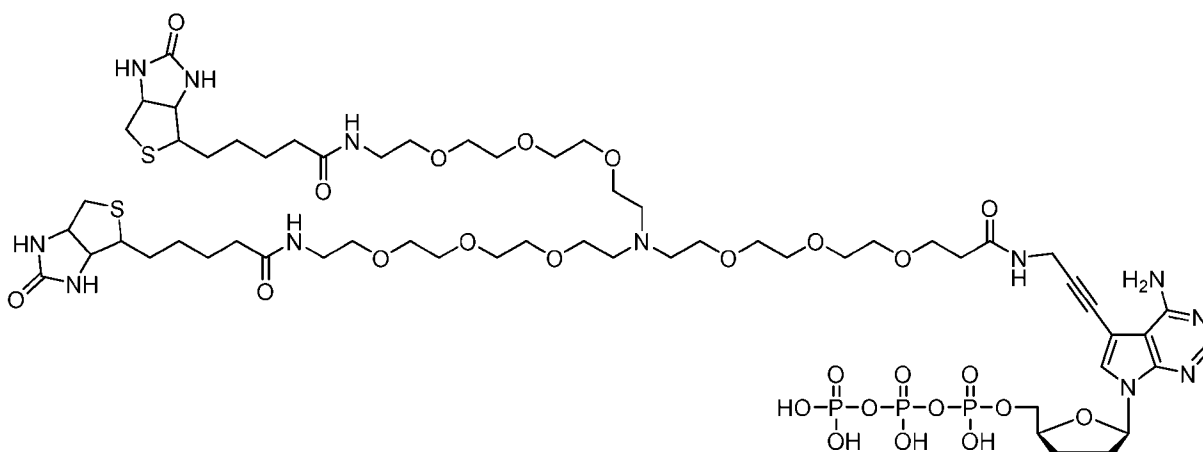
, and



23. The compound of claim 4, wherein the compound is



or



24. A kit, comprising:

a substrate comprising a plurality of capture oligonucleotide probes attached to a surface thereof;

binding unit binding partners, each comprising a label;

at least one compound of formula I according to any one of the preceding claims; and

instructions for use for capturing a target from a sample on the substrate, binding the at least one compound of formula to the captured target as a single base extension of the capture oligonucleotide, and binding the binding unit binding partner to each ARM of the at least one compound of formula I.

25. The kit of claim 24, wherein the substrate is a bead array, and each bead has at least one capture oligonucleotide probe attached to a surface thereof.

26. A kit, comprising:
a bead array comprising a plurality of beads, each bead having at least one capture oligonucleotide probe attached to a surface thereof;
hapten binding partners, each comprising a first fluorescence label;
at least one compound of formula IA according to any one of claims 4 to 23; and
instructions for use for capturing a target oligonucleotide from a sample on the bead array, binding the at least one compound of formula to the captured target oligonucleotide as a single base extension of the capture oligonucleotide, and binding one or more hapten binding partners bind to each ARM of the at least one compound of formula IA.

27. The kit of any one of claims 24 to 26, wherein the at least one compound of formula I or formula IA comprises a first compound of formula I or formula IA comprising ddATP and a second compound of formula I or formula IA comprising ddUTP.

28. The kit of claim 26 or 27 wherein the hapten binding partners are independently selected from anti-DNP antibodies, avidin, anti-DIG antibodies, WGA, SNAP tags, anti-nitrotyrosine antibodies, and His-Tag.

29. A process, comprising:
contacting a substrate with a sample DNA, the substrate comprising a plurality of capture oligonucleotide probes that each bind to a target oligonucleotide in the sample, wherein the target oligonucleotide has a complementary sequence to the capture oligonucleotide probes such that the target oligonucleotide binds to the one or more capture oligonucleotide probes stopping at least one base before a locus of interest thereby providing a substrate having a captured target oligonucleotide;

admixing the substrate having the captured target oligonucleotide with a composition comprising at least a first modified ddNTPs under conditions sufficient to hybridize the first modified ddNTP to one or more of the captured target oligonucleotides as a single base extension of an associated capture oligonucleotide probe, wherein the first is a compound of formula I in accordance with any one of the preceding claims comprising a first binding unit;

contacting the substrate comprising the hybridized first modified ddNTPs with a labeling composition comprising at least a first binding unit binding partner under conditions to bind the first binding unit binding partner to the first binding unit, wherein the first binding unit binding partner comprises a first label, and upon binding of the first binding unit with the first binding unit binding partner, the substrate is labeled; and

analyzing a signal associated with the label.

30. The process of claim 29, wherein the substrate is a bead array, and each bead array comprises a capture oligonucleotide attached to a surface thereof.

31. The process of claim 29 or 30, wherein analyzing the signal associated with the label comprising analyzing the signal to identifying binding pairs.

32. The process of any one of claims 29 to 31, wherein the first binding unit comprises antibodies, proteins, oligonucleotides, DNA, RNA, drugs, small molecules, dyes, and peptides.

33. The process of any one of claims 29 to 32, wherein the first binding unit binding partner comprises one or more of antibodies, proteins, oligonucleotides, DNA, RNA, drugs, small molecules, dyes, and peptides.

34. The process of any one of claims 29 to 33, wherein the labeling composition comprising a plurality of binding unit binding partners each comprising a label, and composition comprising at least the first modified ddNTP comprises a plurality of modified ddNTP comprising a plurality of different binding units, wherein upon contact of the labeling composition with the substrate at least a portion of the plurality of binding unit binding partners bind to one of the plurality of different binding units; and analyzing the signal comprises identifying binding pairs.

35. The process of any one of claims 29 to 34, wherein the label is a fluorescence label.

36. A process for genotyping, comprising:
contacting a bead array with a sample DNA, each bead of the bead array comprising one or more capture oligonucleotide probes that each bind to a target complementary sequence in the sample, wherein the target complementary sequence binds to the one or more capture oligonucleotide probes stopping at least one base before a locus of interest thereby providing a bead array having captured target complementary sequence;

admixing the bead array having the captured target complementary sequences with a composition comprising at least a first hapten modified ddNTPs and a second hapten modified ddNTP under conditions sufficient to hybridize the first and second hapten modified ddNTP to respective ones of the captured target complementary sequences as a single base extension of an associated capture oligonucleotide probe, wherein the first and second haptens are different and at least the first hapten modified ddNTP is a compound of formula I in accordance with any one of the preceding claims;

contacting the bead array comprising the hybridized first and second hapten modified ddNTPs with a labeling composition comprising a first hapten binding partner and a second

hapten binding partner under conditions to bind the first hapten binding partner to the first hapten and the second hapten binding partner to the second hapten, wherein the first hapten binding partner comprises a first fluorescence label, and the second hapten binding partner comprises a second fluorescence label and upon binding of the first and/or second hapten with the respective first and/or second hapten binding partner, the bead array is fluorescently labeled; and

analyzing a fluorescence signal of the bead array after binding of the first hapten binding partner and the second hapten binding partner.

37. The process of claim 36, wherein the first hapten modified ddNTP comprises a first hapten modified ddATP and a first hapten modified ddUTP and the second hapten modified ddNTP comprises a second hapten modified ddGTP and a second hapten modified ddCTP.

38. The process of claim 36 or 37, comprising repeatedly contacting the bead array with the labeling composition.

39. The process of claim 36, wherein the bead array is fluorescently labeled with fewer repetitions of contacting the bead array with the labeling composition as compared to contacting the bead array with a contacting step comprising a hapten modified ddNPT having only a single ARM and at least maintains the same fluorescence signal intensity.

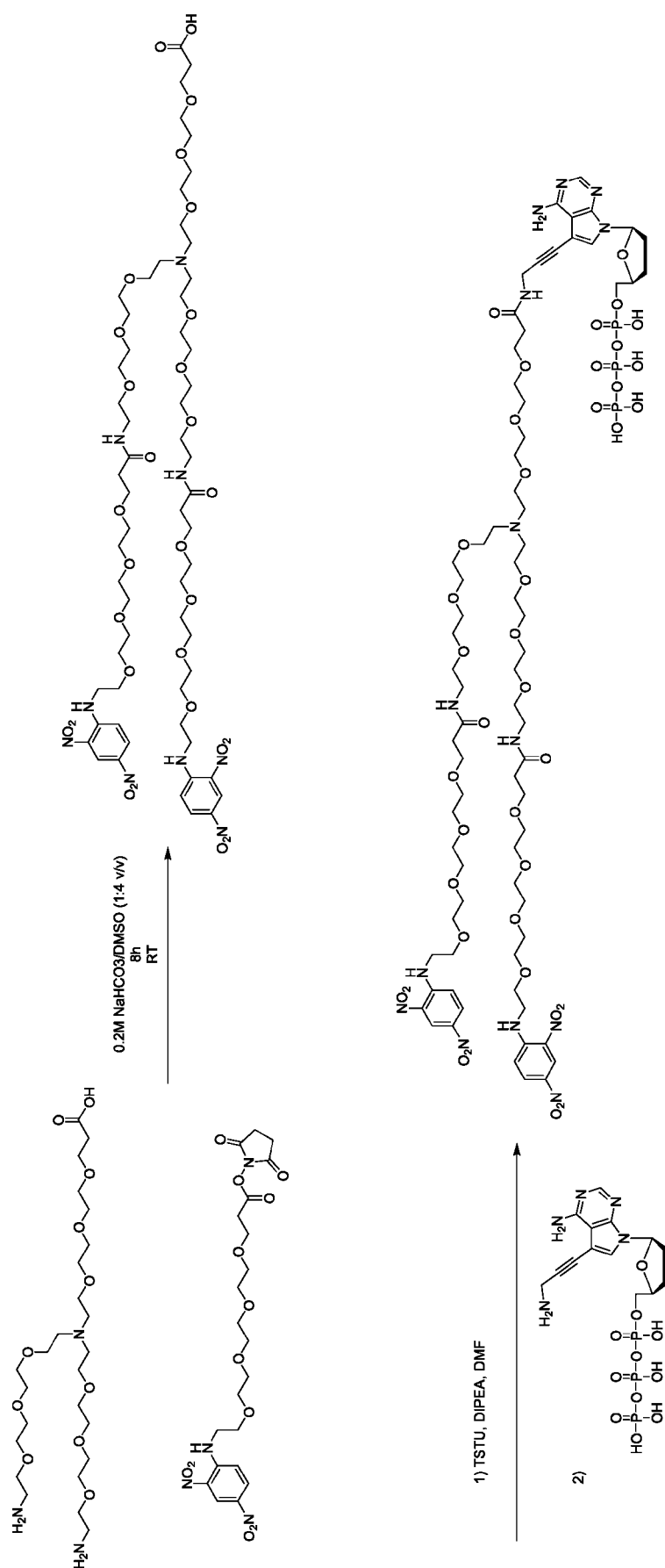


Figure 1

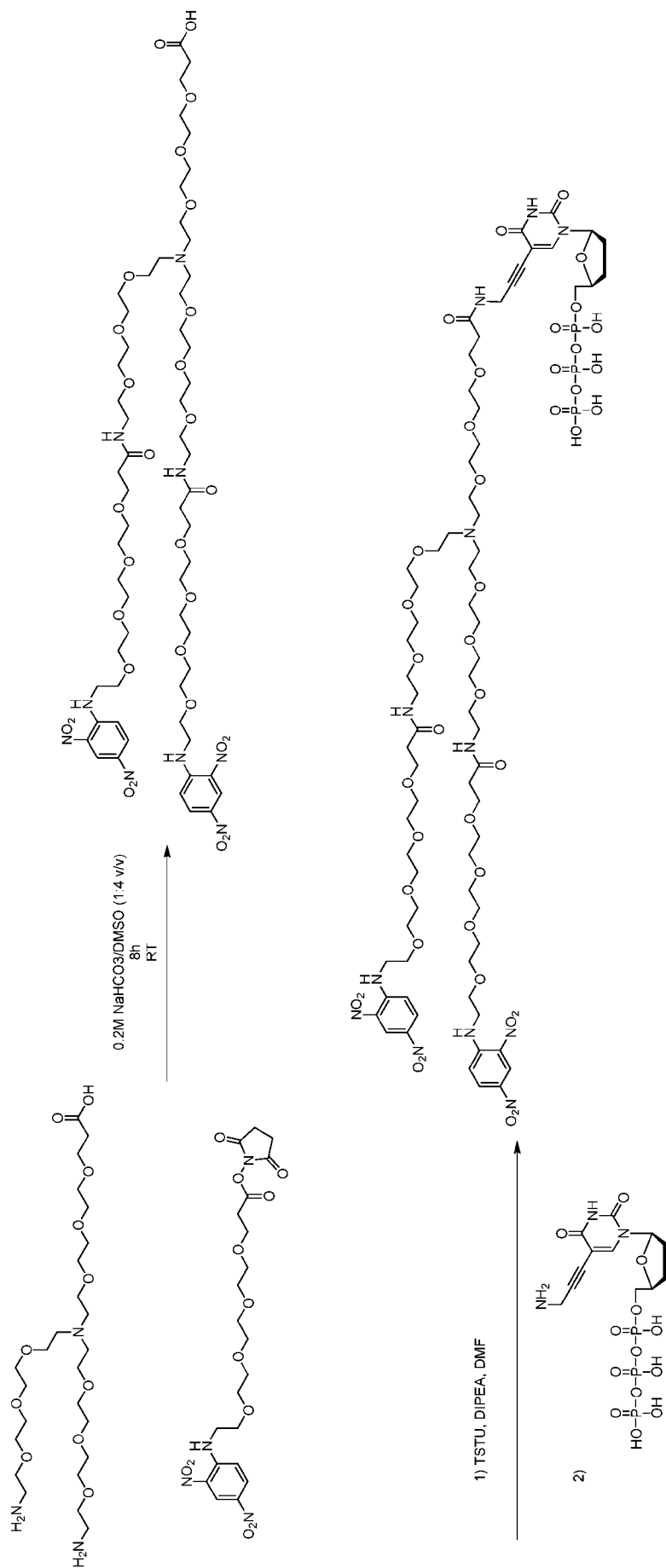


Figure 2

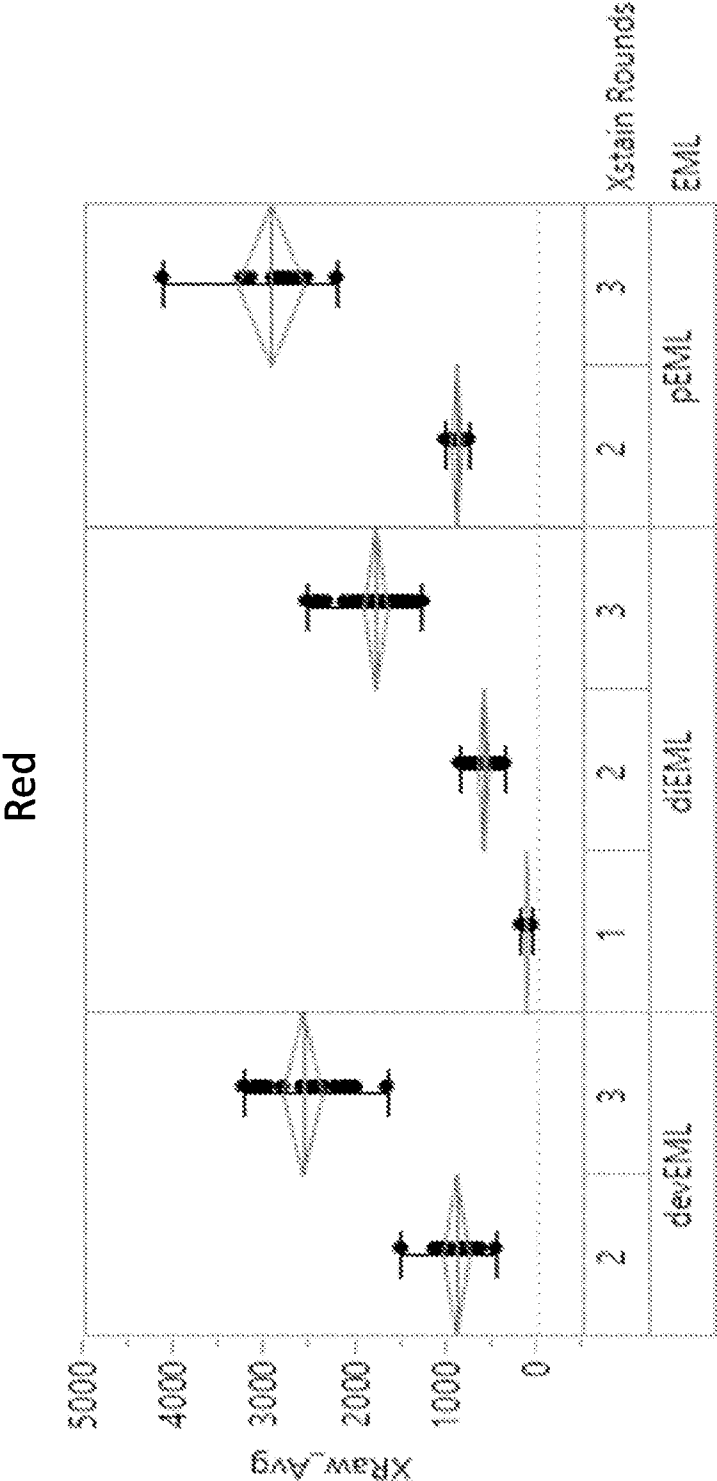


Figure 3

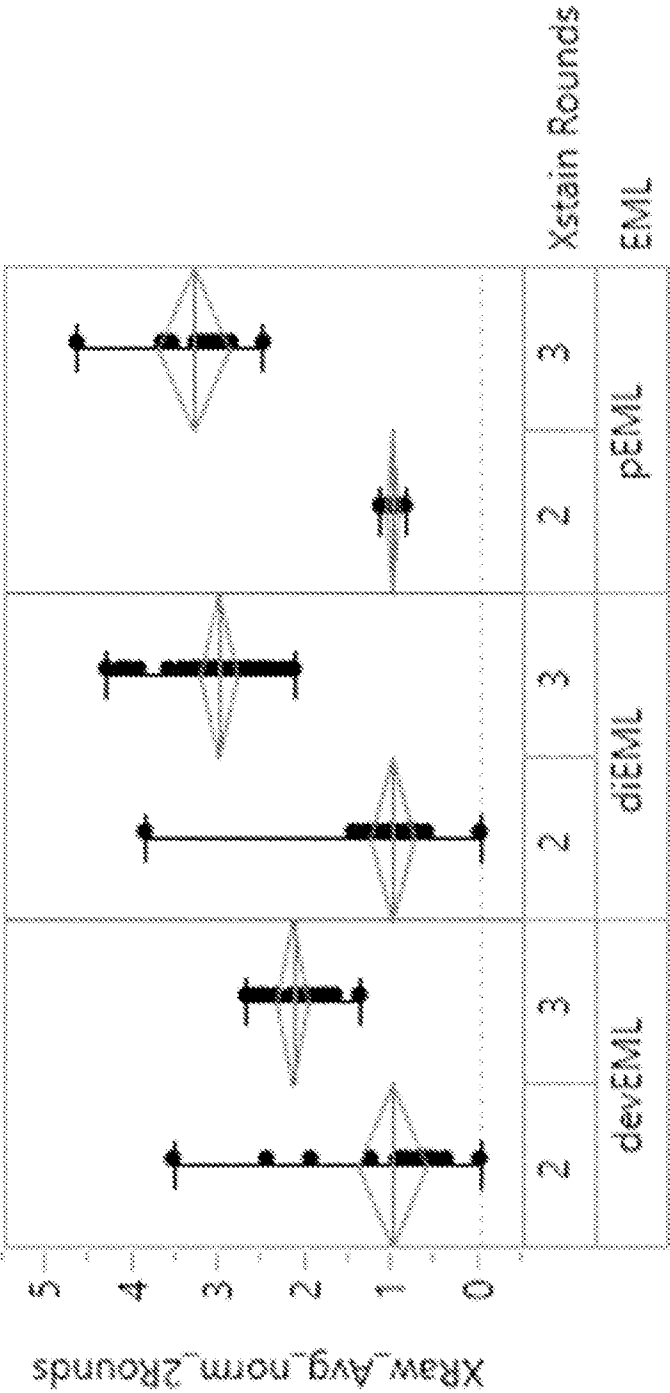


Figure 4

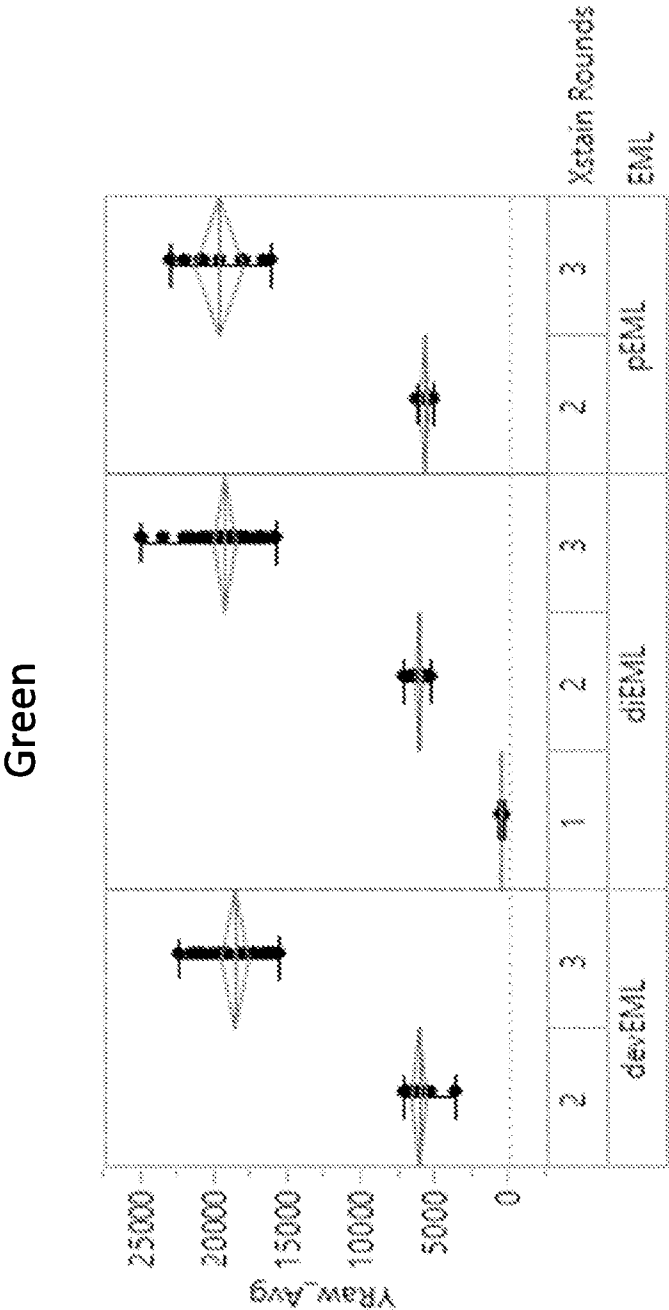


Figure 5

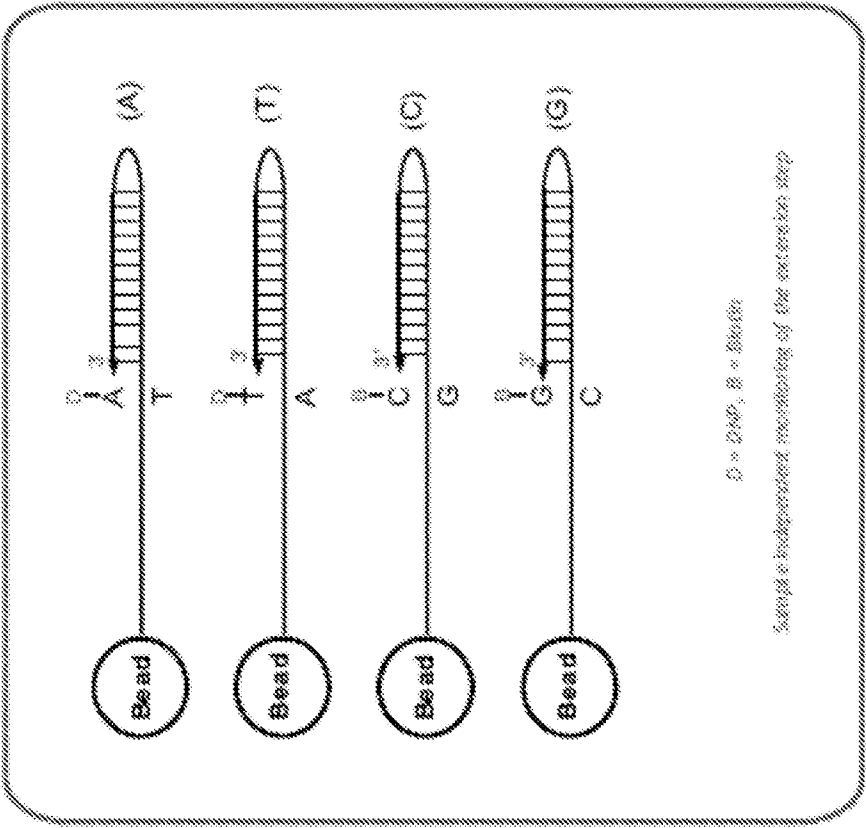


Figure 6

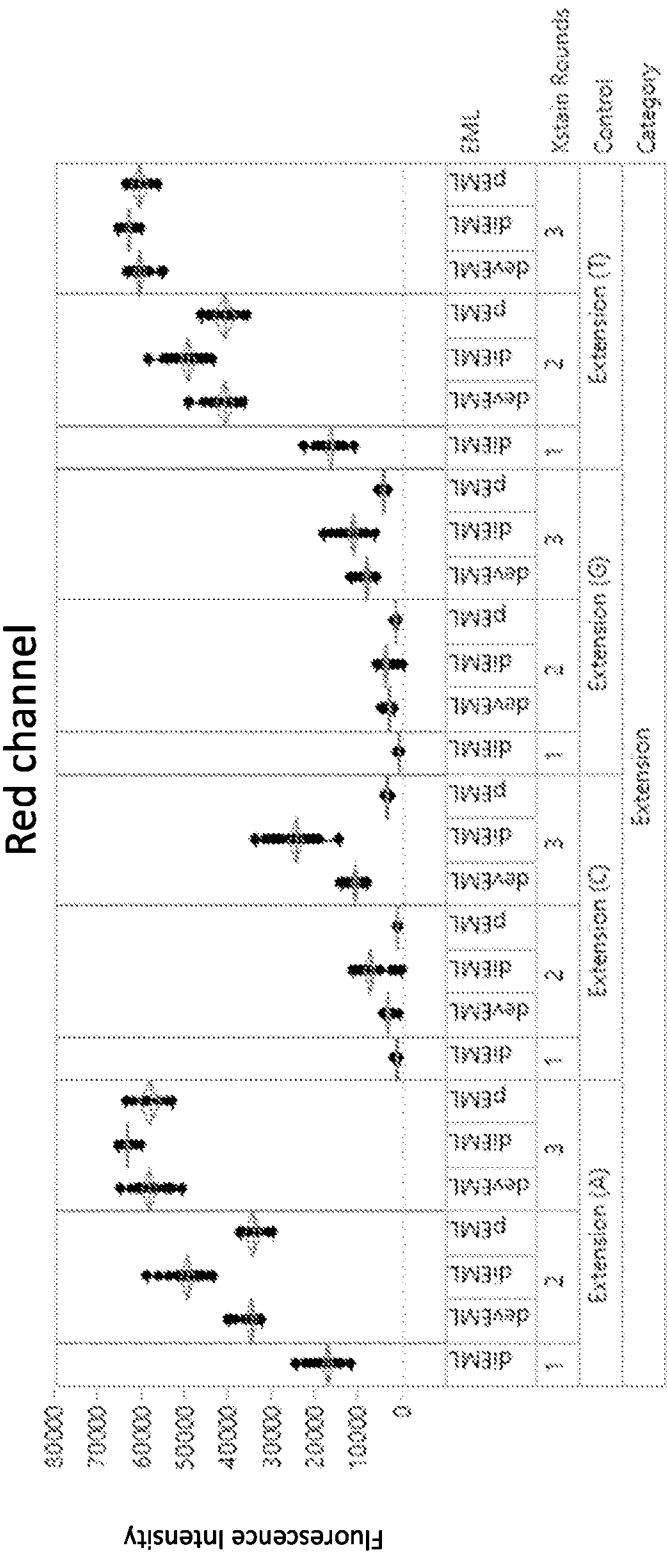


Figure 7A

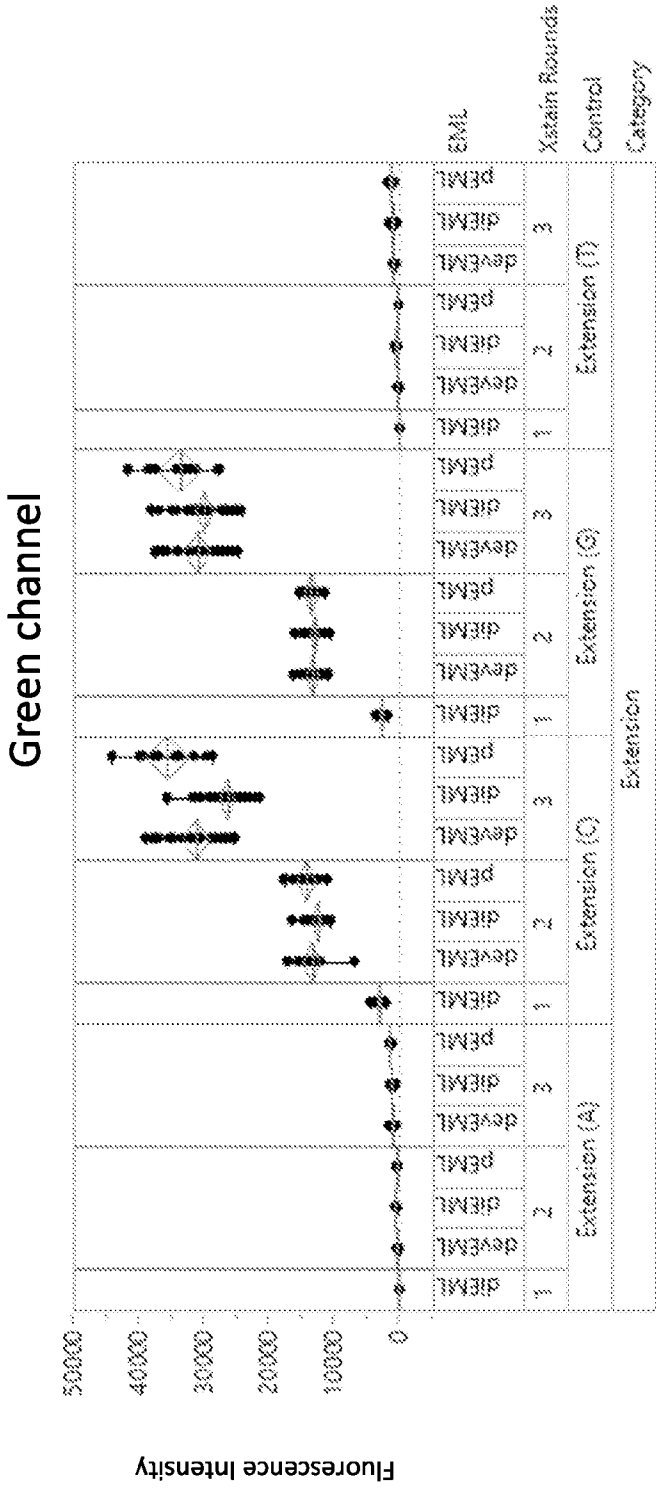


Figure 7B

RED

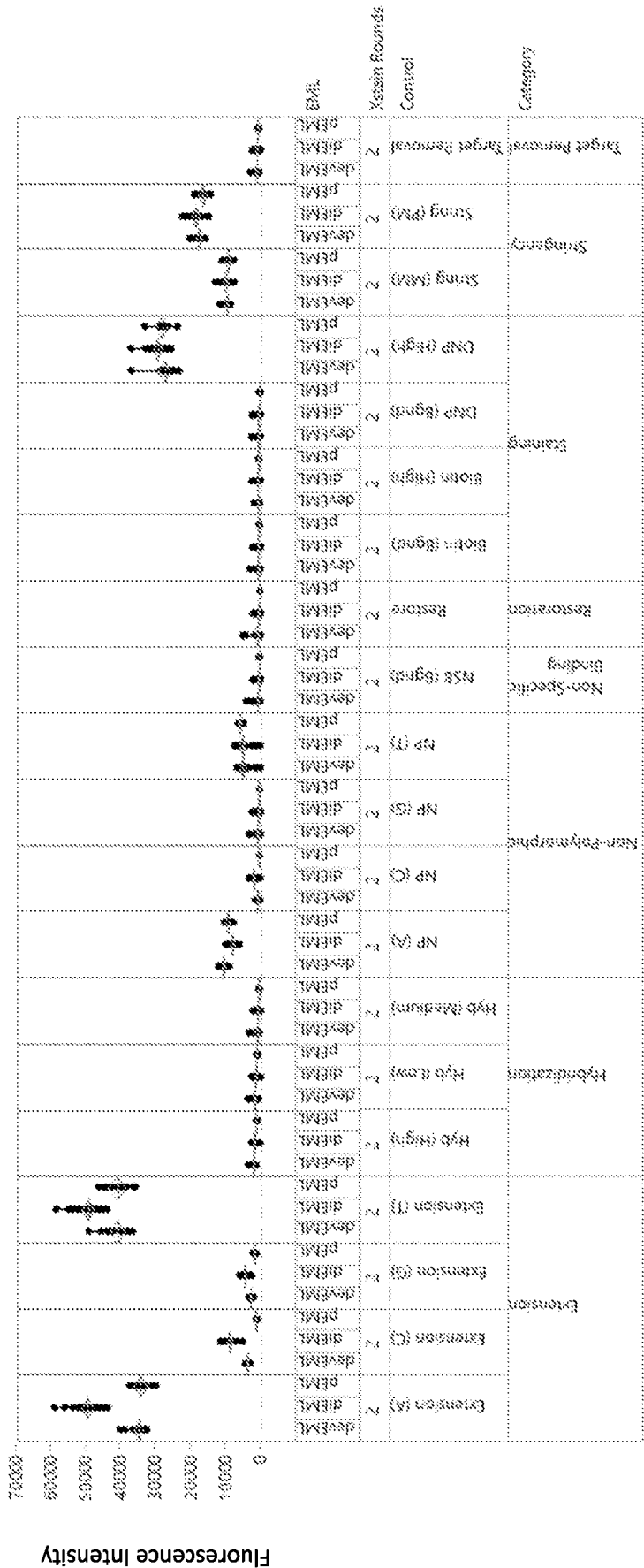


Figure 8

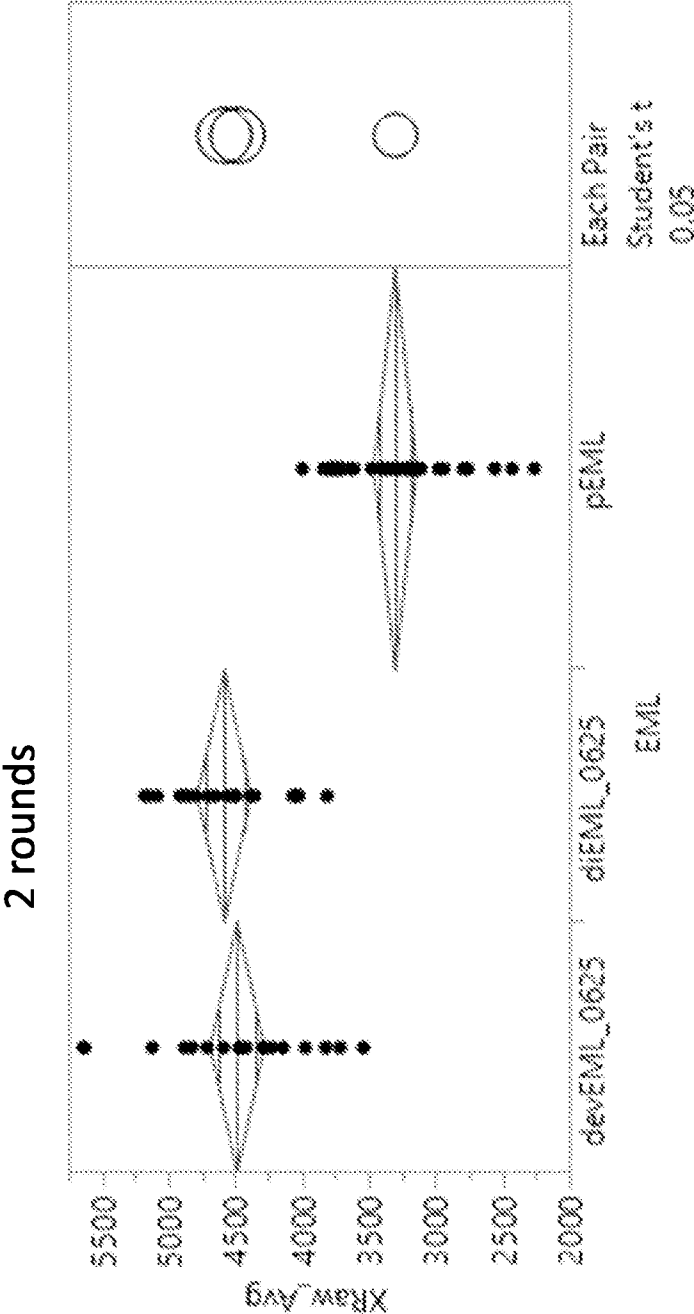


Figure 9A

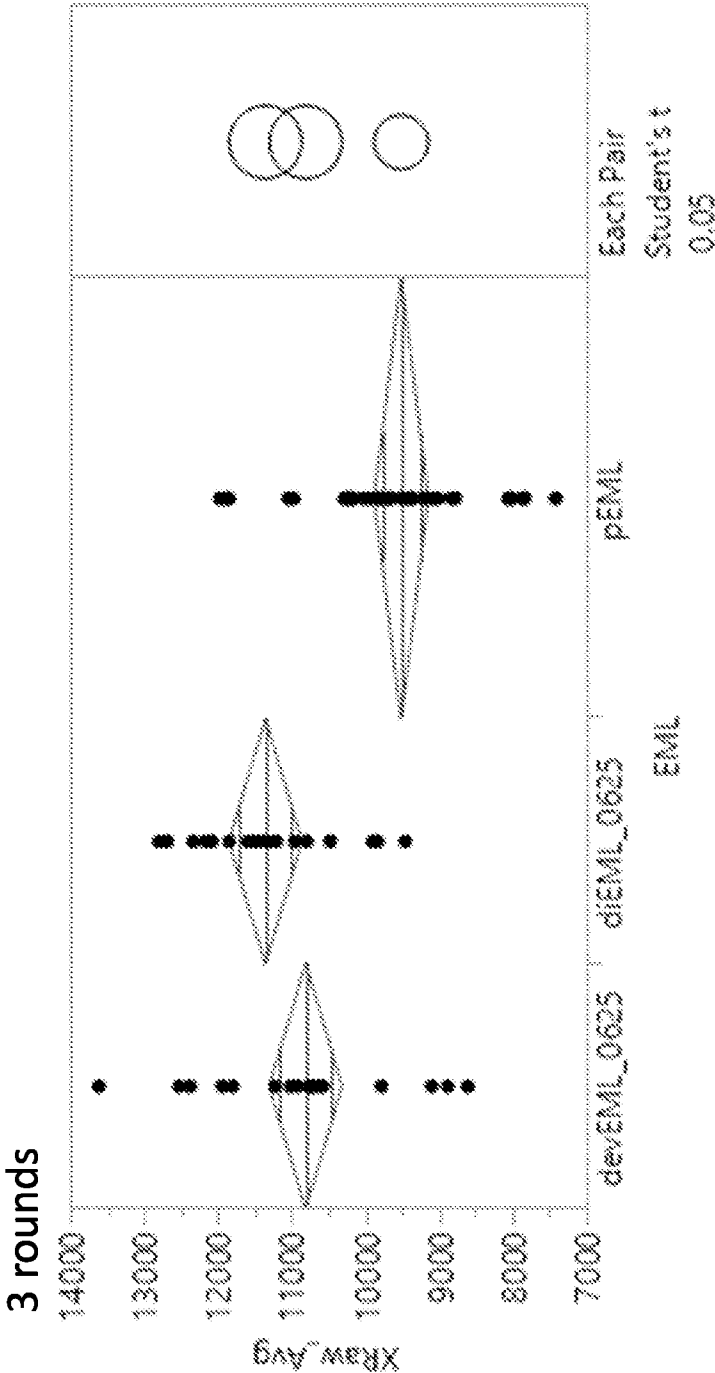


Figure 9B

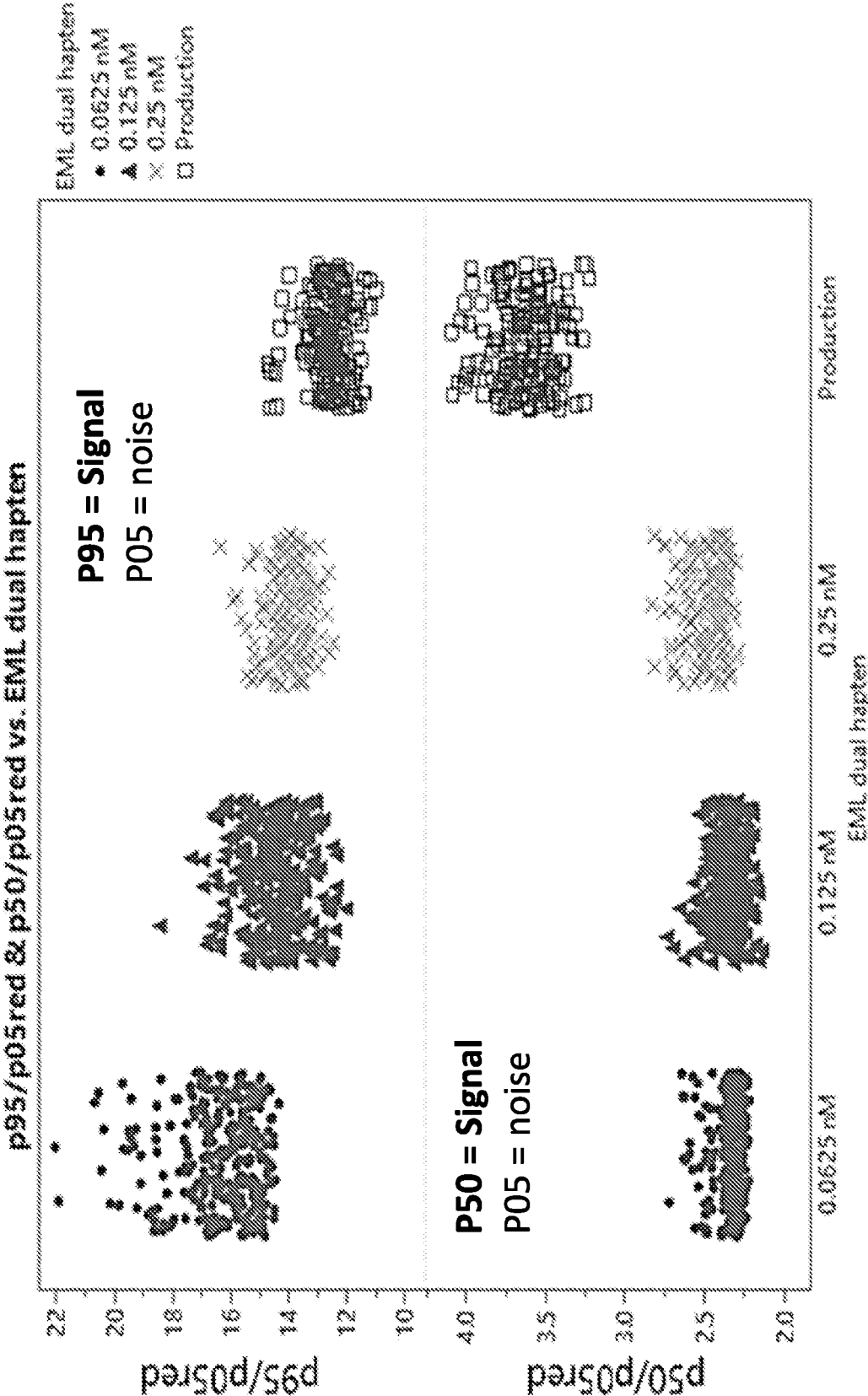


Figure 10

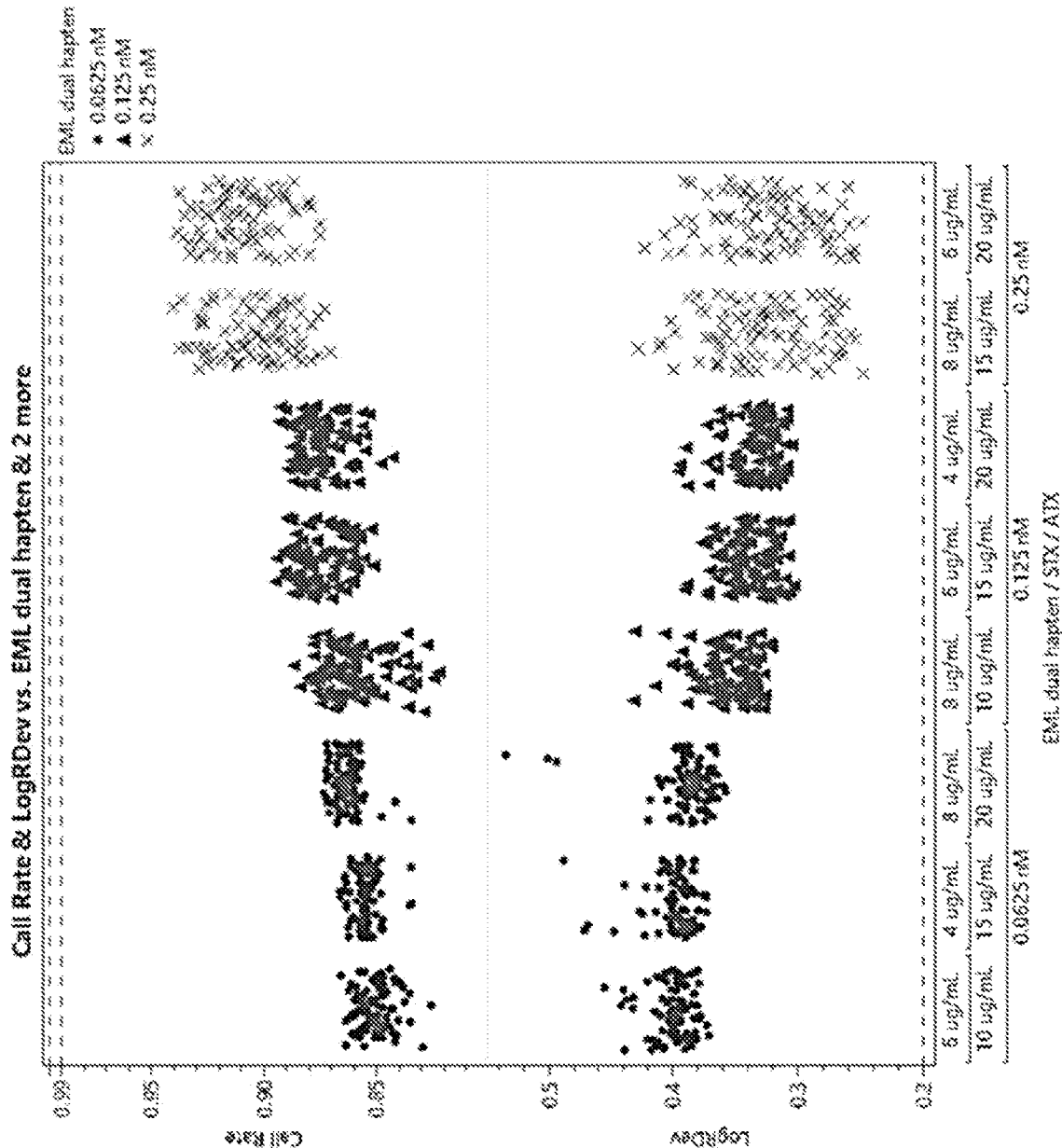


Figure 11

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2024/032844

A. CLASSIFICATION OF SUBJECT MATTER INV. C07H19/10 C07H19/20 C12Q1/68 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) C07H 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		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KUMAR SHIV ET AL: "Synthesis and Energy Transfer Efficiency of FRET Terminators Derived from Different Linkers", NUCLEOSIDES, NUCLEOTIDES & NUCLEIC ACIDS, vol. 22, no. 5-8, 1 October 2003 (2003-10-01), pages 1595-1598, XP093203929, US ISSN: 1525-7770, DOI: 10.1081/NCN-120023042	1-18,22, 24-39
A	Scheme 1 figure 1 ----- - / - -	19-21,23
☒ 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 11 September 2024		Date of mailing of the international search report 01/10/2024
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 Fanni, Stefano

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2024/032844

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RAO T S ET AL: "Four color FRET dye nucleotide terminators for DNA sequencing", NUCLEOSIDES, NUCLEOTIDES & NUCLEIC ACIDS, TAYLOR & FRANCIS, US, vol. 20, no. 4-7, 1 January 2001 (2001-01-01), pages 673-676, XP002244469, ISSN: 1525-7770, DOI: 10.1081/NCN-100002348	1-18,22, 24-39
A	Scheme II -----	19-21,23
X	WO 2009/054922 A1 (UNIV COLUMBIA [US]; JU JINGYUE [US] ET AL.) 30 April 2009 (2009-04-30)	1-18,22, 24-39
A	claim 1 figures 1-2 -----	19-21,23
X	RAO T S ET AL: "Synthesis of novel piperidinyI linker based energy transfer terminators and their potential use in DNA sequencing", NUCLEOSIDES, NUCLEOTIDES & NUCLEIC ACIDS, TAYLOR & FRANCIS, US, vol. 24, no. 5-7, 1 January 2005 (2005-01-01), pages 801-804, XP009090421, ISSN: 1525-7770, DOI: 10.1081/NCN-200060586	1-18,22, 24-39
A	Scheme 1 -----	19-21,23
X	WO 2021/007458 A1 (UNIV COLUMBIA [US]; JU JINGYUE [US] ET AL.) 14 January 2021 (2021-01-14)	1-18,22, 24-39
A	claims 1,70 figures 3-4,15-16,30-31,34-39,50-55,66-67,79-90,95 -99,105- -----	19-21,23
X	NAMPALLI SATYAM ET AL: "FLUORESCENCE RESONANCE ENERGY TRANSFER DYE NUCLEOTIDE TERMINATORS: A NEW SYNTHETIC APPROACH FOR HIGH-THROUGHOUT DNA SEQUENCING", NUCLEOSIDES, NUCLEOTIDES & NUCLEIC ACIDS, vol. 20, no. 4-7, 31 March 2001 (2001-03-31), pages 361-367, XP093128253, US ISSN: 1525-7770, DOI: 10.1081/NCN-100002307	1-18,22, 24-39
A	Scheme 2 -----	19-21,23

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2024/032844

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2009054922 A1	30-04-2009	DK 2725107 T3 EP 2209911 A1 EP 2725107 A1 EP 3431615 A2 US 2011039259 A1 US 2016024574 A1 US 2019390271 A1 US 2022290227 A1 WO 2009054922 A1	02-01-2019 28-07-2010 30-04-2014 23-01-2019 17-02-2011 28-01-2016 26-12-2019 15-09-2022 30-04-2009
WO 2021007458 A1	14-01-2021	CN 114981283 A EP 3997101 A1 US 2022372061 A1 WO 2021007458 A1	30-08-2022 18-05-2022 24-11-2022 14-01-2021