



(51) International Patent Classification:

Not classified

(21) International Application Number:

PCT/SG2022/050672

(22) International Filing Date:

20 September 2022 (20.09.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

10202110408R 21 September 2021 (21.09.2021) SG

(71) Applicant: **NANYANG TECHNOLOGICAL UNIVERSITY** [SG/SG]; 50 Nanyang Avenue, Singapore 639798 (SG).

(72) Inventors: **PU, Kanyi**; c/o Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798 (SG). **LIEW, Si Si**; c/o Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798 (SG). **ZENG, Ziling**; c/o Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798 (SG).

(74) Agent: **KINNAIRD, James Welsh**; Marks & Clerk Singapore LLP, Tanjong Pagar Post Office, P.O. Box 636, Singapore 910816 (SG).

(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, 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).

(54) Title: ACTIVATABLE MOLECULAR PROBES FOR IN VIVO VIRUS IMAGING AND DETECTION

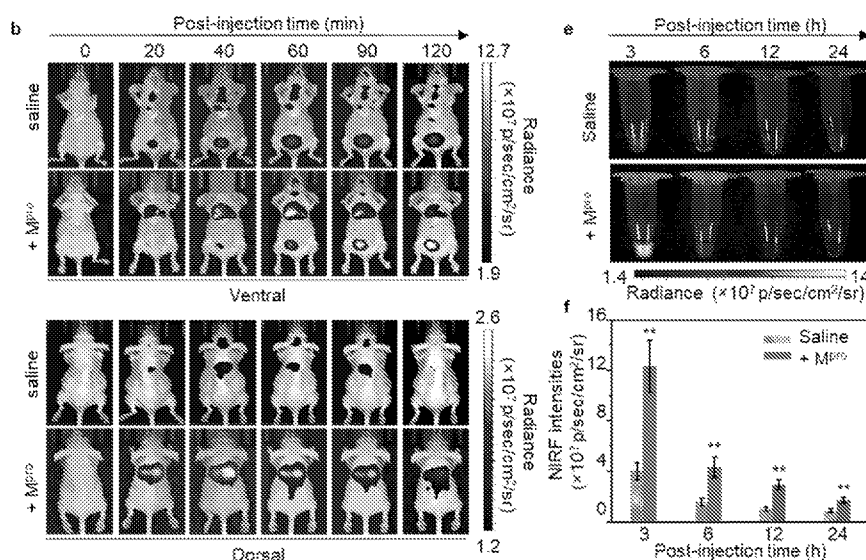


FIG. 5

(57) Abstract: Disclosed herein is a compound of formula I: A-[B]_n-C I where: A represents a peptide group that targets a viral protease, B, when present, represents a self-immolative linking group, C represents a fluorophore or chemilumiphore group covalently linked to the rest of the molecule by an oxygen atom or a NH group, and n represents 0 or 1, or a pharmaceutically acceptable salt or solvate thereof. Also disclosed herein are uses of these compounds and methods of manufacture.



Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *without international search report and to be republished
upon receipt of that report (Rule 48.2(g))*

Activatable Molecular Probes for In Vivo Virus Imaging and Detection

Field of Invention

- 5 The current invention relates to activatable molecular probes that can be used for *in vivo* imaging and detection of viral infections.

Background

- 10 The listing or discussion of a prior-published document in this specification should not necessarily be taken as an acknowledgement that the document is part of the state of the art or is common general knowledge.

Real-time polymerase chain reaction (RT-PCR) is the gold standard for identifying active
15 infections via detecting viral ribonucleic acids with high sensitivity and selectivity. Antigen-rapid tests (ARTs) have also been developed as point-of-care tests to detect ongoing viral infections. However, the high risks of false-negative results render them complementary to RT-PCR. Serological tests are useful for staging past infections by quantifying host IgM/IgG antibodies levels against viral proteins. However, these diagnostic methods are static, unable to
20 distinguish between viable/non-viable virus or reflect viral replication activity. Thereby, multiple patient sample collections are required to monitor the clinical phases of infection. Despite the recent growing body of available and in-development diagnostics, approaches for real-time, dynamic imaging of viral infection remain inchoate.

- 25 Molecular optical probes that activate their signals in the presence of biomarkers are powerful tools for real-time non-invasive imaging of diseases *in vivo*. Although fluorescence molecular probes have been widely applied to detect proteases for diagnosis of kidney/liver injury, cancer, inflammation and neurodegenerative diseases, they have been less exploited for virus detection. A few examples include detection of dengue virus, human immunodeficiency virus
30 (HIV) and SARS-CoV-2 virus. However, they are limited to *in vitro* studies due to the shallow tissue-penetrating wavelengths of these probes. Considering the potential advantages in direct monitoring of viral replication activity, clinicopathologic staging of infection and high throughput inhibitor screening, fluorescence molecular probes capable of real-time dynamic *in vivo* imaging of viral proteases are highly desired.

35

Therefore, there exists a need to discover fluorescence molecular probes capable of real-time dynamic *in vivo* imaging of viral proteases.

Summary of Invention

It has been surprisingly found that specific fluorescent molecular probes can be used to image viral protease activity *in vivo*.

Aspects and embodiments of the invention will now be discussed by reference to the following numbered clauses.

1. A compound of formula I:



wherein:

A represents a peptide group that targets a viral protease;

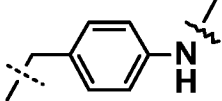
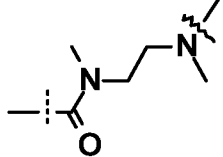
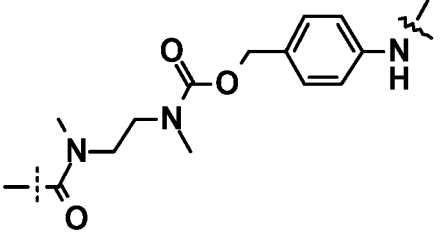
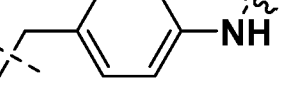
B, when present, represents a self-immolative linking group;

C represents a fluorophore or chemilumiphore group covalently linked to the rest of the molecule by an oxygen atom or a NH group; and

n represents 0 or 1,

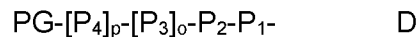
or a pharmaceutically acceptable salt or solvate thereof.

2. The compound according to Clause 1, or a pharmaceutically acceptable salt or solvate thereof, wherein B, when present is selected from:

 ; or more particularly	 ;
 ; and	 ,

where the wavy line represents the point of attachment to A and the dotted line represents the point of attachment to C.

- 5 3. The compound according to Clause 1 or Clause 2, or a pharmaceutically acceptable salt or solvate thereof, wherein A represents a peptide sequence according to fragment D that targets a flavivirus protease:



10 where:

P₁ to P₄, when present, are amino acids that together form an oligo peptide;

o is 0 or 1;

p is 0 or 1; and

PG is a protecting group.

15

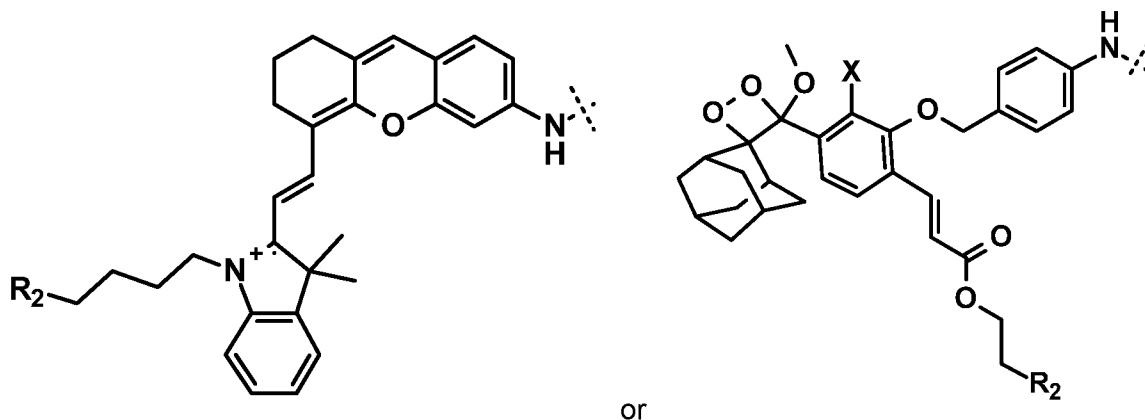
4. The compound according to Clause 3, or a pharmaceutically acceptable salt or solvate thereof, wherein PG is selected from a benzoyl protecting group, an acetyl protecting group or a *tert*-butyloxycarbonyl protecting group.

- 20 5. The compound according to Clause 3 or Clause 4, or a pharmaceutically acceptable salt or solvate thereof, wherein fragment D is selected from the list of:

- (a) Bz-Nle-Lys-Arg-Arg-;
- (b) Bz-Nle-Lys-Arg-Ala-;
- (c) Bz-Nle-Lys-Thr-Arg-;
- 25 (d) Bz-Nle-Thr-Arg-Arg-;
- (e) Bz-Nle-Ala-Arg-Arg-;
- (f) Bz-Nle-Lys-Arg-Phe-;
- (g) Bz-Nle-Lys-Phe-Arg-;
- (h) Bz-Nle-Phe-Arg-Arg-;
- 30 (i) Bz-Thr-Lys-Arg-Arg-;
- (j) Bz-Thr-Thr-Arg-Arg-;
- (k) Bz-Phe-Lys-Arg-Arg-;
- (l) Bz-Ala-Lys-Arg-Arg-;
- (m) Bz-Nle-Lys-Lys-Arg-;
- 35 (n) Bz-Lys-Arg-Arg-;
- (o) Bz-Arg-Arg-;
- (p) Ac-Nle-Lys-Arg-Arg-;

- (q) Ac-Lys-Arg-Arg-;
 (r) Boc-Gly-Arg-Arg-; and
 (s) Boc-Lys-Arg-Arg-.

- 5 6. The compound according to any one of Clauses 3 to 5, or a pharmaceutically acceptable salt or solvate thereof, wherein C is selected from:

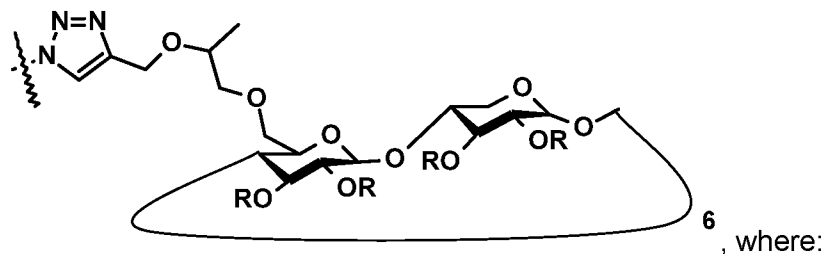
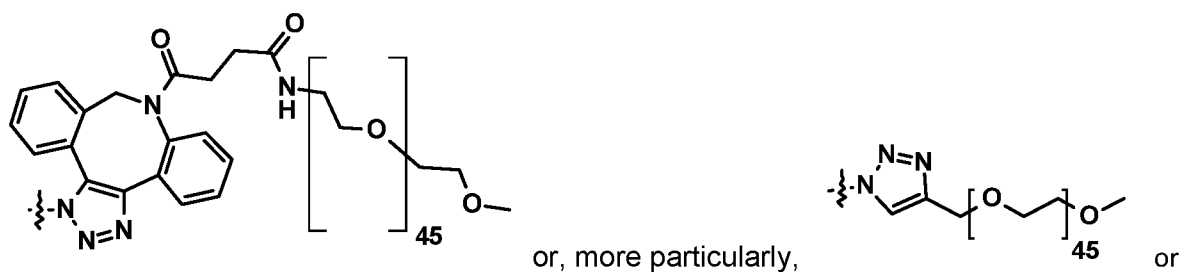


10 where:

the dotted line represents the point of attachment to the rest of the molecule;

X represents H or halo (e.g. Cl); and

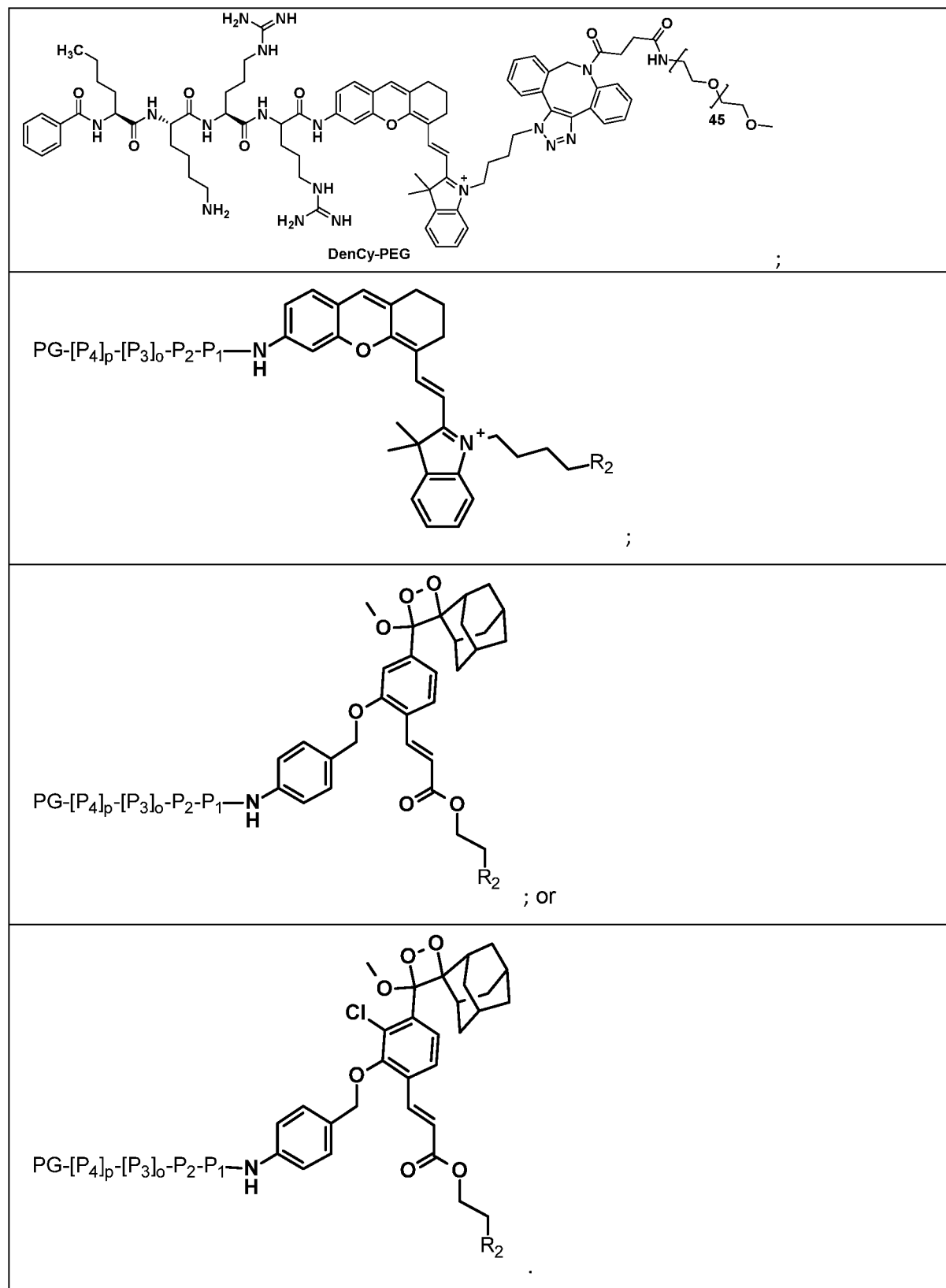
R₂ is selected from:



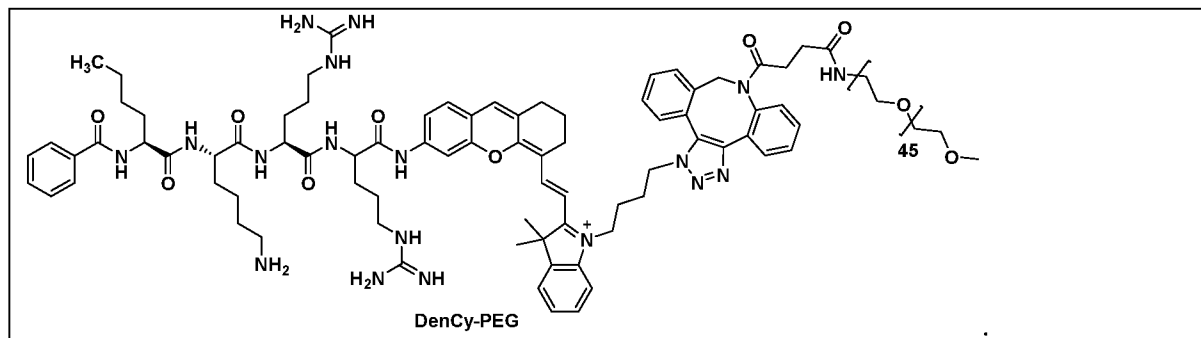
the wiggly line represents the point of attachment to the rest of the molecule; and

R represents H or 2-hydroxypropyl.

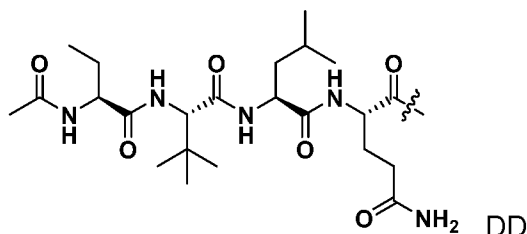
7. The compound according to Clause 6, or a pharmaceutically acceptable salt or solvate thereof, wherein the compound of formula I is selected from:



8. The compound according to Clause 7, or a pharmaceutically acceptable salt or solvate thereof, wherein the compound of formula I is:

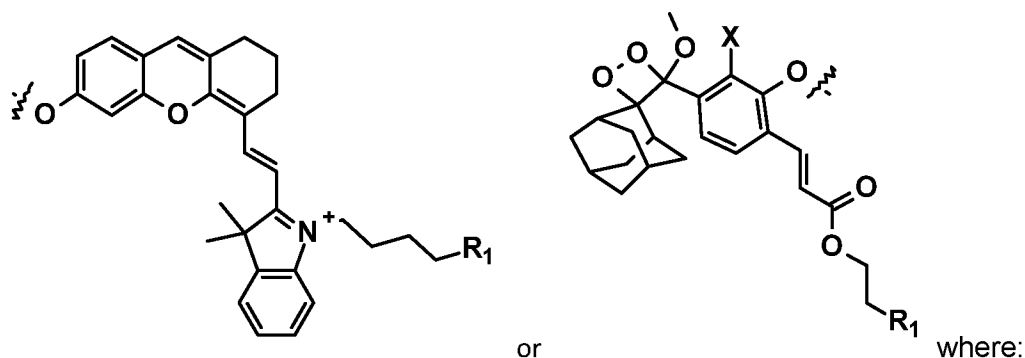


9. The compound according to Clause 1 or Clause 2, or a pharmaceutically acceptable salt or solvate thereof, wherein A represents a peptide sequence according to fragment DD that targets SARS-CoV-2 main protease:

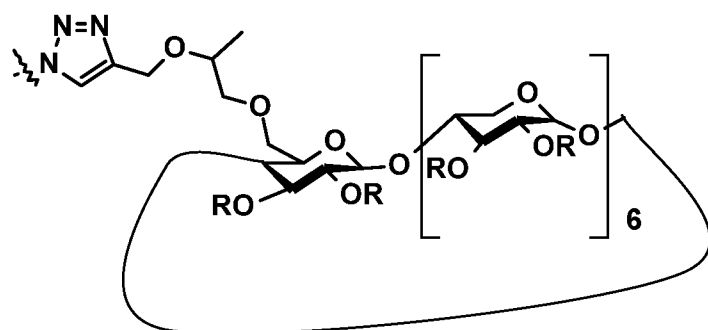


10 where the wavy line represents the point of attachment to the rest of the molecule.

10. The compound according to Clause 9, or a pharmaceutically acceptable salt or solvate thereof, wherein C is selected from:



the wiggly line represents the point of attachment to the rest of the molecule;
X represents H or halo (e.g. Cl); and
R₁ represents:

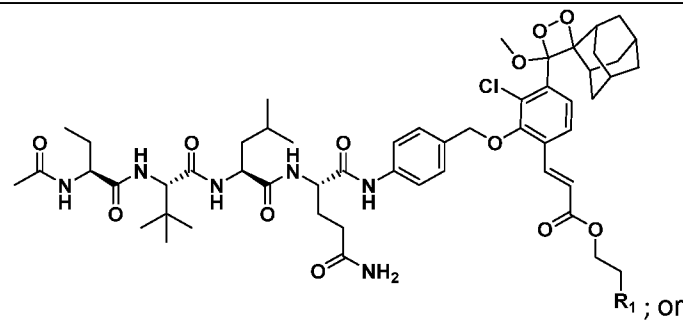
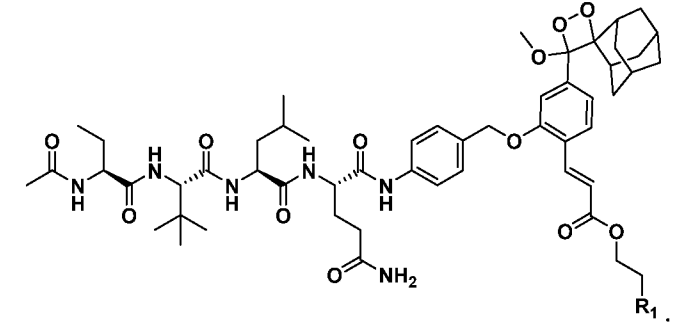


, where:

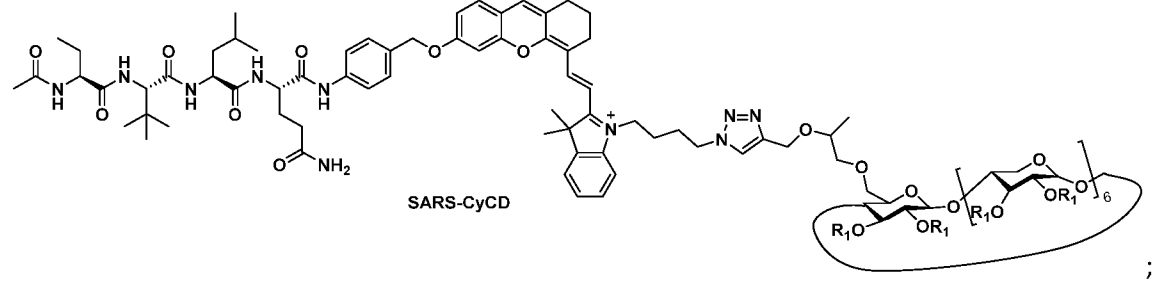
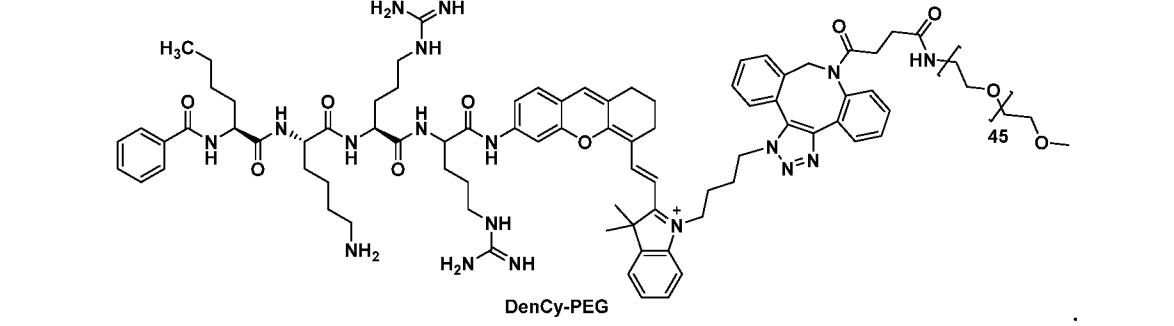
the wiggly line represents the point of attachment to the rest of the molecule; and
R represents H or 2-hydroxypropyl.

- 5 11. The compound according to Clause 9 or Clause 10, or a pharmaceutically acceptable salt or solvate thereof, wherein the compound of formula I is selected from:

A	
B	
C	

D	
E	

12. The compound according to Clause 1, or a pharmaceutically acceptable salt or solvate thereof, wherein the compound of formula I is selected from:

 SARS-CyCD	or  DenCy-PEG
---	---

- 5 13. A method for detection of a viral protease in an analyte, the method comprising the following steps:
- (a) providing an analyte and a fluid comprising a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof, according to any one of Clauses 1 to 12;

- (b) contacting the analyte with the fluid comprising a compound of formula I, or pharmaceutically acceptable salt or solvate thereof, for a period of time; and
- (c) after the period of time detecting any near-infrared fluorescence (NIRF) or, more particularly, chemiluminescence, wherein the presence of the viral protease in the fluid comprising the analyte and the compound of formula I, or pharmaceutically acceptable salt or solvate thereof, is indicated by near-infrared fluorescence (NIRF) or, more particularly, chemiluminescence.

14. A method for detection of viral protease *in vivo*, the method comprising the following steps:

- (ai) administering a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof, according to any one of Clauses 1 to 12 to a subject; and
- (aii) detecting any near-infrared fluorescence (NIRF) or, more particularly, chemiluminescence, wherein the presence of the viral protease *in vivo* is indicated by near-infrared fluorescence (NIRF) or, more particularly, chemiluminescence.

15. Use of a compound of formula I or a salt and/or solvate thereof according to any one of Clauses 1 to 12 in the manufacture of a diagnostic agent for *in vivo* diagnosis of a disease caused by a virus.

16. A compound of formula I or a salt and/or solvate thereof according to any one of Clauses 1 to 12 for use in the *in vivo* diagnosis of a disease caused by a virus.

17. A composition comprising a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof, according to any one of Clauses 1 to 12 in admixture with one or more of a pharmaceutically acceptable adjuvant, diluent and carrier.

Drawings

FIG. 1 depicts the scheme depicting *in vivo* detection of SARS-CoV-2 *via* intratracheal (i.t.) injection of M^{pro}-activatable NIRF probe (SARS-CyCD), followed by NIRF imaging and optical urinalysis.

FIG. 2 depicts the synthetic routes for a) CyOH, b) P-HP β CD c) Cy-CD, and d) SARS-CyCD.

FIG. 3 depicts the *in vitro* characterization of SARS-CyCD in response to SARS-CoV-2 M^{pro}. (a) UV/VIS absorption spectra and (b) fluorescence spectra of SARS-CyCD (10 μ M) in the

absence or presence of SARS-CoV-2 M^{pro} (250 nM) in Tris buffer (20 mM, pH 7.4) at 37 °C; (c) High-performance liquid chromatography (HPLC) traces of SARS-CyCD (10 µM) in the absence or presence of SARS-CoV-2 M^{pro} (250 nM); (d) Nonlinear regression analysis of cleavage rate V (µmol/min) of SARS-CyCD as a function of substrate concentration. Various concentrations of SARS-CyCD (1, 5, 10, 20, 40, 80, 150 µM) were incubated with SARS-CoV-2 M^{pro} (250 nM) at 37 °C for 30 min in Tris buffer (20 mM, pH 7.4). After incubation, the mixture was measured by HPLC; and (e) Near-infrared fluorescence (NIRF, 710 nm) changes and (f) the corresponding NIRF images acquired with an IVIS spectrum imaging system at 720 nm upon excitation at 675 nm of SARS-CyCD (10 µM) after incubation with different enzymes (250 nM) in corresponding buffers at 37 °C for 120 min. 1: Blank, 2: HCV NS4A/NS3-3protease, 3: urokinase, 4: caspase-3, 5: furin, 6: GGT, 7: SARS-CoV-1 M^{pro}, 8: SARS-CoV-2 M^{pro}. Error bars: standard deviation from three separate measurements.

FIG. 4 depicts the nonlinear regression analysis of cleavage rate V (µM min⁻¹) of SARS-Cy/SARS-CyCD as a function of substrate concentration. Various concentrations of SARS-Cy/SARS-CyCD (1, 5, 10, 20, 40, 80, 150 µM) were incubated with SARS-CoV-2 M^{pro} (250 nM) at 37 °C for 30 min in Tris buffer (20 mM, pH 7.4). After incubation, the mixture was measured by HPLC.

FIG. 5 depicts the renal clearance and *in vivo* stability studies of SARS-CyCD. (a) Blood concentration (% ID g⁻¹) of SARS-CyCD after i.t. injection into living mice; (b) Renal clearance efficiency (RCE) as a function of time post i.t. injection of SARS-CyCD (2 µmol kg⁻¹ body weight) in living mice; (c) *In vivo* stability studies of SARS-CyCD through HPLC analysis of excreted components in the urine samples after i.t. administration; and (d) The renal clearance efficiency of SARS-CyCD at 24 h after i.t. administration. Heart (He), liver (Li), spleen (Sp), intestine (In), muscle (Mu), lung (Lu), skin (Sk), kidney (Ki), bladder (Bl). Error bars: standard deviation from three separate measurements.

FIG. 6 depicts (a) fluorescence images acquired with an IVIS spectrum imaging system of excreted CyCD and SARS-CyCD in the urine from living mice at different timepoints post i.t. injection of CyCD or SARS-CyCD (2 µmol kg⁻¹ body weight); and (b) fluorescence spectra of excreted SARS-CyCD in the urine of living mice with SARS-CyCD in phosphate-buffered saline (PBS) as reference.

FIG. 7 depicts (a) timeline for i.t. injection of SARS-CoV-2 M^{pro} and SARS-CyCD, followed by NIRF imaging and optical urinalysis; (b) NIRF images at 0, 20, 40, 60, 90 and 120 mins after i.t. injection of 0.15 mg kg⁻¹ SARS-CoV-2 M^{pro} and 2 µmol kg⁻¹ SARS-CyCD into living mice;

Dynamic NIRF intensities of (c) lungs (ventral) and (d) bladder as a function of time post-injection of SARS-CyCD in living mice; and (e) Fluorescence images acquired with an IVIS spectrum imaging system and (f) enhancement of excreted SARS-CyCD and CyCD in the urine from living mice at different timepoints post i.t. injection. The values relative to the control groups. **p < 0.01 (n = 3).

FIG. 8 depicts (a) dynamic NIRF intensities of lungs (dorsal) as a function of time in saline or SARS-CoV-2 M^{pro} treated living mice (0.15 mg kg⁻¹ body weight), co-injected with SARS-CyCD (2 μmol kg⁻¹ body weight); and (b) fluorescence images of lungs collected from saline or SARS-CoV-2 M^{pro} treated mice (0.15 mg kg⁻¹ body weight), co-injected with SARS-CyCD (2 μmol kg⁻¹ body weight) after 24 h post i.t. injection. The values are relative to the control groups. **p < 0.01 (n = 3).

FIG. 9 depicts the cell viability of NIH3T3 fibroblast cells after 24 h incubation with SARS-CyCD at different concentrations.

FIG. 10 depicts the synthetic route for Den-CyPEG.

Description

Disclosed herein are a series of activatable molecular probes for the detection of viral proteases. The molecular probe may comprise a hemicyanine fluorophore/chemilumiphore caged with a protease peptide substrate and a cyclodextrin/polyethylene glycol (PEG) unit. In the presence of a specific viral protease, cleavage of the peptide substrate results in near-infrared fluorescence (NIRF) or a chemiluminiscent “Turn-ON” response, and the release of the fluorescent or chemiluminiscent fragment. The PEG-ylated fluorophore/chemilumiphore can be visualised *in vivo* via NIRF imaging. By virtue of the high renal clearance of the cyclodextrin moiety, cyclodextrin-modified fluorophore/chemilumiphore can also be excreted into urine for sensitive urinalysis and detection of virus infection.

In a first aspect of the invention, there is provided a compound of formula I:



wherein:

A represents a peptide group that targets a viral protease;

B, when present, represents a self-immolative linking group;

C represents a fluorophore or chemilumiphore group covalently linked to the rest of the

5 molecule by an oxygen atom or a NH group; and

n represents 0 or 1,

or a pharmaceutically acceptable salt or solvate thereof.

10 In embodiments herein, the word “comprising” may be interpreted as requiring the features mentioned, but not limiting the presence of other features. Alternatively, the word “comprising” may also relate to the situation where only the components/features listed are intended to be present (e.g. the word “comprising” may be replaced by the phrases “consists of” or “consists essentially of”). It is explicitly contemplated that both the broader and narrower interpretations can be applied to all aspects and embodiments of the present invention. In other words, the
15 word “comprising” and synonyms thereof may be replaced by the phrase “consisting of” or the phrase “consists essentially of” or synonyms thereof and *vice versa*.

The phrase, “consists essentially of” and its pseudonyms may be interpreted herein to refer to a material where minor impurities may be present. For example, the material may be greater
20 than or equal to 90% pure, such as greater than 95% pure, such as greater than 97% pure, such as greater than 99% pure, such as greater than 99.9% pure, such as greater than 99.99% pure, such as greater than 99.999% pure, such as 100% pure.

As used herein, the singular forms “a,” “an,” and “the” include plural referents unless the
25 context clearly dictates otherwise. Thus, for example, reference to “a composition” includes mixtures of two or more such compositions, and the like.

References herein (in any aspect or embodiment of the invention) to compounds of formula I include references to such compounds *per se*, to tautomers of such compounds, as well as to
30 pharmaceutically acceptable salts or solvates, or pharmaceutically functional derivatives of such compounds.

Pharmaceutically acceptable salts that may be mentioned include acid addition salts and base addition salts. Such salts may be formed by conventional means, for example by reaction of
35 a free acid or a free base form of a compound of formula I with one or more equivalents of an appropriate acid or base, optionally in a solvent, or in a medium in which the salt is insoluble, followed by removal of said solvent, or said medium, using standard techniques (e.g. *in vacuo*,

by freeze-drying or by filtration). Salts may also be prepared by exchanging a counter-ion of a compound of formula I in the form of a salt with another counter-ion, for example using a suitable ion exchange resin.

- 5 Examples of pharmaceutically acceptable salts include acid addition salts derived from mineral acids and organic acids, and salts derived from metals such as sodium, magnesium, or preferably, potassium and calcium.

10 Examples of acid addition salts include acid addition salts formed with acetic, 2,2-dichloroacetic, adipic, alginic, aryl sulphonic acids (e.g. benzenesulphonic, naphthalene-2-sulphonic, naphthalene-1,5-disulphonic and *p*-toluenesulphonic), ascorbic (e.g. L-ascorbic), L-aspartic, benzoic, 4-acetamidobenzoic, butanoic, (+) camphoric, camphor-sulphonic, (+)-(1*S*)-camphor-10-sulphonic, capric, caproic, caprylic, cinnamic, citric, cyclamic, dodecylsulphuric, ethane-1,2-disulphonic, ethanesulphonic, 2-hydroxyethanesulphonic, 15 formic, fumaric, galactaric, gentisic, glucoheptonic, gluconic (e.g. D-gluconic), glucuronic (e.g. D-glucuronic), glutamic (e.g. L-glutamic), α -oxoglutaric, glycolic, hippuric, hydrobromic, hydrochloric, hydriodic, isethionic, lactic (e.g. (+)-L-lactic and ($\pm$)-DL-lactic), lactobionic, maleic, malic (e.g. (-)-L-malic), malonic, ($\pm$)-DL-mandelic, metaphosphoric, methanesulphonic, 1-hydroxy-2-naphthoic, nicotinic, nitric, oleic, orotic, oxalic, palmitic, pamoic, phosphoric, 20 propionic, L-pyroglutamic, salicylic, 4-amino-salicylic, sebacic, stearic, succinic, sulphuric, tannic, tartaric (e.g. (+)-L-tartaric), thiocyanic, undecylenic and valeric acids.

Particular examples of salts are salts derived from mineral acids such as hydrochloric, hydrobromic, phosphoric, metaphosphoric, nitric and sulphuric acids; from organic acids, such 25 as tartaric, acetic, citric, malic, lactic, fumaric, benzoic, glycolic, gluconic, succinic, arylsulphonic acids; and from metals such as sodium, magnesium, or preferably, potassium and calcium.

As mentioned above, also encompassed by formula I are any solvates of the compounds and 30 their salts. Preferred solvates are solvates formed by the incorporation into the solid state structure (e.g. crystal structure) of the compounds of the invention of molecules of a non-toxic pharmaceutically acceptable solvent (referred to below as the solvating solvent). Examples of such solvents include water, alcohols (such as ethanol, isopropanol and butanol) and dimethylsulphoxide. Solvates can be prepared by recrystallising the compounds of the 35 invention with a solvent or mixture of solvents containing the solvating solvent. Whether or not a solvate has been formed in any given instance can be determined by subjecting crystals of the compound to analysis using well known and standard techniques such as

thermogravimetric analysis (TGE), differential scanning calorimetry (DSC) and X-ray crystallography.

The solvates can be stoichiometric or non-stoichiometric solvates. Particularly preferred solvates are hydrates, and examples of hydrates include hemihydrates, monohydrates and dihydrates.

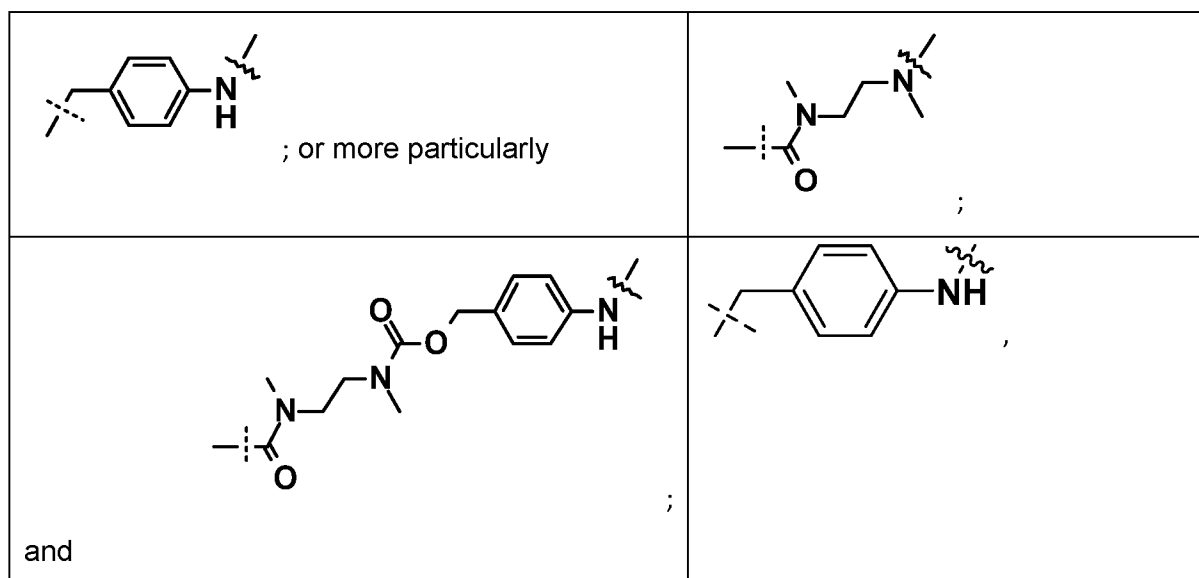
For a more detailed discussion of solvates and the methods used to make and characterise them, see Bryn *et al.*, *Solid-State Chemistry of Drugs*, Second Edition, published by SSCI, Inc of West Lafayette, IN, USA, 1999, ISBN 0-967-06710-3.

Compounds of formula I may contain double bonds and may thus exist as *E* (*entgegen*) and *Z* (*zusammen*) geometric isomers about each individual double bond. All such isomers and mixtures thereof are included within the scope of the invention.

Compounds of formula I may exist as regioisomers and may also exhibit tautomerism. All tautomeric forms and mixtures thereof are included within the scope of the invention.

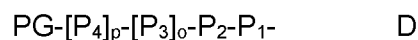
Compounds of formula I may contain one or more asymmetric carbon atoms and may therefore exhibit optical and/or diastereoisomerism. Diastereoisomers may be separated using conventional techniques, e.g. chromatography or fractional crystallisation. The various stereoisomers may be isolated by separation of a racemic or other mixture of the compounds using conventional, e.g. fractional crystallisation or HPLC, techniques. Alternatively, the desired optical isomers may be made by reaction of the appropriate optically active starting materials under conditions which will not cause racemisation or epimerisation (i.e. a 'chiral pool' method), by reaction of the appropriate starting material with a 'chiral auxiliary' which can subsequently be removed at a suitable stage, by derivatisation (i.e. a resolution, including a dynamic resolution), for example with a homochiral acid followed by separation of the diastereomeric derivatives by conventional means such as chromatography, or by reaction with an appropriate chiral reagent or chiral catalyst all under conditions known to the skilled person. All stereoisomers and mixtures thereof are included within the scope of the invention.

In embodiments of the first aspect, B, when it is present, may be selected from:



where the wavy line represents the point of attachment to A and the dotted line represents the point of attachment to C. This may apply to any of the following embodiments in relation to the expression of A in formula I.

- 5 In certain embodiments of the first aspect, A may represent a peptide sequence according to fragment D that targets a flavivirus protease:



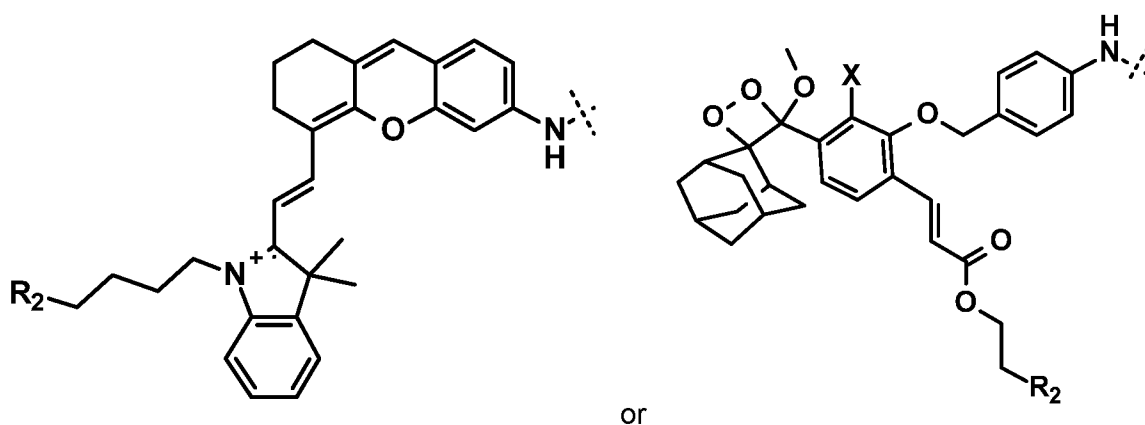
where:

- 10 P_1 to P_4 , when present, are amino acids that together form an oligo peptide;
 o is 0 or 1;
 p is 0 or 1; and
 PG is a protecting group. In certain embodiments that may be mentioned herein, PG may be selected from a benzoyl protecting group, an acetyl protecting group or a *tert*-butoxycarbonyl
 15 protecting group.

In particular embodiments that may be mentioned herein, fragment D may be selected from the list below, where the flavivirus(es) targeted by the peptide sequence are listed in [Square brackets]:

- 20 (a) Bz-Nle-Lys-Arg-Arg- [Dengue, Zika, Yellow Fever, Japanese Encephalitis];
 (b) Bz-Nle-Lys-Arg-Ala- [Yellow Fever];
 (c) Bz-Nle-Lys-Thr-Arg- [Dengue, Zika];
 (d) Bz-Nle-Thr-Arg-Arg- [Dengue, Zika];
 (e) Bz-Nle-Ala-Arg-Arg- [West Nile, Yellow Fever];
 25 (f) Bz-Nle-Lys-Arg-Phe- [Dengue, Zika, Yellow Fever];

- (g) Bz-Nle-Lys-Phe-Arg- [Dengue, Zika, Yellow Fever];
 (h) Bz-Nle-Phe-Arg-Arg- [Dengue, Zika, Yellow Fever];
 (i) Bz-Thr-Lys-Arg-Arg- [Dengue, Zika];
 (j) Bz-Thr-Thr-Arg-Arg- [Dengue, Zika];
 5 (k) Bz-Phe-Lys-Arg-Arg- [Dengue, Zika, West Nile];
 (l) Bz-Ala-Lys-Arg-Arg- [Dengue, Zika, Yellow Fever];
 (m) Bz-Nle-Lys-Lys-Arg- [Dengue, Zika, West Nile, Yellow Fever];
 (n) Bz-Lys-Arg-Arg- [Dengue, Zika, Yellow Fever];
 (o) Bz-Arg-Arg- [Dengue, Zika, West Nile, Yellow Fever];
 10 (p) Ac-Nle-Lys-Arg-Arg- [Dengue, Japanese Encephalitis];
 (q) Ac-Lys-Arg-Arg- [Dengue, Zika, West Nile, Yellow Fever];
 (r) Boc-Gly-Arg-Arg- [Dengue, Zika, West Nile, Japanese Encephalitis]; and
 (s) Boc-Lys-Arg-Arg- [Dengue, Zika, West Nile, Japanese Encephalitis].
- 15 In embodiments of the invention where A is a peptide sequence according to fragment D, C may be selected from:



20

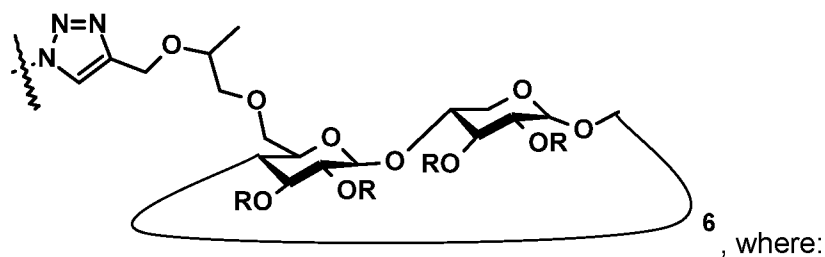
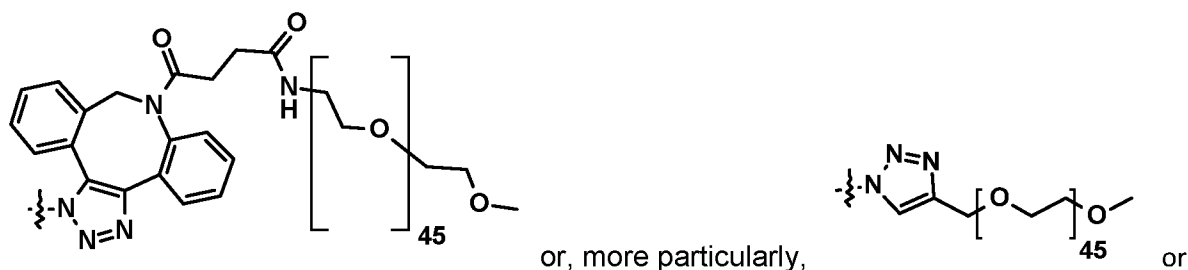
where:

the dotted line represents the point of attachment to the rest of the molecule;

X represents H or halo (e.g. Cl); and

R₂ is selected from:

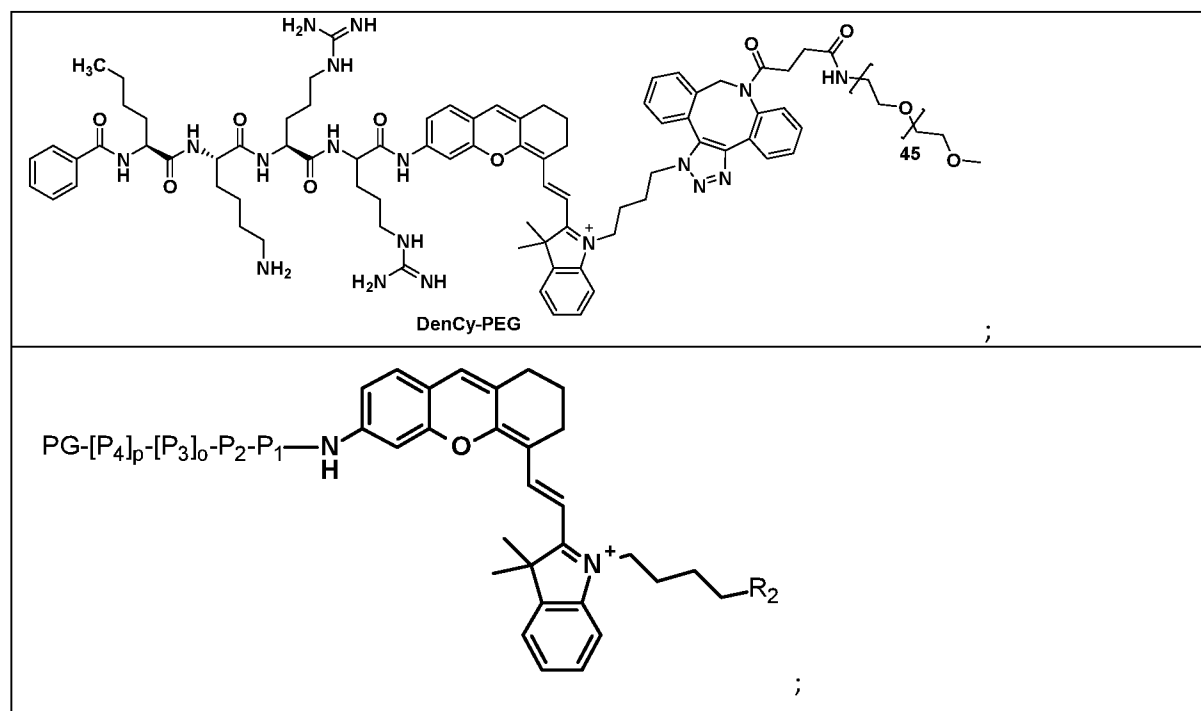
25

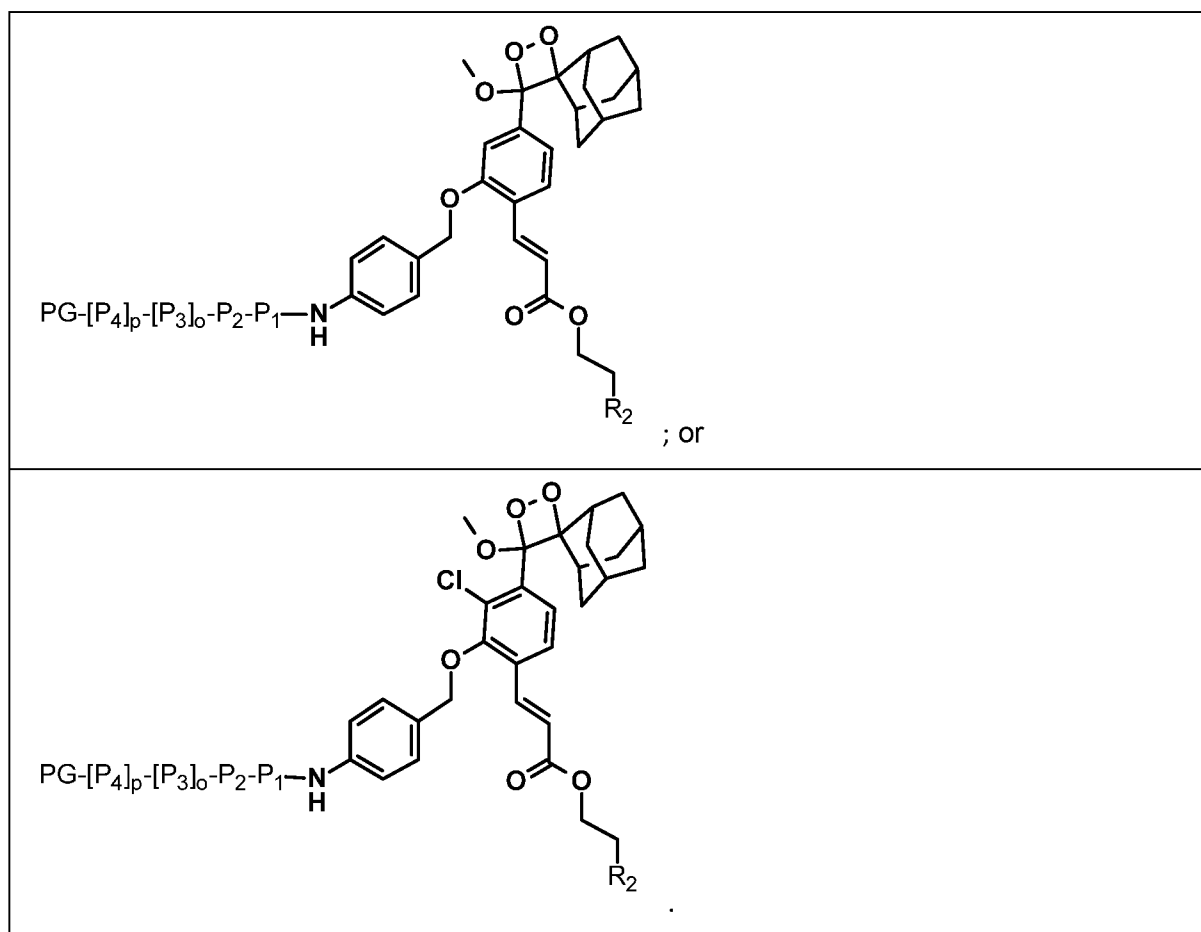


the wiggly line represents the point of attachment to the rest of the molecule; and

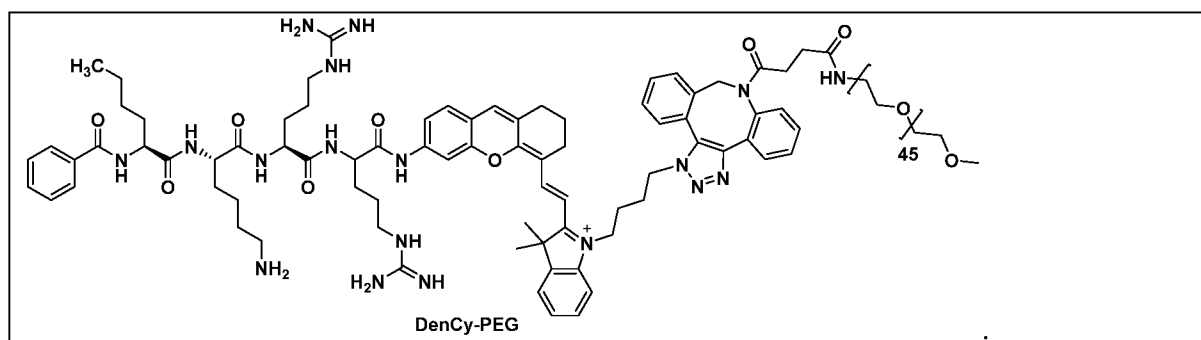
5 R represents H or 2-hydroxypropyl.

In embodiments of the invention where A is a peptide sequence according to fragment D, the compound of formula I, or a pharmaceutically acceptable salt or solvate thereof, may be selected from:

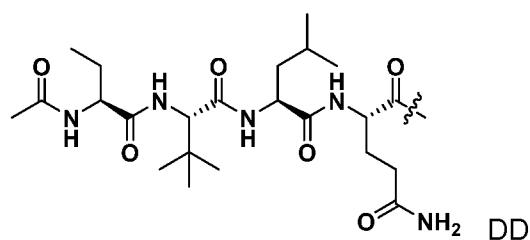




More particularly, the compound of formula I or a pharmaceutically acceptable salt or solvate thereof may be:

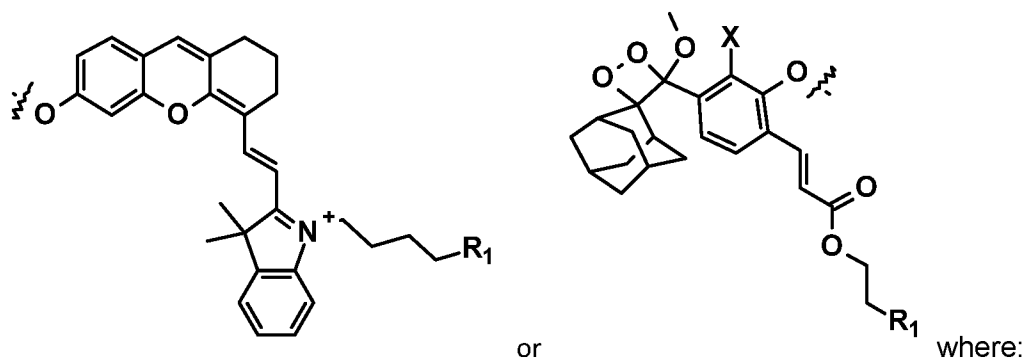


- 5 In alternative embodiments of the invention, A may represent a peptide sequence according to fragment DD that targets the SARS-CoV-2 main protease:



where the wavy line represents the point of attachment to the rest of the molecule.

In embodiments where A represents the fragment DD, C may be selected from:

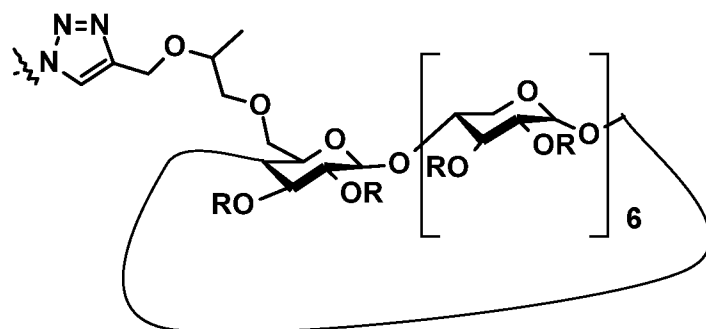


5

the wiggly line represents the point of attachment to the rest of the molecule;

X represents H or halo (e.g. Cl); and

R₁ represents:

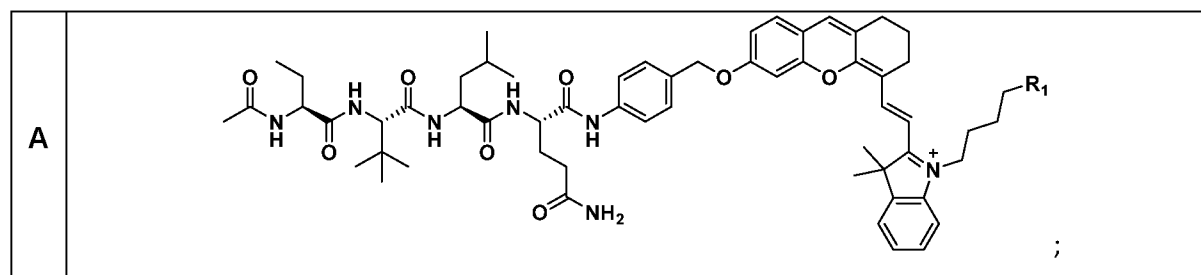


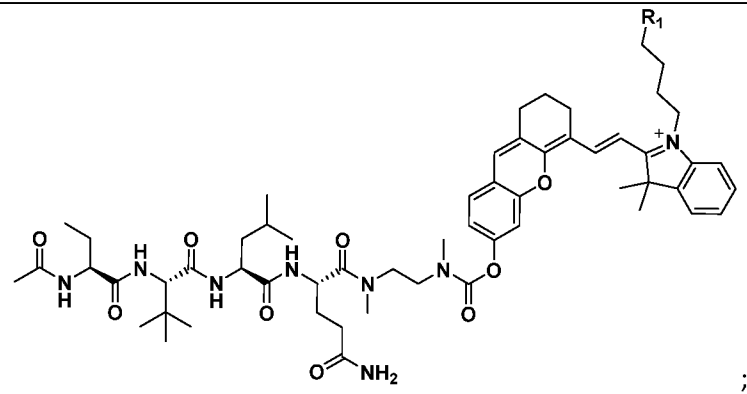
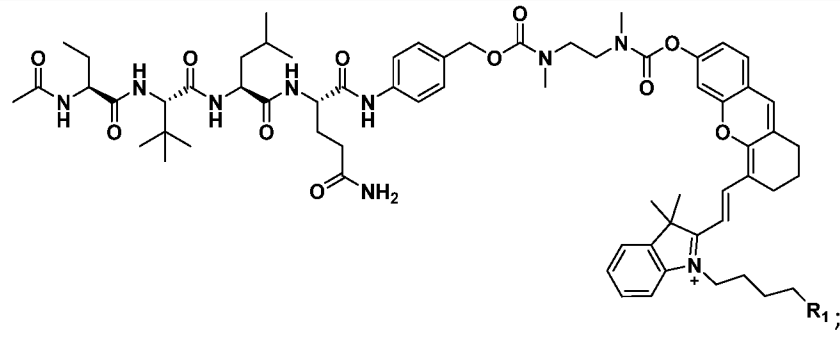
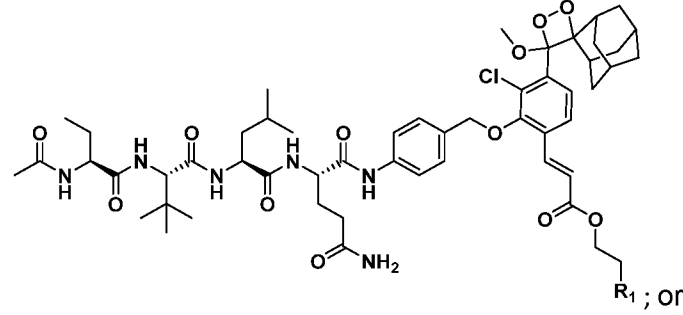
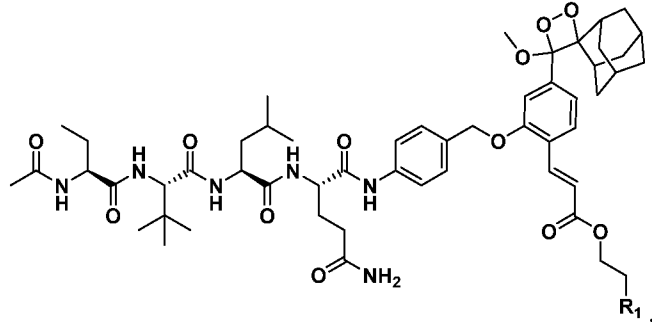
, where:

- 10 the wiggly line represents the point of attachment to the rest of the molecule; and
R represents H or 2-hydroxypropyl.

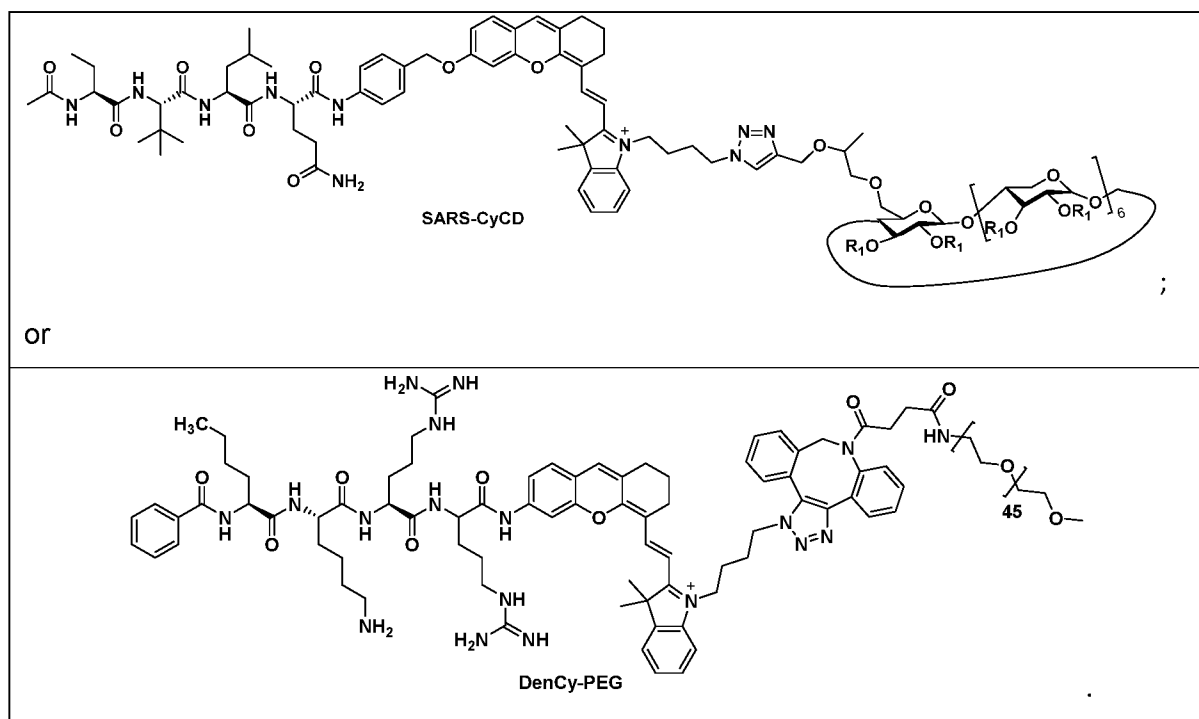
In embodiments where A represents the fragment DD, the compound of formula I or a pharmaceutically acceptable salt or solvate thereof, may be selected from:

15



B	
C	
D	
E	

More particularly, the compound of formula I or a pharmaceutically acceptable salt or solvate thereof, may be selected from:



For the avoidance of doubt, in the context of the present invention, the term “*treatment*” includes references to therapeutic or palliative treatment of patients in need of such treatment, as well as to the prophylactic treatment and/or diagnosis of patients which are susceptible to the relevant disease states.

The terms “*patient*” and “*patients*” include references to mammalian (e.g. human) patients. As used herein the terms “subject” or “patient” are well-recognized in the art, and, are used interchangeably herein to refer to a mammal, including dog, cat, rat, mouse, monkey, cow, horse, goat, sheep, pig, camel, and, most preferably, a human. In some embodiments, the subject is a subject in need of treatment or a subject with a disease or disorder. However, in other embodiments, the subject can be a normal subject. The term does not denote a particular age or sex. Thus, adult and newborn subjects, whether male or female, are intended to be covered.

The term “effective amount” refers to an amount of a compound, which confers a therapeutic effect on the treated patient (e.g. sufficient to treat or prevent the disease). The effect may be objective (i.e. measurable by some test or marker) or subjective (i.e. the subject gives an indication of or feels an effect).

Further embodiments of the invention that may be mentioned include those in which the compound of formula I is isotopically labelled. However, other, particular embodiments of the

invention that may be mentioned include those in which the compound of formula I is not isotopically labelled.

The term "isotopically labelled", when used herein includes references to compounds of formula I in which there is a non-natural isotope (or a non-natural distribution of isotopes) at one or more positions in the compound. References herein to "one or more positions in the compound" will be understood by those skilled in the art to refer to one or more of the atoms of the compound of formula I. Thus, the term "isotopically labelled" includes references to compounds of formula I that are isotopically enriched at one or more positions in the compound.

The isotopic labelling or enrichment of the compound of formula I may be with a radioactive or non-radioactive isotope of any of hydrogen, carbon, nitrogen, oxygen, sulfur, fluorine, chlorine, bromine and/or iodine. Particular isotopes that may be mentioned in this respect include ^2H , ^3H , ^{11}C , ^{13}C , ^{14}C , ^{13}N , ^{15}N , ^{15}O , ^{17}O , ^{18}O , ^{35}S , ^{18}F , ^{37}Cl , ^{77}Br , ^{82}Br and ^{125}I .

When the compound of formula I is labelled or enriched with a radioactive or nonradioactive isotope, compounds of formula I that may be mentioned include those in which at least one atom in the compound displays an isotopic distribution in which a radioactive or non-radioactive isotope of the atom in question is present in levels at least 10% (e.g. from 10% to 5000%, particularly from 50% to 1000% and more particularly from 100% to 500%) above the natural level of that radioactive or non-radioactive isotope.

The compounds of formula I may be prepared for administration to a subject. Thus, in a further aspect of the invention, there is provided a composition comprising a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof as described herein in admixture with one or more of a pharmaceutically acceptable adjuvant, diluent and carrier.

Compounds of formula I may be administered by any suitable route, but may particularly be administered orally, intravenously, intramuscularly, cutaneously, subcutaneously, transmucosally (e.g. sublingually or buccally), rectally, transdermally, nasally, pulmonarily (e.g. tracheally or bronchially), topically, by any other parenteral route, in the form of a pharmaceutical preparation comprising the compound in a pharmaceutically acceptable dosage form. Particular modes of administration that may be mentioned include oral, intravenous, cutaneous, subcutaneous, nasal, intramuscular or intraperitoneal administration.

Compounds of formula I will generally be administered as a pharmaceutical formulation in admixture with a pharmaceutically acceptable adjuvant, diluent or carrier, which may be selected with due regard to the intended route of administration and standard pharmaceutical practice. Such pharmaceutically acceptable carriers may be chemically inert to the active compounds and may have no detrimental side effects or toxicity under the conditions of use. Suitable pharmaceutical formulations may be found in, for example, Remington *The Science and Practice of Pharmacy*, 19th ed., Mack Printing Company, Easton, Pennsylvania (1995). For parenteral administration, a parenterally acceptable aqueous solution may be employed, which is pyrogen free and has requisite pH, isotonicity, and stability. Suitable solutions will be well known to the skilled person, with numerous methods being described in the literature. A brief review of methods of drug delivery may also be found in e.g. Langer, *Science* (1990) 249, 1527.

Otherwise, the preparation of suitable formulations may be achieved routinely by the skilled person using routine techniques and/or in accordance with standard and/or accepted pharmaceutical practice.

The amount of compound of formula I in any pharmaceutical formulation used in accordance with the present invention will depend on various factors, such as the severity of the condition to be treated, the particular patient to be treated, as well as the compound(s) which is/are employed. In any event, the amount of compound of formula I in the formulation may be determined routinely by the skilled person.

For example, a solid oral composition such as a tablet or capsule may contain from 1 to 99 % (w/w) active ingredient; from 0 to 99% (w/w) diluent or filler; from 0 to 20% (w/w) of a disintegrant; from 0 to 5% (w/w) of a lubricant; from 0 to 5% (w/w) of a flow aid; from 0 to 50% (w/w) of a granulating agent or binder; from 0 to 5% (w/w) of an antioxidant; and from 0 to 5% (w/w) of a pigment. A controlled release tablet may in addition contain from 0 to 90 % (w/w) of a release-controlling polymer.

A parenteral formulation (such as a solution or suspension for injection or a solution for infusion) may contain from 1 to 50 % (w/w) active ingredient; and from 50% (w/w) to 99% (w/w) of a liquid or semisolid carrier or vehicle (e.g. a solvent such as water); and 0-20% (w/w) of one or more other excipients such as buffering agents, antioxidants, suspension stabilisers, tonicity adjusting agents and preservatives.

Depending on the disorder, and the patient, to be treated, as well as the route of administration, compounds of formula I may be administered at varying therapeutically effective doses to a patient in need thereof.

- 5 However, the dose administered to a mammal, particularly a human, in the context of the present invention should be sufficient to effect a therapeutic or diagnostic response in the mammal over a reasonable timeframe. One skilled in the art will recognize that the selection of the exact dose and composition and the most appropriate delivery regimen will also be influenced by *inter alia* the pharmacological properties of the formulation, the nature and severity of the condition being treated, and the physical condition and mental acuity of the recipient, as well as the potency of the specific compound, the age, condition, body weight, sex and response of the patient to be treated, and the stage/severity of the disease.

- 15 Administration may be continuous or intermittent (e.g. by bolus injection). The dosage may also be determined by the timing and frequency of administration. In the case of oral or parenteral administration the dosage can vary from about 0.01 mg to about 1000 mg per day of a compound of formula I.

- 20 In any event, the medical practitioner, or other skilled person, will be able to determine routinely the actual dosage, which will be most suitable for an individual patient. The above-mentioned dosages are exemplary of the average case; there can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

- 25 As will be appreciated, the compounds disclosed herein have high renal clearance. As will be appreciated, the compounds disclosed herein may be suitable for use in the diagnosis of viral infections *in vitro* and *in vivo* via the detection of the presence of a viral protease. Thus in a further aspect of the invention, there is provided a method for detection of a viral protease in an analyte, the method comprising the following steps:

- 30 (a) providing an analyte and a fluid comprising a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof, as described hereinbefore;
- (b) contacting the analyte with the fluid comprising a compound of formula I, or pharmaceutically acceptable salt or solvate thereof, for a period of time; and
- (c) after the period of time detecting any near-infrared fluorescence (NIRF) or, more particularly, chemiluminescence, wherein the presence of the viral protease in the fluid
- 35 comprising the analyte and the compound of formula I, or pharmaceutically acceptable

salt or solvate thereof, is indicated by near-infrared fluorescence (NIRF) or, more particularly, chemiluminescence.

In a further aspect of the invention, there is disclosed a method for detection of a viral protease *in vivo*, the method comprising the following steps:

- (ai) administering a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof as described herein to a subject; and
- (a ii) detecting any near-infrared fluorescence (NIRF) or, more particularly, chemiluminescence, wherein the presence of the viral protease *in vivo* is indicated by near-infrared fluorescence (NIRF) or, more particularly, chemiluminescence.

In a further aspect of the invention, there is provided a use of a compound of formula I or a salt and/or solvate thereof as described herein for the manufacture of a diagnostic agent for *in vivo* diagnosis of a disease caused by a virus.

In a yet further aspect, there is provided a compound of formula I or a salt and/or solvate thereof as described herein for use in the *in vivo* diagnosis of a disease caused by a virus.

Again, the diseases that may be detected by viral protease detection are described hereinbefore. As will be appreciated, the skilled person making the diagnosis may then treat the subject according to the presence (or absence) of the disease in question.

Examples

Materials

All commercial reagents were purchased from reputable vendors and used without further purification, unless indicated otherwise. 2-chlorotriyl chloride polystyrene resin, Fmoc-Gln(Trt)-OH, Fmoc-Leu-OH, Fmoc-Tle-OH, Fmoc-Abu-OH, 1-hydroxybenzotriazole hydrate (HOBT), and 3-[bis(dimethylamino)methylumyl]-3H-benzotriazol-1-oxide hexafluorophosphate (HBTU) were purchased from Sangon Biotech for solid phase peptide synthesis (SPPS). SARS-CoV-2 main protease (Cat. No. SAE0172) and γ -glutamyl transferase (GGT) were purchased from Sigma Aldrich. Hepatitis C Virus (HCV) NS4A/NS3-3 protease was purchased from Sigma. SARS-CoV-1 main protease, furin, urokinase (uPA) and caspase-3 were purchased from R&D Systems. Heparinized capillary tubes was purchased from Paul Marienfeld, Germany. Metabolic cages were purchased from Lab Products Inc, USA. Microspray aerosolizer for i.t. administration in mice was purchased from

PenWu, Bio Jane Trading Limited. MTS assay (Promega Cat. no. G3581) was purchased from Promega.

Analytical techniques

5 *Thin layer chromatography (TLC)*

The reaction progress was monitored by TLC on pre-coated silica plates (Merck 60 F254 nm, 250 μ m thickness) and spots were visualized by ultraviolet (UV) light or appropriate staining (e.g phosphomolybdic acid stain (PMA), and basic KMnO_4).

10 *Flash column chromatography*

Flash column chromatography was carried out using 200 or 400 mesh silica gel.

Nuclear magnetic resonance (NMR)

15 All ^1H NMR and ^{13}C NMR spectra were carried out on a Bruker ACF-400 MHz NMR spectrometer. Chemical shifts were reported in parts per million (ppm) relative to residual solvent peaks. ^1H and ^{13}C chemical shifts (δ) were referenced to TMS or residual solvent peaks ($\text{CDCl}_3 = 7.26$ ppm and $(\text{CD}_3)_2\text{SO} = 2.50$ ppm) for ^1H NMR. The following abbreviations were used for reporting ^1H NMR spectra: chemical shift (δ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, dd = doublet of doublets, m = multiplet), and coupling constant
20 (Hz).

Electrospray ionization-mass spectrometry (ESI-MS)

ESI-MS spectra were acquired on a Thermo Finnigan Polaris Q quadrupole ion trap mass spectrometer (ThermoFisher Corporation) equipped with a standard ESI source.

25

HPLC

HPLC analyses were done on an Agilent 1260 system equipped with a G1311B pump, UV detector and an Agilent Zorbax SB-C18 RP (9.4 $\times$ 250 mm) column, with methanol and water as the eluent.

30

UV/Vis spectroscopy

UV/Vis spectra were measured on a Shimadzu UV-2450 spectrophotometer.

Fluorescence spectroscopy

35 Fluorescence measurements were performed on a Fluorolog 3-TCSPC spectrofluorometer (Horiba Jobin Yvon).

Calculation of partition coefficients

In silico calculation of the partition coefficients (Log D at pH 7.4) was calculated using Marvin and JChem calculator plug-ins (ChemAxon, Hungary).

5 General procedure for animal experiments

All animal experiments were performed in accordance with the Guidelines for Care and Use of Laboratory Animals of the Nanyang Technological University-Institutional Animal Care and Use Committee (NTU-IACUC) and approved by the Institutional Animal Care and Use Committee (IACUC) for Animal Experiment, Singapore. Adult female Ncr nude mice (18 – 20
10 g) were housed in a temperature controlled (22 °C) room with 12 h dark light cycles (0700 h on and 1900 h off). The animals were provided ad libitum with food (Tecklad T.2918.CS Irradiated Rodent Diet purchased from Acre Engineering Pte Ltd) and water.

General procedure for statistical analysis

15 The *in vivo* fluorescence intensities were quantified with ROI analysis using Living Image 4.3 Software. Data is mean $\pm$ standard deviation (S.D.) unless stated otherwise. Investigators were blinded to group allocation during experiments. Statistical differences between two groups were tested with a two-tailed Student's t-test and more than three groups were determined by one-way analysis of variance followed by Tukey's post hoc test. For all tests, P
20 values less than 0.05 were considered statistically significant. *P < 0.05, **P < 0.01 and ***P < 0.001. All statistical calculations were performed using GraphPad Prism 6.0, including assumptions of tests used.

Example 1. Synthesis of SARS-CyCD

25

Herein, we report the first protease-activatable NIRF probe (termed SARS-CyCD) for *in vivo* imaging and urinalysis of SARS-CoV-2.

The main protease (M^{pro}), one of the coronavirus proteases that plays a pivotal role in viral polypeptide processing, was chosen as the biomarker for signal activation of SARS-CyCD.
30 M^{pro} has also been validated as an *in vivo* diagnostic marker for SARS-CoV-2. SARS-CyCD comprises a hemicyanine fluorophore caged with the M^{pro} peptide substrate (N-Acetyl-Abu-Tle-Leu-Gln(Trt)-OH) via a *para*-aminobenzylalcohol (PABA) self-immolative linker and a cyclodextrin unit, which functions as the NIRF signaling moiety and renal-clearance enabler,
35 respectively (FIG. 1). After i.t. administration into the lungs of living mice, the peptide substrate of SARS-CyCD is cleaved by SARS-CoV-2 M^{pro} , resulting in NIRF "Turn-ON" response and the release of the fluorescent fragment (CyCD). By virtue of the high renal clearance of

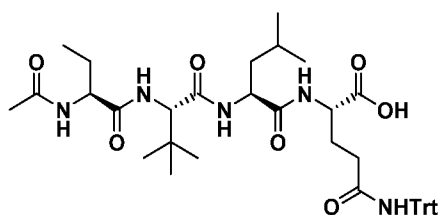
cyclodextrin moiety, CyCD can be excreted into urine for sensitive urinalysis of SARS-CoV-2. Such a delivery strategy enables the probe to avoid the first-pass metabolism with good bioavailability in the lungs, the main route of viral entry where viral load is high. Subsequent *in vivo* NIRF imaging of the lungs and optical urinalysis can therefore permit non-invasive and real-time detection of SARS-CoV-2 (FIG. 1), further highlighting its potential for clinical translation.

SARS-CyCD was prepared as shown in FIG. 2. SARS-CoV-2 peptide substrate (*N*-Acetyl-Abu-Tle-Leu-Gln(Trt)-OH) was first synthesized using standard SPPS, followed by amide coupling with PABA to yield 2. Next, 2 was brominated using PBr₃, followed by nucleophilic substitution reaction with the hydroxyl group of CyOH to obtain compound 4. Removal of trityl (Trt) protecting group on Gln of 4 with trifluoroacetic acid afforded 5. Finally, copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) reaction of SARS-Cy (5) with propynyl-HPβCD afforded the probe, SARS-CyCD.

All reactions were sealed with septa through which a nitrogen atmosphere was introduced unless otherwise stated. All non-aqueous reactions were carried out under a nitrogen atmosphere in oven-dried glassware. Reactions were conducted in round-bottomed flasks containing Teflon-coated magnetic stir bars. Heating of reactions was accomplished with a silicon oil bath on top of a stirring hotplate equipped with an electronic contact thermometer to maintain the indicated temperatures.

CyOH, P-HPβCD and CyCD were synthesised according to reported protocols (Cheng, P. *et al.*, *J. Am. Chem. Soc.* **2019**, *141*, 10581-10584; and Huang, J. *et al.*, *Nat. Mater.* **2019**, *18*, 1133-1143).

Compound 1 (SARS(Trt))



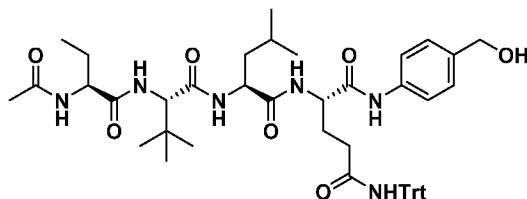
Peptide *N*-Acetyl-Abu-Tle-Leu-Gln(Trt)-OH (1.40 g, 2.0 mmol) was synthesised by SPPS.

¹H NMR (400 MHz, DMSO-*d*₆) δ: 12.54 (s, 1H), 8.60 (s, 1H), 8.03-7.96 (m, 3H), 7.62 (d, *J* = 9.44 Hz, 1H), 7.29-7.16 (m, 15H), 4.36 (q, *J* = 7.6 Hz, 1 H), 4.29-4.23 (m, 2H), 4.17-4.11 (m, 1 H), 2.41-2.27 (m, 2H), 2.00-1.90 (m, 1H), 1.84 (s, 3H), 1.73-1.58 (m, 6H), 1.07-1.01 (m, 1H),

1.00-0.84 (m, 17H). ^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ 173.59, 172.35, 171.90, 171.66, 170.19, 169.70, 162.78, 145.36, 128.98, 127.90, 126.77, 69.68, 60.03, 54.32, 52.03, 51.14, 41.28, 36.24, 34.69, 33.02, 31.24, 27.70, 27.09, 25.56, 24.53, 23.53, 22.89, 21.93, 10.58. MS (ESI): m/z = 742.29 $[\text{M} + \text{H}]^+$.

5

Compound 2 (SARS(Trt)-PABA)



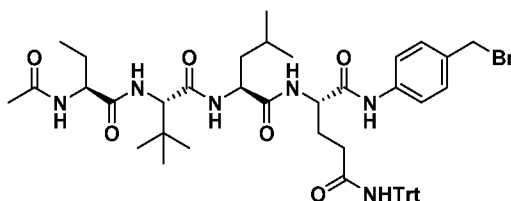
N-Ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline (EEDQ, 197 mg, 0.8 mmol) was added to Compound 1 (300 mg, 0.4 mmol) and the reaction mixture was dissolved in anhydrous dichloromethane (DCM). The reaction mixture was stirred at room temperature for 20 min. Next, 4-aminobenzyl alcohol (99 mg, 0.8 mmol) was added and the reaction was stirred at room temperature for 16 h. The reaction mixture was concentrated under reduced pressure and triturated with ether (50 mL) thrice to obtain pure Compound 2 as a white solid (300 mg, 88%).

15

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ : 9.91 (s, 1H), 8.64 (s, 1H), 8.04-7.95 (m, 3H), 7.63 (d, J = 9.24 Hz, 1H), 7.53 (d, J = 8.08 Hz, 2H), 7.28-7.16 (m, 18H), 4.43 (s, 2H), 4.36-4.23 (m, 4H), 2.37-2.33 (m, 2H), 2.01-2.00 (m, 1H), 1.83 (s, 3H), 1.57-1.44 (m, 6H), 0.93-0.80 (m, 18H). ^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ 172.27, 172.05, 171.79, 170.45, 170.38, 169.78, 145.36, 138.01, 137.93, 129.00, 127.90, 127.34, 126.77, 119.48, 69.71, 65.39, 63.10, 60.22, 55.33, 54.37, 53.80, 51.38, 40.98, 34.65, 33.10, 28.64, 27.11, 25.53, 24.62, 23.55, 23.42, 22.90, 21.85, 15.63, 10.60. MS (ESI): m/z = 847.54 $[\text{M} + \text{H}]^+$.

20

Compound 3 (SARS(Trt)-Br)



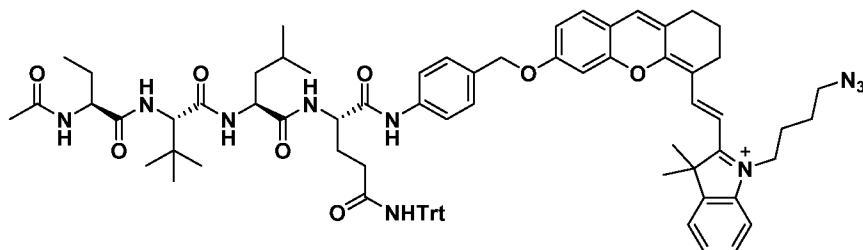
25

Compound 2 (120 mg, 0.14 mmol) was dissolved in anhydrous tetrahydrofuran (THF) and cooled in an ice bath. Next, PBr_3 (27 μL , 0.28 mmol) was added dropwise. The reaction mixture was left to stir at 0 $^\circ\text{C}$ for 2 h or until reaction was completed as indicated by TLC. Following which, the solvent was removed under reduced pressure. The resulting crude solids were dissolved in DCM (100 mL) and washed with saturated NaHCO_3 (50 mL x 3) and brine (50 mL

30

x 1). The organic layer was dried with anhydrous Na₂SO₄ and concentrated under reduced pressure. The crude product was triturated with ether (10 mL x 2) to yield a powdery off-white solid. Crude Compound 3 was used immediately in the next step without further purification.

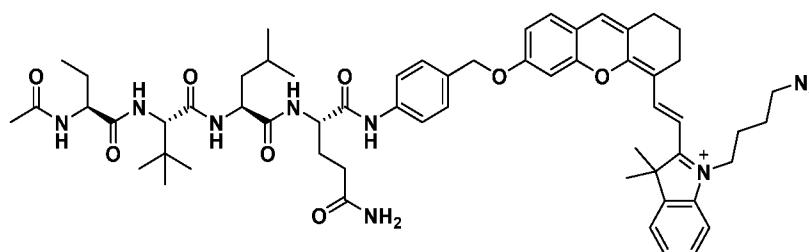
5 Compound 4 (SARS(Trt)-Cy)



CyOH (12 mg, 0.025 mmol) was dissolved in 1 mL anhydrous THF/acetonitrile (1:1), followed by the addition of K₂CO₃ (13.8 mg, 0.10 mmol). The reaction was stirred at room temperature for 10 min. Compound 3 (91 mg, 0.10 mmol) was then added, and the reaction mixture was left to stir at 50 °C for 4 h. After completion, the resulting mixture was concentrated under reduced pressure. The resulting crude solids were dissolved in DCM (50 mL) and washed with water (25 mL x 2) and brine (25 mL x 1). The organic layer was dried with anhydrous Na₂SO₄ and concentrated under reduced pressure. Pure compound 4 (10.0 mg, 30%) was then obtained after preparative HPLC purification.

¹H NMR (400 MHz, Methanol-*d*₄) δ: 8.75 (d, *J* = 16.0 Hz, 1H), 8.15 (d, *J* = 7.60 Hz, 1H), 7.75-7.73 (m, 1H), 7.70 (d, *J* = 8.00 Hz, 1H), 7.65 (d, *J* = 12.0 Hz, 2H), 7.55-7.54 (m, 2H), 7.48-7.40 (m, 5H), 7.26-7.23 (m, 15H), 7.07-7.04 (m, 2H), 6.51 (d, *J* = 16.0 Hz, 1H), 5.24 (s, 2H), 4.45-4.35 (m, 4H), 4.28-4.23 (m, 4H), 3.49-3.43 (m, 3H), 2.80-2.77 (m, 2H), 2.75-2.72 (m, 2H), 2.50-2.49 (m, 3H), 1.97 (s, 3H), 1.82 (s, 6H), 1.80-1.73 (m, 6H), 1.68-1.61 (m, 8H), 0.95-0.87 (m, 17H). ¹³C NMR (100 MHz, Methanol-*d*₄) δ 177.59, 173.28, 173.15, 172.81, 172.77, 172.27, 171.44, 170.55, 162.38, 161.99, 154.39, 145.78, 144.57, 142.05, 141.47, 138.22, 134.15, 132.22, 128.87, 128.65, 127.95, 127.35, 127.06, 126.43, 122.57, 119.98, 119.85, 115.96, 114.51, 112.36, 103.12, 101.38, 70.23, 70.17, 61.28, 55.21, 53.60, 52.05, 50.64, 50.39, 44.13, 40.01, 33.74, 32.44, 29.37, 27.09, 27.07, 25.87, 25.56, 24.51, 24.46, 23.94, 23.65, 22.33, 22.11, 21.02, 20.44, 20.23, 13.01, 9.35. MS (ESI): *m/z* = 1295.65 [M]⁺.

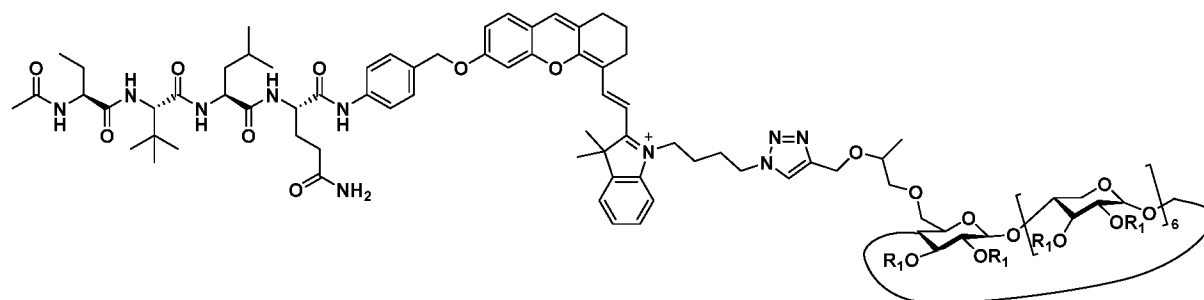
Compound 5 (SARS-Cy)



Compound 4 (13 mg, 0.01 mmol) was dissolved in DCM (1 mL) and cooled in an ice bath. Triethylsilane (25 μ L, 2.5% v/v) was added, followed by trifluoroacetic acid (TFA, 200 μ L, 20% v/v). The reaction was left to stir in an ice bath for 16 h. The reaction was monitored by analytical HPLC. Upon completion, the reaction was diluted with DCM (50 mL) and washed with saturated NaHCO₃ solution (25 mL x 3) and brine (25 mL x 1). The organic layer was dried with anhydrous Na₂SO₄ and concentrated under reduced pressure. Pure SARS-Cy (7.2 mg, 68%) was then obtained after preparative HPLC purification.

¹H NMR (400 MHz, Methanol-*d*₄) δ : 8.78 (d, *J* = 14.72 Hz, 1H), 8.23-8.18 (m, 2H), 7.74-7.69 (m, 4H), 7.65-7.63 (m, 1H), 7.57-7.44 (m, 7H), 7.10-7.06 (m, 2 H), 6.53 (d, *J* = 14.8 Hz, 1H), 5.96-5.86 (m, 2H), 5.36 (t, *J* = 4.64 Hz, 1H), 5.31-5.26 (m, 3H), 5.20 (dd, *J* = 1.28 Hz, 10.48 Hz, 2H), 4.56-4.54 (m, 2H), 4.49-4.46 (m, 1H), 4.42-4.37 (m, 2H), 4.25-4.22 (m, 2H), 3.46 (t, *J* = 6.5 Hz, 2H), 2.82-2.73 (m, 4H), 2.54-2.47 (m, 2H), 2.21 (t, *J* = 8.0 Hz, 2H), 2.1-2.0 (m, 2H), 2.00 (s, 3H), 1.99-1.95 (m, 2H), 1.84 (d, *J* = 2.0 Hz, 5H), 1.85-1.75 (m, 2H), 1.72-1.57 (m, 3H), 1.00-0.96 (m, 9H), 0.94-0.87 (m, 8H). ¹³C NMR (100 MHz, Methanol-*d*₄) δ 177.6275, 173.3062, 173.1101, 172.2337, 170.4758, 162.3822, 161.9813, 154.4020, 145.7650, 142.0564, 141.4598, 138.2226, 134.1084, 132.2178, 128.8447, 128.7012, 128.0307, 127.9603, 127.2451, 126.9818, 122.5178, 120.0715, 119.9645, 115.9419, 114.4215, 114.3063, 112.2939, 103.1012, 101.2972, 70.1755, 68.8361, 61.1539, 55.1961, 53.6366, 52.0928, 50.5805, 50.3936, 44.1057, 40.0117, 33.8934, 33.8291, 31.1836, 29.3431, 28.9009, 28.6118, 27.0505, 25.8024, 25.5863, 25.4928, 24.5075, 24.4272, 23.6523, 22.0044, 20.9464, 20.4729, 20.2578, 9.2921. MS (ESI): *m/z* = 1053.57 [M]⁺.

Compound SARS-CyCD



SARS-Cy (24 mg, 0.02 mmol) and P-HP β CD (40 mg, 0.025 mmol) were first dissolved in 3 mL DMSO/H₂O (2:1). The reaction mixture was purged with nitrogen gas for 5 min. Sodium ascorbate (31.6 mg, 0.16 mmol) and CuSO₄·5H₂O (10.0 mg, 0.04 mmol) were next added. The reaction mixture was purged with nitrogen gas for another 5 min and allowed to stir at room temperature for 24 h. Upon completion, the reaction was dialysed with deionised water. Pure SARS-CyCD (20 mg, 33%) was then obtained after preparative HPLC purification.

¹H NMR (400 MHz, DMSO-*d*₆) δ : 9.95 (s, 1H), 8.03-7.98 (m, 4H), 7.66-7.45 (m, 5H), 7.27-7.23 (m, 4H), 7.11-7.01 (m, 2H), 6.78 (s, 2H), 5.89-5.71 (m, 8H), 5.54-5.46 (m, 3H), 5.10-5.02 (m, 8H), 4.83 (s, 7H), 4.68-4.34 (m, 19H), 4.36-4.14 (m, 10H), 3.75-3.62 (m, 61H), 2.13-2.12 (m, 2H), 1.84 (s, 6H), 1.63-1.60 (m, 2H), 1.50-1.45 (m, 2H), 1.02 (s, 18H), 0.90-0.81 (m, 17H). ¹³C NMR (100 MHz, DMSO-*d*₆) δ 174.02, 173.23, 172.33, 172.04, 170.74, 170.47, 169.83, 129.24, 128.95, 127.91, 126.78, 119.70, 105.89, 91.77, 88.35, 80.78, 77.50, 75.27, 73.73, 65.75, 65.51, 60.18, 58.37, 58.28, 58.03, 56.11, 54.36, 53.70, 51.41, 34.61, 31.85, 31.71, 30.86, 29.47, 29.13, 29.01, 28.26, 28.10, 27.05, 25.56, 25.43, 24.56, 23.50, 22.88, 22.53, 21.80, 20.30, 19.65, 14.39, 13.93, 10.58. MALDI-TOF MS found: 2,600–3,000.

Example 2. *In vitro* response of SARS-CyCD

The *in vitro* response of SARS-CyCD towards SARS-CoV-2 M^{pro} was studied.

Preparation of stock solutions for in vitro tests

SARS-CyCD and CyCD were respectively dissolved in DMSO to obtain a 5 mM stock solution. Enzyme stock solutions of SARS-CoV-2 M^{pro}, γ -glutamyl transferase (GGT), SARS-CoV-1 M^{pro}, Hepatitis C Virus (HCV) NS4A/NS3-3 protease, furin, and caspase-3 were prepared in distilled water.

UV and fluorescence measurements

SARS-CoV-2 M^{pro} was incubated in freshly prepared Tris buffer (20 mM Tris, pH = 7.4, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT)) at 37 °C before 10 μ M of SARS-CyCD (500 μ M DMSO stock) was added. After 60 min, UV/VIS and fluorescence measurements of the solution were performed. Fluorescence images were acquired using the IVIS spectrum imaging system with excitation at 675 $\pm$ 10 nm and emission at 720 $\pm$ 10 nm and an acquisition time of 0.1 s.

Enzyme kinetic studies

Various concentrations of SARS-CyCD (1, 5, 10, 15, 20, 40, 80, 150 μM) were incubated with SARS-CoV-2 M^{pro} (250 nM) at 37 °C for 30 min in a 100 μL system of Tris buffer (20 mM Tris, pH = 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM DTT). After incubation, the mixture was injected into HPLC (methanol/water) for quantification analyses. The initial reaction velocity ($\mu\text{mol/s}$) was calculated, plotted against SARS-CyCD concentration, and fitted to a Michaelis-Menten curve. The kinetic parameters were calculated using Michaelis-Menten equation: $V = V_{\text{max}} \times [S] / (K_m + [S])$, where V is initial velocity, and $[S]$ is substrate concentration. The calculated parameters are as follows: $V_{\text{max}} = 0.208 \mu\text{M min}^{-1}$, $K_m = 11.45 \mu\text{M}$, $k_{\text{cat}} = 0.014 \text{ s}^{-1}$, and $k_{\text{cat}}/K_m = 12.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$.

In vitro selectivity studies

SARS-CyCD was incubated with the indicated enzymes including HCV NS4A/NS3-3 protease, urokinase, caspase 3, furin, GGT, SARS-CoV-1 M^{pro} and SARS-CoV-2 M^{pro} in Tris buffer (20 mM Tris, pH = 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM DTT) at 37 °C for 120 min. Fluorescence images were acquired using the IVIS spectrum imaging system with excitation at $675 \pm 10 \text{ nm}$ and emission at $720 \pm 10 \text{ nm}$ and an acquisition time of 0.1 s.

Results and discussion

First, the optical properties were examined in the presence and absence of SARS-CoV-2 M^{pro}. When caged with the peptide substrate, SARS-CyCD showed two characteristic UV absorption peaks at 625 and 680 nm, and was initially non-fluorescent (FIG. 3a-b). After incubation with SARS-CoV-2 M^{pro}, the absorption peak at 625 nm disappeared, and a new peak appeared at 690 nm (FIG. 3a). In FIG. 3b, fluorescence measurements indicated a 50-fold fluorescence “Turn-ON” enhancement at 710 nm, with the successful uncaging of SARS-CyCD. HPLC analysis confirmed the release of CyCD ($t_R = 18 \text{ min}$) in the presence of SARS-CoV-2 M^{pro} (FIG. 3c). The enzymatic Michaelis–Menten constants (K_m) of SARS-CoV-2 M^{pro} toward SARS-CyCD was calculated to be 11.5 μM (FIG. 3d). The catalytic rate constants (k_{cat}) of SARS-CoV-2 M^{pro} toward SARS-CyCD was 0.014 s^{-1} , whereas the catalytic efficiency (k_{cat}/K_m) of SARS-CoV-2 M^{pro} toward SARS-CyCD was determined to be $12.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ (FIG. 4). More importantly, the NIRF intensity of SARS-CyCD did not show any significant increase when incubated with other enzymes, such as Hepatitis C Virus (HCV) NS4A/NS3-3 protease, urokinase, caspase-3, furin, γ -glutamyl transferase (GGT), and SARS-CoV-1 M^{pro} (FIG. 3e-f), suggesting its high specificity towards SARS-CoV-2 M^{pro}.

Example 3. Pharmacokinetics of SARS-CyCD

The pharmacokinetics of SARS-CyCD was investigated along with its uncaged fragment (CyCD).

Pharmacokinetics studies

5 Mice were anesthetized by isoflurane anesthesia for the entire duration of the experiment. The end of the tail was cut for blood extraction. Blood was sampled in heparinized capillary tubes as a reference before injection. Mice were i.t. injected with SARS-CyCD or CyCD ($2 \mu\text{mol kg}^{-1}$ body weight) and blood was sampled at 1, 20, 40, 60, 80, 120, 180, and 360 min post-injection. Collected blood samples were stored in an ice box to prevent clotting before
10 centrifugation at 4500 r.p.m for 15 min. SARS-CyCD and CyCD in the blood were quantified using the IVIS spectrum imaging system and plotted as a function of time to calculate elimination half-life value ($t_{1/2\beta}$). Calibration curves were established using different concentrations of SARS-CyCD or CyCD in blank plasma. The percentage injected dose per gram (%ID/g) was calculated as: %ID/g = dose in plasma sample/injected dose/blood weight
15 $\times 100\%$. Quantification results were presented as an absorption phase followed by an elimination phase and plotted as a function of time. A bi-exponential decay curve to estimate elimination half life $t_{1/2}$, which can be calculated as: $t_{1/2} = \ln(2)/\text{slope of elimination phase}$.

Results and discussion

20 After i.t. injection, the concentration of SARS-CyCD or CyCD in blood was analyzed. The concentration of CyCD in blood reached 0% of the injected dose (ID) at 110 min post-injection, with an elimination half-life ($t_{1/2}$) of 37-39 min, while SARS-CyCD had a slower elimination (180 mins) with a $t_{1/2}$ of approximately 52-54 min (FIG. 5a). The renal clearance efficiencies (RCE) of both SARS-CyCD and CyCD were determined by fluorescence quantification of the amount
25 of CyCDs in the urine collected from living mice as a function of time after i.t. injection. The RCE of SARS-CyCD was determined to be $59 \pm 5.0\%$ ID at 24 h post-injection, lower than that of CyCD ($79 \pm 1.5\%$ ID) (FIG. 5b). This was attributed to the higher hydrophilicity of CyCD with a lower LogD value (Table 1), relative to SARS- CyCD due to the hydrophobic peptide moiety.

30

Table 1. LogD values of SARS-Cy, SARS-CyCD and CyCD.

Probe	Log D value
SARS-Cy	3.68
SARS-CyCD	-4.21
CyCD	-5.00

Example 4. *In vivo* stability of SARS-CyCD

To investigate the *in vivo* stability of SARS-CyCD, the NIRF images and fluorescence measurements of SARS-CyCD recovered from urine were analyzed alongside the reference in PBS.

In vivo stability studies

The collected urine in PBS (10 mM, pH 7.4) was measured on a fluorescence spectrophotometer, imaged by the IVIS spectrum imaging system, and analysed by HPLC.

Results and discussion

Compared to CyCD, negligible fluorescence “Turn-ON” was observed for excreted SARS-CyCD (FIG. 6a). Furthermore, excreted SARS-CyCD showed almost identical fluorescence as the reference in PBS (FIG. 6b). In FIG. 5c, HPLC analysis of SARS-CyCD also showed an almost identical profile with the reference compound in PBS. Together, these data confirmed that SARS-CyCD underwent negligible *in vivo* metabolism in living mice.

Example 5. SARS-CyCD for *in vivo* detection of SARS-CoV-2

Because of a lack of established mouse models for studying SARS-CoV-2 lung infection, the potential of SARS-CyCD for *in vivo* detection of SARS-CoV-2 was validated in an artificial SARS-CoV-2 M^{pro}-positive mice model established via i.t. injection of commercially available recombinant SARS-CoV-2 M^{pro} as proof-of-concept (FIG. 7a). After 0.5 h injection of the protease, SARS-CyCD was administered *via* i.t. injection. Next, whole-body longitudinal NIRF imaging of the mice was conducted at different timepoints.

I.t. administration

20 µL of SARS-CyCD (10 mg mL⁻¹ in PBS) or SARS-CoV-2 M^{pro} (1 mg mL⁻¹ in PBS) was loaded into the microspray aerosolizer (purchased from PenWu, Bio Jane Trading Limited.). Then, the tip of the aerosolizer was gently inserted down the trachea of the anesthetized mice, followed by i.t. administration of aerosols into the lungs of mice.

In vivo biodistribution studies

Mice were i.t. injected with PBS (control), SARS-CyCD or CyCD (2 µmol kg⁻¹ body weight) and imaged using the IVIS spectrum imaging system at 0, 20, 40, 60, 90, and 120 min post-injection. The abdominal cavity and resected organs from mice were imaged after

ethanization at 24 h post-injection. Fluorescence images were acquired using the IVIS spectrum imaging system with excitation at 675 ± 10 nm and emission at 720 ± 10 nm.

Renal clearance studies

5 Mice were i.t injected with SARS-CyCD/CyCD ($2 \mu\text{mol kg}^{-1}$ body weight) and placed in metabolic cages. Urine was collected at 3, 6, 12, and 24 h post-injection, diluted in PBS and centrifuged at 4500 r.p.m. for 10 min and filtered by $0.22 \mu\text{m}$ syringe filter. SARS-CyCD in the urine was quantified using the IVIS spectrum imaging system and HPLC. The fluorescence spectra were measured for the urine samples. Mice were sacrificed and major organs were
10 collected, homogenized in PBS buffer (10 mM, pH 7.4), and centrifuged at 4500 r.p.m for 15 min to remove insoluble components. The supernatant containing extracted molecules were taken for fluorescence measurements using the IVIS spectrum imaging system.

Real-time in vivo NIRF imaging of SARS-CoV-2 M^{pro} in living mice

15 Real-time NIRF imaging was conducted at $t = 0, 10, 20, 30, 40, 60, 90,$ and 120 min after sequential i.t injection of SARS-CoV-2 M^{pro} (0.15 mg kg^{-1} body weight), followed by SARS-CyCD ($2 \mu\text{mol kg}^{-1}$ body weight). Fluorescence images were acquired using the IVIS spectrum imaging system. NIRF intensities of lungs and bladder in living mice were analyzed by the ROI analysis using the Living Image 4.3 Software.

20

Urinalysis of SARS-CoV-2 M^{pro} in living mice

Urine was collected using metabolic cages from living mice after sequential i.t injection of SARS-CoV-2 M^{pro} (0.15 mg kg^{-1} body weight), followed by SARS-CyCD ($2 \mu\text{mol kg}^{-1}$ body weight) at 3, 6, 12, 24 h post i.t. injection. The collected urine samples were centrifuged at
25 4500 r.p.m. for 8 min, filtered by $0.22 \mu\text{m}$ syringe filter, and measured using the IVIS spectrum imaging system with excitation at 675 ± 10 nm and emission at 720 ± 10 nm.

Cell viability

NIH3T3 cells were seeded in 96-well plates (104 cells per well) for 24 h, and then different
30 concentrations of SARS-CyCD (0, 10, 25, 50, 80 and $100 \mu\text{g/mL}$) were added. Cells were incubated with or without (control) SARS-CyCD, followed by addition of MTS assay ($100 \mu\text{L}$, 0.1 mg/mL) for 4 h. The absorbance of MTS at 490 nm was measured using a microplate reader. The cytotoxic effects (VR) of SARS-CyCD were assessed using the following equation:
VR = $A/A_0 \times 100\%$, where A and A_0 are the absorbance of the experimental group and control
35 group, respectively. The assays were performed in triplicates for each concentration.

Results and discussion

To study residual SARS-CyCD in the body after 24 h urinary recovery, the mice were dissected and major organs were homogenized in PBS. Extraction of SARS-CyCD from each organ and subsequent analysis were conducted. Residual SARS-CyCD mainly accumulated in the lungs (ca. 16% ID), liver (ca. 14% ID), and kidneys (ca. 5% ID). In the other organs, negligible amounts of SARS-CyCD were found (FIG. 5d).

Strong NIRF signals were observed in the lungs (both dorsal and ventral views) of SARS-CoV-2 M^{pro}-positive mice (FIG. 7b). As for the control group without SARS-CoV-2 M^{pro}, negligible NIRF signals were observed. NIRF signals in the lungs of SARS-CoV-2 M^{pro}-positive mice further increased by up to 5.0-fold (ventral) and 7.5-fold (dorsal) at 60 min post i.t. injection (FIG. 7c). However, the negative control group showed only a 2.2-fold (ventral) and 3.5-fold (dorsal) increase in NIRF fluorescence intensity (FIG. 8a). Furthermore, *ex vivo* fluorescence analysis of lungs collected from SARS-CoV-2 M^{pro}-positive mice co-injected with saline and SARS-CyCD after 24 h post i.t. injection indicated 5-fold enhancement (FIG. 8b). These results clearly indicated that the cleavage of SARS-CyCD occurred in SARS-CoV-2 M^{pro}-positive mice, giving rise to a gradual increase in NIRF “Turn-ON” signals from uncaged CyCD.

Due to the rapid and efficient renal clearance of the cleaved product (CyCD), the bladder was easily delineated through NIRF imaging at 60 min post-injection of SARS-CyCD (FIG. 7b); note that the bladders of the control mice had negligible NIRF signals (FIG. 7b). Furthermore, a 2.2-fold increase in NIRF fluorescence intensity was observed in the bladders of SARS-CoV-2 M^{pro}-positive mice, compared to the saline-treated control group (FIG. 7d). Optical urinalysis was also carried out by fluorescence measurement of SARS-CyCD in the urine of mice at various timepoints after i.t. injection. The NIRF signal for the urine of SARS-CoV-2 M^{pro}-positive mice was always higher than that of the control mice at all timepoints. In particular, a 3.0-fold maximum significant NIRF difference relative to the control group was observed 3 h post i.t. injection (FIG. 7e-f). Thus, SARS-CyCD-based optical urinalysis was validated as a potential way for specific detection of SARS-CoV-2 infection.

Therefore, we have developed a M^{pro}-activatable NIRF probe (SARS-CyCD) for *in vivo* detection of SARS-CoV-2. Our *in vitro* results demonstrated that this probe is specifically activated in the presence of SARS-CoV-2 M^{pro} for fluorescence signal “Turn-ON”. Pharmacokinetics studies in Example 3 indicated that SARS-CyCD has high renal clearance efficiency and minimal *in vivo* metabolism. Such hemicyanine-based probes have also been used in animal models and showed no cytotoxicity (FIG. 9). Upon i.t. administration, SARS-CyCD was effectively cleaved by SARS-CoV-2 M^{pro} in the lungs of living mice to liberate its

fluorescent fragment (CyCD), followed by excretion *via* kidneys. Thus, SARS-CyCD can detect SARS-CoV-2 infection through optical urinalysis, showing its high potential for clinical translation.

Claims

1. A compound of formula I:



wherein:

A represents a peptide group that targets a viral protease;

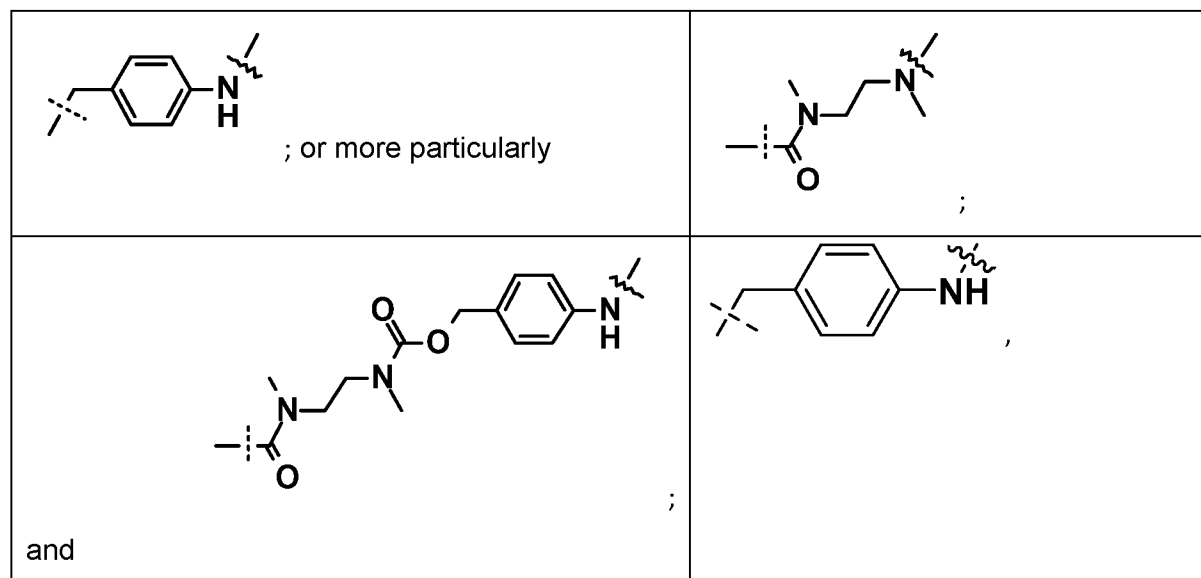
B, when present, represents a self-immolative linking group;

C represents a fluorophore or chemilumiphore group covalently linked to the rest of the molecule by an oxygen atom or a NH group; and

n represents 0 or 1,

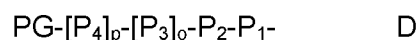
or a pharmaceutically acceptable salt or solvate thereof.

2. The compound according to Claim 1, or a pharmaceutically acceptable salt or solvate thereof, wherein B, when present is selected from:



where the wavy line represents the point of attachment to A and the dotted line represents the point of attachment to C.

3. The compound according to Claim 1 or Claim 2, or a pharmaceutically acceptable salt or solvate thereof, wherein A represents a peptide sequence according to fragment D that targets a flavivirus protease:



where:

P₁ to P₄, when present, are amino acids that together form an oligo peptide;

o is 0 or 1;

p is 0 or 1; and

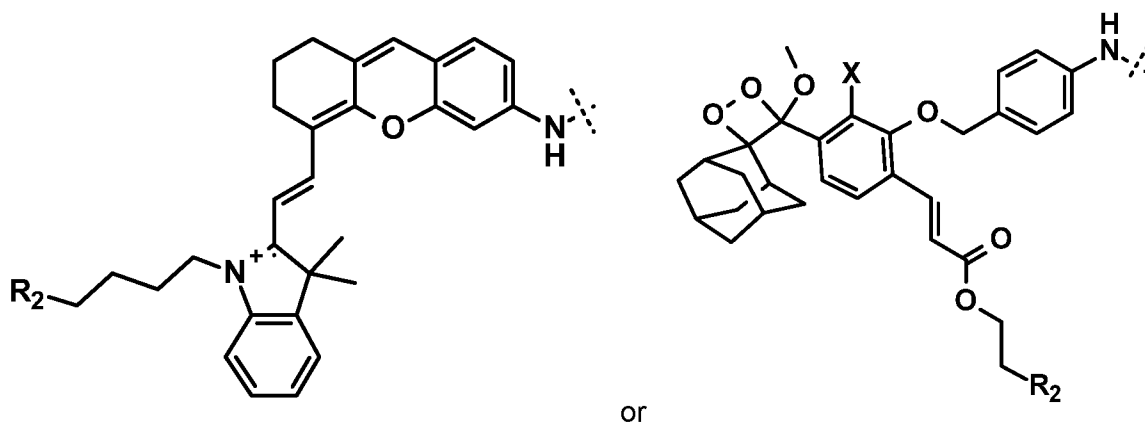
PG is a protecting group.

4. The compound according to Claim 3, or a pharmaceutically acceptable salt or solvate thereof, wherein PG is selected from a benzoyl protecting group, an acetyl protecting group or a *tert*-butoxycarbonyl protecting group.

5. The compound according to Claim 3 or Claim 4, or a pharmaceutically acceptable salt or solvate thereof, wherein fragment D is selected from the list of:

- (a) Bz-Nle-Lys-Arg-Arg-;
- (b) Bz-Nle-Lys-Arg-Ala-;
- (c) Bz-Nle-Lys-Thr-Arg-;
- (d) Bz-Nle-Thr-Arg-Arg-;
- (e) Bz-Nle-Ala-Arg-Arg-;
- (f) Bz-Nle-Lys-Arg-Phe-;
- (g) Bz-Nle-Lys-Phe-Arg-;
- (h) Bz-Nle-Phe-Arg-Arg-;
- (i) Bz-Thr-Lys-Arg-Arg-;
- (j) Bz-Thr-Thr-Arg-Arg-;
- (k) Bz-Phe-Lys-Arg-Arg-;
- (l) Bz-Ala-Lys-Arg-Arg-;
- (m) Bz-Nle-Lys-Lys-Arg-;
- (n) Bz-Lys-Arg-Arg-;
- (o) Bz-Arg-Arg-;
- (p) Ac-Nle-Lys-Arg-Arg-;
- (q) Ac-Lys-Arg-Arg-;
- (r) Boc-Gly-Arg-Arg-; and
- (s) Boc-Lys-Arg-Arg-.

6. The compound according to any one of Claims 3 to 5, or a pharmaceutically acceptable salt or solvate thereof, wherein C is selected from:

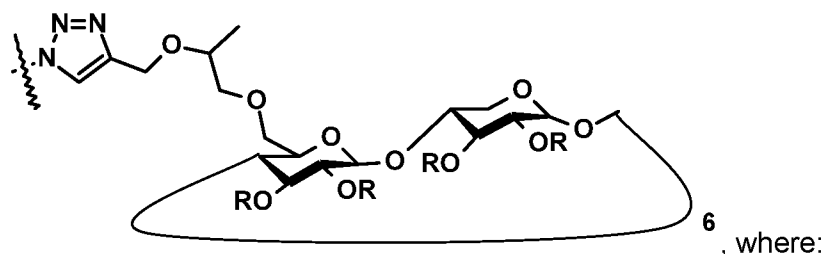
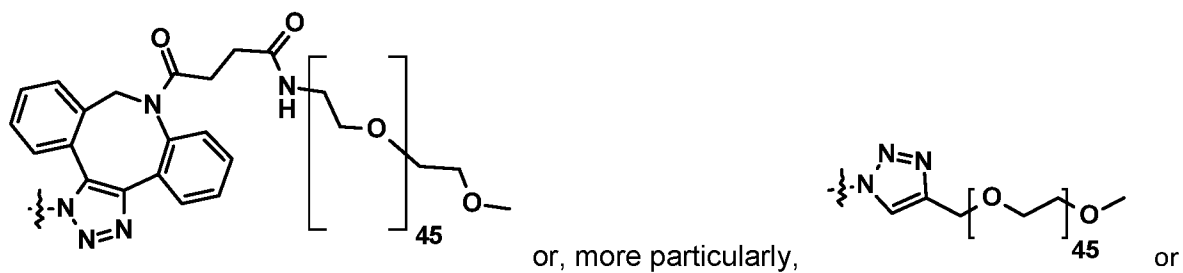


where:

the dotted line represents the point of attachment to the rest of the molecule;

X represents H or halo (e.g. Cl); and

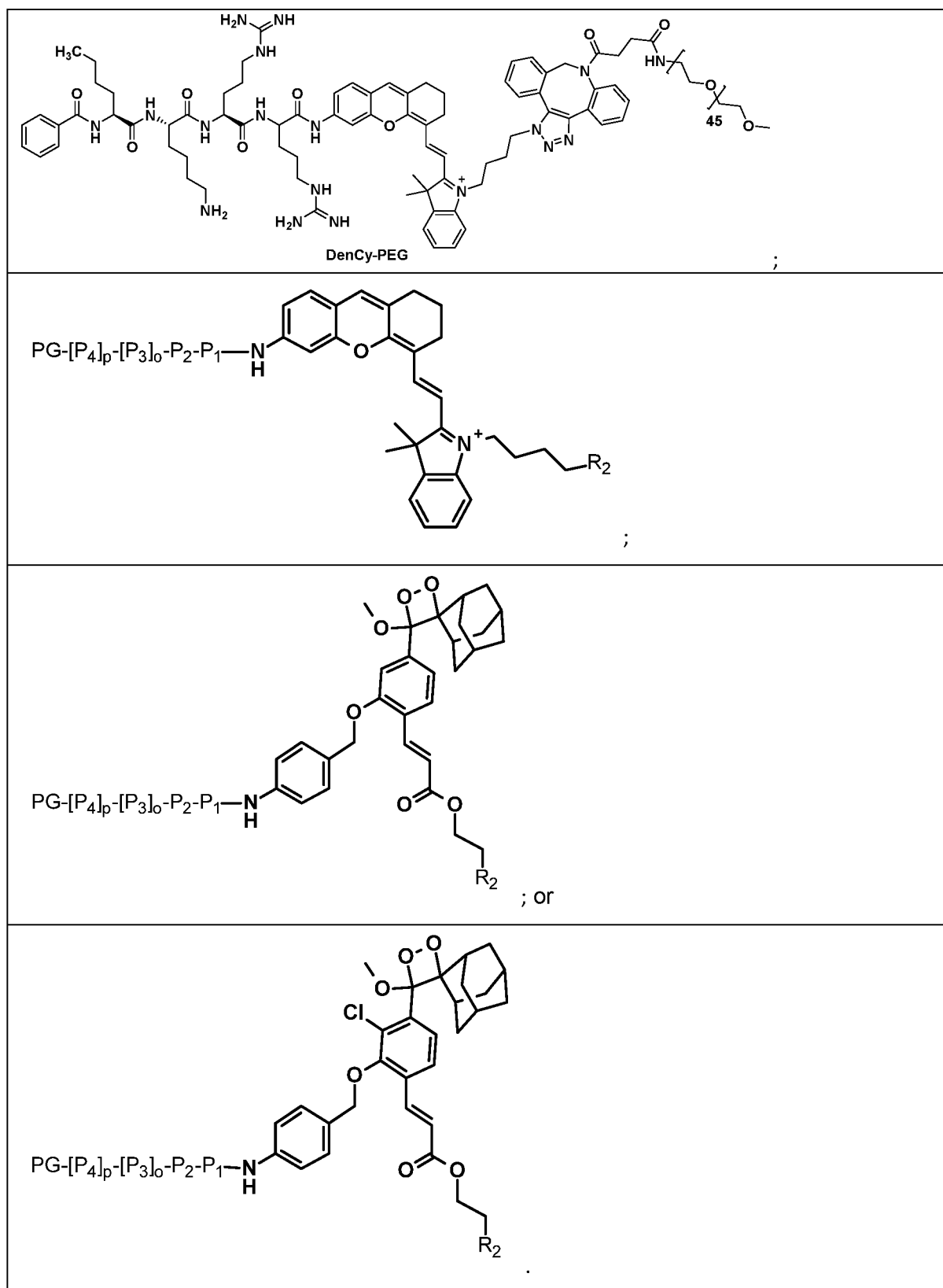
R₂ is selected from:



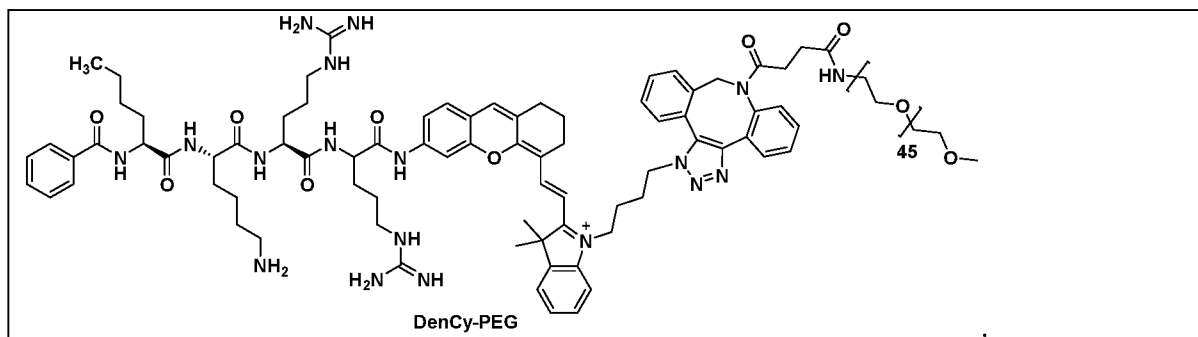
the wiggly line represents the point of attachment to the rest of the molecule; and

R represents H or 2-hydroxypropyl.

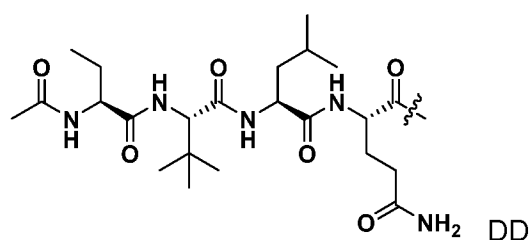
7. The compound according to Claim 6, or a pharmaceutically acceptable salt or solvate thereof, wherein the compound of formula I is selected from:



8. The compound according to Claim 7, or a pharmaceutically acceptable salt or solvate thereof, wherein the compound of formula I is:

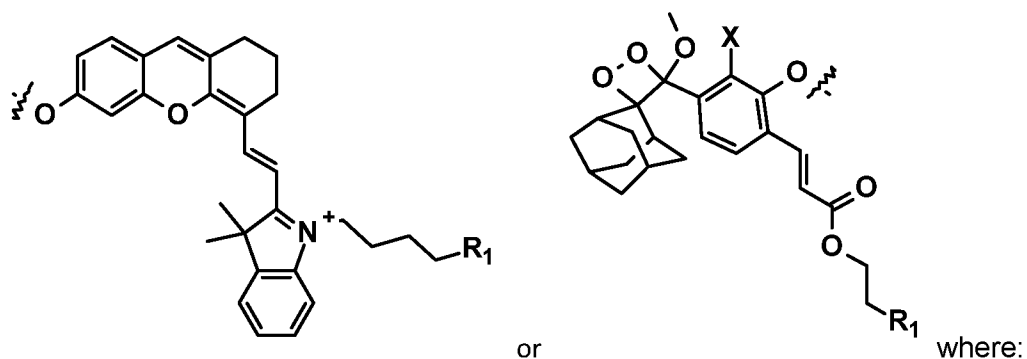


9. The compound according to Claim 1 or Claim 2, or a pharmaceutically acceptable salt or solvate thereof, wherein A represents a peptide sequence according to fragment DD that targets SARS-CoV-2 main protease:



where the wavy line represents the point of attachment to the rest of the molecule.

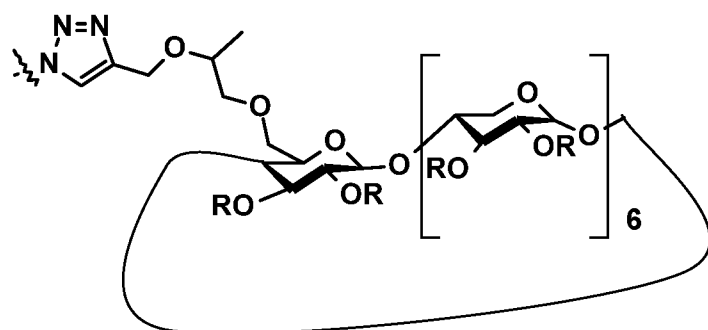
10. The compound according to Claim 9, or a pharmaceutically acceptable salt or solvate thereof, wherein C is selected from:



the wiggly line represents the point of attachment to the rest of the molecule;

X represents H or halo (e.g. Cl); and

R₁ represents:

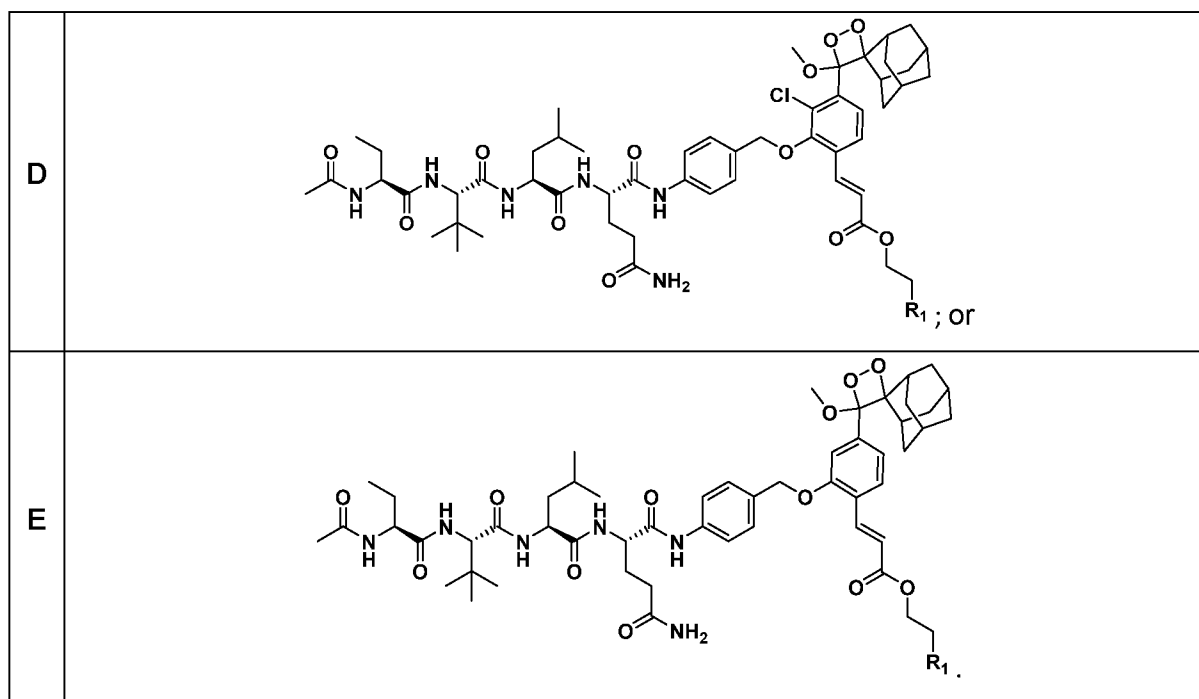


, where:

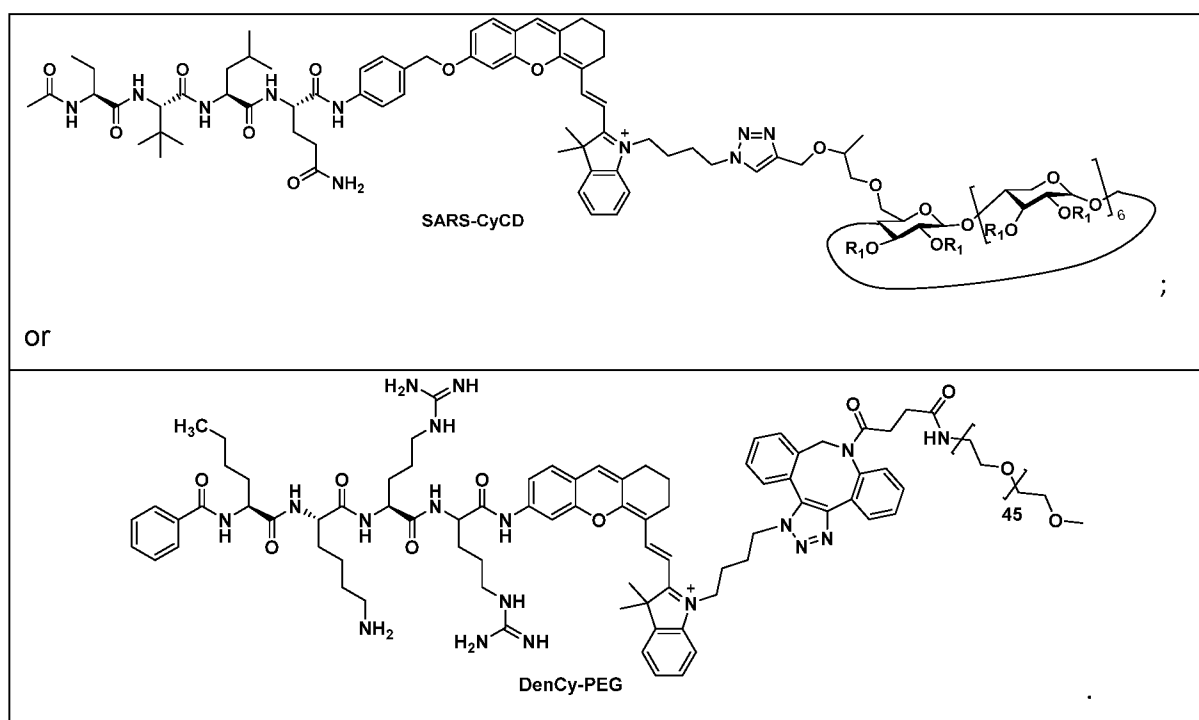
the wiggly line represents the point of attachment to the rest of the molecule; and
R represents H or 2-hydroxypropyl.

11. The compound according to Claim 9 or Claim 10, or a pharmaceutically acceptable salt or solvate thereof, wherein the compound of formula I is selected from:

A	
B	
C	



12. The compound according to Claim 1, or a pharmaceutically acceptable salt or solvate thereof, wherein the compound of formula I is selected from:



13. A method for detection of a viral protease in an analyte, the method comprising the following steps:

- (a) providing an analyte and a fluid comprising a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof, according to any one of Claims 1 to 12;

- (b) contacting the analyte with the fluid comprising a compound of formula I, or pharmaceutically acceptable salt or solvate thereof, for a period of time; and
 - (c) after the period of time detecting any near-infrared fluorescence (NIRF) or chemiluminescence, wherein the presence of the viral protease in the fluid comprising the analyte and the compound of formula I, or pharmaceutically acceptable salt or solvate thereof, is indicated by near-infrared fluorescence (NIRF) or chemiluminescence.
14. A method for detection of a viral protease *in vivo*, the method comprising the following steps:
- (ai) administering a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof, according to any one of Claims 1 to 12 to a subject; and
 - (aii) detecting any near-infrared fluorescence (NIRF) or chemiluminescence, wherein the presence of the viral protease *in vivo* is indicated by near-infrared fluorescence (NIRF) or chemiluminescence.
15. Use of a compound of formula I or a salt and/or solvate thereof according to any one of Claims 1 to 12 in the manufacture of a diagnostic agent for *in vivo* diagnosis of a disease caused by a virus.
16. A compound of formula I or a salt and/or solvate thereof according to any one of Claims 1 to 12 for use in the *in vivo* diagnosis of a disease caused by a virus.
17. A composition comprising a compound of formula I, or a pharmaceutically acceptable salt or solvate thereof, according to any one of Claims 1 to 12 in admixture with one or more of a pharmaceutically acceptable adjuvant, diluent and carrier.

Figures

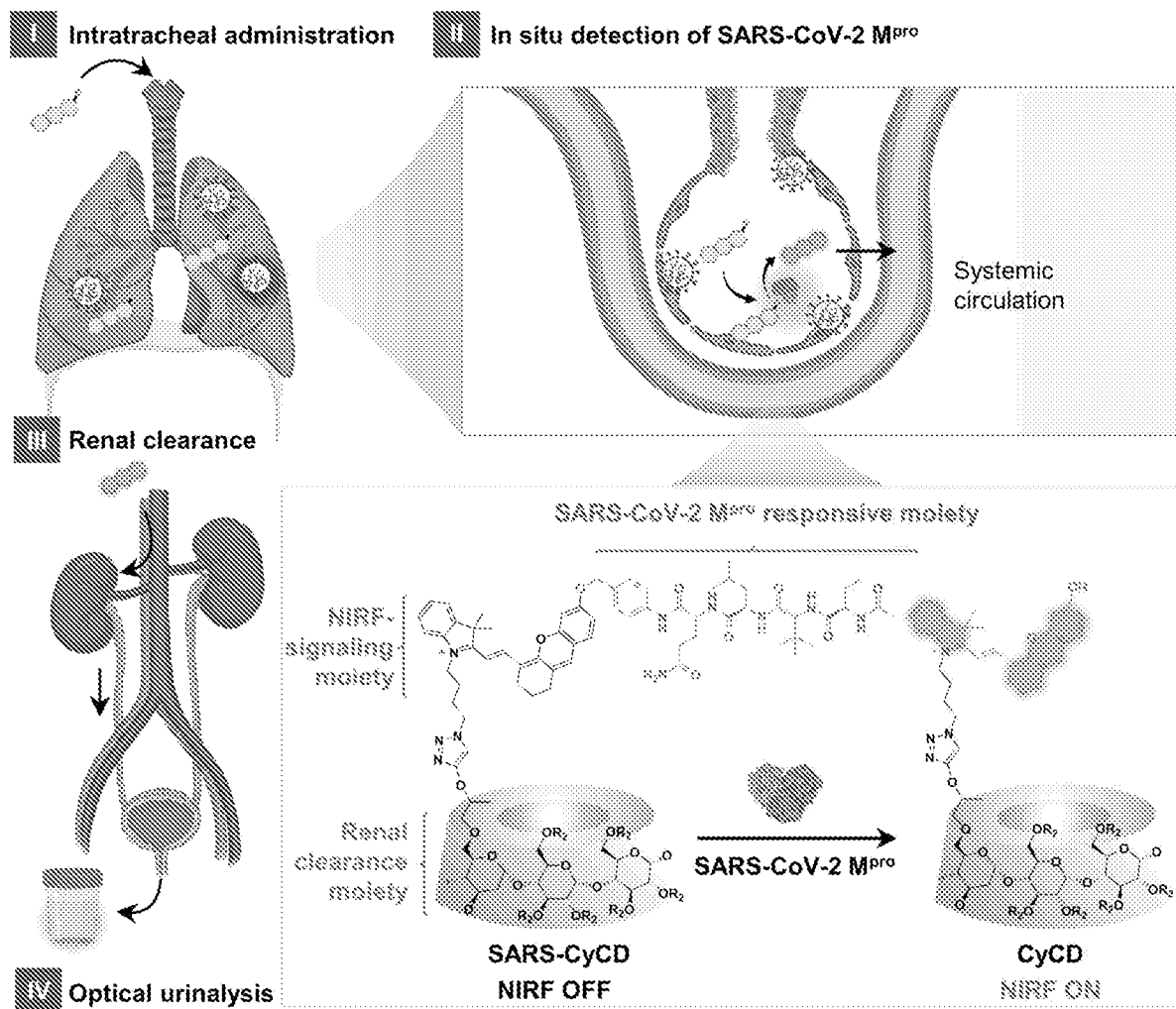
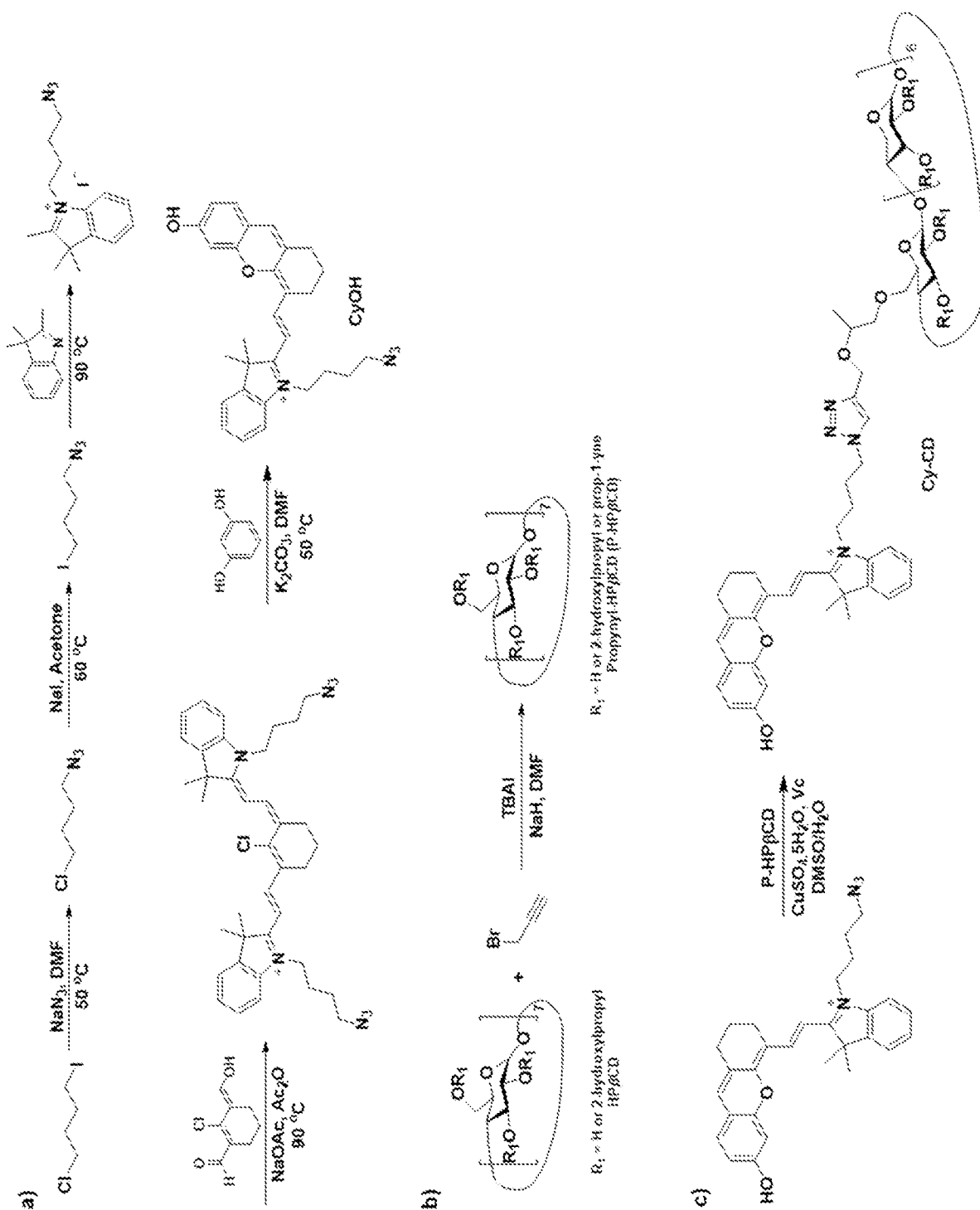


FIG. 1



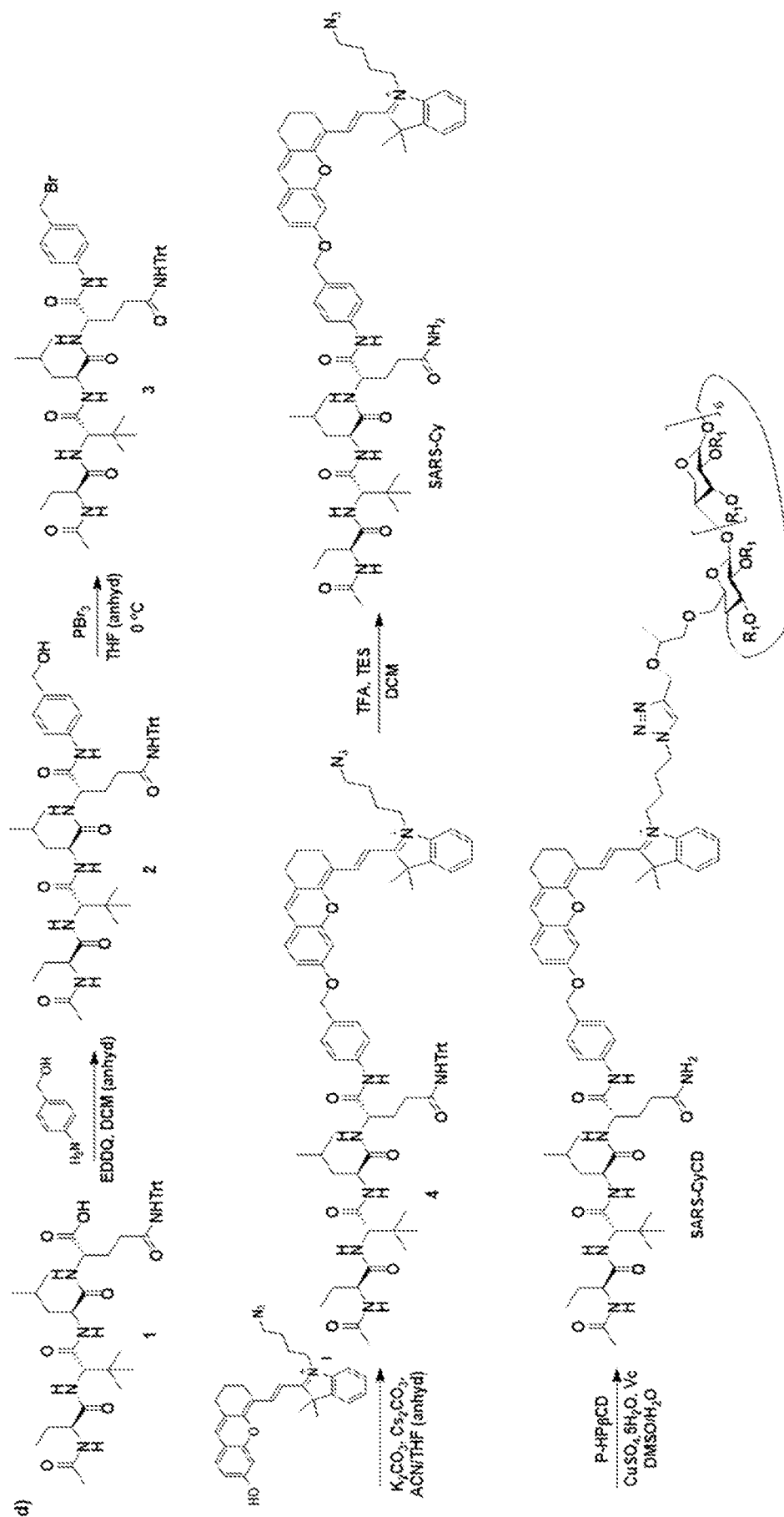


FIG. 2d

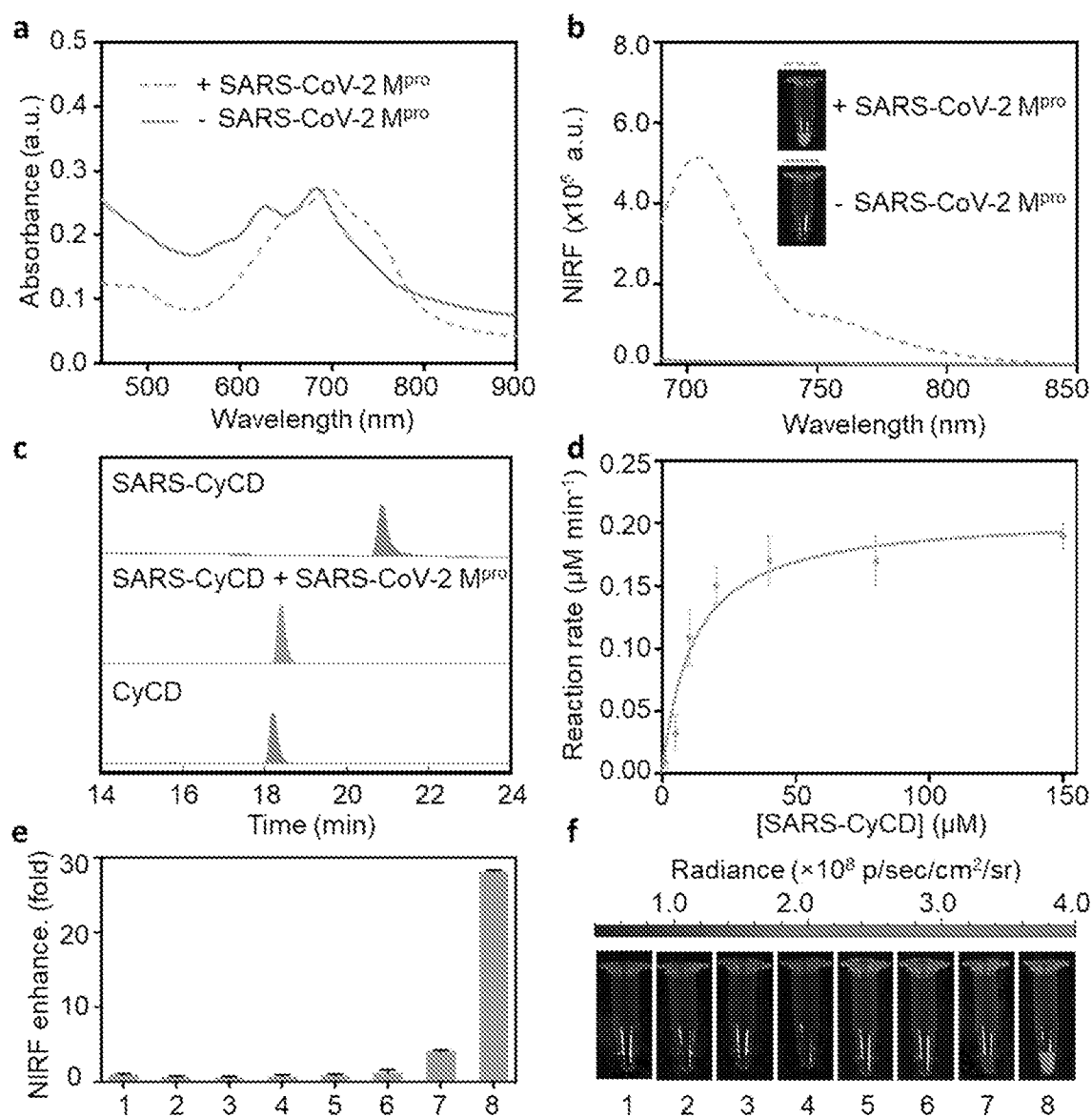
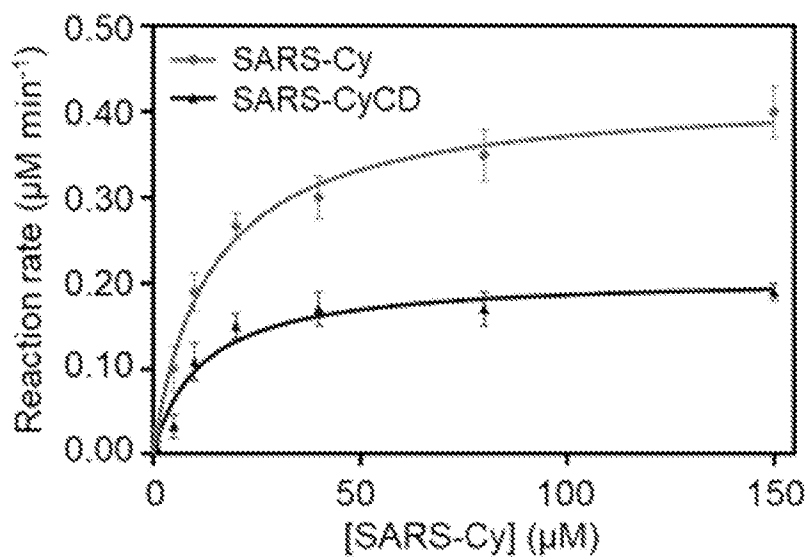


FIG. 3



	SARS-Cy	SARS-CyCD
Michaelis-Menten Constants (K_m)	13.5 μM	11.5 μM
Catalytic Rate Constants (k_{cat})	0.028 s^{-1}	0.014 s^{-1}
Catalytic Efficiency (k_{cat}/K_m)	$20.7 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$	$12.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$

FIG. 4

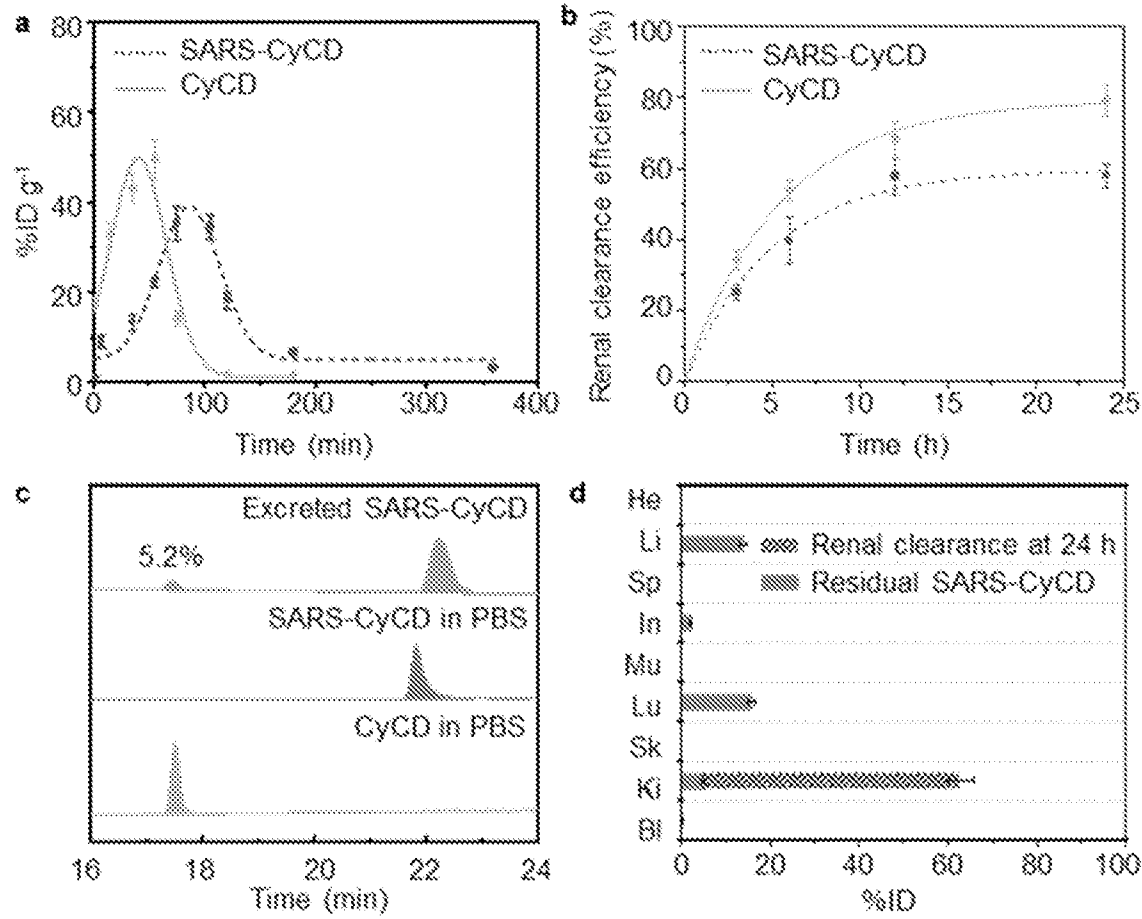


FIG. 5

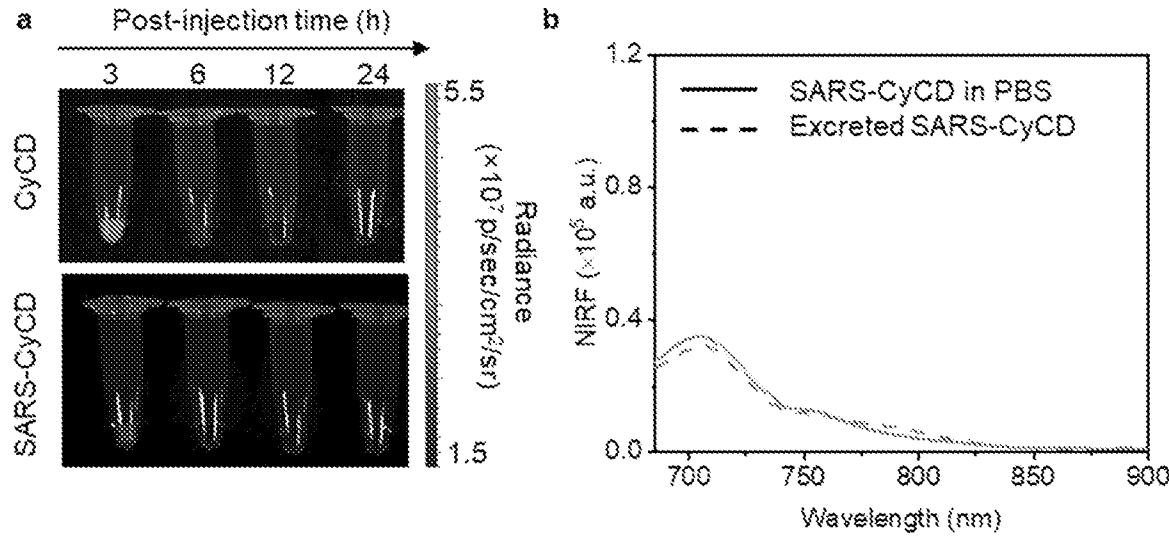


FIG. 6

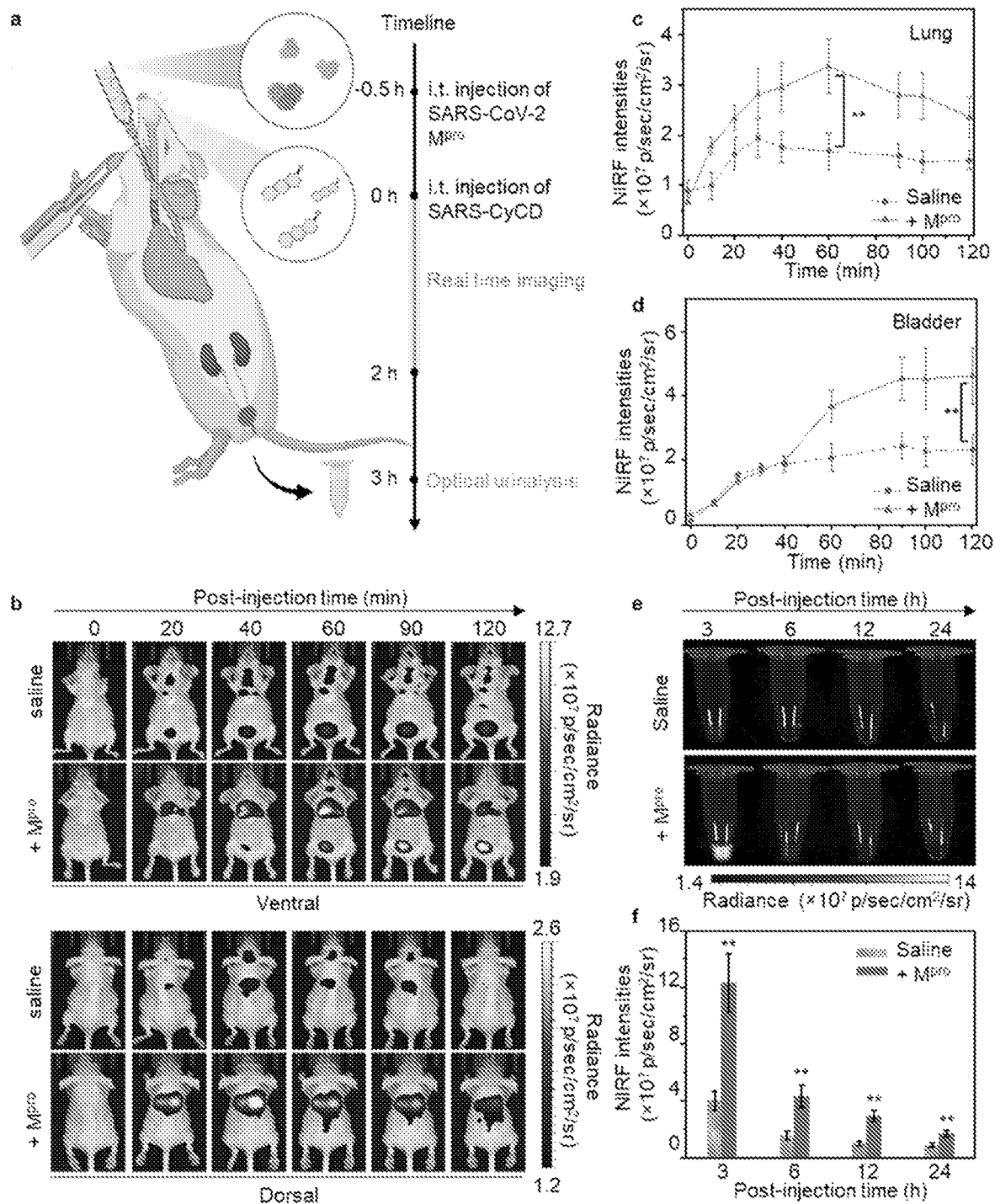


FIG. 7

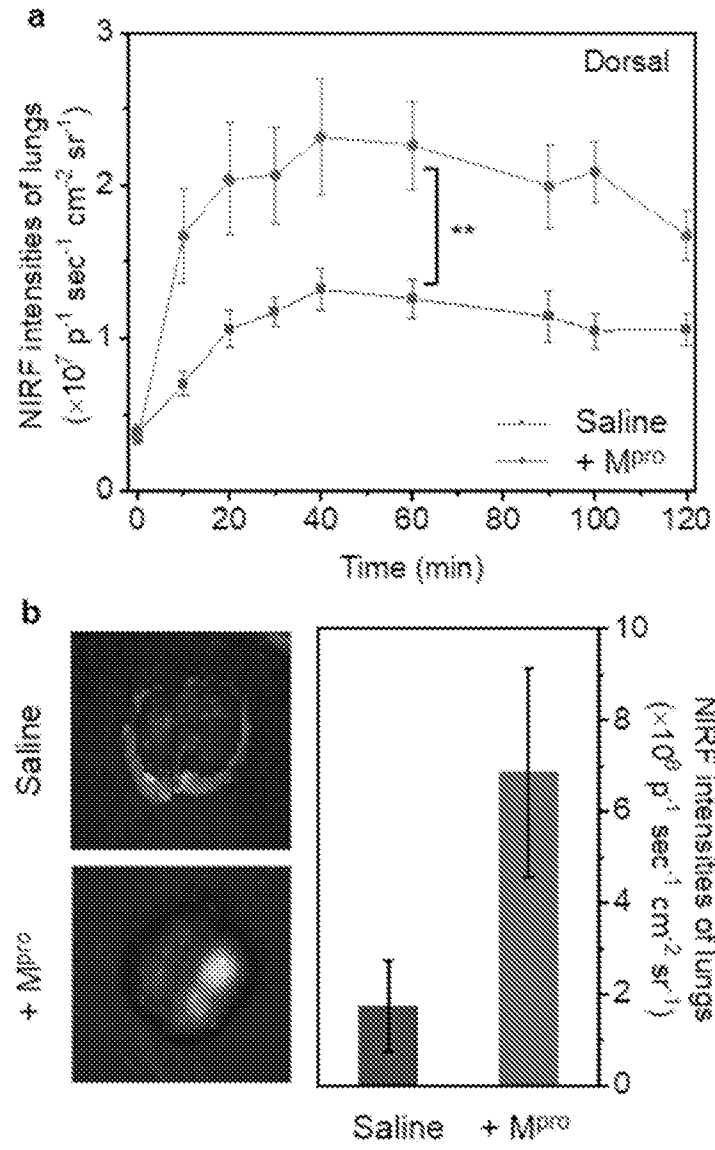


FIG. 8

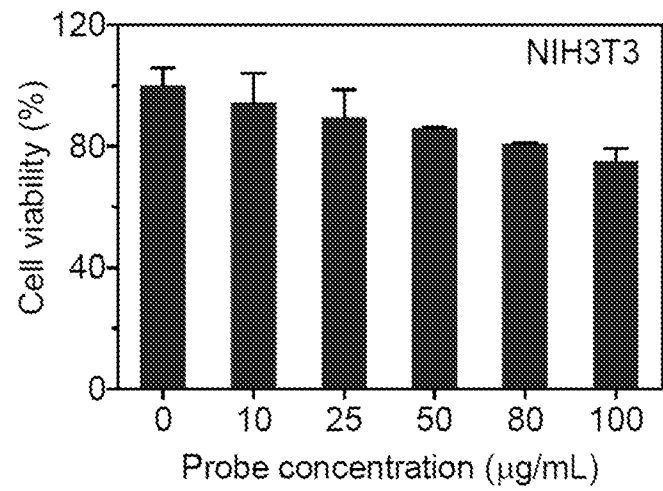


FIG. 9

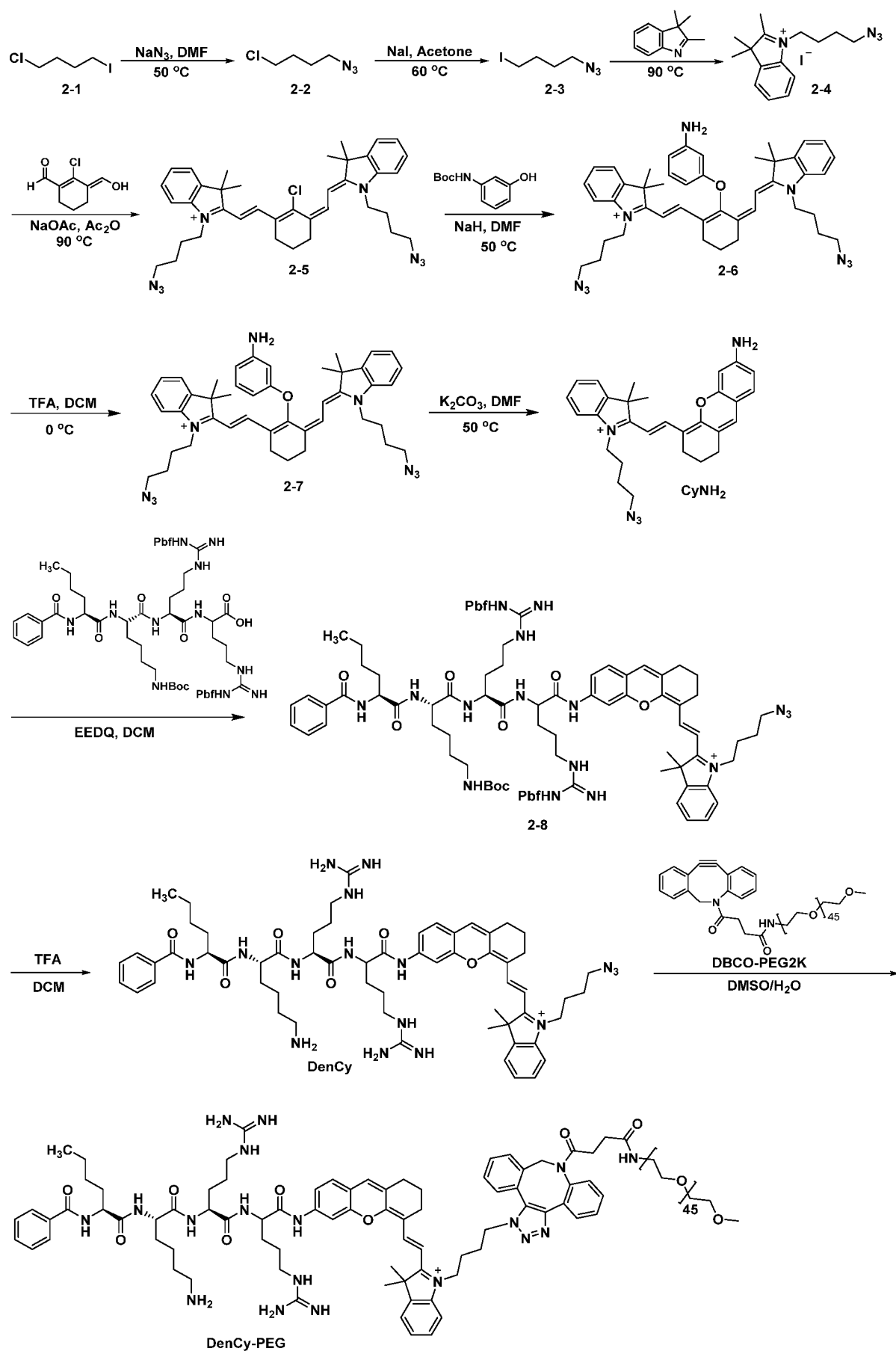


FIG. 10