

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 March 2006 (30.03.2006)

PCT

(10) International Publication Number
WO 2006/032893 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/GB2005/003654

(22) International Filing Date:
22 September 2005 (22.09.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
04104605.3 22 September 2004 (22.09.2004) EP

(74) Agents: MARSHALL, Cameron, John et al.; Carpmals & Ransford, 43-45 Bloomsbury Square, London WC1A 2RA (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(71) Applicant (for all designated States except US): OXFORD GENE TECHNOLOGY IP LIMITED [GB/GB]; Begbroke Business & Science Park, Hirsch Building, Sandy Lane, Yarnton, Oxford OX5 1PF (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): SHCHEPINOV, Mikhail, Sergeevich [RU/GB]; Oxford Gene Technology, Begbroke Business & Science Park, Hirsch Building, Sandy Lane, Yarnton, Oxford OX5 1PF (GB). SOUTHERN, Edwin, Mellor [GB/GB]; Oxford Gene Technology, Begbroke Business & Science Park, Hirsch Building, Sandy Lane, Yarnton, Oxford OX5 1PF (GB). KORSHUN, Vladimir, A. [RU/GB]; Oxford Gene Technology, Begbroke Business & Science Park, Hirsch Building, Sandy Lane, Yarnton, Oxford OX5 1PF (GB).

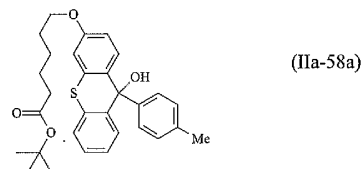
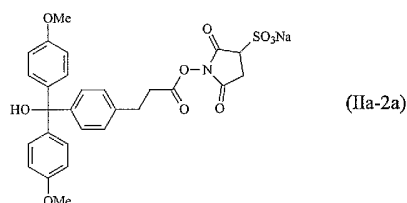
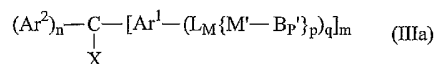
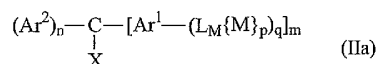
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: TRITYL DERIVATIVES FOR ENHANCING MASS SPECTROMETRY



(57) Abstract: The present invention provides a compound of the formula: (IIa); having a reactive functional group M, capable of reacting with a biopolymer, BP, having at least one group capable of reacting with M to form a covalent linkage, to provide a biopolymer derivative of the formula: (IIIa). The biopolymer derivatives of the invention have enhanced ionisability with respect to free 10 biopolymer (Bp) enabling improved analysis of the biopolymer using mass spectrometry. The invention further provides specific examples of compounds formula (IIa), e.g. compounds of formula: (IIa-2a) and (IIa-58a).

TRITYL DERIVATIVES FOR ENHANCING MASS SPECTROMETRY

All documents cited herein are incorporated by reference in their entirety.

TECHNICAL FIELD

This invention relates to compounds useful in mass spectrometry. In particular, it relates to compounds and solid supports useful in the methods of international patent application WO2005/057207. The invention further relates to derivatised biopolymers and ions obtainable therefrom.

BACKGROUND OF THE INVENTION

Mass spectrometry is a versatile analytical technique possessing excellent detection range and speed of detection with respect to High Performance Liquid Chromatography (HPLC), Gas Chromatography (GC), Infra-Red (IR) and Nuclear Magnetic Resonance (NMR).

However, many biopolymers, such as carbohydrates and proteins, are difficult to analyse using mass spectrometry due to significant difficulties in ionising the biopolymer, even using Matrix Assisted Laser Desorption/Ionisation Time Of Flight (MALDI-TOF) techniques. Despite the considerable resolving power of 2D-PAGE, this technology has fallen far short of the ultimate goal of displaying the whole proteome in a single experiment, as many proteins are resistance to 2D-PAGE analysis (*e.g.* those with low or high molecular masses, membrane proteins, proteins with extreme isoelectric points, *etc.*). Many proteins are thus invisible to 2-D PAGE [Cravatt & Sorensen (2000) Current Opinion in Chemical Biology vol. 4, p. 663-668].

WO2005/057207 discloses methods for improving ionisation of biopolymers, thus allowing improved analysis of biopolymers by mass spectrometry and analysis of biopolymers which may be otherwise difficult or impossible to analyse using known mass spectrometry techniques.

However, there remains a need for new and improved compounds for enhancing mass spectrometry which are useful in the methods of WO2005/057207.

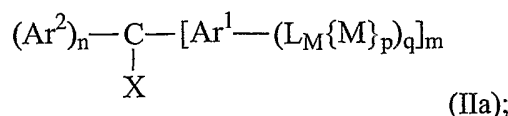
DISCLOSURE OF THE INVENTION

It has been found that covalent attachment of trityl derivatives to biopolymers can improve the ionisation properties of the biopolymer. The invention provides compounds of formulae (IIa) and (IIb) which may be reacted with a biopolymer in the methods of WO2005/057207 to provide biopolymers derivatised as specified in formulae (IIIa) and (IIIb). The biopolymer derivatives of formulae (IIIa) and (IIIb) can be readily ionised to form ions of formula (I), which are particularly suitable for mass spectrometry analysis.

Whereas triphenylmethyl derivatives covalently attached to certain biopolymers (*e.g.* DNA) are known in the prior art [*e.g.* Chem. Soc. Rev. (2003) 32, p. 3-13], the prior art attaches the polymer to the α -triphenylmethyl carbon atom through a non-aromatic linker. In contrast, under the present invention the biopolymer is attached to the α -triarylmethyl carbon atom via an aromatic group

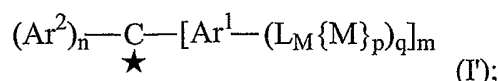
adjacent to the central carbon atom. Consequently, ionisation of the prior art derivatives results in separation of the triphenylmethyl derivative and the biopolymer, whereas according to the present invention the biopolymer remains bound to the trityl derivative on ionisation, thereby allowing analysis of the biopolymer by mass spectrometry.

- 5 In a first aspect of the invention, there is therefore provided a compound of formula (IIa):



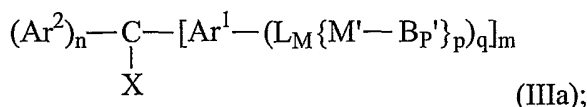
where:

X is a group capable of being cleaved from the a-carbon atom to form an ion of formula (I')



- 10 C $\star$ is a carbon atom bearing a single positive charge or a single negative charge;
M is independently a reactive functional group;
Ar¹ is independently an aromatic group or an aromatic group substituted with one or more A;
Ar² is independently an aromatic group or an aromatic group substituted with one or more A;
optionally wherein (a) two or three of the groups Ar¹ and Ar² are linked together by
15 one or more L⁵, where L⁵ is independently a single bond or a linker atom or group; and/or (b)
two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group
substituted with one or more A;
A is independently a substituent;
L_M is independently a single bond or a linker atom or group;
20 n = 0, 1 or 2 and m = 1, 2, or 3, provided the sum of n+m = 3;
p independently = 1 or more; and
q independently = 1 or more.

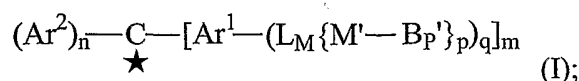
- The compounds of formula (IIa) may be employed in the methods of WO2005/057207 (*e.g.* of
claims 1, 2, 15 or 16) by reacting them with a biopolymer, B_P, having at least one group capable of
25 reacting with M to form a covalent linkage, to provide a biopolymer derivative of the formula (IIIa):



where:

- X, Ar¹, Ar², L_M, n, m, p and q are defined above;
B_P' is independently the biopolymer residue of B_P produced on formation of the covalent
30 linkage; and
M' is independently the residue of M produced on formation of the covalent linkage.

The C—X bond between X and the a-carbon atom of the derivative of formula (IIIa) may be cleaved to form an ion of formula (I):

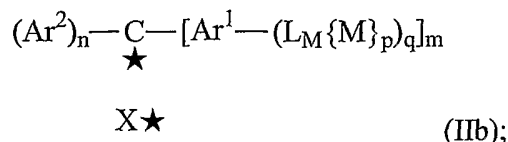


where:

B_P' , M' , Ar^1 , Ar^2 , L_M , n , m , p and q are defined above; and

$\text{C}\star$ is a carbon atom bearing a single positive charge or a single negative charge.

- 5 In a second aspect of the invention, there is provided a compound of formula (IIb):

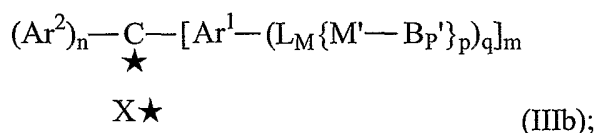


where:

$\text{X}\star$ is a counter-ion to $\text{C}\star$;

and $\text{C}\star$, M , Ar^1 , Ar^2 , L_M , n , m , p and q are as defined above.

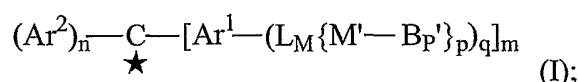
- 10 The compounds of formula (IIb) may be employed in the methods of WO2005/057207 (*e.g.* of claims 1, 2, 15 or 16) by reacting them with a biopolymer, B_P , having at least one group capable of reacting with M to form a covalent linkage, to provide a biopolymer derivative of the formula (IIIb):



where:

- 15 $\text{C}\star$, $\text{X}\star$, Ar^1 , Ar^2 , L_M , n , m , p , q , B_P' and M' are defined above.

The counter-ion $\text{X}\star$ may be dissociated from the derivative of formula (IIIb) to form an ion of formula (I):



where:

- 20 $\text{C}\star$, B_P' , M' , Ar^1 , Ar^2 , L_M , n , m , p and q are defined above.

- In a third aspect of the invention, there is provided biopolymer derivatives of the formula (IIIa) or (IIIb), as defined above. The biopolymer derivatives of the invention have enhanced ionisability with respect to free biopolymer, B_P . Advantageously, the biopolymer derivatives may not require a matrix (*e.g.* as used in MALDI-MS) in order to elicit ionisation, although a matrix may help to enhance ionisation. Preferably, ionisation may be obtained without requiring acid treatment, in particular by direct laser illumination.
- 25

- In a fourth aspect of the invention, there is provided ions of formula (I), as defined above. These ions are stabilised by the resonance effect of the aromatic groups Ar^1 and Ar^2 . Electron-withdrawing groups, when $\text{C}\star$ is an anion, or electron-donating groups, when $\text{C}\star$ is a cation, may optionally be provided on Ar^1 and/or Ar^2 to assist this resonance effect. Consequently, the biopolymer derivatives of the invention readily form ions of formula (I) relative to the native biopolymer, B_P .
- 30

The ions of formula (I) are generally only ever seen on a mass spectrum with a single charge, which is advantageous since it reduces cluttering of the mass spectrum.

The invention provides compounds of the formulae (IIa) and (IIb), as defined above, which are useful for forming ions of formula (I). As the difference in the molecular mass of the ions of formula (I) and that of the free biopolymer can be accurately calculated, the derivatised compounds of the invention allow analysis of the biopolymer B_P, which may be otherwise difficult or impossible to analyse using known mass spectrometrical techniques.

The compounds of formulae (IIa) and (IIb) may form ions of formula (I') by either cleaving the C-X bond between X and the α-carbon atoms in the case of the compounds of formula (IIa) or dissociating X★ in the case of compounds of formula (IIb).

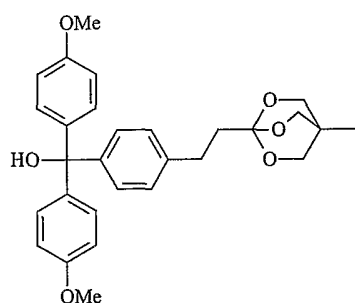
In a fifth aspect of the invention, ions of formula (I'), as defined above, are provided. Ions of formula (I') are stabilised by the resonance effect of the aromatic groups Ar¹ and Ar². Electron-withdrawing groups, when C★ is an anion, or electron-donating groups, when C★ is a cation, may optionally be provided on Ar¹ and/or Ar² to assist this resonance effect.

The compounds of formulae (IIa) and (IIb) are useful in the methods disclosed in WO2005/057207, claiming priority from UK patent application GB 03 284 14.8. The invention therefore provides the methods of WO2005/057207, *e.g.* of claims 1, 2, 15 or 16, comprising a compound of formula (IIa) or (IIb) disclosed herein.

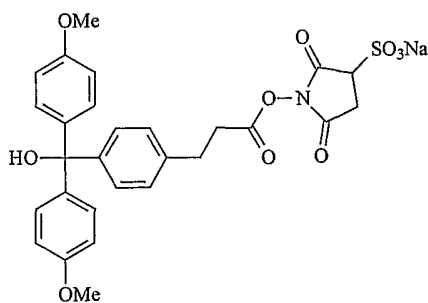
Other advantageous features of the compounds of the invention include more uniformity of the signal intensity between different analytes (useful for quantitative studies) and similar desorption properties between compounds with different, but close, masses, so that techniques such as isotope coded affinity tagging (ICAT) can be employed with the compounds of the invention.

The homogeneous methods of the invention are particularly appropriate for small molecules, *e.g.* amines.

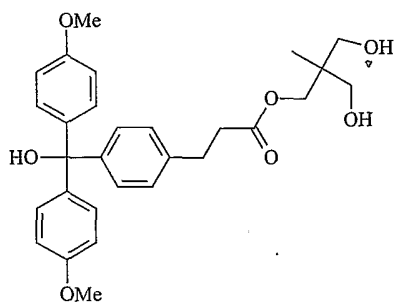
In a sixth aspect of the invention, there is provided compounds of formula (IIa) of formulae (IIa-1a) to (IIa-69):



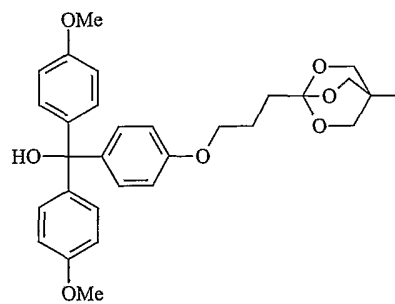
(IIa-1a)



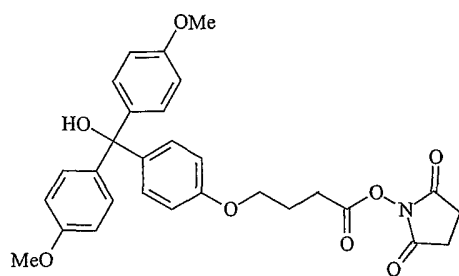
(IIa-2a)



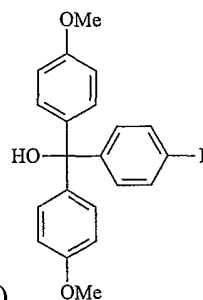
(IIa-3a)



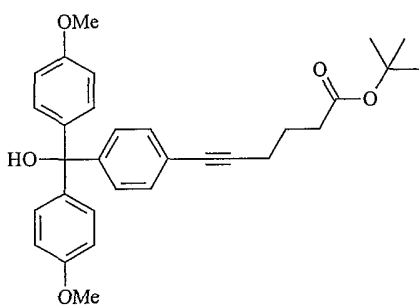
(IIa-4a)



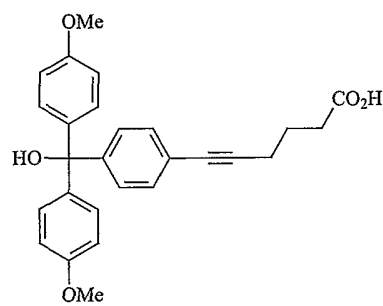
(IIa-5a)



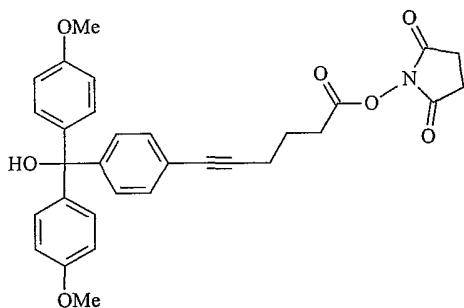
(IIa-6a)



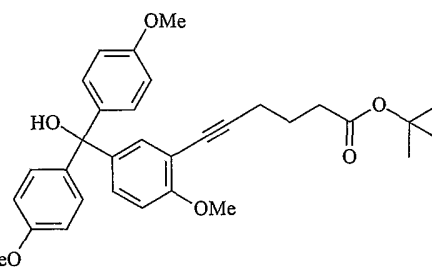
(IIa-6b)



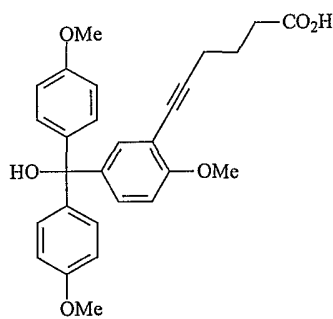
(IIa-6c)



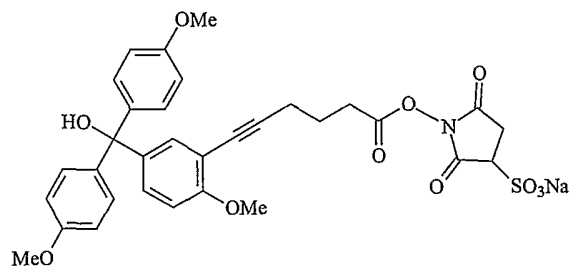
(IIa-6d)



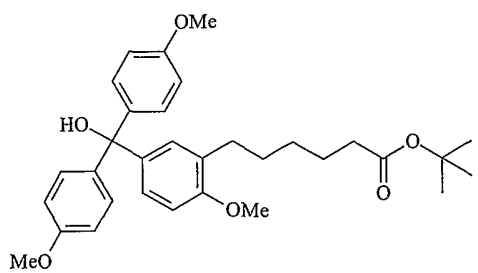
(IIa-7a)



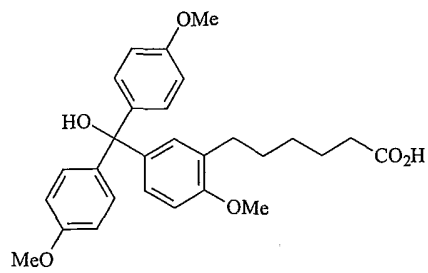
(IIa-7b)



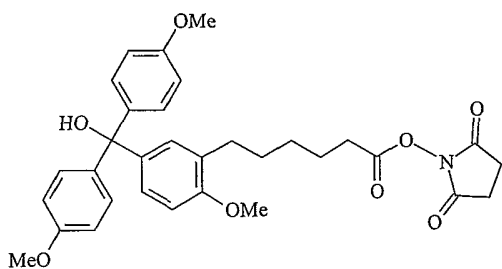
(IIa-7c)



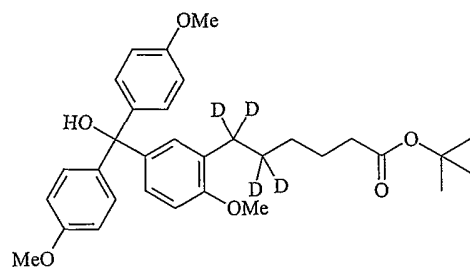
(IIa-8a)



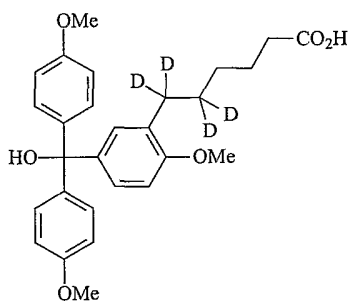
(IIa-8b)



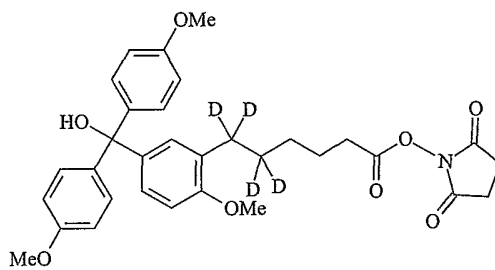
(IIa-8c)



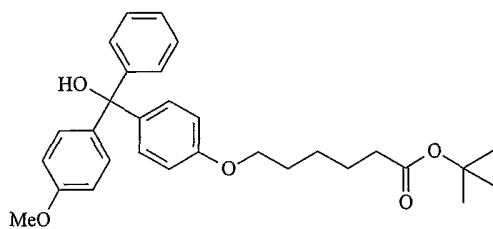
(IIa-9a)



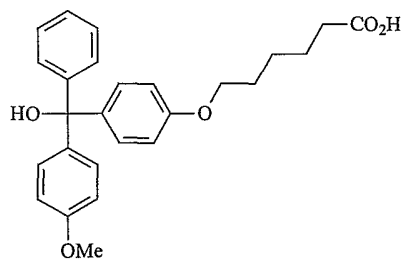
(IIa-9b)



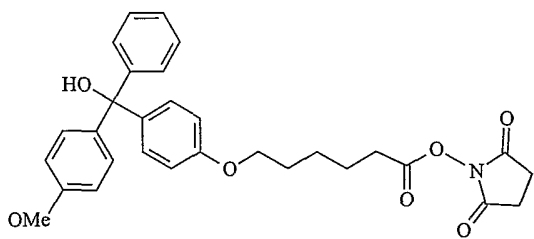
(IIa-9c)



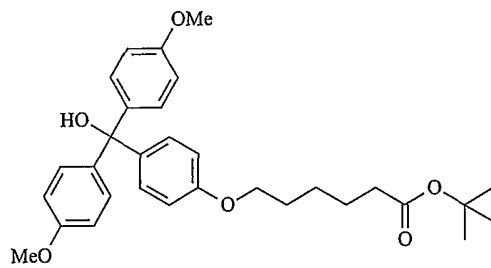
(IIa-10a)



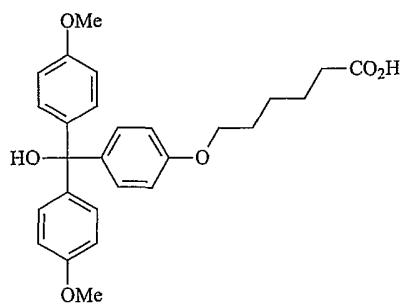
(IIa-10b)



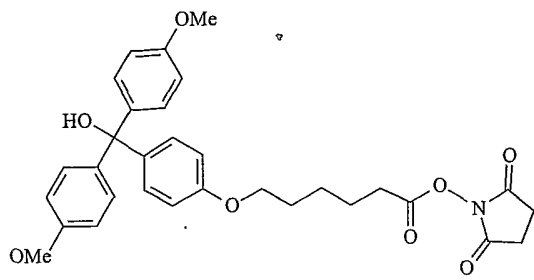
(IIa-10c)



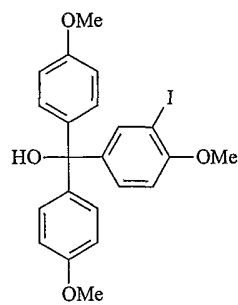
(IIa-11a)



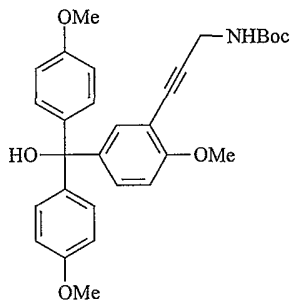
(IIa-11b)



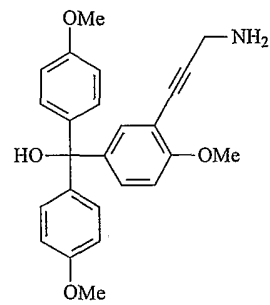
(IIa-11c)



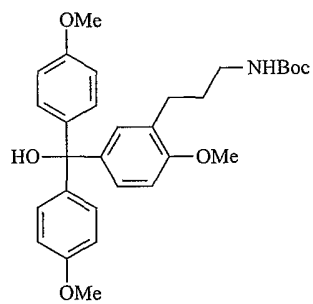
(IIa-12a)



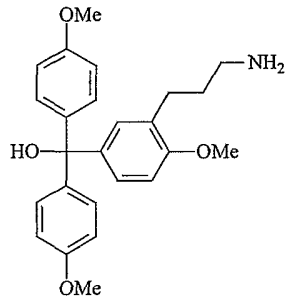
(IIa-12b)



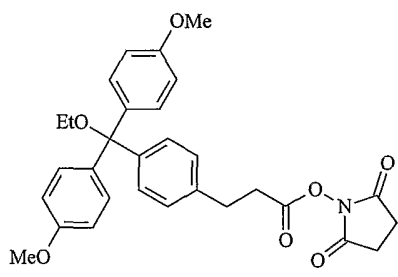
(IIa-12c)



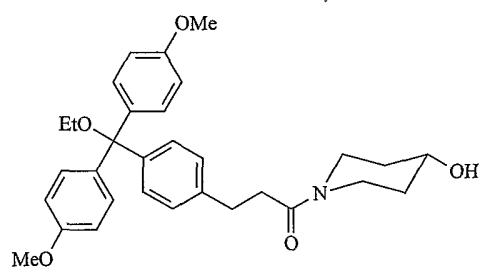
(IIa-13a)



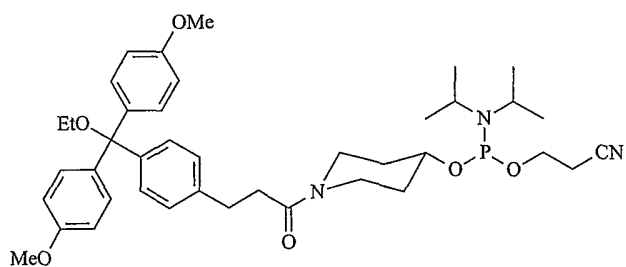
(IIa-13b)



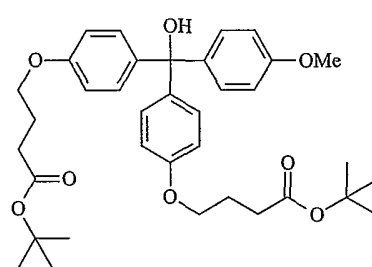
(IIa-14a)



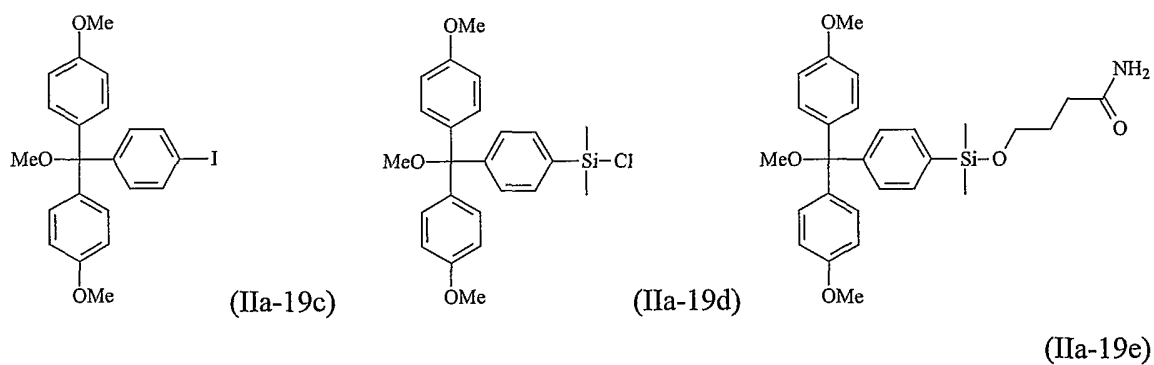
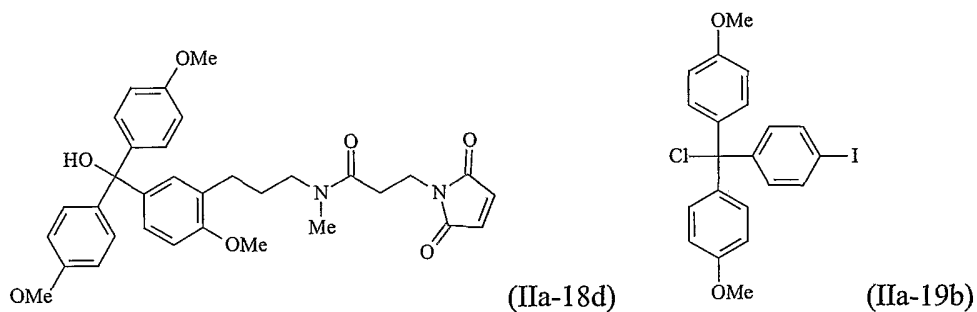
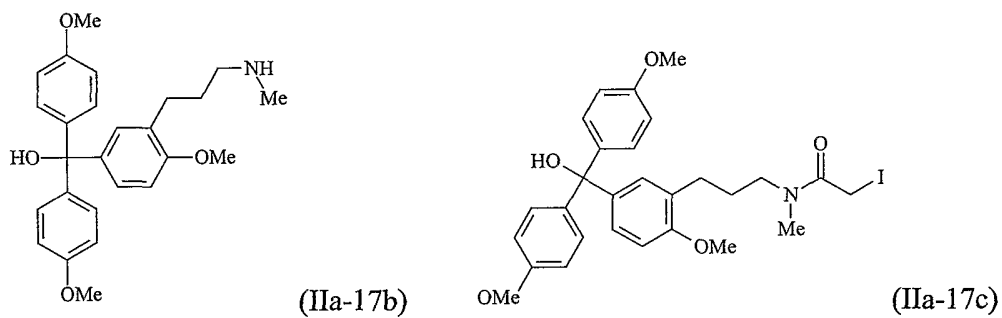
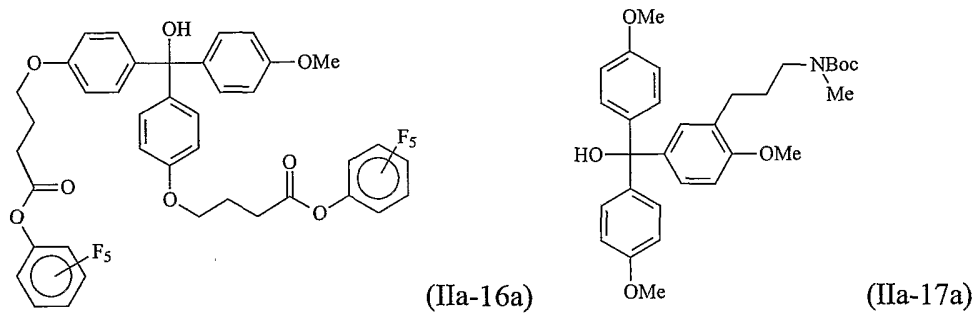
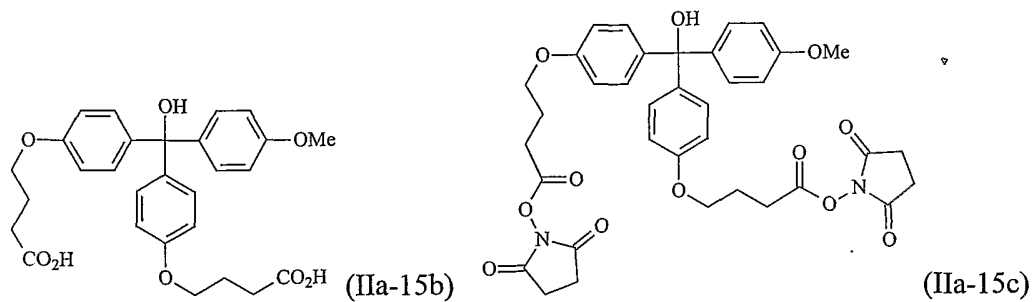
(IIa-14b)

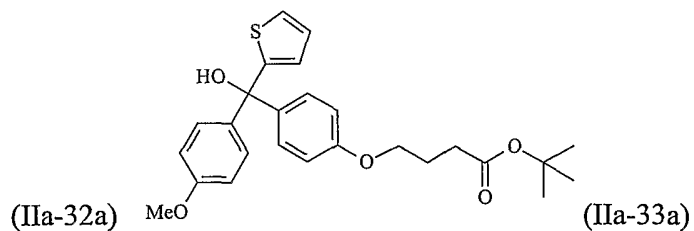
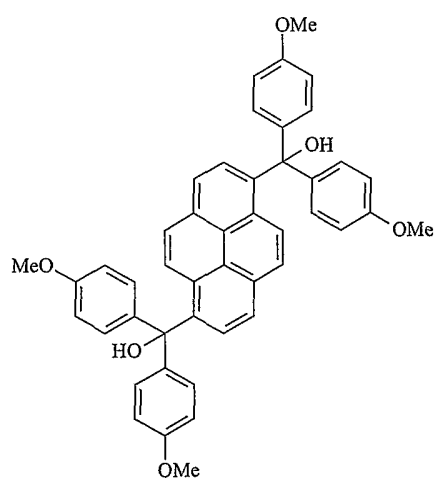
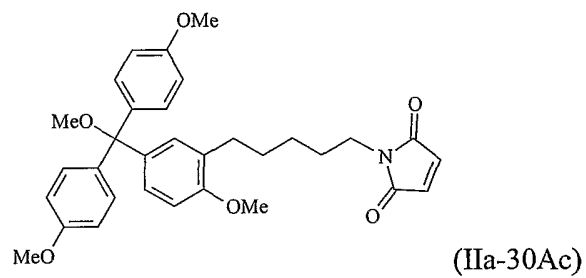
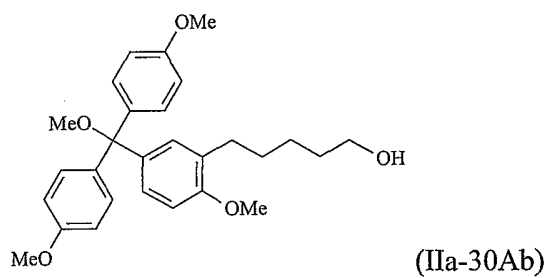
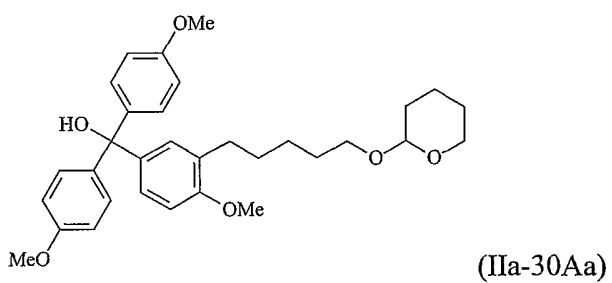
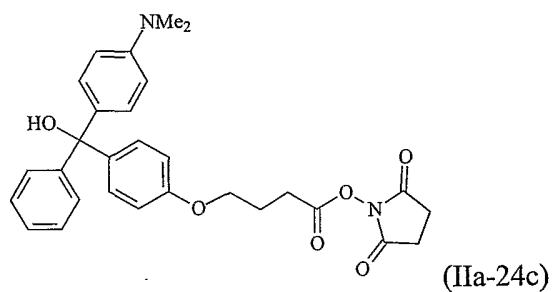
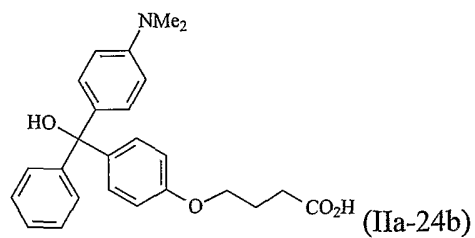
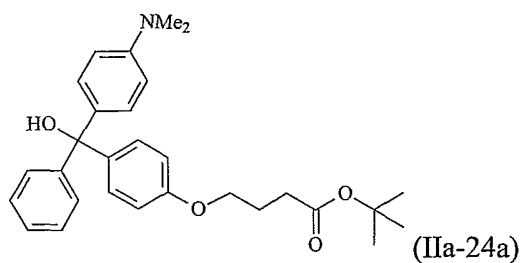
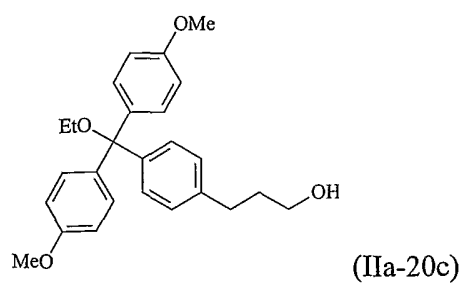
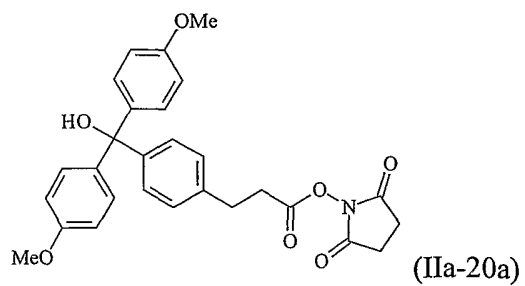


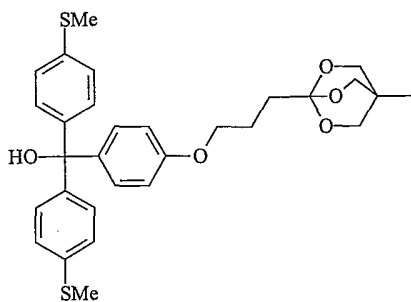
(IIa-14c)



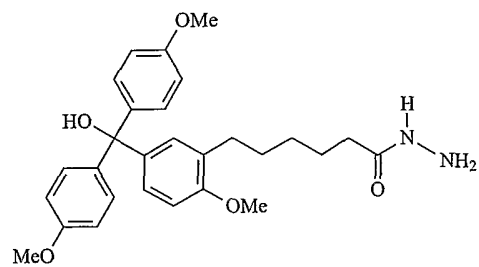
(IIa-15a)



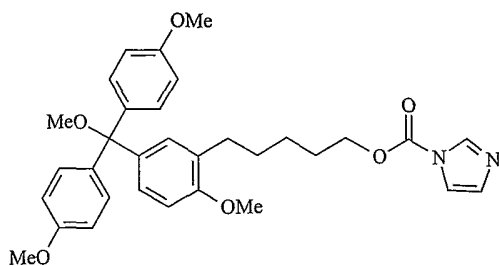




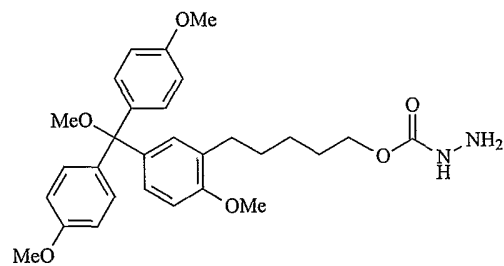
(IIa-34a)



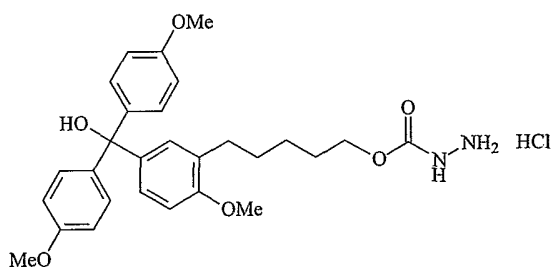
(IIa-35Ab)



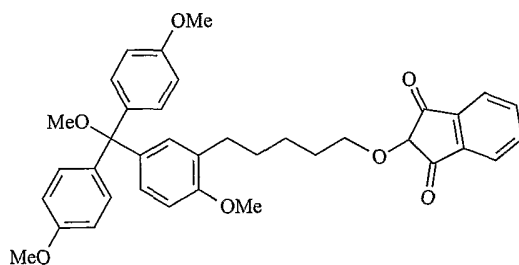
(IIa-35Bb)



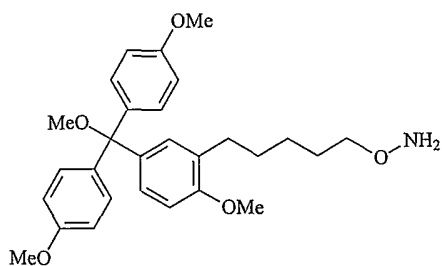
(IIa-35Bc)



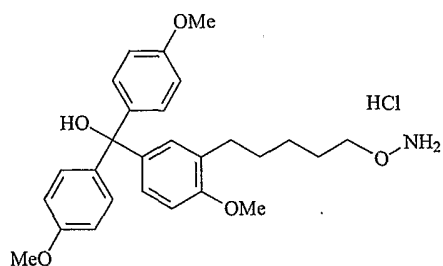
(IIa-35Bd)



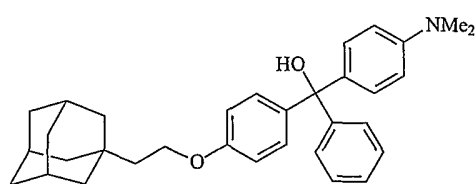
(IIa-35Cb)



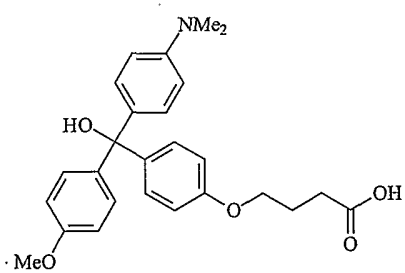
(IIa-35Cc)



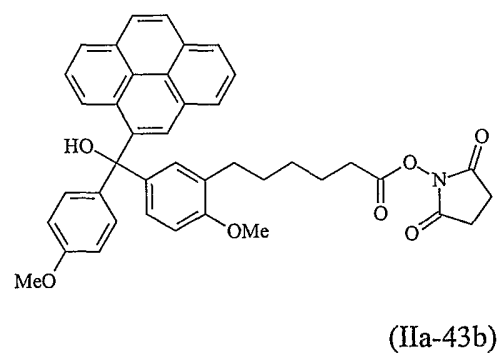
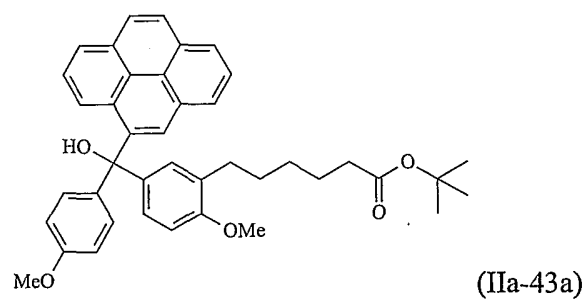
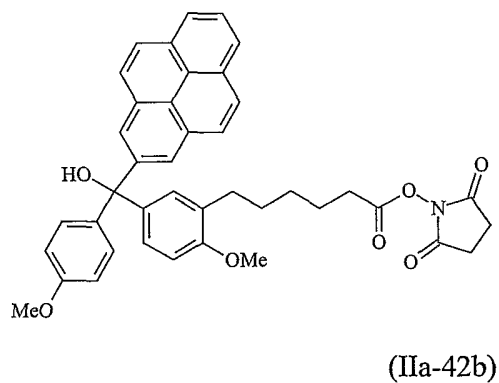
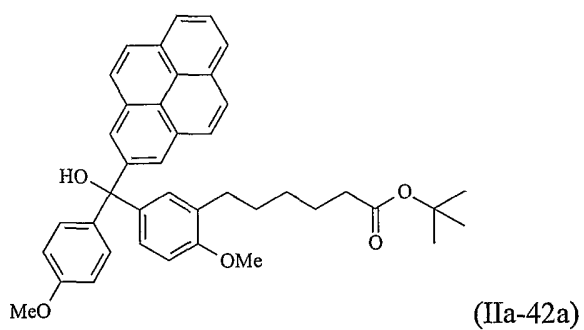
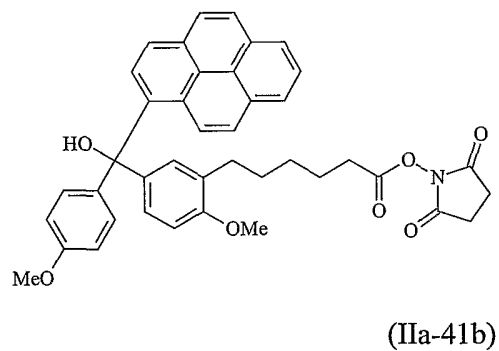
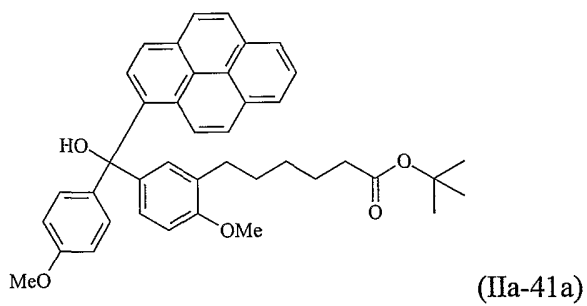
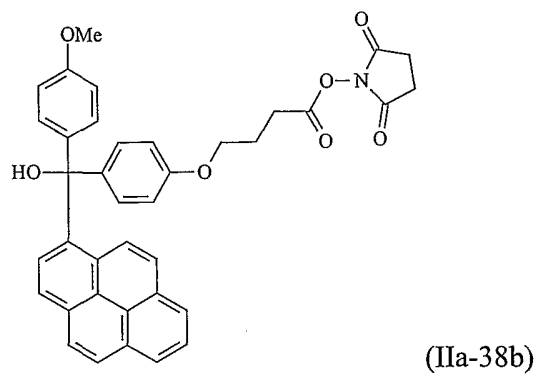
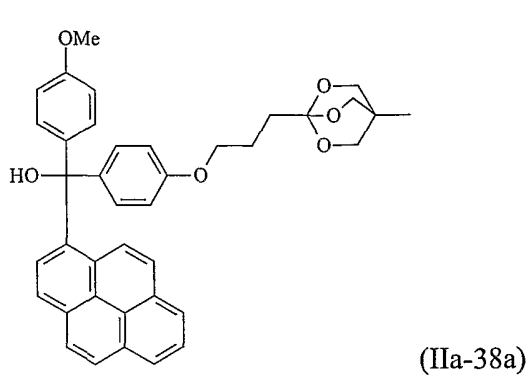
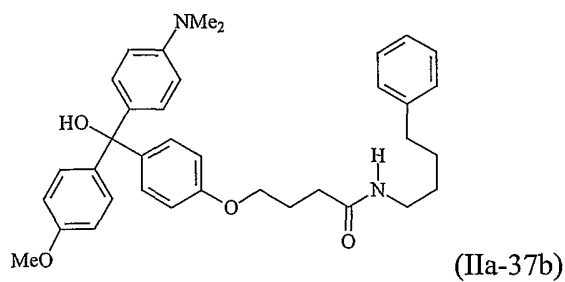
(IIa-35Cd)

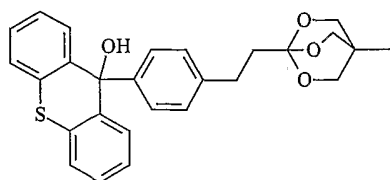


(IIa-36a)

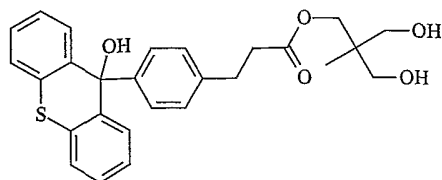


(IIa-37a)

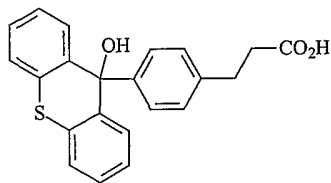




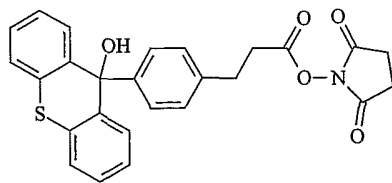
(IIa-48a)



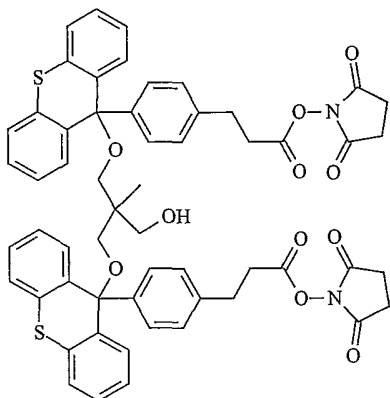
(IIa-48b)



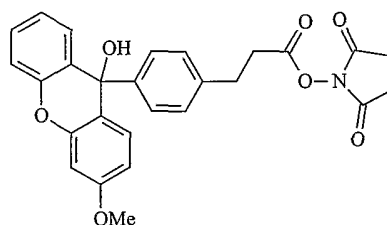
(IIa-48c)



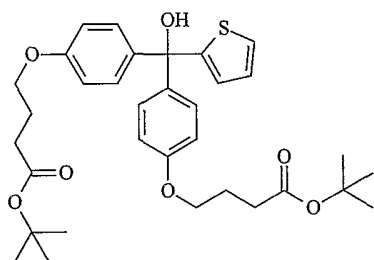
(IIa-48d)



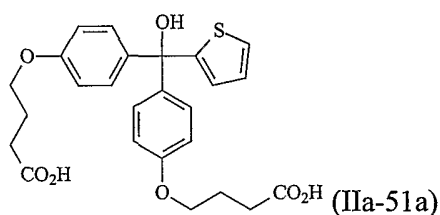
(IIa-48e)



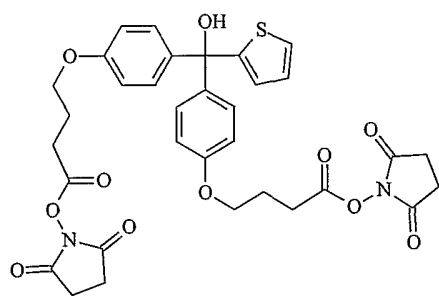
(IIa-49a)



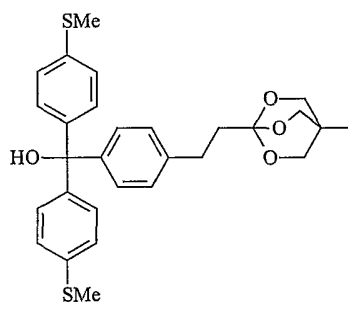
(IIa-50a)



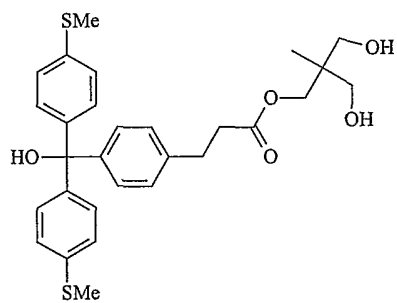
(IIa-51a)



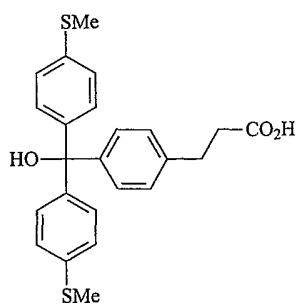
(IIa-51b)



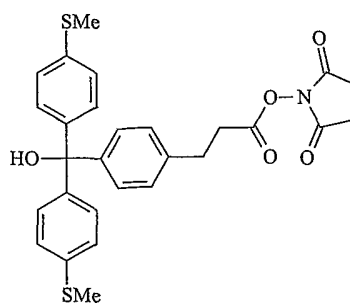
(IIa-53a)



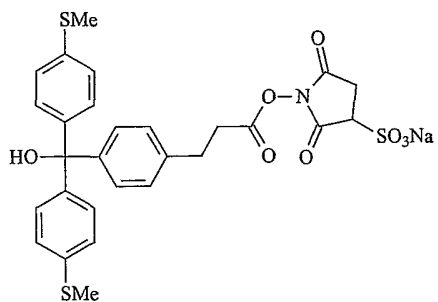
(IIa-53b)



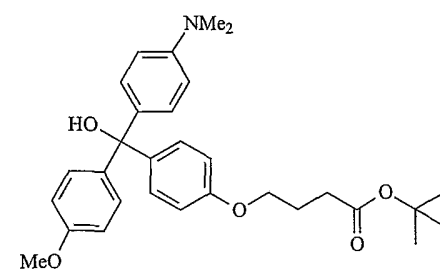
(IIa-53c)



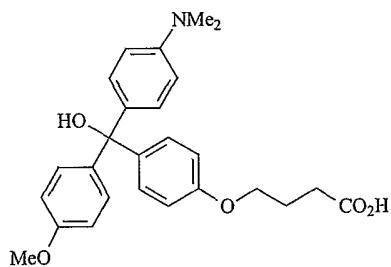
(IIa-53d)



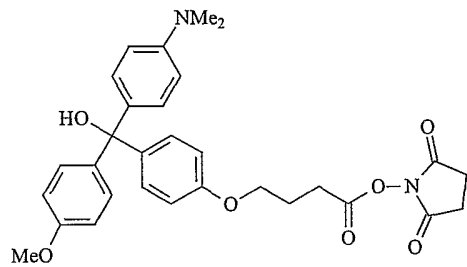
(IIa-53e)



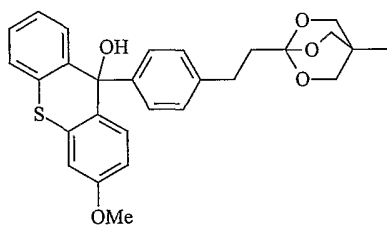
(IIa-54a)



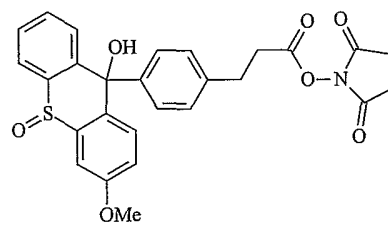
(IIa-54b)



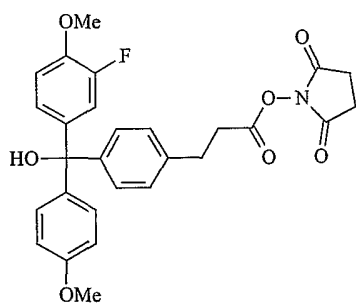
(IIa-54c)



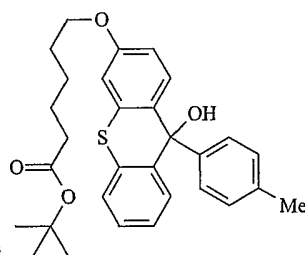
(IIa-56a)



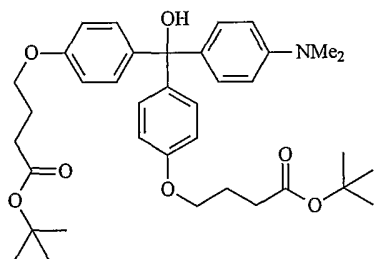
(IIa-56b)



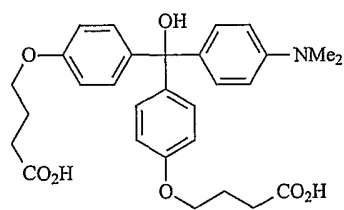
(IIa-57a)



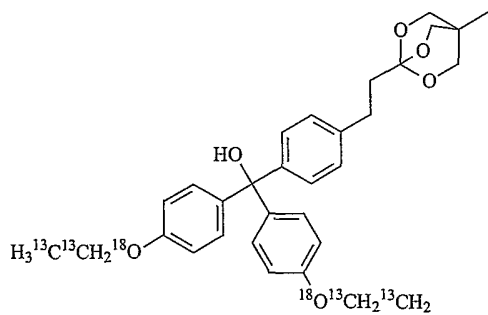
(IIa-58a)



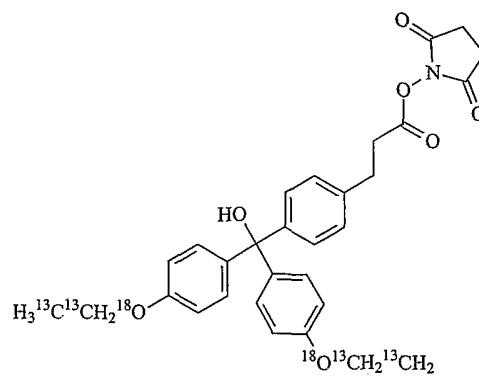
(IIa-59a)



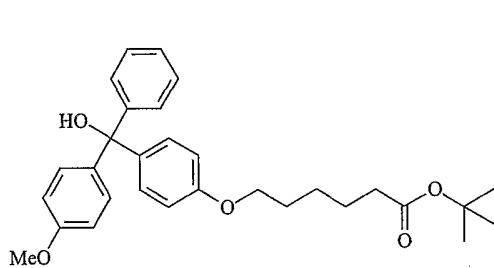
(IIa-59b)



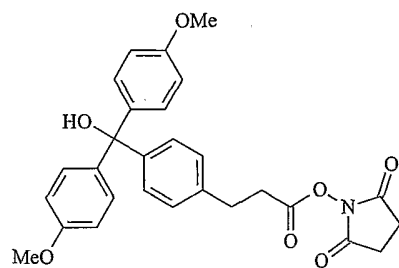
(IIa-60a)



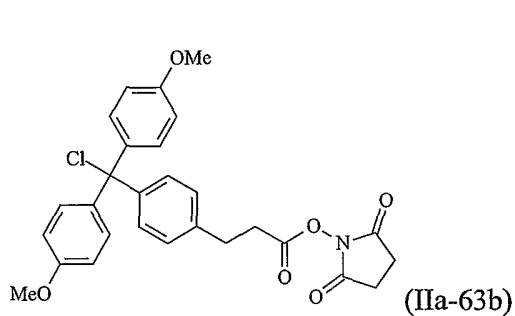
(IIa-60b)



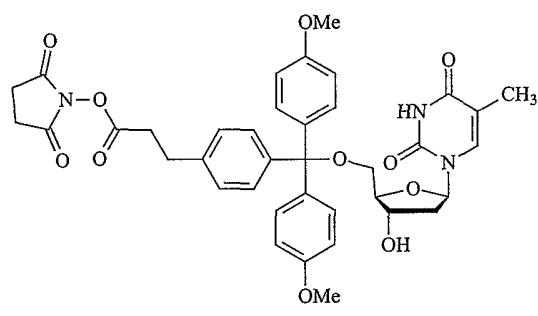
(IIa-61a)



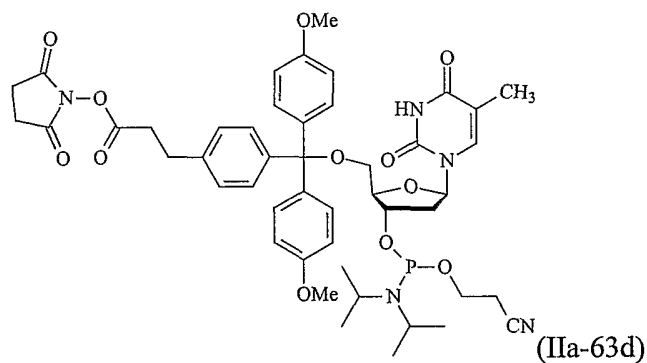
(IIa-62a)



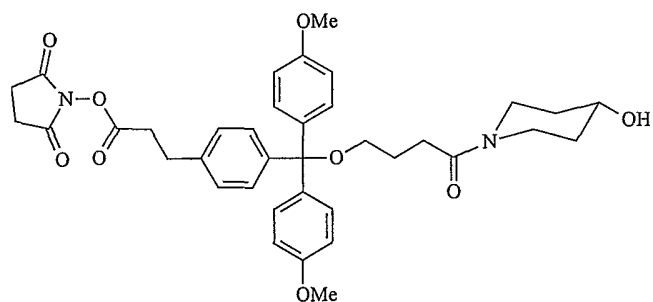
(IIa-63b)



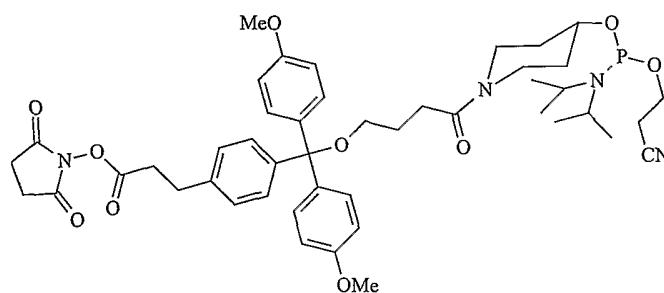
(IIa-63c)



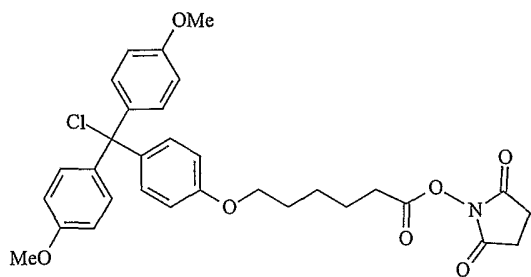
(IIa-63d)



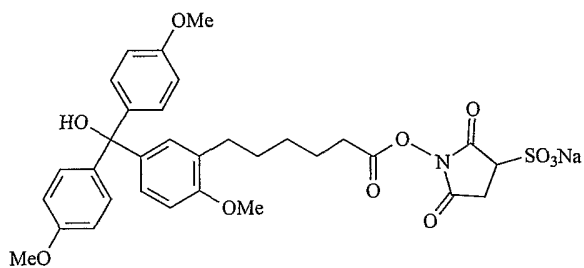
(IIa-63e)



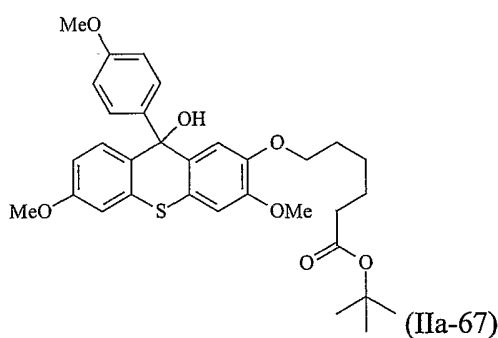
(IIa-63f)



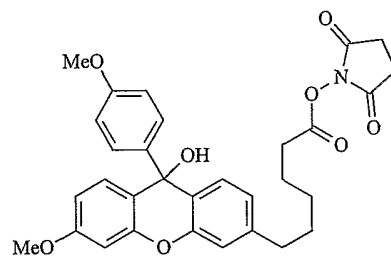
(IIa-64b)



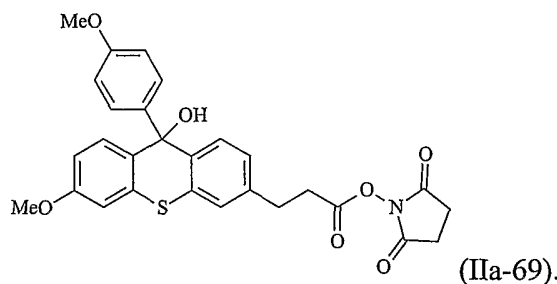
(IIa-66)



(IIa-67)



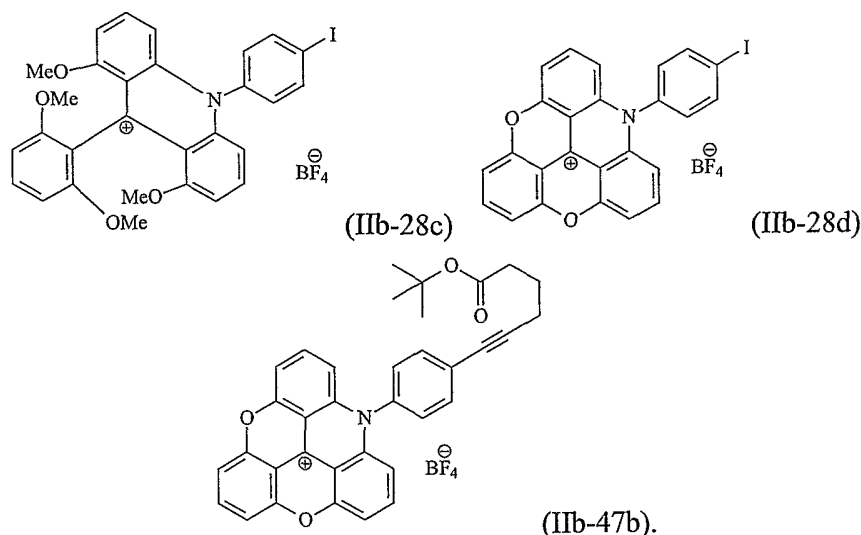
(IIa-68)



(IIa-69).

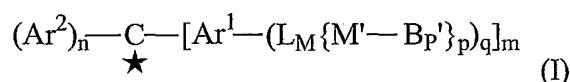
5

In a seventh aspect of the invention, there is provided compounds of formula (IIb) of formulae (IIb-28c), (IIb-28d) and (IIb-47b):



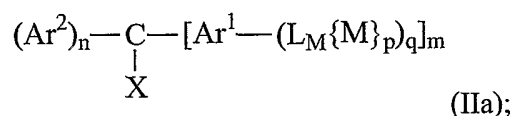
The compounds of formulae (IIa-1a) to (IIa-69), (IIb-28c), (IIb-28d) and (IIb-47b), are particularly useful in the methods of WO2005/057207.

- 5 Thus, in an eighth aspect of the invention, there is provided a method of forming an ion of formula (I):

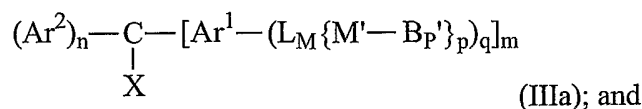


comprising the steps of:

- (i) reacting a compound of the formula (IIa):



with a biopolymer, B_P , having at least one group capable of reacting with M to form a covalent linkage, to provide a biopolymer derivative of the formula (IIIa):



- (ii) cleaving the C—X bond between X and the α -carbon atom of the derivative of formula (IIIa) to form the ion of formula (I);

where:

$\text{C}^\star$ is a carbon atom bearing a single positive charge or a single negative charge;

X is a group capable of being cleaved from the α -carbon atom to form an ion of formula (I);

M is independently a group capable of reacting with B_P to form the covalent linkage;

- 20 B_P' is independently the biopolymer residue of B_P produced on formation of the covalent linkage;

M' is independently the residue of M produced on formation of the covalent linkage;

Ar^1 is independently an aromatic group or an aromatic group substituted with one or more A;

Ar^2 is independently an aromatic group or an aromatic group substituted with one or more A;

optionally wherein (a) two or three of the groups Ar¹ and Ar² are linked together by one or more L⁵, where L⁵ is independently a single bond or a linker atom or group; and/or (b) two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group substituted with one or more A;

5 A is independently a substituent;

L_M is independently a single bond or a linker atom or group;

n = 0, 1 or 2 and m = 1, 2, or 3, provided the sum of n+m = 3;

p independently = 1 or more; and

q independently = 1 or more; and

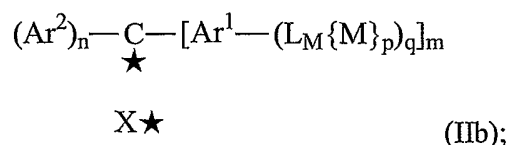
10 wherein the compound of formula (IIa) is selected from the compounds of formulae (IIa-1a) to (IIa-69) of the sixth aspect of the invention.

In a ninth aspect of the invention, there is provided a compound of formula (IIIa) obtainable from a compound of formula (IIa) selected from the compounds of formulae (IIa-1a) to (IIa-69) of the sixth aspect of the invention by the method of the eighth aspect of the invention.

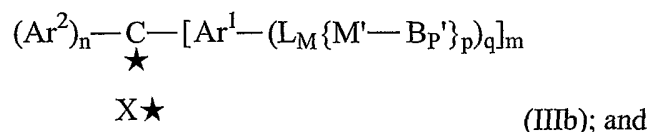
15 In a tenth aspect of the invention, there is provided a compound of formula (I) obtainable from a compound of formula (IIa) selected from the compounds of formulae (IIa-1a) to (IIa-69) of the sixth aspect of the invention by the method of the eighth aspect of the invention.

Furthermore, in an eleventh aspect of the invention, there is provided a method of forming an ion of formula (I), comprising the steps of:

20 (i) reacting a compound of the formula (IIb):



with a biopolymer, B_P, having at least one group capable of reacting with M to form a covalent linkage, to provide a biopolymer derivative of the formula (IIIb):



25 dissociating X★ from the derivative of formula (IIIb), to form the ion of formula (I);
where:

X★ is a counter-ion to C★;

and C★, M, B_P', M', Ar¹, Ar², L_M, n, m, p and q are as defined in the eighth aspect of the invention;

30 wherein the compound of formula (IIb) is selected from the compounds of formulae (IIb-28c), (IIb-28d) and (IIb-47b) of the seventh aspect of the invention.

In a twelfth aspect of the invention, there is provided a compound of formula (IIIb) obtainable from a compound of formula (IIb) selected from the compounds of formulae (IIb-28c), (IIb-28d) and (IIb-47b) of the seventh aspect of the invention by the method of the eleventh aspect of the invention.

In a thirteenth aspect of the invention, there is provided a compound of formula (I) obtainable from a compound of formula (IIb) selected from the compounds of formulae (IIb-28c), (IIb-28d) and (IIb-47b) of the seventh aspect of the invention by the method of the eleventh aspect of the invention.

The compounds of formulae (IIa) or (IIb) may optionally be purified after step (i) of methods of the eighth and eleventh aspects of the invention.

The invention also provides biopolymer derivatives of the formula (IIIa) or (IIIb), as defined above.

The biopolymer derivatives of the invention have enhanced ionisability with respect to free biopolymer, B_P. Advantageously, the biopolymer derivatives may not require a matrix (*e.g.* as used in MALDI-MS) in order to elicit ionisation, although a matrix may help to enhance ionisation. Preferably, ionisation may be obtained without requiring acid treatment, in particular by direct laser illumination.

The invention also provides ions of formula (I), as defined above. These ions are stabilised by the resonance effect of the aromatic groups Ar¹ and Ar². Electron-withdrawing groups, when C★ is an anion, or electron-donating groups, when C★ is a cation, may optionally be provided on Ar¹ and/or Ar² to assist this resonance effect. Consequently, the biopolymer derivatives of the invention readily form ions of formula (I) relative to the native biopolymer, B_P.

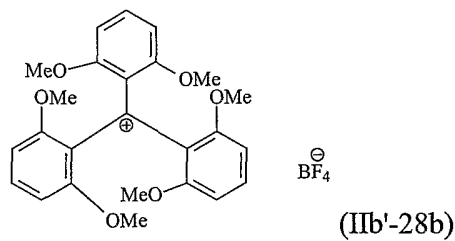
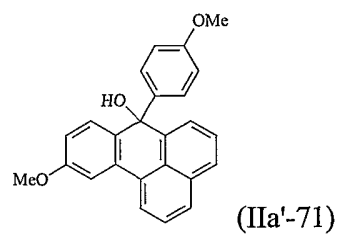
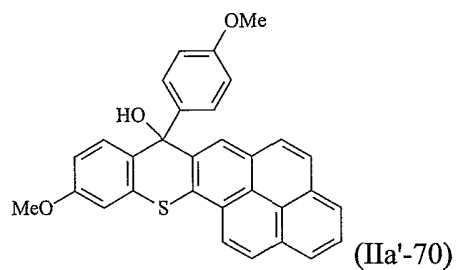
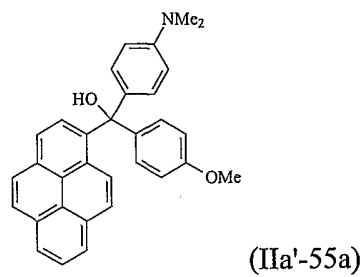
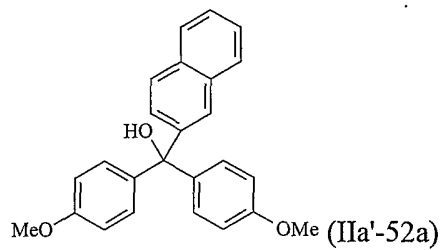
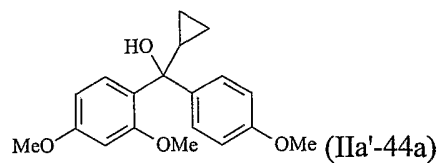
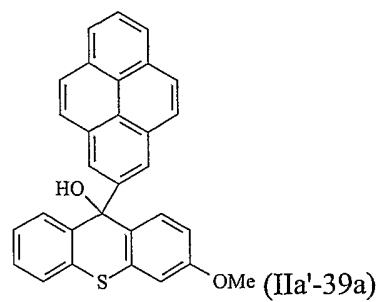
The ions of formula (I) are generally only ever seen on a mass spectrum with a single charge, which is advantageous since it reduces cluttering of the mass spectrum.

The invention also provides compounds of the formula (IIa) and (IIb), as defined above. As mentioned above, these compounds are useful for forming ions of formula (I). As the difference in the molecular mass of the ions of formula (I) and that of the free biopolymer can be accurately calculated, the derivatised compounds of the invention allow analysis of the biopolymer B_P, which may be otherwise difficult or impossible to analyse using known mass spectrometrical techniques.

Other advantageous features of the compounds of the invention include more uniformity of the signal intensity between different analytes (useful for quantitative studies) and similar desorption properties between compounds with different, but close, masses, so that techniques such as isotope coded affinity tagging (ICAT) can be employed with the compounds of the invention.

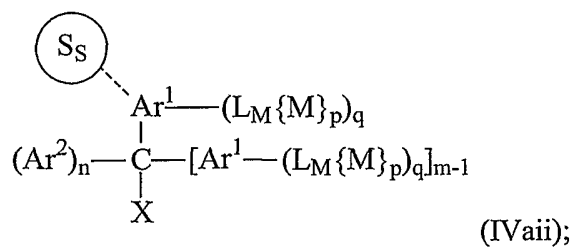
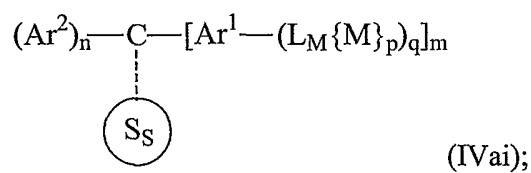
The homogeneous methods of the invention are particularly appropriate for small molecules, *e.g.* amines.

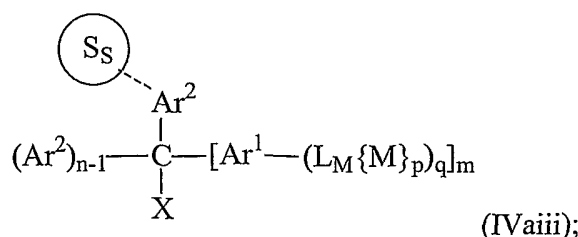
The invention also provides intermediates useful in the synthesis of compounds of formulae (IIa) and (IIb) having the formulae:



5 Solid Supports

The invention also provides solid supports of formula (IVai), (IVaii) or (IVaiii):





where:

X, Ar¹, Ar², L_M, M, n, m, p and q are as defined above;

S_S is a solid support;

- 5 C---S_S comprises a cleavable bond between C and S_S;
 S_S---Ar¹ comprises a cleavable bond between Ar¹ and S_S; and
 S_S---Ar² comprises a cleavable bond between Ar² and S_S.

The cleavable bond of C---S_S, S_S---Ar¹ or S_S---Ar² may be a covalent, ionic, hydrogen, dipole-dipole or van der Waals bond.

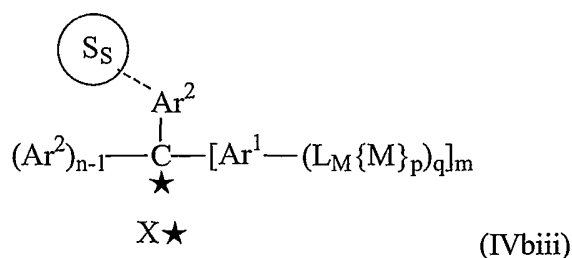
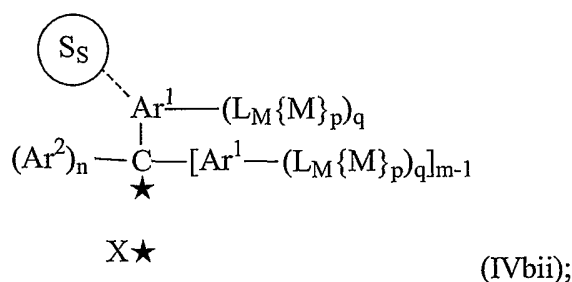
- 10 The solid supports of formula (IVai), (IVaii) and (IVaiii) may form ions of formula (I):

(a) for modified solid supports of formula (IVai) by cleaving the C-S_S bond between the α-carbon atom of the modified solid support of formula (IVai) and the solid support S_S to form the ion of formula (I);

- (b) for modified solid supports of formula (IVaii) by, either simultaneously or
 15 sequentially, cleaving the C-X bond between X and the α-carbon atom and cleaving the S_S---Ar¹ bond between the solid support and the Ar¹ group to form the ion of formula (I); or

(c) for modified solid supports of formula (IVaiii) by, either simultaneously or sequentially, cleaving the C-X bond between X and the α-carbon atom and cleaving the S_S---Ar² bond between the solid support and the Ar² group to form the ion of formula (I).

- 20 The invention also provides solid supports of formula (IVbii) or (IVbiii):



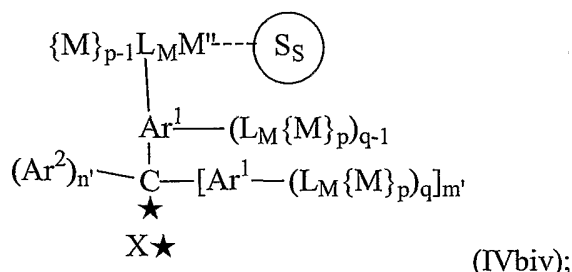
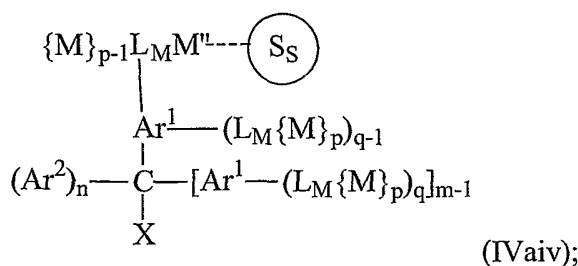
where: X★, Ar¹, Ar², L_M, M, n, m, p, q, S_S, C---S_S, S_S---Ar¹ and S_S---Ar² are as defined above.

The solid supports of formula (IVbii) and (IVbiii) may form ions of formula (I):

(a) for modified solid supports of formula (IVbii) by, either simultaneously or sequentially, dissociating $X^\star$ from the derivative of formula (IVbii) and cleaving the $S_S \cdots \text{Ar}^1$ bond between the solid support and the Ar^1 group to form an ion of formula (I'); or

(b) for modified solid supports of formula (IVbiii) by, either simultaneously or sequentially, dissociating $X^\star$ from the derivative of formula (IVbiii) and cleaving the $S_S \cdots \text{Ar}^2$ bond between the solid support and the Ar^2 group to form an ion of formula (I').

The invention also provides solid supports of formula (IVaiv) or (IVbiv):



10 where:

X , $X^\star$, Ar^1 , Ar^2 , L_M , M , p , q , n , m , and S_S are as defined above;

$M'' \cdots S_S$ comprises a bond between M'' and S_S ; and

M'' is the same as M except that S_S is bound to a portion of M which does not form part of the residue of M'' remaining attached to the ion of formula (I') which residue is produced after reaction of

15 group M'' .

In this embodiment of the invention, the solid support is bound to a part of group M'' which does not go on to form part of the residue of M'' remaining attached to the ion of formula (I') which residue is produced after reaction of group M'' .

The solid supports of formula (IVai), (IVaii), (IVaiii), (IVbii), (IVbiii), (IVaiv) and (IVbiv) are useful

20 in the methods disclosed in WO2005/057207.

Methods of Analysis

The invention also provides a method for analysing a biopolymer, B_p , comprising the steps of:

- (i) reacting the biopolymer B_p with a compound of formula (IIa) or (IIb), wherein the compound of formula (IIa) or (IIb) is selected from the compounds of formulae (IIa-1a) to (IIa-69) or the compounds of formulae (IIb-28c), (IIb-28d) and (IIb-47b) described above;
 - (ii) providing an ion of formula (I); and
- 25

(iii) analysing the ion of formula (I) by mass spectrometry.

The biopolymer will typically have been obtained using a preparative or analytical process. For example, it may have been purified using various separation methods (*e.g.* 1-dimensional or 2-dimensional, reverse-phase or normal-phase separation, by *e.g.* chromatography or electrophoresis) and the separation may be based on any of a number of characteristics (*e.g.* isoelectric point, molecular weight, charge, hydrophobicity, *etc.*). Typical methods include 2D SDS-PAGE, 2D liquid chromatography (*e.g.* Multidimensional Protein Identification Technology, MudPIT, or 2D HPLC methods). The separation method can preferably interface directly with the mass spectrometer.

Known analytical techniques can thus be adapted or improved by the method of the invention. A particularly preferred method involves 2D-PAGE of a biopolymer, or mixture of biopolymers, selection of a spot of interest in the electrophoretogram, and then derivatisation and analysis of that spot using the techniques of the invention. The biopolymer may be proteolytically digested prior to its analysis (typically within the PAGE gel, but optionally digested after extraction from the gel) and/or may itself be the product of a proteolytic digest.

The invention also provides, in a method for analysing a biopolymer, B_P, the improvement consisting of: (i) reacting a biopolymer, B_P with a compound of formula (IIa) or (IIb), wherein the compound of formula (IIa) or (IIb) is selected from the compounds of formulae (IIa-1a) to (IIa-69) or the compounds of formulae (IIb-28c), (IIb-28d) and (IIb-47b) described above; (ii) providing an ion of formula (I); and (iii) analysing the ion by mass spectrometry.

Typically, the analysis by mass spectrometry is carried out in a spectrometer which is suitable for MALDI-TOF spectrometry.

In the spectrometer, the ion source may be a matrix-assisted laser desorption ionisation (MALDI), an electrospray ionisation (ESI) ion source, a Fast-Atom Bombardment (FAB) ion source. Preferably, the ion source is a MALDI ion source. The MALDI ion source may be traditional MALDI source (under vacuum) or may be an atmospheric pressure MALDI (AP-MALDI) source. MALDI is a preferred ionisation method, although the use of a matrix is generally not required

In the spectrometer, the mass analyser may be a time of flight (TOF), quadrupole time of flight (Q-TOF), ion trap (IT), quadrupole ion trap (Q-IT), triple quadrupole (QQQ) Ion Trap or Time-Of-Flight Time-Of-Flight (TOF/TOF) or Fourier transform ion cyclotron resonance (FTICR) mass analyser. Preferably, the mass analyser is a TOF mass analyser.

Preferably, the mass spectrometer is a MALDI-TOF mass spectrometer.

Further Embodiments*M' bound to B_P' by a non-covalent linker*

The above-mentioned embodiments of the invention may also be provided in which M' is bound to B_P' by a non-covalent bond. All the other features of the invention are the same except the groups
5 which relate to the non-covalent bond between M' and B_P'.

The non-covalent bond may be direct between M' and B_P' or may be provided by one or more binding groups present on M' and/or B_P'.

Preferred non-covalent bonds are those having an association constant (K_a) of at least 10^{14} M^{-1} , preferably about 10^{15} M^{-1} .

10 In preferred embodiment, one of M' and B_P' will have a binding group comprising biotin, and the other of M' and B_P' will have a binding group comprising avidin or streptavidin.

Preferably, when the compounds of the invention comprise a non-covalent bond between M' and B_P' and a cleavable bond between C and S_S, Ar¹ and S_S, or Ar² and S_S, these bonds are differentially cleavable. More preferably, the non-covalent bond between M' and B_P' is not cleaved under
15 conditions which the cleavable bond between C and S_S, Ar¹ and S_S, or Ar² and S_S, as appropriate, is cleaved.

L_M bound to Ar¹ by more than one bond

The above-mentioned embodiments of the invention may also be provided in which L_M is bound to Ar¹ by more than one covalent bond (e.g. 2 or 3 bonds) which are either single, double or triple
20 covalent bonds, or one or more multiple bonds (e.g. double or triple covalent bonds). All the other features of the invention are the same except the groups which relate to the bond or bonds between Ar¹ and L_M.

Ionisation of Compounds other than Biopolymers

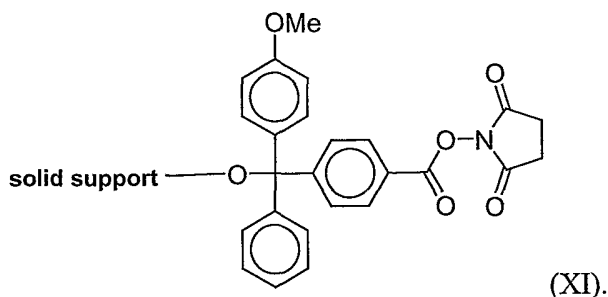
In addition to biopolymers, the present invention may be used for ionising any molecule or complex
25 of molecules which requires mass spectrum analysis. Thus, the above-mentioned embodiments of the invention may also be provided in which B_P is replaced by any molecule or complex having at least one group capable of reacting with M to form a covalent linkage. All the other features of the invention are the same, except group M is group capable of reacting with the molecule to be analysed.

30 Examples of other molecules which may be analysed in the present invention include non-biological polymers (e.g. synthetic polyesters, polyamides and polycarbonates), petrochemicals and small molecules (e.g. alkanes, alkenes, amines, alcohols, esters and amides). Amines are particularly preferred.

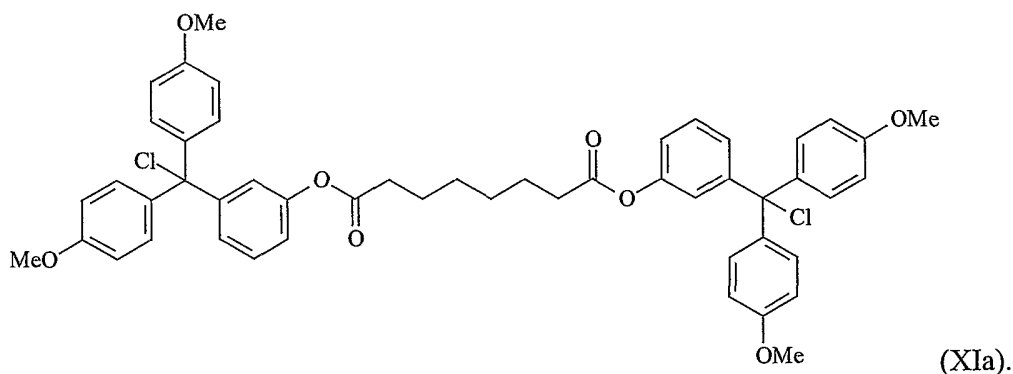
Examples of complexes which may be analysed in the present invention include double- and triple-stranded RNA, DNA and/or peptide nucleic acid (PNA) complexes, enzyme/substrate complexes, multimeric proteins (*e.g.* dimers, trimers, tetramers, pentamers, *etc.*), virions, *etc.*

Disclaimers

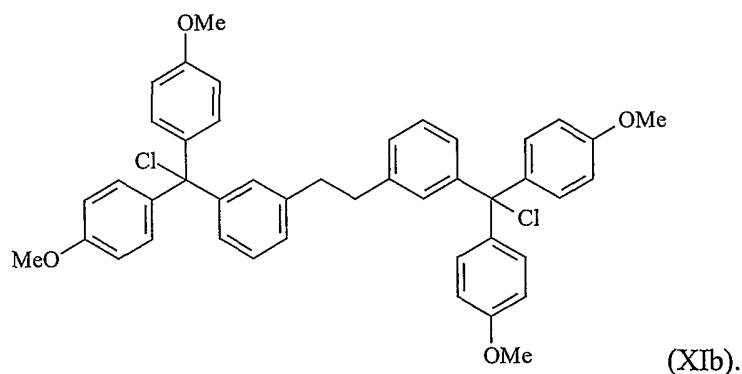
- 5 Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XI) are disclaimed



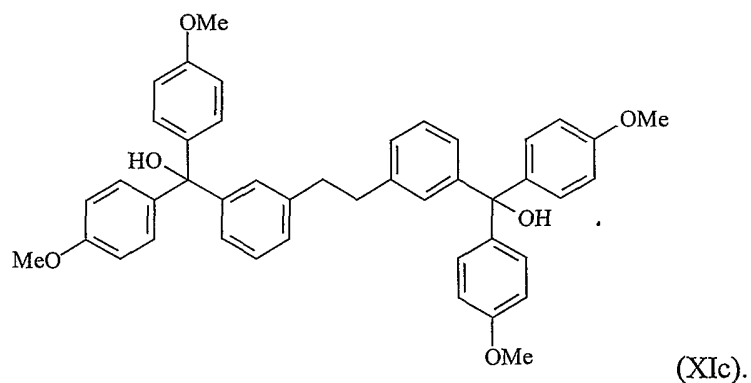
Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XIa) are disclaimed.



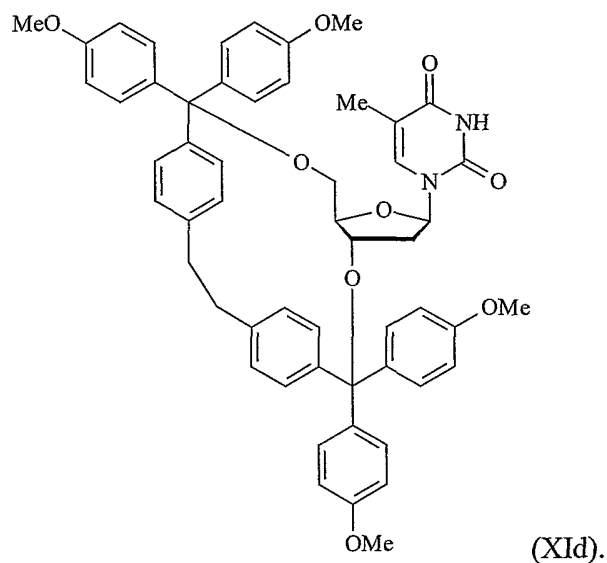
Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XIb) are disclaimed



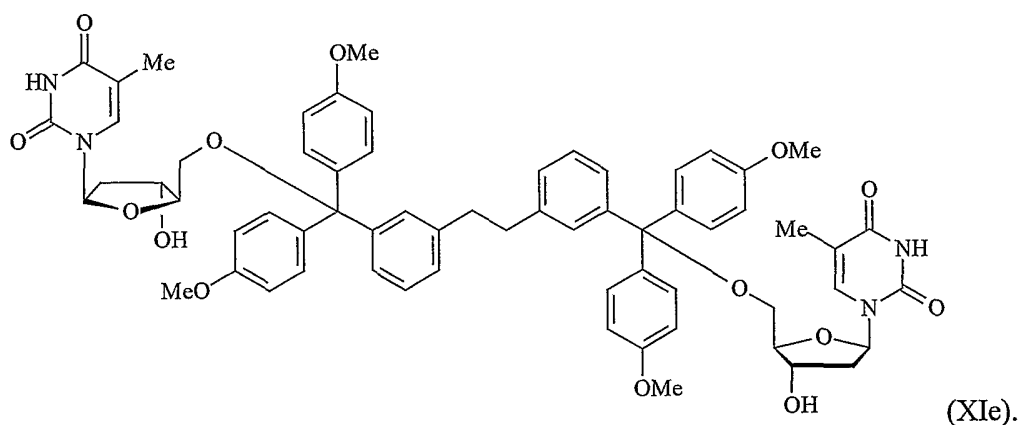
- 15 Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XIc) are disclaimed



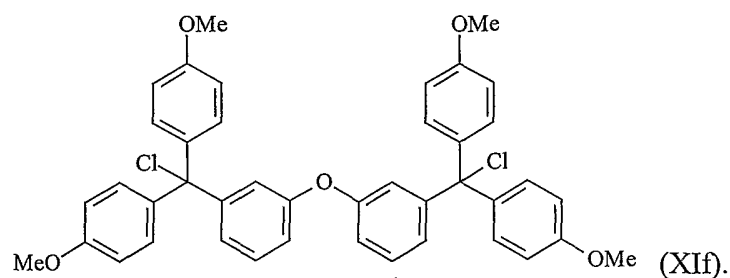
Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XIc) are disclaimed



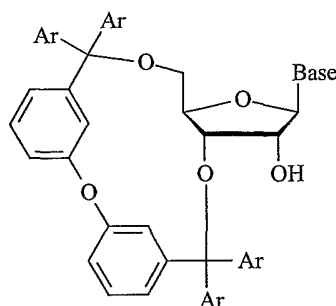
- 5 Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XIe) are disclaimed



Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XIe) are disclaimed



Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XIg-j) are disclaimed

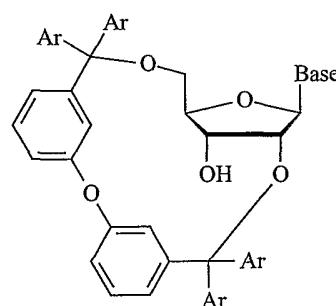


5

Ar = p-anisyl

Formula	Base
XI g	Uridine
XI h	N ⁴ -benzoyl-cytidine
XI i	N ⁶ -benzoyl-adenosine
XI j	N ² -phenylacetyl-guanosine

Preferably, all embodiments of the invention (including products of formulae (I) and (IIa)) involving or relating to the compound of formula (XI k-n) are disclaimed



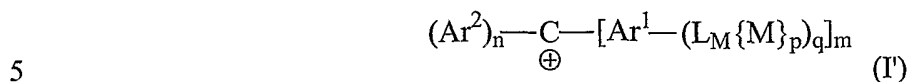
10

Ar = p-anisyl

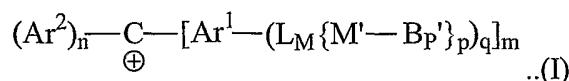
Formula	Base
XI k	Uridine
XI l	N ⁴ -benzoyl-cytidine
XI m	N ⁶ -benzoyl-adenosine
XI n	N ² -phenylacetyl-guanosine

Preferred Embodiments**Definition of C★**

Preferably, C★ bears a single positive charge such that ions of the invention are cations, the ion of formula (I') has the following structure:



the ion of formula (I) has the following structure:



and the compounds of formulae (IIb), (IIIb), (IVbii), (IVbiii) and (IVbiv) have the structures disclosed in table 1.

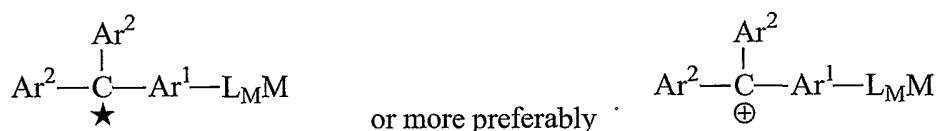
10 ***n, m, p and q***

For the purposes of compounds of the invention having n-1 groups Ar², n may not be less than 1.

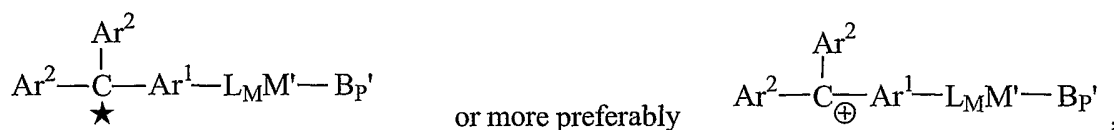
Preferably n = 2 and m = 1.

Preferably p = 1, 2 or 3. Preferably p = 1.

Preferably q = 1, 2 or 3. Preferably q = 1.

15 **Preferably n = 2, m = 1, p = 1 and q = 1. Thus, the ion of formula (I') has the structure:**

the ion of formula (I) has the structure:



and the compounds of formulae (IIa), (IIb), (IIIa), (IIIb), (IVai), (IVaii), (IVaiii), (IVaiv), (IVbii),

20 **(IVbiii) and (IVbiv) have the structures disclosed in table 2.*****X, Ar¹, Ar², L_M, M and L⁵***

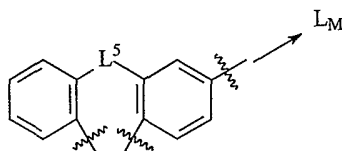
Preferred compounds of formula (IIa) are those wherein at least one (e.g. 1, 2, 3, 4, 5 or 6) of the groups X, Ar¹, Ar², L_M, M and L⁵ (where present) are selected from the groups X, Ar¹, Ar², L_M, M and L⁵ listed in table 3. Particularly preferred compounds of formula (IIa) are those wherein all of the

25 **groups X, Ar¹, Ar², L_M, M and L⁵ are selected from the groups X, Ar¹, Ar², L_M, M and L⁵ listed in table 3.**

Preferred compounds of formula (IIb) are those wherein at least one (e.g. 1, 2, 3, 4, 5 or 6) of the groups $X^\star$, Ar^1 , Ar^2 , L_M , M and L^5 are selected from the $X^\star$, Ar^1 , Ar^2 , L_M , M and L^5 listed in table 4. Particularly preferred compounds of formula (IIa) are those wherein all of the $X^\star$, Ar^1 , Ar^2 , L_M , M and L^5 are selected from the groups $X^\star$, Ar^1 , Ar^2 , L_M , M and L^5 listed in table 4.

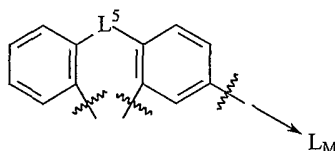
5 Combinations of Ar Groups

In a preferred embodiment, one Ar^1 and one Ar^2 are combined to form the group:



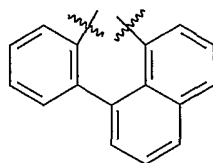
optionally substituted by A. Preferably, L^5 is O (e.g. compound (IIa-68)) or S (e.g. compounds (IIa-58a) and (IIa-69)). Compounds of this embodiment show improved mass spectrometry enhancing properties. Preferred optional substituents A are $-OMe$ (e.g. compounds IIa-68 and IIa-69), preferably para to $C^\star$.

In another preferred embodiment, one Ar^1 and one Ar^2 are combined to form the group:



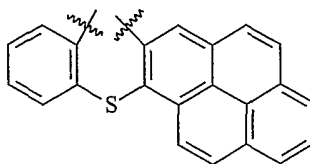
optionally substituted by A. Preferably, L^5 is O or S (e.g. compound (IIa-67)), preferably S. Compounds of this embodiment also show improved mass spectrometry enhancing properties. Preferred optional substituents A are $-OMe$ (e.g. compound IIa-67), preferably para to $C^\star$.

In another preferred embodiment, two Ar^1 or Ar^2 groups (i.e. Ar^1+Ar^1 , Ar^1+Ar^2 , or Ar^2+Ar^2), are linked by one L^5 , wherein one Ar^1 or Ar^2 group is a polycyclic aromatic group (e.g. naphthyl or pyrenyl), preferably a pyrenyl group. Such combinations of Ar groups are fluorescent and allow labelling, e.g. of the biopolymer. An example of such a combination of Ar groups is:

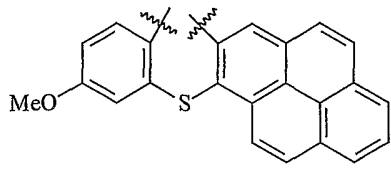


optionally substituted by A, e.g. $-OMe$, wherein when one or more of the Ar groups is Ar^1 , the combination includes an appropriate number of $L_M\{M\}_p$ groups.

It is particularly preferred in this embodiment that L^5 is S. The S atom may be oxidised to $S=O$ without loss of the X group, advantageously allowing modification of the properties (e.g. fluorescent properties) of the combined Ar group. A particularly preferred combination of Ar groups in this embodiment is:



optionally substituted by A, *e.g.* -OMe, *e.g.*



wherein when one or more of the Ar groups is Ar¹, the combination includes an appropriate number
5 of L_M{M}_p groups.

Biopolymers

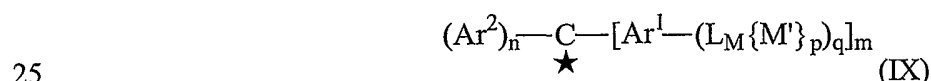
The term 'biopolymer' includes polymers found in biological samples, including polypeptides, polysaccharides, and polynucleotides (*e.g.* DNA or RNA). Polypeptides may be simple copolymers of amino acids, or they may include post-translational modifications *e.g.* glycosylation, lipidation,
10 phosphorylation, *etc.* Polynucleotides may be single-stranded (in whole or in part), double-stranded (in whole or in part), DNA/RNA hybrids, *etc.* RNA may be mRNA, rRNA or tRNA.

Advantageous biopolymers are those which do not readily form a molecular ion in known MALDI-TOF MS techniques, especially those which do not form a molecular ion on illumination of laser light at 340 nm.

15 Biopolymers for use in the invention comprise two or more monomers, which may be the same or different as each other. Preferred biopolymers comprise at least *pp* monomers, where *pp* is 5 or more (*e.g.* 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 250). More preferred biopolymers comprise *ppp* or fewer monomers where *ppp* is 300 or less (*e.g.* 200, 100, 50).

Biopolymers may have a molecular mass of at least *qq* kDa, where *qq* = 0.5 or more (*e.g.* 0.6, 0.7,
20 0.8, 0.9, 1, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 75, 100, *etc.*). Preferred biopolymers are those having a molecular mass within the range of detection of a mass spectrometer. More preferred biopolymers have a molecular mass of *qqq* kDa or less, where *qqq* is 30 or less (*e.g.* 20, 10, 5).

Preferably, the mass, m(IX), of the fragment (IX)



of the cation of formula (I) is significantly less than the mass, m(B_P¹), of the biopolymer residue B_P¹. For example the ratio m(B_P¹) / m(IX) is preferably more than *nn*, where *nn* is at least 2 (*e.g.* 3, 4, 5, 10, 100, 1000, *etc.*).

The invention is suitable for use with purified biopolymers or mixtures of biopolymers. For example, a pure recombinant protein could be derivatised and analysed by MS, or biopolymers within a cellular lysate or extract could be derivatives and then analysed.

Preferred biopolymers are polypeptides. Particularly preferred biopolymers are polypeptides formed
5 after proteolytic digestion of a protein.

Biopolymers bound to solid supports

In preferred embodiments of the invention the biopolymer is bound to a solid support such that it is cleavable from the solid support at least once it has been derivatised by a compound of the invention. B_P is thus derivatised *in situ* while bound to the support, and is then released. As the biopolymer is
10 bound to the solid support, this aspect of the invention is particular relevant to methods involving compounds of formulae (IIa) and (IIb).

The biopolymer may be bound to the solid support by a covalent, ionic, hydrogen, dipole-dipole or van der Waals bond (also known as a dispersion bond or a London forces bond). The covalent, ionic, hydrogen, dipole-dipole or van der Waals bond may be direct between the biopolymer and the solid
15 support or may be provided by one or more binding groups present on the biopolymer and/or solid support. Preferred groups are non-covalent groups.

Examples of groups which can form these types of bond, and methods for cleaving these types of bond, are set out below in connection with C---S_S bonds, *etc.*

In a particularly preferred embodiment, the solid support is provided with $-(NMe_3)^+$ binding groups
20 and the biopolymer has a net negative charge, or *vice versa* (*i.e.* the $-(NMe_3)^+$ is on the biopolymer). In other preferred embodiments, the solid support is provided with anions such as carboxylate, phosphate or sulphate, or anions formed from acid groups, and the biopolymer (*e.g.* a histone) has a net positive charge, or *vice versa*.

Reactivity with group M

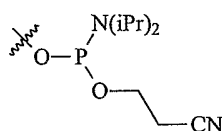
25 The biopolymers have at least one reactive group capable of reacting with M to form a covalent linkage. Such groups typically include naturally occurring groups and groups formed synthetically on the biopolymer.

Naturally occurring groups include lipid groups of lipoproteins (*e.g.* myristoyl, glycosylphosphatidylinositol, ethanolamine phosphoglycerol, palmitate, stearate, S- or N- or O-acyl
30 groups, lipoic acid, isoprenyl, geranylgeranyl, farnesyl, *etc.*), amide, carbohydrate groups of N- and O- glycoproteins, amine groups (*e.g.* on lysine residues or at the N-terminus of a protein), hydroxyl (*e.g.* in β-hydroxyaspartate, β-hydroxyasparagine, 5-hydroxylysine, ¾-hydroxyproline), thiol, sulfhydryl, phosphoryl, sulfate, methyl, acetyl, formyl (*e.g.* on N-terminal methionines from prokaryotes), phenyl, indolyl, guanidyl, hydroxyl, phosphate, methylthio, ADP-ribosyl *etc.*

The reactive group is bound to the biopolymer by one or more covalent bonds (*e.g.* 2 or 3 bonds), which are either single, double or triple covalent bonds (preferably single bonds). Preferably, the reactive group is bound to the biopolymer by one single bond.

- Groups which may be formed naturally or synthetically on the biopolymer and which are bound to the biopolymer by one bond include: $-\text{NR}_2$ *e.g.* $-\text{NHR}$, especially $-\text{NH}_2$; $-\text{SR}$ *e.g.* $-\text{SH}$; $-\text{OR}$ *e.g.* $-\text{OH}$; $-\text{B(R)Y}$; $-\text{BY}_2$; $-\text{C(R)}_2\text{Y}$; $-\text{C(R)Y}_2$; $-\text{CY}_3$; $-\text{C(=Z)Y}$ *e.g.* $-\text{C(=O)Y}$; $-\text{Z-C(=Z)Y}$; $-\text{C(=Z)R}$ *e.g.* $-\text{C(=Z)H}$, especially $-\text{C(=O)H}$; $-\text{C(R)(OH)OR}$; $-\text{C(R)(OR)}_2$; $-\text{S(=O)Y}$; $-\text{Z-S(=O)Y}$; $-\text{S(=O)}_2\text{Y}$; $-\text{Z-S(=O)}_2\text{Y}$; $-\text{S(=O)}_3\text{Y}$; $-\text{Z-S(=O)}_3\text{Y}$; $-\text{P(=Z)(ZR)Y}$ *e.g.* $-\text{P(=O)(OH)Y}$; $-\text{P(=Z)Y}_2$; $-\text{Z-P(=Z)(ZR)Y}$; $-\text{Z-P(=Z)Y}_2$; $-\text{P(=Z)(R)Y}$ *e.g.* $-\text{P(=O)(H)Y}$; $-\text{Z-P(=Z)(R)Y}$; or $-\text{N=C(=Z)}$ *e.g.* $-\text{N=C(=O)}$.
- Another group which may be formed naturally or synthetically on the biopolymer and which is bound to the biopolymer by one bond is $-\text{CN}$.

- Other groups which may be formed naturally or synthetically on the biopolymer and which are bound to the biopolymer by one bond are: $-\text{P(ZR)Y}$ *e.g.* $-\text{P(OH)Y}$; $-\text{PY}_2$; $-\text{Z-P(ZR)Y}$; $-\text{Z-PY}_2$; $-\text{P(R)Y}$ *e.g.* $-\text{P(H)Y}$; $-\text{Z-P(R)Y}$. A particularly preferred group is $-\text{Z-P(ZR)Y}$, especially a phosphoramidite group:



Another example of a group which may be formed naturally or synthetically on the biopolymer and which is bound to the biopolymer by one bond is $-\text{Y}$. In particular, when the reactive group is halo (especially iodo), the reactive group may be bound to an aliphatic or aromatic carbon.

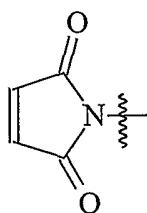
- Groups which may be formed synthetically on the biopolymer and which are bound to the biopolymer by two bonds include $-\text{N(R)-}$ *e.g.* $-\text{NH-}$; $-\text{S-}$; $-\text{O-}$; $-\text{B(Y)-}$; $-\text{C(R)(Y)-}$; $-\text{CY}_2-$; $-\text{C(=O)-}$; $-\text{C(OH)(OR)-}$; $-\text{C(OR)}_2-$.

Groups which may be formed synthetically on the biopolymer and which are bound to the

biopolymer by three bonds include $-\text{C(Y)-}$.

- Preferred groups include nucleophilic groups, either natural or synthetic, *e.g.*: $-\text{NR}_2$ *e.g.* $-\text{NHR}$, especially $-\text{NH}_2$; $-\text{SR}$ *e.g.* $-\text{SH}$; $-\text{OR}$ *e.g.* $-\text{OH}$; $-\text{N(R)-}$ *e.g.* $-\text{NH-}$; $-\text{S-}$; and $-\text{O-}$. The groups $-\text{NH}_2$, $-\text{SH}$ and $-\text{OH}$ are particularly preferred.

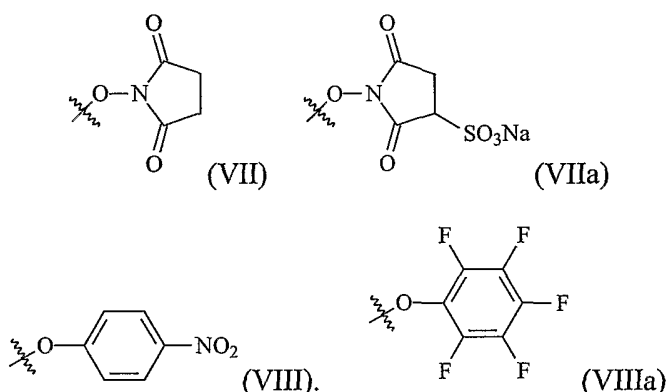
Another preferred reactive group is maleimidyl:



Y is independently a leaving group, including groups capable of leaving in an SN_2 substitution reaction or being eliminated in an addition-elimination reaction with the reactive group of the biopolymer B_p .

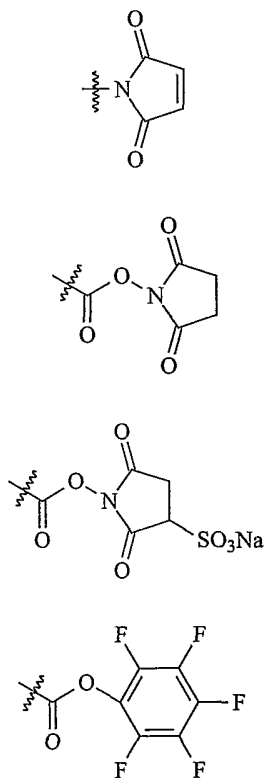
- Preferred examples of Y include halogen (preferably iodo), C_{1-8} hydrocarbyloxy (e.g. C_{1-8} alkoxy), C_{1-8} hydrocarbyloxy substituted with one or more A, C_{1-8} heterohydrocarbyloxy, C_{1-8} heterohydrocarbyloxy substituted with one or more A, mesyl, tosyl, pentafluorophenyl, -O-succinimidyl (formula VII) or a sulfo sodium salt thereof (sulfoNHS – formula VIIa), -S-succinimidyl, or phenyloxy substituted with one or more A e.g. p-nitrophenyloxy (formula VIII) or pentafluorophenoxy (formula VIIIa).

10



Thus, preferred reactive group on the biopolymer are:

15



Other preferred examples of Y include -ZR. Particularly preferred examples of Y are -ZH (e.g. -OH or -NH₂) and -Z- C_{1-8} alkyl groups such as -NH- C_{1-8} alkyl groups (e.g. -NHMe) and -O- C_{1-8} alkyl

groups (e.g. -O-t-butyl). Thus, preferred reactive groups are -C(O)-NH-C₁₋₈alkyl and -C(O)-O-C₁₋₈alkyl (e.g. -C(O)-O-t-butyl).

Other preferred examples of Y include -Z-ZR. Particularly preferred examples include -NR-NR₂, especially -NH-NH₂, and -ONR₂, especially -O-NH₂.

5 Z is independently O, S or N(R). Preferred (=Z) is (=O).

R is independently H, C₁₋₈hydrocarbyl (e.g. C₁₋₈alkyl) or C₁₋₈hydrocarbyl substituted with one or more A.

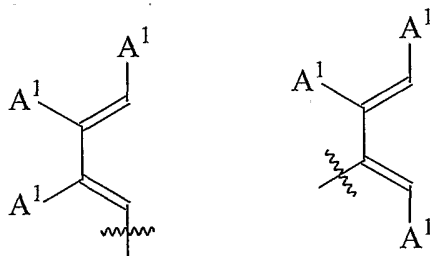
R is preferably H.

Other preferred reactive groups include -C(=O)Y, especially -C(=O)-O-succinimidyl and
10 -C(=O)-O-(p-nitrophenyl).

In a further embodiment, the reactive group may be -Si(R)₂-Y, with Y being halo (e.g. chloro) being especially preferred. Preferred groups R in this embodiment are C₁₋₈alkyl, especially methyl. A particularly preferred reactive group in this embodiment is -Si(Me)₂Cl.

Other groups which may be formed naturally or synthetically on the biopolymer include groups
15 capable of reacting in a cycloaddition reaction, especially a Diels-Alder reaction.

In the case of Diels-Alder reactions, the reactive group on the biopolymer is either a diene or a dienophile. Preferred diene groups are



and multivalent derivatives formally formed by removal of one or more hydrogen atoms, where A¹ is
20 -R¹ or -Z¹R¹, where R¹ and Z¹ are defined below.

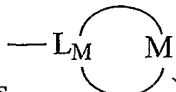
Preferred dienophile groups are -CR¹=CR¹₂, -CR¹=C(R¹)A², -CA²=CR¹₂, -CA²=C(R¹)A² or -CA²=CA²₂, and multivalent derivatives formally formed by removal of one or more hydrogen atoms, where R¹ is defined below and A² is independently halogen, trihalomethyl, -NO₂, -CN, -N⁺(R¹)₂O⁻, -CO₂H, -CO₂R¹, -SO₃H, -SOR¹, -SO₂R¹, -SO₃R¹, -OC(=O)OR¹, -C(=O)H, -C(=O)R¹,
25 -OC(=O)R¹, , -OC(=O)NR¹₂, -N(R¹)C(=O)R¹, -C(=S)NR¹₂, -NR¹C(=S)R¹, -SO₂NR¹₂, -NR¹SO₂R¹, -N(R¹)C(=S)NR¹₂, or -N(R¹)SO₂NR¹₂, where R¹ is defined below. A particularly preferred dienophile group is maleimidyl.

Group M

The group M is a reactive functional group. Reactive functional groups include groups capable of
30 reacting to form a covalent linkage and groups capable of ionic bonding, hydrogen bonding, dipole-

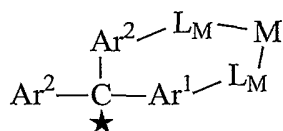
dipole bonding or van der Waals bonding. Particularly preferred groups M are those capable of reacting to form a covalent linkage.

The group M is bound to L_M by one or more covalent bonds (e.g. 2 or 3 bonds, especially 2 such

as , which are either single, double or triple covalent bonds (preferably single bonds).

5 Preferably, M is bound to L_M by one single bond.

Alternatively, or in addition, M is bound by more than one L_M , such L_M either being attached to the same or different Ar^1 or Ar^2 . In a preferred embodiment M is bound by more than one L_M from different Ar^1 or Ar^2 , e.g.:



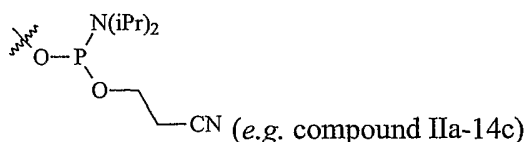
10 Covalent Linkage

Particularly preferred groups M are those capable of reacting to form a covalent linkage. Preferably, the group M is capable of reacting with the reactive group of the biopolymer, B_P , to form a covalent linkage.

Examples of group M bound to L_M by one bond include $-NR_2$ e.g. $-NHR$ (e.g. $-NHMe$ (e.g. compound (IIa-17b)), especially $-NH_2$ (e.g. compounds (IIa-12c) & (IIa-13b)); $-SR$ e.g. $-SH$; $-OR$ e.g. $-OH$ (e.g. compound (IIa-3a)); $-B(R)Y$; $-BY_2$; $-C(R)_2Y$; $-C(R)Y_2$; $-CY_3$; $-C(=Z)Y$ e.g. $-C(=O)Y$; $-Z-C(=Z)Y$; $-C(=Z)R$ e.g. $-C(=Z)H$, especially $-C(=O)H$; $-C(R)(OH)OR$; $-C(R)(OR)_2$; $-S(=O)Y$; $-Z-S(=O)Y$; $-S(=O)_2Y$; $-Z-S(=O)_2Y$; $-S(=O)_3Y$; $-Z-S(=O)_3Y$; $-P(=Z)(ZR)Y$ e.g. $-P(=O)(OH)Y$; $-P(=Z)Y_2$; $-Z-P(=Z)(ZR)Y$; $-Z-P(=Z)Y_2$; $-P(=Z)(R)Y$ e.g. $-P(=O)(H)Y$; $-Z-P(=Z)(R)Y$; or $-N=C(=Z)$ e.g. $-N=C(=O)$.

Another example of a group M bound to L_M by one bond is $-CN$.

Other examples of group M bound to L_M by one bond are $-P(ZR)Y$ e.g. $-P(OH)Y$; $-PY_2$; $-Z-P(ZR)Y$; $-Z-PY_2$; $-P(R)Y$ e.g. $-P(H)Y$; $-Z-P(R)Y$. A particularly preferred group M is $-Z-P(ZR)Y$, especially a phosphoramidite group:



Another example of group M bound to L_M by one bond is $-Y$. In particular, when group M is halo (especially iodo), M may be bound to an aliphatic (e.g. compound (IIa-17c)) or aromatic carbon (e.g. compounds (IIb-28c) & (IIb-28d)). When M is halo (e.g. iodo) and is bound to an aromatic carbon, L_M may, for example, be a single bond.

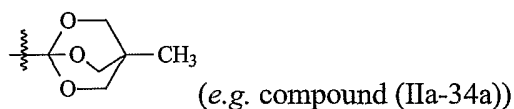
Examples of group M bound to L_M by two bonds include $-N(R)-$ *e.g.* $-NH-$; $-S-$; $-O-$; $-B(Y)-$; $-C(R)(Y)-$; $-CY_2-$; $-C(=O)-$; $-C(OH)(OR)-$; $-C(OR)_2-$.

Examples of group M bound to L_M by three bonds include $\begin{array}{c} | \\ -C(Y)- \end{array}$.

Preferred groups M include electrophilic groups, especially those susceptible to SN_2 substitution reactions, addition-elimination reactions and addition reactions, *e.g.* $-B(R)Y$; $-BY_2$; $-C(R)_2Y$; $-C(R)Y_2$; $-CY_3$; $-C(=Z)Y$ *e.g.* $-C(=O)Y$, especially $-C(O)OH$ (*e.g.* compound 24b) and $-C(O)NH_2$ (*e.g.* compound 19e); $-Z-C(=Z)Y$; $-C(=Z)R$ *e.g.* $-C(=Z)H$, especially $-C(=O)H$; $-C(R)(OH)OR$; $-C(R)(OR)_2$; $-S(=O)Y$; $-Z-S(=O)Y$; $-S(=O)_2Y$; $-Z-S(=O)_2Y$; $-S(=O)_3Y$; $-Z-S(=O)_3Y$; $-P(=Z)(ZR)Y$ *e.g.* $-P(=O)(OH)Y$; $-P(=Z)Y_2$; $-Z-P(=Z)(ZR)Y$; $-Z-P(=Z)Y_2$; $-P(=Z)(R)Y$ *e.g.* $-P(=O)(R)Y$; $-Z-P(=Z)(H)Y$; $-N=C(=Z)$ *e.g.* $-N=C(=O)$; $-B(Y)-$; $-C(R)(Y)-$; $-CY_2-$; $-C(=O)-$; $-C(OH)(OR)-$; $-C(OR)_2-$; or $\begin{array}{c} | \\ -C(Y)- \end{array}$.

Another preferred electrophilic group M is $-CN$.

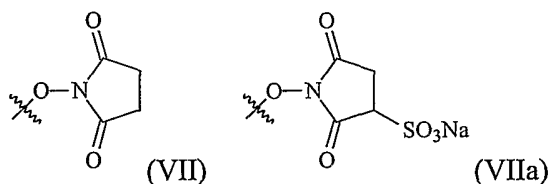
Still further preferred examples of group M are orthoesters, *e.g.* $-C(OR)_3$. In a preferred embodiment, the R groups are linked together to form a hydrocarbyl group, *e.g.* a C_{1-8} alkyl group. A preferred example of group M in this embodiment is:

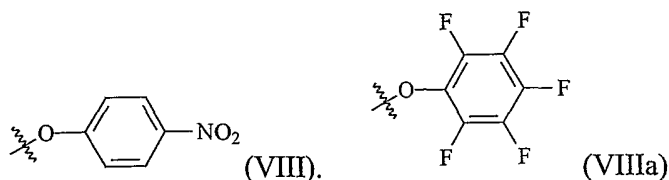


Another preferred group M is maleimido (*e.g.* compound (IIa-18d)).

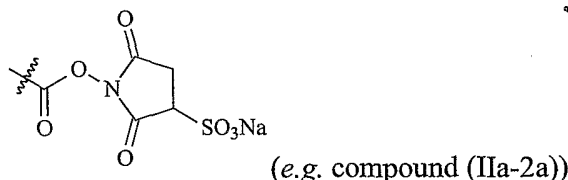
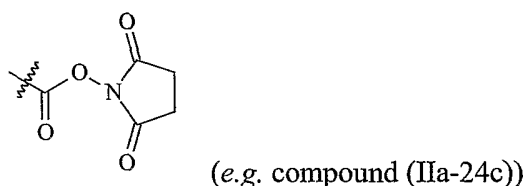
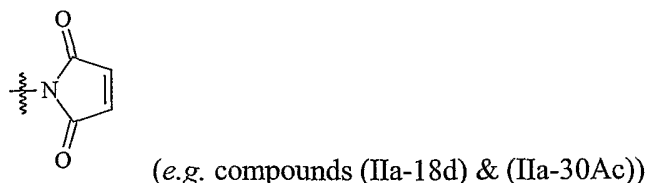
Y is independently a leaving group, including groups capable of leaving in an SN_2 substitution reaction or being eliminated in an addition-elimination reaction.

Preferred examples of Y include halogen (preferably iodo), C_{1-8} hydrocarbyloxy (*e.g.* C_{1-8} alkoxy), C_{1-8} hydrocarbyloxy substituted with one or more A, C_{1-8} heterohydrocarbyloxy, C_{1-8} heterohydrocarbyloxy substituted with one or more A, mesyl, tosyl, pentafluorophenyl, $-O$ -succinimidyl (formula VII) or a sulfo sodium salt thereof (sulfoNHS – formula VIIa), $-S$ -succinimidyl, or phenyloxy substituted with one or more A *e.g.* *p*-nitrophenyloxy (formula VIII) or pentafluorophenoxy (formula VIIa) (*e.g.* compound (IIa-16)).

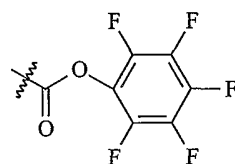




Thus, preferred groups M are:



5



Other preferred examples of Y include -ZR. Particularly preferred examples of Y are -ZH (e.g. -OH or -NH₂) and -Z-C₁₋₈alkyl groups such as -NH-C₁₋₈alkyl groups (e.g. -NHMe) and -O-C₁₋₈alkyl groups (e.g. -O-t-butyl). Thus, preferred groups M are -C(O)-NH-C₁₋₈alkyl (e.g. -C(O)NHMe) and

10 -C(O)-O-C₁₋₈alkyl (e.g. -C(O)-O-t-butyl (e.g. compounds (IIa-24a) & (IIa-33a)).

Other preferred examples of Y include -Z-ZR. Particularly preferred examples include -NR-NR₂, especially -NH-NH₂ (e.g. compounds (IIa-35Ab), (IIa-35Bc) and (IIa-35Bd)), and -ONR₂, especially -O-NH₂ (e.g. compounds (IIa-35Cc) and (IIa-35Cd)).

Z is independently O, S or N(R). Preferred (=Z) is (=O).

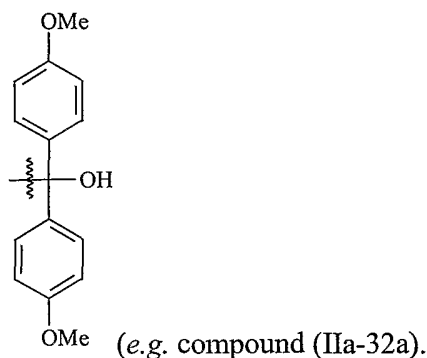
15 R is independently H, C₁₋₈hydrocarbonyl (e.g. C₁₋₈alkyl) or C₁₋₈hydrocarbonyl substituted with one or more A.

R is preferably H.

Particularly preferred groups M include -C(=O)Y, especially -C(=O)-O-succinimidyl and -C(=O)-O-(p-nitrophenyl).

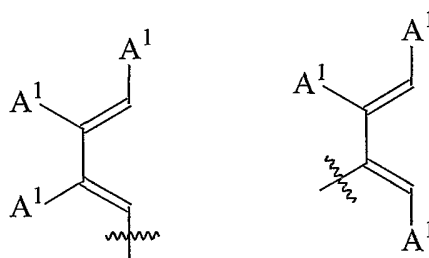
In a further embodiment, M may be $-\text{Si}(\text{R})_2\text{-Y}$, with Y being halo (*e.g.* chloro) being especially preferred. Preferred groups R in this embodiment are C_{1-8} alkyl, especially methyl. A particularly preferred group M in this embodiment is $-\text{Si}(\text{Me})_2\text{Cl}$ (*e.g.* compound (IIa-19d)).

In a further embodiment, M may be $-\text{C}(\text{Ar}^2)_2\text{X}$. Preferred groups Ar^2 and X are set out below. In this
5 embodiment it is preferred that L_M is a bond. A particularly preferred group M in this embodiment is:



Other groups M include groups capable of reacting in a cycloaddition reaction, especially a Diels-Alder reaction.

In the case of Diels-Alder reactions, the group M is either a diene or a dienophile. Preferred diene
10 groups are



and multivalent derivatives formally formed by removal of one or more hydrogen atoms, where A^1 is $-\text{R}^1$ or $-\text{Z}^1\text{R}^1$, where R^1 and Z^1 are defined below.

Preferred dienophile groups are $-\text{CR}^1=\text{CR}^1_2$, $-\text{CR}^1=\text{C}(\text{R}^1)\text{A}^2$, $-\text{CA}^2=\text{CR}^1_2$, $-\text{CA}^2=\text{C}(\text{R}^1)\text{A}^2$ or
15 $-\text{CA}^2=\text{CA}^2_2$, and multivalent derivatives formally formed by removal of one or more hydrogen atoms, where R^1 is defined below and A^2 is independently halogen, trihalomethyl, $-\text{NO}_2$, $-\text{CN}$, $-\text{N}^+(\text{R}^1)_2\text{O}^-$, $-\text{CO}_2\text{H}$, $-\text{CO}_2\text{R}^1$, $-\text{SO}_3\text{H}$, $-\text{SOR}^1$, $-\text{SO}_2\text{R}^1$, $-\text{SO}_3\text{R}^1$, $-\text{OC}(=\text{O})\text{OR}^1$, $-\text{C}(=\text{O})\text{H}$, $-\text{C}(=\text{O})\text{R}^1$, $-\text{OC}(=\text{O})\text{R}^1$, $-\text{OC}(=\text{O})\text{NR}^1_2$, $-\text{N}(\text{R}^1)\text{C}(=\text{O})\text{R}^1$, $-\text{C}(=\text{S})\text{NR}^1_2$, $-\text{NR}^1\text{C}(=\text{S})\text{R}^1$, $-\text{SO}_2\text{NR}^1_2$, $-\text{NR}^1\text{SO}_2\text{R}^1$, $-\text{N}(\text{R}^1)\text{C}(=\text{S})\text{NR}^1_2$, or $-\text{N}(\text{R}^1)\text{SO}_2\text{NR}^1_2$, where R^1 is defined below. A particularly preferred dienophile
20 group is maleimidyl.

Preferred examples of group M are shown in figures 2A and 2B.

Ionic Bonding

Where group M is a reactive functional group capable of ionic bonding, group M typically comprises one or more chelating ligands.

Suitable chelating ligands which can bind anions include polyamines and cryptands.

Suitable chelating ligands which can bind cations include polyacidic compounds (*e.g.* EDTA) and crown ethers.

Hydrogen Bonding

- 5 Where group M is a reactive functional group capable of hydrogen bonding, M will typically bear one or more hydroxy, amino or thio hydrogen atoms or a group bearing an atom having one or more lone pair of electrons (*e.g.* an oxygen, sulphur or nitrogen atom).

Preferred groups capable of hydrogen bonding include biotin, avidin and streptavidin.

Dipole-Dipole Bonding

- 10 Where group M is a reactive functional group capable of dipole-dipole bonding, the dipole-dipole bond may be formed between permanent dipoles or between a permanent dipole and an induced dipole.

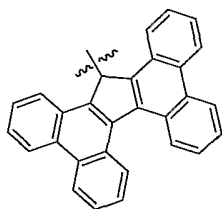
Preferred groups M capable of dipole-dipole bonding comprise acid groups, or $-(NMe_3)^+$, carboxy, carboxylate, phosphate or sulphate groups.

15 *Van der Waals Bonding*

Where group M is a reactive functional group capable van der Waals bonding, M will typically comprise a hydrocarbyl or heterohydrocarbyl group (usually a large hydrocarbyl group having at least ten carbon atoms up to about 50 carbon atoms), optionally substituted with one or more A. Polyfluorinated hydrocarbyl and heterohydrocarbyl groups are particularly preferred. Typically, the

- 20 hydrocarbyl or heterohydrocarbyl groups are aryl or heteroaryl groups or groups of the formula $-C(R^6)_2Ar^3$, $-C(R^6)(Ar^3)_2$ or $-C(Ar^3)_3$, where Ar^3 is independently defined the same as Ar^2 and R^6 is H, C_{1-8} hydrocarbyl, C_{1-8} hydrocarbyl substituted by one or more A, C_{1-8} heterohydrocarbyl or C_{1-8} heterohydrocarbyl substituted by one or more A.

A preferred group capable of van der Waals bonding is tetrabenzofullerene (formula X).



(formula X)

25

Other preferred groups capable of van der Waals bonding are adamantyl (*e.g.* 2-adamantyl (*e.g.* compound (IIa-36a)) and phenyl (*e.g.* example (IIa-37b)).

Preferably, these groups are linked to a hydrocarbylene group (*e.g.* C_{1-8} alkylene group) which forms L_M or a part thereof.

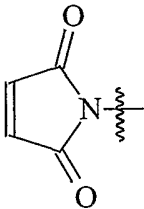
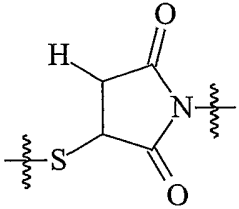
Matching B_P and M

The reactive group on the biopolymer and the group M must be dependently selected in order to form the covalent linkage. For example, where the biopolymer includes the groups $-NH_2$, $-OH$ or $-SH$, M will typically be $-B(R)Y$; $-BY_2$; $-C(R)_2Y$; $-C(R)Y_2$; $-CY_3$; $-C(=Z)Y$ e.g. $-C(=O)Y$; $-Z-C(=Z)Y$; $-C(=Z)R$ e.g. $-C(=Z)H$, especially $-C(=O)H$; $-C(R)(OH)OR$; $-C(R)(OR)_2$; $-S(=O)Y$; $-Z-S(=O)Y$; $-S(=O)_2Y$; $-Z-S(=O)_2Y$; $-S(=O)_3Y$; $-Z-S(=O)_3Y$; $-P(=Z)(ZR)Y$ e.g. $-P(=O)(OH)Y$; $-P(=Z)Y_2$; $-Z-P(=Z)(ZR)Y$; $-Z-P(=Z)Y_2$; $-P(=Z)(R)Y$ e.g. $-P(=O)(H)Y$; $-Z-P(=Z)(R)Y$; $-N=C(=Z)$ e.g. $-N=C(=O)$; $-B(Y)-$; $-C(R)(Y)-$; $-CY_2-$; $-C(=O)-$; $-C(OH)(OR)-$; $-C(OR)_2-$; or $\text{---}\overset{\textstyle |}{\text{C}}(Y)\text{---}$. M may also be $-CN$.

- 10 In a preferred embodiment, one of the reactive group on the biopolymer and group M is a maleimidyl and the other will be a $-SH$ group.

Alternatively, when the covalent linkage is to be formed by a Diels Alder reaction, one of the reactive group on the biopolymer and group M will typically be a diene and the other will be a dienophile.

- 15 Preferred covalent linkages are those produced through the reaction of the following groups:

M	Group on B_P	Obtained Linkage $M'-B_P'$
$-C(=O)-O\text{-succinimidyl}$ [<i>i.e.</i> carboxy-NHS]	$-NH_2$	$-CO-NH-$
$-C(=O)-O\text{-(p-nitrophenyl)}$	$-NH_2$	$-CO-NH-$
$-C(=O)-\text{pentafluorophenyl}$	$-NH_2$	$-CO-NH-$
Biotin	avidin / streptavidin	biotin-(strept)avidin
	$-SH$	
$-N=C=S$ (isothiocyanate)	$-NH_2$	$-NH-CS-NH-$

The covalent residue $M'-B_P'$ is the reaction product of M and B_P . B_P' will generally be the same as B_P except that instead of the reactive group, B_P' will have a residue of the reactive group covalently bound to the residue M' . Depending on the choice of the reactive group and the choice of M , M' and the residue of the reactive group will typically form linkages, in the orientation $L_M-M'-B_P'$, including $-C(R)_2Z-$, $-ZC(R)_2-$, $-C(=Z)Z-$, $-ZC(=Z)-$, $-ZC(=Z)Z-$, $-C(OH)(R)Z-$, $-ZC(OH)(R)-$, $-C(R)(OR)Z-$, $-ZC(R)(OR)-$, $-C(R)(OR)Z-$, $-ZC(R)(OR)-$, $-S(=O)Z-$, $-ZS(=O)-$, $-ZS(=O)Z-$, $-S(=O)_2Z-$, $-ZS(=O)_2-$, $-ZS(=O)_2Z-$, $-S(=O)_3Z-$, $-ZS(=O)_3-$, $-ZS(=O)_3Z-$, $-P(=Z)(ZR)Z-$, $-ZP(=Z)(ZR)-$, $-ZP(=Z)(ZR)Z-$, $-P(=Z)(R)Z-$, $-ZP(=Z)(R)-$, $-ZP(=Z)(R)Z-$, $-NH-C(=Z)-Z-$, where Z and R are as defined above.

Group M''

M'' is the same as M except that S_S is bound to a portion of M which does not form part of the residue of M'' remaining attached to the ion of formula (I') which residue is produced after reaction of group M''. Thus, M'' is a residue of M formable by the conjugation of M and S_S. However, M'' need
 5 not necessarily be formed by the conjugation of M and S_S.

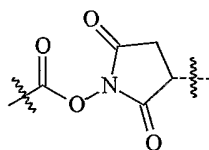
M''--S_S comprises a covalent, ionic, dipole-dipole, hydrogen, or van der Waals bond. The covalent, ionic, hydrogen, dipole-dipole or van der Waals bond may be direct between M'' and S_S or may be provided by one or more binding groups present on M'' and/or S_S.

Examples of groups which can form these types of bond, and methods for cleaving these types of
 10 bond, are set out below in connection with C--S_S bonds, *etc.*

This embodiment of the invention is advantageous, since the derivatisation of the biopolymer will also release the derivatised biopolymer from the solid support. Thus, an additional step of cleaving the biopolymer from the solid support is not required.

Preferred groups M'' are groups M having a leaving group, wherein the group S_S is bound to the
 15 leaving group, *e.g.* groups M mentioned above having a leaving group Y, wherein the group S_S is bound to the leaving group Y.

A particularly preferred group M'' is:

**L_M**

20 Where the group L_M is a linker atom or group, it has a sufficient number of linking covalent bonds to link L_M to the group Ar¹ by a single covalent bond (or more, as appropriate) and to link L_M to the p instances of M groups (which may be attached to L_M by one or more bonds).

The group L_M may be directly bound to the aromatic part of Ar¹, bound to one or more of the substituents A of Ar¹, or both. Preferably, L_M is bound directly to the aromatic part of Ar¹.

25 In an alternative embodiment, L_M may be bound to L₅.

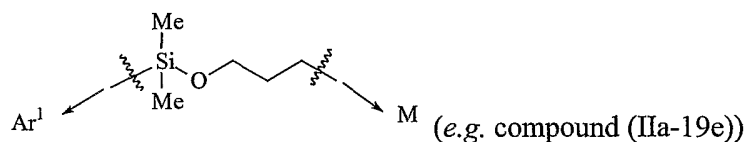
When L_M is a linker atom, preferred linker atoms are O or S, particularly O.

When L_M is a linker group, preferred linker groups, in the orientation Ar¹-(L_M{M})_p, are -E^M-, -(D^M)_t-, -(E^M-D^M)_t-, -(D^M-E^M)_t-, -E^M-(D^M-E^M)_t- or -D^M-(E^M-D^M)_t-, where a sufficient number of linking covalent bonds, in addition to the covalent bonds at the chain termini shown, are provided on
 30 groups E^M and D^M for linking the p instances of M groups.

D^M is independently C_{1-8} hydrocarbylene or C_{1-8} hydrocarbylene substituted with one or more A. Preferred D^M are C_{1-8} alkylene, C_{1-8} alkenylene and C_{1-8} alkynylene, especially C_{1-8} alkylene and C_{1-8} alkynylene, each optionally substituted with one or more A (preferably unsubstituted). A preferred substituent A is 2H . Preferred L_M in the orientation $Ar^1-(L_M\{M\}_p)_q$ are: $-CH_2CH_2-$ (e.g. compounds 1a & 2a); $-C=C-CH_2CH_2CH_2-$ (e.g. compounds (IIa-6b), (IIa-6c), (IIa-6d), (IIa-7a), (IIa-7b) & (IIa-7c)); $-(CH_2)_5-$ (e.g. compounds (IIa-8a), (IIa-8b) & (IIa-8c)); $-CD_2CD_2CH_2CH_2CH_2-$; $-C=C-CH_2-$ (e.g. compounds (IIa-12b) & (IIa-12c)) and $-CH_2CH_2CH_2-$ (e.g. compounds (II-4a), (IIa-5a), (IIa-13a) & (IIa-13b)).

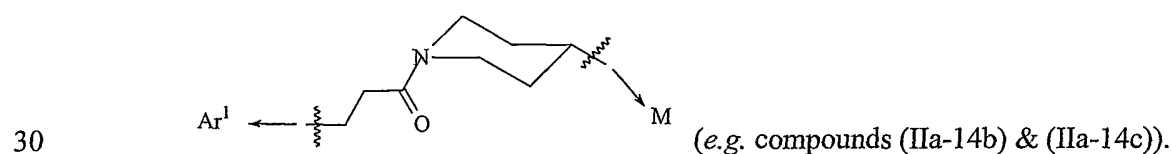
E^M , in the orientation $Ar^1-(L_M\{M\}_p)_q$, is independently $-Z^M-$, $-C(=Z^M)-$, $-Z^MC(=Z^M)-$, $-C(=Z^M)Z^M-$, $-Z^MC(=Z^M)Z^M-$, $-S(=O)-$, $-Z^MS(=O)-$, $-S(=O)Z^M-$, $-Z^MS(=O)Z^M-$, $-S(=O)_2-$, $-Z^MS(=O)_2-$, $-S(=O)_2Z^M-$, $-Z^MS(=O)_2Z^M-$, where Z^M is independently O, S or $N(R^M)$ and where R^M is independently H, C_{1-8} hydrocarbyl (e.g. C_{1-8} alkyl) or C_{1-8} hydrocarbyl substituted with one or more A. Preferably E^M is, in the orientation $Ar^1-(L_M\{M\}_p)_q$, $-O-$, $-S-$, $-C(=O)-$, $-C(=O)O-$, $-C(=S)-$, $-C(=S)O-$, $-OC(=S)-$, $-C(=O)S-$, $-SC(=O)-$, $-S(O)_2-$, $-NR^M-$, $-C(=O)N(R^M)-$, $-C(=S)N(R^M)-$, $-N(R^M)C(=O)-$, $-N(R^M)C(=S)-$, $-S(=O)N(R^M)-$, $-N(R^M)S(=O)-$, $-S(=O)_2N(R^M)-$, $-N(R^M)S(=O)_2-$, $-OC(=O)O-$, $-SC(=O)O-$, $-OC(=O)S-$, $-N(R^M)C(=O)O-$, $-OC(=O)N(R^M)-$, $-N(R^M)C(=O)N(R^M)-$, $-N(R^M)C(=S)N(R^M)-$, $-N(R^M)S(=O)N(R^M)-$ or $-N(R^M)S(=O)_2N(R^M)-$.

Alternative groups E^M to those defined above, in the orientation $Ar^1-(L_M\{M\}_p)_q$, are $-Z^M-Si(R^M)_2-Z^M-$, $-Si(R^M)_2-Z^M-$ and $-Z^M-Si(R^M)_2-$. The group $-Si(R^M)_2-Z^M-$ is particularly preferred. Z^M is preferably O. R^M is preferably C_{1-8} alkyl, preferably methyl. These groups E^M are particularly preferred in the groups $-(E^M-D^M)_t-$, especially when $t=1$ and D^M is C_{1-8} alkylene. The following group is especially preferred:



In addition to the above definition of D^M , D^M may also be C_{1-8} heterohydrocarbylene or C_{1-8} heterohydrocarbylene substituted with one or more A. In this embodiment,

C_{1-8} cycloheteroalkylene groups are particularly preferred, e.g.: Thus, preferred L_M groups $-D^M-E^M-D^M-$ are, in the orientation $Ar^1-(L_M\{M\}_p)_q$, $-C_{1-8}$ alkylene- $C(O)-C_{1-8}$ cycloheteroalkylene (preferably where the hetero atom is N and is bound to the carboxy), especially:



$t = 1$ or more, *e.g.* from 1 to 50, 1 to 40, 1 to 30, 1 to 20 or 1 to 10. Preferably $t = 1, 2, 3, 4, 5, 6, 7, 8, 9$, or 10.

Preferably, L_M links one group M to Ar^1 , M is linked to L_M by a single covalent bond and therefore no additional bonds are required (*e.g.* $L_M\{M\}_1$ may be $-E^M-\{M\}$, $-(D^M)_t-\{M\}$, $-(E^M-D^M)_t-\{M\}$,

5 $-(D^M-E^M)_t-\{M\}$, $-E^M-(D^M-E^M)_t-\{M\}$ or $-D^M-(E^M-D^M)_t-\{M\}$).

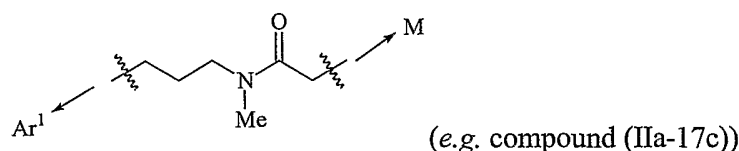
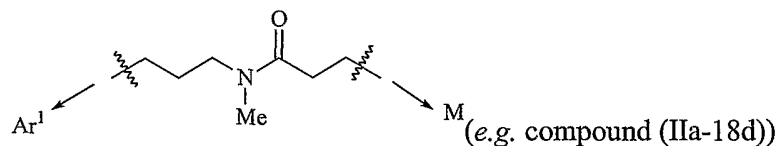
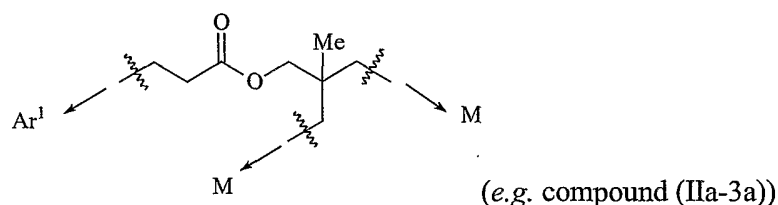
Where L_M includes a group which also falls within the definition of group M , the group M is preferably more reactive than the group included in L_M .

L_M is preferably $-(D^M)_t-$, $-(E^M-D^M)_t-$, or $-D^M-(E^M-D^M)_t-$.

When group L_M is $-(D^M)_t-$, t is preferably 1. D^M is preferably C_{1-8} alkylene, preferably C_{1-5} alkylene, preferably methylene or ethylene.

When group L_M is $-(E^M-D^M)_t-$, or $-D^M-(E^M-D^M)_t-$, E^M is preferably (in the orientation $Ar^1-(L_M\{M\}_p)_q$), $-C(=O)N(R^M)-$ (*e.g.* $-C(=O)NH-$) or O (preferably O), and D^M is preferably C_{1-8} alkylene, preferably ethylene, propylene, butylene or pentylene. t is preferably 1. Especially preferred L_M are, in the orientation $Ar^1-(L_M\{M\}_p)_q$, $-O-CH_2CH_2CH_2-$ (*e.g.* compounds (IIa-15a), (IIa-15b), (IIa-15c) & (IIa-16a)) and $-O-CH_2CH_2CH_2CH_2CH_2-$ (*e.g.* compounds (IIa-10a), (IIa-10b), (IIa-10c), (IIa-11a), (IIa-11b) & (IIa-11c)).

Another preferred group $-D^M-(E^M-D^M)_t-$ is where D^M is C_{1-8} alkylene and t is 1. Preferred E^M in this group, in the orientation $Ar^1-(L_M\{M\}_p)_q$, are $-Z^MC(=Z^M)-$ (especially $-N(R^M)C(=O)-$, *e.g.* $-N(Me)C(=O)-$) and $-C(=Z^M)Z^M-$ (especially $-C(=O)O-$). Particularly preferred L_M groups are:



The group $-(E^M-D^M)_t-$ is preferred, a particularly preferred example of which is (in the orientation $Ar^1-(L_M\{M\}_p)_q$) $-C(=O)NH-CH_2CH_2CH_2-O-CH_2CH_2-O-CH_2CH_2-O-CH_2CH_2CH_2-$.

The group $-(D^M-E^M)_t-$ is also preferred when D^M is C_{1-8} alkylene and t is 1. Preferred E^M in this group, in the orientation $Ar^1-(L_M\{M\}_p)_q$ are $-Z^MC(=Z^M)-$ and $-C(=Z^M)Z^M-$, especially $-Z^MC(=Z^M)-$ (particularly $-N(R^M)C(=O)-$, e.g. $-N(Me)C(=O)-$). A particularly preferred example is $-CH_2CH_2CH_2N(Me)C(O)-$.

- 5 In an alternative embodiment it is preferred that L_M is a single covalent bond.

When Ar^2 is phenyl, L_M is preferably provided in a position ortho or para to $C^\star$. When Ar^2 is other than phenyl, L_M is preferably attached to an atom which bears the charge in at least one of the resonance structures of the ions of formula (I').

- Where $C^\star$ is a cation, L_M is preferably an electron-donating group. Where $C^\star$ is an anion, L_M is preferably an electron-withdrawing group.

Preferred examples of L_M are shown in figures 1A and 1B.

C--S_S, S_S--Ar¹ and S_S--Ar² Bonds

- $C--S_S$, S_S--Ar^1 and S_S--Ar^2 comprise a cleavable covalent, ionic, hydrogen, dipole-dipole or van der Waals bond (also known as a dispersion bond or a London forces bond). The covalent, ionic, hydrogen, dipole-dipole or van der Waals bond may be direct between C and S_S , Ar^1 and S_S , or Ar^2 and S_S , or may be provided by one or more binding groups present on C and/or S_S , Ar^1 and/or S_S , or Ar^2 and/or S_S , respectively.

Covalent Bonding

- Where the bond is covalent, the bond may be direct (e.g. $C-S_S$, Ar^1-S_S or Ar^2-S_S , respectively) or may be provided by a linker atom or group L^4 (e.g. $C-L^4-S_S$, $Ar^1-L^4-S_S$ or $Ar^2-L^4-S_S$, respectively).

When L^4 is a linker group, preferred linker groups are $-E^4-$, $-(D^4)_t--$, $-(E^4-D^4)_t--$, $-(D^4-E^4)_t--$, $-E^4-(D^4-E^4)_t--$ or $-D^4-(E^4-D^4)_t--$.

D^4 is independently C_{1-8} hydrocarbylene or C_{1-8} hydrocarbylene substituted with one or more A.

- E^4 is, in the orientation $C-L^4-S_S$, independently $-Z^4-$, $-C(=Z^4)-$, $-Z^4C(=Z^4)-$, $-C(=Z^4)Z^4-$, $-Z^4C(=Z^4)Z^4-$, $-S(=O)-$, $-Z^4S(=O)-$, $-S(=O)Z^4-$, $-Z^4S(=O)Z^4-$, $-S(=O)_2-$, $-Z^4S(=O)_2-$, $-S(=O)_2Z^4-$, $-Z^4S(=O)_2Z^4-$, where Z^4 is independently O, S or $N(R^4)$, and where R^4 is independently H, C_{1-8} hydrocarbyl (e.g. C_{1-8} alkyl) or C_{1-8} hydrocarbyl substituted with one or more A. Preferably E^4 is, in the orientation $C-L^4-S_S$, $-O-$, $-S-$, $-C(=O)-$, $-C(=O)O-$, $-C(=S)-$, $-C(=S)O-$, $-OC(=S)-$, $-C(=O)S-$, $-SC(=O)-$, $-S(O)-$, $-S(O)_2-$, $-N(R^4)-$, $-C(=O)N(R^4)-$, $-C(=S)N(R^4)-$, $-N(R^4)C(=O)-$, $-N(R^4)C(=S)-$, $-S(=O)N(R^4)-$, $-N(R^4)S(=O)-$, $-S(=O)_2N(R^4)-$, $-N(R^4)S(=O)_2-$, $-OC(=O)O-$, $-SC(=O)O-$, $-OC(=O)S-$, $-N(R^4)C(=O)O-$, $-OC(=O)N(R^4)-$, $-N(R^4)C(=O)N(R^4)-$, $-N(R^4)C(=S)N(R^4)-$, $-N(R^4)S(=O)N(R^4)-$ or $-N(R^4)S(=O)_2N(R^4)-$.

$t'' = 1$ or more, e.g. from 1 to 50, 1 to 40, 1 to 30, 1 to 20 or 1 to 10. Preferably $t'' = 1, 2, 3, 4, 5, 6, 7, 8, 9$, or 10.

Where L^4 includes a group which also falls within the definition of group M, the group M is preferably more reactive than the group included in L^5 .

L^4 is preferably a linker atom, preferably O or S, particularly O.

When the solid support S_S is gold, L^4 is preferably covalently attached to the S_S by a sulphide or
5 disulphide group.

Ionic Bonding

Where the bond is ionic, the bond is typically direct (e.g. $C^\star S_S^\star$, where $S_S^\star$ is a solid support counterion to $C^\star$).

Alternatively, it may be provided by binding groups, e.g. chelating ligands, present on C or S_S , Ar^1 or
10 S_S , or Ar^2 or S_S , respectively. In the case of C- - S_S bonds, the chelating ligand is typically only present on S_S and chelates with $C^\star$.

Suitable chelating ligands which can bind anions include polyamines and cryptands.

Suitable chelating ligands which can bind cations include polyacidic compounds (e.g. EDTA) and crown ethers.

15 *Hydrogen Bonding*

Where the bond is a hydrogen bond, the bond is usually provided by binding groups present on C or S_S , Ar^1 or S_S , or Ar^2 or S_S , respectively.

Typically, in order to form the hydrogen bond, one of C or S_S , Ar^1 or S_S , or Ar^2 or S_S , as appropriate, will have a binding group bearing one or more hydroxy, amino or thio hydrogen atoms, and the other
20 of C or S_S , Ar^1 or S_S , or Ar^2 or S_S , respectively, will have a binding group bearing an atom having one or more lone pair of electrons (e.g. an oxygen, sulphur or nitrogen atom). Preferably, one of C or S_S , Ar^1 or S_S , or Ar^2 or S_S , as appropriate, will have a binding group comprising biotin, and the other of C or S_S , Ar^1 or S_S , or Ar^2 or S_S , respectively, will have a binding group comprising avidin or streptavidin.

25 Alternatively, the hydrogen bond may be direct.

Dipole-Dipole Bonding

Where the bond is a dipole-dipole bond, it may be formed between permanent dipoles or between a permanent dipole and an induced dipole.

Typically, in order to form the dipole-dipole bond, one of S_S and the compound of the invention has
30 a permanent dipole and the other of S_S and the compound of the invention has an induced dipole or a permanent dipole, the attraction between the dipoles forming a dipole-dipole bond.

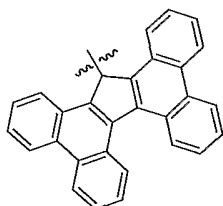
Preferably, S_s comprises binding groups (*e.g.* acid groups, $-(NMe_3)^+$, carboxy, carboxylate, phosphate or sulphate groups) which produce a dipole at the surface of the solid support to bind the compound of the invention.

Van der Waals Bonding

- 5 Where the bond is a van der Waals bond, the bonding is usually provided by binding groups present on C or S_s , Ar^1 or S_s , or Ar^2 or S_s , respectively.

Typically, in order to form the van der Waals bond, at least one, but preferably both, of C or S_s , Ar^1 or S_s , or Ar^2 or S_s , as appropriate, will have a hydrocarbyl or heterohydrocarbyl group (usually a large hydrocarbyl group having at least ten carbon atoms up to about 50 carbon atoms), optionally
 10 substituted with one or more A. Polyfluorinated hydrocarbyl and heterohydrocarbyl groups are particularly preferred. Typically, the hydrocarbyl or heterohydrocarbyl groups are aryl or heteroaryl groups or groups of the formula $-C(R^6)_2Ar^3$, $-C(R^6)(Ar^3)_2$ or $-C(Ar^3)_3$, where Ar^3 is independently defined the same as Ar^2 and R^6 is H, C_{1-8} hydrocarbyl, C_{1-8} hydrocarbyl substituted by one or more A, C_{1-8} heterohydrocarbyl or C_{1-8} heterohydrocarbyl substituted by one or more A.

- 15 A preferred binding group is tetrabenzofullerene (formula X).



(formula X)

Alternatively, the van der Waals bond may be direct.

Bond Cleavage

Preferably, the ions of formula (I') have a pK_{r+} value of at least zz , where zz is 0 or more (*e.g.* 0, 1, 2,
 20 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14). More preferably, zz is 1 or more, still more preferably 2 or more, still more preferably 3 or more.

Preferably, the ions of formula (I) have a pK_{r+} value of at least zz , where zz is defined above.

Preferably, the compounds of formula (IIa), (IIb), (IIIa) or (IIIb) or the solid supports of formula (IVai), (IVaii), (IVaiii), (IVbii), (IVbiii), (IVaiv) or (IVbiv) provide ions of formula (I') having a pK_{r+}
 25 value of at least zz , where zz is defined above.

C-X Bonds

The C-X bonds are cleavable by irradiation, electron bombardment, electrospray, fast atom bombardment (FAB), inductively coupled plasma (ICP) or chemical ionisation. Preferably, the C-X bonds are cleavable by irradiation or chemical ionisation.

The term 'irradiation' includes, for example, laser illumination, in particular as used in MALDI mass spectrometry. Laser light of about 340 nm is particularly preferred because it is typically used in MALDI mass spectrometers.

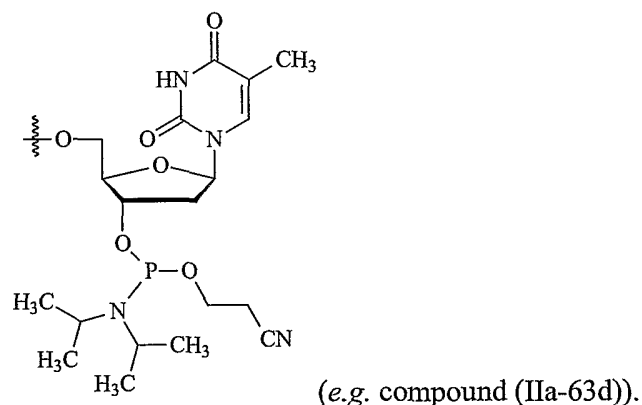
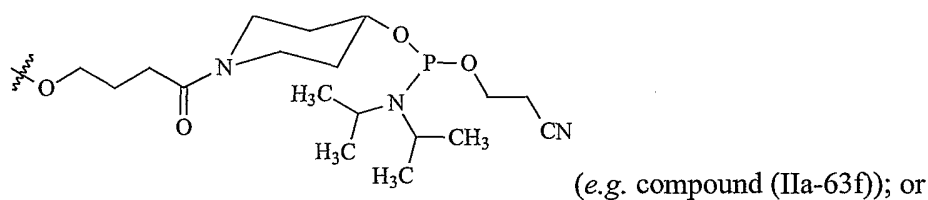
The term 'electron bombardment' includes, for example, bombardment with electrons having energy
5 of about 70 ev.

Chemical ionisation can be effected, for example, by treatment with acid or acidic matrices (*e.g.* acidic matrices used in MALDI analysis).

Preferably group X is halogen, hydroxy, C₁₋₈hydrocarbyloxy, C₁₋₈hydrocarbyloxy substituted with one or more A, C₁₋₈heterohydrocarbyloxy, C₁₋₈heterohydrocarbyloxy substituted with one or more
10 A, mesyl, tosyl, pentafluorophenyl, -O-succinimidyl -S-succinimidyl, or phenyloxy substituted with one or more A *e.g.* p-nitrophenyloxy. The groups pentafluorophenyl, -O-succinimidyl, -S-succinimidyl, and p-nitrophenyloxy are preferred.

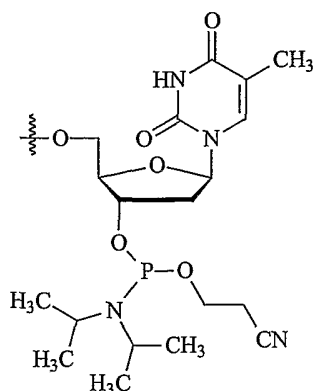
Particularly preferred groups X are halogen, hydroxy, C₁₋₈hydrocarbyloxy. Especially preferred groups are hydroxy (*e.g.* compounds (IIa-61a) & (IIa-62a)), ethoxy (*e.g.* compound (IIa-14a)) and
15 chloro (*e.g.* compound (IIa-64b)) groups.

Other preferred groups X are alkyl ethers, *e.g.*:

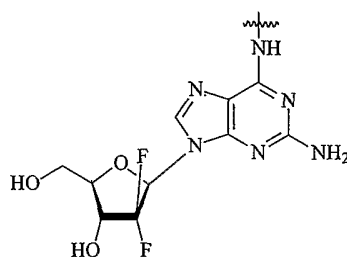


Group X may also be a -Q-oligonucleotide, where Q is O, S or N(R), where R is H, C₁₋₈hydrocarbyl
20 or C₁₋₈hydrocarbyl substituted with one or more A. Q is preferably O.

Group X may also be a nucleoside, preferably where the nucleoside is bound via its 5' end, *e.g.*:



In some embodiments of the invention, where B_p is an antibody (particularly where it is a monoclonal antibody that recognises a tumour-associated antigen), X is not:



- 5 or, optionally, X is not any other 2,6-diaminopurine nucleoside prodrug group.

In some embodiments of the invention, X is not H. If X is H, preferably at least one of Ar^1 and Ar^2 is polycyclic, heterocyclic or unsubstituted.

Preferred examples of group X are shown in figure 4.

Ionic $C \star X \star$ Bonds

- 10 $X \star$ is any counterion for forming salts with compounds of the invention.

$X \star$ includes ions having single charges and multiple charges. Typically ions having multiple charges will be associated with an appropriate number of compounds of formula (IIb), (IVbii), (IVbiii) or (IVbiv), in order to balance the charge. Ions having multiple charges include doubly charged ions (e.g. SO_4^{2-}) and triply charged ions. $X \star$ preferably has a single charge.

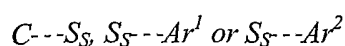
- 15 The counterion $X \star$ may be dissociated from the derivative of formula (IIb), (IVbii), (IVbiii), (IVbiv) or (Vbii) by irradiation, electron bombardment, electrospray, fast atom bombardment (FAB), inductively coupled plasma (ICP) or chemical ionisation. Preferably, the counterion $X \star$ may be dissociated by irradiation.

When $X \star$ is a cation, $X \star$ is preferably H^+ or Li^+ , especially Li^+ .

- 20 When $X \star$ is an anion, $X \star$ is preferably, BF_4^- or ClO_4^- , especially BF_4^- (e.g. compounds (IIb-28b), (IIb-28c) & (IIb-28d)).

It is preferred that $X \star$ is an anion.

Preferred examples of group X★ are shown in figure 4.



The C - -S_S, S_S - -Ar¹ or S_S - -Ar² bonds are cleavable by irradiation, electron bombardment, electrospray, fast atom bombardment (FAB), inductively coupled plasma (ICP) or chemical
5 ionisation. Preferably, the C - -S_S, S_S - -Ar¹ or S_S - -Ar² bonds are cleavable by irradiation or chemical ionisation.

Where appropriate, the C - -S_S, S_S - -Ar¹ or S_S - -Ar² bonds may be cleaved simultaneously or sequentially with the cleaving of the C-X bond or the dissociation of X★, as appropriate, by selection of suitable cleaving/dissociating conditions.

10 In one embodiment of the invention, the C - -S_S bond in the solid support of formula (Vai) may be cleaved in sub-steps of step (iia) so that in a first sub-step a residue X (where X is the leaving group defined above) is provided and in a second subsequent sub-step the C-X bond is cleaved thereby forming the ion of formula (I). If desired, the second sub-step may be carried out substantially (*e.g.* seconds, minutes, hours or even days) after the first sub-step.

15 *Ar¹ and Ar²*



Ar² is independently an aromatic group or an aromatic group substituted with one or more A and is preferably independently cyclopropyl, cyclopropyl substituted with one or more A, aryl, aryl substituted with one or more A, heteroaryl, or heteroaryl substituted with one or more A.

20 Where aryl or substituted aryl, Ar² is preferably C₆₋₃₀ aryl or substituted C₆₋₃₀ aryl. Where heteroaryl or substituted heteroaryl, Ar² is preferably C₆₋₃₀ heteroaryl or substituted C₆₋₃₀ heteroaryl.

Examples of aryl and heteroaryl are monocyclic aromatic groups (*e.g.* phenyl or pyridyl), fused polycyclic aromatic groups (*e.g.* naphthyl, such as 1-naphthyl or 2-naphthyl) and unfused polycyclic aromatic groups (*e.g.* monocyclic or fused polycyclic aromatic groups linked by a single bond, a
25 double bond, or by a -(CH=CH)_r- linking group, where r is one or more (*e.g.* 1, 2, 3, 4 or 5).

Other examples of aryl groups are monovalent derivatives of aceanthrylene, acenaphthylene, acephenanthrylene, anthracene, azulene, chrysene, coronene, fluoranthene, fluorene, *as*-indacene, *s*-indacene, indene, naphthalene, ovalene, perylene, phenalene, phenanthrene, picene, pleiadene, pyrene, pyranthrene and rubicene, which groups may be optionally substituted by one or more A.

30 Other examples of heteroaryl groups are monovalent derivatives of acridine, carbazole, β -carboline, chromene, cinnoline, furan, imidazole, indazole, indole, indolizine, isobenzofuran, isochromene, isoindole, isoquinoline, isothiazole, isoxazole, naphthyridine, perimidine, phenanthridine, phenanthroline, phenazine, phthalazine, purine, pyran, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole, pyrrolizine, quinazoline, quinoline, quinolizine, quinoxaline, thiophene and
35 xanthene, which groups may be optionally substituted by one or more A. Preferred heteroaryl groups

are five- and six-membered monovalent derivatives, such as the monovalent derivatives of furan, imidazole, isothiazole, isoxazole, pyran, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole, pyrrolizine and thiophene. The five-membered monovalent derivatives are particularly preferred, *i.e.* the monovalent derivatives of furan, imidazole, isothiazole, isoxazole, pyrazole, pyrrole and thiophene. The heteroaryl groups may be attached to the remainder of the compound by any carbon or hetero (e.g. nitrogen) atom.

- Ar² is preferably C₆₋₃₀aryl substituted by one or more A, preferably phenyl or naphthyl (*e.g.* 1-naphthyl or 2-naphthyl, especially 2-naphthyl) substituted by one or more A, more preferably phenyl substituted by one or more A. When Ar² is phenyl, A is preferably provided in a position ortho or para to C★.
- 10 When Ar² is other than phenyl, A is preferably attached to an atom which bears the charge in at least one of the resonance structures of the ions of formula (I).

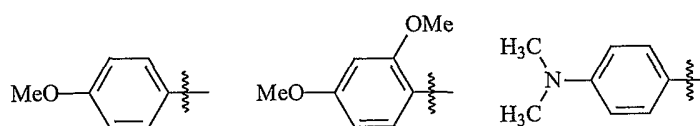
Fused polycyclic aromatic groups, optionally substituted with one or more A, are particularly preferred.

- A particularly preferred Ar² is unsubstituted pyrenyl or pyrenyl substituted with one or more A.
- 15 Unsubstituted pyrenyl is preferred. The pyrenyl group may be 1-pyrenyl (*e.g.* compounds (IIa-38a), (IIa-38b), (IIa-39a), (IIa-41a) & (IIa-41b)), 2-pyrenyl (*e.g.* compounds (IIa-42a) & (IIa-42b)) or 4-pyrenyl (*e.g.* compounds (IIa-43a) & (IIa-43b)).

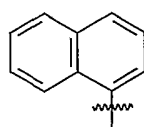
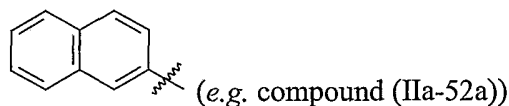
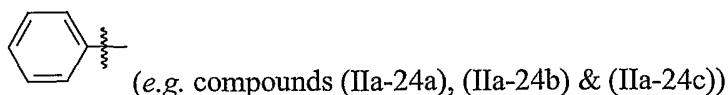
Preferred heteroaryl Ar² groups, whether substituted or unsubstituted, are pyridyl, pyrrolyl, thienyl and furyl, especially thienyl.

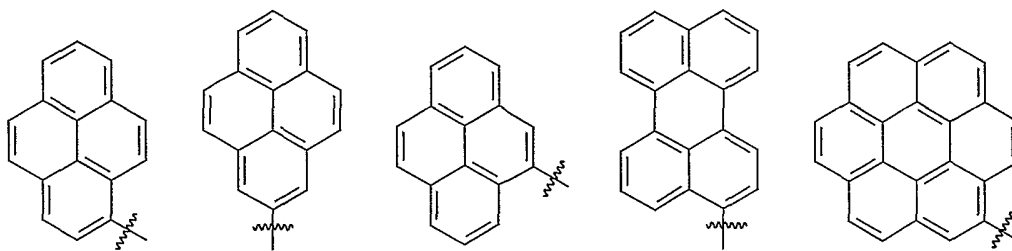
- 20 A preferred Ar² group is thiophenyl or thiophenyl substituted with one or more A. Unsubstituted thiophenyl is preferred. Examples of thiophenyl are thiophen-2-yl and thiophen-3-yl, with thiophen-2-yl being especially preferred (*e.g.* compounds 50a, 51a & 51b).

When substituted, Ar² is preferably substituted by 1, 2 or 3 A. Ar² is preferably:



- 25 When unsubstituted, Ar² is preferably:





In another preferred embodiment, Ar^2 is cyclopropyl or cyclopropyl substituted with one or more A. Unsubstituted cyclopropyl is preferred (*e.g.* compound (IIa-44a)). One or more, preferably one, of Ar^2 may be cyclopropyl.

- 5 Preferred examples of group Ar^2 are shown in figures 3A and 3B.

Ar^1

- Ar^1 is independently an aromatic group or an aromatic group substituted with one or more A. The definition of Ar^1 is the same as Ar^2 (as defined above), except that the valency of the group Ar^1 is adapted to accommodate the q instances of the linker L_M . Preferred Ar^2 groups are also preferred Ar^1 groups, (as defined above), except that the valency of the group Ar^1 is adapted to accommodate the q instances of the linker L_M .
- 10

When $q = 1$, Ar^1 is a divalent radical and is preferably independently cyclopropylene, cyclopropylene substituted with one or more A, arylene, arylene substituted with one or more A, heteroarylene, or heteroarylene substituted with one or more A.

- 15 Where arylene or substituted arylene, Ar^1 is preferably C_{6-30} arylene or substituted C_{6-30} arylene. Where heteroarylene or substituted heteroarylene, Ar^1 is preferably C_{6-30} heteroarylene or substituted C_{6-30} heteroarylene.

- Examples of arylene and heteroarylene are monocyclic aromatic groups (*e.g.* phenylene or pyridylene), fused polycyclic aromatic groups (*e.g.* naphthylene) and unfused polycyclic aromatic groups (*e.g.* monocyclic or fused polycyclic aromatic groups linked by a single bond, a double bond, or by a $-(\text{CH}=\text{CH})_r-$ linking group, where r is one or more (*e.g.* 1, 2, 3, 4 or 5)).
- 20

- Other examples of arylene groups are polyvalent derivatives (where the valency is adapted to accommodate the q instances of the linker L_M) of aceanthrylene, acenaphthylene, acephenanthrylene, anthracene, azulene, chrysene, coronene, fluoranthene, fluorene, *as*-indacene, *s*-indacene, indene, naphthalene, ovalene, perylene, phenalene, phenanthrene, picene, pleiadene, pyrene, pyranthrene and rubicene, which groups may be optionally substituted by one or more A.
- 25

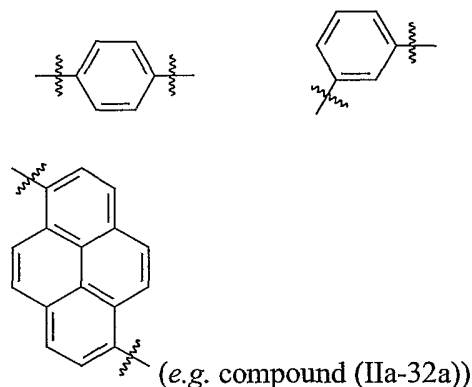
- Other examples of heteroarylene groups are polyvalent derivatives (where the valency is adapted to accommodate the q instances of the linker L_M) of acridine, carbazole, β -carboline, chromene, cinnoline, furan, imidazole, indazole, indole, indolizine, isobenzofuran, isochromene, isoindole, isoquinoline, isothiazole, isoxazole, naphthyridine, perimidine, phenanthridine, phenanthroline, phenazine, phthalazine, purine, pyran, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole,
- 30

pyrrolizine, quinazoline, quinoline, quinolizine, quinoxaline, thiophene and xanthene, which groups may be optionally substituted by one or more A. Preferred heteroaryl groups are five- and six-membered polyvalent derivatives, such as the polyvalent derivatives of furan, imidazole, isothiazole, isoxazole, pyran, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole, pyrrolizine and thiophene. The five-membered polyvalent derivatives are particularly preferred, *i.e.* the polyvalent derivatives of furan, imidazole, isothiazole, isoxazole, pyrazole, pyrrole and thiophene. The heteroaryl groups may be attached to the remainder of the compound by any carbon or hetero (e.g. nitrogen) atom.

Ar¹ is preferably C₆₋₃₀arylene substituted by one or more A, preferably phenylene or naphthylene substituted by one or more A, more preferably phenylene substituted by one or more A. When Ar¹ is phenylene, A is preferably provided in a position ortho or para to C★. When Ar¹ is other than phenylene, A is preferably attached to an atom which bears the charge in at least one of the resonance structures of the ions of formula (I).

When substituted, Ar¹ is preferably substituted by 1, 2 or 3 A.

15 When unsubstituted, preferred Ar¹ are:



Preferred examples of group Ar¹ are shown in figures 3A and 3B.

Combinations of Ar

20 Optionally two or three of the groups Ar¹ and Ar² are linked together by one or more L⁵, where L⁵ is independently a single bond or a linker atom or group; and/or two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group substituted with one or more A.

When L⁵ is a linker group, preferred linker groups are -E⁵-, -(D⁵)_t-, -(E⁵-D⁵)_t-, -(D⁵-E⁵)_t-, -E⁵-(D⁵-E⁵)_t- or -D⁵-(E⁵-D⁵)_t-.

25 D⁵ is independently C₁₋₈hydrocarbylene or C₁₋₈hydrocarbylene substituted with one or more A.

E⁵ is independently -Z⁵-, -C(=Z⁵)-, -Z⁵C(=Z⁵)-, -C(=Z⁵)Z⁵-, -Z⁵C(=Z⁵)Z⁵-, -S(=O)-, -Z⁵S(=O)-, -S(=O)Z⁵-, -Z⁵S(=O)Z⁵-, -S(=O)₂-, -Z⁵S(=O)₂-, -S(=O)₂Z⁵-, -Z⁵S(=O)₂Z⁵-, where Z⁵ is independently O, S or N(R⁵) and where R⁵ is independently H, C₁₋₈hydrocarbyl or C₁₋₈hydrocarbyl substituted with one or more A. Preferably E⁵ is -O-, -S-, -C(=O)-, -C(=O)O-, -C(=S)-, -C(=S)O-, -OC(=S)-,

-C(=O)S-, -SC(=O)-, -S(O)-, -S(O)₂-, -N(R⁵)-, -C(=O)N(R⁵)-, -C(=S)N(R⁵)-, -N(R⁵)C(=O)-, -N(R⁵)C(=S)-, -S(=O)N(R⁵)-, -N(R⁵)S(=O)-, -S(=O)₂N(R⁵)-, -N(R⁵)S(=O)₂-, -OC(=O)O-, -SC(=O)O-, -OC(=O)S-, -N(R⁵)C(=O)O-, -OC(=O)N(R⁵)-, -N(R⁵)C(=O)N(R⁵)-, -N(R⁵)C(=S)N(R⁵)-, -N(R⁵)S(=O)N(R⁵)- or -N(R⁵)S(=O)₂N(R⁵)-.

- 5 t' = 1 or more, *e.g.* from 1 to 50, 1 to 40, 1 to 30, 1 to 20 or 1 to 10. Preferably t' = 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10. Most preferably t'=1.

Where L⁵ includes an atom or group which also falls within the definition of group M, the group M is preferably more reactive than the group included in L⁵.

L⁵ is preferably a linker atom, preferably O or S, particularly O.

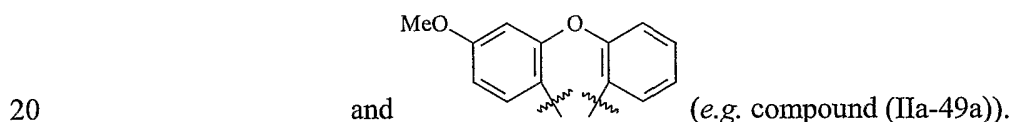
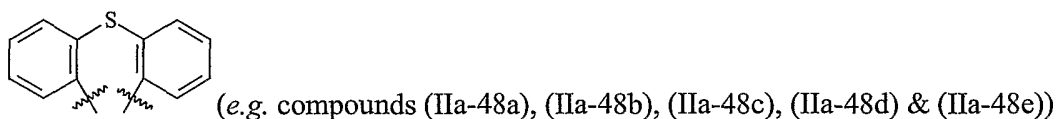
- 10 When L⁵ is a linker group, a preferred L⁵ is -N(R⁵)-.

In another embodiment in which L⁵ is a linker group, L⁵ is -S(=O)- (*e.g.* compound (IIa-56b))

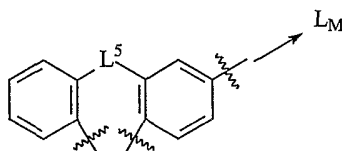
When two of the groups Ar¹ and Ar² are linked together by one or more (*e.g.* 2, 3 or 4) L⁵, they are preferably linked together by one L⁵, preferably O.

- Preferred combinations of Ar are two Ar² (*e.g.* two Ar² phenyl groups) linked together by one L⁵
15 (*e.g.* O or S).

Particularly preferred combinations of Ar are two Ar² phenyl groups, optionally substituted by one or more A (preferably unsubstituted), linked together by one L⁵ (*e.g.* O or S), where L⁵ is ortho to C★ with respect to both phenyl groups. Especially preferred combinations of two Ar² groups are:

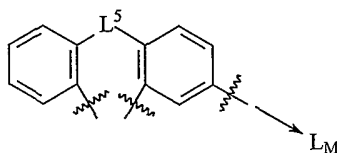


In another embodiment, a preferred combination of one Ar¹ and one Ar² is:



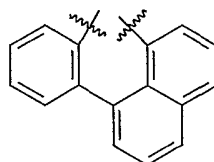
- optionally substituted by A. Preferably, L⁵ is O (*e.g.* compound (IIa-68)) or S (*e.g.* compounds (IIa-58a) and (IIa-69)). Compounds of this embodiment show improved mass spectrometry
25 enhancing properties. Preferred optional substituents A are -OMe (*e.g.* compounds IIa-68 and IIa-69), preferably para to C★.

In another embodiment, a preferred combination of one Ar¹ and one Ar² is:



optionally substituted by A. Preferably, L^5 is O or S (*e.g.* compound (IIa-67)), preferably S. Compounds of this embodiment also show improved mass spectrometry enhancing properties. Preferred optional substituents A are $-OMe$ (*e.g.* compound IIa-67), preferably para to $C^\star$.

- 5 In another embodiment, a preferred combination of Ar are two Ar^1 or Ar^2 groups (*i.e.* Ar^1+Ar^1 , Ar^1+Ar^2 , or Ar^2+Ar^2), linked by one L^5 , wherein one Ar^1 or Ar^2 group is a polycyclic aromatic group (*e.g.* naphthyl or pyrenyl), preferably a pyrenyl group. Such combinations of Ar groups are fluorescent and allow labelling, *e.g.* of the biopolymer. An example of such a combination of Ar groups is:

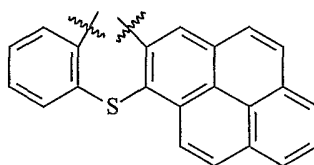


10

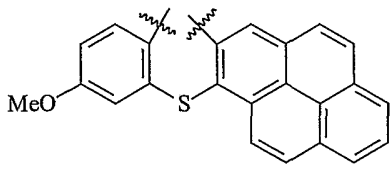
optionally substituted by A, *e.g.* $-OMe$, wherein when one or more of the Ar groups is Ar^1 , the combination includes an appropriate number of $L_M\{M\}_p$ groups.

It is particularly preferred in this embodiment that L^5 is S. The S atom may be oxidised to $S=O$ without loss of the X group, advantageously allowing modification of the properties (*e.g.* fluorescent properties) of the combined Ar group. A particularly preferred combination of Ar groups in this

15



optionally substituted by A, *e.g.* $-OMe$, *e.g.*

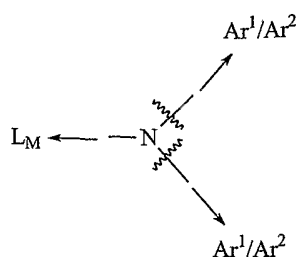


- 20 wherein when one or more of the Ar groups is Ar^1 , the combination includes an appropriate number of $L_M\{M\}_p$ groups.

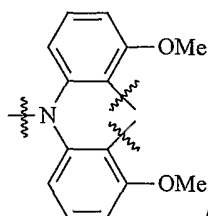
In another embodiment, at least one L_M is linked to an atom or group L^5 . In this embodiment, the preferred L^5 mentioned above are, where appropriate, modified to remove substituents R^5 in order to accommodate L_M , *e.g.* the R^5 substituent of the group $-N(R^5)-$ is replaced by L_M . In this embodiment,

25

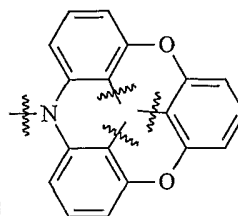
the L^5 group to which L_M is bound is preferably:



Preferred combinations of Ar¹ and/or Ar² in this embodiment are:



(*e.g.* compound (IIb-28c)) and



(*e.g.* compounds (IIb-28d) & (IIb-47b))

- 5 When two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group substituted with one or more A, the aromatic group may be a carbocyclic aromatic group or a carbocyclic aromatic group in which one or more carbon atoms are each replaced by a hetero atom. Typically, in an aromatic group in which one or more carbon atoms are each replaced by a hetero atom, up to three carbons are so replaced, preferably up to two carbon atoms, more preferably one
- 10 carbon atom.

Preferred hetero atoms are O, Se, S or N, more preferably O, S or N.

When two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group substituted with one or more A, preferred aromatic groups are C₈₋₅₀ aromatic groups.

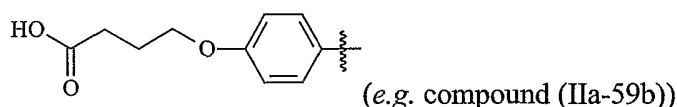
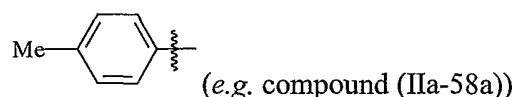
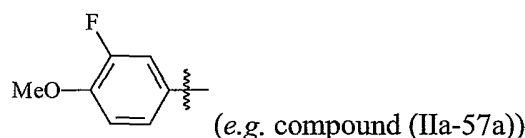
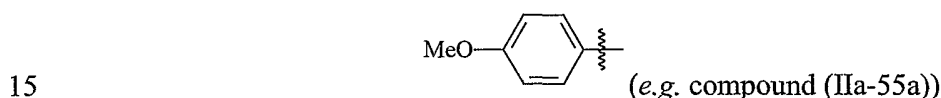
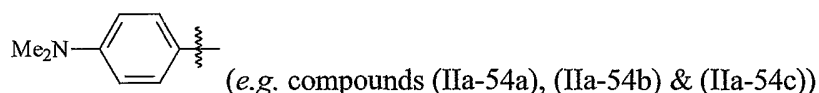
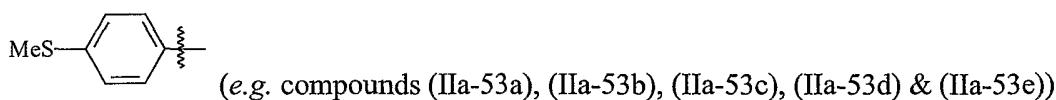
- The aromatic groups may be monocyclic aromatic groups (*e.g.* radicals of suitable valency derived from benzene), fused polycyclic aromatic groups (*e.g.* radicals of suitable valency derived from naphthalene) and unfused polycyclic aromatic groups (*e.g.* monocyclic or fused polycyclic aromatic groups linked by a single bond, a double bond, or by a $-(CH=CH)_r-$ linking group, where *r* is one or more (*e.g.* 1, 2, 3, 4 or 5).
- 15

- When two or three of the groups Ar¹ and Ar² together form a carbopolycyclic fused ring aromatic group, preferred groups are radicals of suitable valency obtained from naphthalene, anthracene or phenanthracene, chrysene, aceanthrylene, acenaphthylene, acephenanthrylene, azulene, fluoranthene, fluorene, *as*-indacene, *s*-indacene, indene, phenalene, and pleiadene.
- 20

- When two or three of the groups Ar¹ and Ar² together form a carbopolycyclic fused ring aromatic group in which one or more carbon atoms are each replaced by a hetero atom, preferred groups are radicals of suitable polyvalency obtained from acridine, carbazole, β -carboline, chromene, cinnoline, indole, indolizine, isobenzofuran, isochromene, isoindole, isoquinoline, naphthyridine, perimidine, phenanthridine, phenanthroline, phenazine, phthalazine, pteridine, purine, pyrrolizine, quinazoline, quinoline, quinolizine and quinoxaline.
- 25

Substitution of Ar¹ and Ar² – Anions and Cations

When C★ is a cation, A is preferably an electron-donating group, including -R¹ or -Z¹R¹, where R¹ and Z¹ are defined below. Preferably, R¹ is C₁₋₈hydrocarbyl, more preferably C₁₋₈alkyl, especially methyl. Z¹ is preferably O, S or NR¹. R¹ may be substituted with one or more S_{ub}², but is preferably unsubstituted. When C★ is a cation, A is preferably -OMe (*e.g.* compound (IIa-55a)), -SMe (*e.g.* compounds (IIa-53a), (IIa-53b), (IIa-53c), (IIa-53d) & (IIa-53e)), -N(Me)₂ (*e.g.* compounds (IIa-54a), (IIa-54b) & (IIa-54c)) or Me (*e.g.* compound (IIa-58a)). When C★ is a cation, A, when an electron-donating group, is preferably provided (especially in relation to Ar¹ or Ar² being phenyl) in a position ortho or para to C★, preferably para. Furthermore, when C★ is a cation, A, when an electron-withdrawing group (*e.g.* F (*e.g.* compound (IIa-57a))), is preferably provided (especially in relation to Ar¹ or Ar² being phenyl) in a position meta to C★. Thus, preferred groups Ar¹ and Ar² are as follows:



When C★ is an anion, A is preferably an electron-withdrawing group, including halogen, trihalomethyl, -NO₂, -CN, -N⁺(R¹)₂O⁻, -CO₂H, -CO₂R¹, -SO₃H, -SOR¹, -SO₂R¹, -SO₃R¹, -OC(=O)OR¹, -C(=O)H, -C(=O)R¹, -OC(=O)R¹, -C(=O)NH₂, -C(=O)NR¹₂, -N(R¹)C(=O)OR¹, -N(R¹)C(=O)NR¹₂, -OC(=O)NR¹₂, -N(R¹)C(=O)R¹, -C(=S)NR¹₂, -NR¹C(=S)R¹, -SO₂NR¹₂, -NR¹SO₂R¹, -N(R¹)C(=S)NR¹₂, or -N(R¹)SO₂NR¹₂, where R¹ is defined below. When C★ is an anion, A, when an electron-withdrawing group, is preferably provided (especially in relation to Ar¹ or Ar² being phenyl) in a position ortho or para to C★, preferably para. Furthermore, when C★ is an anion, A, when an electron-donating group, is preferably provided (especially in relation to Ar¹ or Ar² being phenyl) in a position meta to C★.

The group A may also comprise one or more isotopes of the atoms making up group A (*e.g.* example 60), thus, as discussed in more detail below, allowing the masses of the compounds of the invention to be varied. Preferred isotopes are ^{13}C , ^{18}O and ^2H . When providing a series of compounds which differ only in their masses, ^{13}C and ^{18}O are particularly preferred as ^2H atoms may cause a substantial change in the chemical properties of the compound due to the kinetic isotope effect.

Solid Supports

‘Solid supports’ for use with the invention include polymer beads, metals, resins, columns, surfaces (including porous surfaces) and plates (*e.g.* mass-spectrometry plates).

The solid support is preferably one suitable for use in a mass spectrometer, such that the invention can be conveniently accommodated into existing MS apparatus. Ionisation plates from mass spectrometers are thus preferred solid supports, *e.g.* gold, glass-coated or plastic-coated plates. Solid gold supports are particularly preferred.

Resins or columns, such as those used in affinity chromatography and the like, are particularly useful for receiving solutions of biopolymers (purified or mixtures). For example, a cellular lysate could be passed through such a column of formula (IVai), (IVaii), (IVaiii), (IVaiv), (IVbii), (IVbiii) or (IVbiv) followed by cleavage of the support to leave compounds of formula (I).

Solid supports of formulae (IVai), (IVaii), (IVaiii), (IVaiv), (IVbii), (IVbiii) or (IVbiv) will generally present exposed groups M capable of reacting with a biopolymer, B_p. For MS analysis, ions preferably have a predictable mass to charge (*m/e*) ratio. If a biopolymer reacts with more than one M group, however, then it will carry more than one positive charge once ionised, and its *m/e* ratio will decrease. Advantageously, therefore, the groups M are arranged such that any biopolymer molecule will covalently link with only a single group M. Consequently, each biopolymer will, on ionisation, carry a single positive charge and thus have a predictable mass to charge ratio.

Typically, the surface density of the solid supports of (IVai), (IVaii), (IVaiii), (IVaiv), (IVbii), (IVbiii) or (IVbiv) will be provided so that a biopolymer molecule can only covalently link with one group M and thus to prevent the formation of multiply derivatised biopolymers.

Varying the mass of compounds of the invention

Within the general formulae (I), (IIa), (IIb), (IIIa), (IIIb), (IVai), (IVaii), (IVaiii), (IVaiv), (IVbii), (IVbiii), (IVbiv), (Vai), (Vaii), (Vaiii), (Vaiv), (Vbii), (Vbiii) and (Vbiv), there is much scope for variation. There is thus much scope of variation in the mass of these compounds. In some embodiments of the invention, it is preferred to use a series of two or more (*e.g.* 2, 3, 4, 5, 6 or more) compounds with different and defined molecular masses.

The masses of the compounds of the invention can be varied via L_M, Ar¹ and/or Ar². Preferably, the masses of the compounds of the invention are varied by varying A on the groups Ar¹ and/or Ar².

In this aspect of invention, compounds of the invention advantageously comprise one or more of F or I as substituents A of the groups Ar¹, Ar² or Ar³. F and I each only have one naturally occurring isotope, ¹⁹F and ¹²⁷I respectively, and thus by varying the number of F and I atoms present in the structure of the compounds, can provide a series of molecular mass labels having substantially identical shaped peaks on a mass spectrum.

Compounds of the invention may also include one or more ²H atoms, preferably as a substituent A or a part thereof of the groups L_M, Ar¹, Ar² or Ar³ (in particular L_M), in order to vary the masses of the compounds of the invention. The compounds of the invention may include isotopes of ¹³C and ¹⁸O, preferably as a substituent A or a part thereof of the groups L_M, Ar¹, Ar² or Ar³ (in particular Ar¹, Ar² or Ar³), in order to vary the masses of the compounds of the invention. Compounds comprising ²H, ¹³C and ¹⁸O may also be used to provide a series of molecular mass labels having substantially identical shaped peaks on a mass spectrum, by varying the number of ²H, ¹³C and ¹⁸O atoms present in the structure of the compounds. When providing a series of compounds which differ only in their masses, ¹³C and ¹⁸O are particularly preferred as ²H atoms may cause a substantial change in the chemical properties of the compound due to the kinetic isotope effect.

In order to increase the molecular mass of the compounds of the invention and to increase the number of available sites for substitution by A, especially F and I, one or more of Ar¹ and Ar² may be substituted by one or more dendrimer radicals of appropriate valency, either as substituent A or group L_M.

Preferred dendrimer radicals are the radicals obtained from the dendrimers of US 6,455,071 and PAMAM dendrimers.

The compounds of the invention may advantageously be used in the method of analysing a biopolymer disclosed herein, in particular in a method for following a reaction involving a biopolymer, B_P, since the abundance of a species of may be determined by mass spectrometry by measuring the intensity of the relevant peak in an obtained mass spectrum.

Specifically, there is provided a method for analysing biopolymer B_P, comprising the steps of:

(i) reacting a first sample comprising biopolymer B_P with a compound of formula (IIa) or (IIb), wherein the compound of formula (IIa) or (IIb) is selected from the compounds of formulae (IIa-1a) to (IIa-69) or the compounds of formulae (IIb-28c), (IIb-28d) and (IIb-47b) described above, at a time t₁;

(ii) reacting a second sample comprising biopolymer B_P with a compound of formula (IIa) or (IIb), wherein the compound of formula (IIa) or (IIb) is selected from the compounds of formulae (IIa-1a) to (IIa-69) or the compounds of formulae (IIb-28c), (IIb-28d) and (IIb-47b) described above, at a later time t₂;

(iii) preparing and analysing cations of formula (I) from the first and second samples; and

- (iv) comparing the results of the analysis from step (iii).

If levels of the biopolymer B_P decrease between times t_1 and t_2 then there will be a decrease in detected ion; if levels of the biopolymer B_P increase between times t_1 and t_2 then there will be an increase in detected ion. The effects of stimuli on transcription and/or translation can therefore be monitored.

Advantageously, different compounds of formula (IIa) or (IIb) are used at different times in order to facilitate simultaneous and parallel analysis of the first and second samples. For example, if the two compounds used at times t_1 and t_2 differ only by a ^1H to ^{19}F substitution then the relative abundance of B_P at the two times can be determined by comparing peaks separated by 18 units.

- Advantageously, the reaction of the biopolymer with the compound of formula (IIa) or (IIb) will fix the biopolymer to prevent it reacting further and the steps of providing and analysing the cations may be carried out at a later convenient time. Alternatively, if the reaction of the biopolymer with the compound of formula (IIa) or (IIb) does not quench the reaction of the biopolymer being followed, a cation of formula (I) from the reaction product of step (i) or step (v) should be obtained as soon as possible after reaction of the biopolymer with the compound of formula (IIa) or (IIb).

Compounds of Formulae (IIa) and (IIb)

Compounds of Formulae (IIa-1) to (IIa-69), (IIb-28c), (IIb-28d) and (IIb-47b)

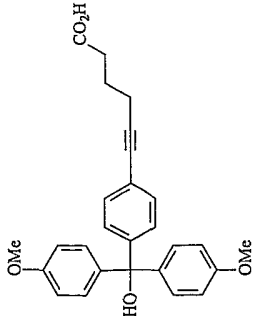
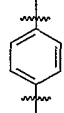
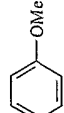
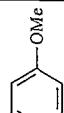
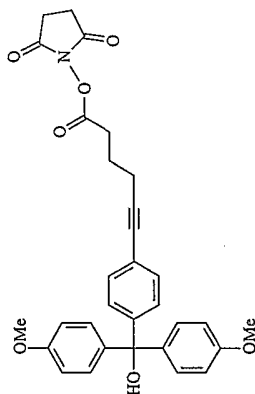
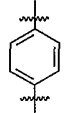
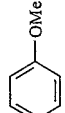
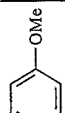
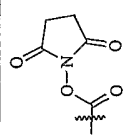
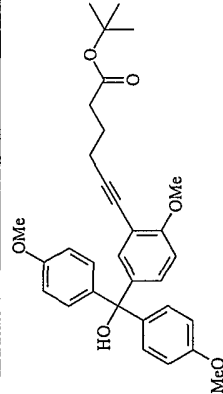
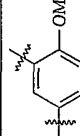
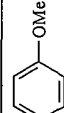
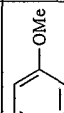
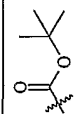
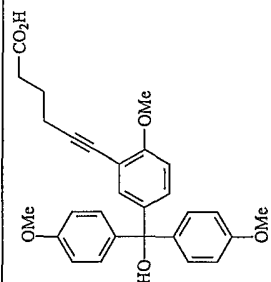
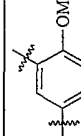
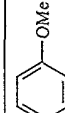
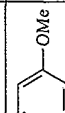
The present invention is particularly directed to compounds of formula (IIa) of the formulae (IIa-1) to (IIa-69) set out in table 3 and to compounds of formula (IIb) of the formulae (IIb-28c), (IIb-28d) and (IIb-47b) set out in table 4.

Table 3

Compounds of formulae (IIa-1) to (IIa-69)

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
					(left hand side attached to central carbon)				(left hand side attached to Ar ¹)		
IIa-1a		-OH	1	2				1	-CH ₂ CH ₂ -	1	
IIa-2a		-OH	1	2				1	-CH ₂ CH ₂ -	1	
IIa-3a		-OH	1	2				1	-CH ₂ CH ₂ -C(O)O-CH ₂ C(Me)-(CH ₂) ₂	2	M(#1) = OH M(#2) = OH

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-4a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-5a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-6a		-OH	1	2				1	single bond	1	-I
IIa-6b		-OH	1	2				1	-C≡CCH ₂ CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-6c		-OH	1	2				1	-C≡CCH ₂ CH ₂ CH ₂ -	1	-CO ₂ H
IIa-6d		-OH	1	2				1	-C≡CCH ₂ CH ₂ CH ₂ -	1	
IIa-7a		-OH	1	2				1	-C≡CCH ₂ CH ₂ CH ₂ -	1	
IIa-7b		-OH	1	2				1	-C≡CCH ₂ CH ₂ CH ₂ -	1	-CO ₂ H

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-7c		-OH	1	2				1	-C≡CCH ₂ CH ₂ CH ₂ -	1	
IIa-8a		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-8b		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	-CO ₂ H
IIa-8c		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-9a		-OH	1	2				1	-CD ₂ CD ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-9b		-OH	1	2				1	-CD ₂ CD ₂ CH ₂ CH ₂ CH ₂ -	1	-CO ₂ H
IIa-9c		-OH	1	2				1	-CD ₂ CD ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-10a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-10b		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	-CO ₂ H
IIa-10c		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-11a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-11b		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	-CO ₂ H

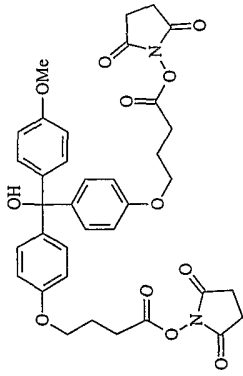
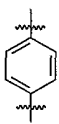
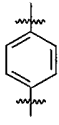
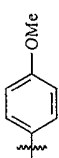
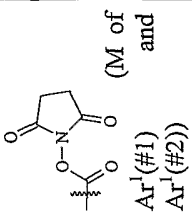
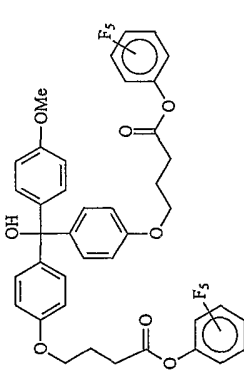
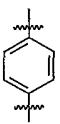
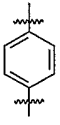
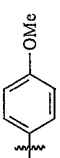
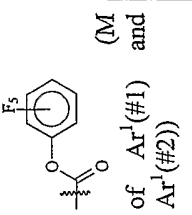
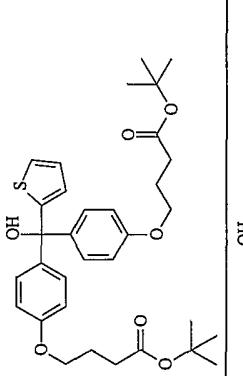
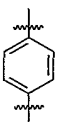
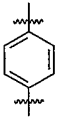
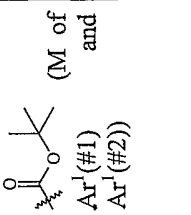
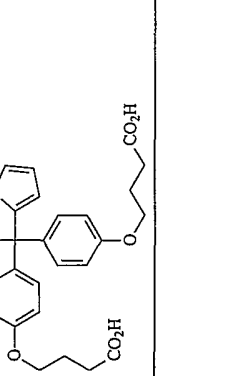
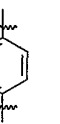
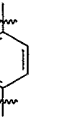
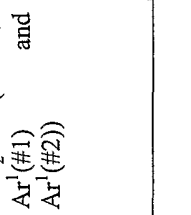
Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-11c		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-12a		-OH	1	2				1	single bond	1	-I
IIa-12b		-OH	1	2				1	-C≡CCCH ₂ -	1	-NHBoc
IIa-12c		-OH	1	2				1	-C≡CCCH ₂ -	1	-NH ₂

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-13a		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ -	1	-NHBoc
IIa-13b		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ -	1	-NH ₂
IIa-14a		-OEt	1	2				1	-CH ₂ CH ₂ -	1	
IIa-14b		-OEt	1	2				1	-CH ₂ CH ₂ CO- N((CH ₂) ₂) ₂ CH-	1	-OH

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
Ila-14c		-OEt	1	2				1	-CH ₂ CH ₂ CO- N((CH ₂) ₂) ₂ CH-	1	

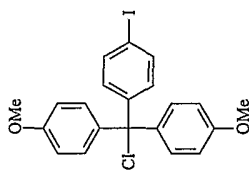
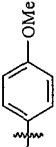
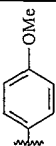
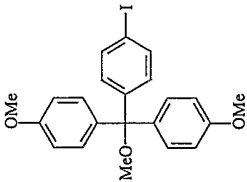
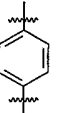
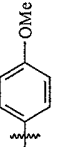
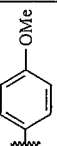
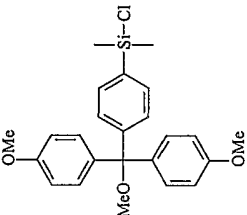
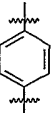
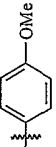
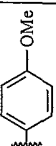
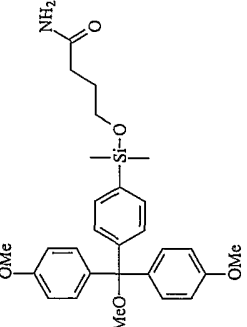
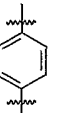
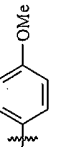
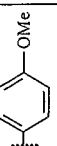
Boc = t-Butyloxycarbonyl

Cpd. no.	Structure	X	m	n	Ar ¹ (#1)	Ar ¹ (#2)	Ar ² (1)	q	L _M	p	M
					(left hand side attached to central carbon)				(left hand side attached to Ar ¹)		
IIa-15a		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	(M of Ar ¹ (#1) and Ar ¹ (#2))
IIa-15b		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	(M of Ar ¹ (#1) and Ar ¹ (#2))

Cpd. no.	Structure	X	m	n	Ar ¹ (#1)	Ar ¹ (#2)	Ar ² (1)	q	L _M	p	M
IIa-15c		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	 (M of Ar ¹ (#1) and Ar ¹ (#2))
IIa-16a		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	 (M of Ar ¹ (#1) and Ar ¹ (#2))
IIa-50a		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	 (M of Ar ¹ (#1) and Ar ¹ (#2))
IIa-51a		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	 (M of Ar ¹ (#1) and Ar ¹ (#2))

Cpd. no.	Structure	X	m	n	Ar ¹ (#1)	Ar ¹ (#2)	Ar ² (1)	q	L _M	p	M
Ila-51b		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	
Ila-59a		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	
Ila-59b		-OH	2	1				1	-OCH ₂ CH ₂ CH ₂ - (L _M of Ar ¹ (#1) and Ar ¹ (#2))	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
					(left hand side attached to central carbon)				(left hand side attached to Ar ¹)		
IIa-17a		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ -	1	-NH-Boc-Me
IIa-17b		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ -	1	-NH-Me
IIa-17c		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ N(Me)-C(=O)-CH ₂ -	1	-I
IIa-18d		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ N(Me)-C(=O)-CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-19b		-Cl	1	2				1	single bond	1	-I
IIa-19c		-OMe	1	2				1	single bond	1	-I
IIa-19d		-OMe	1	2				1	single bond	1	-Si(Me) ₂ Cl
IIa-19e		-OMe	1	2				1	-Si(Me) ₂ -O-(CH ₂) ₃ -	1	-C(O)NH ₂

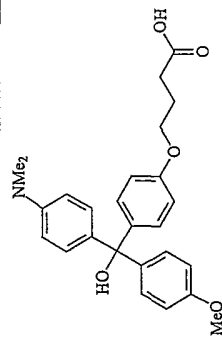
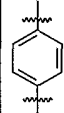
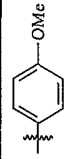
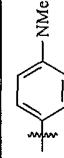
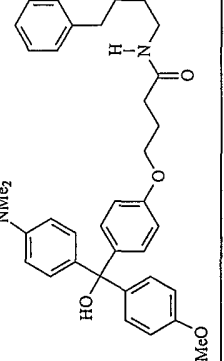
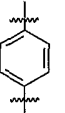
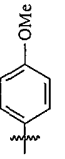
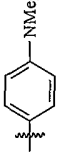
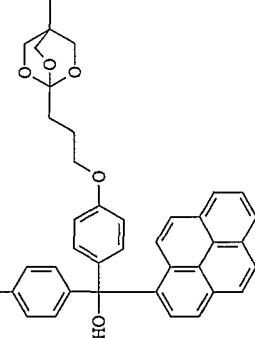
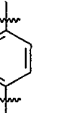
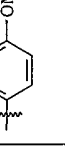
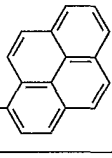
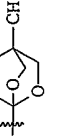
Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-20a		-OH	1	2				1	-CH ₂ CH ₂ -	1	
IIa-20c		-OEt	1	2				1	-CH ₂ CH ₂ CH ₂ -	1	-OH
IIa-24a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	
IIa-24b		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	-CO ₂ H

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-24c		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	
IIa-30Aa		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-30Ab		-OMe	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	-OH
IIa-30Ac		-OMe	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-32a		OH	1	2				1	single bond	1	
IIa-33a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	
IIa-34a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-35Ab		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -CO-	1	-NH-NH ₂
IIa-35Bb		-OMe	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -OC(O)-	1	
IIa-35Bc		-OMe	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -OC(O)-	1	-NH-NH ₂
IIa-35Bd		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -OC(O)-	1	-NH-NH ₂

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
Ila-35Cb		-OMe	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
Ila-35Cc		-OMe	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	-O-NH ₂
Ila-35Cd		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	-O-NH ₂
Ila-36a		-OH	1	2				1	-OCH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-37a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	-CO ₂ H
IIa-37b		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -C(O)NH-	1	-(CH ₂)Ph
IIa-38a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-38b		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	
IIa-41a		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-41b		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ -	1	

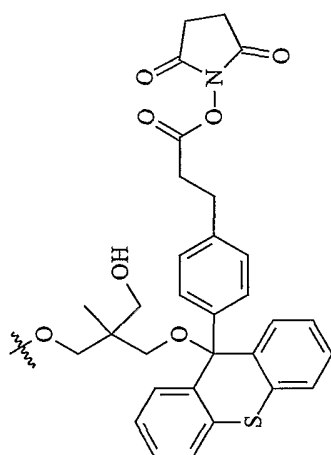
Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-42a		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-42b		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-43a		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	
IIa-43b		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	L ⁵	q	L _M	p	M
IIa-48a		-OH	1	2				S	1	-CH ₂ CH ₂ -	1	
IIa-48b		-OH	1	2				S	1	-CH ₂ CH ₂ -C(O)O-CH ₂ C(Me)-(CH ₂) ₂	2	M(#1) = OH M(#2) = OH
IIa-48c		-OH	1	2				S	1	-CH ₂ CH ₂ -	1	-CO ₂ H
IIa-48d		-OH	1	2				S	1	-CH ₂ CH ₂ -	1	
IIa-48e		-X(IIa-48e)	1	2				S	1	-CH ₂ CH ₂ -	1	

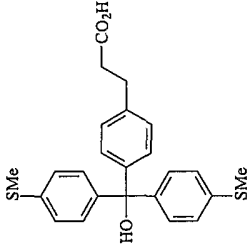
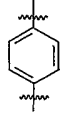
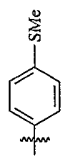
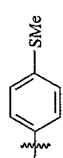
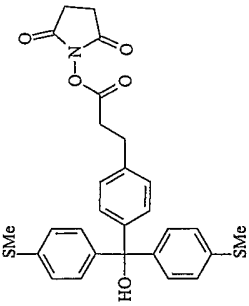
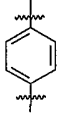
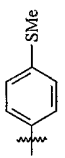
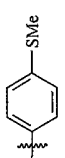
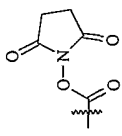
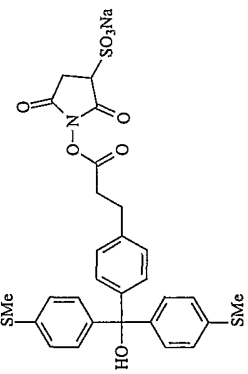
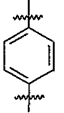
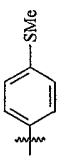
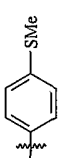
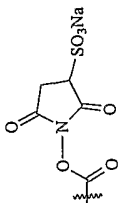
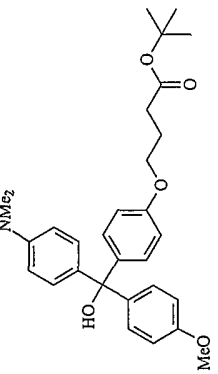
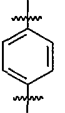
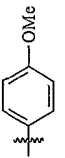
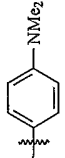
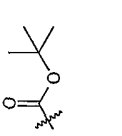
Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	L ^S	q	L _M	p	M
IIa-49a		-OH	1	2				O	1	-CH ₂ CH ₂ -	1	
IIa-56a		-OH	1	2				S	1	-CH ₂ CH ₂ -	1	
IIa-56b		-OH	1	2				S=O	1	-CH ₂ CH ₂ -	1	
IIa-58a		-OH	1	2				S	1	-O(CH ₂) ₅ -	1	

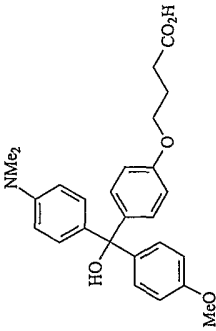
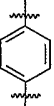
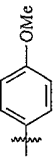
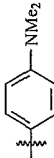
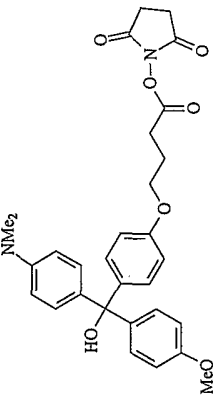
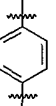
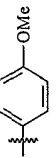
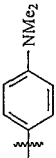
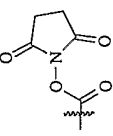
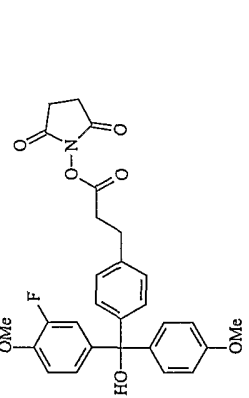
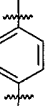
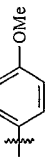
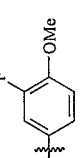
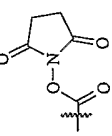
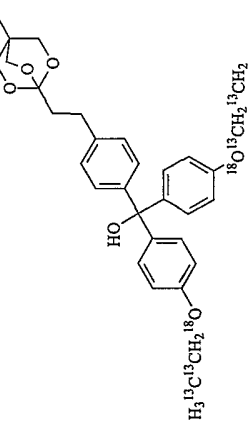
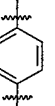
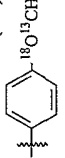
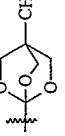
Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	L ^s	q	L _M	p	M
IIa-67		-OH	1	2				S	1	-O(CH ₂) ₅ -	1	
IIa-68		-OH	1	2				O	1	-(CH ₂) ₅ -	1	
IIa-69		-OH	1	2				S	1	-(CH ₂) ₂ -	1	

Group X of formula -X(IIa-48e) has the following structure:



Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
						(left hand side attached to central carbon)			(left hand side attached to Ar ¹)	.	
IIa-53a		-OH	1	2				1	-CH ₂ CH ₂ -	1	
IIa-53b		-OH	1	2				1	-CH ₂ CH ₂ -C(O)O-CH ₂ C(Me)-(CH ₂) ₂	2	M(#1) = OH M(#2) = OH

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-53c		-OH	1	2				1	-CH ₂ CH ₂ -	1	-CO ₂ H
IIa-53d		-OH	1	2				1	-CH ₂ CH ₂ -	1	
IIa-53e		-OH	1	2				1	-CH ₂ CH ₂ -	1	
IIa-54a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ -	1	

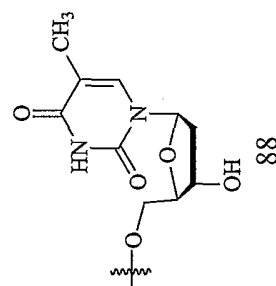
Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-54b		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	-CO ₂ H
IIa-54c		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -	1	
IIa-57a		-OH	1	2				1	-CH ₂ CH ₂ -	1	
IIa-60a		-OH	1	2		Both Ar ² (#1) and Ar ² (#2) 		1	-CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-60b		-OH	1	2		Both Ar ² (#1) and Ar ² (#2) are 4-(18O 13CH2 13CH2)phenyl groups.		1	-CH ₂ CH ₂ -	1	
IIa-61a		-OH	1	2				1	-OCH ₂ CH ₂ CH ₂ -CH ₂ CH ₂ -	1	
IIa-62a		-OH	1	2				1	-CH ₂ CH ₂ -	1	
IIa-63b		-Cl	1	2				1	-CH ₂ CH ₂ -	1	

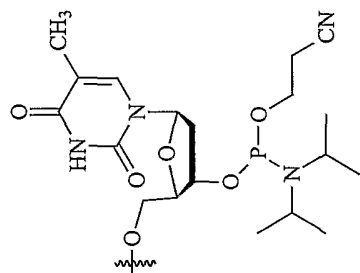
Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-63c		-X(IIa-63c)	1	2				1	-CH ₂ CH ₂ -	1	
IIa-63d		-X(IIa-63d)	1	2				1	-CH ₂ CH ₂ -	1	
IIa-63e		-X(IIa-63e)	1	2				1	-CH ₂ CH ₂ -	1	

Cpd. no.	Structure	X	m	n	Ar ¹	Ar ² (#1)	Ar ² (#2)	q	L _M	p	M
IIa-63f		-X(IIa-63f)	1	2				1	-CH ₂ CH ₂ -	1	
IIa-64b		-Cl	1	2				1	-OCH ₂ CH ₂ CH ₂ CH ₂ CH ₂ - CH ₂ -	1	
IIa-66		-OH	1	2				1	-CH ₂ CH ₂ CH ₂ CH ₂ - CH ₂ CH ₂ -	1	

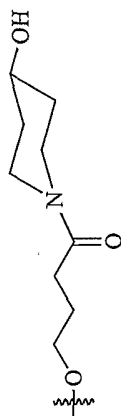
Group X of formula -X(IIa-63c) has the following structure:



Group X of formula -X(IIa-63d) has the following structure:



Group X of formula -X(IIa-63e) has the following structure:



5 Group X of formula -X(IIa-63f) has the following structure:

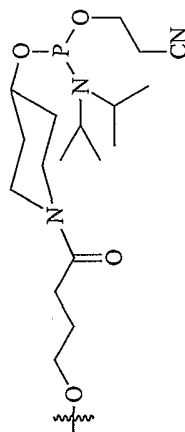


Table 4

Compounds of formulae (IIb-28c), (IIb-28d) and (IIb-47b)

Cpd. no.	Structure	X★	m	n	Ar ¹ (#1)	Ar ¹ (#2)	Ar ²	L ⁵	q	L _M	p	M
IIb-28c		BF ₄ ⁻	2	1				N	1		1	-I
IIb-28d		BF ₄ ⁻	2	1				L ⁵ (#1) = N L ⁵ (#2) = O L ⁵ (#3) = O	1		1	-I
IIb-47b		BF ₄ ⁻	2	1				L ⁵ (#1) = N L ⁵ (#2) = O L ⁵ (#3) = O	1		1	

Methods for synthesising compounds of formulae (IIa-1) to (IIa-64b), (IIb-28c), (IIb-28d) and (IIb-47b) are described in detail in European patent application 04 104 605.3, published as EP 1 506 959 A, as summarised in table 5 below:

Table 5

Compound	Synthetic Route – Example of EP 1 506 959 A
IIa-1a	Example 1 (compound 1a)
IIa-2a	Example 2 (compound 2a = compound 25a)
IIa-3a	Example 3 (compound 3a)
IIa-4a	Example 4 (compound 4a)
IIa-5a	Example 5 (compound 5a)
IIa-6a	Example 6 (compound 6a = compound 19a)
IIa-6b	Example 6 (compound 6b)
IIa-6c	Example 6 (compound 6c)
IIa-6d	Example 6 (compound 6d)
IIa-7a	Example 7 (compound 7a)
IIa-7b	Example 7 (compound 7b)
IIa-7c	Example 7 (compound 7c)
IIa-8a	Example 8 (compound 8a)
IIa-8b	Example 8 (compound 8b)
IIa-8c	Example 8 (compound 8c = compound 35Aa)
IIa-9a	Example 9
IIa-9b	Example 9
IIa-9c	Example 9
IIa-10a	Example 10 (compound 10a)
IIa-10b	Example 10 (compound 10b)
IIa-10c	Example 10 (compound 10c)
IIa-11a	Example 11 (compound 11a)
IIa-11b	Example 11 (compound 11b)
IIa-11c	Example 11 (compound 11c = compound 64a)
IIa-12a	Example 12 (compound 12a)
IIa-12b	Example 12 (compound 12b)
IIa-12c	Example 12 (compound 12c)
IIa-13a	Example 13 (compound 13a)
IIa-13b	Example 13 (compound 13b)
IIa-14a	Example 14 (compound 14a = compound 20b = compound 62b)

Compound	Synthetic Route – Example of EP 1 506 959 A
IIa-14b	Example 14 (compound 14b)
IIa-14c	Example 14 (compound 14c)
IIa-15a	Example 15 (compound 15a)
IIa-15b	Example 15 (compound 15b)
IIa-15c	Example 15 (compound 15c)
IIa-16a	Example 15 (compound 16a)
IIa-17a	Example 17 (compound 17a = compound 18a)
IIa-17b	Example 17 (compound 17b = compound 18b)
IIa-17c	Example 17 (compound 17c = compound 18c)
IIa-18d	Example 18 (compound 18d)
IIa-19b	Example 19 (compound 19b)
IIa-19c	Example 19 (compound 19c)
IIa-19d	Example 19 (compound 19d)
IIa-19e	Example 19 (compound 19e)
IIa-20a	Example 20 (compound 20a)
IIa-20c	Example 20 (compound 20c)
IIa-24a	Example 24 (compound 24a)
IIa-24b	Example 24 (compound 24b)
IIa-24c	Example 24 (compound 24c)
IIb-28c	Example 28 (compound 28c)
IIb-28d	Example 28 (compound 28d = compound 47a)
IIa-30Aa	Example 30A (compound 30Aa)
IIa-30Ab	Example 30A (compound 30Ab = compound 35Ba = compound 35Ca)
IIa-30Ac	Example 30A (compound 30Ac)
IIa-32a	Example 32 (compound 32a = compound 40a)
IIa-33a	Example 33 (compound 33a)
IIa-34a	Example 34 (compound 34a)
IIa-35Ab	Example 35A (compound 35Ab)
IIa-35Bb	Example 35B (compound 35Bb)
IIa-35Bc	Example 35B (compound 35Bc)
IIa-35Bd	Example 35B (compound 35Bd)
IIa-35Cb	Example 35C (compound 35Cb)
IIa-35Cc	Example 35C (compound 35Cc)
IIa-35Cd	Example 35C (compound 35Cd)
IIa-36a	Example 36 (compound 36a)

Compound	Synthetic Route – Example of EP 1 506 959 A
IIa-37a	Example 37 (compound 37a)
IIa-37b	Example 37 (compound 37b)
IIa-38a	Example 38 (compound 38a)
IIa-38b	Example 38 (compound 38b)
IIa-41a	Example 41 (compound 41a)
IIa-41b	Example 41 (compound 41b)
IIa-42a	Example 42 (compound 42a)
IIa-42b	Example 42 (compound 42b)
IIa-43a	Example 43 (compound 43a)
IIa-43b	Example 43 (compound 43b)
IIb-47b	Example 47 (compound 47b)
IIa-48a	Example 48 (compound 48a)
IIa-48b	Example 48 (compound 48b)
IIa-48c	Example 48 (compound 48c)
IIa-48d	Example 48 (compound 48d)
IIa-48e	Example 48 (compound 48e)
IIa-49a	Example 49 (compound 49a)
IIa-50a	Example 50 (compound 50a)
IIa-51a	Example 51 (compound 51a)
IIa-51b	Example 51 (compound 51b)
IIa-53a	Example 53 (compound 53a)
IIa-53b	Example 53 (compound 53b)
IIa-53c	Example 53 (compound 53c)
IIa-53d	Example 53 (compound 53d)
IIa-53e	Example 53 (compound 53e)
IIa-54a	Example 54 (compound 54a)
IIa-54b	Example 54 (compound 54b)
IIa-54c	Example 54 (compound 54c)
IIa-56a	Example 56 (compound 56a)
IIa-56b	Example 56 (compound 56b)
IIa-57a	Example 57 (compound 57a)
IIa-58a	Example 58 (compound 58a)
IIa-59a	Example 59 (compound 59a)
IIa-59b	Example 59 (compound 59b)
IIa-60a	Example 60 (compound 60a)

Compound	Synthetic Route – Example of EP 1 506 959 A
IIa-60b	Example 60 (compound 60b)
IIa-61a	Example 61 (compound 61a)
IIa-62a	Example 63 (compound 62a = compound 63a)
IIa-63b	Example 63 (compound 63b)
IIa-63c	Example 63 (compound 63c)
IIa-63d	Example 63 (compound 63d)
IIa-63e	Example 63 (compound 63e)
IIa-63f	Example 63 (compound 63f)
IIa-64b	Example 64 (compound 64b)

Compound IIa-66 may be synthesised similarly to compound IIa-8c by example 8 of EP 1 506 959 A, but by utilising N,N-disulfosuccinimidyl carbonate in place of N,N-disuccinimidyl carbonate.

- 5 Syntheses for compounds of formulae (IIa-67) and (IIa-68) are described in examples 3 and 4 herein, respectively.

The compound of formula (IIa-69) may be synthesised by the route described in example 5 herein.

Intermediates of Formulae (IIa') and (IIb')

- Methods for synthesising compounds of formulae (IIa'-39a), (IIa'-44a), (IIa'-52a), (IIa'-55a) and
 10 (IIb'-28b) are described in detail in European patent application 04 104 605.3, published as EP 1 506 959 A, as summarised in table 6 below:

Table 6

Compound	Synthetic Route – Example of EP 1 506 959 A
(IIa'-39a)	Example 39 (compound 39a)
(IIa'-44a)	Example 44 (compound 44a)
(IIa'-52a)	Example 52 (compound 52a)
(IIa'-55a)	Example 55 (compound 55a)
(IIb'-28b)	Example 28 (compound 28b)

- Syntheses for compounds of formulae (IIa'-70) and (IIa'-71) are described in examples 1 and 2
 15 herein, respectively.

Intermediates of the invention may be modified into compounds of formulae (IIa) or (IIb), *e.g.* by the addition of one or more groups $L_M\{M\}_p$, by the procedures disclosed in EP 1 506 959 A and the documents mentioned below describing the synthesis of compounds of formulae (IIa) or (IIb).

Preparation of Compounds of Formula (IIa) or (IIb)

The compounds of formula (IIa) or (IIb) are available commercially or may be synthesised by known techniques.

Commercially available compounds of formula (IIa) or (IIb) are disclosed, for example in the
5 Molecular Probes Catalogue, 2002. Commercially available trityls, and derivatives and analogues thereof, may also be derivatised with the groups $(L_M\{M\}_p)_q$ by known techniques.

Methods for synthesis of compounds of formula (IIa) or (IIb) useful in the present invention are described in Chem. Soc. Rev. (2003) 32, p. 3-13, scheme 2 and "1. introduction", last two paragraphs. Groups $(L_M\{M\}_p)_q$ are usually introduced into the intermediates and the compounds are
10 then assembled using the appropriate pathways. Alternatively, the groups $(L_M\{M\}_p)_q$ may be added after assembly of the aromatic groups and a-carbon of the compounds.

Methods for synthesis of compounds of formulae (IIa) or (IIb) are also described in WO99/60007.

Chemical Groups

The ions of the invention are stabilised by the resonance effect of the aromatic groups Ar^1 and Ar^2 .
15 The term 'C★' is a carbon atom bearing a single positive charge or a single negative charge' therefore not only includes structures having the charge localised on the carbon atom but also resonance structures in which the charge is delocalised from the carbon atom.

The term 'linker atom or group' includes any divalent atom or divalent group.

The term 'aromatic group' includes quasi and/or pseudo-aromatic groups, e.g. cyclopropyl and
20 cyclopropylene groups.

The term 'halogen' includes fluorine, chlorine, bromine and iodine.

The term 'hydrocarbyl' includes linear, branched or cyclic monovalent groups consisting of carbon and hydrogen. Hydrocarbyl groups thus include alkyl, alkenyl and alkynyl groups, cycloalkyl (including polycycloalkyl), cycloalkenyl and aryl groups and combinations thereof, e.g.
25 alkylcycloalkyl, alkylpolycycloalkyl, alkylaryl, alkenylaryl, cycloalkylaryl, cycloalkenylaryl, cycloalkylalkyl, polycycloalkylalkyl, arylalkyl, arylalkenyl, arylcycloalkyl and arylcycloalkenyl groups. Preferred hydrocarbyl are C_{1-14} hydrocarbyl, more preferably C_{1-8} hydrocarbyl.

Unless indicated explicitly otherwise, where combinations of groups are referred to herein as one moiety, e.g. arylalkyl, the last mentioned group contains the atom by which the moiety is attached to
30 the rest of the molecule.

The term 'hydrocarbylene' includes linear, branched or cyclic divalent groups consisting of carbon and hydrogen formally made by the removal of two hydrogen atoms from the same or different (preferably different) skeletal atoms of the group. Hydrocarbylene groups thus include alkylene, alkenylene and alkynylene groups, cycloalkylene (including polycycloalkylene), cycloalkenylene and
35 arylene groups and combinations thereof, e.g. alkylene-cycloalkylene, alkylene-polycycloalkylene,

alkylenearylene, alkenylenearylene, cycloalkylenealkylene, polycycloalkylenealkylene, arylenealkylene and arylenealkenylene groups. Preferred hydrocarbylene are C₁₋₁₄ hydrocarbylene, more preferably C₁₋₈ hydrocarbylene.

The term 'hydrocarbyloxy' means hydrocarbyl-O-.

- 5 The terms 'alkyl', 'alkylene', 'alkenyl', 'alkenylene', 'alkynyl', or 'alkynylene' are used herein to refer to both straight, cyclic and branched chain forms. Cyclic groups include C₃₋₈ groups, preferably C₅₋₈ groups.

The term 'alkyl' includes monovalent saturated hydrocarbyl groups. Preferred alkyl are C₁₋₈, more preferably C₁₋₄ alkyl such as methyl, ethyl, n-propyl, i-propyl or t-butyl groups.

- 10 Preferred cycloalkyl are C₅₋₈ cycloalkyl.

The term 'alkoxy' means alkyl-O-.

The term 'alkenyl' includes monovalent hydrocarbyl groups having at least one carbon-carbon double bond and preferably no carbon-carbon triple bonds. Preferred alkenyl are C₂₋₄ alkenyl.

- 15 The term 'alkynyl' includes monovalent hydrocarbyl groups having at least one carbon-carbon triple bond and preferably no carbon-carbon double bonds. Preferred alkynyl are C₂₋₄ alkynyl.

The term 'aryl' includes monovalent aromatic groups, such as phenyl or naphthyl. In general, the aryl groups may be monocyclic or polycyclic fused ring aromatic groups. Preferred aryl are C_{6-C₁₄}aryl.

- 20 Other examples of aryl groups are monovalent derivatives of aceanthrylene, acenaphthylene, acephenanthrylene, anthracene, azulene, chrysene, coronene, fluoranthene, fluorene, *as*-indacene, *s*-indacene, indene, naphthalene, ovalene, perylene, phenalene, phenanthrene, picene, pleiadene, pyrene, pyranthrene and rubicene.

The term 'alkylene' includes divalent saturated hydrocarbylene groups. Preferred alkylene are C₁₋₄ alkylene such as methylene, ethylene, n-propylene, i-propylene or t-butylene groups.

Preferred cycloalkylene are C₅₋₈ cycloalkylene.

- 25 The term 'alkenylene' includes divalent hydrocarbylene groups having at least one carbon-carbon double bond and preferably no carbon-carbon triple bonds. Preferred alkenylene are C₂₋₄ alkenylene.

The term 'alkynylene' includes divalent hydrocarbylene groups having at least one carbon-carbon triple bond and preferably no carbon-carbon double bonds. Preferred alkynylene are C₂₋₄ alkynylene.

- 30 The term 'arylene' includes divalent aromatic groups, such phenylene or naphthylene. In general, the arylene groups may be monocyclic or polycyclic fused ring aromatic groups. Preferred arylene are C_{6-C₁₄}arylene.

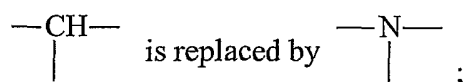
Other examples of arylene groups are divalent derivatives of aceanthrylene, acenaphthylene, acephenanthrylene, anthracene, azulene, chrysene, coronene, fluoranthene, fluorene, *as*-indacene, *s*-

indacene, indene, naphthalene, ovalene, perylene, phenalene, phenanthrene, picene, pleiadene, pyrene, pyranthrene and rubicene.

The term 'heterohydrocarbyl' includes hydrocarbyl groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced
 5 independently by O, S, Se or N, preferably O, S or N. Heterohydrocarbyl groups thus include heteroalkyl, heteroalkenyl and heteroalkynyl groups, cycloheteroalkyl (including polycycloheteroalkyl), cycloheteroalkenyl and heteroaryl groups and combinations thereof, *e.g.*
 heteroalkylcycloalkyl, alkylcycloheteroalkyl, heteroalkylpolycycloalkyl, alkylpolycycloheteroalkyl, heteroalkylaryl, alkylheteroaryl, heteroalkenylaryl, alkenylheteroaryl, cycloheteroalkylaryl,
 10 cycloalkylheteroaryl, heterocycloalkenylaryl, cycloalkenylheteroaryl, cycloalkylheteroalkyl, cycloheteroalkylalkyl, polycycloalkylheteroalkyl, polycycloheteroalkylalkyl, arylheteroalkyl, heteroarylalkyl, arylheteroalkenyl, heteroarylalkenyl, arylcycloheteroalkyl, heteroarylpolycycloalkyl, arylheterocycloalkenyl and heteroarylpolycycloalkenyl groups. The heterohydrocarbyl groups may be attached to the remainder of the compound by any carbon or hetero (*e.g.* nitrogen) atom.

15 The term 'heterohydrocarbylene' includes hydrocarbylene groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N. Heterohydrocarbylene groups thus include heteroalkylene, heteroalkenylene and heteroalkynylene groups, cycloheteroalkylene (including polycycloheteroalkylene), cycloheteroalkenylene and heteroarylene groups and combinations thereof,
 20 *e.g.* heteroalkylenecycloalkylene, alkylenecycloheteroalkylene, heteroalkylenepolycycloalkylene, alkylenepolycycloheteroalkylene, heteroalkylenearylene, alkyleneheteroarylene, heteroalkenylenearylene, alkenyleneheteroarylene, cycloalkyleneheteroalkylene, cycloheteroalkylenealkylene, polycycloalkyleneheteroalkylene, polycycloheteroalkylenealkylene, aryleneheteroalkylene, heteroarylenealkylene, aryleneheteroalkenylene, heteroarylenealkenylene
 25 groups. The heterohydrocarbylene groups may be attached to the remainder of the compound by any carbon or hetero (*e.g.* nitrogen) atom.

Where reference is made to a carbon atom of a hydrocarbyl or other group being replaced by an O, S, Se or N atom, what is intended is that:



30 —CH= is replaced by —N=; or

—CH₂— is replaced by —O—, —S— or —Se—.

The term 'heteroalkyl' includes alkyl groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N.

The term 'heteroalkenyl' includes alkenyl groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N.

The term 'heteroalkynyl' includes alkynyl groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N.

The term 'heteroaryl' includes aryl groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N. Preferred heteroaryl are C₅₋₁₄heteroaryl. Examples of heteroaryl are pyridyl, pyrrolyl, thienyl or furyl.

Other examples of heteroaryl groups are monovalent derivatives of acridine, carbazole, β -carboline, chromene, cinnoline, furan, imidazole, indazole, indole, indolizine, isobenzofuran, isochromene, isoindole, isoquinoline, isothiazole, isoxazole, naphthyridine, perimidine, phenanthridine, phenanthroline, phenazine, phthalazine, purine, pyran, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole, pyrrolizine, quinazoline, quinoline, quinolizine, quinoxaline, thiophene and xanthene. Preferred heteroaryl groups are five- and six-membered monovalent derivatives, such as the monovalent derivatives of furan, imidazole, isothiazole, isoxazole, pyran, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole, pyrrolizine and thiophene. The five-membered monovalent derivatives are particularly preferred, *i.e.* the monovalent derivatives of furan, imidazole, isothiazole, isoxazole, pyrazole, pyrrole and thiophene.

The term 'heteroalkylene' includes alkylene groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N.

The term 'heteroalkenylene' includes alkenylene groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N.

The term 'heteroalkynylene' include alkynylene groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N.

The term 'heteroarylene' includes arylene groups in which up to three carbon atoms, preferably up to two carbon atoms, more preferably one carbon atom, are each replaced independently by O, S, Se or N, preferably O, S or N. Preferred heteroarylene are C₅₋₁₄heteroarylene. Examples of heteroarylene are pyridylene, pyrrolylene, thienylene or furylene.

Other examples of heteroarylene groups are divalent derivatives (where the valency is adapted to accommodate the q instances of the linker L_M) of acridine, carbazole, β -carboline, chromene, cinnoline, furan, imidazole, indazole, indole, indolizine, isobenzofuran, isochromene, isoindole,

isoquinoline, isothiazole, isoxazole, naphthyridine, perimidine, phenanthridine, phenanthroline, phenazine, phthalazine, purine, pyran, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole, pyrrolizine, quinazoline, quinoline, quinolizine, quinoxaline, thiophene and xanthene. Preferred heteroarylene groups are five- and six-membered divalent derivatives, such as the divalent derivatives of furan, imidazole, isothiazole, isoxazole, pyran, pyrazine, pyrazole, pyridazine, pyridine, pyrimidine, pyrrole, pyrrolizine and thiophene. The five-membered divalent derivatives are particularly preferred, *i.e.* the divalent derivatives of furan, imidazole, isothiazole, isoxazole, pyrazole, pyrrole and thiophene.

Substitution

10 A is independently a substituent, preferably a substituent S_{ub}^1 . Alternatively, A may be 2H .

S_{ub}^1 is independently halogen, trihalomethyl, $-NO_2$, $-CN$, $-N^+(R^1)_2O^-$, $-CO_2H$, $-CO_2R^1$, $-SO_3H$, $-SOR^1$, $-SO_2R^1$, $-SO_3R^1$, $-OC(=O)OR^1$, $-C(=O)H$, $-C(=O)R^1$, $-OC(=O)R^1$, $-NR^1_2$, $-C(=O)NH_2$, $-C(=O)NR^1_2$, $-N(R^1)C(=O)OR^1$, $-N(R^1)C(=O)NR^1_2$, $-OC(=O)NR^1_2$, $-N(R^1)C(=O)R^1$, $-C(=S)NR^1_2$, $-NR^1C(=S)R^1$, $-SO_2NR^1_2$, $-NR^1SO_2R^1$, $-N(R^1)C(=S)NR^1_2$, $-N(R^1)SO_2NR^1_2$, $-R^1$ or $-Z^1R^1$.

15 Z^1 is O, S, Se or NR^1 .

R^1 is independently H, C_{1-8} hydrocarbyl, C_{1-8} hydrocarbyl substituted with one or more S_{ub}^2 , C_{1-8} heterohydrocarbyl or C_{1-8} heterohydrocarbyl substituted with one or more S_{ub}^2 .

S_{ub}^2 is independently halogen, trihalomethyl, $-NO_2$, $-CN$, $-N^+(C_{1-6}alkyl)_2O^-$, $-CO_2H$, $-CO_2C_{1-6}alkyl$, $-SO_3H$, $-SOC_{1-6}alkyl$, $-SO_2C_{1-6}alkyl$, $-SO_3C_{1-6}alkyl$, $-OC(=O)OC_{1-6}alkyl$, $-C(=O)H$, $-C(=O)C_{1-6}alkyl$, $-OC(=O)C_{1-6}alkyl$, $-N(C_{1-6}alkyl)_2$, $-C(=O)NH_2$, $-C(=O)N(C_{1-6}alkyl)_2$, $-N(C_{1-6}alkyl)C(=O)O(C_{1-6}alkyl)$, $-N(C_{1-6}alkyl)C(=O)N(C_{1-6}alkyl)_2$, $-OC(=O)N(C_{1-6}alkyl)_2$, $-N(C_{1-6}alkyl)C(=O)C_{1-6}alkyl$, $-C(=S)N(C_{1-6}alkyl)_2$, $-N(C_{1-6}alkyl)C(=S)C_{1-6}alkyl$, $-SO_2N(C_{1-6}alkyl)_2$, $-N(C_{1-6}alkyl)SO_2C_{1-6}alkyl$, $-N(C_{1-6}alkyl)C(=S)N(C_{1-6}alkyl)_2$, $-N(C_{1-6}alkyl)SO_2N(C_{1-6}alkyl)_2$, $C_{1-6}alkyl$ or $-Z^1C_{1-6}alkyl$.

25 Where reference is made to a substituted group, the substituents are preferably from 1 to 5 in number, most preferably 1.

Preferred examples of substituent group A are shown in figure 5.

Miscellaneous

A may optionally be a monovalent dendrimer radical or a monovalent dendrimer radical substituted with one or more substituents S_{ub}^1 .

General

The term "comprising" means "including" as well as "consisting" *e.g.* a composition "comprising" X may consist exclusively of X or may include something additional *e.g.* X + Y.

The term "about" in relation to a numerical value x means, for example, $x \pm 10\%$.

The word “substantially” does not exclude “completely” *e.g.* a composition which is “substantially free” from Y may be completely free from Y. Where necessary, the word “substantially” may be omitted from the definition of the invention.

Tables

5 Table 1 — $C^\star$ is a cation

Formula	Structure
Formula (I')	$(Ar^2)_n - \underset{\oplus}{C} - [Ar^1 - (L_M\{M\}_p)_q]_m$
Formula (I)	$(Ar^2)_n - \underset{\oplus}{C} - [Ar^1 - (L_M - \{M' - B_{P'}\}_p)_q]_m$
Formula (IIb)	$(Ar^2)_n - \underset{\oplus}{C} - [Ar^1 - (L_M\{M\}_p)_q]_m$ $X^\ominus$
Formula (IVbii)	$\begin{array}{c} \textcircled{S_S} \\ \vdots \\ Ar^1 - (L_M\{M\}_p)_q \\ \\ (Ar^2)_n - \underset{\oplus}{C} - [Ar^1 - (L_M\{M\}_p)_q]_{m-1} \end{array}$ $X^\ominus$
Formula (IVbiii)	$\begin{array}{c} \textcircled{S_S} \\ \vdots \\ Ar^2 \\ \\ (Ar^2)_{n-1} - \underset{\oplus}{C} - [Ar^1 - (L_M\{M\}_p)_q]_m \end{array}$ $X^\ominus$
Formula (IVbiv)	$\begin{array}{c} \{M\}_{p-1} L_M M'' - \textcircled{S_S} \\ \\ Ar^1 - (L_M\{M\}_p)_{q-1} \\ \\ (Ar^2)_{n'} - \underset{\oplus}{C} - [Ar^1 - (L_M\{M\}_p)_q]_{m'} \end{array}$ $X^\ominus$

Table 2 — $n = 2, m = 1, p = 1$ and $q = 1$

Formula	Structure
Formula (I')	$\begin{array}{c} Ar^2 \\ \\ Ar^2 - \underset{\star}{C} - Ar^1 - L_M M \end{array}$

Formula (I)	$\begin{array}{c} \text{Ar}^2 \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^1 - \text{L}_M \text{M}' - \text{B}_{\text{P}}' \\ \star \end{array}$
Formula (IIa)	$\begin{array}{c} \text{Ar}^2 \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^1 - \text{L}_M \text{M} \\ \\ \text{X} \end{array}$
Formula (IIb)	$\begin{array}{c} \text{Ar}^2 \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^1 - \text{L}_M \text{M} \\ \star \\ \text{X} \star \end{array}$
Formula (IIIa)	$\begin{array}{c} \text{Ar}^2 \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^1 - \text{L}_M \text{M}' - \text{B}_{\text{P}}' \\ \\ \text{X} \end{array}$
Formula (IIIb)	$\begin{array}{c} \text{Ar}^2 \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^1 - \text{L}_M \text{M}' - \text{B}_{\text{P}}' \\ \star \\ \text{X} \star \end{array}$
Formula (IVai)	$\begin{array}{c} \text{Ar}^2 \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^1 - \text{L}_M \text{M} \\ \\ \text{S}_S \end{array}$
Formula (IVaii)	$\begin{array}{c} \text{S}_S \\ \\ \text{Ar}^1 - \text{L}_M \text{M} \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^2 \\ \\ \text{X} \end{array}$
Formula (IVaiii)	$\begin{array}{c} \text{S}_S \\ \\ \text{Ar}^2 \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^1 - \text{L}_M \text{M} \\ \\ \text{X} \end{array}$
Formula (IVaiv)	$\begin{array}{c} \text{L}_M \text{M}'' - \text{S}_S \\ \\ \text{Ar}^1 \\ \\ \text{Ar}^2 - \text{C} - \text{Ar}^2 \\ \\ \text{X} \end{array}$

Formula (IVbii)	
Formula (IVbiii)	
Formula (IVbiv)	

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B show preferred examples of group L_M and compounds of the invention.

Figures 2A and 2B show preferred examples of group M and compounds of the invention.

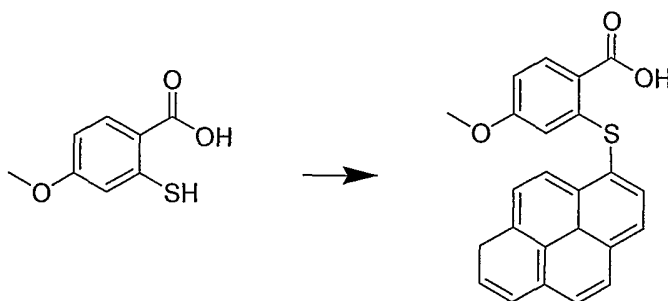
5 Figures 3A and 3B show preferred examples of groups Ar^1 and Ar^2 and compounds of the invention.

Figure 4 shows preferred examples of groups X and X^* and compounds of the invention.

Figure 5 shows preferred examples of substituent group A and compounds of the invention.

MODES FOR CARRYING OUT THE INVENTION

Example 1 – Preparation of Intermediate (IIa'-70)

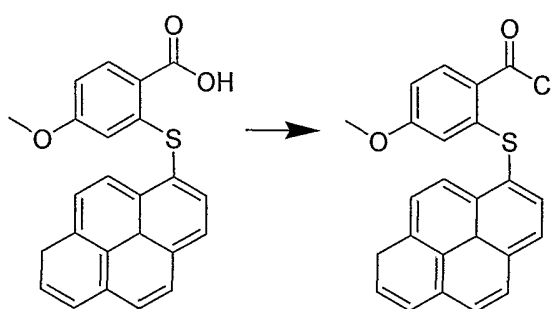


10

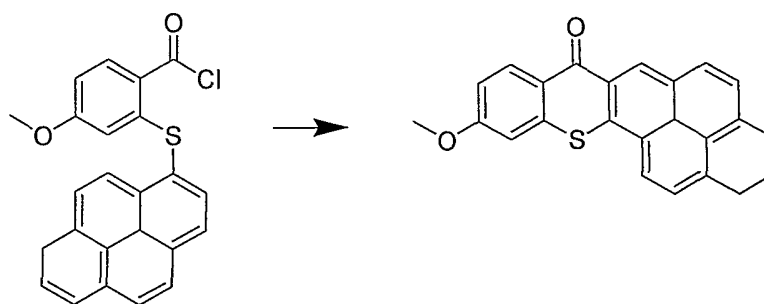
2-(1,3-dihydropyren-8-ylthio)-4-methoxybenzoic acid. 10 g of 2-mercapto-4-methoxybenzoic acid (184.21 mwt, 0.0542 mol), 1-bromopyrene (281.16 mwt, 2.41 g, 1 eqt, 0.0542 mol) were placed in 100 ml round bottom flask. Potassium carbonate (138.21 mwt, 1.18 g, 1 eqt, 0.0542 mol) and 300 mg

of Cu were also added, followed by 70 ml of dry DMF. The reaction was refluxed for 4 hours. The reaction mixture poured into 300 ml of 1N HCl, extracted ethyl acetate (300 ml x 2). The combined organic phases were then washed with water (150 ml x 3). Organic phase was filtered and dried over sodium sulphate, the solvent removed in vacuo. to give 0.474 g, 15 % yield of a single compound by
 5 TLC.

^1H NMR (400 MHz, DMSO- d_6): δ = 8.5 - 8.3 (m, 8H), 8.25 - 8.12 (t, J = 7.65 Hz, 1H), 8.1 - 8.0 (d, J = 8.7 Hz, 1H), 6.8 - 6.7 (d, J = 8.8 Hz, 1H), 5.7 - 5.6 (d, J = 2.4 Hz, 1H), 3.5 - 3.4 (br OH, 1H), 3.35 (s, 3H). ^{13}C NMR (400 MHz, DMSO- d_6): δ = 167.95, 162.98, 145.59, 135.73, 134.73, 134.04, 133.23, 131.49, 131.15, 130.22, 129.75, 128.12, 127.74, 127.12, 126.76, 126.44, 125.48, 125.18,
 10 124.34, 119.95, 113.49, 109.61. HRMS (ESI): m/z calcd for $\text{C}_{24}\text{H}_{16}\text{NaO}_3\text{S}$ [$\text{M} + \text{Na}^+$]: 407.0718; found 407.0717.



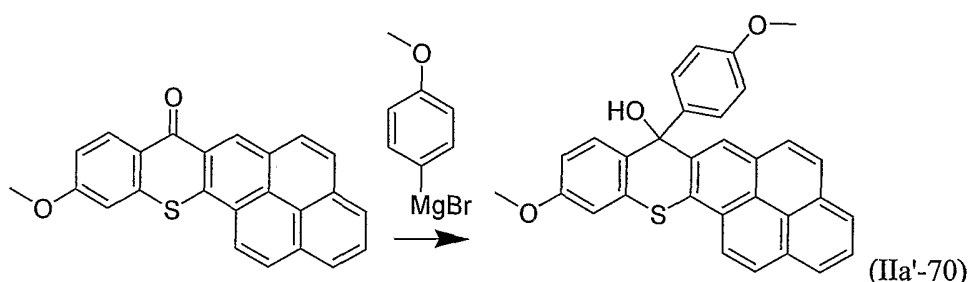
2-(1,3-dihydropyren-8-ylthio)-4-methoxybenzoyl chloride. 10-methoxy-1H-phenaleno[1,9-*bc*]thioxanthen-7(14H)-one (384.44 mwt, 0.474 g, 1.23 mmol), placed in a 50 ml round bottom flask
 15 under an argon atmosphere. 50 ml of dry dichloromethane was added followed by a few drops of dimethylformamide. Oxalyl chloride (126.43 mwt, d 1.455, 4 eqt. 0.63 g, 0.42 ml, 4.93 mmol) was added drop wise with stirring. The reaction was stirred for 1 hour or longer until the suspension dissolved. The product was concentrated under reduced pressure and azeotroped with toluene (5 ml x 3) and then dried under high vacuum to give a foamy solid, 0.472 g, 95%), product was used
 20 immediately.



10-methoxy-1H-phenaleno[1,9-*bc*]thioxanthen-7(14H)-one. The acid chloride (402.89 mwt, 0.472 g, 1.17 mmol) was placed in a 100 ml round bottom flask, 30 ml of dry dichloromethane was added under an argon atmosphere. Aluminium chloride (133.34 mwt, 1.5 eqt., 0.234 g, 1.75 mmol) was
 25 added slowly and the reaction stirred at room temperature. The reaction was complete within 1 hour.

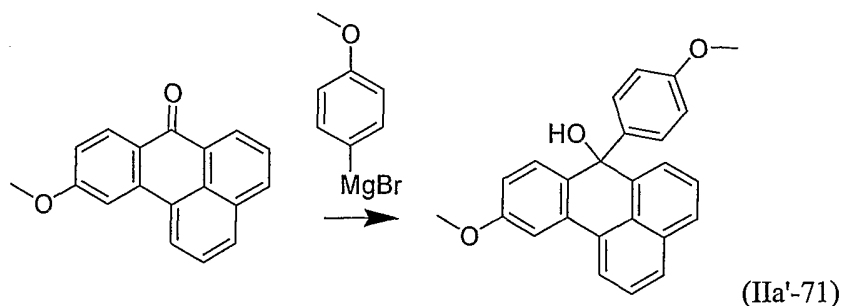
The reaction was slowly quenched with 15 ml of water and extracted with dichloromethane (50 ml x 2). The organic phases were combined and dried over sodium sulphate, filtered and reduced to give the crude product. Product purified via column chromatography, silica gel, hexane: ethyl acetate gradient elution. (8:1 - 1:1). Product obtained as faint yellow solid, 0.35 g, 81% yield.

- 5 Alternatively the product can be purified via recrystallisation, from hexane and ethyl acetate. Reaction repeated to give the product in yields ranging from 80 – 93 %. ¹H NMR (200 MHz, CDCl₃): δ = 8.9 (s, 1H), 8.7 – 8.6 (d, *J* = 8.9 Hz, 1H), 8.59 – 8.48 (d, *J* = 9.3 Hz, 1H), 8.35 – 7.9 (m, 6H), 7.24 – 7 (m, 2H), 4 (s, 3H). ¹³C NMR (500 MHz, CDCl₃): δ = 180.14, 162.71, 138.72, 132.18, 131.98, 131.73, 131.01, 129.15, 128.44, 128.37, 127.79, 127.56, 126.79, 126.50, 126.21, 125.87, 125.78, 125.49, 124.20, 122.58, 122, 115.33, 108.53, 55.84. HRMS (ESI): *m/z* calcd for C₂₄H₁₅O₂S [M + H]: 267.0793; found 267.0797.



- 10-Methoxy-7-(4-methoxyphenyl)-7H-phenaleno[1,9-*bc*]thioxanthen-7-ol 0.09g of ketone (368.44 mwt, 0.244 mmol) was placed in a 100 ml round bottom flask under a positive atmosphere of argon. 10 ml of dry THF was added, followed by 4-methoxyphenyl magnesium bromide (0.5 M solution in THF, 5 eqt., 1.22 mmol, 2.44 ml). The suspension was then refluxed over night. The reaction was slowly quenched with 10 ml of water and extracted with ethyl acetate (50 ml x 2). The organic phases were combined and washed with water (50 ml), dried over sodium sulphate. Filtered and concentrated *in vacuo* to give the crude product. Product purified via column chromatography, silica gel, hexane: ethyl acetate gradient elution. (3:1). Product obtained as faint yellow solid, 0.045 g, 39% yield. HRMS (MALDI): *m/z* calcd for C₃₁H₂₂NaO₃S [M - OH]: 457.5696; found 457.0418.

Example 2 – Preparation of Intermediate (IIa'-71)

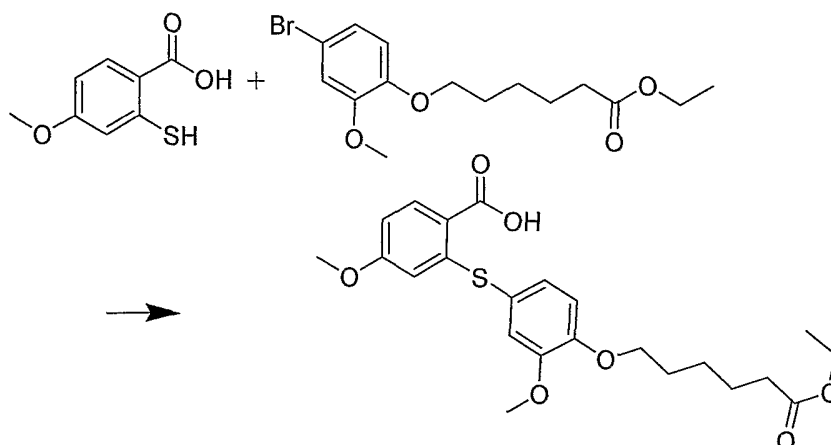


- 10-Methoxy-7-(4-methoxyphenyl)-7H-benzo(*de*)anthracen-7-ol. 3.14 g of 7H-benzo[*de*]anthracen-7-one (230.27 mwt, 0.01 mol) was placed in a 250 ml round bottom flask under a positive atmosphere of argon. 40 ml of dry THF was added, followed by 4-methoxyphenyl

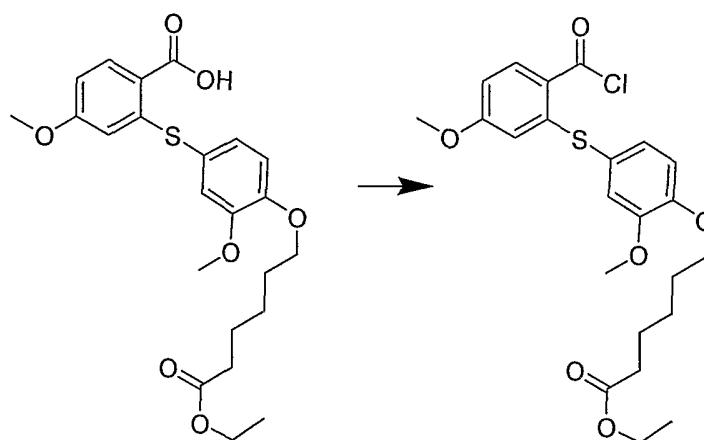
magnesium bromide (0.5 M solution in THF, 1.5 eqt., 0.02 mol, 40.9 ml). The reaction was then stirred over night at room temperature. The reaction was slowly quenched with 50 ml of water and extracted with ethyl acetate (150 ml x 2). The organic phases were combined and washed with water (150 ml), dried over sodium sulphate. Filtered and concentrated *in vacuo* to give the crude product.

- 5 Product purified via column chromatography, silica gel, hexane: ethyl acetate gradient elution. (5:1). Product obtained as faint yellow solid, 3.4 g, 68 % yield. 60 mg of sample purified by preparative TLC (hexane: ethyl acetate). MALDI: m/z calcd for $C_{25}H_{20}NaO_3$ [M^{+}]: 368.1412; found 337.9097.

Example 3 – Preparation of Compound (IIa-67)

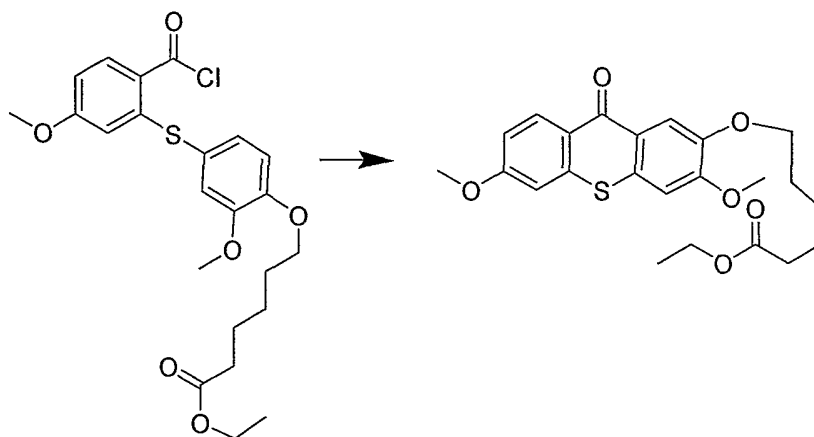


- 10 **2-((4-(5-(ethoxycarbonyl)pentyloxy)-3-methoxyphenyl)sulfanyl)-4-methoxybenzoic acid.** 6.64 g of 2-mercapto-4-methoxybenzoic acid (184.21 mwt, 0.036 mol), 12.83 g ethyl 6-(4-bromo-2-methoxyphenoxy) hexanoate (345.22 mwt, 1eqt, 0.036 mol) were placed in 250 ml round bottom flask. Potassium carbonate (138.21 mwt, 4.95 g, 1 eqt, 0.036 mol) and 0.6 g of Cu were also added, followed by 100 ml of dry DMF. The reaction was refluxed for 4 hours. The reaction mixture was
- 15 poured into 300 ml of 1N HCl, extracted with of ethyl acetate (300 ml x 2), washed with water, (150 ml x 3). Organic phases were combined and dried over sodium sulphate, filtered and the solvent removed in *vacuo*. 14.46 g, 89.4 % of a single compound obtained.



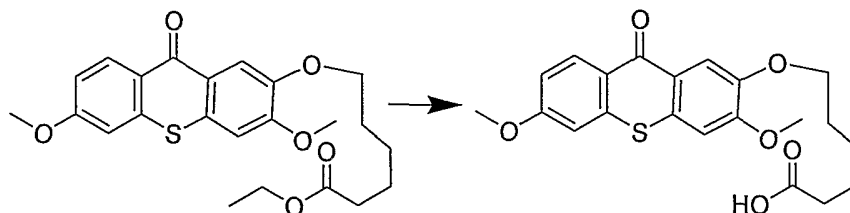
- Ethyl 6-(3,6-dimethoxy-9-oxo-9H-thioxanthen-2-yloxy)hexanoyl chloride.** The acid (448.52 mwt, 5.89 g, 0.013 mol), placed in 100 ml round bottom flask under an argon atmosphere. 50 ml of dry
- 20

dichloromethane was added followed by a few drops of dimethylformamide. Oxalyl chloride (126.43 mwt, d 1.455, 2 eqt. 3.32 g, 2.28 ml, 0.0262 mol) was added dropwise with stirring. The reaction was stirred for 1 hour or longer until the suspension dissolved. The product was concentrated under reduced pressure and azeotroped with toluene (5 ml x 3) and then dried under high vacuum to give a
 5 foamy solid, 6.11 g, 100%. Product used without further purification.



Ethyl 6-(3,6-dimethoxy-9-oxo-9H-thioxanthen-2-yl)oxy)hexanoate. The acid chloride (466.97 mwt, 6.1 g, 0.013 mol) was placed in a 250 ml round bottom flask, 80 ml of dry dichloromethane was added under an argon atmosphere. The reaction was stirred at room temperature and aluminium
 10 chloride (133.34 mwt, 1.5 eqt., 2.61 g, 0.019 mol) was slowly added. The reaction was complete within 1 hour. The reaction was slowly quenched with 30 ml of water and extracted with dichloromethane, (100 ml x 2). The organic phases were then combined and then washed with a solution of sodium chloride (100 ml), dried over sodium sulphate, filtered and reduced to give the crude product.

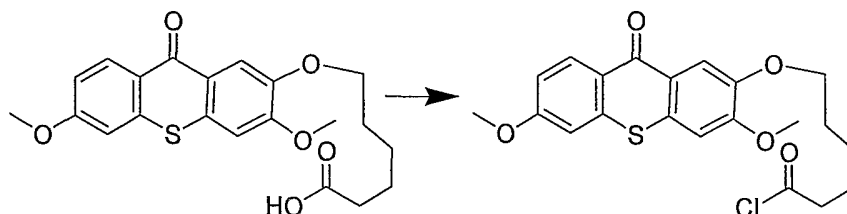
15 Product purified via column chromatography, silica gel, hexane: ethyl acetate gradient elution. (3:1). Product obtained as a pale yellow solid 3.38 g, 60 % yield. Alternatively the product can be purified via recrystallisation, from hexane and ethyl acetate. MALDI: m/z calcd for $C_{23}H_{27}O_6$ [$M + H$]: 430.1450; found 431.0.



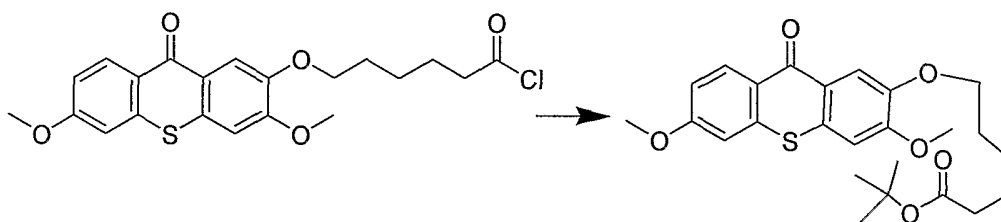
20

6-(3,6-dimethoxy-9-oxo-9H-thioxanthen-2-yl)oxy) hexanoic acid. 2 g (430.51 mwt, 4.46 mmol) of starting material placed in a 100 ml round bottom flask, 20 ml of tetrahydrofuran and methanol respectively were added. Lithium hydroxide (23.95 mwt, 4 eqt., 0.445 g, 18.58 mmol) was added and the reaction heated at reflux for 5 hours. The reaction was allowed to cool to room temperature. The

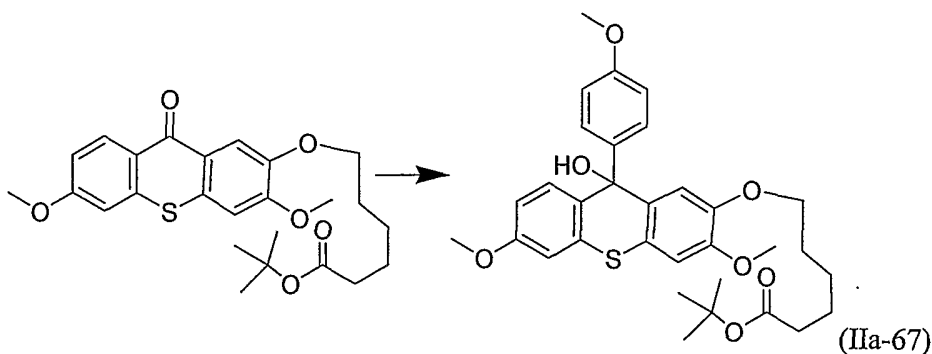
crude reaction mixture was concentrated under reduced pressure too 1/3 the original volume and added to cold 1N HCl. The precipitate generated was filtered and dried under high vacuum to give a white solid, 1.8 g, 96 % yield MALDI: m/z calcd for $C_{21}H_{23}O_6S[M + H]^+$: 402.1137; found 403.0.



- 5 **6-(3,6-dimethoxy-9-oxo-9H-thioxanthen-2-yloxy)hexanoyl chloride.** 1.5 g of acid placed in a dry 100 ml round bottom flask, dry dichloromethane (40ml) was added under an atmosphere of argon. A few drops of dry dimethylformamide was added to the suspension, followed by oxalyl chloride dropwise (126.63 mwt, d 1.455, 3 eqt, 1.41 g, 0.973 ml, 11.18 mmol). The suspension slowly dissolves after 2 hours of stirring. The acid chloride was concentrated under reduced pressure and
- 10 azeotroped with toluene (5 ml x 3). The product was then dried thoroughly under high vacuum and used immediately.

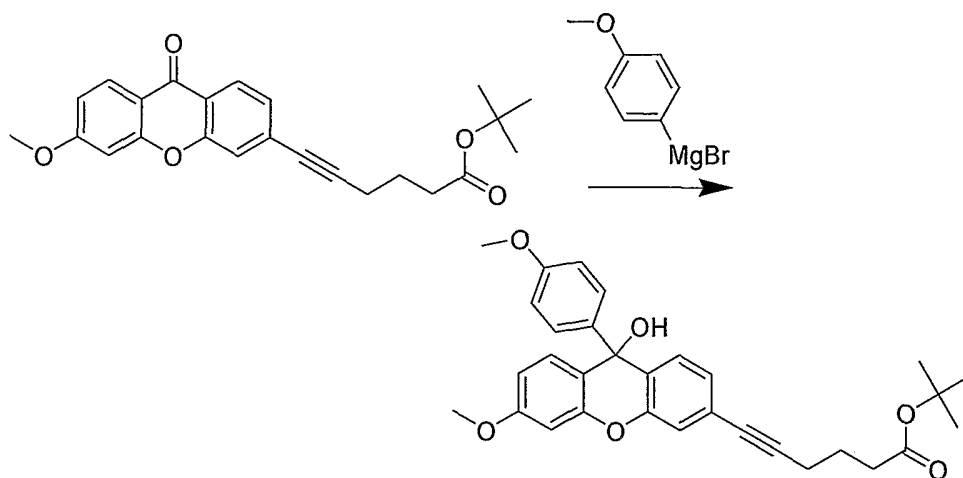


- tert-butyl 6-(3,6-dimethoxy-9-oxo-9H-thioxanthen-2-yloxy)hexanoate.** The acid chloride (420.9 mwt, 1.57 g, 3.73 mmol) was placed in dry 100 ml round bottom flask, dichloromethane and *tert*-butanol, 20 ml and 30 ml respectively were added, followed by triethylamine (101.19 mwt, d 0.726, 2 eqt., 0.76 g, 1.1 ml, 7.46 mmol). The reaction was stirred overnight (TLC control). The reaction mixture was concentrated under reduced pressure and diluted with 100 ml of dichloromethane. The organic phase was washed with sodium bicarbonate solution (50 ml x 3), water (50 ml x 2) and the organic phase dried over sodium sulphate. The product was then filtered and concentrated under
- 20 reduced pressure to give a solid, 1.4 g, 84 %. MALDI: m/z calcd for $C_{21}H_{23}O_6S[M + H]^+$: 402.1137; found 402.87.

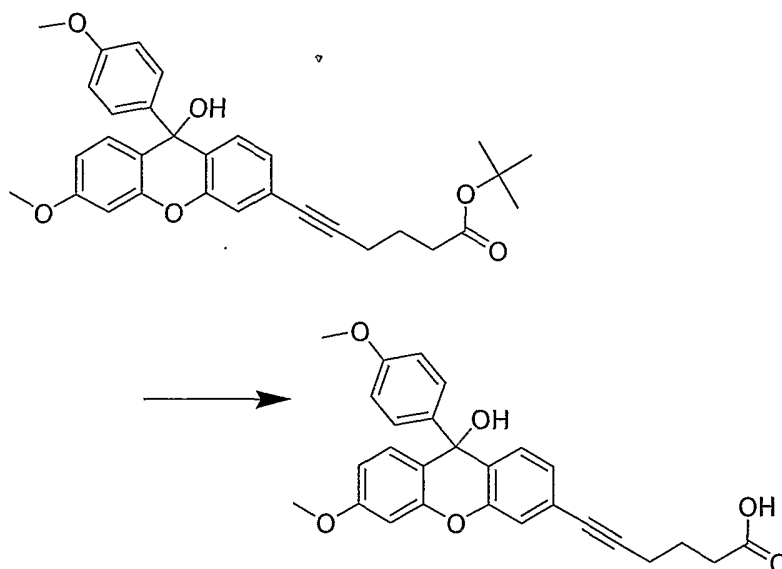


tert-butyl 6-(9-hydroxy-3,6-dimethoxy-9-(4-methoxyphenyl)-9H-thioxanthen-2-yl)hexanoate. 1.395 g of ketone (446.55 mwt, 3.12 mmol) was placed in a dry 100 ml round bottom flask, dry tetrahydrofuran (40 ml) was added. 4-methoxyphenyl magnesium bromide (0.5 M, 2 eqt, 12.49 ml, 6.24 mmol) was added and the reaction mixture was refluxed for 4 hours under an argon atmosphere. (TLC control). The reaction was quenched with 10 ml of water and stirred for 10 minutes. The reaction mixture was concentrated under reduced pressure. Ethyl acetate (50 ml) was added and the organic phase washed with sodium chloride solution (30 ml), water, (30 ml x 2). The organic phase was dried over sodium sulphate, filtered to give the product as a foamy solid, 1.3 g, 73.4 % yield. MALDI: m/z calcd for $C_{32}H_{37}O_6S$ [$M^+ OH$]: 549.2305; found 549.0457.

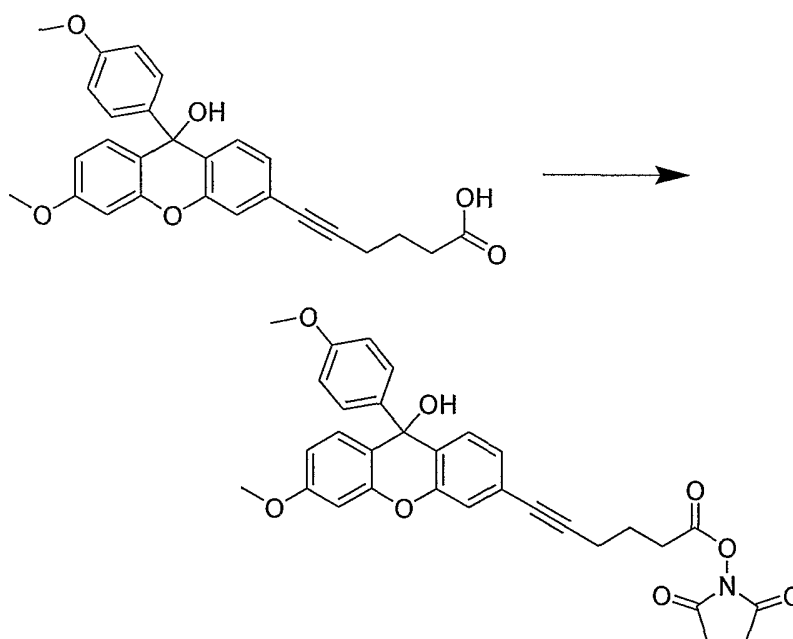
10 Example 4 – Preparation of Compound (IIa-68)



tert-butyl 6-(9-hydroxy-3-methoxy-9-(4-methoxyphenyl)-9H-xanthen-6-yl)hex-5-ynoate. 0.720 g of starting material (392.44 mwt, 1.83 mmol) was added to a dry 100 ml round bottom flask, dry THF (30 ml) was added under an argon atmosphere. 4-methoxyphenyl magnesium bromide (0.5 M solution in THF, 2 eqt, 3.66 mmol, 7.35 ml) was added dropwise to the reaction mixture at room temperature. The reaction was stirred overnight. The reaction mixture was quenched with water (10 ml), concentrated *in vacuo*. Ethyl acetate (150 ml) was added and the organic phase was washed with water (100 ml x 2). The product dried over sodium sulphate, filtered and concentrated *in vacuo*. Crude product purified by column chromatography, hexane:ethyl acetate, gradient elution (4:1). 0.524 g, 57.3 % yield.

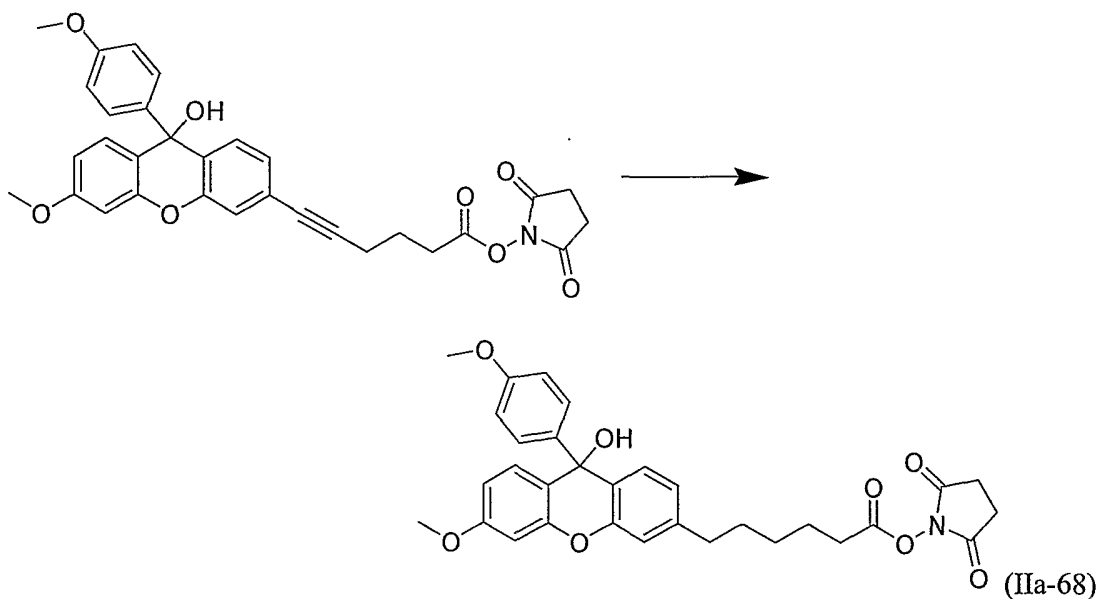


6-(9-hydroxy-3-methoxy-9-(4-methoxyphenyl)-9H-xanthen-6-yl)hex-5-ynoic acid. 0.520 g of starting material (500.58 mwt, 1.038 mmol) was added to a dry 100 ml round bottom flask, DCM: TFA (6 ml respectively) was added and the reaction was stirred overnight. The reaction mixture was concentrated *invacuo*. The product was then azeotroped with toluene (5 ml x 4) until traces of TFA was complete removed. Product isolated as a viscous oil, 0.461 g, 100 % yield.



6-(9-hydroxy-3-methoxy-9-(4-methoxyphenyl)-9H-xanthen-6-yl)hex-5-ynoate-N-hydroxysuccinimide. 0.461 g of starting material (444.47 mwt, 1.037 mmol) was added to a dry 100 ml round bottom flask. Acetonitrile (30 ml) was added followed by N,N'-disuccinimidyl carbonate (256.17 mwt, 1.25 eqt, 1.296 mmol, 0.332 g) and triethylamine (101.19 mwt, d 0.721, 4 eqt., 4.148 mmol, 0.420 g, 0.58 ml). The reaction was stirred overnight. The reaction mixture was concentrated *invacuo*. The crude product dissolved in ethyl acetate (100 ml), organic phase was washed with water (50 ml x 2). The organic was dried over sodium sulphate, filtered and concentrated in *vacuo* to give a

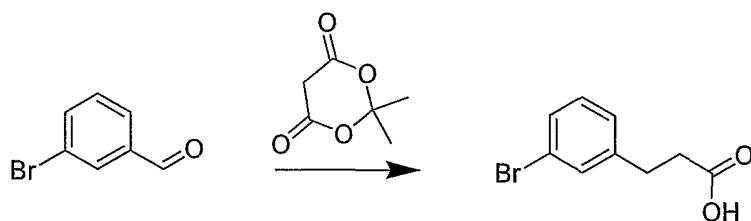
very pure product, isolated as a viscous oil, 0.561 g, 100 % yield. MALDI: m/z calcd for $C_{31}H_{26}NO_7$ $[M^- OH]$: 524.1704; found 524.04.



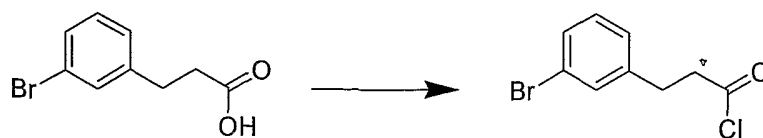
6-(9-hydroxy-3-methoxy-9-(4-methoxyphenyl)-9H-xanthen-6-yl)hexanoate-N-

- 5 **hydroxysuccinimide.** 0.275 g of starting material (541.54 mwt, 0.5078 mmol) was added to a dry 100 ml round bottom flask. Dry ethyl acetate (30 ml) was added followed by Palladium, 10% on carbon (106.4 mwt, 1 eqt, 0.5078 mmol, 0.054 g). The reaction was then purged with hydrogen (3 times), The reaction was stirred for 4 days under a positive pressure of hydrogen. The crude product was then filtered through a short pad of silica and concentrated invacuo. Product, isolated as a
- 10 viscous oil, 0.277 g, 100 % yield. MALDI: m/z calcd for $C_{31}H_{30}NO_7$ $[M^- OH]$: 528.2017; found 527.98.

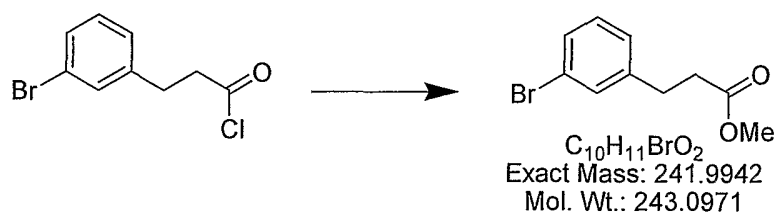
Example 5 – Preparation of Compound (IIa-69)



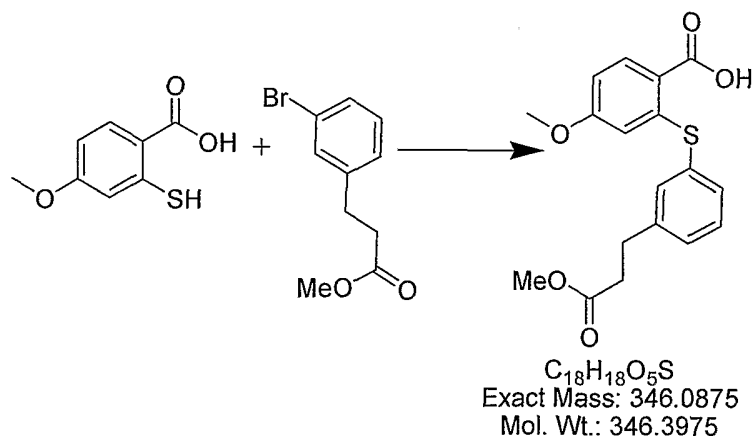
- 15 **3-(3-bromophenyl)propanoic acid.** 450 ml of triethylamine was added dropwise to an ice cold solution of formic acid (300 ml). 3-Bromobenzaldehyde (46.25 g, 0.249 mol, 1 eqt.) and meldrums acid (36 g, 0.249 mol, 1 eqt.) were added. The reaction mixture was refluxed for 20 hours. The reaction mixture was cooled to room temperature and poured into 500 ml of 6N HCl. The precipitate was collected by filtration. The solid was dissolved in 300 ml of chloroform, the organic phase was washed with water (200 ml x 2). The organic phase was dried over magnesium sulphate, filtered and
- 20 concentrated invacuo to give a white solid, 34 grams, 59.4 %



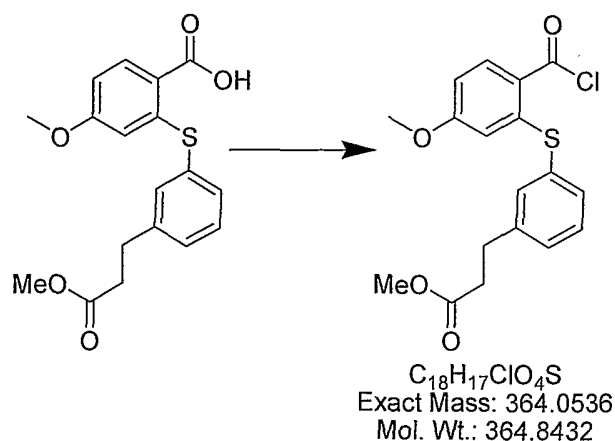
3-(3-bromophenyl)propanoyl chloride. 29.4 grams of 3-(3-bromophenyl)propanoic acid (229.07 mwt, 0.128 mol) was added to a 250 ml round bottom flask. 100 ml of dry dichloromethane was added, followed by a cat. amount of DMF. Oxalyl chloride (126.93 mwt, d 1,478, 1.5 eqt., 24.44 g, 16.53 ml, 0.192 mol) was added slowly at room temperature. The reaction was stirred for 2 hours. The reaction was filtered and concentrated under reduced pressure. The product was azeotroped with toluene (5 ml x 3) to give a viscous oil 31.77 g, 100 %.



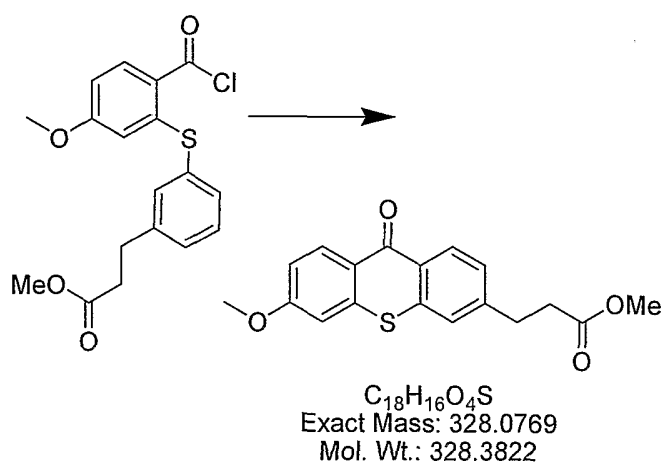
Methyl 3-(3-bromophenyl)propanoate. 30 g of 3-(3-bromophenyl)propanoyl chloride (247.51 mwt, 0.121 mol) was placed in a 250 ml round bottom flask. 60 ml of dry dichloromethane was added followed by the slow addition of dry methanol (100 ml). The reaction was stirred for 2 hours at room temperature. The reaction was concentrated under reduced pressure to give a viscous oil 25.82 g, 87.6 %.



2-((3-(2-(methoxycarbonyl(ethyl)phenyl)sulfanyl)-4-methoxybenzoic acid. 18 g of 2-mercapto-4-methoxybenzoic acid (184.21 mwt, 0.097 mol), 23.75 g methyl 3-(3-bromophenyl)propanoate (243.09 mwt, 1eqt, 0.097 mol) were placed in 250 ml round bottom flask. Potassium carbonate (138.21 mwt, 13.5 g, 1 eqt, 0.097 mol) and 1.0 g of Cu were also added, followed by 100 ml of dry DMF. The reaction was refluxed for 4 hours. The reaction mixture was poured into 300 ml of 1N HCl, extracted with ethyl acetate (300 ml x 2), washed with water, (150 ml x 3). Organic phases were combined and dried over sodium sulphate, filtered and the solvent removed in vacuo. 21.9 g, 65 % of a single compound obtained.

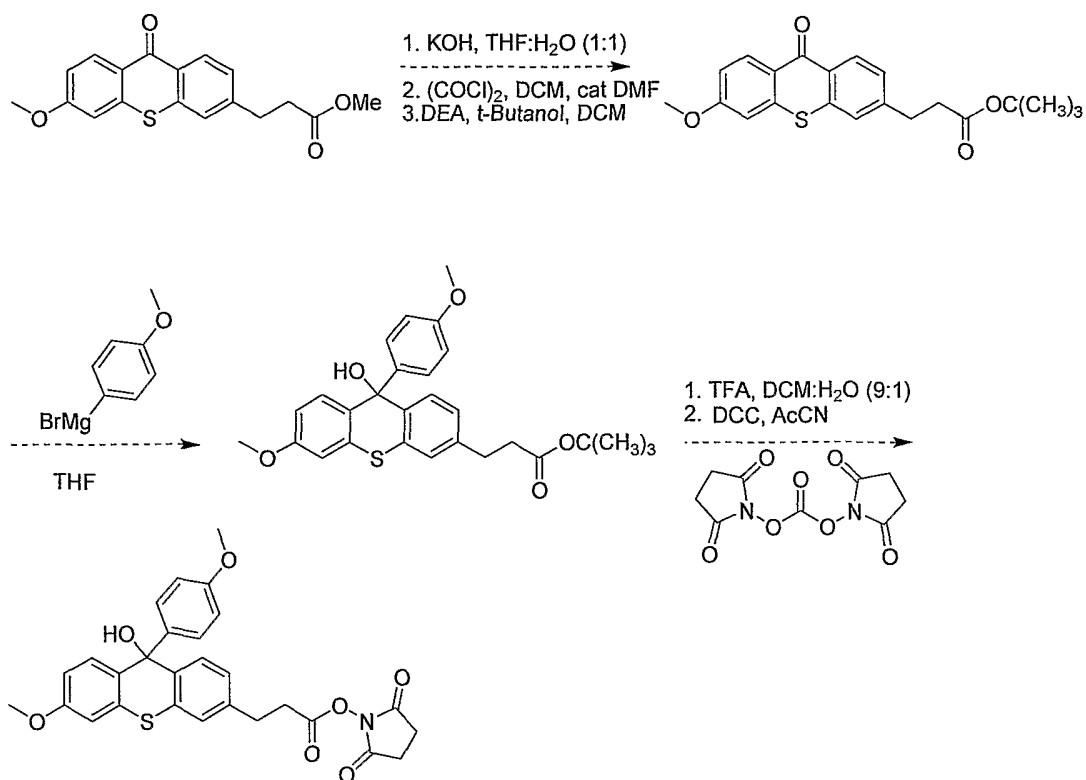


Methyl 3-(3-(2-(chlorocarbonyl)-5-methoxyphenyl thio (phenyl) propanoate. The acid (346.39 mwt, 21.9 g, 0.063 mol) was placed in a 100 ml round bottom flask under an argon atmosphere. 50 ml of dry dichloromethane was added followed by a few drops of dimethylformamide. Oxalyl chloride (126.43 mwt, d 1.455, 2 eqt. 15.98 g, 10.98 ml, 0.126 mol) was added dropwise with stirring. The reaction was stirred for 1 hour or longer until the suspension dissolved. The product was concentrated under reduced pressure and azeotroped with toluene (5 ml x 3) and then dried under high vacuum to give a foamy solid, 23 g, 100%.



- 10 **Methyl 3-(3-methoxy-9-oxo-9H-thioxanthen-6-yl) propanoate.** The acid chloride (364.84 mwt, 23 g, 0.063 mol) was placed in a 100 ml round bottom flask, 100 ml of dry dichloromethane was added under an argon atmosphere. The reaction was stirred at room temperature, aluminium chloride (133.34 mwt, 1.5 eqt., 12.6 g, 0.0945 mol) was slowly added. The reaction was complete within 1 hour. The reaction was slowly quenched with water (30 ml) and extracted with dichloromethane (100 ml x 2). The organic phases were combined and dried over sodium sulphate, filtered and concentrated under reduced pressure to give the crude product. Product purified via column chromatography, silica gel, hexane: ethyl acetate gradient elution. (3:1). Product obtained as faint yellow solid, 6.67 g, 32 % yield.

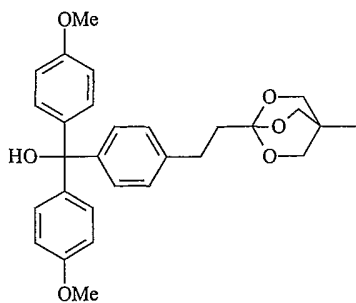
Compound (IIa-69). From the intermediate above, compound (IIa-69) may be prepared as follows:



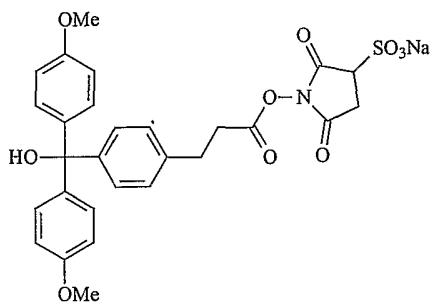
It will be understood that the invention is described above by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention.

CLAIMS

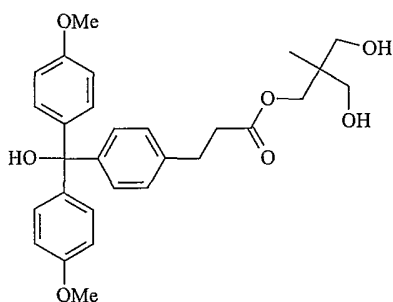
1. A compound of formula:



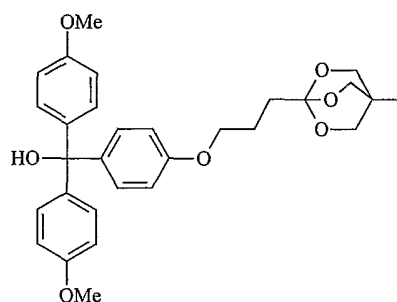
(IIa-1a);



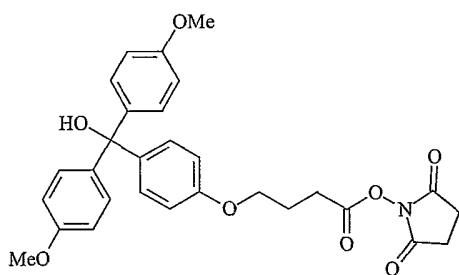
(IIa-2a);



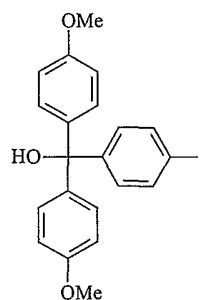
(IIa-3a);



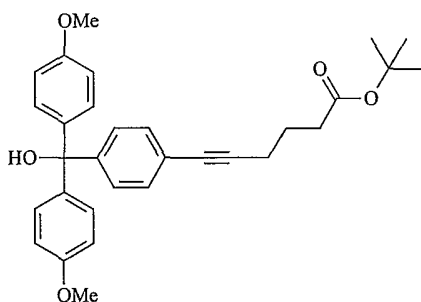
(IIa-4a);



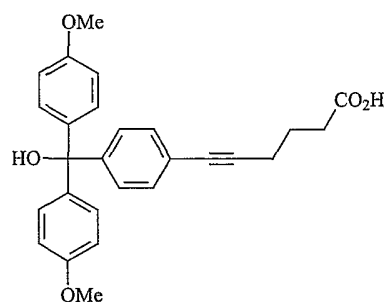
(IIa-5a);



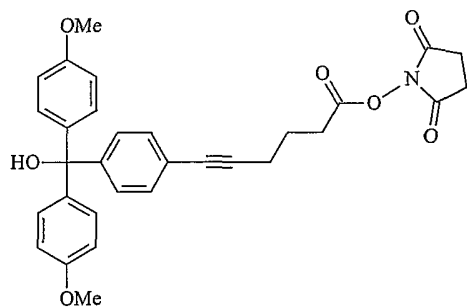
(IIa-6a);



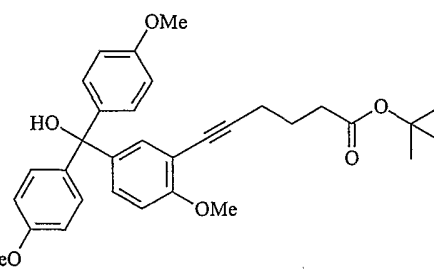
(IIa-6b);



(IIa-6c);

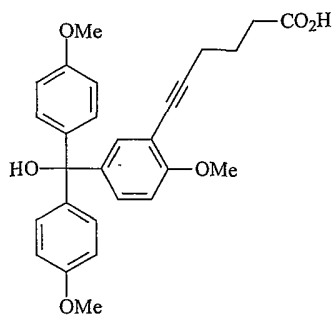


(IIa-6d);

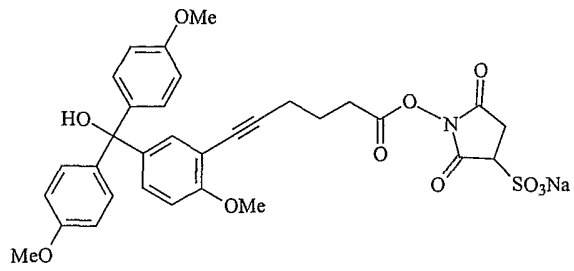


(IIa-7a);

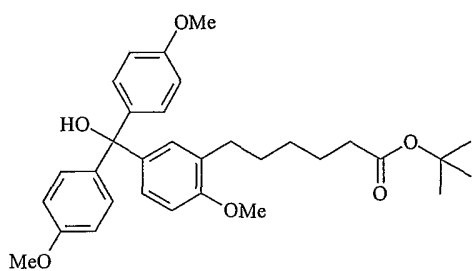
5



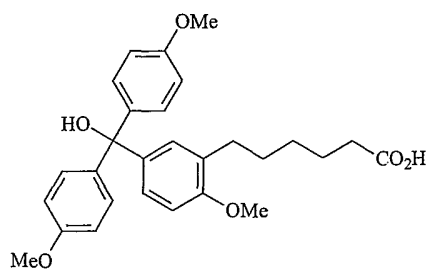
(IIa-7b);



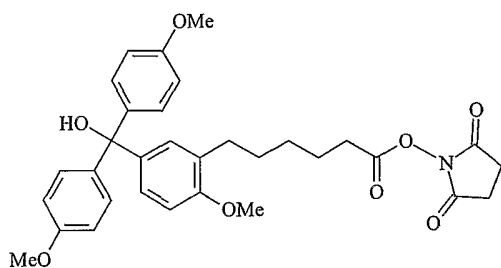
(IIa-7c);



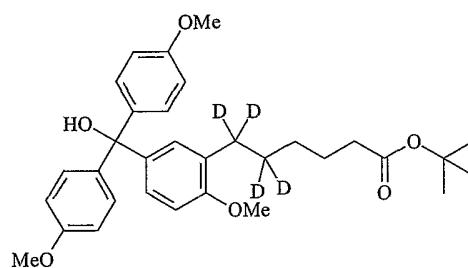
(IIa-8a);



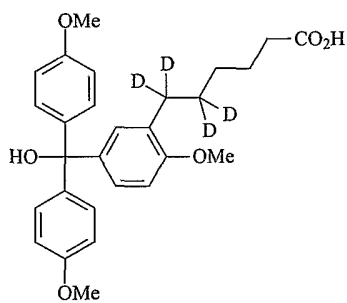
(IIa-8b);



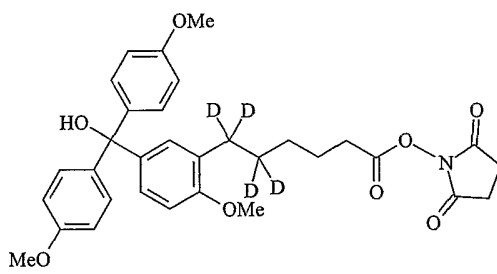
(IIa-8c);



(IIa-9a);

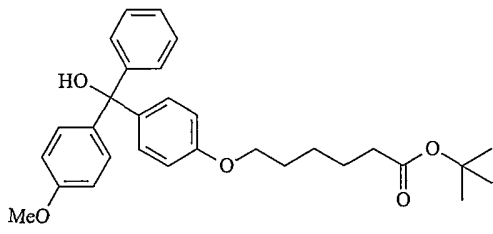


(IIa-9b);

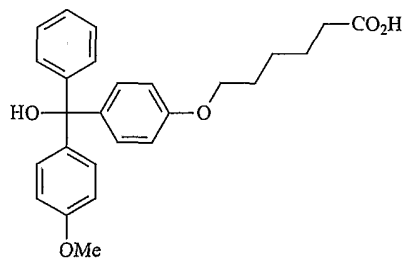


(IIa-9c);

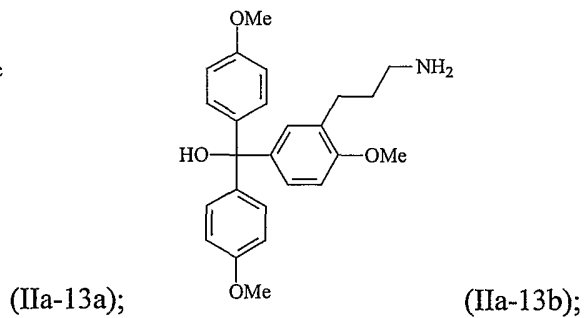
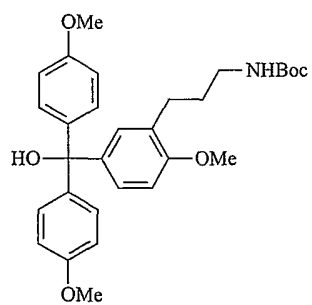
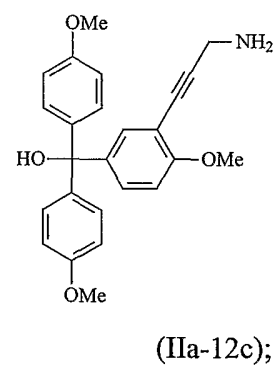
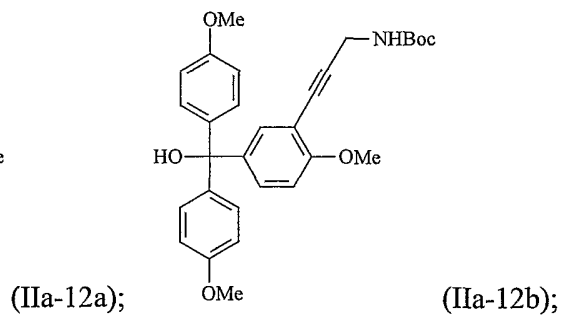
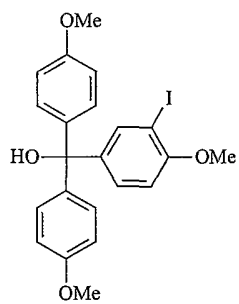
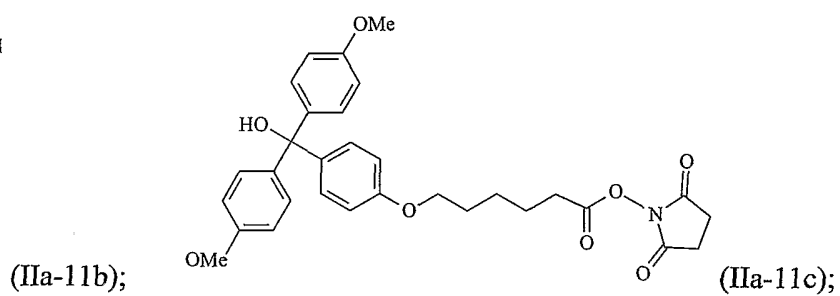
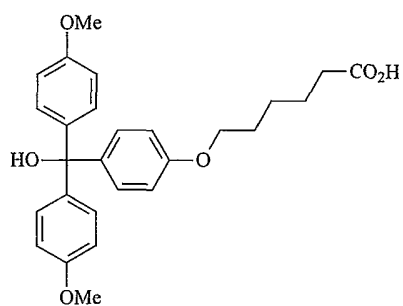
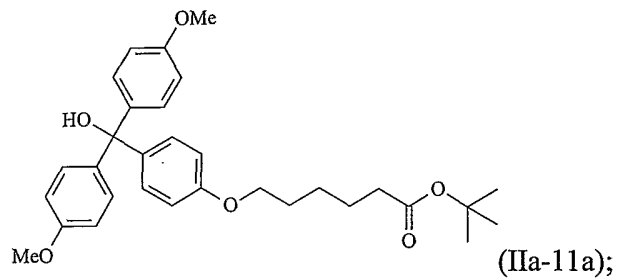
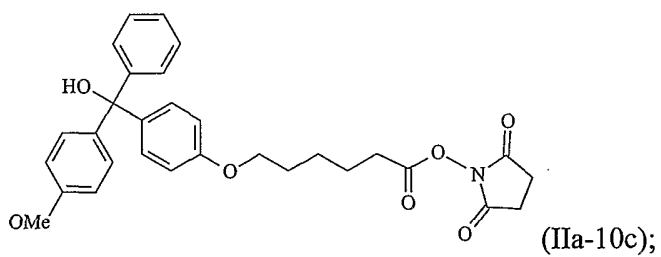
5

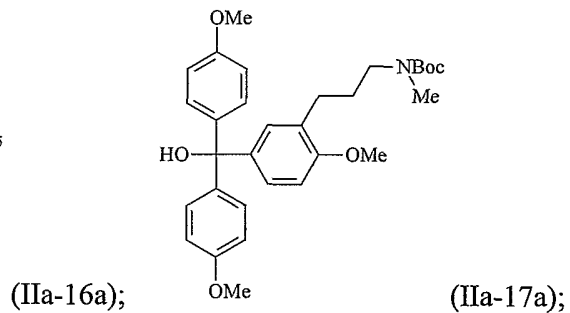
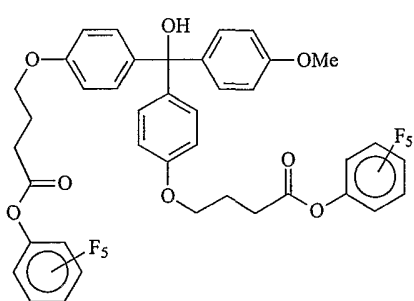
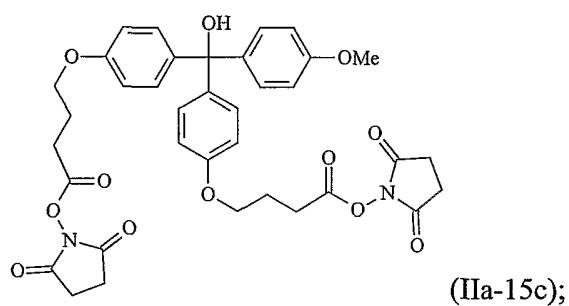
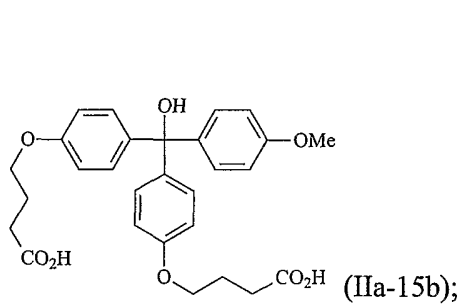
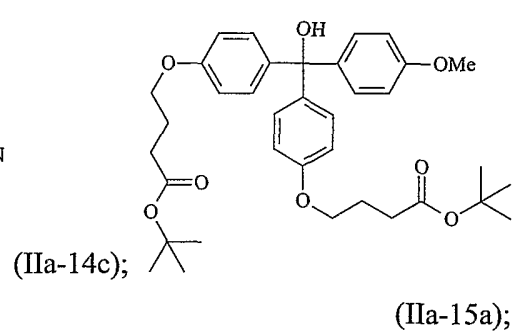
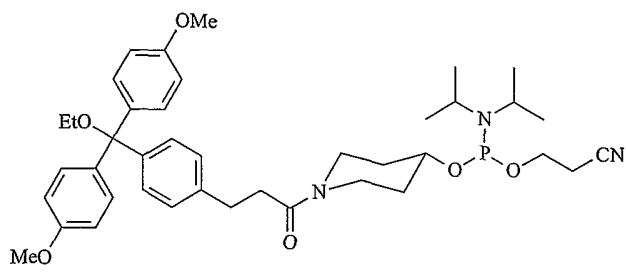
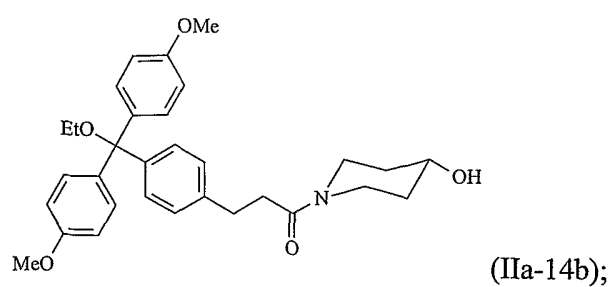
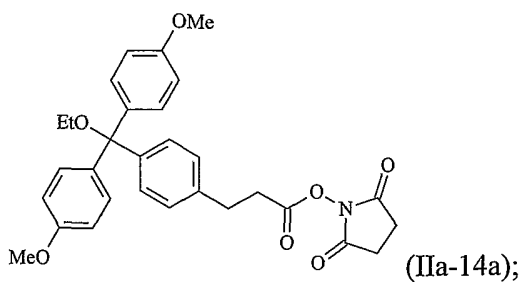


(IIa-10a);

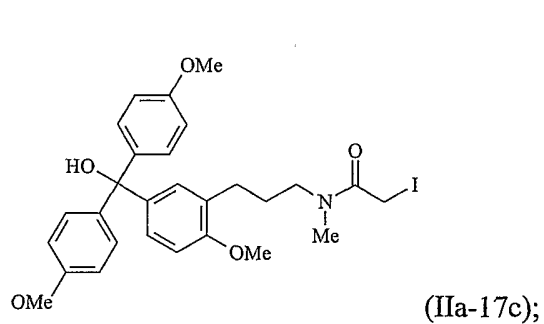
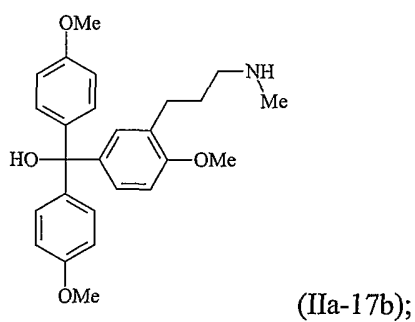


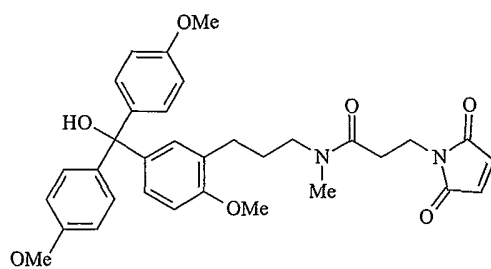
(IIa-10b);



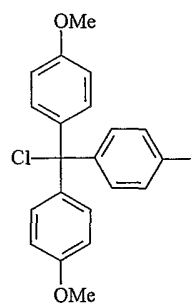


5

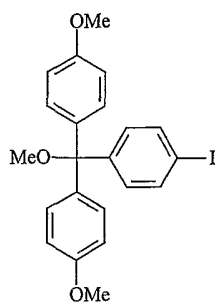




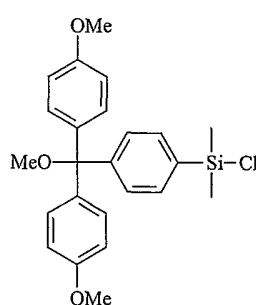
(IIa-18d);



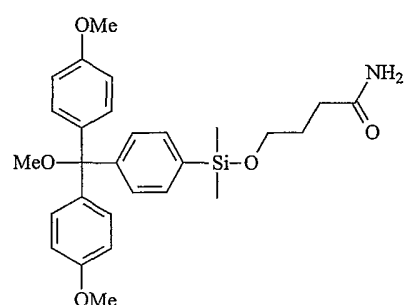
(IIa-19b);



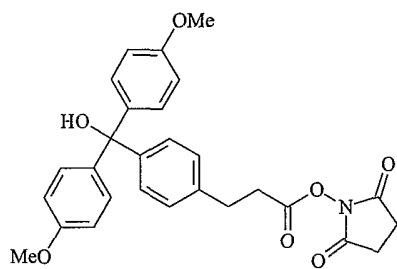
(IIa-19c);



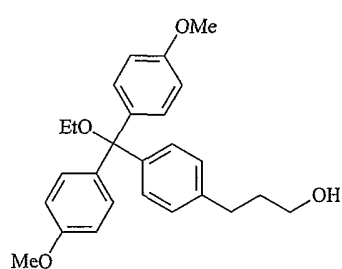
(IIa-19d);



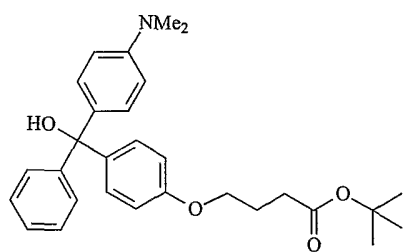
(IIa-19e);



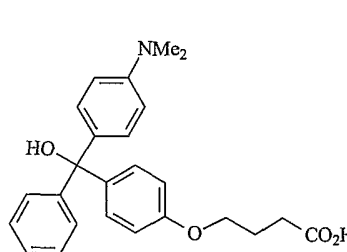
(IIa-20a);



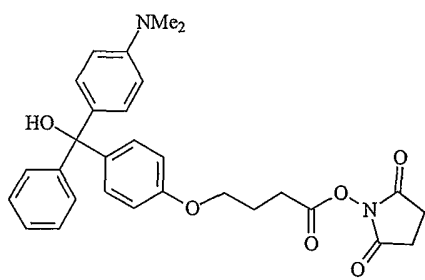
(IIa-20c);



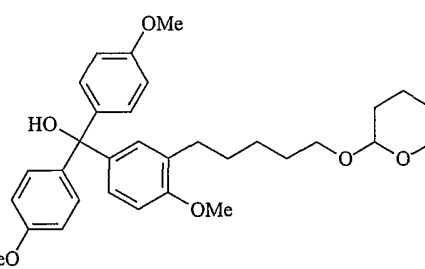
(IIa-24a);

^H (IIa-24b);

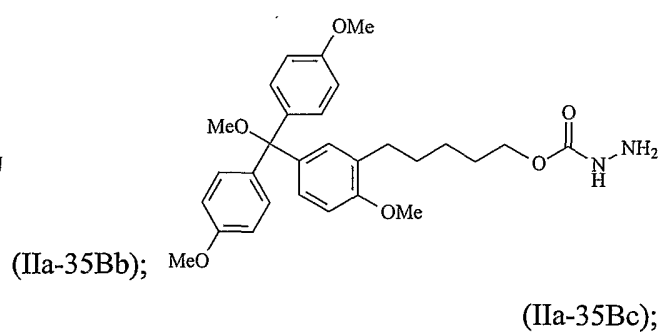
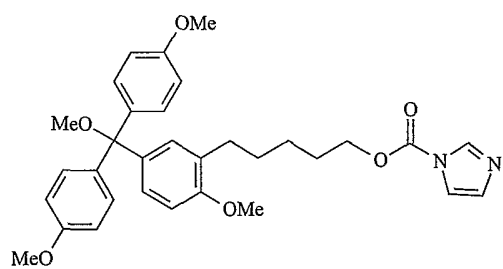
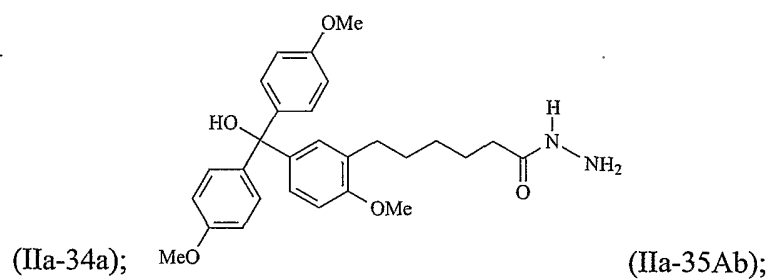
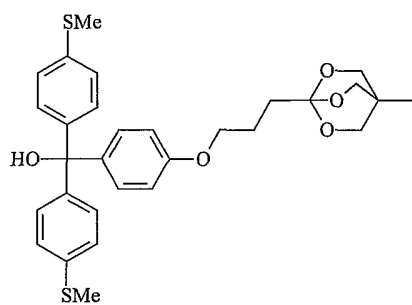
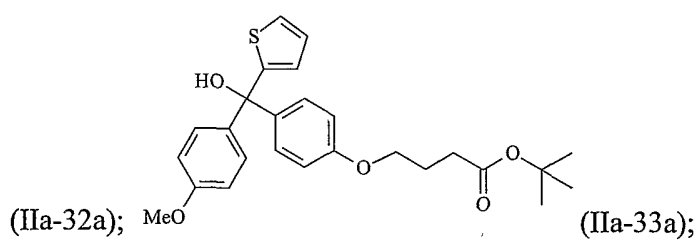
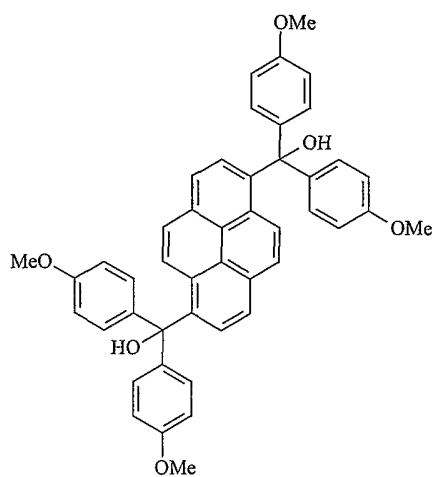
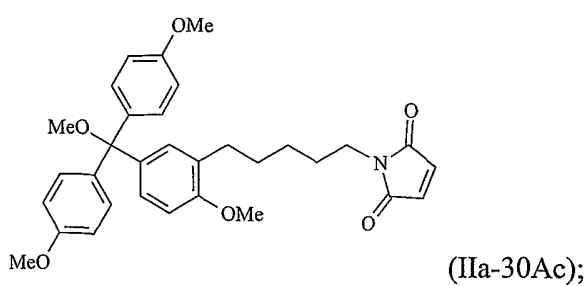
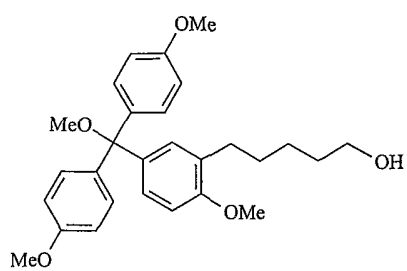
5

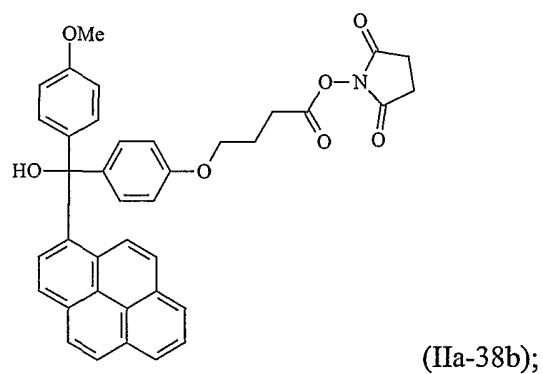
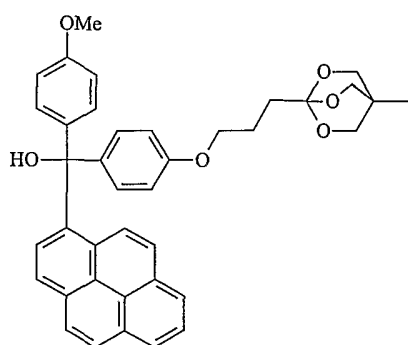
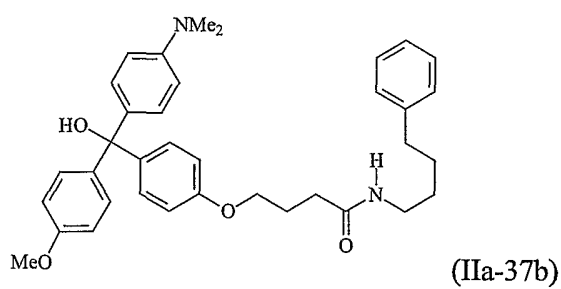
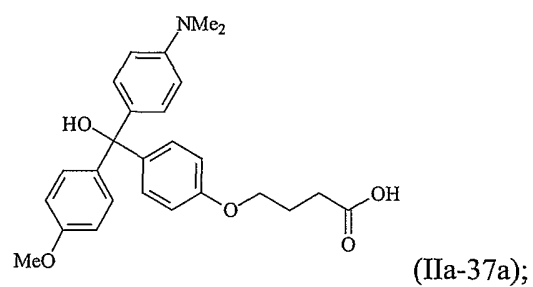
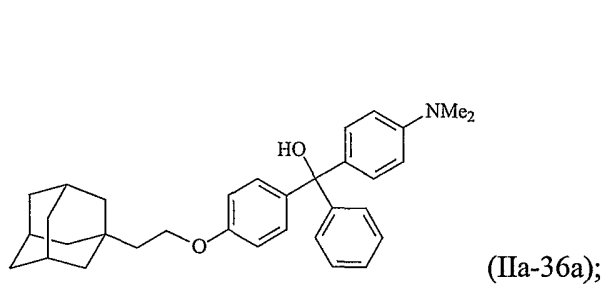
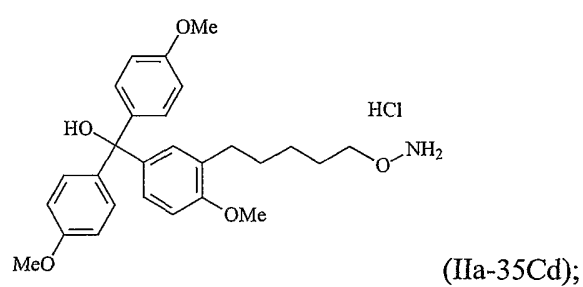
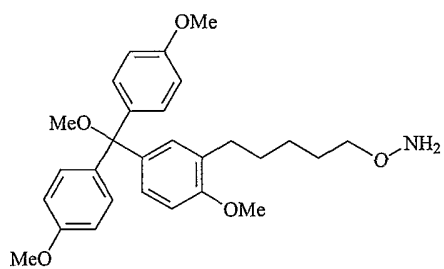
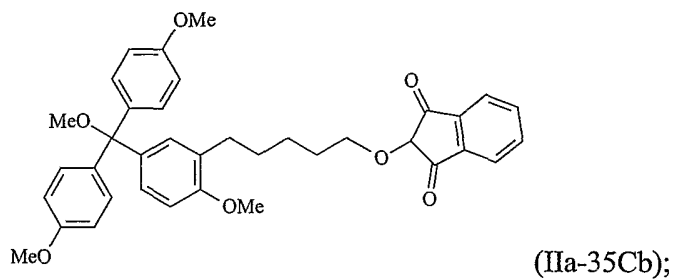
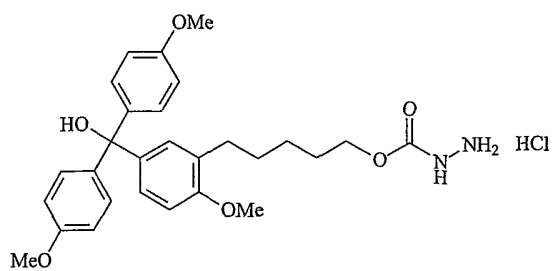


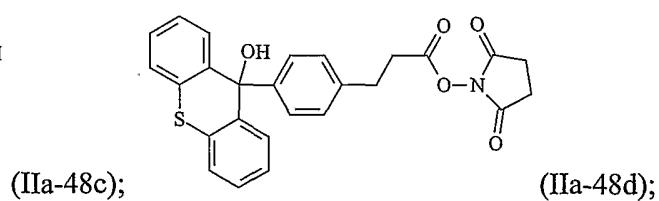
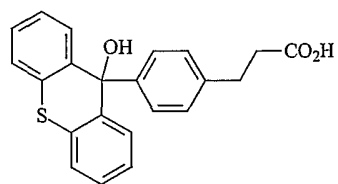
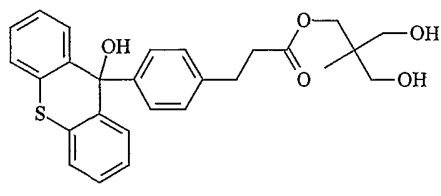
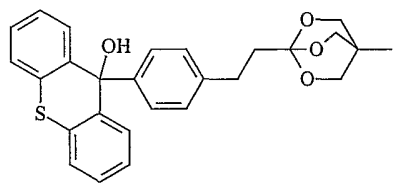
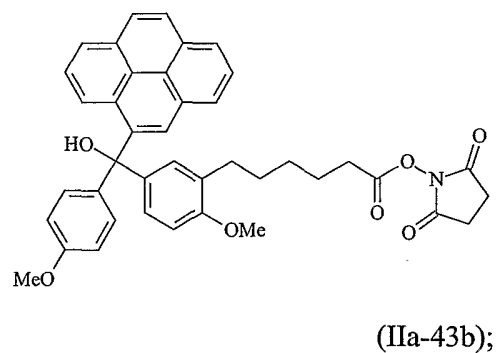
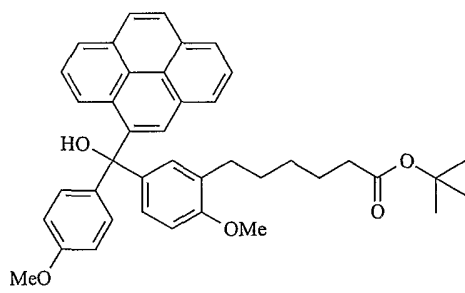
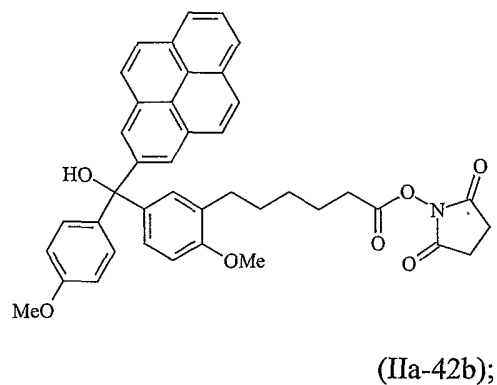
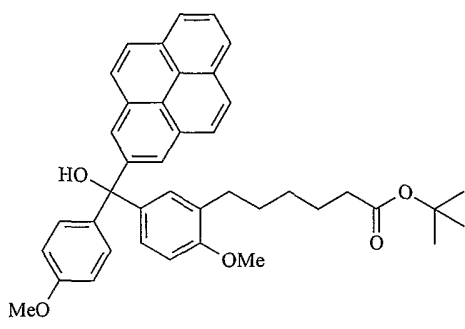
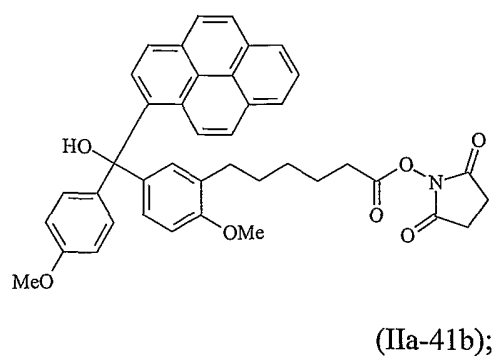
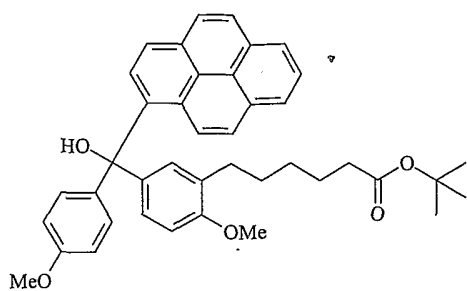
(IIa-24c);

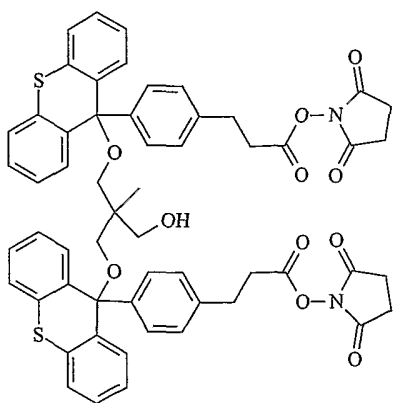


(IIa-30Aa);

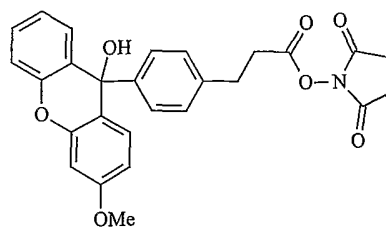




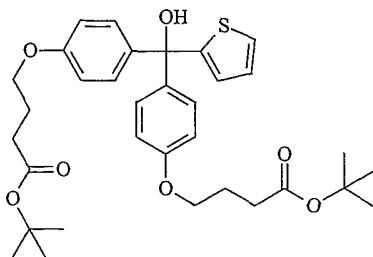




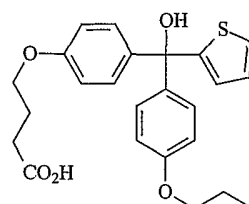
(IIa-48e);



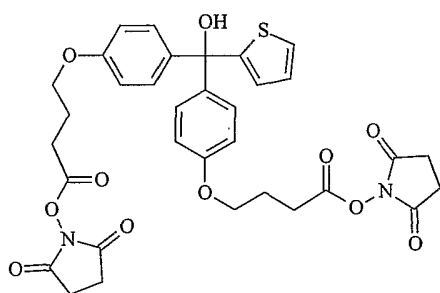
(IIa-49a);



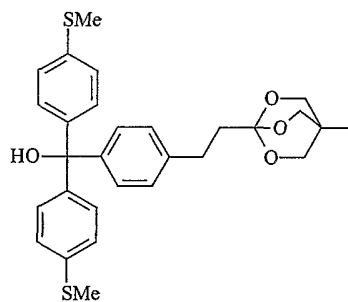
(IIa-50a);



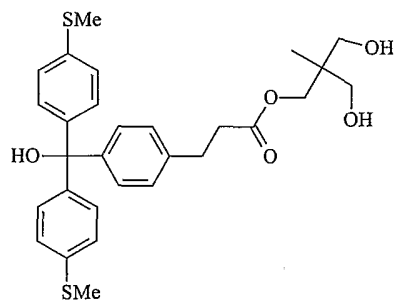
(IIa-51a);



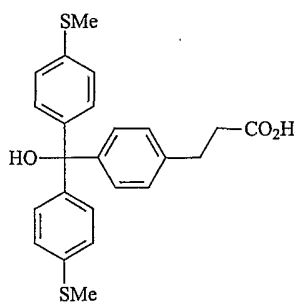
(IIa-51b);



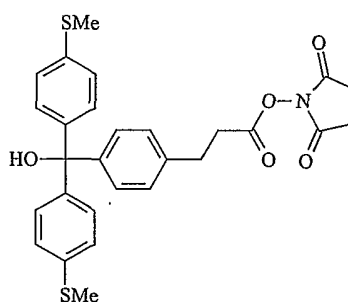
(IIa-53a);



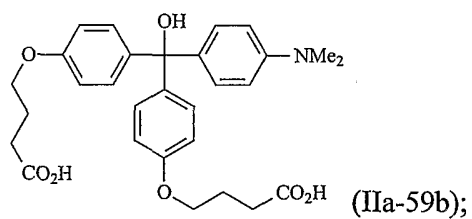
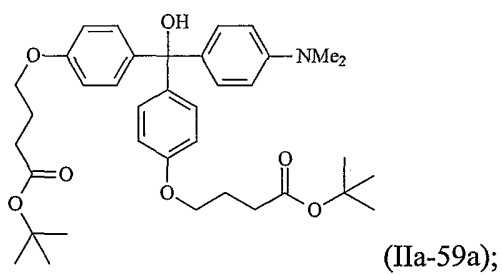
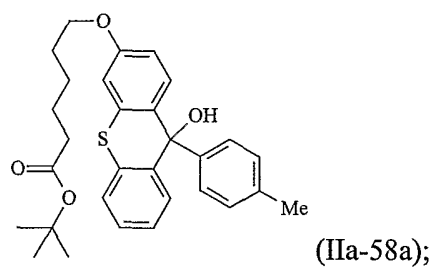
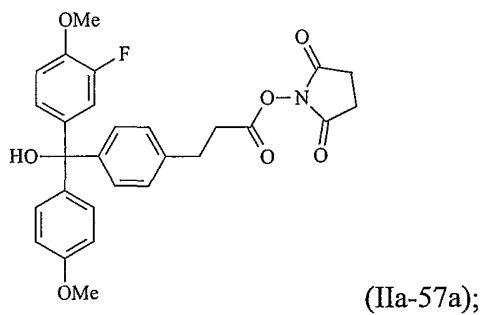
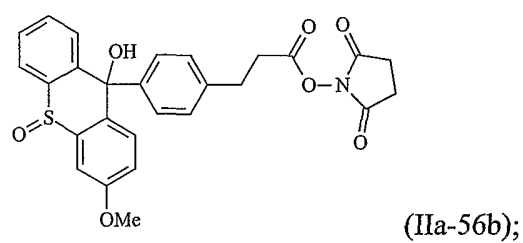
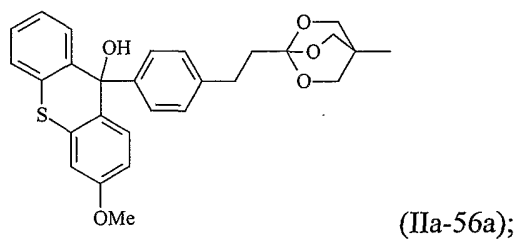
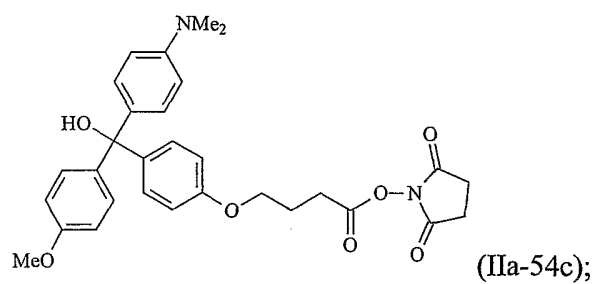
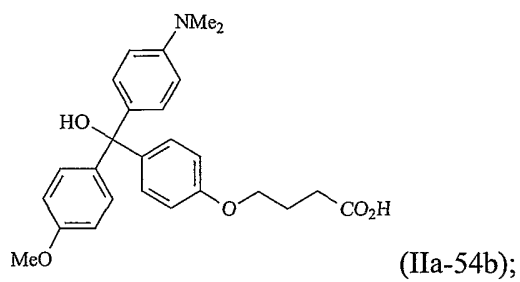
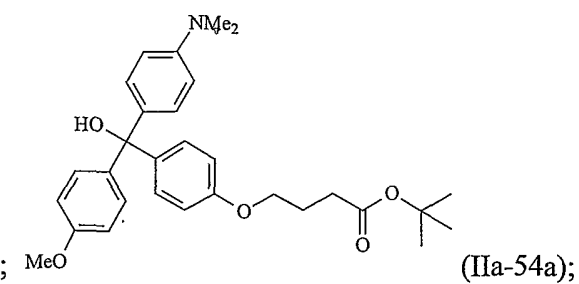
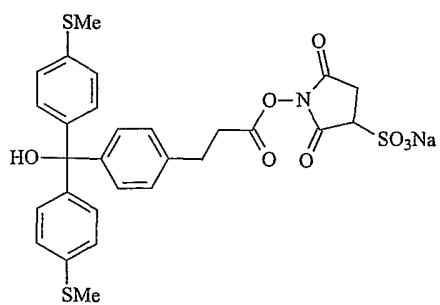
(IIa-53b);

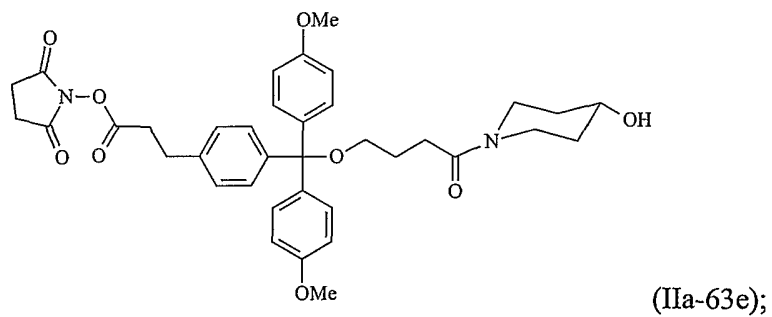
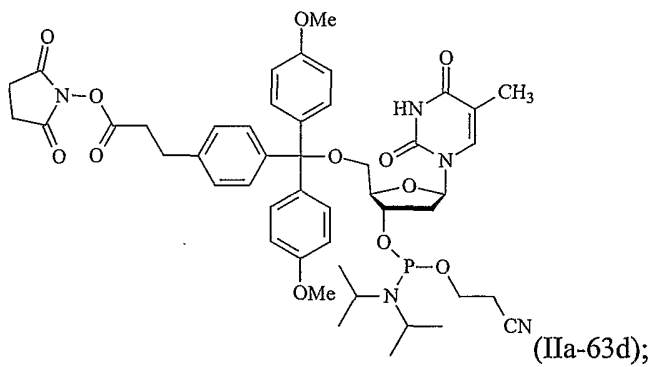
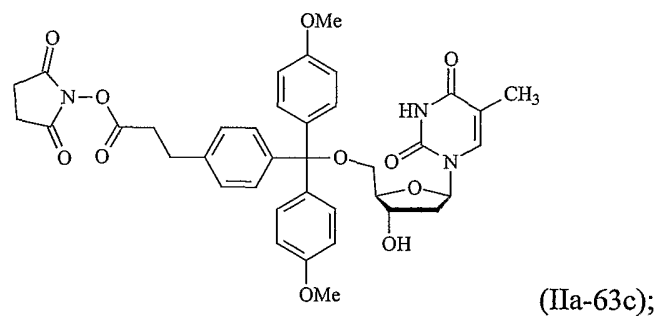
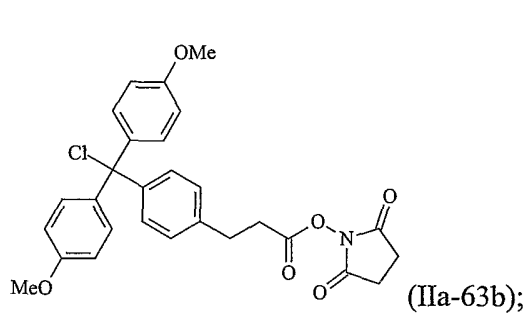
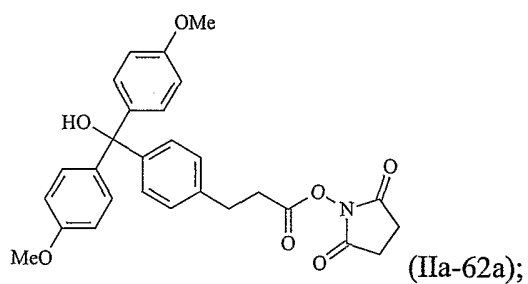
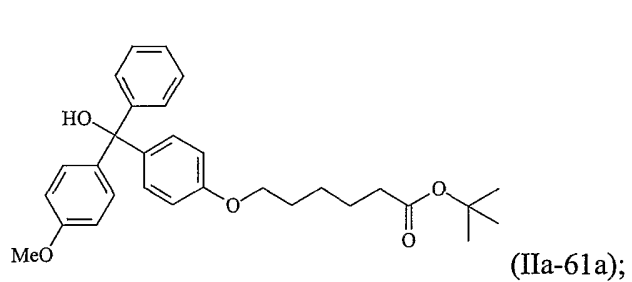
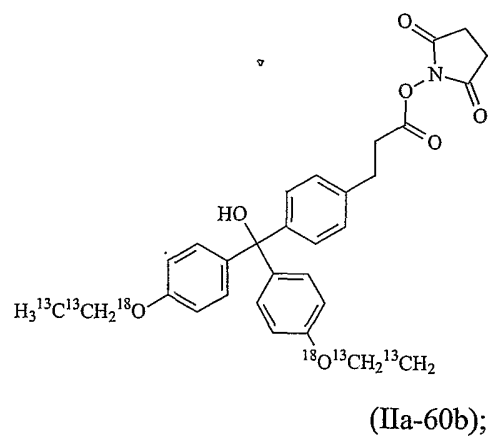
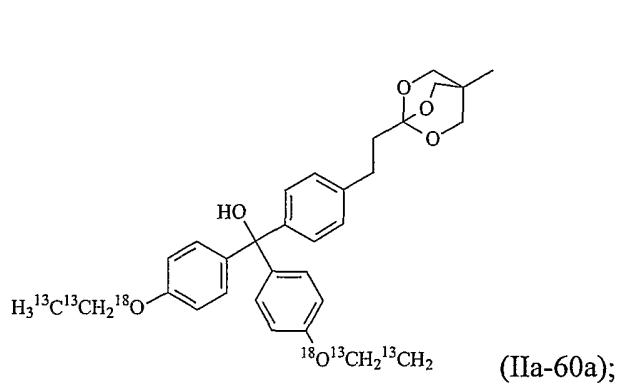


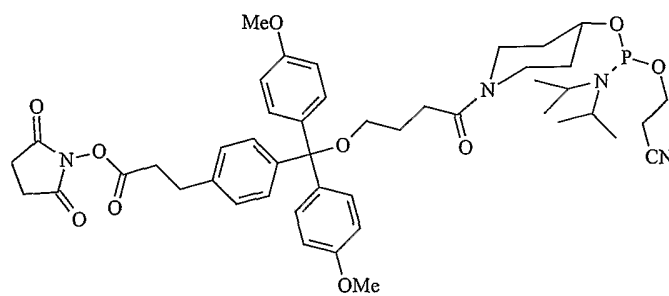
(IIa-53c);



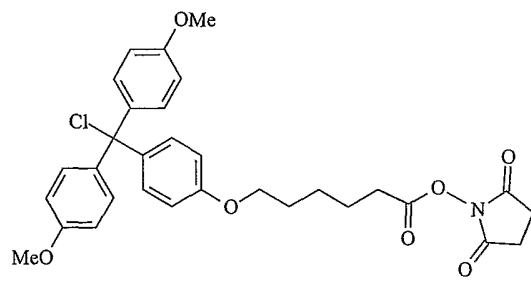
(IIa-53d);



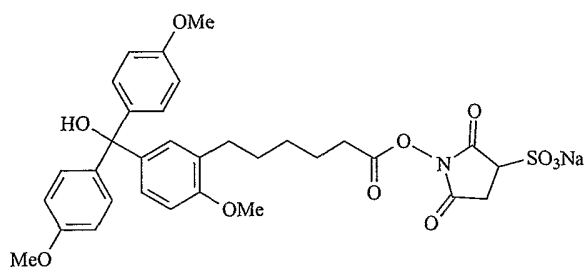




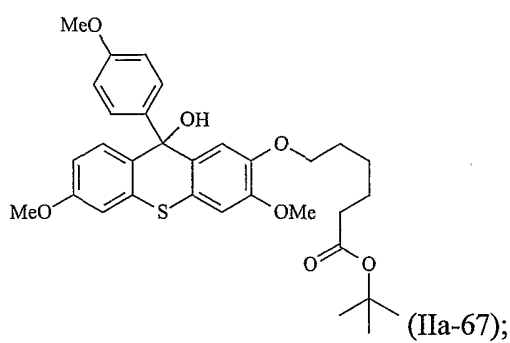
(IIa-63f);



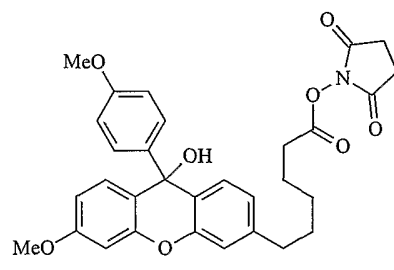
(IIa-64b);



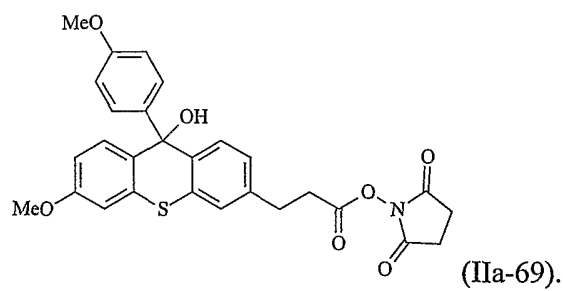
(IIa-66);



(IIa-67);

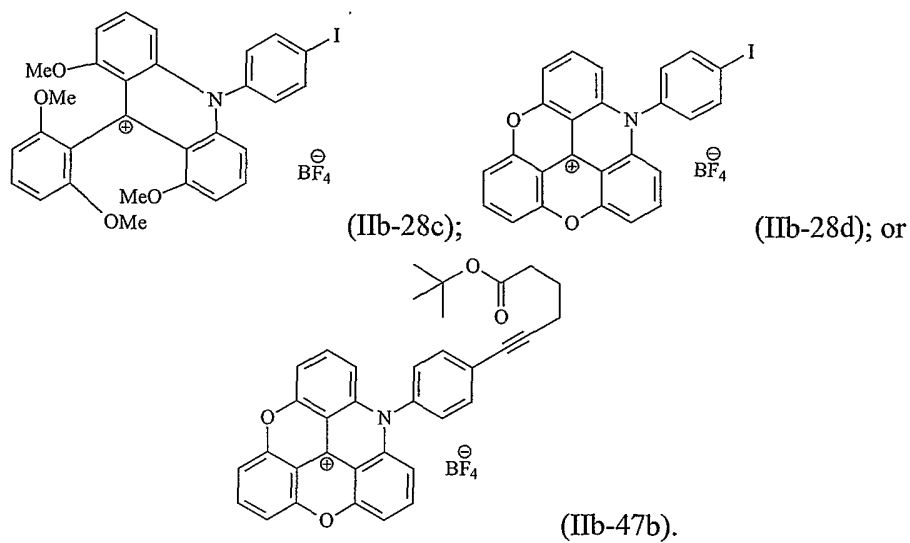


(IIa-68); or

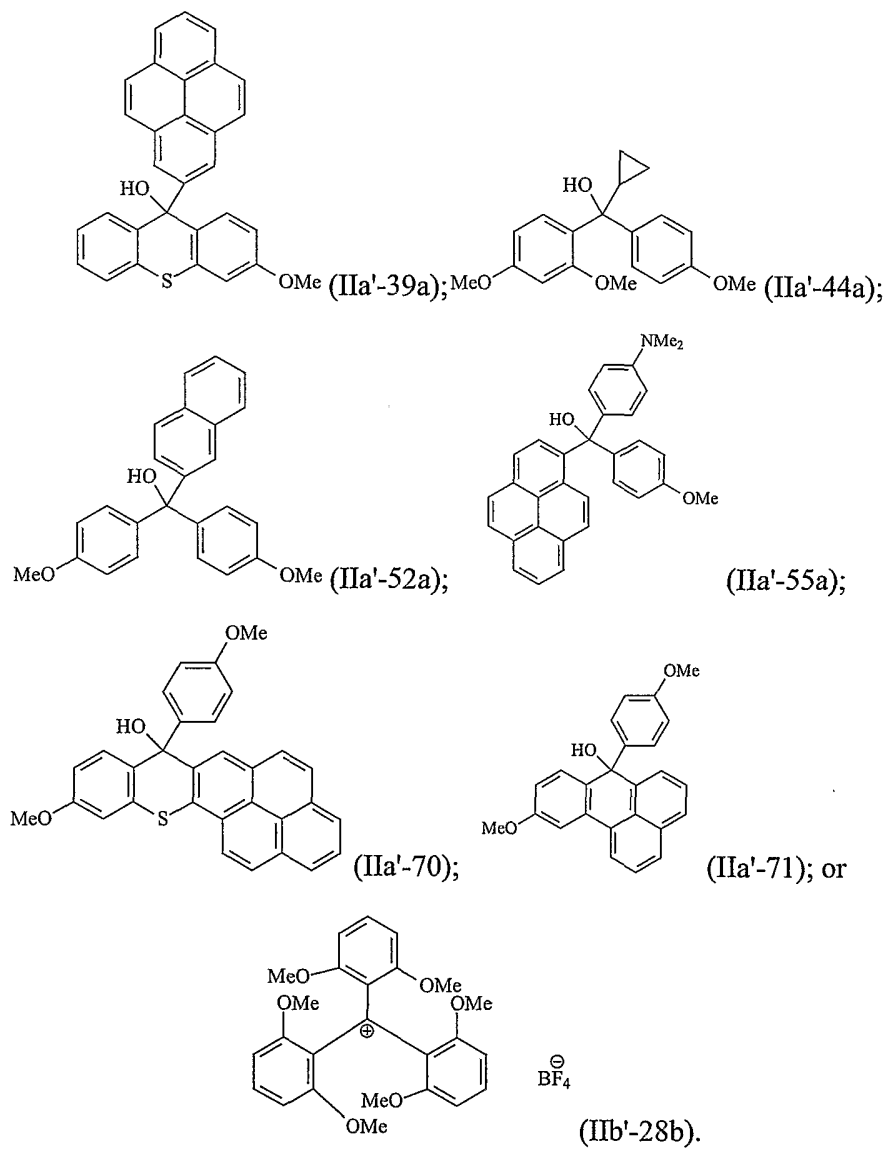


(IIa-69).

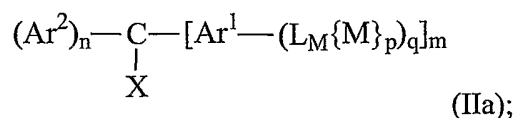
2. A compound of formula:



5 3. A compound of formula:

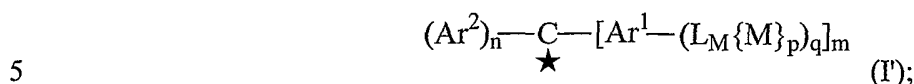


4. A compound of formula (IIa):



where:

X is a group capable of being cleaved from the a-carbon atom to form an ion of formula (I')



C★ is a carbon atom bearing a single positive charge or a single negative charge;

M is independently a reactive functional group;

Ar¹ is independently an aromatic group or an aromatic group substituted with one or more A;

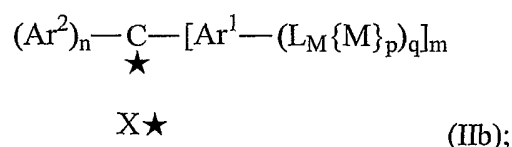
Ar² is independently an aromatic group or an aromatic group substituted with one or more A;

- 10 optionally wherein (a) two or three of the groups Ar¹ and Ar² are linked together by one or more L⁵, where L⁵ is independently a single bond or a linker atom or group; and/or (b) two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group substituted with one or more A;

A is independently a substituent;

- 15 L_M is independently a single bond or a linker atom or group;
n = 0, 1 or 2 and m = 1, 2, or 3, provided the sum of n+m = 3;
p independently = 1 or more; and
q independently = 1 or more.

- 20 5. A compound of formula (IIb):



where:

X★ is a counter-ion to C★;

C★ is a carbon atom bearing a single positive charge or a single negative charge;

- 25 M is independently a reactive functional group;

Ar¹ is independently an aromatic group or an aromatic group substituted with one or more A;

Ar² is independently an aromatic group or an aromatic group substituted with one or more A;

optionally wherein (a) two or three of the groups Ar¹ and Ar² are linked together by one or more L⁵, where L⁵ is independently a single bond or a linker atom or group; and/or (b)

two or three of the groups Ar^1 and Ar^2 together form an aromatic group or an aromatic group substituted with one or more A;

A is independently a substituent;

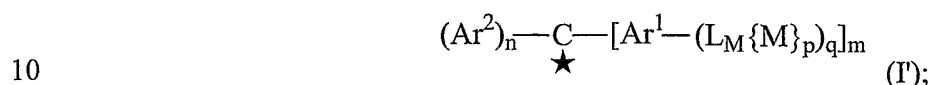
L_M is independently a single bond or a linker atom or group;

5 $n = 0, 1$ or 2 and $m = 1, 2$, or 3 , provided the sum of $n+m = 3$;

p independently = 1 or more; and

q independently = 1 or more.

6. An ion of formula (I'):



where:

$\text{C}\star$ is a carbon atom bearing a single positive charge or a single negative charge;

M is independently a reactive functional group;

Ar^1 is independently an aromatic group or an aromatic group substituted with one or more A;

15 Ar^2 is independently an aromatic group or an aromatic group substituted with one or more A;

optionally wherein (a) two or three of the groups Ar^1 and Ar^2 are linked together by one or more L^5 , where L^5 is independently a single bond or a linker atom or group; and/or (b) two or three of the groups Ar^1 and Ar^2 together form an aromatic group or an aromatic group substituted with one or more A;

20 A is independently a substituent;

L_M is independently a single bond or a linker atom or group;

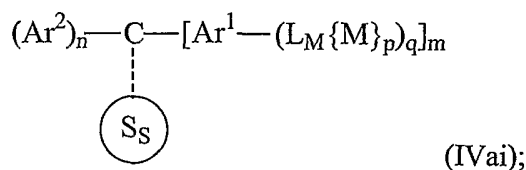
$n = 0, 1$ or 2 and $m = 1, 2$, or 3 , provided the sum of $n+m = 3$;

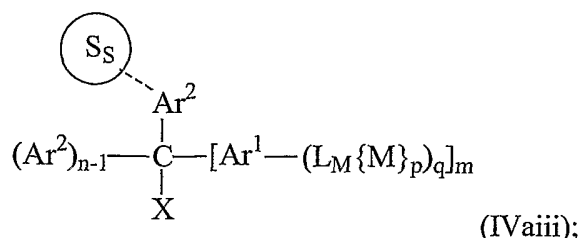
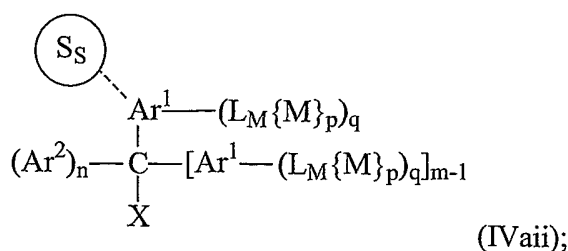
p independently = 1 or more; and

q independently = 1 or more.

25

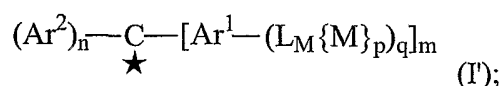
7. A solid support of formula (IVai), (IVaii) or (IVaiii):





where:

X is a group capable of being cleaved from the a-carbon atom of the compound of formula
 5 (II) to form an ion of formula (I')



C $\star$ is a carbon atom bearing a single positive charge or a single negative charge;

M is independently a reactive functional group;

Ar¹ is independently an aromatic group or an aromatic group substituted with one or more A;

10 Ar² is independently an aromatic group or an aromatic group substituted with one or more A;

optionally wherein (a) two or three of the groups Ar¹ and Ar² are linked together by one or more L⁵, where L⁵ is independently a single bond or a linker atom or group; and/or (b) two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group substituted with one or more A;

15 A is independently a substituent;

L_M is independently a single bond or a linker atom or group;

n = 0, 1 or 2 and m = 1, 2, or 3, provided the sum of n+m = 3;

p independently = 1 or more;

q independently = 1 or more;

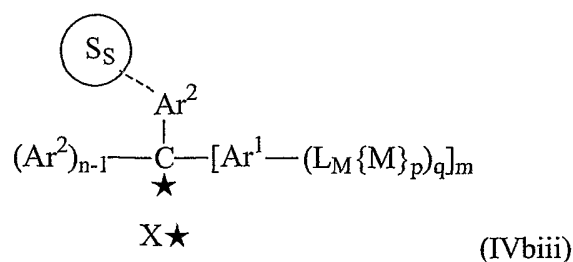
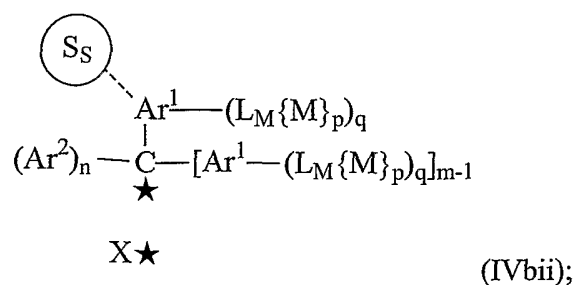
20 S_S is a solid support;

C--S_S comprises a cleavable bond between C and S_S;

S_S--Ar¹ comprises a cleavable bond between Ar¹ and S_S; and

S_S--Ar² comprises a cleavable bond between Ar² and S_S.

8. A solid support of formula (IVbii) or (IVbiii):



where: $\text{X}^{\star}$, Ar^1 , Ar^2 , L_M , M , n , m , p , q , S_S , $\text{C}^{\star} - \text{S}_\text{S}$, $\text{S}_\text{S} - \text{Ar}^1$ and $\text{S}_\text{S} - \text{Ar}^2$ are as defined above.

5 $\text{X}^{\star}$ is a counter-ion to $\text{C}^{\star}$;

$\text{C}^{\star}$ is a carbon atom bearing a single positive charge or a single negative charge;

M is independently a reactive functional group;

Ar^1 is independently an aromatic group or an aromatic group substituted with one or more A;

Ar^2 is independently an aromatic group or an aromatic group substituted with one or more A;

10 optionally wherein (a) two or three of the groups Ar^1 and Ar^2 are linked together by one or more L^5 , where L^5 is independently a single bond or a linker atom or group; and/or (b) two or three of the groups Ar^1 and Ar^2 together form an aromatic group or an aromatic group substituted with one or more A;

A is independently a substituent;

15 L_M is independently a single bond or a linker atom or group;

$\text{n} = 0, 1$ or 2 and $\text{m} = 1, 2$, or 3 , provided the sum of $\text{n} + \text{m} = 3$;

p independently = 1 or more;

q independently = 1 or more;

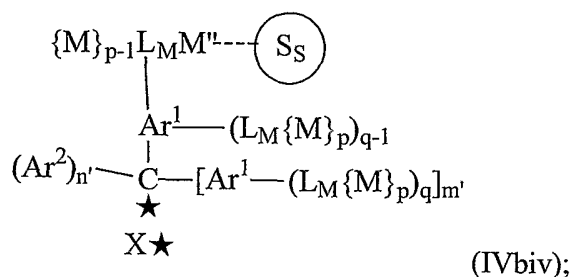
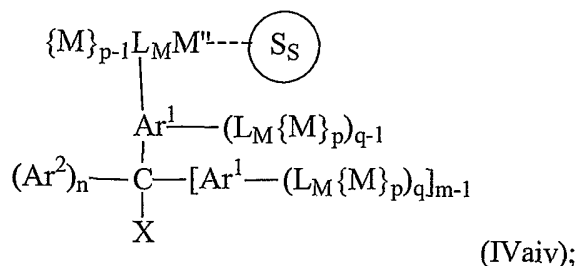
S_S is a solid support;

20 $\text{C}^{\star} - \text{S}_\text{S}$ comprises a cleavable bond between C and S_S ;

$\text{S}_\text{S} - \text{Ar}^1$ comprises a cleavable bond between Ar^1 and S_S ; and

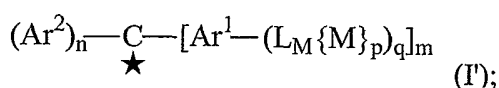
$\text{S}_\text{S} - \text{Ar}^2$ comprises a cleavable bond between Ar^2 and S_S .

9. A solid support of formula (IVaiv) or (IVbiv):



where:

- 5 X is a group capable of being cleaved from the α -carbon atom of the compound of formula (II) to form an ion of formula (I')



X $\star$ is a counter-ion to C $\star$;

C $\star$ is a carbon atom bearing a single positive charge or a single negative charge;

- 10 M is independently a reactive functional group;

Ar¹ is independently an aromatic group or an aromatic group substituted with one or more A;

Ar² is independently an aromatic group or an aromatic group substituted with one or more A;

optionally wherein (a) two or three of the groups Ar¹ and Ar² are linked together by one or more L⁵, where L⁵ is independently a single bond or a linker atom or group; and/or (b)

- 15 two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group substituted with one or more A;

A is independently a substituent;

L_M is independently a single bond or a linker atom or group;

n = 0, 1 or 2 and m = 1, 2, or 3, provided the sum of n+m = 3;

- 20 p independently = 1 or more;

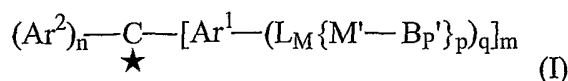
q independently = 1 or more;

S_S is a solid support;

M''--S_S comprises a bond between M'' and S_S; and

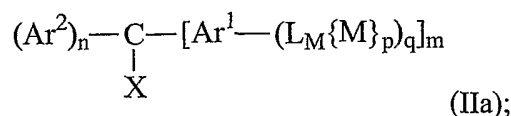
M'' is the same as M except that S₅ is bound to a portion of M which does not form part of the residue of M'' remaining attached to the ion of formula (I') which residue is produced after reaction of group M''.

- 5 10. A method of forming an ion of formula (I):

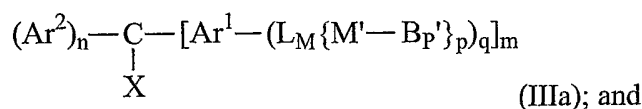


comprising the steps of:

- (i) reacting a compound of the formula (IIa):



- 10 with a biopolymer, B_P, having at least one group capable of reacting with M to form a covalent linkage, to provide a biopolymer derivative of the formula (IIIa):



- (ii) cleaving the C—X bond between X and the α-carbon atom of the derivative of formula (IIIa) to form the ion of formula (I);

- 15 where:

C★ is a carbon atom bearing a single positive charge or a single negative charge;

X is a group capable of being cleaved from the α-carbon atom to form an ion of formula (I);

M is independently a group capable of reacting with B_P to form the covalent linkage;

- 20 B_P' is independently the biopolymer residue of B_P produced on formation of the covalent linkage;

M' is independently the residue of M produced on formation of the covalent linkage;

Ar¹ is independently an aromatic group or an aromatic group substituted with one or more A;

Ar² is independently an aromatic group or an aromatic group substituted with one or more A;

- 25 optionally wherein (a) two or three of the groups Ar¹ and Ar² are linked together by one or more L⁵, where L⁵ is independently a single bond or a linker atom or group; and/or (b) two or three of the groups Ar¹ and Ar² together form an aromatic group or an aromatic group substituted with one or more A;

A is independently a substituent;

L_M is independently a single bond or a linker atom or group;

$n = 0, 1$ or 2 and $m = 1, 2$, or 3 , provided the sum of $n+m = 3$;

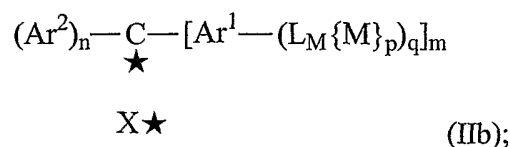
p independently = 1 or more; and

q independently = 1 or more.

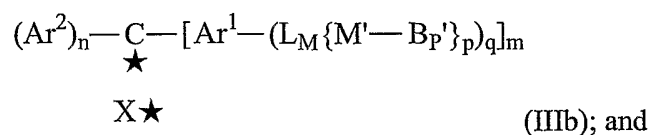
5 11. The method of claim 10 wherein the compound of formula (IIa) is a compound of claim 1.

12. A method of forming an ion of formula (I), comprising the steps of:

(i) reacting a compound of the formula (IIb):



10 with a biopolymer, B_p , having at least one group capable of reacting with M to form a covalent linkage, to provide a biopolymer derivative of the formula (IIIb):



dissociating $\text{X} \star$ from the derivative of formula (IIIb), to form the ion of formula (I);

where:

15 $\text{X} \star$ is a counter-ion to $\text{C} \star$;

and $\text{C} \star$, M , B_p , M' , Ar^1 , Ar^2 , L_M , n , m , p and q are as defined in claim 11.

13. The method of claim 12 wherein the compound of formula (IIb) is a compound of claim 2.

Figure 1A

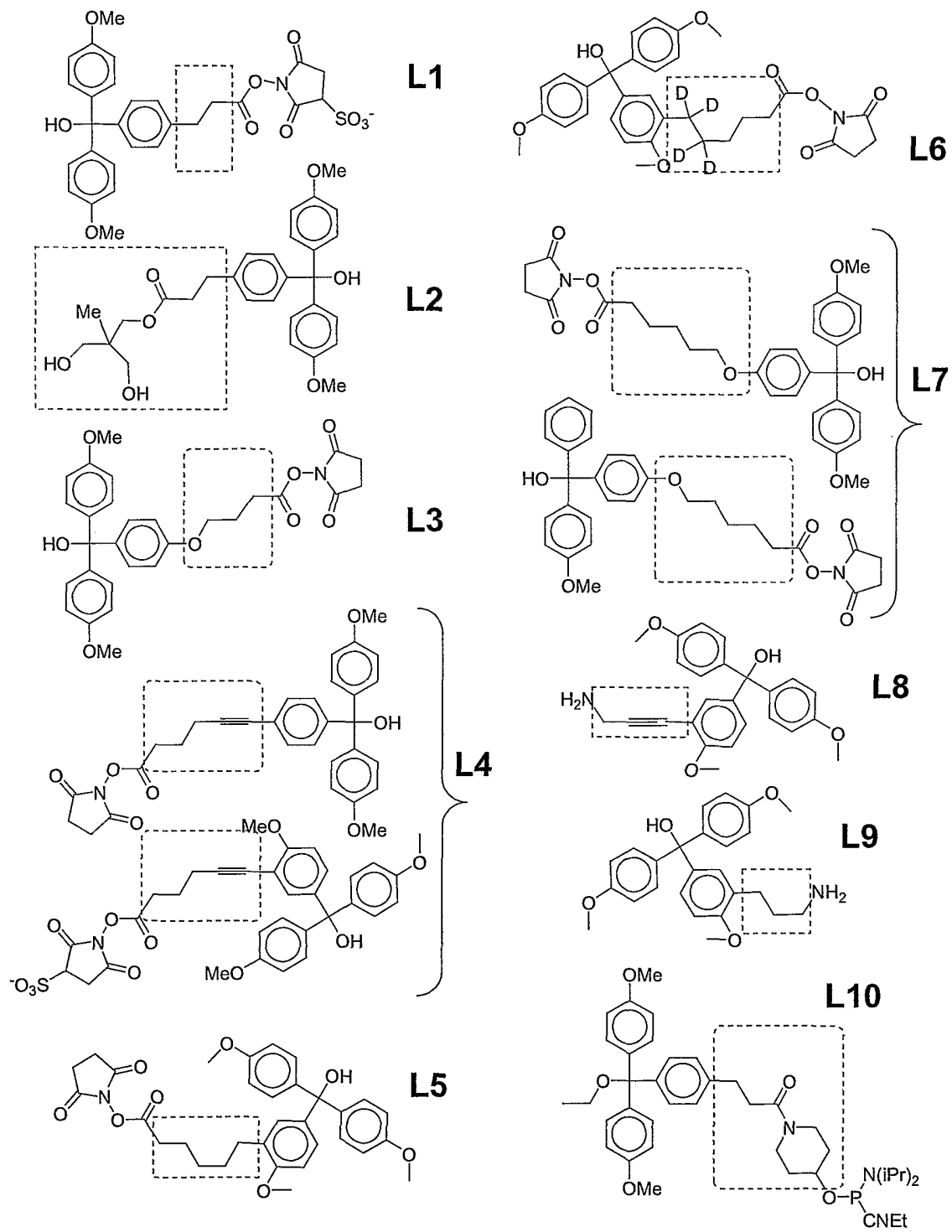


Figure 1B

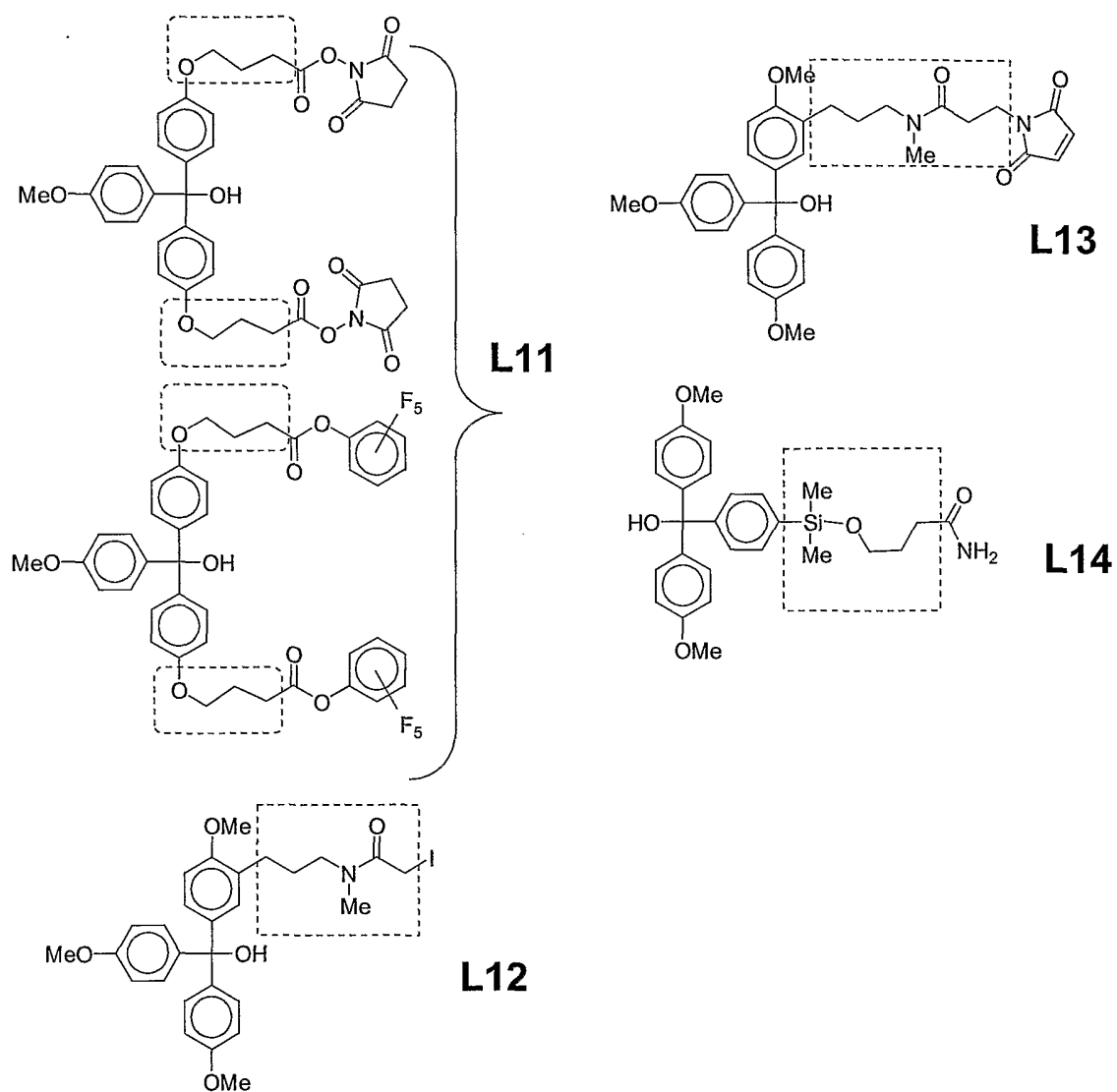


Figure 2A

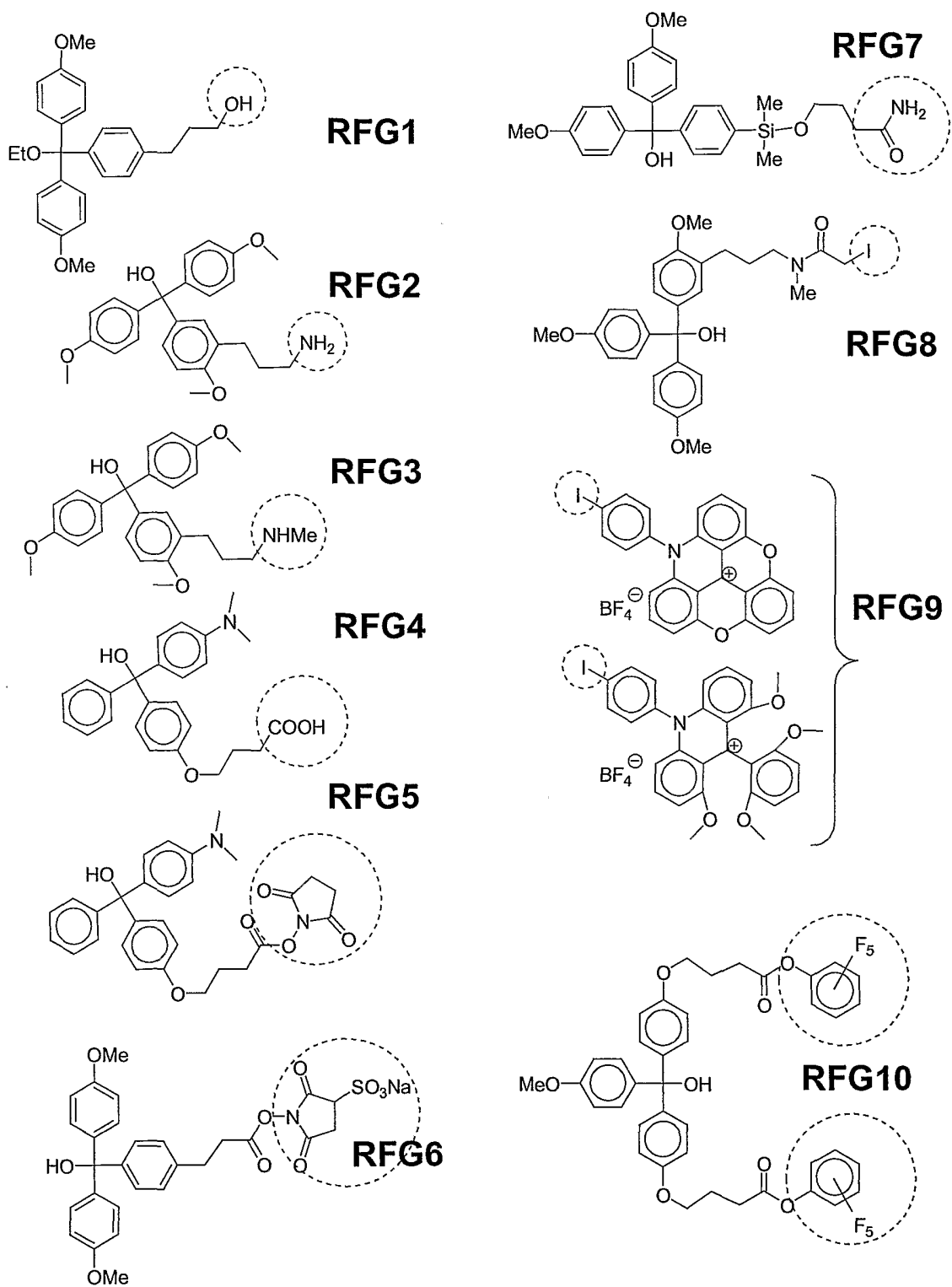


Figure 2B

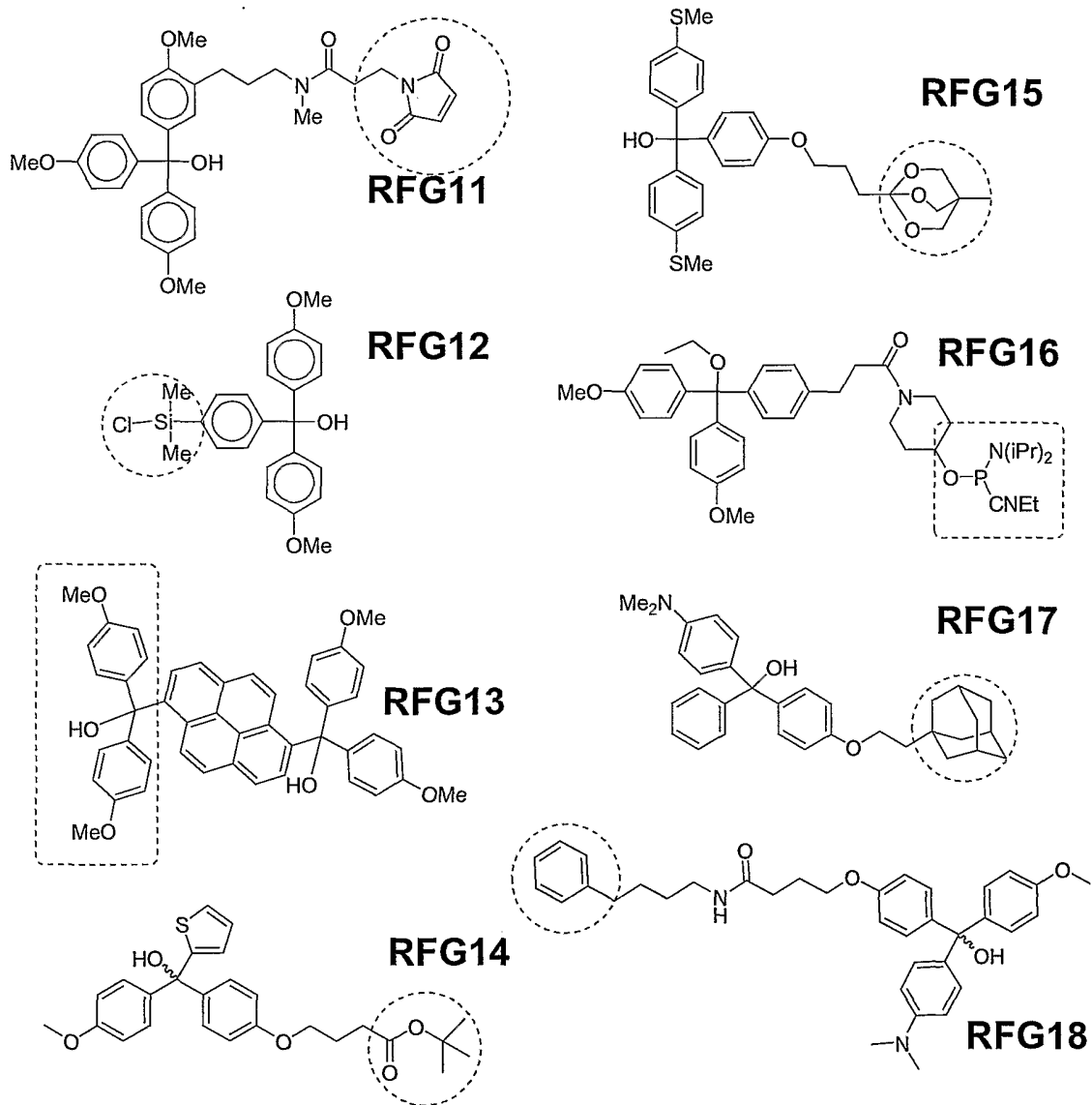
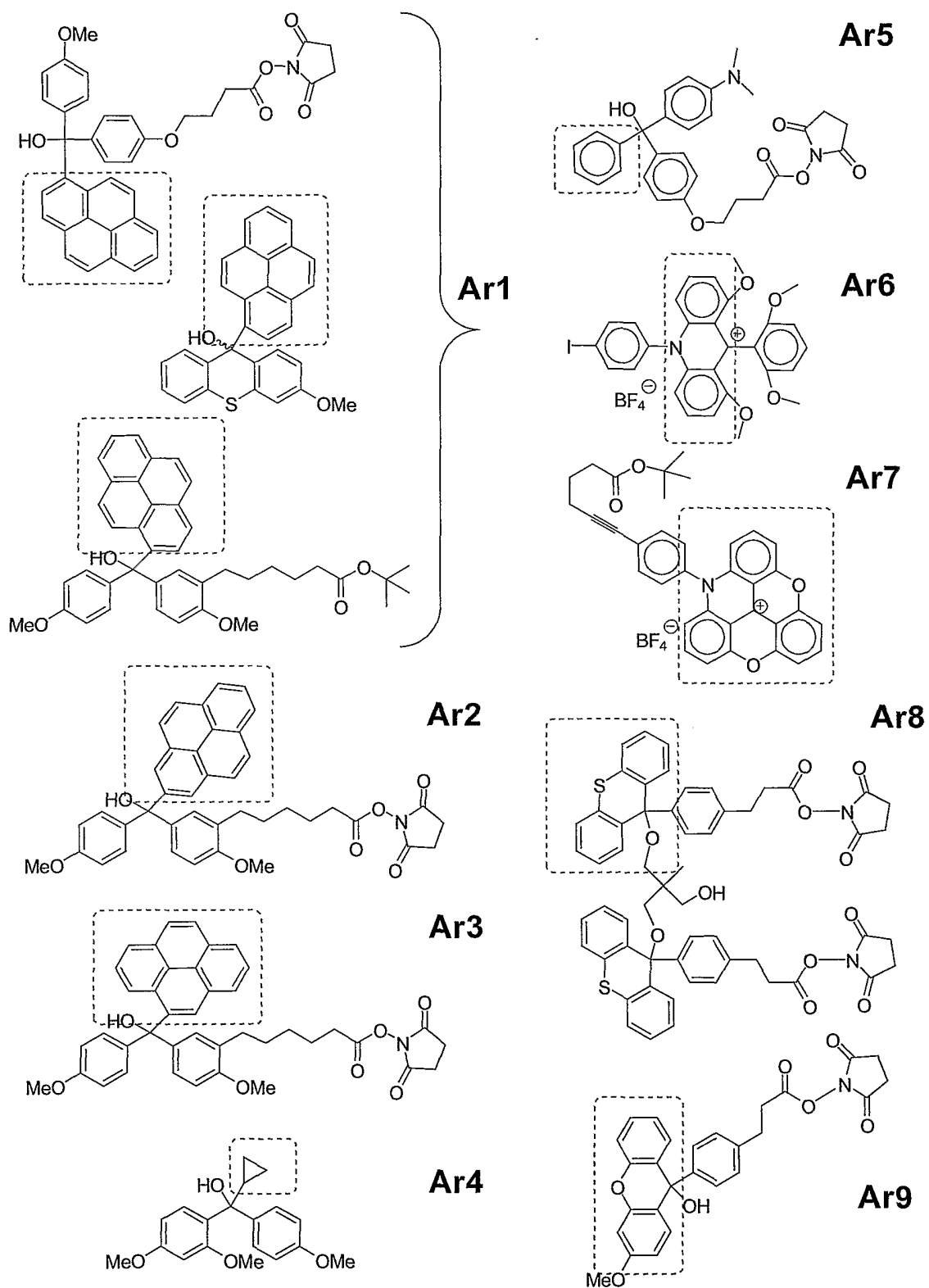
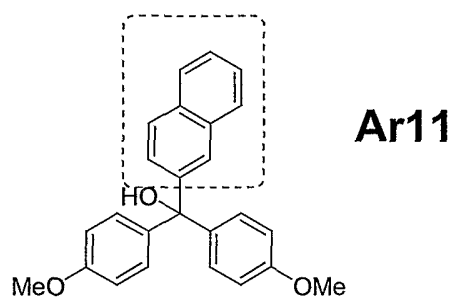
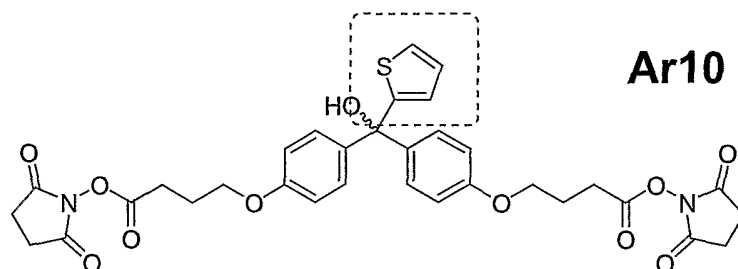


Figure 3A



6/8

Figure 3B

7/8

Figure 4

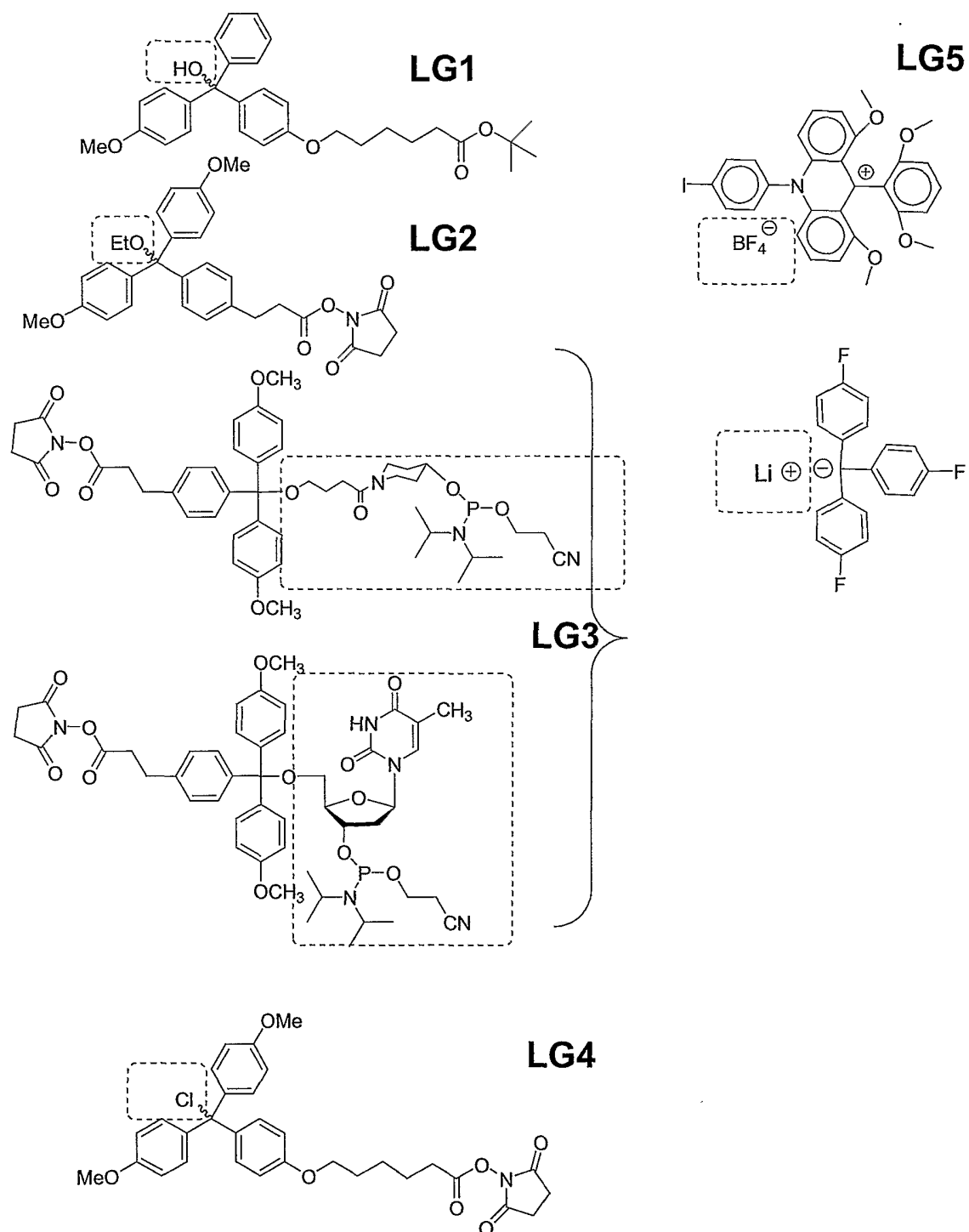


Figure 5

