



(51) International Patent Classification:

B01D 15/08 (2006.01) C07D 341/00 (2006.01)
C08G 75/0209 (2016.01) C07K 1/14 (2006.01)

(21) International Application Number:

PCT/EP2017/064022

(22) International Filing Date:

08 June 2017 (08.06.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

16305674.0 08 June 2016 (08.06.2016) EP

(71) Applicants: UCB BIOPHARMA SPRL [BE/BE]; 60 Allée de la Recherche, 1070 BRUXELLES (BE). UNIVERSITÉ CLAUDE BERNARD LYON 1 [FR/FR]; 43, boulevard du 11 novembre 1918, 69100 VILLEURBANNE (FR). CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE [FR/FR]; 3, rue Michel Ange, 75016 PARIS (FR). INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON [FR/FR]; 20 Avenue Albert Einstein, 69100 VILLEURBANNE (FR). CPE LYON FORMATION CONTINUE ET RECHERCHE - CPE LYON FCR SAS [FR/—]; 43 Boulevard du 11 Novembre 1918, VILLEURBANNE, 69100 (FR).

(72) Inventors: LECLAIRE, Julien; 53 impasse des Ifs, 69620 LE BOIS D'OINGT (FR). LECOMTE NORRANT, Edith; Avenue Franklin Roosevelt 156, 1050 BRUXELLES (BE). PERRET, Florent; 4 rue Pierre Corneille, 69006 LYON (FR). DUMARTIN, Melissa; 9 rue Édouard Vaillant, 69100 VILLEURBANNE (FR). OURRI, Benjamin; 36 rue Villon, 69008 LYON (FR).

(74) Agent: DOMENEGO, Bertrand et al.; Cabinet Lavoix, 2, place d'Estienne d'Orves, 75441 PARIS CEDEX 09 (FR).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,

TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: METHODS FOR IDENTIFYING PROTEINS BY USING SYNTHETIC RECEPTORS

(57) Abstract: The invention relates to oligomeric macrocycles and to uses thereof as receptors for recognition of protein post-translational modifications (PTM) or specific motifs in proteins.

METHODS FOR IDENTIFYING PROTEINS BY USING SYNTHETIC RECEPTORS

The invention relates to oligomeric macrocycles and to uses thereof as receptors for recognition of protein post-translational modifications (PTM) or specific motifs in proteins.

Proteins are complex organic molecules susceptible to numerous post-translational modifications occurring spontaneously *in vivo*, during aging or as a consequence of physiologic or pathologic processes, thereby contributing to the diversity of the proteome. *In vitro*, post-translational modification of proteins also occurs due to e.g. enzymatic processing, oxidation or hydrolysis.

Protein oxidation is a covalent modification of a protein induced either directly by reactive oxygen species or indirectly by reaction with secondary by-products of oxidative stress. Oxidative changes to proteins can lead to diverse functional consequences, such as inhibition of enzymatic and binding activities, increased susceptibility to aggregation and proteolysis, increased or decreased uptake by cells, and altered immunogenicity (Shacter, Drug Metabolism Reviews, 32(3&4), 307–326 (2000)). Oxidative modifications in synthetic or recombinant proteins occur mainly during production and storage. In particular, methionine side chains are prone to oxidative modification, thereby producing Met(O) residues.

Deamidation is a non-enzymatic (hydrolytic) post-translational modification of asparagine to aspartic acid or isoaspartic acid, or glutamine to glutamic acid. Deamidation notably consists in replacement of an amine (–NH₂) group with a hydroxyl (–OH group) via hydrolysis. Deamidation occurs in solution, possibly at all Asn and Gln residues, and in all proteins eventually. Deamidation rates vary dramatically with the primary and higher-order structure of proteins, and results in a mixture of protein variants that are relatively difficult to differentiate.

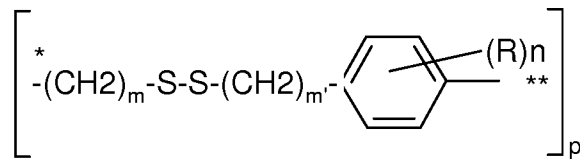
Oxidative and hydrolytic modifications of amino acids in proteins being associated with loss of biological activity, means for detecting oxidized and/or hydrolyzed forms of a protein, quantifying oxidized and/or hydrolyzed forms of a protein, or for collecting oxidized and/or hydrolyzed forms of a protein (by selective binding) are desired, in particular when the protein is a therapeutic protein.

Vial et al (Chem. Commun. 2016, DOI: 10.1039/c6cc07713g (accepted 15th November 2016) report that the asymmetric deformation of a dyn[4]arene upon the binding of various lysine derivatives leads to distinct induced circular dichroism outputs.

Summary of the invention

The present invention concerns method for differentiating native form from oxidized and/or hydrolyzed forms of a protein comprising:

a. contacting a composition likely to comprise native, oxidized and/or hydrolyzed forms of a protein with a oligomeric macrocycle of general formula (A):



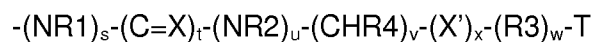
Where

m and m' identical or different are an integer independently chosen from 0 or 1;

n is an integer comprised between 1 and 4;

p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 50;

each R identical or different may be represented by the following formula:



where T is a terminal group and may be chosen from the groups consisting in H, N₃, alkyl, heteroaryl, a solid phase support;

s, t, u, v, w identical or different may be independently 0 or 1;

R₁ represents H or -alkyl optionally substituted by one or more of halogen atoms, OH, OR₁₁, NR₁₁R₁₂, CN;

X represents O, NH or S;

X' is O, NH or S;

R₂ represents H, -alkyl optionally substituted by one or more of halogen atoms, OH, OR₁₁, NR₁₁R₁₂, CN;

R₃ represents -alkyl-, -heteroaryl- or -aryl- optionally substituted by one or more of -R₁₁ or -OR₁₁, -N=C(NR₁₁)₂-NR₁₁-, -(OCH₂CH₂)_q-(CH₂CH₂O)_q-, -alkyl-NH-C(=O)-NH-, -aryl-NH-C(=O)-NH-, -alkenyl-OCO-NH-, -alkyl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -aryl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -NR₁₁-, -N(R₁₁)-C(=NH)-NH-, -O-, -C(O)O-, -N(R₁₃)-(CH₂)_x-NR₁₁-, -SO₂-aryl-;

R₄ is H or COOH;

R11, R12 identical or different are independently selected from the group consisting in H, alkyl optionally substituted by one or more of halogen atoms, OH, CN;

R13 is -alkyl optionally substituted by NR11R12;

q is an integer comprised between 1 and 200;

x is 0 or 1;

where alkyl is C1-C6 alkyl; aryl is a mono or bicyclic C6-C10 aromatic ring system; heteroaryl is a 5 to 10 membered mono or bicyclic aromatic ring system comprising one to 4 heteroatoms chosen from N, O or S.

It being understood that R may be different for each unit and/or for each n;

whereby the cyclic form is achieved by the bond between the S (*) of the first unit and the carbon of the phenyl group (**) of the last unit;

as well as its various stereoisomers;

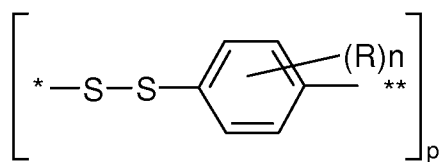
and

b. detecting or collecting complexes formed with the oligomeric macrocycle of general formula (A), wherein said complexes preferentially comprise native form of the protein or oxidized and/or hydrolyzed forms of the protein.

According to an embodiment, $m=m'=0$.

According to an embodiment, the method comprises

a. contacting a composition likely to comprise native, oxidized and/or hydrolyzed forms of a protein with a oligomeric macrocycle of general formula (I):



(I)

where n is an integer comprised between 1 and 4;

p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 50;

each R identical or different may be represented by the following formula:



where T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl, a solid phase support;

s, t, u, v, w identical or different may be independently 0 or 1;

R1 represents H or -alkyl optionally substituted by one or more of halogen atoms, OH, OR11, NR11R12, CN;

X represents O, NH or S;

R2 represents H, -alkyl optionally substituted by one or more of halogen atoms, OH, OR11, NR11R12, CN;

R3 represents -alkyl-, -heteroaryl- or -aryl- optionally substituted by -R11 or -OR11, -N=C(NR11₂)-NR11-, -(CH₂CH₂O)_q-, -alkyl-NH-C(=O)-NH-, -aryl-NH-C(=O)-NH-, -alkenyl-OCO-NH-, -alkyl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -aryl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -NR11-, -N(R11)-C(=NH)-NH-, -O-, -N(R13)-(CH₂)_x-NR11-, -SO₂-aryl;

R11, R12 identical or different are independently selected from the group consisting in H, alkyl optionally substituted by one or more of halogen atoms, OH, CN;

R13 is -alkyl optionally substituted by NR11R12;

q is an integer comprised between 1 and 200;

where alkyl is C1-C6 alkyl; aryl is a mono or bicyclic C6-C10 aromatic ring system; heteroaryl is a 5 to 10 membered mono or bicyclic aromatic ring system comprising one to 4 heteroatoms chosen from N, O or S.

It being understood that R may be different for each unit and/or for each n;

whereby the cyclic form is achieved by the bond between the S (*) of the first unit and the carbon of the phenyl group (**) of the last unit;

as well as its various stereoisomers;

and

- b. detecting or collecting complexes formed with the oligomeric macrocycle of general formula (I), wherein said complexes preferentially comprise native form of the protein or oxidized and/or hydrolyzed forms of the protein.

The present invention also concerns a method of purification of a protein which comprises:

- a) Contacting a composition comprising a target protein with a oligomeric macrocycle of formula (A) as defined above

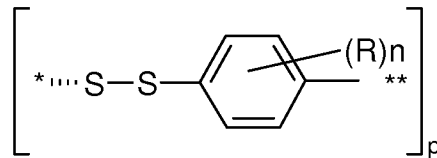
and

b) Recovering the target protein bound to the oligomeric macrocycle of general formula (A) as defined above

wherein the target protein or peptide comprises a lysine at position 1 or 2 of its amino acid sequence.

5 According to an embodiment, the method of purification comprises

a) Contacting a composition comprising a target protein with a oligomeric macrocycle of formula (I''):



(I'')

10 where :

n is 2;

p is 4;

R is $-(\text{NR}1)_s-(\text{C}=\text{X})_t-(\text{NR}2)_u-(\text{CH}_2)_v-(\text{R}3)_w-\text{T}$

Where s is 0:

15 t is 1 and X is O;

u is 0 or 1 and R2 is H, v is 0, 1 or 2;

w is 0 or 1 with R3 is chosen from -heteroaryl- or -aryl- optionally substituted by -R11 or -OR11, -N=C(NR11₂)-NR11-, -O-, -NR11-, -N(R11)-C(=NH)-NH-, -N(R13)-(CH₂)_x-NR11-; with R11 is H or alkyl;

20 T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl.

and

b) Recovering the target protein bound to the oligomeric macrocycle of general formula (I'') as defined above

25

wherein the target protein or peptide comprises a lysine at position 1 or 2 of its amino acid sequence.

The present inventions also provides for a compound of formula (A) or (I)

30

As defined above,

with the exception of the compounds where

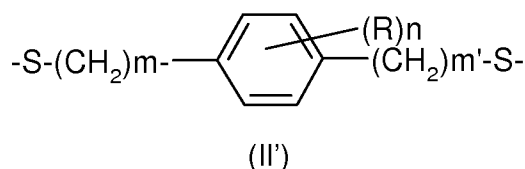
$m=m'=0$ and

$p=4$, $n=2$ and $R=\text{COOH}$, or

$p=4$, $n=2$ and $R=\text{NH}_2$,

5 It is also provided a process of preparation of the compound of formula (A) such as (I) of the invention as defined above,

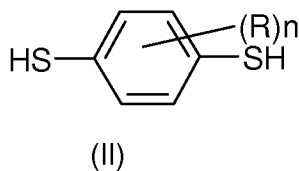
said process comprising the step of self-assembling a compound of formula (II'):



Where R is defined above
or a salt thereof.

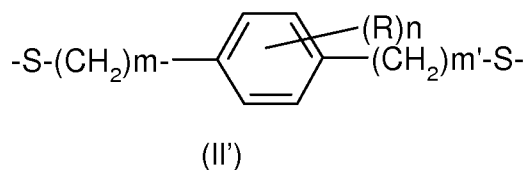
15 According to an embodiment, it is also provided a process of preparation of the compound of formula (I) of the invention as defined above,

said process comprising the step of self-assembling a compound of formula (II):



Where R is defined above
or a salt thereof.

The present invention also concerns a compound of formula (II'):



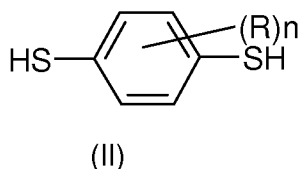
where n and R are as defined above

with the exception of the compounds where $m=m'=0$ and

$n=2$ and $R=\text{COOH}$, or

30 $n=2$ and $R=\text{NH}_2$.

According to an embodiment, the compounds (II') are of formula (II):



where n and R are as defined above
with the exception of the compounds where
n=2 and R=COOH, or
n=2 and R=NH₂.

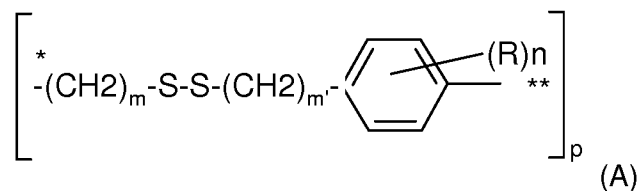
Detailed description

The invention relates to oligomeric macrocyclic receptors that display selective binding for altered vs unaltered peptides or proteins. These oligomeric macrocyclic receptors can be used in methods of separation of altered vs unaltered peptides or proteins, at a polishing stage in analytical and preparative scale.

Some oligomeric macrocyclic receptors also display specificity for Lys residue, thereby enabling their use as specific receptor for Lys-tagged peptides or proteins.

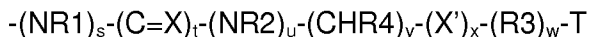
Oligomeric macrocycles

The oligomeric macrocycles that can be used in the frame of the methods and uses according to the invention are subjects of the invention are of formula (A) or (I) as follows:



Where

m and m' identical or different are an integer independently chosen from 0 or 1;
n is an integer comprised between 1 and 4;
p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 50;
each R identical or different may be represented by the following formula:



where T is a terminal group and may be chosen from the groups consisting in H, N₃, alkyl, heteroaryl, a solid phase support;

s, t, u, v, w identical or different may be independently 0 or 1;

R1 represents H or -alkyl optionally substituted by one or more of halogen atoms, OH, OR11, NR11R12, CN;

X represents O, NH or S;

X' is O, NH or S;

R2 represents H, -alkyl optionally substituted by one or more of halogen atoms, OH, OR11, NR11R12, CN;

R3 represents -alkyl-, -heteroaryl- or -aryl- optionally substituted by one or more of -R11 or -OR11, -N=C(NR11₂)-NR11-, -(OCH₂CH₂)_q-(CH₂CH₂O)_q-, -alkyl-NH-C(=O)-NH-, -aryl-NH-C(=O)-NH-, -alkenyl-OCO-NH-, -alkyl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -aryl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -NR11-, -N(R11)-C(=NH)-NH-, -O-, -C(O)O-, -N(R13)-(CH₂)_x-NR11-, -SO₂-aryl-;

R4 is H or COOH;

R11, R12 identical or different are independently selected from the group consisting in H, alkyl optionally substituted by one or more of halogen atoms, OH, CN;

R13 is -alkyl optionally substituted by NR11R12;

q is an integer comprised between 1 and 200;

x is 0 or 1;

where alkyl is C1-C6 alkyl; aryl is a mono or bicyclic C6-C10 aromatic ring system; heteroaryl is a 5 to 10 membered mono or bicyclic aromatic ring system comprising one to 4 heteroatoms chosen from N, O or S.

It being understood that R may be different for each unit and/or for each n;

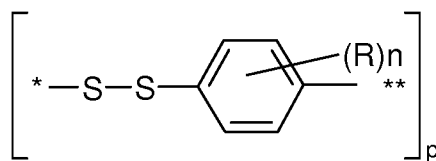
whereby the cyclic form is achieved by the bond between the S (*) of the first unit and the carbon of the phenyl group (**) of the last unit;

as well as its various stereoisomers;

and

b. detecting or collecting complexes formed with the oligomeric macrocycle of general formula (A), wherein said complexes preferentially comprise native form of the protein or oxidized and/or hydrolyzed forms of the protein;

According to an embodiment, the macrocycles are of formula (I)



(I)

where n is an integer comprised between 1 and 4;

p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 50;

each R identical or different may be represented by the following formula:



where T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl, a solid phase support;

s, t, u, v, w identical or different may be independently 0 or 1;

R1 represents H or -alkyl optionally substituted by one or more of halogen atoms, OH, OR11, NR11R12, CN;

X represents O, NH or S;

R2 represents H, -alkyl optionally substituted by one or more of halogen atoms, OH, OR11, NR11R12, CN;

R3 represents -alkyl-, -heteroaryl- or -aryl- optionally substituted by -R11 or -OR11, -N=C(NR11₂)-NR11-, -(CH₂CH₂O)_q-, -alkyl-NH-C(=O)-NH-, -aryl-NH-C(=O)-NH-, -alkenyl-OCO-NH-, -alkyl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -aryl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -NR11-, -N(R11)-C(=NH)-NH-, -O-, -N(R13)-(CH₂)_x-NR11-, -SO₂-aryl-;

R11, R12 identical or different are independently selected from the group consisting in H, alkyl optionally substituted by one or more of halogen atoms, OH, CN;

R13 is -alkyl optionally substituted by NR11R12;

q is an integer comprised between 1 and 200;

where alkyl is C1-C6 alkyl; aryl is a mono or bicyclic C6-C10 aromatic ring system; heteroaryl is a 5 to 10 membered mono or bicyclic aromatic ring system comprising one to 4 heteroatoms chosen from N, O or S.

whereby the cyclic form is achieved by the bond between the S (*) of the first unit and the carbon of the phenyl group (**) of the last unit;

as well as its various stereoisomers.

It is to be understood that R may be different for each unit and/or for each n: when all R are identical, the oligomer of formula (I) will be called homooligomer. When at least one R is different from the others, the oligomer will be called heterooligomer.

5 According to an embodiment, $m=m'=0$

According to a particular embodiment, in formula (I) or (A)

n is 2;

p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 5;

10 each R identical or different may be represented by the following formula:



15 where T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl, a solid phase support;

s, t, u, v, w identical or different may be independently 0 or 1;

R1 represents H or -alkyl;

X represents O, NH or S;

R2 represents H, -alkyl;

20 R3 represents -heteroaryl- or -aryl- optionally substituted by -R11 or -OR11, -N=C(NR11₂)-NR11-, -(CH₂CH₂O)_q-, -NR11-, -N(R11)-C(=NH)-NH-, -N(R13)-(CH₂)_x-NR11-, -SO₂-aryl-, -alkyl-NH-C(=O)-NH-, -aryl-NH-C(=O)-NH-, -alkenyl-OCO-NH-, -alkyl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -aryl-NH-C(=O)-NH-(CH₂CH₂O)_q;

25 R11, R12 identical or different are independently selected from the group consisting in H, alkyl;

R13 is -alkyl optionally substituted by NR11R12;

q is an integer comprised between 1 and 200.

According to a further embodiment, in formula (I) or (A) :

30 n is 2;

p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 5;

each R identical or different may be represented by the following formula:



35 And where

When s is 1:

R1 is H;

t is 0 or 1 and X is O;

u is 0 or 1 and R2 is H,

v is 0;

w is 0 or 1 and R3 is $-(CH_2CH_2O)_q-$, $-alkyl-NH-C(=O)-NH-$, $-aryl-NH-C(=O)-NH-$, $-alkenyl-OCO-NH-$, $-alkyl-NH-C(=O)-NH-(CH_2CH_2O)_q-$, $-aryl-NH-C(=O)-NH-(CH_2CH_2O)_q-$, with q is comprised between 1 and 200; $-SO_2-aryl-$;

T is a terminal group and may be chosen from the groups consisting in H, alkyl or a solid phase support;

or

When s is 0:

t is 1 and X is O;

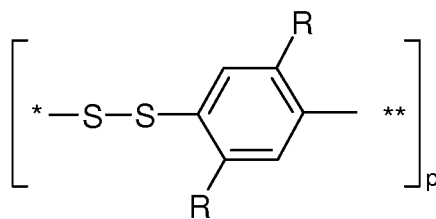
u is 0 or 1 and R2 is H, v is 0, 1 or 2;

w is 0 or 1 with R3 is chosen from $-heteroaryl-$ or $-aryl-$ optionally substituted by $-R_{11}$ or $-OR_{11}$, $-N=C(NR_{11})_2-NR_{11}-$, $-O-$, $-NR_{11}-$, $-N(R_{11})-C(=NH)-NH-$, $-N(R_{13})-(CH_2)_x-NR_{11}-$; with R11 is H or alkyl;

T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl.

Typically, the solid support may be chosen from polystyrene (PS), polyacrylamide, amylose, cellulose, silica, glass, resorcinarene.

According to an embodiment, the compound of formula (A) or (I) may typically comprise disubstituted units, such as in the following formula (I'):

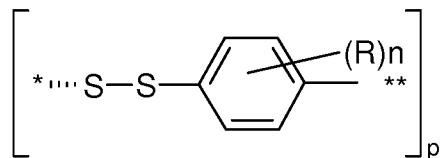


(I')

where each R identical or different, and p is defined as in formula (I).

The compound of formula (I) may exhibit an axial chirality, in that the free rotation of the units respective to each other may be hindered by the bulk of R groups. This may result in the presence of stereoisomers. Such axial chirality and the diastereoselective synthesis thereof has been reported in particular by Ogoshi et al Chem. Eur. J. 2012, 18, 7493-7500.

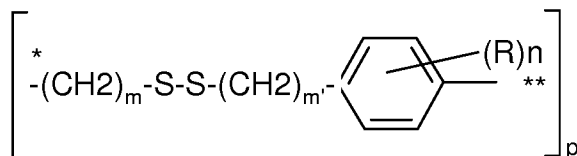
According to an embodiment, compounds of formula (A) such as (I) may thus be in the form of the stereoisomer of formula (I'') :



(I'')

where R, n and p are defined as in formula (A).

According to a further object, the present invention also concerns a compound of formula (A):



(A)

As defined above,

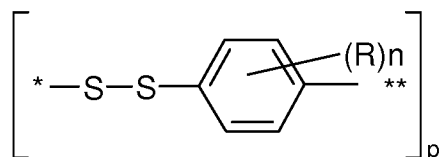
with the exception of the compounds where

$m=m'=0$ and

$p=4$, $n=2$ and $R=\text{COOH}$, or

$p=4$, $n=2$ and $R=\text{NH}_2$.

According to an embodiment, the present invention also concerns a compound of formula (I) :



(I)

As defined above,
with the exception of the compounds where
p=4, n=2 and R=COOH, or
p=4, n=2 and R=NH₂.

5

As used herein, the term "alkyl" refers to a branched or straight hydrocarbon chain of 1 to 6 carbon atoms, which is formed by the removal of one hydrogen atom. In other preferred embodiments, the alkyl group contains from 1 to 4 carbon atoms. A designation such as "C₁-C₆ alkyl" refers to an alkyl radical containing from 1 to 6 carbon atoms. Examples
10 include methyl, ethyl, n-propyl, isopropyl, n-butyl, isobutyl, sec-butyl, t-butyl, pentyl, 2-methylpentyl, hexyl, etc.

As used herein, the term "aromatic" or "aryl" in aryl or heteroaryl refers to a cyclically carbocyclic aryl or heteroaryl system as defined herein, which satisfies the Hückel (4n+2) rule and/or with a stability due to delocalization significantly greater than
15 that of a hypothetical localized structure.

As used herein, the term "aryl" refers to an aromatic carbo, mono-, bi- or multicyclic hydrocarbon ring containing from 6 to 10, preferably 6 carbon atoms, which is formed by removal of one hydrogen atom. Examples include phenyl, naphthyl, indenyl, etc.

As used herein, the term "heteroaryl" refers to a 5 to 10, preferably 6 membered aromatic hetero, mono-, bi- or multicyclic ring, comprising one or more heteroatoms
20 chosen from N, O and/or S. Examples include pyrrolyl, pyridyl, pyrazolyl, thienyl, pyrimidinyl, pyrazinyl, tetrazolyl, indolyl, quinolyl, purinyl, imidazolyl, thienyl, thiazolyl, benzothiazolyl, furanyl, benzofuranyl, 1,2,4-thiadiazolyl, isothiazolyl, triazolyl, tetrazolyl, isoquinolyl, benzothienyl, isobenzofuryl, pyrazolyl, carbazolyl, benzimidazolyl, isoxazolyl,
25 pyridyl-N-oxyde, quinolyl, benzoimidazolyl, benzothiazolyl groups, etc.

As used herein, "Hal" refers to a halogen atom which in turn includes fluoro, chloro, iodo, bromo.

"Alkyl", "aryl", "heteroaryl" refer also to the corresponding "alkylene", "arylene", "heteroarylene" which are formed by the removal of two hydrogen atoms.

30 It will be appreciated that the compounds of the present invention may exhibit one or more stereoisomers due to the axial chirality discussed above. Thus, all stereoisomers of the oligomers of the invention are intended and encompassed herein.

Synthesis

35 The present invention also concerns the process of preparation of compounds of formula (I) as defined above.

The compounds of the present invention may be prepared in a number of ways well known to those skilled in the art. The compounds can be synthesized, for example, by application or adaptation of the methods described below, or variations thereon as appreciated by the skilled person. The appropriate modifications and substitutions being readily apparent and well known or readily obtainable from the scientific literature to those skilled in the art.

In particular, such methods can be found in R.C. Larock, *Comprehensive Organic Transformations*, VCH publishers, 1989.

Compounds of the present invention may be prepared by a variety of synthetic routes. The reagents and starting materials are commercially available, or readily synthesized by well-known techniques by one of ordinary skill in the arts. All substituents, unless otherwise indicated, are as previously defined.

In the reactions described hereinafter it may be necessary to protect reactive functional groups, for example hydroxy, amino, imino, thio or carboxy groups, where these are desired in the final product, to avoid their unwanted participation in the reactions. Conventional protecting groups may be used in accordance with standard practice, for examples see T.W. Green and P. G. M. Wuts in *Protective Groups in Organic Chemistry*, John Wiley and Sons, 1991; J. F. W. McOmie in *Protective Groups in Organic Chemistry*, Plenum Press, 1973.

As used herein, a "leaving group" corresponds to a group which may easily be cleaved from a molecule by breaking a heterolytic bond, with departure of electronic pair. This group may then easily be replaced by another functional group, during a substitution reaction, for example. Such leaving groups may consist in halogen atoms or activated hydroxy groups, such as mesylate, tosylate, triflate or acetyl groups, etc. Examples of leaving groups, as well as references relating to their preparation, are given in « *Advances in Organic Chemistry* », J. March, 3rd Edition, Wiley Interscience, p. 310-316.

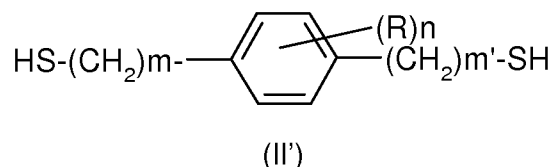
Some reactions may be carried out in the presence of a base. There is no particular restriction on the nature of the base to be used in this reaction, and any base conventionally used in reactions of this type may equally be used here, provided that it has no adverse effect on other parts of the molecule. Examples of suitable bases include: sodium hydroxide, potassium carbonate, cesium carbonate, cesium hydroxide, triethylamine, alkali metal hydrides, such as sodium hydride and potassium hydride; alkyllithium compounds, such as methyllithium and butyllithium; and alkali metal alkoxides, such as sodium methoxide and sodium ethoxide.

Usually, reactions are carried out in a suitable solvent. A variety of solvents may be used, provided that it has no adverse effect on the reaction or on the reagents involved. Examples of suitable solvents include: hydrocarbons, which may be aromatic, aliphatic or cycloaliphatic hydrocarbons, such as hexane, cyclohexane, benzene, toluene and xylene; amides, especially fatty acid amides, such as dimethylformamide, dimethylacetamide; and ethers, such as diethyl ether, diphenylether and tetrahydrofuran.

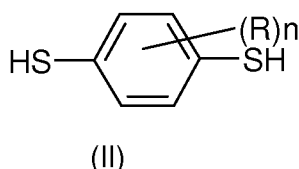
The reactions can take place over a wide range of temperatures. In general, it may be found convenient to carry out the reaction at a temperature of from about 0°C to about 250°C (more preferably from about room temperature to about 100°C). The time required for the reaction may also vary widely, depending on many factors, notably the reaction temperature and the nature of the reagents.

More specifically, according to the process of the invention, the compounds of formula (I) may be prepared by one of the following embodiments:

According a first embodiment, the compounds of formula (A) such as (I) may be prepared by self-assembly of compounds of formula (II'):



Such as compounds of formula (II):

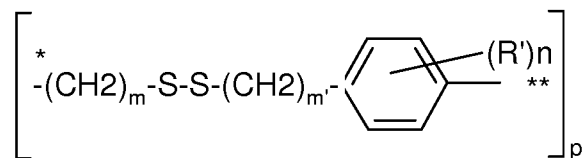


Where n, R are defined as in formula (A).

Typically, said reaction may be conducted by application or adaptation of the methodology disclosed by Skowron et al Journal of Organic Chemistry, 2016, 81, 654-661. More specifically the self-assembly may be conducted by dissolving a compound of formula (II) in an aqueous solution at a pH comprised between 6 and 8, preferably between about 7 and 8, typically under stirring. The precipitate so formed may then be filtered, washed and dried.

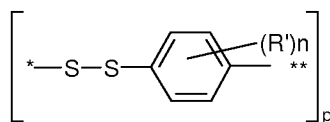
This embodiment is particularly suitable for preparing compounds of formula (A) where R is $-(\text{NR}_1)_s-(\text{C}=\text{X})_t-(\text{NR}_2)_u-(\text{CH}_2)_v-(\text{R}_3)_w-\text{T}$, and t is 1.

According to a second embodiment, the compounds of formula (A) such as (I) may be obtained by derivatization of a compound of formula (III'):



(III')

Such as (III):



(III)

where

m and m' are as defined above,

n is comprised between 1 to 4;

R' represents a -NH₂ or -COOH group and n and p are defined as in formula (A)

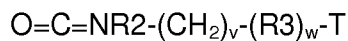
so as to introduce the appropriate R group.

This embodiment may lead to monofunctionalized (only one R' is functionalized into R, the other R' being unamended) or multifunctionalized (more than one, optionally all R' are functionalized into R) oligomers (A)

This embodiment is particularly suitable for compounds where s=1, or s=0 and t=1.

Said reaction may be conducted in a suitable solvent such as DMSO or DMF.

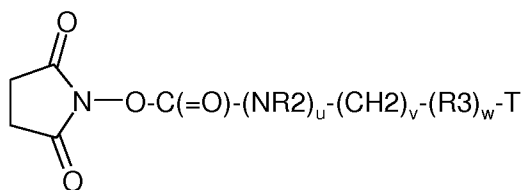
As a representative example of such derivatization reaction, when in formula (A) such as (I) s=1, t=1, X=O, u=1, said derivatization may be carried out by reacting said compound of formula (III) where R'=NH₂, with an isocyanate derivative of formula (IV) :



(IV)

where v, w, R₂, R₃ and T are defined as in formula (A).

As another example of such derivatization, when in formula (A) $s=1$, $t=1$, $X=O$, $u=1$, said derivatization may be carried out by reacting said compound of formula (III) where $R'=NH_2$, with a succinimide derivative of formula (IV') :



(IV')

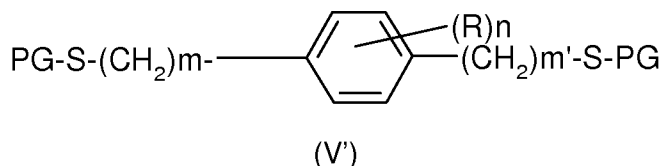
where $u=1$ and v , w , R_2 , R_3 and T are defined as in formula (A).

As an alternative example of such derivatization, when in formula (I) $s=0$, $t=1$, said derivatization may be carried out by reacting said compound of formula (II') where $R'=\text{COOH}$ with a compound of formula (IV''):



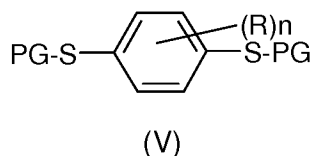
Where u , R_3 , w and T are defined as in formula (A).

The intermediates of formula (II') may be prepared by deprotection of a compound of formula (V'):



(V')

Such as (V)



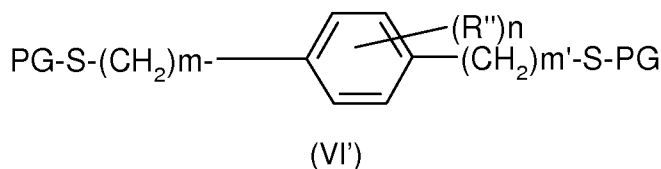
(V)

where R , m , m' are defined as in formula (II') and PG represents a protecting group of the thiol (SH) function.

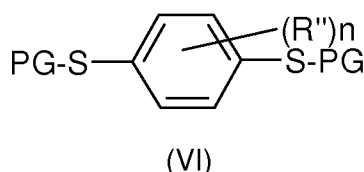
Suitable PG groups are defined in Greene et al as discussed above, and may in particular include $-\text{Alkyl-CN}$, $-\text{C}(=\text{O})-\text{NAlkyl}_2$ groups or trityl ($-\text{C}(\text{phenyl})_3$) group as disclosed in the thesis manuscript of Hamieh, 9-Jan-2015 (Synthetic receptors for ammonium ions using dynamic combinatorial chemistry), e.g. on page 149 and 186-187.

Said deprotection may be carried out by adaptation of known procedures. For example, cyanoethyl protecting groups may be deprotected in the presence of a base, such as cesium hydroxide; $-C(=O)-NAlkyl_2$ groups may be cleaved in the presence of a strong base such as NaOH or KOH (in accordance with the method disclosed by Skowron et al J. Org. Chem. 2016, 81, 654-661) and trityl groups may be deprotected by an acid such as trifluoroacetic acid (see Hamieh, 9-Jan-2015, as discussed above).

The compound of formula (V') may be prepared by functionalization of a compound of formula (VI'):

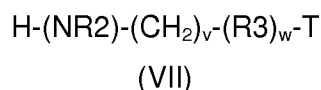


Such as (VI):



where R'' represents a $-(NR_1)_s-C(=O)OH$ group, where R_1 , s are defined as in formula (A).

Typically, when $u=1$, this reaction may be carried out by reacting an appropriate reagent such as a compound of formula (VII):

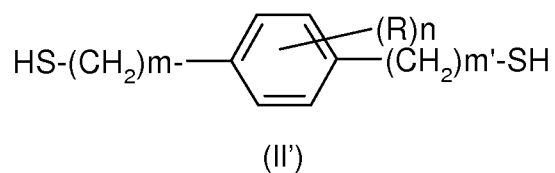


This reaction may be conducted in the presence of a coupling agent, such as 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide, and in a suitable solvent, such as dichloromethane (DCM).

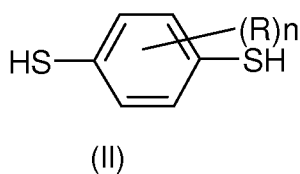
The compounds thus prepared may be recovered from the reaction mixture by conventional means. For example, the compounds may be recovered by distilling off the solvent from the reaction mixture or, if necessary after distilling off the solvent from the reaction mixture, pouring the residue into water followed by extraction with a water-immiscible organic solvent and distilling off the solvent from the extract. Additionally, the product can, if desired, be further purified by various well techniques, such as recrystallization, reprecipitation or the various chromatography techniques, notably column chromatography or preparative thin layer chromatography.

Starting compounds and reactants, unless otherwise indicated, are commercially available or described in literature, or can be prepared according to methods described in literature or known to one of skill in the art.

5 According to a further object, the present invention also concerns the compound of formula (II'):



Such as :



where n and R are defined as above
with the exception of the compounds where
m=m'=0 and
n=2 and R=COOH, or
n=2 and R=NH₂.

Method for detecting or collecting oxidized forms of a protein

The invention also relates to a method for differentiating native form of a protein from oxidized and/or hydrolyzed forms of said protein which comprises:

- a) contacting a composition likely to comprise native, oxidized and/or hydrolyzed forms of a protein with a oligomeric macrocycle of general formula (A), as defined above; and
- b) detecting or collecting complexes formed with the oligomeric macrocycle of general formula (A), wherein said complexes preferentially comprise native form of the protein or oxidized and/or hydrolyzed forms of the protein.

In an embodiment, said method for differentiating native from oxidized and/or hydrolyzed forms of a protein is a method of quantifying native, or oxidized and/or hydrolyzed forms of a protein which comprises:

- a) contacting a composition likely to comprise native, oxidized and/or hydrolyzed forms of a protein with a oligomeric macrocycle of general formula (A), as defined above; and
- b) measuring the quantity of complexes formed with the oligomeric macrocycle of general formula (A), wherein said complexes preferentially comprise native form of the protein, or oxidized and/or hydrolyzed forms of the protein; and
- c) optionally comparing the quantity of complexes measured at step b) with a control.

The control can be for instance a quantity of complexes formed with the oligomeric macrocycle of general formula (A) measured in an aliquot of said composition likely to comprise native, oxidized and/or hydrolyzed forms of a protein at an earlier point in time. The control can also be a quantity of complexes formed with the oligomeric macrocycle of general formula (A) measured in a composition comprising said protein and which is essentially devoid of oxidized and/or hydrolyzed forms of the protein. The control can also be a quantity of complexes formed with the oligomeric macrocycle of general formula (A) measured in a composition comprising said protein and which is essentially devoid of native form of the protein.

In an embodiment, the method is for detecting, quantifying or collecting native forms or a protein. In an embodiment, the method is for detecting, quantifying or collecting oxidized forms or a protein. In another embodiment, the method is for detecting, quantifying or collecting hydrolyzed forms or a protein. In still another embodiment, the

method is for detecting, quantifying or collecting oxidized and hydrolyzed forms of a protein.

According to an embodiment, the discrimination between for native form of a protein from oxidized and/or hydrolyzed forms of said protein may be carried out by induced circular dichroism (ICD).

"Protein" and "peptide" as indifferently used herein may be any natural or synthetic protein or peptide, of at least 2 amino acids linked together by a peptide bond. A peptide or protein may be of any suitable length, e.g., from 2, 5 or 10 amino acids up to 100, 200 or 500 amino acids, or more. The protein or peptide preferably consists of a therapeutic protein, such as an antibody.

As used herein, an "oxidized form of a protein" denotes a protein in which one or more amino acid side chain(s) has(have) been modified post-translationally by oxidation. In an embodiment, an oxidized form of the protein is a methoxidated form of the protein. A "methoxidated form of a protein" is a protein in which one or more methionine residue(s) has(have) been modified by methoxidation. "Methoxidation" (or oxidation of Met) denotes oxidation of a methionine of a protein to methionine sulfoxide (MetO).

As used herein, an "hydrolyzed form of a protein" denotes a protein in which one or more amino acid side chain(s) has(have) been modified post-translationally by hydrolysis. Hydrolyzed forms of a protein do not include forms in which the protein has been modified by peptide bond hydrolysis.

In an embodiment, a "hydrolyzed form of the protein" is a deamidated form of the protein. A "deamidated form of a protein" is a protein post-translationally modified by modification of one or more asparagine to aspartic acid or isoaspartic acid, and/or modification of one or more glutamine to glutamic acid.

The term "native protein" denotes a non-methoxydated non-deamidated form of a protein, and preferably a non-oxidized non-hydrolyzed form of a protein. MetO is a modified amino acid which is not present in native proteins. Asp and Glu may be present in a protein as a result of post-translation modification of Asn and Gln residue(s), or as a result of the translation of a nucleic acid sequence encoding the protein. Accordingly a native protein will contain less Asp or Glu residues compared with a deamidated form thereof.

Native proteins in solution can undergo hydrolytic and/or oxidative modification(s), resulting in a mixture of protein isomers. The method of the invention is intended to

differentiate native proteins from modified isoforms, by selectively binding, or preferentially specifically binding, modified isoforms (oxidized and/or hydrolyzed forms) of the protein.

In an embodiment said oxidized and/or hydrolyzed forms of the protein are present in a composition comprising oxidized and/or hydrolyzed forms of the protein, and native protein. The composition is in particular a liquid composition.

Preferably, the method for differentiating native from oxidized and/or hydrolyzed forms of a protein comprises:

- a) contacting a composition comprising oxidized and/or hydrolyzed forms of the protein, and native protein, with a oligomeric macrocycle of general formula (A), as defined above; and
- b) detecting or collecting complexes formed with the oligomeric macrocycle of general formula (A), wherein said complexes preferentially comprise native form of the protein, or oxidized and/or hydrolyzed forms of the protein;

wherein the method further comprises c) recovering the forms of the protein which are bound to said oligomeric macrocycle of general formula (A).

In another aspect, the invention relates to a method of purifying native protein from a composition comprising oxidized and/or hydrolyzed forms of the protein, and native protein, the method comprising:

- a) contacting a composition comprising oxidized and/or hydrolyzed forms of the protein, and native protein, with a oligomeric macrocycle of general formula (A), as defined above; and
- b) recovering the forms of the protein which are bound or not bound to said oligomeric macrocycle of general formula (A).

In the step of contacting a composition likely to comprise oxidized and/or hydrolyzed forms of a protein with an oligomeric macrocycle of general formula (A), contacting is performed with one or more oligomeric macrocycle of general formula (A).

The composition likely to comprise native, oxidized and/or hydrolyzed forms of a protein is a liquid composition, preferably an aqueous composition.

In an embodiment, said oligomeric macrocycle of general formula (A) is also in solution with the native, oxidized and/or hydrolyzed forms of the protein likely contained in the composition. The isoelectric point (pI) of the protein differing from the pI of the complexes comprising the oligomeric macrocycle of formula (A) and the protein, recovering the forms of the protein which are bound or unbound to said oligomeric

macrocycle of general formula (A) can be performed for instance by capillary electrophoresis.

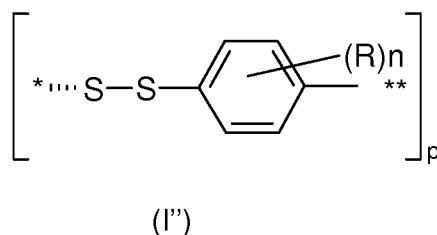
In another embodiment, said oligomeric macrocycle of general formula (A) is functionalized and is covalently bound to a solid support (i.e. T is a solid phase), e.g. a chromatographic resin or column.

In still another embodiment, said oligomeric macrocycle of general formula (A) comprises a PEG moiety, and the forms of the protein which are bound to the functionalized oligomeric macrocycle of general formula (A) and are present in aqueous solution are separated from the unbound forms of the protein by liquid-liquid extraction using a saline aqueous phase.

Detecting or collecting complexes formed by the oxidized and/or hydrolyzed forms of the protein with the oligomeric macrocycle of general formula (A) can be readily achieved by the skilled person.

Affinity of oligomeric macrocycles of general formula (A) for various amino acids or dipeptides has been measured. From this measurement, selectivity for dipeptides comprising native versus oxidized or hydrolyzed residues has been determined. The selectivity determined on dipeptides with the oligomeric macrocycles of general formula (A) is predictive for the selectivity of said oligomeric macrocycles of general formula (A) towards oligopeptides (i.e. peptides comprising up to 10 amino acids). Selectivity for oligopeptides as well as for larger proteins can also be determined empirically by exposing native peptide or protein to an oxidizing and/or hydrolytic treatment, and determining the affinity of the oligomeric macrocycles of general formula (A) for the native peptide or protein comparatively to the peptide or protein exposed to oxidizing and/or hydrolytic treatment.

For differentiating native and methoxydated forms of a protein, the oligomeric macrocycle of general formula (A) is preferably an oligomeric macrocycle of formula (I'') :



where :

n is 2;

p is 4;

R is $-(\text{NR}_1)_s-(\text{C}=\text{X})_t-(\text{NR}_2)_u-(\text{CH}_2)_v-(\text{R}_3)_w-\text{T}$

Where s is 0:

t is 1 and X is O;

u is 0 or 1 and R₂ is H, v is 0, 1 or 2;

w is 0 or 1 with R₃ is chosen from -heteroaryl- or -aryl- optionally substituted by -R₁₁ or -OR₁₁, -N=C(NR₁₁)₂-NR₁₁-, -O-, -NR₁₁-, -N(R₁₁)-C(=NH)-NH-, -N(R₁₃)-(CH₂)_x-NR₁₁-; with R₁₁ is H or alkyl;

T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl.

Such compounds are referred to as "compounds (I')-A".

In particular an illustrative compound for compounds (I')-A is compound **(15)₄** of formula (I'') where in formula (I''):

n is 2;

p is 4;

R is -(NR₁)_s-(C=X)_t-(NR₂)_u-(CH₂)_v-(R₃)_w-T

Where s is 0:

t is 1 and X is O;

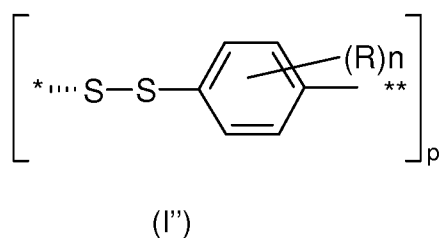
u= v= 0

w=1 with R₃ is -O-, and

T is H.

The oligomeric macrocycle of formula **(15)₄** preferentially binds to a methionine sulfoxide in C-ter position to a Lys residue, compared to in particular to a methionine in a motif Lys-Met. The oligomeric macrocycle of formula **(15)₄** has two-fold selectivity for dipeptide Lys-Met compared to Lys-Met(O). The oligomeric macrocycle of formula **(15)₄** preferentially binds native forms of oligopeptides, compared to methoxydated forms

For differentiating native from deamidated forms of a protein, the oligomeric macrocycle of general formula (A) is preferably an oligomeric macrocycle of formula (I''):



Where

n is 2;

p is 4;

R is $-(NR_1)_s-(C=X)_t-(NR_2)_u-(CH_2)_v-(R_3)_w-T$

And where

When s is 1:

R1 is H;

t is 0 or 1 and X is O;

u is 0 or 1 and R2 is H,

v is 0;

w is 0 or 1 and R3 is $-(CH_2CH_2O)_q-$, $-alkyl-NH-C(=O)-NH-$, $-aryl-NH-C(=O)-NH-$, $-alkenyl-OCO-NH-$, $-alkyl-NH-C(=O)-NH-(CH_2CH_2O)_q-$, $-aryl-NH-C(=O)-NH-(CH_2CH_2O)_q-$, with q is comprised between 1 and 200; $-SO_2-aryl-$;

T is a terminal group and may be chosen from the groups consisting in H, alkyl or a solid phase support;

(herein referred to as "compounds (I''-B)")

or

When s is 0:

t is 1 and X is O;

u is 0 or 1 and R2 is H, v is 0, 1 or 2;

w is 0 or 1 with R3 is chosen from $-heteroaryl-$ or $-aryl-$ optionally substituted by $-R_{11}$ or $-OR_{11}$, $-N=C(NR_{11_2})-NR_{11}-$, $-O-$, $-NR_{11}-$, $-N(R_{11})-C(=NH)-NH-$, $-N(R_{13})-(CH_2)_x-NR_{11}-$; with R11 is H or alkyl;

T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl;

(herein referred to as "compounds (I''-A)".)

Illustrative compounds are :

- compound **(15)₄** as defined above which is illustrative for compounds (I''-A) and
- compound **(19)₄** which is illustrative for compounds (I''-B) and which corresponds to formula (I'') where :

n is 2;

p is 4;

R is $-(NR_1)_s-(C=X)_t-(NR_2)_u-(CH_2)_v-(R_3)_w-T$

where s is 1, R1 is H, t=u=w= 0 and T is H.

Oligomeric macrocycle of formula **(15)₄** preferentially binds to a Gln residue adjacent to a Lys residue, in particular to a Gln in a Lys-Gln motif. The oligomeric macrocycle of

formula (15)₄ has six-fold selectivity for Lys-Gln compared to Lys-Glu. The oligomeric macrocycle of formula (15)₄ preferentially binds native forms of oligopeptides, compared to deamidated forms.

Oligomeric macrocycle of formula (19)₄ has a two-fold increase in affinity for Met compared to Met(O), and a three-fold increase in affinity for Glu compared to Gln. The oligomeric macrocycle of formula (19)₄ preferentially binds native forms of oligopeptides, compared to methoxydated or deamidated forms.

Deamidation of Asn and Gln residues in protein is considered to disrupt the protein's hydrogen bonding network near deamidation sites, thereby affecting the global structure of the protein. In particular regions adjacent to deamidation sites are expected to become more flexible, allowing for spatial rearrangement (Soulby et al., Protein Science, 2015, 24: 850-860). Accordingly, the above oligomeric macrocycles not only to detect Glu or Asp residue adjacent to aromatic amino acids or to lysine in the primary structure of a protein, but they can also detect Glu or Asp residue that became adjacent (e.g. no more than 5Å apart) to aromatic amino acids or to lysine in the tertiary or quaternary structure of a protein, due conformational transition induced by the deamidation.

Accordingly, in an embodiment the (native) protein, for which detection of oxidized and/or hydrolyzed forms is sought, comprises at least one Methionine, Glutamine or Asparagine which is adjacent to a Lysine in the primary, tertiary, or quaternary structure of said protein.

Method of purification of a protein or peptide

As reported herein, some oligomeric macrocycle of formula (A), and in particular compound (15)₄ bind basic aminoacids such as lysine and arginine on dipeptide sequences. Compound (15)₄ actually displays a strong preference for lysine vs arginine. Furthermore modification of a peptide by N-terminal addition of Lys or Met-Lys was shown to induce a ten-fold increase in affinity for Compound (15)₄ compared to the affinity of the unmodified peptide. The properties of the oligomeric macrocycle (15)₄ make it particularly attractive as selective receptor for N-Lys-tagged peptides or proteins, or more generally for proteins or peptides comprising a lysine at position 1 or 2 of their amino acid sequences.

As already mentioned above, the terms "protein" and "peptide" are indifferently used herein.

Accordingly, the invention also relates to a method of purification of a protein or peptide which comprises:

- a) Contacting a composition comprising a target protein or peptide with a oligomeric macrocycle **(15)₄** as defined above; and
- b) Recovering the target protein or peptide bound to the oligomeric macrocycle **(15)₄**;

5 wherein the target protein or peptide comprises a lysine at position 1 or 2 of its amino acid sequence.

10 All peptides and proteins produced in cells, naturally or recombinantly, initially contain a N-terminal Methionine, which is encoded by the initiation codon of the protein or peptide coding sequence.

In an embodiment, the target protein or peptide comprises a lysine immediately downstream of (or in C-ter to) a N-terminal methionine (i.e. the lysine is at position 2 of the primary sequence of the target protein or peptide). In an embodiment, the native coding sequence of the protein or peptide encodes an amino acid sequence comprising a lysine
15 at position 2; accordingly said target protein or peptide naturally comprises a lysine at position 2, and the protein or peptide can be recombinant or not. In another embodiment, the protein or peptide is recombinant and its coding sequence has been modified to encode an amino acid sequence comprising an insertion of a lysine at position 2 of its primary structure.

20 In another embodiment, the target protein or peptide comprises a Lysine at position 1 of its amino acid sequence. Preferably, the target protein or peptide comprises, at position 2 of its amino acid sequence, a residue selected from the group consisting of Met, Lys, Arg, Gly.

In an embodiment, the target peptide or protein is synthetic, i.e. it has been
25 produced by chemical synthesis, e.g. by stepwise elongation of an amino acid chain, fragment condensation, chemical ligation, or any other method well known to the skilled person. Peptides and proteins consisting of 2 to 100 amino acids, or more, can readily be produced by chemical synthesis.

30 In an embodiment, the target peptide or protein is the mature form of a pro-peptide or pro-protein. In this embodiment the sequence of the target peptide or protein results from proteolytic maturation of the pro-form of the target peptide or protein, wherein the proteolytic maturation cleaves the pro-form sequence in a motif X-Lys and thereby generates an amino acid sequence starting by a lysine.

35 The composition comprising a target protein or peptide is a liquid composition, preferably an aqueous composition.

In an embodiment, said oligomeric macrocycle (15)₄ is also in solution with the comprising a target protein or peptide contained in the composition.

In another embodiment, said oligomeric macrocycle (15)₄ has been functionalized and is covalently bound to a solid support (i.e. T is a solid phase), e.g. a chromatographic resin or column.

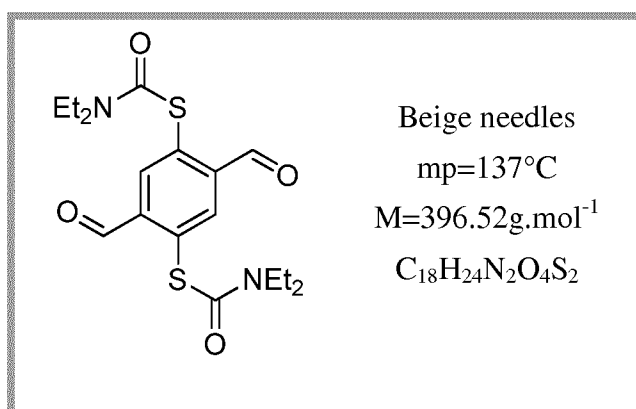
The invention will be further illustrated by the following examples.

EXAMPLES

Example 1: Synthetic Procedures

1.1 Synthesis of the monomers

2,5-di(diethylthiocarbamoylsulfanyl)-1,4-dibenzaldehyde



NMR ¹H (CDCl₃, 300MHz): (ppm) = 10.23 (s, 2H), 8.19 (s, 2H), 3.46 (dd, *J*=26.1 and 6.6Hz, 8H) and 1.24 (dt, *J*=55.8 and 6.2Hz, 12H)

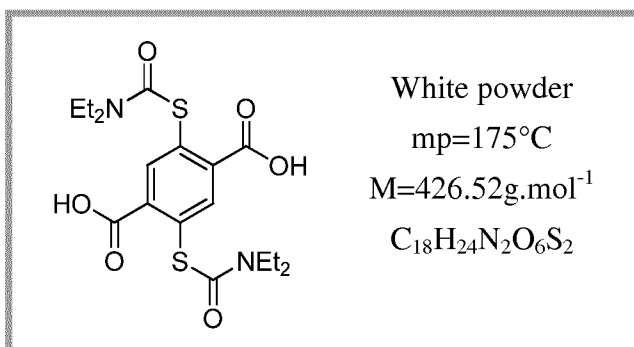
NMR ¹³C (CDCl₃, 75MHz): (ppm) = 189.79, 163.21, 140.74, 137.07, 134.45, 43.09, 14.16 and 13.22.

HRMS (ESI) *m/z* calculated for C₁₈H₂₄N₂NaO₄S₂: 419.1070, measured: 419.1064

In a microwave tube was poured a suspension of 2,5-di(diethylthiocarbamoyloxy)-1,4-dibenzaldehyde (0.1g ; 0.25mmol) and previously degassed dry NMP (2.5mL).

The rearrangement was performed by microwave at 220°C during 3mn. Then the mixture was poured into 100mL of ethyl acetate and 100mL water was added. The organic layer was washed with 3x100mL of water in order to eliminate the NMP and then dried and evaporated to give a crude solid. Then cold ethanol was added, the solid was triturated and placed in the refrigerator during 2 hours. The pure compound is filtered off and washed with cold ethanol to afford beige needles (0.065g, 65%).

Dichloromethane/Cyclohexane (8:2) R_f = 0.5

2,5-di(diethylthiocarbamoylsulfanyl)terephthalic acid

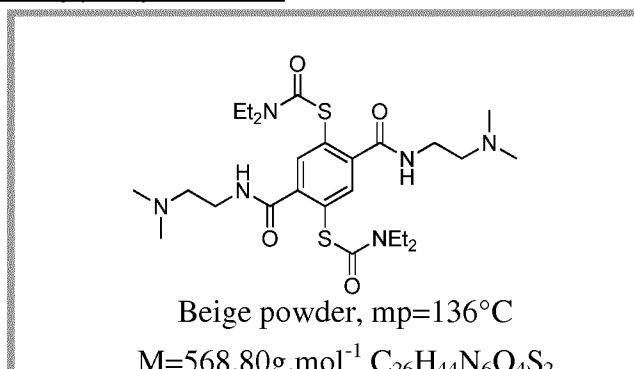
NMR ¹H (MeOD, 300MHz): (ppm) = 7.57 (s, 2H), 3.51-3.45 (br, 8H), 1.31 (br, 6H) and 1.16 (br, 6H)

NMR ¹³C (MeOD, 75MHz): (ppm) = 175.34, 167.94, 147.78, 136.81, 127.47, 43.96, 43.66, 14.13 and 13.48.

HRMS (ESI) *m/z* calculated for C₁₈H₂₄N₂NaO₆S₂: 451.0968, **measured:** 451.0964

In a round bottom flask was poured the 2,5-di(diethylthiocarbamoylsulfanyl)-1,4-dibenzaldehyde (0.430g; 1.08mmol) and a 5:3 mixture (20mL) of dioxane and water. Then permanganate potassium (1.58g; 10mmol) was added by portions and the mixture stirred during 5H. The reaction was monitored by TLC chromatography and the pH was adjusted to 3 with a 10% HCl solution. Then oxalic acid (2.25g; 25mmol) was added dropwise by portions under vigorous stirring and the mixture was stirred overnight at room temperature. The precipitate was then collected by filtration and washed with water. The solid was dried under vacuum and dissolved in methanol. The residual salts were filtrate off and the filtrate was evaporated under vacuum to give a white powder (0.368g; 0.78mmol; 80%).

Ethyl acetate/Petroleum spirit/TFA (7:3:0.1) R_f = 0.3

2,5(diethylthiocarbamoylsulfanyl)terephthalamide

NMR ^1H (CDCl_3 , 300MHz): (ppm) = 7.75 (s, 2H); 7.24 (br t, 2H); 3.47 (m, 12H); 2.46 (t, $J=6.6\text{Hz}$, 4H); 2.24 (s, 12H); 1.27 (br, 6H) and 1.15 (br, 6H).

NMR ^{13}C (CDCl_3 , 75MHz): (ppm) = 167.24, 165.69, 144.21, 137.90, 127.56, 58.01, 45.48, 43.13, 42.66, 37.91, 14.08 and 13.36.

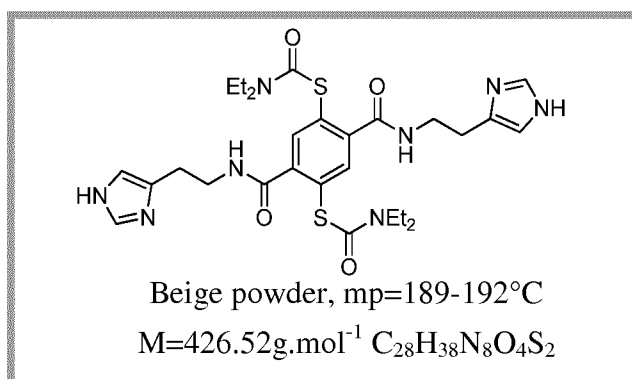
HRMS (ESI) m/z calculated for $\text{C}_{26}\text{H}_{45}\text{N}_6\text{O}_4\text{S}_2$: 569.2938, measured: 569.2929

In a round bottom flask was poured the 2,5-(diethylthiocarbamoylsulfanyl)terephthalic acid (0.200g; 0.47mmol), followed by hydroxybenzotriazole hydrate (0.165g; 1.22mmol). The solids were dissolved in dry DCM (20mL)* under inert atmosphere and cooled to 0°C with an ice bath to give a heterogeneous mixture. Then 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (0.189g; 1.22mmol; 0.215mL) was added and the ice bath removed. The reaction mixture was stirred 1h at room temperature to give an homogeneous solution. Whereupon N,N-dimethylethylenediamine (0.107g; 1.22mmol; 0.133mL) was added, previously dissolved in anhydrous freshly distilled and stabilized over amylene DCM (10mL)* and diisopropylethylamine (0.158g; 1.22mmol; 0.213mL) at 0°C . Then the ice bath was removed and the reaction mixture was stirred overnight at room temperature.

The reaction was monitored by TLC chromatography and the DCM was evaporated. The crude product was then purified by flash chromatography using a mixture of chloroform and 20% of methanol to give a white solid. (0.216g, 80%)

Chloroform + 20% Methanol: R_f = 0.2

2,5(diethylthiocarbamoylsulfanyl)terephthalamide



NMR ^1H (CD_3OD , 300MHz): (ppm) = 7.63 (s, 2H); 7.61 (d, $J=1.1\text{Hz}$, 2H); 6.92 (br, 2H); 3.60 (t, $J=7.2\text{Hz}$, 4H); 3.47 (br, 8H); 2.89 (t, $J=7.2\text{Hz}$, 4H); 1.31 (br, 6H) and 1.17 (br, 6H).

NMR ^{13}C (CD_3OD , 75MHz): (ppm) = 169.60, 166.26, 144.26, 137.59, 136.14, 129.56, 118.65, 112.49, 44.09, 43.88, 40.89, 14.13 and 13.37.

HRMS (ESI) m/z calculated for $\text{C}_{28}\text{H}_{39}\text{N}_8\text{O}_4\text{S}_2$: 615.2530, **measured:** 615.2534

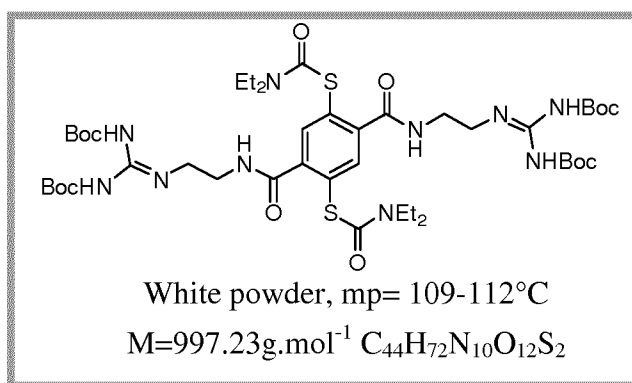
In a round bottom flask was poured the 2.5-(diethylthiocarbamoylsulfanyl)terephthalic acid (0,200g; 0,47mmol), followed by hydroxybenzotriazole hydrate (0,165g; 1,22mmol). The solids were dissolved in dry DCM (20mL)* under inert atmosphere and cooled to 0°C with an ice bath to give an heterogeneous mixture. Then 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (0,189g; 1,22mmol; 0,215mL) was added and the ice bath removed. The reaction mixture was stirred 1H at room temperature to give an homogeneous solution. Whereupon histamine (0,208g; 1,88mmol) was added, previously dissolved in anhydrous freshly distilled and stabilized over amylene DCM (10mL)* and diisopropylethylamine (0,158g; 1,22mmol; 0,213mL) at 0°C. Then the ice bath was removed and the reaction mixture was stirred overnight at room temperature.

The reaction was monitored by TLC chromatography and the DCM was evaporated. The residue obtained was then dissolved in chloroform (40mL) and the organic layer washed with sodium hydrogenocarbonate (3x20mL). Then the organic layer was dried over magnesium sulfate and evaporated to give a crude product.

The crude product was then purified by flash chromatography C18 using a mixture of water and methanol to give a white solid. (0.132g, 65%) (gradient 100%water to 100%methanol)

Chloroform/Methanol (8:2) (Beforehand: Neutralized silica with Et_3N): $R_f = 0.3$

2.5(diethylthiocarbamoylsulfanyl)terephthalamide



NMR ^1H (CDCl_3 , 300MHz): (ppm) = 11.54 (s, 2H); 8.54 (t, $J=5.1\text{Hz}$, 2H); 7.77 (s, 2H); 7.41 (t, $J=5.1\text{Hz}$, 2H); 3.61 (m, 4H); 3.57 (m, 4H); 3.42 (m, 8H); 1.48 (s, 36H); 1.28 (br, 6H) and 1.15 (br, 6H).

NMR ^{13}C (CDCl_3 , 75MHz): (ppm) = 167.75, 165.85, 163.55, 156.58, 153.07, 144.27, 137.84, 127.65, 83.26, 79.39, 43.23, 42.75, 40.38, 39.38, 28.39, 28.20, 13.98 and 13.32.

HRMS (ESI) m/z calculated for $\text{C}_{44}\text{H}_{74}\text{N}_{10}\text{O}_{12}\text{S}_2$: 499.2459, **measured:** 499.2466

In a round bottom flask was poured the 2,5-di(diethylthiocarbamoylsulfanyl)terephthalic acid (0.100g; 0.23mmol), followed by hydroxybenzotriazole hydrate (0.081g; 0.6mmol). The solids were dissolved in dry DCM (10mL)* under inert atmosphere and cooled to 0°C with an ice bath to give an heterogeneous mixture. Then 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (0.093g; 0.6mmol; 0.106mL) was added and the ice bath removed. The reaction mixture was stirred 1h at room temperature to give an homogeneous solution. Whereupon 2-(2-Aminoethyl)-1,3-di-Boc-guanidine (0.181g; 0.6mmol) was added, previously dissolved in anhydrous freshly distilled and stabilized over amylene DCM (10mL)* and diisopropylethylamine (0.077g; 0.6mmol; 0.104mL) at 0°C. Then the ice bath was removed and the reaction mixture was stirred overnight at room temperature.

The reaction was monitored by TLC chromatography and the DCM was evaporated. The residue obtained was then dissolved in chloroform (40mL) and the organic layer washed with sodium hydrogenocarbonate (3x20mL). Then the organic layer was dried over magnesium sulfate and evaporated to give a crude product.

The crude product was then purified by flash chromatography (silica first neutralized with EtNH_3 and washed with eluent in order to eliminate residual trimethylamine) using a mixture of chloroform and 20% of methanol to give a white solid. (0.140g, 60%).

Chloroform/Methanol (8:2) (Beforehand: Neutralized silica with Et_3N): R_f = 0.9

Dialkylthiocarbamate deprotection into thiophenols (in accordance with Skowron et al J. Org. Chem. 2016, 81, 654-661))

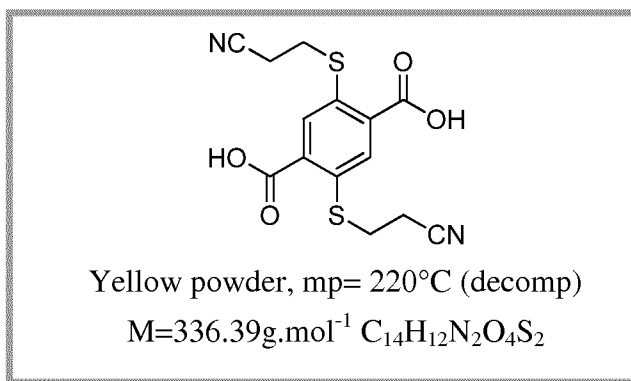
Protected building block was suspended under argon in 3 mL of a 4.2 M degassed solution of KOH in MeOH/H₂O (5/1). The mixture was then refluxed for 1 h 30 min. The reaction mixture was allowed to cool to room temperature and degassed water (13 mL) was added to the solution followed by HCl (10%, 5 mL) to afford a white precipitate. The precipitate was filtered, washed with degassed water, and dried under vacuum to give the deprotected building block as a solid.

To 20 mL of a 1.3 M solution of KOH in H₂O/EtOH (1/1) degassed solution was added the protected building block. The mixture was purged with argon for 15 min and then heated to reflux overnight. The reaction mixture was allowed to cool to room temperature, and concentrated HCl was added to afford a yellow precipitate. The precipitate was filtered,

washed with degassed water, and dried under vacuum to give the pure deprotected building block as a solid.

2,5-di((2-cyanoethyl)thio)terephthalic acid

5



10

NMR ¹H (DMSO, 300MHz): (ppm)= 13.63 (s, 2H), 7.81 (s, 2H), 3.26 (t, *J*=6.6 Hz, 4H), 2.87 (t, *J*=6.6Hz, 4H)

NMR ¹³C (DMSO, 75MHz): (ppm) = 166.79, 134.16, 132.92, 128.53, 119.23, 27.14 and 16.59.

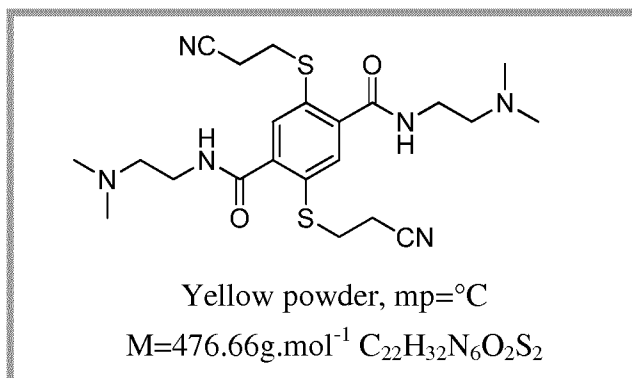
15 **HRMS (ESI) *m/z* calculated for C₁₄H₁₂N₂NaO₄S₂: 359.0131, measured: 359.0130**

The 2,5-dimercaptoterephthalic acid (100mg, 0.42mmol) was dissolved in 3 mL of dry and degassed THF. This solution was added in a round bottom flask under nitrogen which contained NaH 60% in mineral oil (337mg, 8.4mmol) previously washed from its oil with degassed hexane. The addition was performed dropwise at 0°C, then the ice bath was removed and the solution was stirred for 2H at room temperature. Then the 3-Bromopropionitrile (0.696mL, 8.4mmol) was added dropwise and the mixture was heated at reflux for 24H. Additional 3-Bromopropionitrile (0.348mL, 4.2mmol) was added after 24H and the mixture was heated at reflux 24H more.

The reaction was monitored by HPLC.

25 THF was evaporated and the product was suspended in pentane, triturated and filtered. Then water and HCl 4 M were added, the precipitate was filtered off and washed with water and dried over vacuum. The crude solid was then dissolved with the minimum of DMSO and HCl 10% in water was added, the precipitate was filtered off and finally washed with a small amount of water and dried under vacuum overnight. This afforded a yellow powder (0.128mg, 90%).

30

2,5-di((2-cyanoethyl)thio)terephthalamide

NMR ^1H (CDCl_3 , 400MHz): (ppm)= 7.63 (s, 2H), 7.05 (br, 2H), 3.52 (q, $J=11.2\text{Hz}$, 4H), 3.16 (t, $J=7.0\text{Hz}$, 4H), 2.65 (t, $J=7.0\text{Hz}$, 4H), 2.53 (t, $J=5.9\text{Hz}$, 4H) and 2.26 (s, 12H).

NMR ^{13}C (CDCl_3 , 100MHz): (ppm) = 166.65, 140.47, 132.08, 131.73, 117.97, 57.50, 45.22, 37.57, 30.43 and 18.31.

HRMS (ESI) m/z calculated for $\text{C}_{22}\text{H}_{33}\text{N}_6\text{O}_2\text{S}_2$: 477.2101, measured: 477.2105

In a round bottom flask was poured the 2,5-di((2-cyanoethyl)thio)terephthalic acid (0.158g; 0.47mmol), followed by hydroxybenzotriazole hydrate (0.165g; 1.22mmol). The solids were dissolved in dry DCM (20mL)* under inert atmosphere and cooled to 0°C with an ice bath to give an heterogeneous mixture. Then 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (0.189g; 1.22mmol; 0.215mL) was added and the ice bath removed. The reaction mixture was stirred 1h at room temperature to give an homogeneous solution. Whereupon N,N-dimethylethylenediamine (0.107g; 1.22mmol; 0.133mL) was added, previously dissolved in anhydrous freshly distilled and stabilized over amylene DCM (10mL)* and diisopropylethylamine (0.158g; 1.22mmol; 0.213mL) at 0°C. Then the ice bath was removed and the reaction mixture was stirred overnight at room temperature.

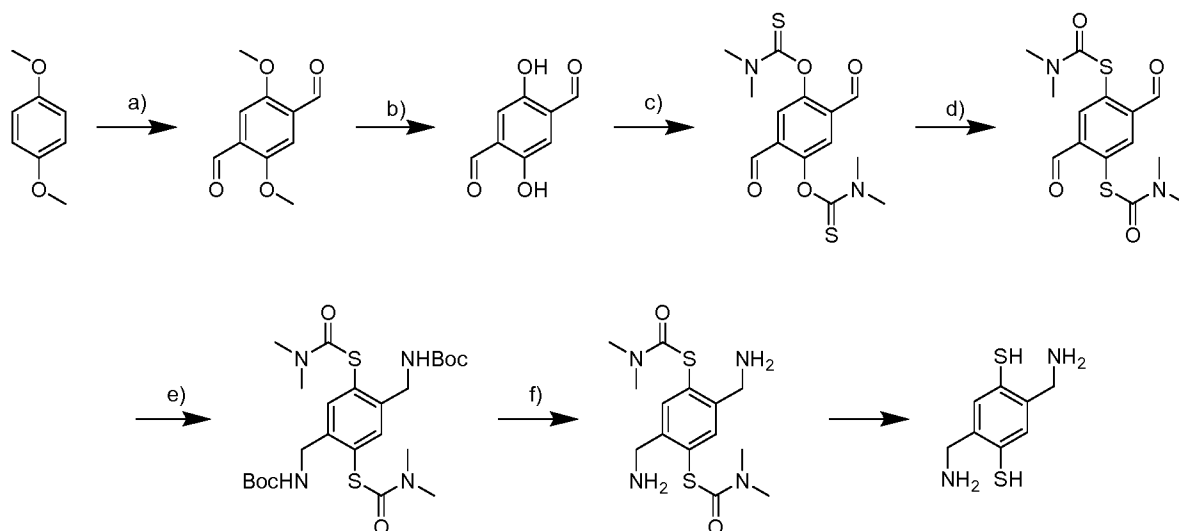
Then the DCM was evaporated and the crude product purified by C18 chromatography using a methanol/water gradient to give a light yellow solid.

Cyanoethyl deprotection :

In a 50mL flask, cyanoethyl-protected building block (0.145 mmol) and cesium hydroxide (488 mg, 2.90 mmol) were placed under nitrogen atmosphere. 7.5 mL of degassed anhydrous THF and 7.5 mL of degassed anhydrous methanol were added and the mixture was stirred at room temperature for 3 h. HCl 3 M was then added until pH 4. A precipitate (CsCl) was formed at the same time as a sticky solid was obtained. The yellow precipitate

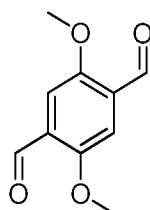
was removed by filtration while the paste was dissolved in water and added to the filtrate. Methanol and THF were evaporated. The remaining water solution was purified by automatic flash-chromatography, affording the desired product as a yellow solid. Building block for deamidation recognition

5 Amino building block



Cationic building block synthesis

- 10 a) 1. *n*-BuLi, TMEDA, Et₂O, reflux, 16 h. 2. 1-formylpiperidine, r.t., 1 h. 51 % ; b) 1. BBr₃, DCM, r.t., 3 h. 2. H₂O, r.t., 1 h. 68 % ; c) dimethylthiocarbamoyl chloride, DABCO, DMA, r.t., 48 h. 71 % ; d) 210°C, 9 mn, NMP, 55 % ; e) BocNH₂, TFA, Et₃SiH, CHCl₃, 40°C, 5 d., 67 % ; f) TFA, tetrachloroethane, r.t., 20 mn, 81 % ; g) NaOH/isopropanol or LiAlH₄/DME

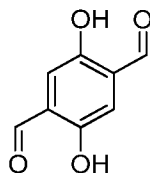


15 1-2 2,5-dimethoxyterephthalaldehyde

To a solution of 1-1 1,4-dimethoxybenzene (11.5 g, 83.26 mmol, 1 eq) in distilled Et₂O (276 mL) was added TMEDA (37.4 mL, 249.78 mmol, 3 eq). At 0°C, *n*BuLi at 2.5 M in hexane (100 mL, 249.78 mmol, 3 eq) was added dropwise. The mixture was stirred at reflux overnight under inert atmosphere. At 0°C, 1-formylpiperidine (27.74 mL, 249.78 mmol, 3 eq). was added dropwise. The mixture was stirred at room temperature for 1 h. Following the addition of distilled water (300 mL) and HCl 3 M (57.5 mL), the mixture was

extracted with hot CHCl_3 (300 mL $\times$ 4). The organic phase was dried over anhydrous Na_2SO_4 and filtered. After removing the solvent, the remaining residue was purified by recrystallization in CHCl_3 to give an orange solid (9.51 g, 51 %).

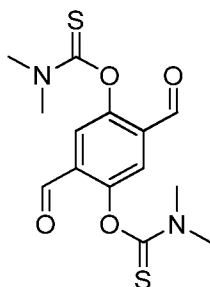
^1H NMR (300 MHz, CDCl_3) δ ppm = 10.51 (s, 2 H, CHO), 7.46 (s, 2 H, Ar), 3.95 (s, 6 H, CH_3) data matched with literature reference



1-3 2,5-dihydroxyterephthalaldehyde

To a solution of **1-2** 2,5-dimethoxyterephthalaldehyde (6.59 g, 33.93 mmol, 1 eq) in distilled DCM (156 mL) was added at 0°C BBr_3 at 1 M in distilled DCM (133.7 mL, 133.7 mmol, 4 eq). The mixture was stirred for 3 hour under N_2 . Distilled water (300 mL) was then added at 0°C . The mixture was extracted with hot CHCl_3 (300 mL $\times$ 4) and washed with Rochelle's salt (300 mL) and water (2 $\times$ 300 mL). The organic phase was dried over anhydrous Na_2SO_4 and filtered. After removing the solvent, the remaining residue was purified by recrystallization in CHCl_3 to give a yellow solid (3.82 g, 68 %).

^1H NMR (300 MHz, CDCl_3) δ ppm = 10.23 (s, 2 H, CHO), 9.97 (s, 2 H, OH), 7.25 (s, 2 H, Ar) data matched with literature reference

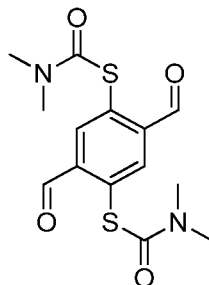


1-4 O,O'-(2,5-diformyl-1,4-phenylene) bis(dimethylcarbamothioate)

To a solution of **1-3** 2,5-dihydroxyterephthalaldehyde (423 mg, 2.546 mmol, 1 eq) and DABCO (1.142 g, 10.184 mmol, 4 eq) in dry DMA (11.3 mL) was added at 0°C dimethylthiocarbamoyl chloride (1.252 g, 10.184 mmol, 4 eq) in 3.7 mL of dryDMA). The mixture was stirred at room temperature for 48 h under inert atmosphere. The white precipitate was filtrated and washed with water. Drying of the resulting solid gave a white powder (618 mg, 71 %).

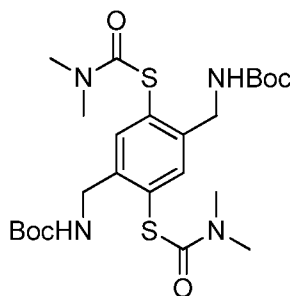
White solid; mp: $226.1 - 228.6^\circ\text{C}$; HRMS (ESI) $[\text{M} + \text{H}]^+$ found 341.0623, calculated 341.0624 for $[\text{C}_{14}\text{H}_{17}\text{N}_2\text{O}_4\text{S}_2]^+$; ^1H NMR (300 MHz, CDCl_3) δ ppm = 10.07 (s, 2 H, CHO), 7.67 (s, 2 H, Ar), 3.52 - 3.48 (m, 6 H, CH_3), 3.45 (s, 6 H, CH_3); ^1H NMR (400 MHz, DMSO) δ ppm = 10.03 (s, 2 H, CHO), 7.65 (s, 2 H, Ar), 3.43 (s, 6 H, CH_3), 3.41 (s, 6 H, CH_3); ^{13}C

NMR (400 MHz, DMSO) δ ppm = 187.3 (CHO), 185.5 (CS), 151.8 ($C_{Ar}S$), 133.1 ($C_{Ar}-CHO$), 123.9 ($C_{Ar}H$), 42.9 (CH_3), 38.5 (CH_3)



1-5 S,S'-(2,5-diformyl-1,4-phenylene) bis(dimethylcarbamothioate)

- 5 In a sealed tube was charged **1-4** O,O'-(2,5-diformyl-1,4-phenylene) bis(dimethylcarbamothioate) (400 mg, 1.175 mmol, 1 eq) In dry NMP (10 mL). The mixture was heated at 210°C for 9 mn under microwave. After cooling at 0°C for 1 h, the pink precipitate formed was filtrated and washed with ethanol. Drying of the solid gave a pink powder (220 mg, 55 %).
- 10 Pink solid; mp: 221.9 – 224.3 °C; HRMS (ESI) $[M + H]^+$ found 341.0625, calculated 341.0624 for $[C_{14}H_{17}N_2O_4S_2]^+$; 1H NMR (300 MHz, $CDCl_3$) δ ppm = 10.26 (s, 2 H, CHO), 8.19 (s, 2 H, Ar), 3.19 (s, 6 H, CH_3), 3.04 (s, 6 H, CH_3); 1H NMR (400 MHz, DMSO) δ ppm = 10.18 (s, 2 H, CHO), 8.03 (s, 2 H, Ar), 3.06 (br. s., 12 H, CH_3); ^{13}C NMR (400 MHz, DMSO) δ ppm = 189.4 (CHO), 162.9 (CO), 139.9 ($C_{Ar}-CHO$), 136.0 (C_{Ar}), 133.5 ($C_{Ar}S$),
- 15 36.5 (CH_3)

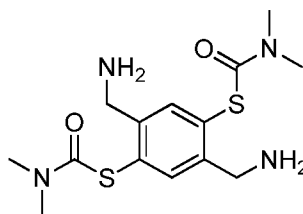


1-6 di-tert-butyl ((2,5-bis((dimethylcarbamoyl)thio)-1,4-phenylene)bis(methylene))dicarbamate

- To the solution of compound **1-5** S,S'-(2,5-diformyl-1,4-phenylene) bis(dimethylcarbamothioate) (50 mg, 0.147 mmol, 1 eq) and *t*-butyl carbamate (100 mg, 0.852 mmol, 5.8 eq) in the mixture of MeCN (1.5 mL) and $CHCl_3$ (3 mL) was added triethylsilane (0.136 μ L, 0.852 mmol, 5.8 eq) and TFA (43 μ L, 0.56 mmol, 3.8 eq). The resulting mixture was stirred 72 hours at room temperature. Following the addition of saturated aqueous $NaHCO_3$ (1 mL), the mixture was extracted with $CHCl_3$ (20 mL $\times$ 4). The
- 20 combined organic phase was dried over anhydrous Na_2SO_4 and filtered. After removing
- 25

the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane : ethyl acetate = 8:2) to give a white solid (57 mg, 71 %).

White solid; mp: 53.2-55.7°C; HRMS (ESI) [M + H]⁺ found 565.2121, calculated 565.2125 for [C₂₄H₃₈N₄NaO₆S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.60 (s, 2 H, Ar), 5.20 (br. s., 2 H, NH), 4.40 (s, 4 H, CH₂), 3.11 (br. s., 6 H, NCH₃), 3.00 (br. s., 6 H, NCH₃), 1.44 (s, 18 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 166.1 (N-CO-S, 155.9 (N-CO-O), 142.9 (C_{Ar}-CH₂), 138.2 (C_{Ar}-S), 130.3 (C_{Ar}H), 79.3 (C_{Boc}), 42.9 (CH₂), 37.1 (NCH₃), 28.7 (NCH₃), 28.4 (CH₃ Boc).



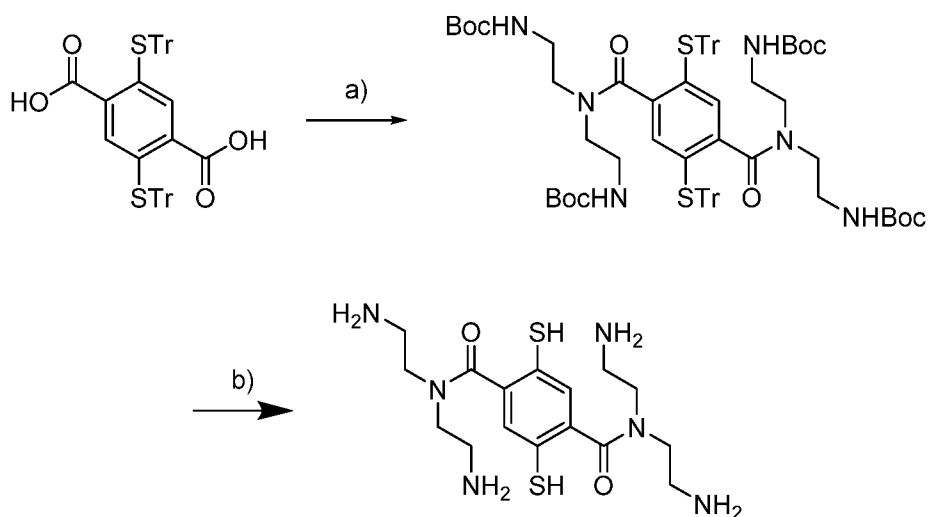
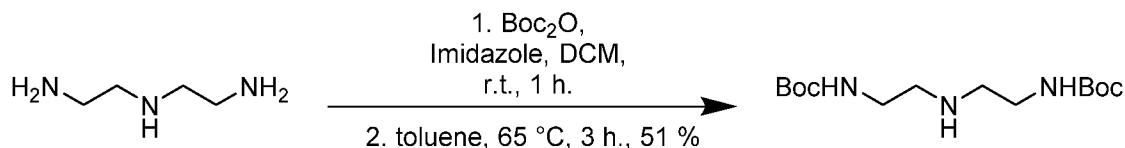
1-7 S,S'-(2,5-bis(aminomethyl)-1,4-phenylene) bis(dimethylcarbamothioate) (di TFA salt)

To a solution of **1-6** di-tert-butyl ((2,5-bis((dimethylcarbamoyl)thio)-1,4-phenylene)bis(methylene))dicarbamate (200 mg, 0.586 mmol, 1 eq) in tetrachloroethane (16.5 mL) was added TFA (5.50 mL). The mixture was stirred at room temperature for 20 mn. The mixture was then evaporated to dryness, yielding a white powder (468 mg, 81 %).

White solid; mp: 204.9-206.1 °C; HRMS (ESI) [M + H]⁺ found 343.1252, calculated 343.1257 for [C₁₄H₂₃N₄O₂S₂]⁺; ¹H NMR (400 MHz, DMSO) δ ppm = 8.36 (br. s., 2 H, H_{TFA}), 7.83 (s, 2 H, Ar), 7.20 (t, J = 51.1 Hz, 2 H, NH), 4.13 (s, 4 H, CH₂), 3.11 (br. s., 6 H, CH₃), 2.94 (br. s., 6 H, CH₃); ¹³C NMR (400 MHz, DMSO) δ ppm = 164.1 (SCON), 139.5 (C_{Ar}-CH₂), 137.9 (C_{Ar}-S), 131.2 (C_{Ar}), 40.5 (CH₂), 37.2 (CH₃), 37.2 (CH₃)

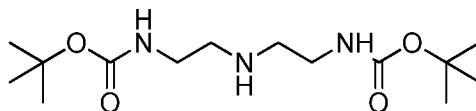
Peptide coupling for amino building block

39



Pre-functionalisation strategy for cationic target

a) HOBt, EDC, **1-14**, DCM, r.t., 48 h, 66 %; b) TFA, Et_3SiH , DCM, r.t., 3 h., 99 %



5

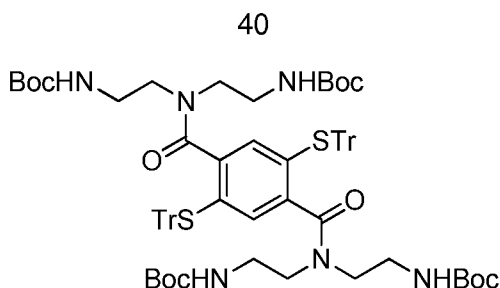
1-14 di-tert-butyl (azanediylbis(ethane-2,1-diyl))dicarbamate

To a solution of imidazole (1.566 g, 23 mmol, 1 eq) in DCM (10 mL) was added di-tert-butyl dicarbonate (5.24 g, 24 mmol, 1.05 eq) portion wise. The reaction mixture was stirred for one hour at room temperature. The reaction mixture was washed with 3*20 mL water, dried over Na_2SO_4 , filtered and the volatiles were removed under reduced pressure. The residue was dissolved in 15 mL toluene and diethylenetriamine (1.2 mL, 11 mmol, 0.5 eq) was added. The reaction mixture was stirred for two hours at 60°C. Then, 20 mL DCM was added, and the organic phase was washed with 2*20 mL water. The organic phase was dried over Na_2SO_4 , filtered and reduced under reduced pressure. The remaining residue was purified by column chromatography (SiO_2 , DCM : MeOH : NH_3 = 10:1:0.1) to give a colorless oil (1.713 g, 51 %).

10

15

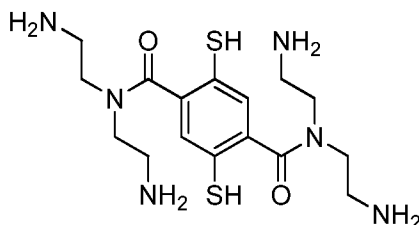
^1H NMR (300 MHz, CDCl_3) δ ppm = 5.04 (br. s., 2 H, (CO)NH), 3.24 - 3.09 (m, 4 H, $\text{CH}_2\text{-NH(CO)}$), 2.69 (t, J = 5.9 Hz, 4 H, $\text{CH}_2\text{-NH}$), 1.41 (s, 18 H, CH_3); ^{13}C NMR (300 MHz, CDCl_3) δ ppm = 156.1 (CO), 79.0 (C_{tBu}), 48.7 ($\text{CH}_2\text{-N}$), 40.3 ($\text{CH}_2\text{-NH(CO)}$), 28.3 (CH_3), data matched with literature reference



1-15 tetra-tert-butyl (((2,5-bis(tritylthio)terephthaloyl)bis(azanetriyl))tetrakis(ethane-2,1-diyl))tetracarbamate

To a solution of **2-1** 2,5-bis(tritylthio)terephthalic acid (500 mg, 0.7 mmol, 1 eq) with HOBT (247 mg, 1.82 mmol, 2.6 eq) in dry DCM (35 mL) was added at 0°C EDC (322 µL, 1.82 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of the amine **1-14** (552 mg, 1.82 mmol, 2.6 eq) in dry DCM (15 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 1:1) to give a white solid (598 mg, 66 %).

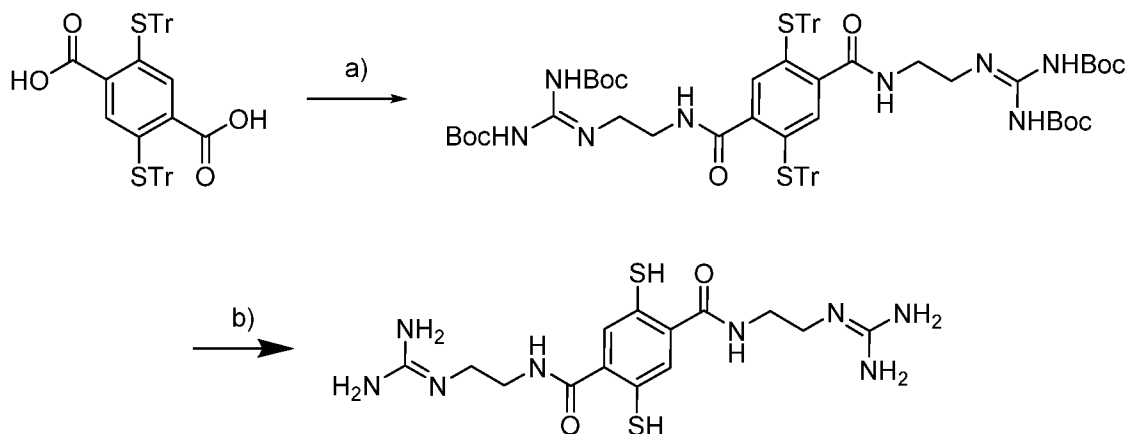
White solid; mp: degradation 225.9 °C; HRMS (ESI) [M + H]⁺ found 1285.6062, calculated 1285.6076 for [C₇₉H₈₉N₆O₁₀S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.35 - 7.22 (m, 32 H, Ar_{trityl}, Ar), 3.56 - 3.17 (m, 8 H, CH₂-NH(CO)), 2.93 - 2.59 (m, 8 H, CH₂-NH), 1.48 - 1.39 (m, 36 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 155.9 (Ar-CONH), 143.5 ((CO)_{Boc}), 129.9 (C_{Ar} trityl), 129.9 (C_{Ar} trityl), 127.9 (C_{Ar}-CONH), 128.1 (C_{Ar} trityl), 128.1 (C_{Ar} trityl), 127.5 (C_{Ar}-S), 127.4 (C_{Ar}H), 77.2 (C_{Boc}), 45.8 (CH₂-CH₂-NHBoc), 39.2 (CH₂-NHBoc), 28.4 (CH₃)



1-16 N₁,N₁,N₄,N₄-tetrakis(2-aminoethyl)-2,5-dimercaptoterephthalamide

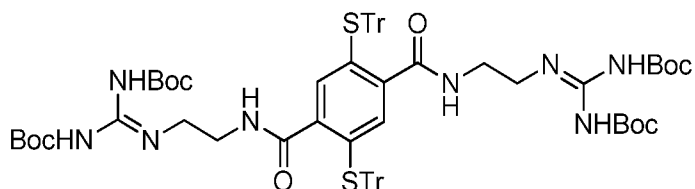
To a solution of **1-15** tetra-tert-butyl (((2,5-bis(tritylthio)terephthaloyl)bis(azanetriyl))tetrakis(ethane-2,1-diyl))tetracarbamate (200 mg, 0.155 mmol, 1 eq) in a mixture of degassed DCM/TFA (3.6 mL / 4.95 mL) was added Et₃SiH (0.45 mL, 2.9 mmol, 19 eq) was added dropwise and the mixture was stirred for 3 h. Solvent and volatile compounds were evaporate. The remaining residue was dissolved in MeOH (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness, affording a pale yellow powder as a TFA salt (133 mg, 99 %).

Pale yellow solid; mp: 86.6 – 89.1°C; HRMS (ESI) $[M + H]^+$ found 401.1785, calculated 401.1788 for $[C_{16}H_{29}N_6O_2S_2]^+$; 1H NMR (400 MHz, D_2O) δ ppm = 7.49 (s, 2 H, Ar), 3.92 - 3.80 (m, 4 H, CH_2-NH_2), 3.67 - 3.50 (m, 4 H, CH_2-NH_2), 3.35 (t, J = 6.5 Hz, 4 H, CH_2-N), 3.15 (t, J = 7.1 Hz, 4 H, CH_2-N); ^{13}C NMR (400 MHz, D_2O) δ ppm = 162.7 (CO), 129.6 ($C_{Ar}H$), 117.6 ($C_{Ar}-CO$), 114.8 ($C_{Ar}-S$), 46.2 (CH_2-NH_2), 43.0 (CH_2-NH_2), 37.0 (CH_2-N), 36.6 (CH_2-N)



Guanidine derivate functionalisation

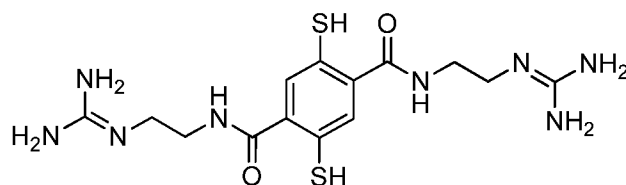
- 10 a) HOBT, EDC, 2-(2-aminoethyl)-1,3-di-Boc-guanidine, DCM, r.t., 48 h, 58 %; b) TFA, Et_3SiH , DCM, r.t., 3 h., 83 %



1-17 (2,5-bis(tritylthio)terephthaloyl)bis(2-(2-aminoethyl)-1,3-di-Boc-guanidine)

- 15 To a solution of **2-1** 2,5-bis(tritylthio)terephthalic acid (500 mg, 0.7 mmol, 1 eq) with HOBT (247 mg, 1.82 mmol, 2.6 eq) in dry DCM (35 mL) was added at 0°C EDC (322 μ L, 1.82 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N_2 . A solution of the 2-(2-aminoethyl)-1,3-di-Boc-guanidine (550 mg, 1.82 mmol, 2.6 eq) in dry DCM (15 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was
- 20 added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na_2SO_4 and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO_2 , cyclohexane: ethyl acetate = 1:1) to give a white solid (529 mg, 58 %).

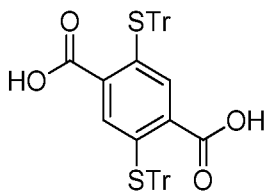
White solid; mp: degradation 115.0 °C; HRMS (ESI) $[M + 2H]^{2+}$ found 642.2853, calculated 642.2870 for $[C_{72}H_{84}N_8O_{10}S_2]^+$; 1H NMR (400 MHz, $CDCl_3$) δ ppm = 11.45 (s, 2 H, NHBoc), 8.40 (t, J = 6.1 Hz, 2 H, NHCO), 7.27 - 7.25 (m, 30 H, Ar_{trityl}), 6.90 (s, 2 H, Ar), 5.07 (t, J = 5.7 Hz, 2 H, NHBoc), 3.35 (s, 4 H, $N-CH_2$), 3.24 - 3.17 (m, 4 H, CH_2-NHCO), 1.49 (s, 36 H, tBu); ^{13}C NMR (400 MHz, $CDCl_3$) δ ppm = 166.8 (C_{BocO}), 163.4 (C_{BocO}), 156.4 (CONH), 153.0 (NNC_{guanidine}N), 143.6 ($C_{Artrityl}$), 139.5 ($C_{Ar-CONH}$), 132.8 (C_{Ar-S}), 132.5 ($C_{Ar trityl}$), 130.1 ($C_{Ar trityl}$), 127.9 ($C_{Ar trityl}$), 127.1 (C_{ArH}), 83.3 (C_{tBu}), 79.2 (C_{tBu}), 77.2 (C_{tBu}), 71.0 ($C_{trityl-S}$), 39.9 (CONH- CH_2), 39.0 (CH_2 -guanidine), 28.3 (CH_3), 28.2 (CH_3), 28.0 (CH_3), 26.9 (CH_3)



1-18 N₁,N₃-bis(2-((diaminomethylene)amino)ethyl)-2,5-dimercaptoterephthalamide

To a solution of **1-17** (2,5-bis(tritylthio)terephthaloyl)bis(2-(2-aminoethyl)-1,3-di-Boc-guanidine) (340 mg, 0.262 mmol, 1 eq) in a mixture of degassed DCM/TFA (6.1 mL / 8.4 mL) was added Et_3SiH (0.76 mL, 4.9 mmol, 19 eq) was added dropwise and the mixture was stirred for 3 h. Solvent and volatile compounds were evaporate. The remaining residue was dissolved in MeOH (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness, affording an orange powder as a TFA salt (136 mg, 83 %).

Orange solid; mp: degradation 226.4°C; HRMS (ESI) $[M + H]^+$ found 399.1367, calculated 399.1380 for $[C_{14}H_{23}N_8O_2S_2]^+$; 1H NMR (400 MHz, D_2O) δ ppm = 7.48 (s, 2 H, Ar), 3.59 - 3.54 (m, 4 H, $N-CH_2$), 3.48 - 3.42 (m, 4 H, CH_2-NHCO); ^{13}C NMR (400 MHz, D_2O) δ ppm = 170.2 ($C_{guanidine}$), 163.4 (C_{TFA}), 162.7 (C_{TFA}), 157.0 (CONH), 135.6 ($C_{Ar-CONH}$), 130.2 (C_{ArH}), 128.1 (C_{Ar-S}), 117.7 (C_{TFAF_3}), 114.8 (C_{TFAF_3}), 40.4 (CH_2-N), 38.5 (CH_2-N)



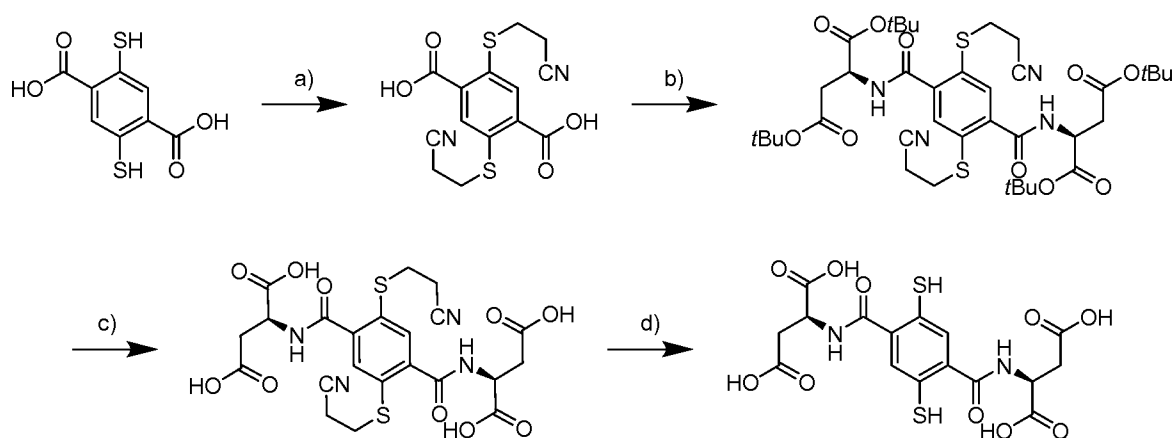
2-1 2,5-bis(tritylthio)terephthalic acid

A solution of **A** (500 mg, 2.17 mmol, 1 eq) and trityl chloride (1.575 g, 5.65 mmol, 2.6 eq) in degassed and dry DMF (15 mL) was stirred at room temperature under an inert atmosphere for 48 h. The resulting precipitate was filtrated, washed with water (3*20 mL), chloroform (3*20 mL) and cyclohexane (3*20 mL). The product was dried under vacuum, yielding compound **1-9** as a yellow solid (1.251 mg, 81 %).

Yellow solid; mp: degradation 405.0 °C; HRMS (ESI) $[M + Na]^+$ found 737.1791, calculated 737.1791 for $[C_{46}H_{34}NaO_4S_2]^+$; 1H NMR (400 MHz, DMSO) δ ppm = 7.31 - 7.21 (m, 30 H, Ar_{trityl}), 7.14 (s, 2 H, Ar); ^{13}C NMR (400 MHz, DMSO) δ ppm = 166.2 (CO), 147.7 ($C_{Ar-trityl}-C$), 143.0 ($C_{trityl}-S$), 129.6 ($C_{Ar-trityl}$), 127.9 ($C_{Ar}-S$), 127.7 ($C_{Ar-trityl}$), 127.5 ($C_{Ar-trityl}$), 126.6 ($C_{Ar}-CO$)

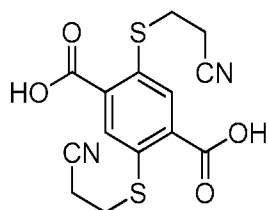
Building block for methionine oxidation recognition

Cyanopropyl protection



Cyanopropyl protection strategy

a) 3-bromopropenenitrile, NaH, THF, reflux, 48 h. 94 %; b) HOBT, EDC, DIPEA, L-Aspartic acid di-tert-butyl ester hydrochloride, DCM, r.t., 48 h, 78 %; c), TFA, Et_3SiH , DCM, r.t., 16 h., 55 %; d) CsOH, THF/MeOH, r.t., 3 h., 99 %

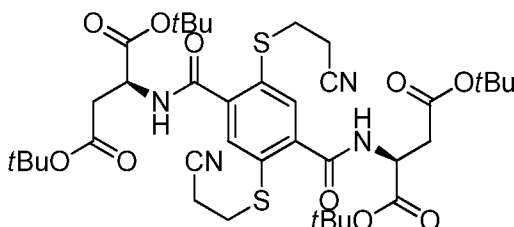


2-4 2,5-bis((2-cyanoethyl)thio)terephthalic acid

To a solution of NaH (313 mg, 13 mmol, 6 eq) in degassed and dry THF (5 mL) was added a solution of **A** (500 mg, 2.17 mmol, 1 eq) in degassed and dry THF (10 mL). The mixture was stirred at room temperature under an inert atmosphere for 1 h. Then, 3-bromopropenenitrile (3.6 mL, 43.4 mmol, 20 eq) was added dropwise. The mixture was then heated at reflux under inert atmosphere for 48 h. The solvent was evaporated, and chloroform was added (50 mL). After sonication, the precipitate was filtrated, and sonicate

with 60 mL of HCl 1 M. The solid was then filtrated, and washed with distilled water (100 mL). The solid was dried under vacuum to yield a yellow powder (688 mg, 94 %).

Yellow solid; mp: 234.6 - 237.4 °C; HRMS (ESI) $[M + Na]^+$ found 359.0133, calculated 359.0131 for $[C_{14}H_{12}NaO_4S_2]^+$; 1H NMR (400 MHz, DMSO) δ ppm = 13.59 (br. s., 2 H, COOH), 7.80 (s, 2 H, Ar), 3.26 (t, J = 6.8 Hz, 4 H, S-CH₂), 2.87 (t, J = 6.8 Hz, 4 H, CH₂-CN); ^{13}C NMR (400 MHz, DMSO) δ ppm = 166.7 (CO), 134.1 (C_{Ar}-CO), 132.9 (C_{Ar}-S).5 (C_{Ar}H), 119.2 (CN), 27.2 (CH₂-S), 16.6 (C-CN)

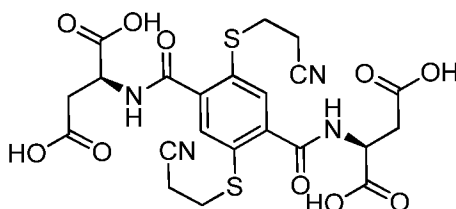


2-5 (2S,2'S)-tetra-tert-butyl 2,2'-((2,5-bis((2-

cynoethyl)thio)terephthaloyl)bis(azanediyl))disuccinate

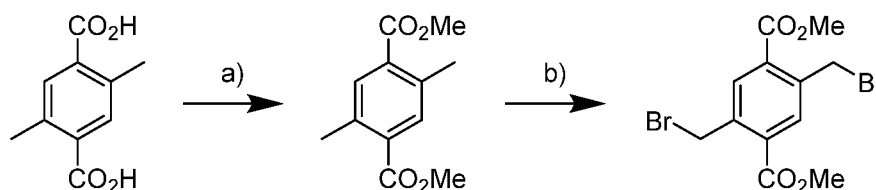
To a solution of **2-4** 2,5-bis((2-cyanoethyl)thio)terephthalic acid (100 mg, 0.297 mmol, 1 eq) with HOBt (104 mg, 0.773 mmol, 2.6 eq) in dry DCM (22 mL) was added at 0°C EDC (134 μ L, 0.773 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of the amine hydrochloride (218 mg, 0.773 mmol, 2.6 eq) and DIPEA (0.400 mL, 2.33 mmol, 7.85 eq) in dry DCM (11 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 1:1) to give a white solid (182 mg, 78 %).

White solid; mp: 134.3 – 137.2 °C; HRMS (ESI) $[M + H]^+$ found 791.3318, calculated 791.3354 for $[C_{38}H_{55}N_4O_{10}S_2]^+$; 1H NMR (400 MHz, CDCl₃) δ ppm = 7.79 (s, 2 H, Ar), 7.61 (d, J = 8.1 Hz, 2 H, NH), 4.86 (td, J = 4.2, 8.8 Hz, 2 H, CH-N), 3.26 - 3.12 (m, 4 H, S-CH₂), 2.99 (dd, J = 4.2, 17.4 Hz, 2 H, AB system, CH₂-CO₂tBu), 2.87 (dd, J = 4.4, 17.4 Hz, 2 H, AB system, CH₂-CO₂tBu), 2.67 (t, J = 7.3 Hz, 4 H, CH₂-CN), 1.49 (s, 18 H, CH₃), 1.45 - 1.44 (m, 18 H, CH₃); ^{13}C NMR (400 MHz, CDCl₃) δ ppm = 170.3 (CH₂-CO-OtBu), 169.3 (CH-CO-OtBu), 165.6 (Ar-CONH), 139.4 (C_{Ar}-CONH), 132.8 (C_{Ar}-S), 132.2 (C_{Ar}H), 117.6 (CN), 82.7 (C_{tBu}, CH-COOtBu), 81.6 (C_{tBu}, CH₂-COOtBu), 49.7 (CH), 37.2 (CH₂-CH), 30.6 (CH₂-S), 28.0 (CH₃, CH₂-COOtBu), 27.9 (CH₃ CH-COOtBu), 18.0 (C-CN)

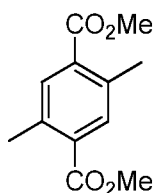


2-6 (2S,2'S)-2,2'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(azanediyl))disuccinic acid

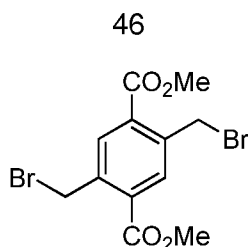
To a solution of **2-5** (2S,2'S)-tetra-tert-butyl 2,2'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(azanediyl))disuccinate (100 mg, 0.126 mmol, 1 eq) in dry DCM (12 mL) was added TFA (8 mL) and triethylsilane (0.200 mL, 1.26 mmol, 10 eq) at 0°C. The mixture was stirred at room temperature for 4 h. Solvent and volatile compounds were evaporate. The remaining residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give a white solid (39 mg, 55 %). White solid, mp: 208.7 decomposition; HRMS (ESI) [M + Na]⁺ found 589.0657, calculated 589.0670 for [C₂₂H₂₂N₄NaO₁₀S₂]⁺; ¹H NMR (400 MHz, MeOD) δ ppm = 7.66 (s, 2 H, Ar), 4.99 - 4.97 (m, 2 H, CH-N), 3.25 (dt, J = 1.5, 7.0 Hz, 4 H, S-CH₂), 3.05 (dd, J = 5.3, 16.9 Hz, 2 H, AB system, CH₂-CO₂H), 2.95 (dd, J = 7.0, 16.7 Hz, 2 H, AB system, CH₂-CO₂H), 2.78 (t, J = 7.0 Hz, 4 H, CH₂-CN); ¹³C NMR (400 MHz, MeOD) δ ppm = 174.1 (CH₂-COOH), 173.8 (CH-COOH), 169.6 (Ar-CONH), 141.7 (C_{Ar}-CONH), 133.4 (C_{Ar}-S), 132.3 (C_{Ar}-H), 119.8 (CN), 51.0 (CH), 36.9 (CH₂-CH), 31.2 (CH₂-S), 18.7 (C-CN)



a) SOCl₂, MeOH, reflux, 16 h., 99 %; b) NBS, AIBN, DCM, reflux, 16 h, 61 %;

**4-3 dimethyl 2,5-dimethylterephthalate**

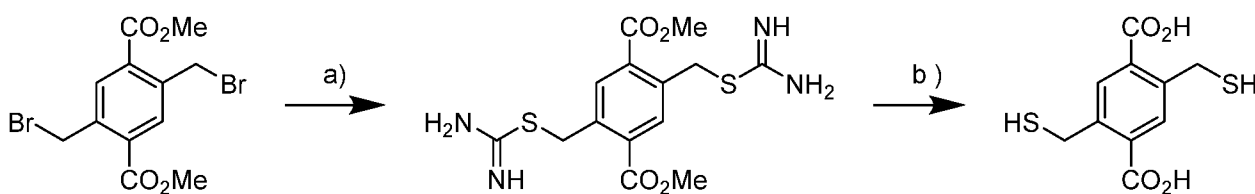
To a suspension of **4-1** 2,5-dimethylterephthalic acid (800 mg, 4.12 mmol, 1 eq) in dry methanol (12.5 mL) was added dropwise at 0°C thionyl chloride (1.25 mL, 17.1 mmol, 3.5 eq). The mixture was heated at reflux and stirred overnight under inert atmosphere. After cooling, the resulting solid was washed with methanol (10 mL) and twice with hexane (2*10 mL). The solid was dried under vacuum to give a white powder (915 mg, 99 %). White solid; mp: 113.7-115.1 °C; HRMS (ESI) [M + H]⁺ found 223.0954, calculated 223.0965 for [C₁₂H₁₅O₄]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.76 (s, 2 H, Ar), 3.91 (s, 6 H, COOCH₃), 2.57 (s, 6 H, Ar-CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 167.5 (CO), 137.0 (C_{Ar}-CH₃), 133.5 (C_{Ar}-CO), 132.4 (C_{Ar}-H), 52.0 (OCH₃), 20.9 (Ar-CH₃)



4-4 dimethyl 2,5-bis(bromomethyl)terephthalate

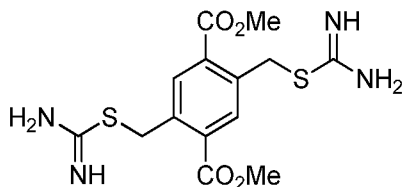
A solution of **4-3** dimethyl 2,5-dimethylterephthalate (600 mg, 2.70 mmol, 1 eq), NBS (1.05 g, 6 mmol, 2.2 eq), AIBN (36 mg, 0.220 mmol, 0.08 eq) in dry DCM (7.5 mL) was heated at reflux until conversion of starting material. The reaction is monitored by NMR control. After cooling, the organic phase was washed a solution of NaHCO₃ (2*10 mL) and brine (2*10 mL). The organic phase was dried with MgSO₄ and evaporate to dryness. The crude mixture was purified by column chromatography (SiO₂, cyclohexane: DCM = 1:1) to give a white powder (360 mg, 61 %).

White solid; mp: 109.3-111.7 °C; HRMS (ESI) [M + Na]⁺ found 400.8989, calculated 400.8995 for [C₁₂H₁₂Br₂NaO₄]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 8.06 (s, 2 H, Ar), 4.93 (s, 4 H, CH₂), 3.98 (s, 6 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.7 (CO), 139.4 (C_{Ar}-CH₂), 134.4 (C_{Ar}H), 132.3 (C_{Ar}-CO), 52.7 (CH₂), 30.0 (CH₃)



Scheme 1: 4-8 synthesis

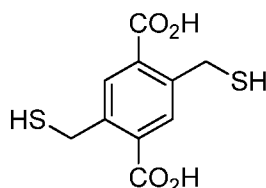
a) thiourea, MeOH, r.t., 16h., 88 % ; b) KOH, MeOH/H₂O, reflux, 2 h., 44 %



4-7 dimethyl 2,5-bis((carbamimidoylthio)methyl)terephthalate

A suspension of **4-4** dimethyl 2,5-bis(bromomethyl)terephthalate (275 mg, 0.715 mmol, 1 eq) and thiourea (115 mg, 1.66 mmol, 2.33 eq) in MeOH (11 mL) was stirred at room temperature for 16 h. The solvent was evaporated, and the crude mixture was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 8:2). Evaporation of the solvent gave a white powder (234 mg, 88 %).

White solid; mp: degradation 250.3 °C; HRMS (ESI) $[M + H]^+$ found 371.0836, calculated 371.0842 for $[C_{14}H_{19}N_4O_4S_2]^+$; 1H NMR (400 MHz, DMSO) δ ppm = 9.40 (br. d., 6 H, NH_2), 8.16 (s, 2 H, Ar), 4.76 (s, 4 H, CH_2), 3.91 (s, 6 H, CH_3); ^{13}C NMR (400 MHz, DMSO) δ ppm = 169.0 (CO), 165.6 (C=NH), 136.9 ($C_{Ar}-CH_2$), 133.4 ($C_{Ar}H$), 132.3 ($C_{Ar}-CO_2$), 52.9 (CH_3), 32.4 (CH_2-S)



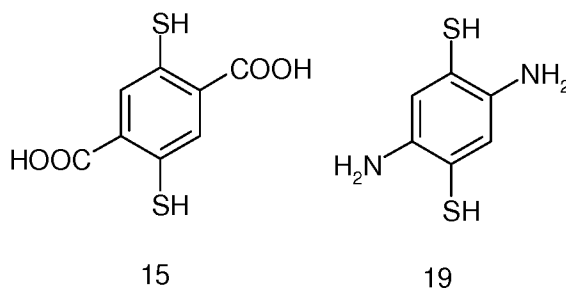
4-8 2,5-bis(mercaptomethyl)terephthalic acid

A solution of **4-7** dimethyl 2,5-bis((carbamimidoylthio)methyl)terephthalate (100 mg, 0.27 mmol, 1 eq) in degassed 2 M KOH (896 mg, 16 mmol, 60 eq) in MeOH/H₂O (1:1, 8 mL) was refluxed under an inert atmosphere for 2 h. The reaction mixture was cooled in ice, and concentrated TFA (1 mL) was added until pH 1. The methanol was evaporated, and the water phase was cooled. A white yellow precipitate was formed, filtered, and washed with water, yielding compound **4-8** as a white solid (31 mg, 44 %).

White solid; mp: degradation 314.1 °C; HRMS (ESI) $[M + Na]^+$ found 280.9902, calculated 280.9913 for $[C_{10}H_{10}NaO_4S_2]^+$; 1H NMR (400 MHz, MeOD) δ ppm = 7.93 (s, 2 H, Ar), 4.06 (s, 4 H, CH_2); ^{13}C NMR (400 MHz, MeOD) δ ppm = 169.5 (CO), 144.0 ($C_{Ar}-CH_2$), 134.4 ($C_{Ar}H$), 133.7 ($C_{Ar}-CO_2H$), 27.4 (CH_2)

1.2 Synthesis of the oligomers

Examples of Oligomers synthesis/ self-assembly from compounds **15** or **19**



10 mM stock solution of building block **15** or **19** were prepared by dissolution in 200 mM Tris buffer pH = 7.4 in Milli-Q water. In a 2mL HPLC vial equipped with a stir bar, 400 μ L of each of these solutions was mixed with either 400 μ L of a 10 mM template solution in the same buffer and 200 μ L of buffer solution or 600 μ L of buffer solution. As an alternative, building blocks can be suspended in Milli-Q water and 1M solution

of metal hydroxide (lithium, sodium, potassium) can be used to adjust carefully the pH to the desired value.

Tetramers 15_4 and 19_4 are respectively obtained from building blocks 15 and 19.

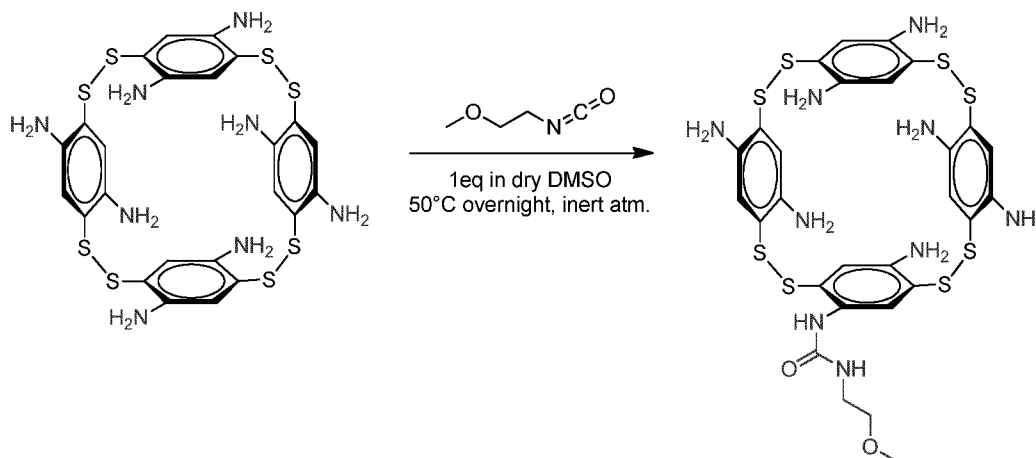
Synthesis can be conducted in accordance with Skowron et al J.Org.Chem. 2016, 81, 654-661.

For hydrophobic building blocks, libraries can be alternatively set in organic solvent (CDCl₃, dmso, etc..) or mixed media . Concentrations are typically millimolar (4 mM) and a catalytic amount of base is required, (0.2 mM Et₃N) both in the absence and presence of template (typically 0.25 eq).

1.3 Synthesis of the functionalized oligomers

1.3.1 Urea functionalization of Dyn[4]arene amine

1.3.1.i with 2-methoxyethyl isocyanate

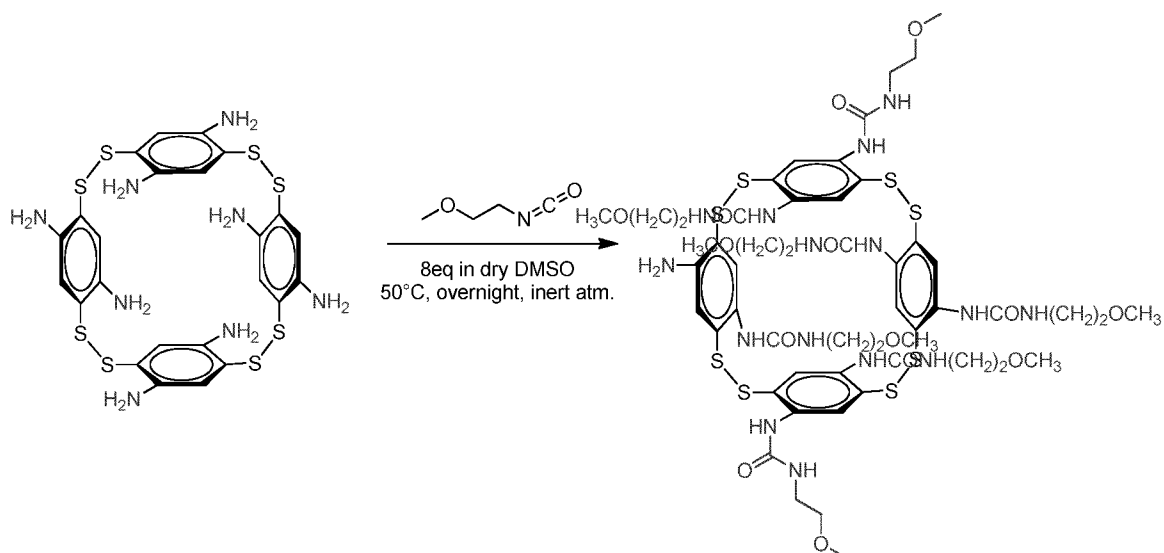


Scheme A. Mono-functionalization.

To a solution of **19**₄ (50mg; 0.073mmol) in 5mL of dry DMSO was added drop by drop over 30mn a solution of 2-methoxyethyl isocyanate (1eq) in 10mL of dry DMSO under inert atmosphere. The solution was stirred at 50°C overnight protected from light. Then an aliquot was taken and send to mass spectrometry analysis. The formation of the mono-urea is observed.

LRMS (ESI, positive):

m/z calculated for $C_{28}H_{32}N_9O_2S_8^+(MH^+)=782.04$; measured= 781.9



Scheme A'. Multi-functionalization.

To a solution of **19₄** (50mg; 0.073mmol) in 5mL of dry DMSO was added the 2-methoxyethyl isocyanate (60mg; 0.58mmol; 62μL) under inert atmosphere. The solution was stirred at 50°C overnight, protected from light. Then an aliquot was taken and send to mass spectrometry analysis. The formation of three species is observed: 5, 6 and 7 ureas.

LRMS (ESI, positive):

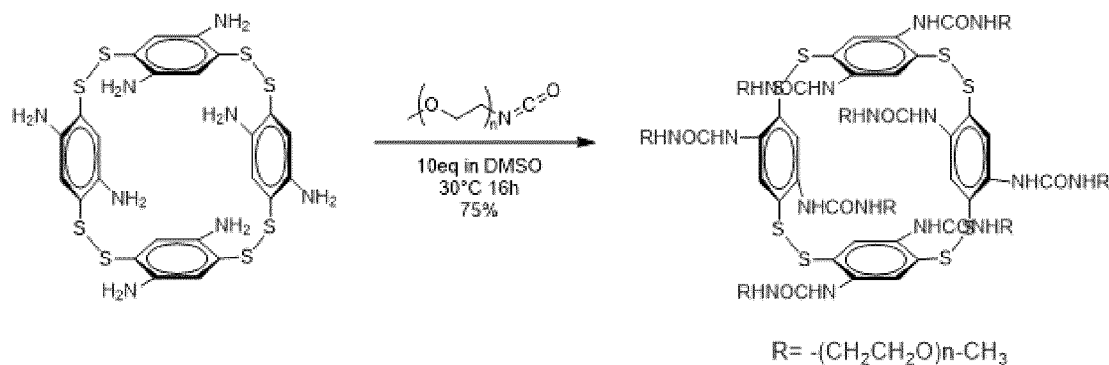
m/z calculated for $C_{52}H_{73}N_{15}NaO_{14}S_8^+(MNa^+)=1410.31$; measured= 1410.1

m/z calculated for $C_{48}H_{66}N_{14}NaO_{12}S_8^+(MNa^+)=1309.26$; measured= 1309.1

m/z calculated for $C_{44}H_{59}N_{13}NaO_{10}S_8^+(MNa^+)=1208.22$; measured= 1208.1

m/z calculated for $C_{44}H_{60}N_{13}O_{10}S_8^+(MH^+)=1186.23$; measured= 1186.0

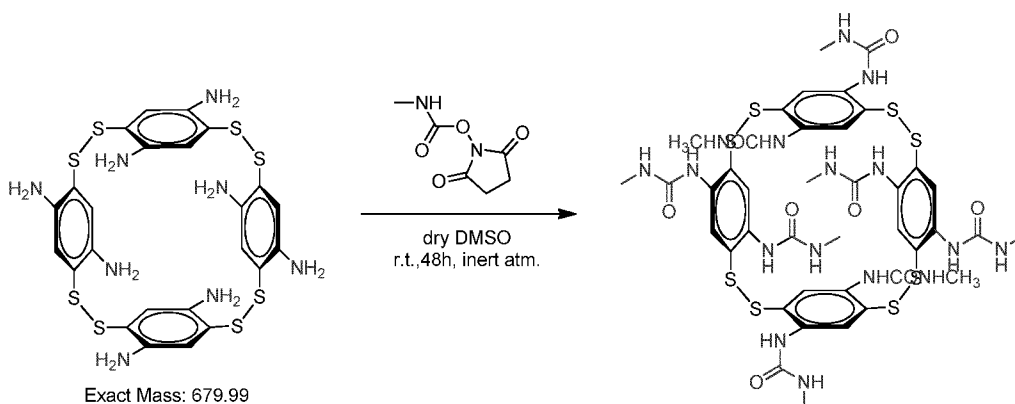
1.3.1.ii with PEG-isocyanate



Scheme B. PEG-isocyanate.

To a solution of **19₄** (8.5mg; 0.025mmol) in 5mL of dry DMSO was added the PEG-isocyanate (500mg; 0.100mmol; MW 5000) under inert atmosphere. The solution was stirred at 30°C overnight, protected from light. Then the solution was poured into 500mL of water, the residual precipitate was filtered off and the filtrate lyophilized. This step was repeated until no precipitate appears when water is added. Finally a light green powder water-soluble is observed.

1.3.1.iii. with N-methyl-succinimidyl



Scheme B'. Multi-functionalization.

To a solution of **19₄** (50mg; 0.073mmol) in 5mL of dry DMSO was added by solid portions the N-succinimidyl N-methylcarbamate (10eq) under inert atmosphere. The solution was stirred at room temperature during 48h protected from light. Then an aliquot was taken and send to mass spectrometry analysis. We observed the formation of multi functionalized species.

LRMS (ESI, positive):

m/z calculated for $C_{26}H_{28}N_9OS_8^+(MH^+)$ = 738.02; measured = 738.0 (1 urea)

m/z calculated for $C_{28}H_{31}N_{10}O_2S_8^+(MH^+)$ = 795.04; measured = 795.0 (2 ureas)

m/z calculated for $C_{30}H_{34}N_{11}O_3S_8^+(MH^+)$ = 852.06; measured = 852.0 (3 ureas)

m/z calculated for $C_{32}H_{37}N_{12}O_4S_8^+(MH^+)$ = 909.08; measured = 909.0 (4 ureas)

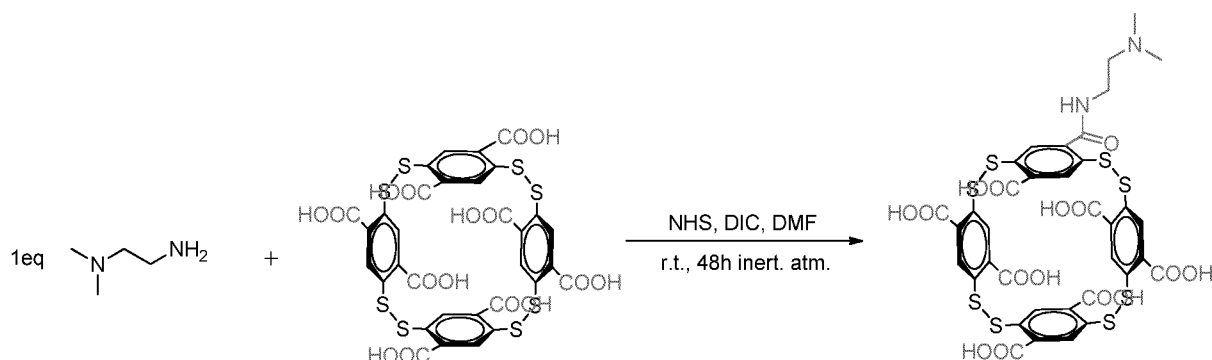
m/z calculated for $C_{34}H_{40}N_{13}O_5S_8^+(MH^+)$ = 966.1; measured = 966.0 (5 ureas)

m/z calculated for $C_{36}H_{43}N_{14}O_6S_8^+(MH^+)$ = 1023.13; measured = 1023.0 (6 ureas)

m/z calculated for $C_{38}H_{46}N_{15}O_7S_8^+(MH^+)$ = 1080.15; measured = 1080.0 (7 ureas)

m/z calculated for $C_{40}H_{49}N_{16}O_8S_8^+(MH^+)$ = 1137.17; measured = 1137.0 (8 ureas)

1.3.2 Functionalization of carboxy oligomers



Scheme C.

LRMS (ESI) m/z calculated for $C_{36}H_{27}N_2O_{15}S_8$: 982.9, measured: 983.0

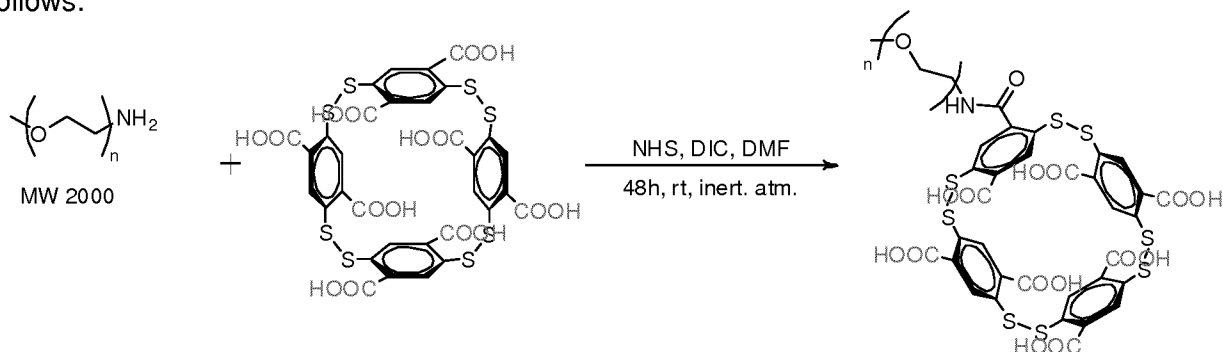
LRMS (ESI) m/z calculated for $C_{40}H_{37}N_4O_{14}S_8$: 1053.0, measured: 1053.0

15 In a round bottom flask was poured the dyn[4]arene carboxylic acid (0.100g; 0.099mmol), followed by 10mL of dry dimethylformamide and placed under inert atmosphere. Then were added drop by drop the N,N'-diisopropylcarbodiimide (0.012g; 0.099mmol; 15.3 μ L), the N-hydroxysuccinimide (0.011g; 0.099mmol) and the N,N-dimethylethylenediamine (8.73mg; 0.099mmol; 10.8 μ L) previously dissolved in dry DMF (5mL). The reaction mixture was stirred 48h at room temperature to give an heterogeneous solution.

20 The reaction was monitored by HPLC reverse phase and the starting material disappearance was observed giving a complex mixture.

The DMF was then co-evaporated with heptane and the crude product was analyzed by mass spectrometry to give two coupling products (1 amide and 2 amides).

25 A similar procedure can be applied to achieve functionalization with PEG derivatives, as follows:



Scheme C'.

1.3.3 DCC study

Library preparation

In a typical experiment, the building blocks were individually dissolved in an aqueous buffer (200mM) at a given pH or in organic media to obtain building block stock solutions at [2-10] mM. The stock solution are freshly prepared and quickly engaged in the DCCs: in 1.5 mL vials equipped with a turbulent were placed 500 μ L of the building block 1 and 500 μ L of building block 2 solution stocks. The vials were left open to the air and stirred at room temperature for 48 hours.

HPLC analysis

Method A: column Agilent Eclipse Plus C₈, 3.5 μ m, 4.6 x 150 mm, V_{inj} = 3 μ L, flow 1 mL/mn. Gradient water + 0.05 % TFA / MeCN + 0.05 % TFA in 25 mn method (t₀ 80/20, t₅ 60/40, t₈ 0/100).

Method B: Agilent Poroshell 120 EC C₈, 2.7 μ m, 2.1 x 50 mm flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 22 mn method (t₀ 100/0, t₁₅ 70/30, t₂₀ 100/0, t₂₂ 100/0).

UPLC/HRMS analysis

Analysis are performed using a U3000 Thermo Fisher Scientific / QTOF Impact II Bruker system. The column used was an Agilent Poroshell 120 EC C₈, 2.7 μ m, 5 x 2.1 mn, V_{inj} = 3 μ L, flow 0.8 mL/mn, Gradient water + 0.1 % formic acid / MeCN:MeOH (50/50) (or MeCN) + 0.1 % formic acid. Mass spectrometry parameters depend on the analysis, UV detection performed at 250 and 350 nm.

The benzylic thiol **4-8** was engaged in a DCC study with **15** (I''-A) building block

Library number	4-8 (eq)	(I''-A) (eq)	Spermine (eq)	Observed species at 72 h in LC/MS
1	1	/	/	Cycle: [(4-8) ₂ -3S+2O-H ₂], (4-8) ₂ (2 species, 13 % for one specie), [(4-8) ₂ +2O], [(4-8) ₂ -S+3O], [(4-8) ₂ +S-H ₂], [(4-8) ₃ -3S+2O-H ₂], [(4-8) ₃ -S+3O], [(4-8) ₃ -2S+O], [(4-8) ₃ -S], (4-8) ₃ (2 species), (4-8) ₄ , [(4-8) ₄ -2S+2O-H ₂], (4-8) ₅ , [(4-8) ₅ -2S+2O-H ₂], [(4-8) ₅ -S-H ₂], [(4-8) ₅ +S], (4-8) ₆ , [(4-8) ₆ +S], (4-8) ₈ , [(4-8) ₈ +S]

2	1	/	0.25	Cycle: (4-8) ₂ (13 %), [(4-8) ₂ +S-H ₂], [(4-8) ₂ +S], (4-8) ₃ (19 %), [(4-8) ₃ +S-H ₂], [(4-8) ₄ -S], (4-8) ₄ (18 %), [(4-8) ₄ +S], (4-8) ₅ (32 %), [(4-8) ₅ +S], (4-8) ₆ , (4-8) ₆ +S, (4-8) ₉
3	1	1	/	Cycle: [A(4-8)-S+3O], [A(4-8)+O], (4-8) ₂ , [A(4-8) ₂], (4-8) ₃ , [A(4-8) ₂ -S+3O], [A(4-8) ₃], (4-8) ₄ , [A ₂ (4-8) ₃], (4-8) ₅ , [A ₂ (4-8) ₄], [A(4-8) ₅], (4-8) ₆ , [A(4-8) ₆], [A(4-8) ₇], (4-8) ₉ , Linear: [A ₂ +4O], [A ₂ (4-8)+4O] No major species
4	1	1	0.5	Cycle: [A ₂ +4O], [A(4-8)-S+3O], [A(4-8)], [(4-8) ₂ +3O], [A(4-8) ₂ +O], (4-8) ₃ (13 %), [A ₂ (4-8)], [A(4-8) ₄ -S+3O], (4-8) ₄ (33 %), [A(4-8) ₃], (4-8) ₅ (22 %), [(4-8) ₅ +S], [A ₁ (4-8) ₅], (4-8), [A(4-8) ₇], (4-8) ₉ , (4-8) ₁₀

DCC study at 72 h between 4-8 and A, in Tris 200 mM pH 7.4, at a total concentration of 4 mM of building blocks

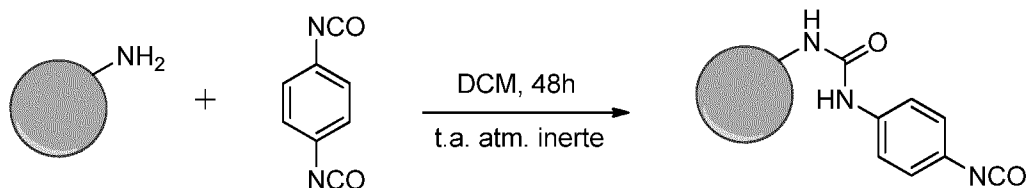
–S: sulfur extrusion on a benzylic position, +S: trisulfide, sulfur addition + O: oxidation, +H₂: insaturation formation, for some heteromacrocycles, several arrangements are possible but only one specie was detected for each case, if necessary, major compounds are highlighted, relative % given are based on total integration

These libraries illustrated the power of DCC. Many structures are created in 72 h. Kinetics of oxidation was found to be slow in Tris buffer. 4-8 alone gave macrocycles of various sizes, among them 3, 4 and 5. Combined with spermine, size 3 and 4 were slightly favoured and stabilised.

1.4 Grafting on solid phase

1.4.1 with Isocyanate

A. 1,4-phenylene diisocyanate



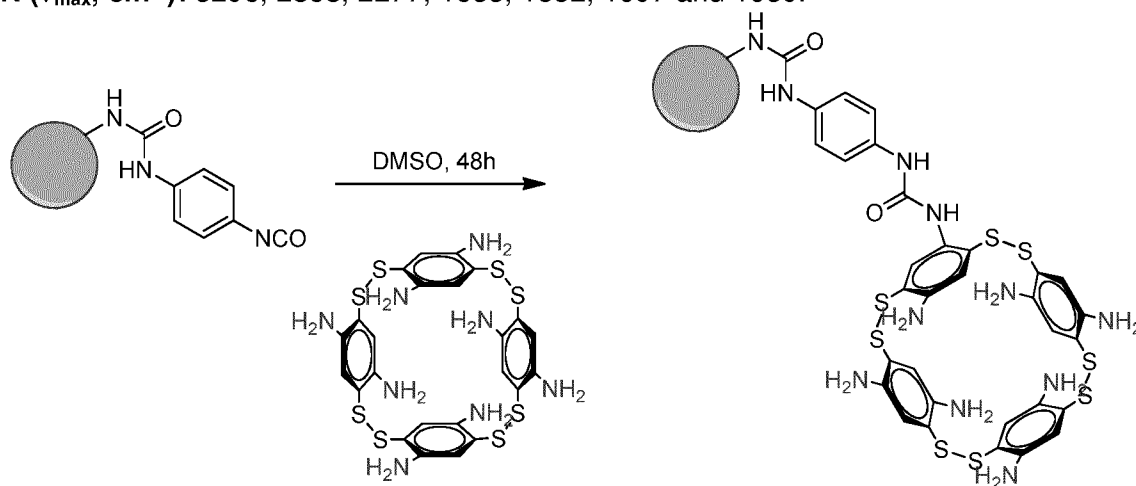
Scheme D.

PEGA-NH₂ resin (0.4mmol/g) was dried under vacuum at 200°C in order to eliminate the residual water. The glassware was also dried on the oven and the round bottom flask flushed with nitrogen. 100mg (0.04mmol) of resin were introduced into the flask and soaked with dry acetonitrile (2mL). Then the 1,4-phenylene diisocyanate (10eq; 0.4mmol; 56mg; 51uL) was diluted in 1mL of dry acetonitrile and slowly added (over 30mn). The mixture was stirred at room temperature during 3 days.

Then the resin was filtered off and washed successively with dry acetonitrile (3x10mL), tetrahydrofuran (3x10mL) and dichloromethane (3x10mL). The light beige beads obtained were dried over vacuum.

A Kaiser test was performed to give a positive result, weaker blue color, attesting the presence of residual amine.

IR ($\nu_{\max}$, cm⁻¹): 3296, 2868, 2277, 1633, 1552, 1097 and 1039.



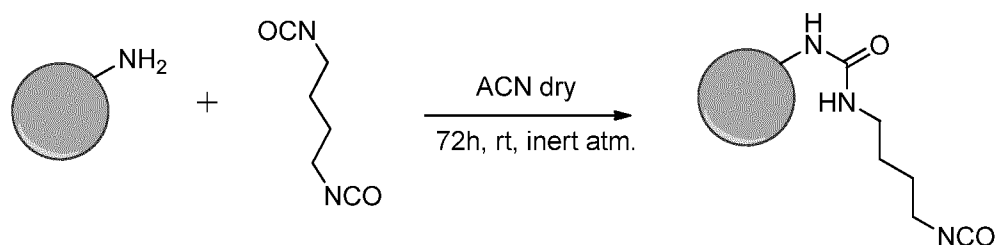
Scheme D'.

The previous beads (100mg) were placed into a dry round bottom flask flushed with nitrogen. And dry DMSO was added (5mL) followed by dyn[4]arene amino by solid portions (eq; mmol; mg). Then the mixture was stirred at room temperature during 2 days protected from light.

Then the resin was filtered off and washed with dry DMSO (4x10mL) and acetonitrile (2x10mL). The green beads obtained were dried over vacuum.

IR (ν_{max} , cm^{-1}): 3292, 2869, 1632, 1517, 1093 and 1003.

B. 1,4-diisocyanatobutane



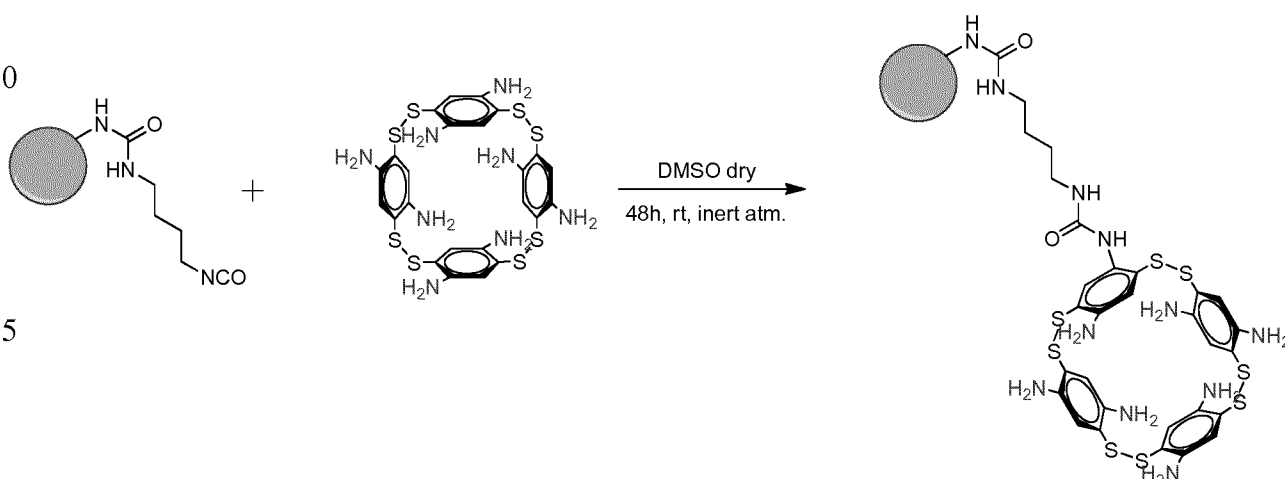
Scheme E.

PEGA-NH₂ resin (0.4mmol/g) was dried under vacuum at 200°C in order to eliminate the residual water. The glassware was also dried on the oven and the round bottom flask flushed with nitrogen. 100mg (0.04mmol) of resin were introduced into the flask and soaked with dry acetonitrile (2mL). Then the 1,4-diisocyanatobutane (10eq; 0.4mmol; 56mg; 51uL) was diluted in 1mL of dry acetonitrile and slowly added (over 30mn). The mixture was stirred at room temperature during 3 days.

Then the resin was filtered off and washed successively with dry acetonitrile (3x10mL), tetrahydrofuran (3x10mL) and dichloromethane (3x10mL). The light beige beads obtained were dried over vacuum.

A Kaiser test was performed to give a positive result, weaker blue color, attesting the presence of residual amine.

IR (ν_{max} , cm^{-1}): 3317, 2939, 2862, 2350, 1610, 1556, 1271 and 1079.

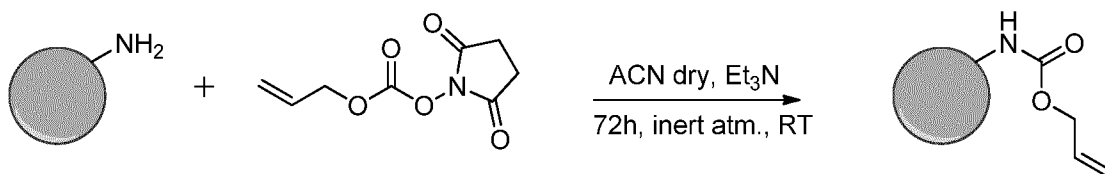


Scheme E'.

The previous beads (100mg) were placed into a dry round bottom flask flushed with nitrogen. And dry DMSO was added (5mL) followed by dyn[4]arene amino by solid portions (eq; mmol; mg). Then the mixture was stirred at room temperature during 2 days protected from light.

Then the resin was filtered off and washed with dry DMSO (4x10mL) and acetonitrile (2x10mL). The green beads obtained were dried over vacuum.

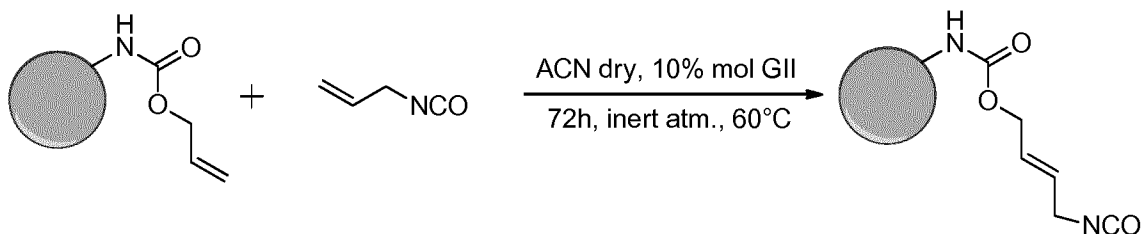
IR ($\nu_{\max}$, cm^{-1}): 3316, 2939, 2862, 2351, 1608, 1552, 1273 and 1080.

1.4.2 Metathesis**Scheme F.**

PEGA- NH_2 resin (0.4mmol/g) was dried under vacuum at 200°C in order to eliminate the residual water. The glassware was also dried on the oven and the round bottom flask flushed with nitrogen. 100mg (0.04mmol) of resin were introduced into the flask and soaked with dry acetonitrile (3mL). Then the N-(Allyloxycarbonyloxy)succinimide (5eq; 0.2mmol; 40mg; 31uL) was added, followed by triethylamine (10eq; 0.4mmol; 53uL). The mixture was stirred at room temperature during 3 days.

Then the resin was filtered off and washed successively with dry acetonitrile (3x10mL), tetrahydrofuran (3x10mL) and dichloromethane (3x10mL). The light salmon color beads obtained were dried over vacuum.

A Kaiser test was performed to give a negative result and the absorbance was measured to give the absence of the blue colored complex, giving a total conversion of the amine on the resin surface.

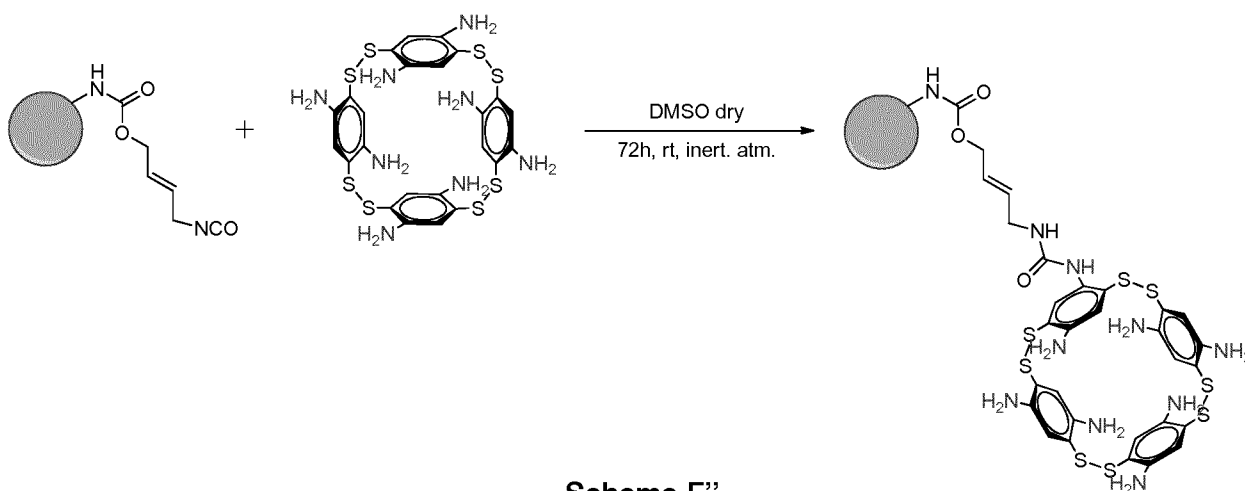


Scheme F'.

The previous beads were dried over vacuum in order to eliminate the residual water and the glassware was also dried on the oven. The beads (100mg) were placed into a schlenk flask flushed with nitrogen and soaked with dry acetonitrile (5mL). Then the Grubb's II catalyst was added (10%mol; 3.4mg; 0.004mmol) followed by the slow addition (30mn) of a solution of allyl isocyanate (20eq; 66.4mg, 0.8mmol; 70.6uL) diluted in 2mL of dry acetonitrile. A constant nitrogen flow was maintained and the mixture was stirred at room temperature during 24 hours. Then the nitrogen balloon was removed and the mixture refluxed during 48 hours under stirring.

Then the resin was filtered off and washed successively with dry acetonitrile (3x10mL), tetrahydrofurane (3x10mL) and dichloromethane (3x10mL). The white beads obtained were dried over vacuum.

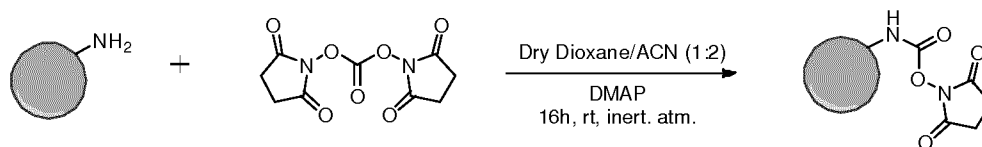
IR (ν_{max} , cm^{-1}): 3502, 2857, 1617 and 1093

**Scheme F''.**

The previous beads (100mg) were placed into a dry round bottom flask flushed with nitrogen. And dry DMSO was added (5mL) followed by an excess of dyn[4]arene amino by solid portions (200mg). Then the mixture was stirred at room temperature during 2 days protected from light.

Then the resin was filtered off and washed with dry DMSO (4x10mL) and acetonitrile (2x10mL). The beads obtained were dried over vacuum.

1.4.3 N-succinimidyl-PEGA



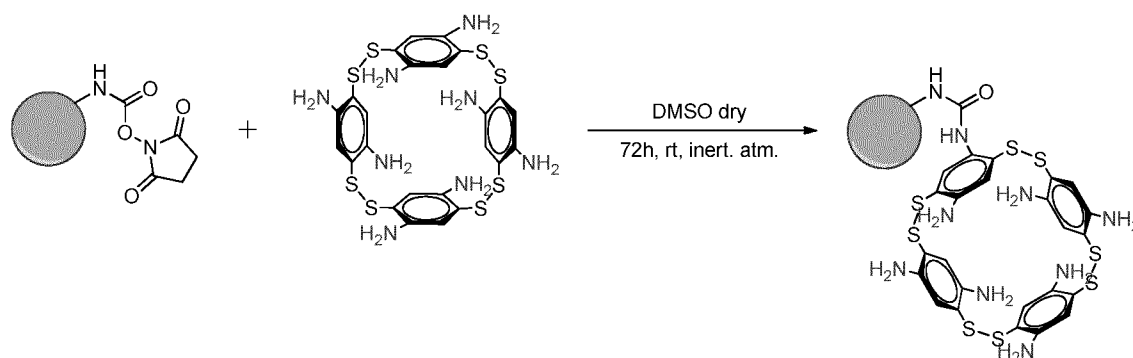
Scheme G.

PEGA-NH₂ resin (0.4mmol/g) was dried under vacuum at 200°C in order to eliminate the residual water. The glassware was also dried on the oven and the round bottom flask flushed with nitrogen. 250mg (0.1mmol) of resin were introduced into the flask and soaked with dry dioxane (5mL). Then the di(N-succinimidyl)carbonate (10eq; 1mmol; 256mg) was added in dry acetonitrile (5mL), followed by DMAP (10eq; 1mmol; 122mg) in dry acetonitrile (5mL) over 15mn. The mixture was stirred at room temperature overnight.

Then the solvents were evaporated and the residual solid triturated with dry diethyl ether. Then the product was filtered off and washed with dry acetonitrile (3x20mL). The beige beads obtained were dried over vacuum.

A Kaiser test was performed to give a negative result and the absorbance was measured to give the absence of the blue colored complex, giving a total conversion of the amine on the resin surface.

IR (ν_{max} , cm⁻¹): 3521, 2859, 1739, 1647 and 1080



Scheme G'.

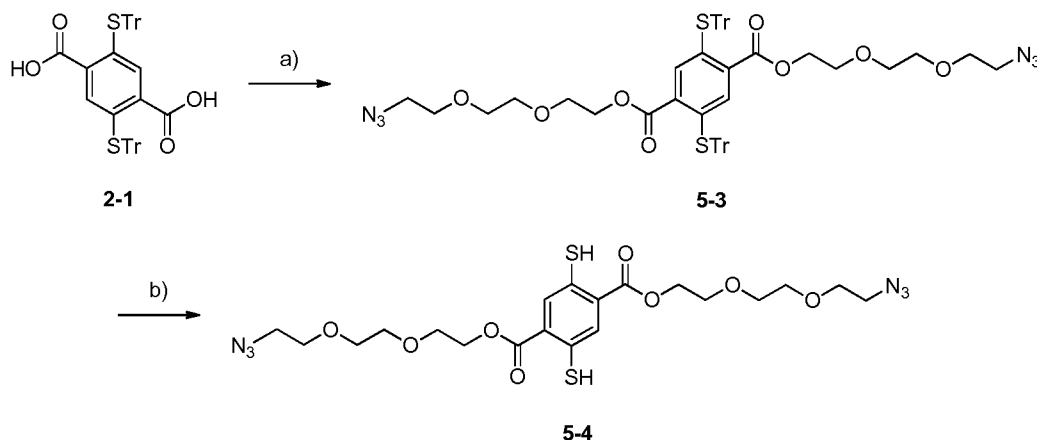
The previous beads (100mg) were placed into a dry round bottom flask flushed with nitrogen. And dry DMSO was added (5mL) followed by an excess of dyn[4]arene amino by solid portions (200mg). Then the mixture was stirred at room temperature during 2 days protected from light.

Then the resin was filtered off and washed with dry DMSO (4x10mL) and acetonitrile (2x10mL). The beads obtained were dried over vacuum.

IR ($\nu_{\max}$, cm^{-1}): 3520, 3860, 1734, 1649, 1072, 1033 and 954.

5

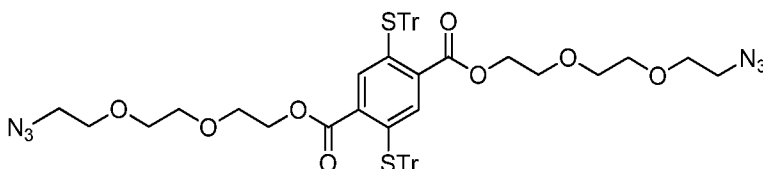
Click chemistry of terminal alkynes and azide is a useful tool for construction of silica based functionalized material



Scheme 2: Clickable building block synthesis

10

a) HOBt, EDC, **5-2**, DCM, r.t., 48 h, 47 %, b) TFA, Et_3SiH , DCM, r.t., 45 mn, 72 %



5-3 bis(2-(2-(2-azidoethoxy)ethoxy)ethyl) 2,5-bis(tritylthio)terephthalate

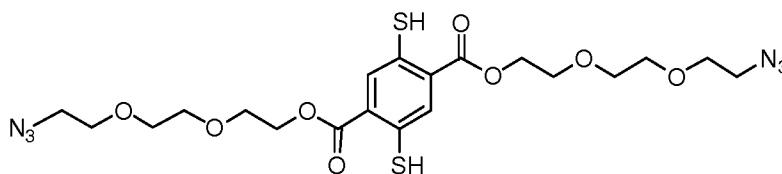
15

To a solution of **2-1** 2,5-bis(tritylthio)terephthalic acid (674 mg, 0.94 mmol, 1 eq) with HOBt (385 mg, 2.45 mmol, 2.6 eq) in dry DCM (50 mL) was added EDC at 0°C (452 μL , 2.45 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N_2 . **5-3** 2-(2-(2-azidoethoxy)ethoxy)ethanol (430 mg, 2.45 mmol, 2.6 eq) was then added dropwise at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na_2SO_4 and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO_2 , cyclohexane: ethyl acetate = 8:2) to give a yellow oil (454 mg, 47 %).

20

HRMS (ESI) $[\text{M} + \text{Na}]^+$ found 1051.3443, calculated 1051.3493 for $[\text{C}_{58}\text{H}_{56}\text{N}_6\text{NaO}_8\text{S}_2]^+$; ^1H NMR (400 MHz, CDCl_3) δ ppm = 7.39 - 7.14 (m, 30 H, $\text{Ar}_{\text{Trityl}}$), 4.16 (t, $J = 5.1$ Hz, 4 H, $\text{CH}_2\text{-COO}$), 3.68 - 3.64 (m, 4 H, $\text{CH}_2\text{-CH}_2\text{-N}_3$), 3.62 - 3.58 (m, 4 H, $\text{CH}_2\text{-O}$), 3.52 - 3.48 (m,

4 H, CH₂-O), 3.47 (t, J = 4.9 Hz, 4 H, CH₂-CH₂-COO), 3.38 (t, J = 5.1 Hz, 4 H, CH₂-N₃);
¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.0 (Ar-COO), 143.3 (C_{Ar}-C-S), 136.0 (C_{Ar}-COO),
 132.8 (C_{Ar} trityl), 130.7 (C_{Ar} trityl), 130.2 (C_{Ar} trityl), 127.7 (C_{Ar}-S), 126.9 (C_{Ar}H), 70.8 (C_{trityl}), 70.6
 (CH₂-O), 70.4 (CH₂-O), 70.1 (CH₂-O), 68.7 (CH₂-C-COO), 64.1 (CH₂-COO), 50.6 (N₃-CH₂)



5-4 bis(2-(2-(2-azidoethoxy)ethoxy)ethyl) 2,5-dimercaptoterephthalate

To a solution of **5-3** bis(2-(2-(2-azidoethoxy)ethoxy)ethyl) 2,5-bis(tritylthio)terephthalate (100 mg, 0.097 mmol, 1 eq) in degassed and dry DCM (10 mL) was added TFA (0.4 mL, 5.22 mmol, 54 eq). The mixture was stirred at room temperature for 30 mn. Then, Et₃SiH (0.1 mL, 0.63 mmol, 6 eq) was added dropwise and the mixture was stirred for 15 mn. Solvent and volatile compounds were evaporate. The remaining residue was dissolved in a 90/10 degassed mixture of MeOH/water (20 mL). This methanolic phase was washed with degassed heptane (4*15 mL), and then evaporated to dryness, affording a yellow oil (38 mg, 72 %).

Example 2 : evaluation of complexation properties of oligomeric macrocycle (19)₄ by UV-Vis titration

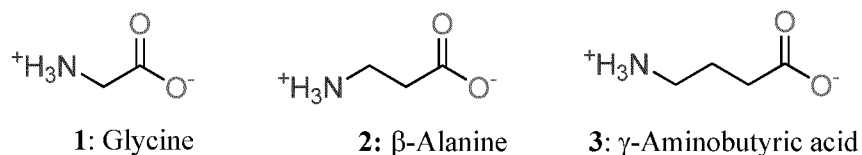
We investigated the complexation properties of (19)₄ by UV-Vis titrations. All spectra were carried out in quartz UV cuvettes.

A solution of Host (H) ([B4]=5.10⁻⁵M in 95% DMSO/ 5% AcONH₄ at 0.1M) was placed in a cuvette equipped with a stirrer and the spectrum was measured at ambient temperature. Then the guest was added in aliquots (10x10μL followed by 3-5x100μL) and the spectrum was measured after stirring. In each case, the addition of guest led to a progressive decrease of absorption. All titrations were repeated at least three times to provide reproducible results.

After data correction by the dilution factor, binding constants were obtained by curve-fitting of the corrected titration data at λ_{max} using a specific binding site model (1:1) with Prism software (Glycine, β-Alanine and γ-Aminobutyric acid) and with Origin software (Gln/Glu and Asn/Asp).

2.1. Gly, β-Ala & GABA

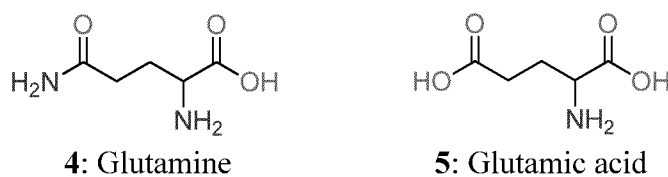
All spectra were carried out in 1 cm quartz UV cuvettes using Zwitterionic guests 1-3 as described below .



Scheme 1. Zwitterionic guests 1-3

2.2. Gln/Glu

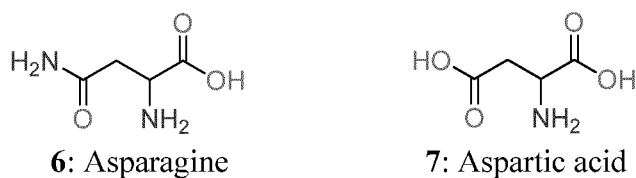
All spectra were carried out in 0.5cm quartz UV cuvettes.



Scheme 2. Amino acids guests 4-5

2.3. Asn/Asp

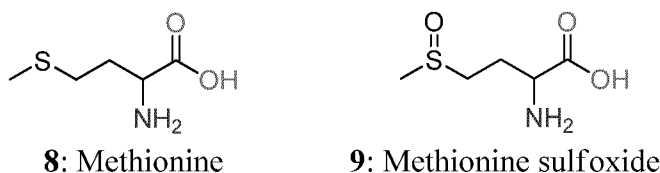
All spectra were carried out in 0.5cm quartz UV cuvettes.



Scheme 3. Amino acids guests 6-7

2.4. Met/Met(Ox)

All spectra were carried out in 0.5cm quartz UV cuvettes.



Scheme 4. Amino acids guests **8-9**

5

2.5. Results: Binding abilities of **(19)₄**

Affinity and selectivity of **(19)₄** for monomeric substrates of methoxidation and deamidation was precisely evaluated by UV-Vis titration in aqueous dmsO (95/5: dmsO/tris buffer 200 mM pH 7.4). Results, summarized in Table 1 show a micromolar affinity for

10

Ligands	$K_a(M^{-1})$	$\Delta K_a(M^{-1})$
Glycine	no binding	-
β -Alanine	$6.66 \cdot 10^4$	$6.66 \cdot 10^3$
γ -Aminobutyric acid	$7.21 \cdot 10^4$	$7.35 \cdot 10^3$
Glutamine	$1.10 \cdot 10^6$	$5.9 \cdot 10^5$
Glutamic acid	$2.98 \cdot 10^5$	$9.16 \cdot 10^4$
Asparagine	$1.63 \cdot 10^5$	$4.01 \cdot 10^4$
Aspartic acid	$1.79 \cdot 10^5$	$4.93 \cdot 10^4$
Methionine	$3.22 \cdot 10^5$	$1.36 \cdot 10^5$
Methionine sulfoxide	$1.16 \cdot 10^5$	$3.59 \cdot 10^4$

Table 1. Binding constant average values of compound **19₄** toward **1-9 Guests** from UV-Vis titration in aqueous solution $AcO^-NH_4^+(1.10^{-1}M)/DMSO$ (5/95)

15

Example 3 : evaluation of complexation properties of oligomeric macrocycle **(15)₄ by Isothermal Titration Calorimetry**

3.1. Isothermal Titration Calorimetry

Isothermal Titration Calorimetry (ITC) experiments were performed at 298 K using a Malvern ITC200 instrument. In a standard experiment, the host solution in Tris buffer pH 7.4 (200mM) was placed into the calorimeter cell (200 μ l) and 20 successive aliquots (2 μ l) of guest solution (10 times more concentrated) were added via a computer-automated injector at 2 min intervals. Heat changes were recorded after each addition. Heats of dilution were subtracted from the titration data prior to curve fitting. The first injection was

20

discarded from each dataset to remove the effect of guest diffusion across the syringe tip during the equilibration process. Titrations curves were fitted with the one binding site model using Origin v.5.0 software supplied by MicroCal.

3.2. Affinity of (15)₄ for Lys-tagged peptide

(15)₄ has been described in Skowron et al., J. Org. Chem. 2016, 81, 654–661. (15)₄ displays a strong preference for lysine vs arginine.

Bradykinin, an octapeptide of sequence Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg (SEQ ID NO:1) which is released from kininogen by plasma kallikrein, has been used as a generic backbone in order to evaluate the utility of the couple [Lysine/(15)₄] as a tag/receptor couple for the purification of proteins.

Accordingly, affinity of (15)₄ for bradykinin, N-Lys-tagged bradykinin (SEQ ID NO:2) and N-Met-Lys-tagged bradykinin (SEQ ID NO:3) has been assayed by Isothermal titration calorimetry (ITC) measurements in Tris 200 mM pH 7.4 at 298 K. Results of the measurement are shown in Table 2.

Ligands	K _a (M ⁻¹)
bradykinin	1.84.10 ⁴
N-Lys-bradykinin	1.18.10 ⁵
N-Met-Lys- bradykinin	1.21.10 ⁵

Table 2. Binding constant of compound (15)₄ toward bradykinin, N-Lys-tagged bradykinin and N-Met-Lys-tagged bradykinin.

These results show that the Lys tag induces an increase in increase in affinity by a factor 10, which is barely attenuated by the presence of a terminal methionine. Accordingly, the couple [Lysine/(15)₄] represent a valuable surrogate to metal-(His)_n tag which is widely used in molecular biology for protein purification but is strongly immunogenic.

In order to investigate the use of homochiral (15)₄ for the binding and extraction of Lys-tagged proteins, GFP was chosen for its ease of use and high stability. In this perspective, two GFP mutants were designed:

- MKQ-GFP-His6: A first mutant incorporating in N-terminal position a Lysine preceded by a Methionine was chosen due to the high susceptibility of N-terminal lysine bearing protein to degradation. Indeed, the presence of a N-terminal lysyl highly

decreases the half-life of the protein in vivo. Accordingly, a Methionine residue was chosen to mask the terminal Lysine, as we could show on Bradykinin that it moderately affects the binding affinity. Moreover, the Lys-Gln sequence was validated in the previous round of measurements to be strongly bound by the receptor and should display an efficient binding moiety.

- MAA-GFP-His6: A second mutant, acting as a negative control was designed, incorporating a Methionine and two Alanines thereby providing a TAG which should not interact with the receptor.

Both mutants were prepared, by recombinant technology, production in a fermenter and purification by IMAC column. C-Terminal His-tag was used to simplify the purification step. The resulting proteins solutions conditioned in Tris Buffer 20 mM pH=7.5, were titrated by UV-visible spectrometry in order to precisely determine their concentration.

Comparative binding experiments were conducted by ITC titrations (concentration of proteins: 0.27 mM). Although, binding seem to occur in both cases, the thermograms obtained are non-conventional. A sigmoid profile was obtained for MKQ-GFP while only few points could be fitted for MAA-GFP. The enthalpy value extracted from a 1:1 binding model for MKQ-GFP is twice as high as for MAA-GFP thereby providing a first encouraging trend toward the preferential binding of a N-Met-Lys tag. The apparent binding constants obtained are nevertheless quite high with $K_a(\text{MKQ-GFP})=3.07 \times 10^5 \text{M}^{-1}$ and $K_a(\text{MAA-GFP})=3.94 \times 10^4 \text{M}^{-1}$. In agreement with the variation in enthalpy, the increase of the affinity is one order of magnitude higher with the N-Met-Lys tag.

CD spectrum of MKQ-GFP and MAA-GFP were recorded with successively 1 equivalent and 10 equivalents of homochiral **(15)**₄. The resulting spectrums did not display any differences, thus suggesting that the binding event does not induces dramatic changes in the protein structure. The complexation of the N-terminal Methionine-Lysine sequence is coherent with this observation, as it should not interfere with the protein structure. Yet, no cotton effect resulting from the folding of the cyclophane is observed.

From these ITC measurements, it appears that the presence of a N-Met-Lys tag may induce a preferred binding of MKQ-GFP, however a significant affinity was also obtained in the case of MAA-GFP. As the Methionine-Alanine sequence should not provide any affinity, it may come from the rest of the protein and in particular from the C-terminal Histidines. This His tag was chosen for purification issues but may be non-orthogonal toward **(15)**₄. This feature will be investigated in upcoming tests. Binding measurement by ITC revealed no affinity between **(15)**₄ and HisNH₂ but it should be extended to a poly(His) guest to provide a clear answer.

3.3. Selective detection of methoxidation/deamidation by (15)₄

Binding of (15)₄ to a series of dipeptides has been assayed by Isothermal titration calorimetry (ITC) measurements in Tris 200 mM pH 7.4 at 298 K. Results of the measurement are shown in Table 3.

Ligands	K _a (M ⁻¹)
Lys-Gly	1.34.10 ⁵
Lys-Arg	1.22.10 ⁵
Lys-Lys	1.03.10 ⁵
Lys-Phe	3.81.10 ⁴
Lys-Pro	1.29.10 ⁴
Arg-Lys	3.84.10 ⁴
Met-Lys	1.56.10 ⁴
Met(O)-Lys	5.97.10 ³
Lys-Glu	1.34.10 ⁴
Lys-Asn	3.33.10 ⁴
Lys-Asp	2.43.10 ⁴
Lys-Gln	8.12.10 ⁴
Lys-Met	6.34.10 ⁴
Lys-Met(O)	9.95.10 ⁴
Glu-Lys	4.26.10 ³
Asp-Lys	5.18.10 ³

Table 3. Binding constant of compound (15)₄ toward dipeptides.

These results show that (15)₄ has two-fold selectivity for Lys-Met(O) compared to Lys-Met. (15)₄ also has six-fold selectivity for Lys-Gln compared to Lys-Glu.

GFP proteins were chosen for their easy use and their stability. A full evaluation of the affinity between A₄ and a set of GFP was conducted (see table 4). The GFP were prepared by recombinant technology, production in a fermenter and purification by IMAC column. μ-FFE, ITC and Octet® (label-free measurement of biomolecular interactions within the interactome) were used to characterise the binding mode.

GFP mutant	"MKQ"	"MAA"	"ΔK"
------------	-------	-------	------

C-terminal		6 Histidine in C-terminal	6 Histidine in C-terminal	Ø
N-terminal		Methionine-lysine-glutamine	Methionine-alanine-alanine	Methionine-alanine-alanine
Interest		Lysine specificity positive control	Lysine specificity negative control	Observation of a potential second binding site
ITC	N (1:1 model)	0.736	0.473	0.577
	K (L/mol)	$6.19 \cdot 10^3$	$1.04 \cdot 10^4$	$1.80 \cdot 10^3$
	ΔH (cal/mol)	-1919	-1644	-3357
	ΔS (cal/mol/°)	10.9	12.9	3.64

Table 4: Summary of the study of the interaction between A_4 and GFPs

ITC study showed a submolar affinity for each GFP. Yet, selectivity factor ranges from 2 to 10 allowing to envisage the implementation of the technology in mutistage liquid-liquid and liquid-solid extraction.

5

3.4. Affinity for large protein

Affinity studies on full antibody were further conducted. The antibody consisted of a di-Fab' fragment comprising:

A variable light chain of sequence: *MKKTAIAIAVALAGFATVAQ*
 10 *ADIQMTQSPSSLSASVGDRVTITCRASQDIAGSLNWLQQKPGKA**IKRLIYATSSLD**SGVP*
*KRFSGSRSGSDYTLTISSLQPEDFATYYCLQYGSFPPTFGQGT**KVEIKRTVAAPSVFIFPP*
*SDEQLKSGTASV**VCLLN**NFY**PREAKVQWKVDNALQSGNSQESVTEQ**DSKD**STYSL**SST*
*LTL**SKADYEKH* *KVYACEVTHQGLSSPVTKSFNRGEC* (SEQ ID NO:4), in which
 characters in italic denote the signal peptide;

15

A variable heavy chain of sequence:
*MKKTAIAIAVALAGFATVAQA**EVQLVESGGGLVQPGGSLRL**SCAASGFTFSSYGMSWVR*
*QAPGKGLEWVATITSGGSYTYVDSVKGRFTISRDN**AKNTLYLQMNSLRAEDTAVYYCV*
*RIGEDALDYWGQGT**LVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEP**VT*
 20 *SWNSGALTSGVHTFPAVLQSSGLYSLSSV**TV**PSSSLGTQTYICNVNHKPSNTKVDKKV*
EPKSCDKTHTCAA (SEQ ID NO:5), in which characters in italic denote the signal
 peptide.

Hydrolytic/oxidative treatment of antibody

A sample of the antibody was submitted to hydrolytic stress by treatment for one month at 37°C in Tris buffer 2M pH = 9 at the CIP. Affinity between **(15)₄** and this oxidized sample was evaluated by ITC measurement in Tris buffer pH = 7.4 after preliminary dialysis. A sample of the antibody was submitted to oxidative stress by treatment for one day at 37°C in Tris buffer 2M pH = 9.

Affinity between **(15)₄** and this oxidized sample was evaluated by ITC measurement in Tris buffer pH = 7.4 in the presence of H₂O₂ (0.2%) after preliminary dialysis. Non-stressed antibody was used as a reference compound for ITC titration.

ITC measurements

The calorimeter cell (200 µl) was filled with the protein in Tris buffer pH=7.5 (100mM) and host solution in the same buffer (protein dialysate) was placed into the syringe and 20 successive aliquots (2 µl) of host solution (10 times more concentrated) were added via a computer-automated injector at 2 min intervals. Heat changes were recorded after each addition. Heats of dilution of host (syringe) in a calorimeter cell filled with Tris buffer 7.5 (100mM) were subtracted from the titration data prior to curve fitting. The first injection was discarded from each dataset to remove the effect of guest diffusion across the syringe tip during the equilibration process.

Results

Results, summarized in Table 4 reveal a binding of **(15)₄** for the degraded antibody (K_d of about 10 µM) while not noticeable binding could be detected for the native antibody.

Antibody	K _f (M ⁻¹)
Hydrolyzed (1 month at pH 9)	K ₁ =7,86.10 ⁴ M ⁻¹ and K ₂ =2,50.10 ⁵ M ⁻¹
Oxidized (H ₂ O ₂ 0.2%, 24h)	No binding
Native form	no binding

Table 4. Binding constant of **(15)₄** for the degraded or native antibody

Analysis of the primary structure of the antibody reveals that two potential deamidation sites have a neighboring Lys residue which is the binding target of **(15)₄**. These may be preferential binding sites from which affinity and selectivity observed originates.

Peptide digestion and mapping was conducted to evaluate the levels of deamidation and methoxidation (Table 5) occurring upon oxidative stress and localize

their position on the sequence of the antibody, or at least the peptide fragments that are the most affected (Table 6).

Summary of post-translational modifications observed for all samples		
	Control (non stressed)	stressed
<i>Methionine oxidation</i>		
HC T002	0.4%	0.6%
HC T008	0.2%	0.1%
LC T001	0.3%	0.1%
<i>Deamidation</i>		
HC T008	0.2%	6.9%
HC T010		1.0%
HC T013		12.3%
LC T003	0.2%	0.4%
LC T013	0.0%	0.2%
LC T016	0.6%	6.2%
LC T020	0.1%	0.5%

Table 5. Qualitative analysis of deamidation and methoxidation levels and their

5 location on peptide fragments obtained after trypsin digestion.

LC T020	[191-207]	VYACEVTH <u>Q</u> GLSSPVTK (SEQ ID NO :6)
LC T016	[150-169]	VDNALQSGNS <u>Q</u> ESVTEQDSK (SEQ ID NO :7)
LC T013	[127-142]	SGTASVVCLLN <u>F</u> YPR (SEQ ID NO :8)
HC T013	[148-210]	DYFPEPVTVSWNSGALTSGVHTFPAVL <u>Q</u> SSGLYSLSSVVTVP SSSLGTQTYICNV NHK PSNTK (SEQ ID NO :9)
HC T010	[99-121]	IGEDALDYWG <u>Q</u> GTLVTVSSASTK (SEQ ID NO :10)
HC T008	[77-87]	<u>N</u> TLYLQMN <u>S</u> LR (SEQ ID NO :11)
LC T003	[25-42]	ASQDIAGSLNWL QQ KPGK (SEQ ID NO :12)

Table 6. correspondence between deamidation sites (underlined) and biding sites (bold).

Coupled (binding-deamidation) are indicated with a box.

10 The oxidative stress applied to the antibody mainly induces significant levels of deamidation and almost no methoxidation. Deamidation is mainly localized on three peptide fragments obtained after trypsin digestion (HC T008, LC T016 and HC T013). None of them directly contains a binding site (Lys) adjacent to a deamidation site. Yet, HC

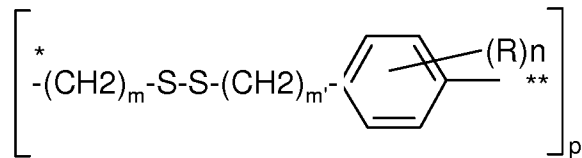
T013 which undergoes the highest levels of deamidation contains multiple deamidation sites, including a 'NXK' sequence.

It seems likely that oxidative stress may alter the secondary and tertiary structure thereby exposing additional accessible lysine residues for reinforced binding.

CLAIMS

1. A method for differentiating native form from oxidized and/or hydrolyzed forms of a protein comprising:

a. contacting a composition likely to comprise native, oxidized and/or hydrolyzed forms of a protein with a oligomeric macrocycle of general formula (A):



(A)

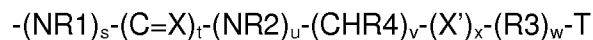
Where

m and m' identical or different are an integer independently chosen from 0 or 1;

where n is an integer comprised between 1 and 4;

p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 50;

each R identical or different may be represented by the following formula:



where T is a terminal group and may be chosen from the groups consisting in H, N₃, alkyl, heteroaryl, a solid phase support;

s, t, u, v, w identical or different may be independently 0 or 1;

R₁ represents H or -alkyl optionally substituted by one or more of halogen atoms, OH, OR₁₁, NR₁₁R₁₂, CN;

X represents O, NH or S;

X' is O, NH or S;

R₂ represents H, -alkyl optionally substituted by one or more of halogen atoms, OH, OR₁₁, NR₁₁R₁₂, CN;

R₃ represents -alkyl-, -heteroaryl- or -aryl- optionally substituted by one or more of

-R₁₁ or -OR₁₁, -N=C(NR₁₁)₂-NR₁₁-, -(OCH₂CH₂)_q-(CH₂CH₂O)_q-, -alkyl-NH-C(=O)-NH-, -aryl-NH-C(=O)-NH-, -alkenyl-OCO-NH-, -alkyl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -aryl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -NR₁₁-, -N(R₁₁)-C(=NH)-NH-, -O-, -C(O)O- -N(R₁₃)-(CH₂)_x-NR₁₁-, -SO₂-aryl-;

R4 is H or COOH;

R11, R12 identical or different are independently selected from the group consisting in H, alkyl optionally substituted by one or more of halogen atoms, OH, CN;

R13 is -alkyl optionally substituted by NR11R12;

5 q is an integer comprised between 1 and 200;

x is 0 or 1;

where alkyl is C1-C6 alkyl; aryl is a mono or bicyclic C6-C10 aromatic ring system; heteroaryl is a 5 to 10 membered mono or bicyclic aromatic ring system comprising one to 4 heteroatoms chosen from N, O or S.

10 It being understood that R may be different for each unit and/or for each n;

whereby the cyclic form is achieved by the bond between the S (*) of the first unit and the carbon of the phenyl group (**) of the last unit;

as well as its various stereoisomers;

and

15

b. detecting or collecting complexes formed with the oligomeric macrocycle of general formula (A), wherein said complexes preferentially comprise native form of the protein or oxidized and/or hydrolyzed forms of the protein.

20 2. The method of claim 1 wherein in formula (a):

n is 2;

p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 5;

each R identical or different may be represented by the following formula:

25



where T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl, a solid phase support;

30 s, t, u, v, w identical or different may be independently 0 or 1;

R1 represents H or -alkyl;

X represents O, NH or S;

R2 represents H, -alkyl;

R3 represents -heteroaryl- or -aryl- optionally substituted by -R11 or -OR11, -N=C(NR11₂)-NR11-, -(CH₂CH₂O)_q-, -NR11-, -N(R11)-C(=NH)-NH-, -N(R13)-(CH₂)_x-

35

NR11-, -SO₂-aryl-, -alkyl-NH-C(=O)-NH-, -aryl-NH-C(=O)-NH-, -alkenyl-OCO-NH-, -alkyl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -aryl-NH-C(=O)-NH-(CH₂CH₂O)_q-;

R11, R12 identical or different are independently selected from the group consisting in H, alkyl;

R13 is -alkyl optionally substituted by NR11R12;

q is an integer comprised between 1 and 200.

3. The method according to claim 1 or 2 wherein in formula (A):

n is 2;

p is the number of monomer units in the oligomeric macrocycle and p is comprised between 3 and 5;

each R identical or different may be represented by the following formula:



And where

When s is 1:

R1 is H;

t is 0 or 1 and X is O;

u is 0 or 1 and R2 is H,

v is 0;

w is 0 or 1 and R3 is -(CH₂CH₂O)_q-, -alkyl-NH-C(=O)-NH-, -aryl-NH-C(=O)-NH-, -alkenyl-OCO-NH-, -alkyl-NH-C(=O)-NH-(CH₂CH₂O)_q-, -aryl-NH-C(=O)-NH-(CH₂CH₂O)_q-, with q is comprised between 1 and 200; -SO₂-aryl-;

T is a terminal group and may be chosen from the groups consisting in H, alkyl or a solid phase support;

or

When

Where s is 0:

t is 1 and X is O;

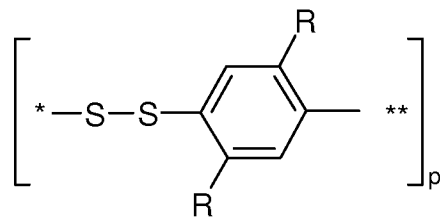
u is 0 or 1 and R2 is H, v is 0, 1 or 2;

w is 0 or 1 with R3 is chosen from -heteroaryl- or -aryl- optionally substituted by -R11 or -OR11, -N=C(NR11₂)-NR11-, -O-, -NR11-, -N(R11)-C(=NH)-NH-, -N(R13)-(CH₂)_x-NR11-; with R11 is H or alkyl;

T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl.

4. The method according to anyone of claims 1 to 3, wherein the solid support is chosen from polystyrene (PS), polyacrylamide, amylose, cellulose.

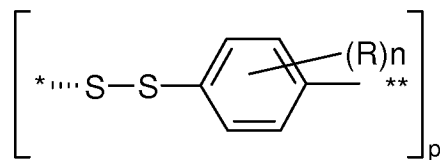
5. The method according to anyone of claims 1 to 4 wherein said compound of formula (A) has the following formula (I'):



(1')

where each R identical or different, and p is defined as in any one of claims 1 to 3.

6. The method according to anyone of the preceding claims wherein the compound of formula (I) is in the form of the stereoisomer of formula (I'') :



(1'')

where R, n and p are defined as in anyone of claims 1 to 3.

7. The method according to any one of the preceding claims, wherein said oxidized and/or hydrolyzed forms of the protein are methoxydated and/or deamidated forms of the protein.

8. The method according to any one of the preceding claims, wherein said protein comprises at least one Methionine, Glutamine or Asparagine which is adjacent to a Lysine in the primary, tertiary, or quaternary structure of said protein.

9. The method according to any one of the preceding claims, wherein said oxidized and/or hydrolyzed forms of the protein are present in a composition comprising oxidized

and/or hydrolyzed forms of the protein and native protein, and wherein the method further comprises recovering the forms of the protein which are not bound or unbound to said oligomeric macrocycle of general formula (A).

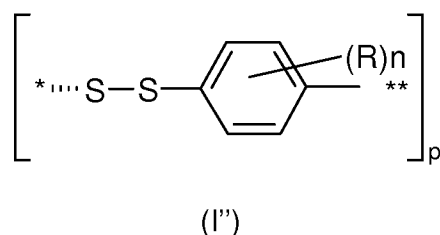
10. A method of purification of a protein which comprises:

a. Contacting a composition comprising a target protein with a oligomeric macrocycle of formula (A) as defined in anyone of the preceding claims and

b. Recovering the target protein bound to the oligomeric macrocycle of general formula (I'') as defined above

wherein the target protein or peptide comprises a lysine at position 1 or 2 of its amino acid sequence.

11. The method of purification according to claim 11 wherein it is conducted with a compound of formula (I'')



where:

n is 2;

p is 4;

R is $-(NR_1)_s-(C=X)_t-(NR_2)_u-(CH_2)_v-(R_3)_w-T$

Where s is 0:

t is 1 and X is O;

u is 0 or 1 and R₂ is H, v is 0, 1 or 2;

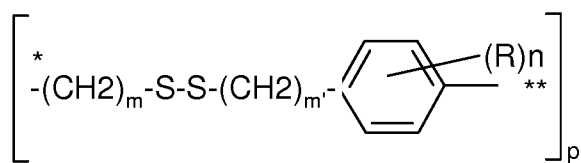
w is 0 or 1 with R₃ is chosen from -heteroaryl- or -aryl- optionally substituted by -R₁₁ or -OR₁₁, -N=C(NR₁₁)₂-NR₁₁-, -O-, -NR₁₁-, -N(R₁₁)-C(=NH)-NH-, -N(R₁₃)-(CH₂)_x-NR₁₁-; with R₁₁ is H or alkyl;

T is a terminal group and may be chosen from the groups consisting in H, alkyl, heteroaryl.

12. The method of purification according to claim 10 or 11, wherein said target protein is a recombinant protein which comprises a lysine immediately downstream of a N-terminal methionine.

13. The method of purification according to claim 11, wherein said target protein is a synthetic or mature protein which comprises a N-terminal lysine.

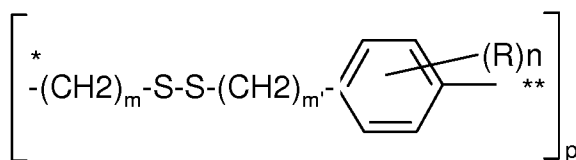
14. A compound of formula (A):



(A)

As defined in anyone of claims 1 to 6,
with the exception of the compounds where
 $m=m'=0$ and
 $p=4$, $n=2$ and $R=\text{COOH}$,
 $p=4$, $n=2$ and $R=\text{NH}_2$;
 $p=4$, $n=2$ and $R=\text{OMethyl}$;
 $p=4$, $n=2$ and $R=\text{OEthyl}$.

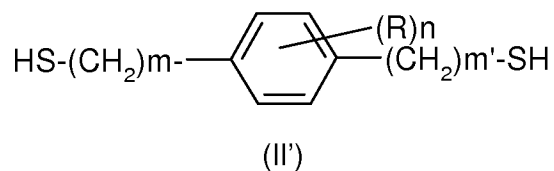
15. The process of preparation of the compound of formula (A) as defined in claim 14



(A)

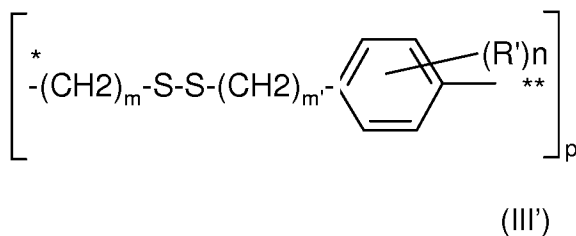
where R , n , p are as defined in anyone of claims 1 to 6,
with the exception of the compounds where
 $m=m'=0$, and
 $p=4$, $n=2$ and $R=\text{COOH}$,
 $p=4$, $n=2$ and $R=\text{NH}_2$;

said process comprising the step of self-assembling a compound of formula (II'):



5 where R, m and m' are defined as in claims 1 to 6 or a salt thereof.

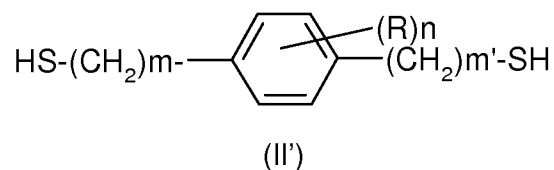
16. The process of preparation of the compound of formula (A) as defined in claim 14, comprising the step of derivatizing a compound of formula (III'):



where

each R' represents a -NH₂ group and ,p is defined as in formula (A)
so as to introduce the appropriate R group.

17. The compound of formula (II'):



Where m, m' n and R are defined as in claims 1 to 6
with the exception of the compounds where
m=m'=0 and
n=2 and R=COOH, or
n=2 and R=NH₂.

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/064022

A. CLASSIFICATION OF SUBJECT MATTER
INV. B01D15/08 C08G75/0209 C07D341/00 C07K1/14
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

B01D C08G G01N C07K C07D

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PIERRE-THOMAS SKOWRON ET AL: "On-Demand Cyclophanes: Substituent-Directed Self-Assembling, Folding, and Binding", THE JOURNAL OF ORGANIC CHEMISTRY, vol. 81, no. 2, 21 December 2015 (2015-12-21), pages 654-661, XP055303558, US ISSN: 0022-3263, DOI: 10.1021/acs.joc.5b02605 figure 1; compounds 1-3 ----- -/-	1-13



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

24 August 2017

Date of mailing of the international search report

02/11/2017

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Vogt, Titus

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SALEH HAMIEH ET AL: "A "Dial-A-Receptor" Dynamic Combinatorial Library", ANGEWANDTE CHEMIE INTERNATIONAL EDITION, vol. 52, no. 47, 2 October 2013 (2013-10-02), pages 12368-12372, XP055303551, DE ISSN: 1433-7851, DOI: 10.1002/anie.201305744 figure 1	1-13
Y	----- LAURENT VIAL ET AL: "Controlling the Biological Effects of Spermine Using a Synthetic Receptor", JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 128, no. 31, 1 August 2006 (2006-08-01), pages 10253-10257, XP055303554, US ISSN: 0002-7863, DOI: 10.1021/ja062536b figure 2; compound 6	1-13
A	----- MAHENDRA P. SONAWANE ET AL: "Synthesis and structural exploration of disulfide bridged [2n] pillararene-like molecules", CHEMICAL COMMUNICATIONS - CHEMCOM., vol. 49, no. 56, 1 January 2013 (2013-01-01), page 6310, XP055303586, ISSN: 1359-7345, DOI: 10.1039/c3cc42984a Scheme 1; compounds 3a,3b	1-13
X,P	----- LAURENT VIAL ET AL: "Chirality sensing and discrimination of lysine derivatives in water with a dyn[4]arene", CHEMICAL COMMUNICATIONS - CHEMCOM., vol. 52, no. 99, 15 November 2016 (2016-11-15), pages 14219-14221, XP055399919, ISSN: 1359-7345, DOI: 10.1039/C6CC07713G the whole document	1-13
A	----- CHIN-YANG YU ET AL: "Synthesis and Ring-Opening Metathesis of Tetraalkoxy-Substituted [2.2]Paracyclophane-1,9-dienes", CHEMISTRY - A EUROPEAN JOURNAL, vol. 17, no. 25, 14 June 2011 (2011-06-14), pages 6991-6997, XP055399946, ISSN: 0947-6539, DOI: 10.1002/chem.201003147 Scheme 1; compound 2	1-13
	----- -/--	

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ANNA BARATTUCCI ET AL: "Stereoselective Synthesis of Dithia[3.3]cyclophane S , S' -Dioxides with Planar and Central Chirality : Stereoselective Synthesis of Dithia[3.3]cyclophane S , S' -Dioxides", EUROPEAN JOURNAL OF ORGANIC CHEMISTRY, vol. 2014, no. 10, 3 February 2014 (2014-02-03), pages 2099-2104, XP055399955, DE ISSN: 1434-193X, DOI: 10.1002/ejoc.201301636 Scheme 3; compound 1	1-13
A	CHIN-YANG YU ET AL: "Synthesis and through-space charge transfer of dioctyloxy diperfluorohexyl substituted [2.2]paracyclophane-1,9-diene", NEW JOURNAL OF CHEMISTRY, vol. 38, no. 10, 1 January 2014 (2014-01-01), pages 5003-5008, XP055399957, GB ISSN: 1144-0546, DOI: 10.1039/C4NJ01045K Scheme 1; compound 2	1-13
A	OMER K. RASHEED ET AL: "A Modular Synthesis of Multidentate S-, N- and O-Containing Meta- and Paracyclophanes : Multidentate S-, N- and O-Containing Meta- and Paracyclophanes", EUROPEAN JOURNAL OF ORGANIC CHEMISTRY, vol. 2015, no. 32, 8 October 2015 (2015-10-08), pages 6988-6993, XP055399959, DE ISSN: 1434-193X, DOI: 10.1002/ejoc.201501058 Scheme 2; compound 17	1-13
A	BENJAMIN J. LIDSTER ET AL: "Alkyl substituted [2.2]paracyclophane-1,9-dienes", ORGANIC & BIOMOLECULAR CHEMISTRY, vol. 14, no. 25, 1 January 2016 (2016-01-01), pages 6079-6087, XP055399984, GB ISSN: 1477-0520, DOI: 10.1039/C6OB00885B Scheme 2; compounds 10,11	1-13
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/064022

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MARY S. COLLINS ET AL: "Pnictogen-directed synthesis of discrete disulfide macrocycles", CHEMICAL COMMUNICATIONS - CHEMCOM., vol. 49, no. 59, 5 June 2013 (2013-06-05), page 6599, XP055400086, ISSN: 1359-7345, DOI: 10.1039/c3cc43524e Scheme 2; figure 3 -----	1-13
A	US 2016/137598 A1 (JOHNSON DARREN W [US] ET AL) 19 May 2016 (2016-05-19) page 13, paragraph 104 -----	1-13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2017/064022

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-13

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-13

Methods of protein investigation, such as: - differentiating between native and non-native forms of a protein, and - protein purification, comprising the use of a para-cyclophane of formula (A).

2. claims: 14-16

Cyclophane of formula (A) and process of its preparation

3. claim: 17

Compound of formula (II'), suitable for use in the preparation of cyclophanes.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2017/064022

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2016137598	A1	19-05-2016	NONE
