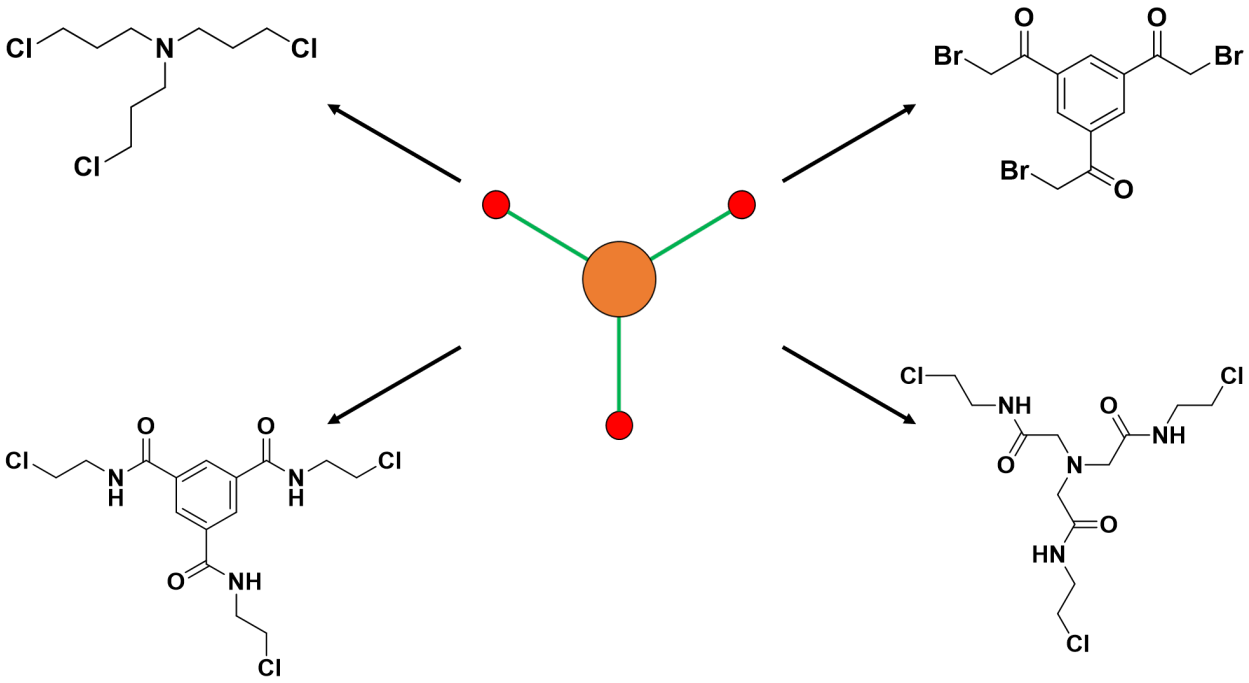


Synthesis of C_3 symmetric, tri-functional, thiol reactive small molecule linkers for bicyclization of peptides

Master's thesis in Materials Chemistry, MPMCN

RENGARAJ GOPAL

Department of Chemistry and Chemical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2020

Master's thesis 2020: KBTX12

Synthesis of C_3 symmetric, tri-functional, thiol reactive small molecule linkers for bicyclization of peptides

RENGARAJ GOPAL

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Chemistry and Chemical Engineeering
Division of Chemistry and Biochemistry, Organic Chemistry
Chalmers University of Technology
Gothenburg, Sweden 2020

Synthesis of C_3 symmetric, tri-functional, thiol reactive small molecule linkers for bicyclization of peptides
RENGARAJ GOPAL

Supervisor:
Tor Svensson, Senior Research Scientist, New Modality Chemistry
Research and Early Development, Respiratory & Immunology (R&I)
AstraZeneca, Gothenburg, Sweden
tor.s.svensson@astrazeneca.com

Examiner:
Jerker Mårtensson, Professor, Division of Chemistry and Biochemistry,
Department of Chemistry and Chemical Engineering,
Chalmers University of Technology.
jerker@chalmers.se

Master's Thesis 2020: KBTX12
Department of Chemistry and Chemical Engineering
Division of Chemistry and Biochemistry, Organic Chemistry
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Four different C_3 symmetric, trifuctional, thiol reactive linkers synthesized using the general design of having a three side chains (with halide leaving groups) linked to a central core.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2020

Synthesis of C_3 symmetric, tri-functional, thiol reactive small molecule linkers for bicyclization of peptides

RENGARAJ GOPAL
Department of Chemistry and Chemical Engineering
Chalmers University of Technology

Abstract

Bicyclic peptides are next-generational therapeutics due to high conformational rigidity, metabolic stability and better cellular permeability. They combine the favourable properties of macromolecular drugs (high target specificity and affinity) with small molecules (stability and good tissue penetration) making them ideal for therapeutics with low toxicity profiles against a diverse set of targets. Bicyclic peptides can be synthesized from a linear precursor by using linkers, a tri-functionalized small molecule. The primary role of a linker in a cyclic peptide is to join the ends of the peptide which leads to reduced backbone flexibility. Linkers have also been shown to influence the structure of the bicyclic peptides by positioning the ends of the peptide at specific distances or orientations, non-covalent interactions or steric exclusion that lock peptide backbones into specific conformers. In this work, the design and synthesis of four polar linkers that are C_3 symmetric, tri-functionalized and thiol reactive with different chemical substituents have been discussed. The linkers were synthesized in low to good yields and the threefold rotational symmetry of the linker ensures the formation of a single product isomer upon conjugation with a peptide. The linkers are made of a central nitrogen atom or a benzene core with three side arms comprising of polar/aliphatic functional groups. The linkers contain a halide leaving group that is sulfhydryl/thiol reactive allowing the reaction with peptides containing cysteine residues. Compared to earlier linkers reported in literature, the synthesized linkers have high hydrophilic nature which could aid in solubility of the bicyclic peptide in aqueous solvent and also lead to formation of non-covalent bonds with the amino-acids of the bicyclic peptide. Thus, introducing additional conformational rigidity leads to lower entropy loss when binding to targets.

Keywords: C_3 symmetry, thiol reactive, linkers, bicyclic peptide, peptide cyclization, non-covalent interactions, polar linkers.

Acknowledgements

I would like to thank AstraZeneca for accepting me as a thesis worker and for allowing me to do a Master's thesis in organic chemistry.

This thesis was possible due to a lot of help from people involved in the project and around me. Hence, I would like to take this opportunity to sincerely thank those people:

Tor Svensson for being my supervisor and guiding me through the project, for showing me around all the lab facilities, patiently teaching new synthetic techniques, helping me troubleshoot chemical reactions, analyzing analytical data and for all the support in helping me write the thesis.

Wu Su for helping me out with administrative stuff and the entire RIA team for making me feel welcome and part of the team. Especially, for the great company during fika breaks where majority of ideas for troubleshooting chemical reactions were born.

Maria Ölwegård Halvarsson for helping me out with any laboratory related question I had and for offering valuable practical chemistry insight.

Kartik, Varsha, Sumukha, Roopa, Adarsh, Ashwin and Ashok for keeping me sane during the hard times and giving me the much needed pep talks. Amma and Appa for all the love and support but especially for all the Indian food recipes which gave me a sense of warmth and comfort of being at home. I have been lucky enough to have people who care and keep regular tabs on me, for which I am extremely grateful.

Tejas for helping me structure the thesis and making it coherent.

Lastly, I would like to thank Jerker Mårtensson for being my examiner and helping me with any questions I had on how to proceed with the thesis and writing it.

Rengaraj Gopal, Gothenburg, October 2020

Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

ACN	Acetonitrile
BCPA	Bis(3-chloropropyl)amine
BTBE	1,1',1"-(Benzene-1,3,5-triyl)tris(2-bromoethan-1-one)
$CDCl_3$	Deuterated chloroform
DCM	Dichloromethane
DIPEA	*N,N*-Diisopropylethylmine
DMF	*N,N*-Dimethylformamide
DMSO	Dimethylsulfoxide
Et_2O	Diethyl ether
EtOAc	Ethylacetate
eq.	Equivalent
GC	Gas Chromatography
LC	Liquid Chromatography
MS	Mass spectroscopy
NATBA	*N,N',N",N"'*-(Adamantane-1,3,5,7-tetrayl)tetrakis(2-bromoacetamide)
NBS	*N*-Bromosuccinimide
NMP	*N*-Methylpyrrolidone
NMR	Nuclear Magnetic Resonance
NTCA	2,2',2"-Nitrilotris(*N*-(2-chloroethyl)acetamide)
PPI/PPIs	Protein-Protein Interaction(s)
TAAB	*N,N',N"*-Benzene-1,3,5-triyltrisprop-2-enamide
TATA/TAT	1,3,5-Triacryloylhexahydro-1,3,5-triazine
TATB	1,1',1"-(1,3,5-Triazinane-1,3,5-triyl)tris(2-bromoethanone)
TBAB	*N,N',N"*-(Benzene-1,3,5-triyl)tris(2-bromoacetamide)
TBCA	N^1,N^3,N^5-Tris(2-chloroethyl)benzene-1,3,5-tricarboxamide
TBMB	1,3,5-Tris(bromomethyl)benzene
TBMT	2,4,6-Tris(bromomethyl)-*s*-triazine
TCPA	Tris(3-chloropropyl)amine
TEA	Triethyl amine
TPSMB	1,3,5-Tris((pyrididn-2-yldisulfanyl)methyl)benzene
UV	Ultraviolet

1

Introduction

1.1 Small molecule drugs

Acetaminophen more commonly known as Tylenol or Paracetamol is one of the most commonly prescribed analgesic used to treat pain conditions. It has a mass of 151.06 Da and it is a traditional small molecule drug. Small molecule drugs are usually chemically synthesized and have a molecular weight less than 900 Da. Modern therapeutics were traditionally equated with small molecule drugs and these constitute the most usable FDA approved drugs [1][8]. These are easily absorbed by the GI tract which makes it easier for non-invasive delivery, can traverse different organs and enter various cells. This property also imparts non-specificity and increases incidence of side effects. They are also limited to targeting proteins with well-defined binding pockets such as enzymes, ion channels and ion receptors [3].

Protein-protein interactions (PPIs) represent a large class of drug targets and are of central importance in most biochemical pathways including those involved in disease processes. PPIs usually involve large, flat interfaces, which makes it a challenging or undruggable target for small molecules. Most small molecules do not make enough points of contact with the PPIs to impart high affinity or specificity. Proteins involved in protein-protein interactions (PPIs) lead the protein of interest to not be easily available which affects the efficacy of the small molecule drugs [1]. Additionally these drugs are rapidly eliminated from the body requiring frequent dosing to maintain the required therapeutic levels.

1.2 Macromolecules

The limitation of small molecule are usually overcome by macromolecules with molecular weight, MW>5000 Da [1]. Macromolecules have emerged as a powerful class of drugs due to their site specific activity and reduced side-effects. Antibodies possess large binding surfaces and are capable of making multiple contacts with a PPI's target surface. However, antibodies or protein based drugs are impermeable to the mammalian cell membrane and are generally limited to targeting extracellular proteins [23][1]. These drugs are poorly absorbed in the GI tract due to their large molecular mass, hydrophilic nature and charge thus posing a challenge for non-invasive delivery [3]. Furthermore, they are easily metabolized and chemically as well as biologically unstable. Peptides on the other hand are in between the small molecules and macromolecules in terms of MW, but are biochemically and

therapeutically distinct from both of them [1].

1.3 Peptides

When compared to small molecules, peptides have a more exquisite affinity, specificity for a unique molecular target, are less toxic and do not accumulate in organs. This is because after acting on the target they get degraded into non-toxic amino acids by proteolytic fission [18]. Peptide based drugs have a large enough surface of interaction to obtain high potency and selectivity but still small enough to penetrate tissues [2]. Peptides are hard to be orally administered as they are hard to be absorbed in GI tract and they are rapidly metabolized [18].

1.3.1 Stability of peptides

Peptides with a sequence of α-helices populate multiple random-like structures in solution instead of a stable three-dimensional structure [10]. Short peptides in isolation do not retain their native conformation and binding capability due to the lack of structural reinforcement [19]. Due to the lack of structural conformity in solution, peptides do not bind efficiently to the intended target and are susceptible to proteolytic degradation. Reinforcing the native alpha-helical conformation can be carried out by peptide stapling, hydrogen bond surrogates, beta-peptides or metal chelating. Peptide stapling is one of the techniques in which two amino acids lying on the same face of the helix are substituted by non-native amino acid residues that have side chains which can be covalently linked or 'stapled' together. This macrocyclization process helps impart structural rigidity to the peptide and also reinforces the alpha-helical structure. [19]

Peptide α-helices are stabilized by the incorporation of covalent and non-covalent linkages between amino acid chains. Linkages commonly used are amide bonds to form lactams, disulfide bridges between cysteine residues, hydrogen bond surrogates, cysteine alkylation using α-haloacetamide derivatives, metal chelation, click chemistry, hydrocarbon linkers formed by olefin metathesis, thioether bridges, triazole linkages formed between azides and alkynes and linkers formed by oxime formation. The inclusion of a semi-rigid cross-linker is a very successful method to stabilize peptide α-helices. [10]

1.4 Peptide cyclization

Peptides can be broadly classified as linear and cyclic. Cyclic peptides are peptides where the polypeptide chains form a ring. These cyclic structures are formed by linking one end of the peptide to another with amide, lactone, ether, thioether or disulfide bonds. A looped structure in a peptide can be achieved by head-to-tail, side chain-to-side chain and side chain-to-terminus cyclization (see **Figure** 1.1). If the cyclized amino acids are only connected via amide bonds, then these are called "homodetic". When any other linkage is involved (disulfide or depsipeptide bonds) it is referred to as a "heterodetic" peptide. [11]

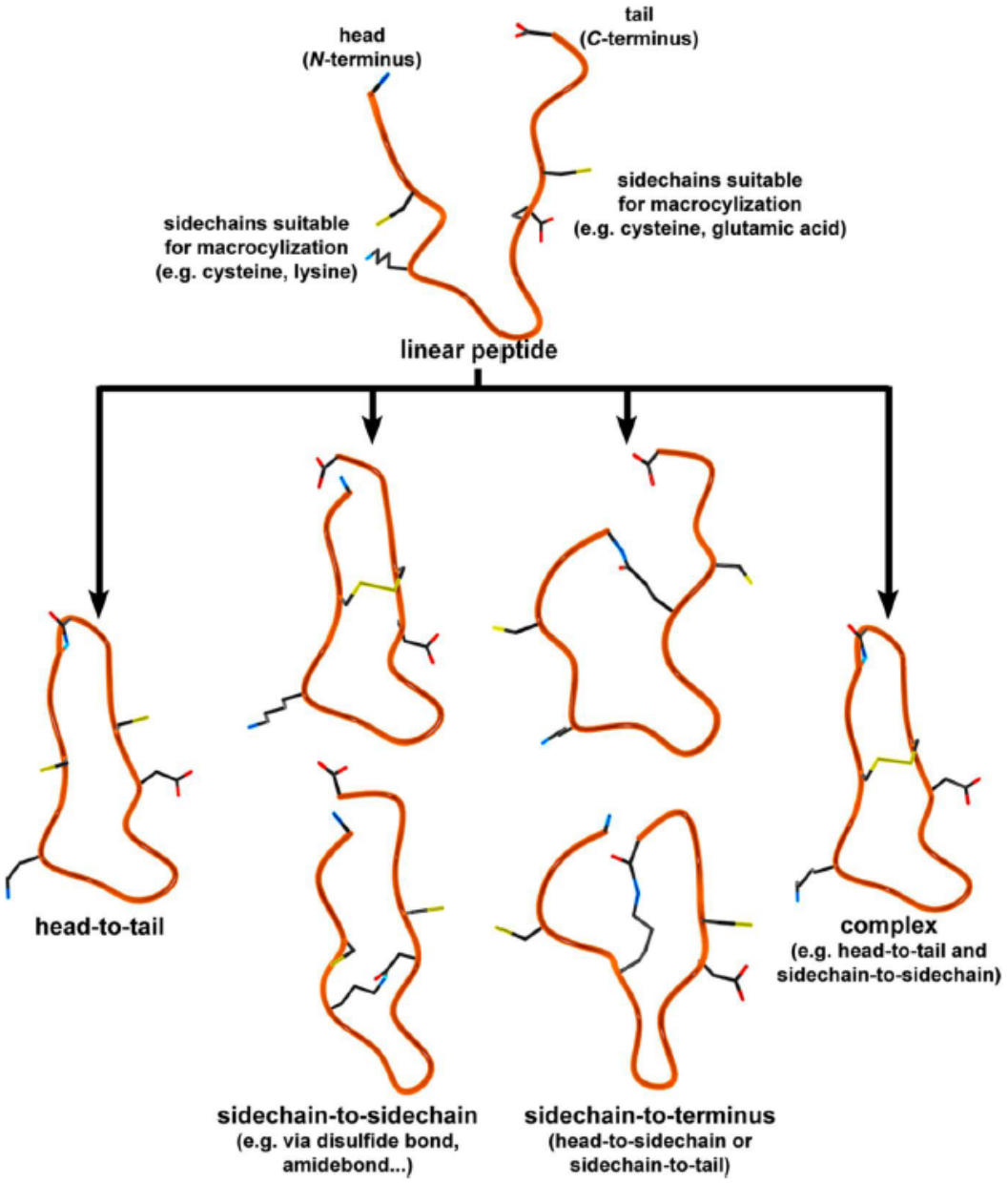


Figure 1.1: Common modes of macrocyclization shown for the example of a generic peptide chain containing two cysteines, a lysine, and a glutamic acid. All other residues are not shown. The peptide backbone is given in orange, carbon in black, nitrogen in blue, oxygen in red, and sulfur in yellow [11].

Carboxyamidation or C-terminal amidation generally improves the activity of the peptide by removing the negative charge of the carboxyterminal. This specific modification can also help stabilize the structure of some peptides but have a negligible effect against proteolytic stability [25]. Biologically active peptides are usually formed by head-to-tail cyclization (N-to-C) which is a amide bond formed between amino and carboxyl termini. This makes cyclic peptides more stable to hydrolysis by exopeptidases. [18]

1.4.1 Properties of cyclic peptides

The conformation rigidity of the cyclic peptides imparts better biological activity compared to their linear counterparts. Cyclic peptides can be used as a drug delivery tool, therapeutic, surfactant, imaging agent or biosensor. Cyclic peptides as a therapeutic are known to have bactericidal, immunosuppressive and anti-angiogenic activity [18].

Macrocyclic compounds have molecular weights between 500 and 2000 Da and have gained much attention as PPI inhibitors. Due to their large sizes and ability to make multiple contacts with a flat surface, they effectively compete with proteins for binding to flat surfaces while retaining many properties of the small molecules such as membrane permeability. Unlike small molecule drugs, macrocycles have greater metabolic stability, less likelihood of eliciting an immune response and low cost of production.[23]

The backbone of cyclic peptides are conformationally less flexible than linear peptides, this imparts the cyclic structures with favorable properties. Constrained peptides have a smaller loss on entropy when target binding leading to higher binding affinities. Cyclic peptides can efficiently target PPIs since they are able to interact even with flat, featureless surfaces of protein which are hard to be targeted by small molecules. Cyclic peptides are also more selective because of the reduced number of conformations. Lastly, cyclic peptides have lower flexibility preventing the peptides from fitting well to the active sites of peptidases. This imparts higher metabolic stability to cyclic peptides as compared to linear peptides.[6]

Monocyclic peptides of small ring sizes (<10 amino acids) are known to be relatively resistant to proteases and have demonstrated oral bioavailability. However, as the ring size increases they lose their conformational rigidity and become increasingly susceptible to proteolytic degradation. One possible solution to increase the proteolytic stability of macrocyclic peptides is to convert them to bicyclic structures. This reduces the size of the ring as compared to the original size of a given monocyclic peptide, thus leading to better conformational rigidity, target binding affinity as well as selectivity.[27]

1.5 Bicyclic peptides

Bicyclic peptides consist of two macrocyclic rings which impart higher conformation rigidity as compared to monocyclic peptides. Bicyclic peptides can exhibit properties typical of proteins as they have a large interface for interaction with the target and form multiple hydrogen and electrostatic bonds. These properties lead to the enhanced binding affinity and exquisite specificity observed for bicyclic peptides.[2] Bicyclic peptides can broadly be classified into two groups: natural and synthetic.

Natural bicyclic peptides

Natural bicyclic peptides are widely distributed in nature and have diverse biological functions. Phallotoxins and amatoxins are some that are isolated from poisonous mushrooms of the genus amantia. Phalloidin, which inhibits cytokinesis is a phallotoxin isolated from the death cap mushrooms. While amatoxins are bicyclic octapeptides that selectively inhibit mammalian RNA polymerase.[27] Naturally occurring bicyclic peptides have diverse biological functions and have a potential to be used in the development of novel therapeutics. In nature, bacteria, fungi, plants and animals produce natural product bicyclic peptides with diverse biological functions. Actinomycin D and romidepsin are two such naturally occurring bicyclic peptides that are used in cancer therapy. Phalloidin, a toxin extracted from death cap mushrooms is used as a cell imaging reagent.[6]

1.5.1 Properties of bicyclic peptides

Bicyclic peptides have limited number of possible conformers due to their reduced flexibility which lead to better target binding affinity, improved metabolic stability (proteolytic stability) and better permeability [6][1][7][30]. Bicyclic peptides also, due to their conformed spatial arrangement, have the ability to form more hydrogen bonds as compared to their linear and monocyclic counterparts. [2]

Each of the rings of the bicyclic structure can have independent function allowing these peptides to be bifunctional. Bicyclic peptides combine the favorable properties of macromolecules (high target specificity and affinity) with small molecules (stability and good tissue penetration) making them ideal for therapeutics with low toxicity profiles against a diverse set of targets.[1]

1.5.2 Chemical methods for synthesis of bicyclic peptides

There are various strategies that are available for cyclization of peptides. The first reported synthesis of a bicyclic peptide in 1978 which was structurally similar to phallotoxins, was formed by a thioether bond between a cysteine residue and an oxidized tryptophan derivative, subsequently followed by a head-to-tail cyclization. Other cyclization methods make use of intramolecular thioester ligation between a N-terminal cysteine and a C-terminal thioester (N-to-C cyclization), followed by a intramolecular disulfide linkage with an internal cysteine residue. [1]
Ring closing olefin and alkyne metathesis, palladium catalyzed C-H arylation process, Michael addition between thiols and maleimides, amide bonds between carbonyl group of glutamate and N-terminus are other methods to form the bicyclic peptides from a linear precursor. Simpler ways to synthesize bicyclic peptides are by linking two monocyclic peptides with a linker. Another strategy is using a linker (a small molecule) to perform a ring closing reaction of a linear precursor. [1]. Bicyclic peptides can be synthesized by using a tri-functionalized linker which can be tethered to the linear precursor. The tether between the linker and linear peptide can be formed via a variety of bonds but thioether and disulfide linkages are the most commonly used. [1]

1.6 Linkers

The primary role of a linker in a cyclic peptide is to join the ends of the peptide which leads to reduced backbone flexibility. Linkers have been shown to influence the structures of the peptides by positioning the ends of the peptide at specific distances or orientations, non-covalent interactions or steric exclusion that lock peptide backbones into specific conformers.

The most frequently used cyclization strategy is based on the formation of a disulfide bridge between two cysteine residues. Though disulfide bonds, due to their small size and flexibility, cannot confer any additional structural constraint to the bicyclic peptide apart from that achieved by connecting the two ends of the peptide.[7] Using thiol reactive functional groups could help overcome the problems of using disulfide linkages, since thioether bonds are much stronger as compared to disulfide linkages. The high nucleophilicity of the sulfhydryl moiety of cysteine under physiological conditions enables the selective reaction with electrophilic trivalent molecules to generate bicyclic peptides.

1.6.1 Rotational symmetry

Threefold rotational symmetry of the linker ensures that a single product isomer is formed upon conjugation to the peptide. This leads to the formation of a unique structural and spatial isomer of the bicyclic peptide.[7]

1.6.2 Effect of linkers on properties of bicyclic peptide

The nature of the chemical linker plays a vital role in determining the actual conformation(s) and hence the binding affinity of the bicyclic peptides. Higher polarity of the chemical linker can also be beneficial for its solubility. [30] Different functional groups and geometries of the linkers provide diverse environment and affect the properties of the peptides [7]. Bulky substituents on the cyclic core of a linker could hinder the S_N2 substitution reaction with the thiol groups and significantly slow down the reaction of the linker with the peptide.

Linkers impose different peptide backbone conformations in bicyclic peptides. The linkers have a strong structural impact in spite of the flexible cysteine side chains that tether the peptides to the linkers. This effect of the relatively small linkers on the peptide could be explained by its central localization, forming the branching point of the bicyclic peptide. However, larger substituents could significantly slow down the S_N2 substitution reaction of the linker and peptide due to steric hindrance. [7]

The structure of the linker/scaffold plays an important role in determining the shape of the bicyclic peptide. A planar linker would bias the bicyclic peptide towards an overall planar geometry as opposed to globular. The surface area of the molecules and therefore their ability to interact with flat protein targets increases due to the overall planar geometry of the bicyclic peptide. [23]

1.6.3 Non-covalent interactions between linker and peptide loops

Bicyclic peptides as compared to antibodies are nearly 100-fold smaller and thus provide better access to chemical synthesis, efficient diffusion into tissues and oral availability. However, bicyclic peptides have a rather flexible structure and they do not have a folded structure in solution. This conformational flexibility may limit the binding affinity of bicyclic peptides against targets. The conformational rigidity of the bicyclic peptides can be increased by non-covalent interactions between the linker and the peptide loops.[5]

Bicyclic peptides with a central hydrophobic core does not form noncovalent interactions with the peptide loops and does not act as a nucleating scaffold (see **Figure**1.2). However, non-covalent interactions between polar linker and the peptide loops of the bicyclic peptide might pre-organize the peptide loops. This helps reduce the entropy loss of the bicyclic peptide upon binding to a target. The formation of the hydrogen bonds between the organic linkers and the amino acids of the peptide loops can be promoted by using hydrophilic linker molecules containing polar functional groups.[5]

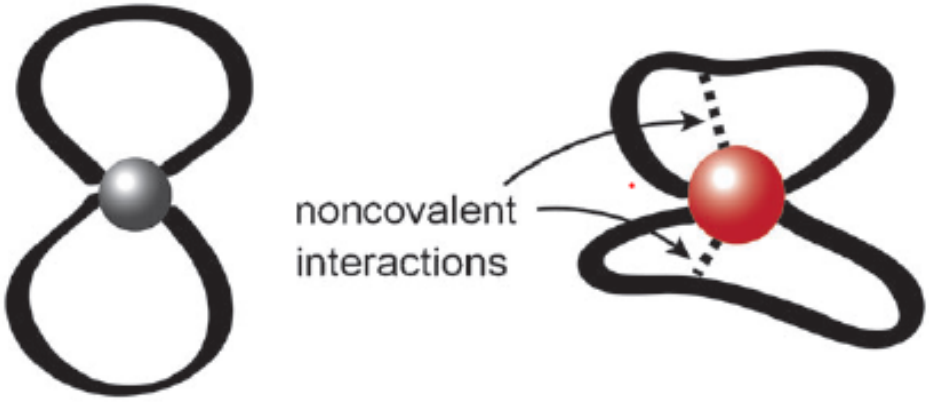


Figure 1.2: Central chemical linker (grey ball) does not form any non-covalent interactions with the peptide loops, while a polar peptide linker (red ball) forms H-bonds with the peptide loops.[5]

2
Aim and Objectives

2.1 Aim

The aim of this project is to synthesize C_3 symmetric, trifunctionalized, thiol reactive linkers for bicyclization of peptides.

2.2 Objective

- A literature study on the current linkers used for peptide bicyclization to understand the effect of linker structure and chemical environment on the properties of the bicyclic peptide.

- Synthesize linkers using halide leaving groups to facilitate the reaction with free thiols of cysteine residues of peptides.

- Synthesize linkers that have not been described in literature or have not been used for bicyclization of peptides.

3
Theory

3.1 Review of linkers

In this section a review of various linkers used in literature for generating bicyclic peptides has been discussed. The major motivation for a review of linkers developed was to understand the various properties and functions that the linkers impart to the bicyclic peptide. The different ways the linker and peptide are conjugated, their chemistry and the applications are discussed below.

3.1.1 TAT/TATA: 1,3,5-triacryloylhexahydro-1,3,5-triazine

TAT is a stable, inexpensive, commercially available and C_3 symmetric compound that possesses three electron deficient olefin groups.

3.1.1.1 Structure

TATA contains a central triazinane core and three acrylamide functional groups at 1-, 3- and 5-positions (see **Figure** 3.1).

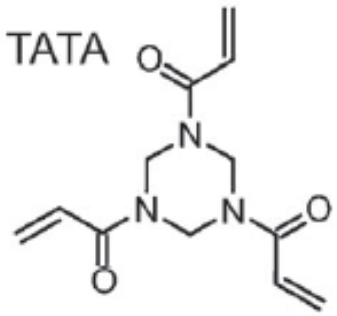


Figure 3.1: Structure of 1,3,5-triacryloylhexahydro-1,3,5-triazine (TATA) molecule [28]

Triazine ring exists in the chair conformation with the three acrylamide groups pointing out toward the same side of the ring to form a bowl like cavity (see **Figure** 3.2). The acrylamide groups invert upon flipping of the chair conformation.[28] The amide bond of the acrylamide substituents has two possible conformations and can isomerize between them. This was also corroborated from NMR experiments. The NMR spectra of a bicyclic peptide with TATA in solution show a pair of spin

systems, which suggests the presence of two conformers of the bicyclic peptide. This was attributed to the conformations of the linker molecule, which could be due to flipping of the triazinane ring or due to isomerization of the amide bonds in the linker. [7]

Figure 3.2: Three-dimensional structure of TATA. [28]

3.1.1.2 Non-covalent interactions

TATA has three potential H-bond acceptor groups (see **Figure** 3.3 but was shown not to form non-covalent bonds with the peptide loops of the bicycle peptide (with a sequence H-A<u>C</u>SDRFRN<u>C</u>PADEAL<u>C</u>G-NH_2).

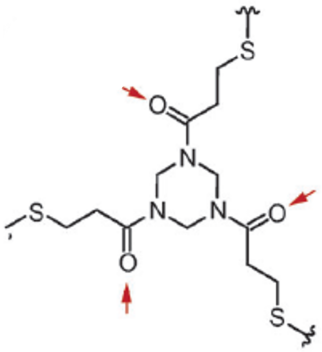


Figure 3.3: Structure of TATA after reacting with thiol containing compounds. Potential H-bond acceptors are indicated by red arrows) [5].

3.1.1.3 Reaction of linker with thiols

TATA possesses three electron deficient olefin groups that can undergo the thiol-ene click reaction. This compound was shown to react stoichiometrically with a wide range of aliphatic and aromatic side chain thiols, the general reaction scheme is shown in **Figure** 3.4.

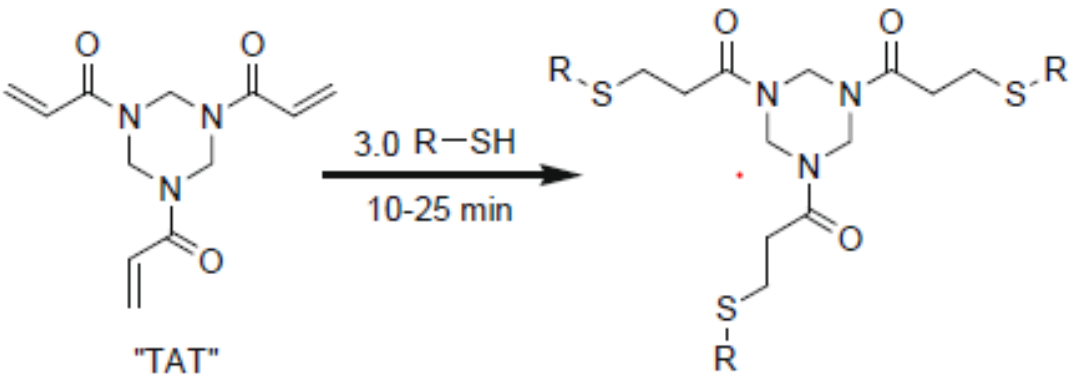


Figure 3.4: General reaction scheme for thiol-ene click reaction between TAT/ TATA and thiols containing various side chain substituents, R (aliphatic or aromatic) [28]

Reaction conditions

The thiol-ene click reaction was carried out at room temperature and resulted in the product in high purity and yield. TATA though insoluble in the solvent (methanol), was shown to disappear a few minutes after the start of reaction. The click reaction was quick (<10 min) and required very minimal purification. Addition of a catalytic amount of *n*-propylamine was necessary for reaction with aliphatic terminal alcohol thiols. No catalyst was required if amino groups are present which can self-catalyze this reaction. Amino groups present in the starting compound was not shown to react with TATA. The amine catalyzed reactions were not successful for thiols containing a carboxylic acid. A probable reason was attributed to the amine catalyst being protonated by the carboxylic acid groups. Deprotonation by heating the thiols in a solvent (methanol) in the presence of potassium bicarbonate resulted in good yields. No amine catalyst was required for the thiol-ene reaction occurred spontaneously under these reactions. [28]

3.1.1.4 Reaction with cysteine

One of the compounds that was tested by the authors was the thiol-ene click reaction between the amino acid DL-cysteine and TATA. A methanol-water solvent system was necessary to aid in the solubility of cysteine. An amine catalyst was not required and a spontaneous reaction occurred between the deprotonated cysteine molecule and TATA resulting in 88% yield. [28] The thiol-ene click reaction resulted in high product yields, minimal workup and short reaction time. Moreover, the there was no necessity for air or water free reaction conditions which makes it an attractive reaction for use in linker design.[28]

3.1.1.5 Reaction with model peptides containing thiol reactive groups

Treatment of a model peptide (H-AC̲SDRFRNC̲PADEALC̲G-NH_2) containing three cysteine residues with the linker TATA resulted in formation of desired bicyclic peptide (see **Figure** 3.5) with mass corresponding to the respective peptide. TATA

undergoes thiol-ene click reactions with cysteine and coupling of one thiol to the acrylamide groups would not result in a change in mass. Hence, it is not possible from a change in mass to derive the number of acrylamide functional groups that had reacted with the peptide. To determine if the all three acrylamide functional groups of the linker had reacted to thiol groups of the peptide an indirect method can be used by adding an excess of free cysteine to the reaction mixture. The unreacted acrylamide groups would react with the free cysteine resulting in a change of mass.[7]

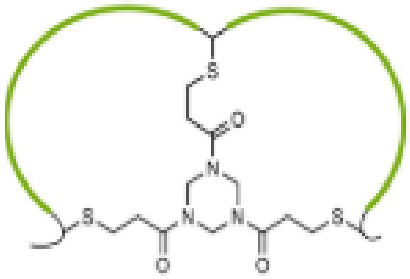


Figure 3.5: Bicyclic peptide formed via cyclization of the model peptide with TATA. [30]

TATA at 10 µM concentration reacted quantitatively with the model linear peptide to form the corresponding bicyclic peptide (see **Figure**3.6. [7]

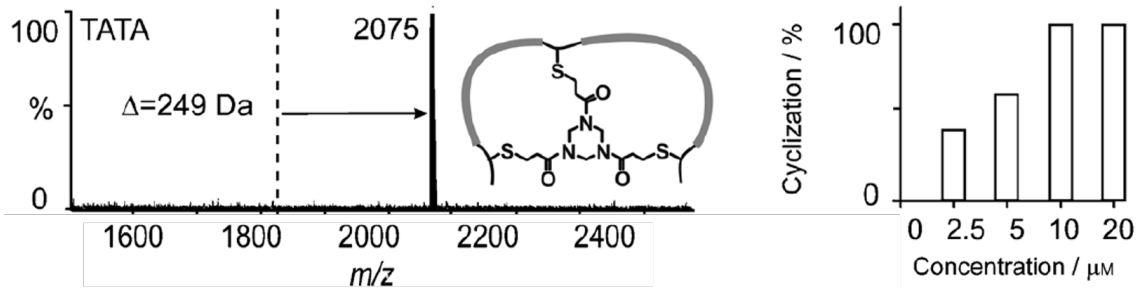


Figure 3.6: The mass spectrum for the reaction products of TATA and a model peptide, H-A$\underline{C}$SDRFRN$\underline{C}$PADEAL$\underline{C}$G-NH_2 and estimations of percentages of cyclized peptides for various linker concentrations. [7]

Only the desired mass bicyclic peptide corresponding to each linker was observed and no other product masses were detected. No intermolecular reaction products were observed wherein multiple peptides were crosslinked by the linkers. Even at excess linker concentrations over peptides, every linker molecule had tethered to only one peptide molecule.[7]

Reaction with peptide containing only two thiol groups

Treatment of a peptide with only two cysteine groups and TATA results in two products: a monocyclic peptide with one unreacted acrylamide functional group and a bicyclic peptide with TATA linked to two cysteine units and another amino acid. The third acrylamide group remains unreacted in some of the reaction product

indicating that TATA is more reactive towards thiols than towards other nucleophiles present in the peptide.[7]

3.1.2 TPSMB: 1,3,5-tris((pyrididn-2-yldisulfanyl)methyl) benzene

TPSMB is a planar and trivalent linker (see **Figure** 3.7) that is sulfhydryl specific that facilitates reversible cyclization and linearization via disulfide bond formation and cleavage of bicyclic peptides of the format CX_nCX_nC, where C is cysteine and X is any amino acid except cysteine. [12]

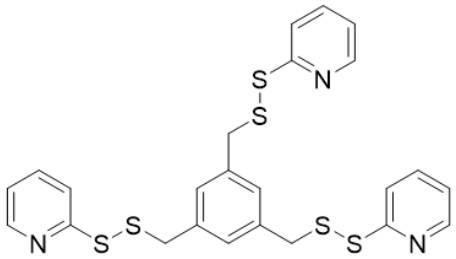


Figure 3.7: Structure of TPSMB molecule

The major application of TPSMB as suggested by the authors, was the idea of switching a peptide molecule, *in vivo* between an active linear form and an inactive bicyclic configuration under physiological conditions.[12]

3.1.2.1 Reversible cyclization

TPSMB enables the generation of bicyclic peptides via three pyridyl-activated disulfide groups.
It rapidly undergoes a highly sulfhydryl specific reaction with suitable linear peptides under physiological conditions to form bicyclic peptides. The linear peptide can be obtained back by reduction of disulfide bond by mild reducing agents like glutathione and cleavage of the linker from the bicyclic peptide. This reversible disulfide mediated cyclization provides a general strategy for delivering linear peptide into mammalian cells. In the oxidizing extracellular or endosomal environment, the peptide exists as a cyclic and conformed peptide with enhanced proteolytic stability and cell permeability. The cyclic peptide upon entering the intracellular space converts into the linear peptide due to the reduction of the disulfide bonds by high levels of glutathione present and subsequently cleaving the linker. [12]

To replicate the conditions of the *in vivo* strategy, an *in vitro* method was investigated to study the reductive stability of a TPSMB modified peptide against glutathione levels found in human plasma and in the cytosol (see **Figure** 3.8) structure in 1 h under the cytosol conditions. While, under the plasma glutathione levels, 85% of the peptides remained in their bicyclic structure after 1 h and only 50% had been converted to the linear form after 6 h. This technique presents itself as a viable option for providing peptides with enhanced proteolytic stability and cell

permeability due to the bicyclic structure which can be reversibly converted to the linear form by cleaving the TBSMB linker molecule. [12]

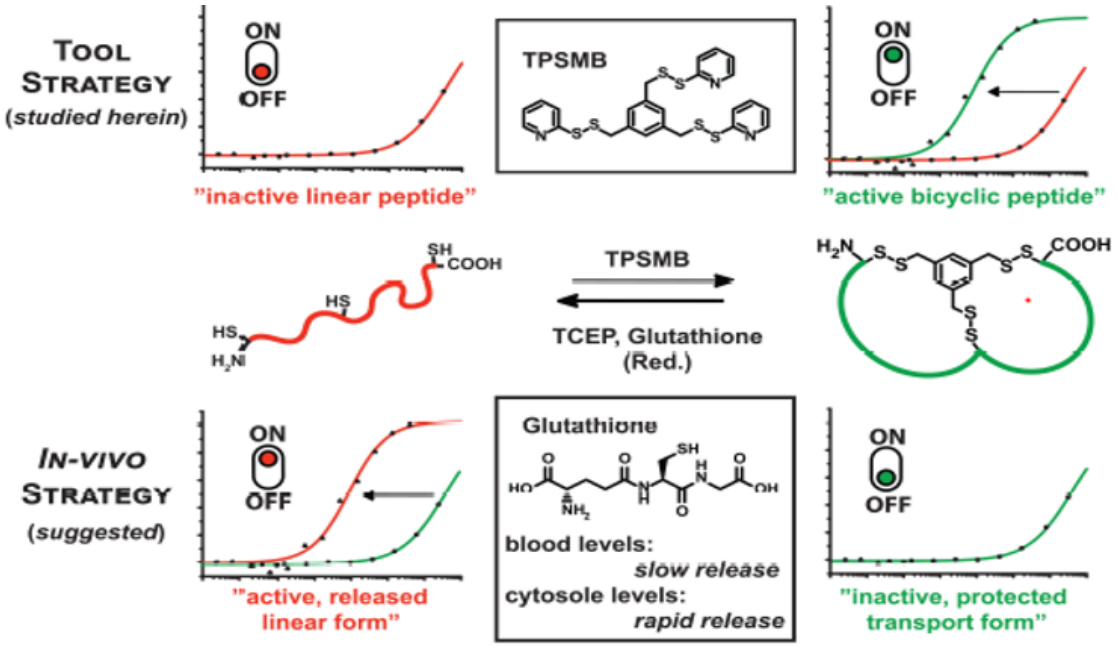


Figure 3.8: Graphical representation of the reaction between TPSMB and a peptide containing thiol groups. The bicyclic peptide formed can be converted back into the linear peptide by selectively reducing the disulfide bonds. Based on the additive *in vitro* or the cellular environment the equilibrium can be switched from linear to cyclized and vice versa.[12]

3.1.2.2 Applications

Possible applications of this technique include capturing/releasing the target on beads, columns of surfaces since TPSMB allows for reversible switch between a bicyclic conformation and a linear conformation upon cleavage of a linker as compared to other classical trivalent linkers which lead to a permanent modification of the peptide to form a bicyclic structure. [12]

3.1.3 TBMB: 1,3,5-tris(bromomethyl)benzene

TBMB contains three thiol reactive groups (see **Figure**3.9) and reacts with cysteine residue in peptides under mild conditions to form bicyclic peptides. It can be used as a scaffold to anchor peptides containing three cysteine residues.

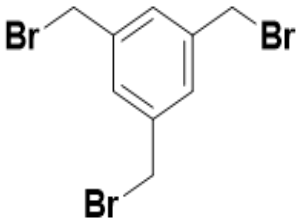


Figure 3.9: Structure of TBMB containing three bromine leaving groups which are thiol reactive.

The threefold rotational symmetry of the TBMB molecule leads to the formation of a unique structural and spatial product isomer upon reaction with a linear peptide, which was confirmed by NMR.[17] The NMR spectra of the bicyclic peptide cyclized with TBMB in solution shows a single spin system indicating a single conformer.[7]

3.1.3.1 Reaction with a model peptide

Treatment of a model peptide (H-A$\underline{C}$SDRFRN$\underline{C}$PADEAL$\underline{C}$G-NH_2) containing three cysteine residues with the linker TBMB resulted in formation of desired bicyclic peptide (see **Figure** 3.10) with mass corresponding to the respective peptide.[7]

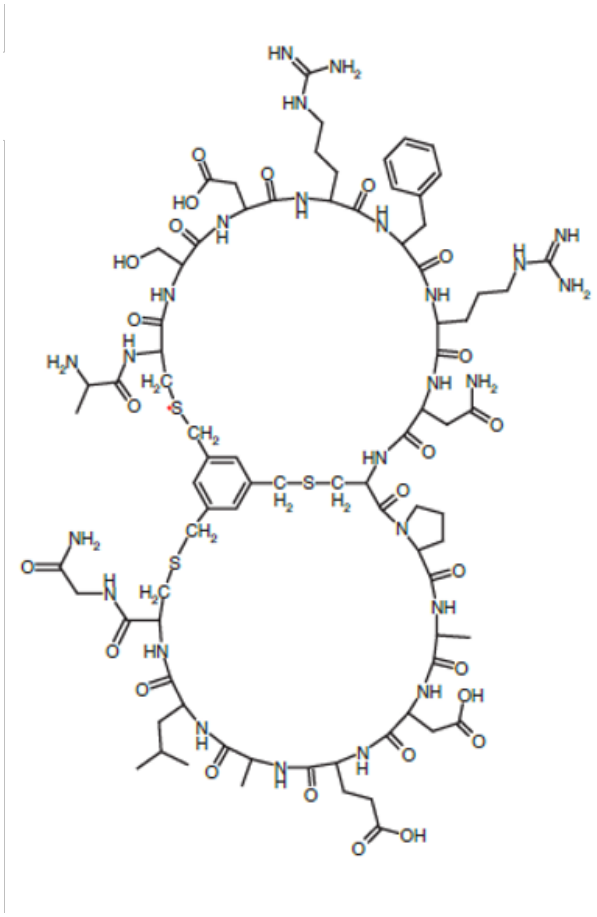


Figure 3.10: Structure of bicyclic peptide product formed by the reaction of the linker TBMB with the linear peptide H-A$\underline{C}$SDRFRN$\underline{C}$PADEAL$\underline{C}$G-NH_2.[7]

TBMB react with the thiol groups via a S_N2 substitution reaction with bromine as the leaving group. Therefore, the masses of the products could indicate the number of thiols that had reacted with the linkers.[7]
TBMB at 2.5 µM concentration reacted quantitatively with the model linear peptide to form the corresponding bicyclic peptide (see **Figure** 3.11).[7]

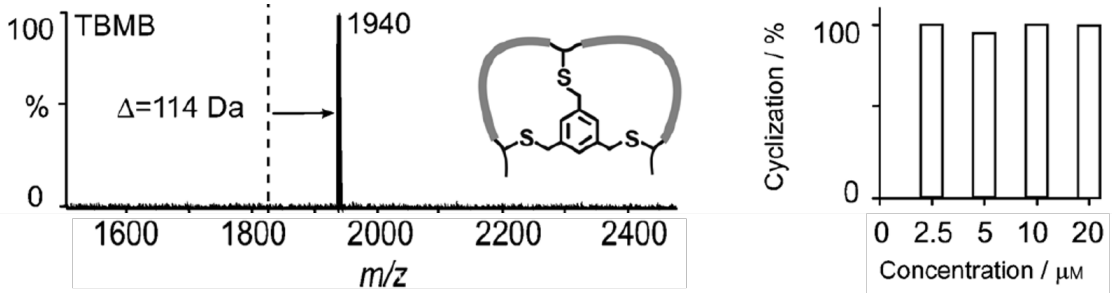


Figure 3.11: The mass spectrum for the reaction products of TBMB and a model peptide, H-ACSDRFRNCPADEALCG-NH_2 and estimation of percentages of cyclized peptides for various linker concentrations.[7]

Only the desired mass bicyclic peptide corresponding to each linker was observed and no other product masses were detected. No intermolecular reaction products were observed wherein multiple peptides were crosslinked by the linker. Even at excess linker concentrations over peptides, every linker molecule had tethered to only one peptide molecule.[7]

Intramolecular side reactions

Excess TBMB (1.5 eq.) reacts intramolecularly with bicyclic peptide upon extended reaction time (>10 min).
TBMB was shown to preferentially react with thiol groups of cysteine but in the absence of cysteine also reacted with primary amines.[17] In the reaction of TBMB with peptides containing only 2 cysteine residues, the third bromomethyl group of the linker reacts intramolecularly with the primary amino group of the N-terminus to form a bicyclic peptide as the major product. Other minor products formed could be the monocyclic peptide with a bromomethyl group that has either not reacted or has been hydrolyzed.[7]

3.1.3.2 Non-covalent interactions

TBMB is a linker with a with a hydrophobic benzene core and does not form any hydrogen bonds or other non-covalent interactions with the amino acids of the peptide loops. Hence, the mesitylene core of TBMB is not an ideal candidate to serve as a structural scaffold.
Bicyclic peptide with the TBMB linker showed no covalent interactions between peptide loops and the mesitylene core.[7]

3.1.3.3 T3: 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene

T3 is a linker with similar structure to TBMB with additional substituents on the benzene core. T3 has three methyl substituents at 2-, 4- and 6-positions of the benzene ring (see **Figure** 3.12).

3.1.3.4 Reaction of T3 with a model peptide

The reaction of the organic scaffold, 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (T3) with the linear peptide (<30 amino acids long) containing 3 cysteine residues was shown to be an extremely fast and clean reaction (see **Figure** 3.12). Treatment of 0.05 mM solution of linear peptide (**5**, see **Figure** 3.12) consisting of 3 thioether groups with 1.05 equiv. of T3 linker gave the corresponding product (**6**, see **Figure** 3.12) in 5 min with a 90% yield. The C_3 symmetry of the linker leads to the formation of a single product with no regioisomers. [29]

Figure 3.12: Overall reaction scheme for the reaction of a linear peptide with 3 thioether groups and 1,3,5-tris(bromomethyl)-2,4,6-trimethylbenzene (T3) linker to form the bicyclic peptide [29]

T3 was also determined to be compatible with all possible side chain functional groups except free cysteine. Hence, avoiding the need for complex synthetic strategies and making this a versatile reaction with a wide scope. [29]

Stepwise cyclization

The cyclization occurs in a stepwise manner: initial formation of the linear intermediate with the linker (**7**, **Figure** 3.13) and intramolecular cyclization to give the monocyclic peptide (**8**, **Figure** 3.13) followed by the bicyclic peptide. The high reactivity of T3 linker and the intermediates (**7** and **8**, **Figure** 3.13) was attributed to the electron donating effect of the three methyl groups on the benzene ring. [29]

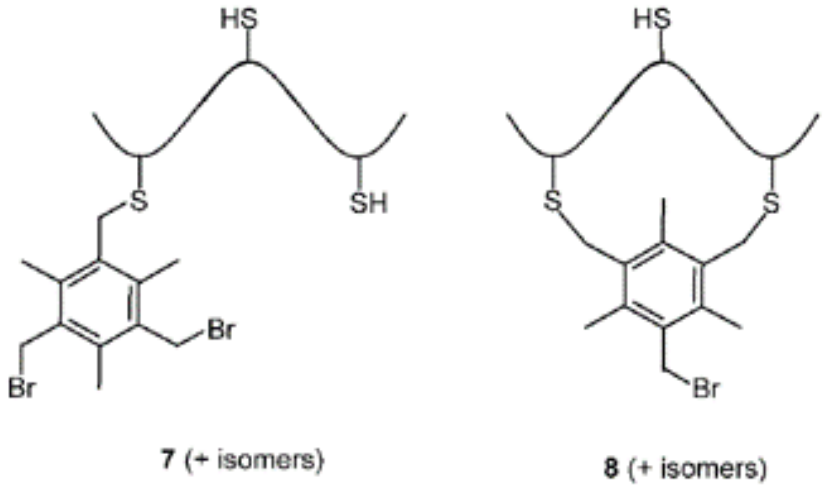


Figure 3.13: Structure of various intermediates formed during the peptide cyclization reaction between linker T3 and a model peptide.[29]

3.1.4 TBAB:*N,N',N"*-(benzene-1,3,5-triyl)tris(2-bromoacetamide)

3.1.4.1 Structure

TBAB contains a central benzene ring containing three bromoacetamide functional groups at 1-, 3- and 5-positions (see **Figure** 3.14. The linker is not entirely rigid since the amide groups linked to the benzene core can rotate around the C-N bonds. However, due to steric effects and overlapping of benzene ring and amide orbitals, some rotational angles are energetically favored.[7]

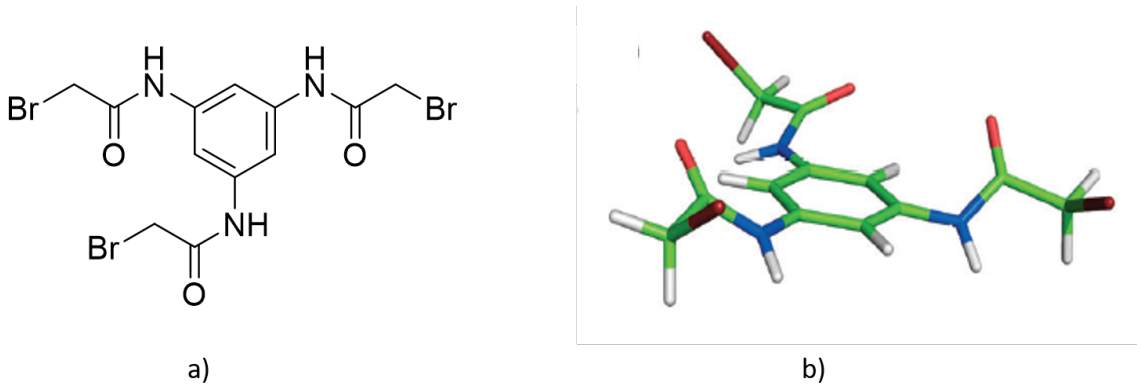


Figure 3.14: a) Chemical structure and b)three-dimensional structure of TBAB consisting of a benzene core and three bromoacetamide side chains [7].

3.1.4.2 Reaction with model peptide

Treatment of a model peptide (H-A$\underline{C}$SDRFRN$\underline{C}$PADEAL$\underline{C}$G-NH_2) containing three cysteine residues with the linker TBAB resulted in formation of desired bicyclic peptide with mass corresponding to the respective peptide. [7]
TBAB at 2.5 µM concentration reacted quantitatively with the model linear peptide to form the corresponding bicyclic peptide (see **Figure**3.15). [7]

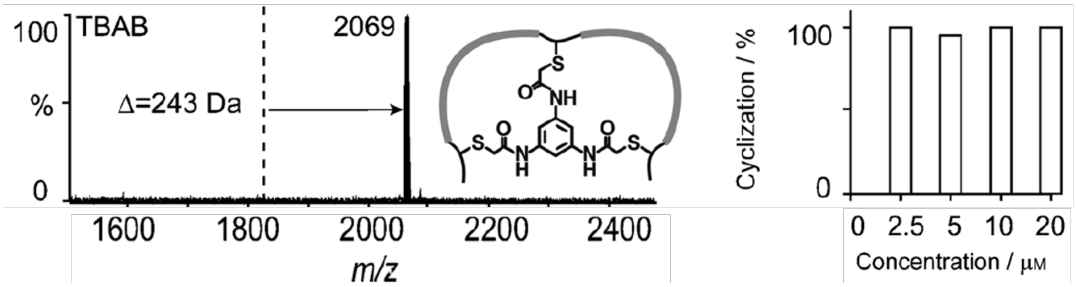


Figure 3.15: The mass spectrum for the reaction products of TBAB and a model peptide, H-A$\underline{C}$SDRFRN$\underline{C}$PADEAL$\underline{C}$G-NH_2 and estimation of percentages of cyclized peptides for various linker concentrations. [7]

TBAB reacts with the thiol groups via a S_N2 substitution reaction with bromine as the leaving group. Therefore, the masses of the products could indicate the number of thiols that had reacted with the linkers. The NMR spectra of the bicyclic peptide cyclized with TBAB in solution shows a single spin system indicating a single conformer. [7]

The bromoacetamide functional group specifically reacts thiols in aqueous buffer at moderate temperature [7]. Hence, only the bicyclic peptide with desired mass corresponding to each linker was observed and no other product masses were detected. No intermolecular reaction products were observed wherein multiple peptides were crosslinked by the linkers. Even at excess linker concentrations over peptides, every linker molecule had tethered to only one peptide molecule. [7]

Reaction with model peptide containing 2 cysteine residues

Treatment of a peptide containing only 2 cysteine groups with TBAB resulted in a single product. However, the product mass does not correspond to a monocyclic peptide with either an unreacted or a hydrolyzed bromoactamide group. The possible explanation given by the authors was due to a modification of the bromoacetamide group but the nature of the product could not be determined from the mass. [7]

3.1.4.3 Non-covalent interactions

TBAB has a large number of polar groups (see **Figure** 3.16) and has a high potential for forming non-covalent bonds with the amino acids of the bicyclic peptide loops. TBAB was shown to form multiple non-covalent interaction between the linker and

the peptide. The planar amide groups of TBAB form H-bonds with amino acids in the peptide.[5]

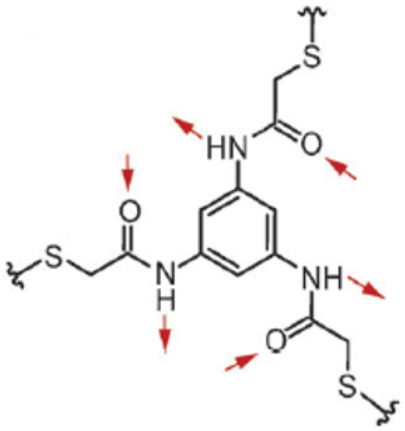


Figure 3.16: Structure of TBAB after reacting with thiol containing compounds. Potential H-bond donors and acceptors are indicated by red arrows.[5]

Hydrophilic linker molecules can form intramolecular non-covalent interactions with the peptide loops of the bicyclic peptide. These could potentially help stabilize the three dimensional structure of the bicyclic peptide in solution. The intramolecular H-bonds formed between TBAB and the peptide loops, might allow the transient pre-organization of the peptide in solution rather than stabilize the peptide to adopt a defined structure in solution. This can help the bicyclic peptides to more closely resemble its target and bind better than entirely flexible peptides. [7]

3.1.5 TAAB: *N,N',N"*-benzene-1,3,5-triyltrisprop-2-enamide

3.1.5.1 Structure

TAAB contains a central benzene ring containing three acrylamide functional groups at 1-, 3- and 5-positions (see **Figure** 3.17). [7]

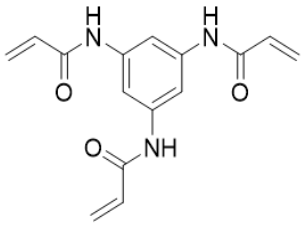


Figure 3.17: Structure of TAAB.

3.1.5.2 Reaction with model peptide

Treatment of a model peptide (H-ACSDRFRNCPADEALCG-NH_2) containing three cysteine residues with the linker TAAB resulted in formation of desired bicyclic

peptide with mass corresponding to the respective peptide. [7]

To determine if all three acrylamide functional groups of the linker had reacted with thiol groups of the peptide, an indirect method can be used by adding an excess of free cysteine to the reaction mixture. The unreacted acrylamide groups would react with the free cysteine resulting in a change of mass. [7] TAAB could only modify 10% of the model peptide even at 20 µM concentration of the linker to form the corresponding bicyclic peptide (see **Figure** 3.18.[7]

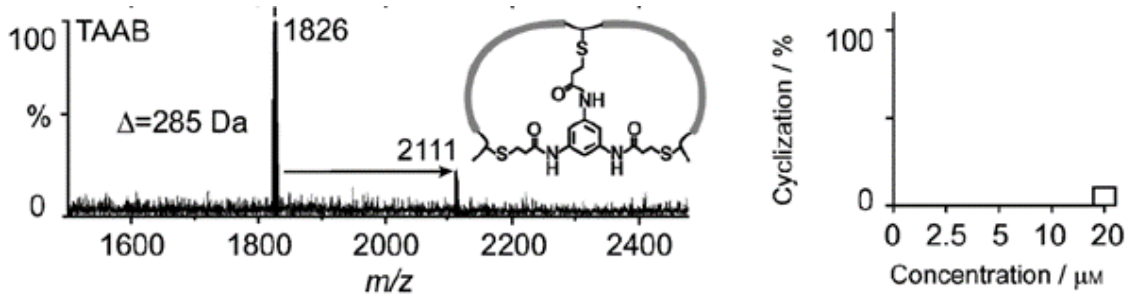


Figure 3.18: The mass spectrum for the reaction products of TAAB and a model peptide, H-AĊSDRFRNĊPADEALĊG-NH_2 and estimation of percentages of cyclized peptides for various linker concentrations. [7]

Only the desired mass bicyclic peptide corresponding to each linker was observed and no other product masses were detected. No intermolecular reaction products were observed wherein multiple peptides were crosslinked by the linkers. Even at excess linker concentrations over peptides, every linker molecule had tethered to only one peptide molecule. [7]

3.1.6 TTA: triacryl-tris-alkyne

3.1.6.1 Structure

TTA is a small organic linker consisting of four functional groups: three spatially isometric thiol reactive groups to form two loops with a linear peptide and another one for chemically coupling with molecular cargoes via click chemistry (see **Figure** 3.19) [22]

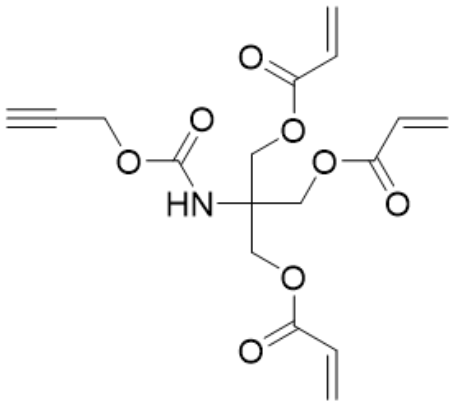


Figure 3.19: Structure of tetra-functional linker TTA.

3.1.6.2 Advantage of using TTA

The presence of multiple amino groups on a bicyclic peptide while conjugation with biomolecules would result in a mixture of biomolecule labelled bicyclic products. TTA helps overcome this problem by formation of a single bicyclic peptide-biomolecule conjugate when there are multiple amino groups in the peptide (see **Figure** 3.20). [22]

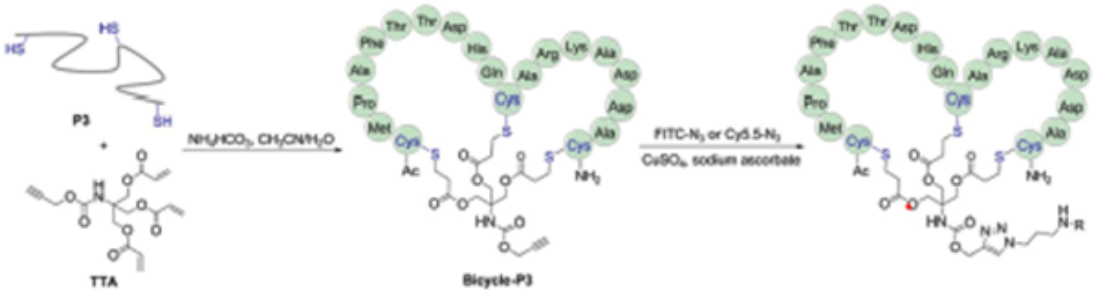


Figure 3.20: Illustration of stepwise reaction between a linear peptide with three cysteine residues (P3) and tetra-functional linker (TTA) to form bicyclic peptide (Bicyclic-P3) followed by conjugating with fluorescent dyes via click chemistry to provide FITC/Cy5.5-monolabeled bicyclic peptides (FITC-bicycle-P3 and Cy-bicycle-P3).[22]

3.1.7 TATB: 1,1',1"-(1,3,5-triazinane-1,3,5-triyl) tris(2-bromoethan-1-one)

3.1.7.1 Structure

TATB has threefold rotational symmetry and contains three sulfhydryl reactive bromoacetyl groups attached to a triazinane core (see **Figure** 3.21)

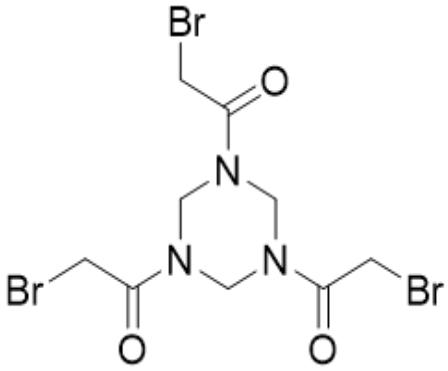


Figure 3.21: Structure of TATB linker.

The triazinane core adopts a chair like conformation with the three amide groups pointing in the same direction to form a bowl-like conformation (see **Figure** 3.22). [30]

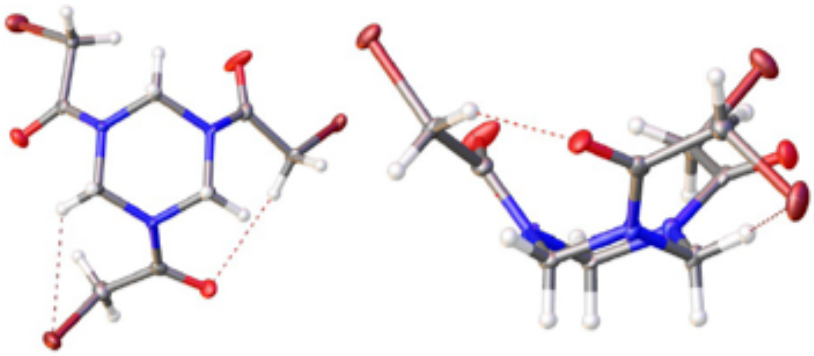


Figure 3.22: X-ray structure displaying the bowl shape of TATB linker. [30]

3.1.7.2 Reaction with model peptide

Treatment of a model peptide (H-AC̲SDRFRNC̲PADEALC̲G-NH_2) containing three cysteine residues with the linker TATB resulted in formation of desired bicyclic peptide with mass corresponding to the respective peptide (see **Figure** 3.23). TATB reacts with the thiol groups of the peptide via a S_N2 substitution reaction with bromine as the leaving group. Therefore, the masses of the products could indicate the number of thiols that had reacted with the linkers due to the absence of the leaving group. [7]

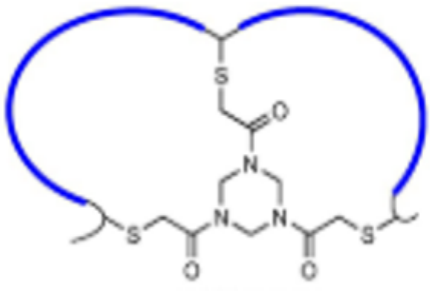


Figure 3.23: Bicyclic peptide formed via cyclization of the model peptide with TATB. [30]

3.1.7.3 Comparison with other linkers

Bicyclic peptide created using TATB, had an inhibitory constant (K_i) value that was of 100 times higher than the bicyclic peptide created using TBMB with the same model peptide, H-AĊSDRFRNĊPADEALĊG-NH_2.[30]

TATB has a similar polarity as TATA, but was shown to be considerably less active, which can be due to the presence of the an extra methylene unit in each of the arms of TATA giving it additional flexibility.[30]

3.1.8 TBMT: 2,4,6-tris(bromomethyl)-*s*-triazine

3.1.8.1 Structure

TBMT has threefold symmetry and has three sulfhydryl reactive groups bromomethyl groups attached at the 1-, 3- and 5-position of *s*-triazine core (see **Figure** 3.24a).

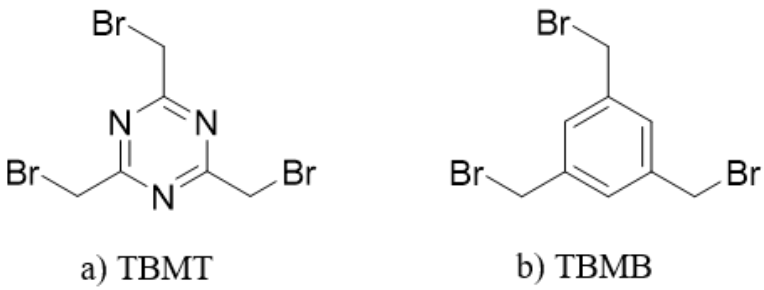


Figure 3.24: Structure of linkers: a) TBMT and b) TBMB linker.

TBMT resembles the widely used TBMB linker (see **Figure** 3.24b) with the difference in the cores of the linkers. TBMB has a benzene core while TBMT has a *s*-triazine core. The core of TBMT is more polar, could impart higher solubility and participate in H-bonding as compared to the hydrophobic mesitylene core of TBMB. [30]

3.1.8.2 Reaction with model peptide

Treatment of a model peptide (H-ACSDRFRNCPADEALCG-NH_2) containing three cysteine residues with the linker TBMT resulted in formation of desired bicyclic peptide with mass corresponding to the respective peptide (see **Figure** 3.25).

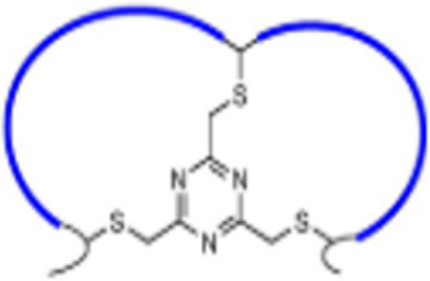


Figure 3.25: Bicyclic peptide formed via cyclization of the model peptide with TBMT. [30]

TBMT react with the thiol groups via a S_N2 substitution reaction with bromine as the leaving group. Therefore, the masses of the products could indicate the number of thiols that had reacted with the linkers. [7]

The bicyclic peptide created using TBMT, had a inhibitory constant (K_i) value that was 10 times higher than the bicyclic peptide created using TBMB with the same model peptide, H-ACSDRFRNCPADEALCG-NH_2. [30]

3.1.9 NATBA: *N,N',N",N"'*-(adamantane-1,3,5,7-tetrayl) tetrakis(2-bromoacetamide)

3.1.9.1 Structure

NATBA is a tetravalent, tetrahedral symmetric linker with four reactive bromoacetyl groups linked to a adamantane core. The bromoacetyl groups are linked to the four bridgehead carbon atoms of the adamantane structure by nitrogens carrying acyl groups attached by amide bonds (see **Figure** 3.26). Three of the bromoacetyl groups of NATBA can be reacted with cysteine residues of a peptide via thioether bonds to form a bicyclic peptide, leaving the fourth group unmodified. The fourth bromoacetyl group can be functionalized by various sulfhydryl bearing groups, linkers, tools or drugs. [13]

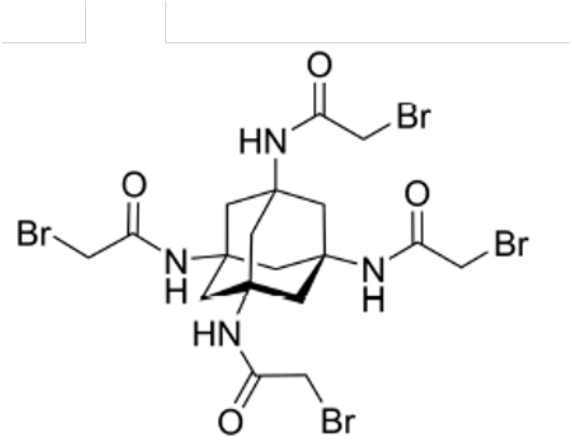


Figure 3.26: Structure of NATBA. [13]

3.1.9.2 Reaction with model peptide

A model peptide with a sequence H-ACEGMINSCEKSDYECG-NH_2 was used for cyclization with NATBA. The reaction rate of cyclization was found to high at pH values closer to that of pK_a of cysteine. However, buffering at a pH of 8 increased the reaction rate and reduced the reaction time (<10 min) [13]. After cyclization of the peptide, the fourth unreacted bromoacetyl group of the NATBA linker can be reacted with additive like FlcSH, BiotinSH, etc. to functionalize the bicycle peptide (see **Figure** 3.27).

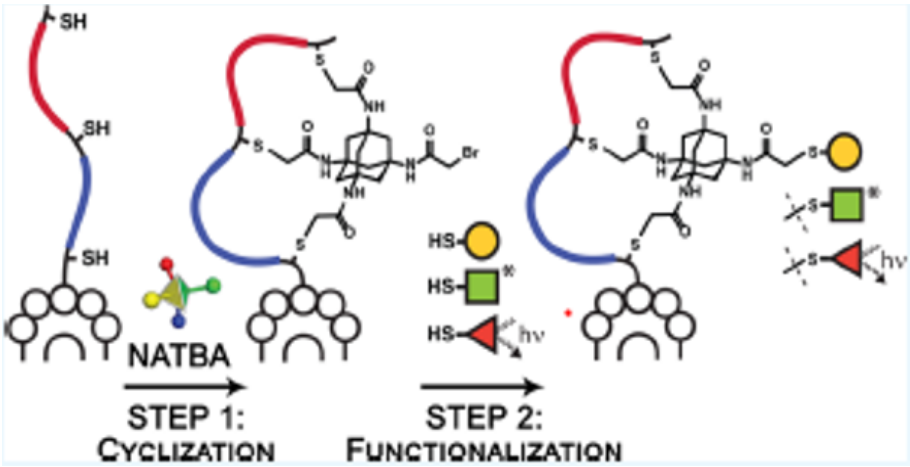


Figure 3.27: General reaction scheme and application of NATBA for peptide cyclization and functionalization. [13]

Pre-coupling the NATBA linker with an additive before cyclization was not preferred due to multiple coupled by-products, low yield and difficulty in isolating the singly coupled product. [13]

TBMB reacts with peptides with three cysteine residues to only generate one single

species of bicyclic peptides. While the reaction of NATBA, a tetravalent linker with the same type of peptide results in two stereoisomers. [13]

3.1.10 Trimesic acid

Trimesic acid (see **Figure** 3.28) is a planar molecule and has been employed in peptide cyclization.

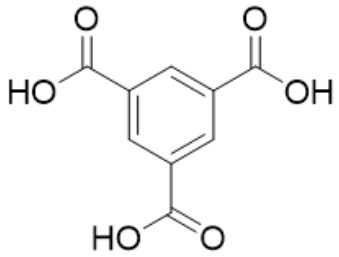


Figure 3.28: Structure of trimesic acid.

Cyclization was mediated by the formation of three amide bonds between trimesic acid and the peptide (see **Figure** 3.29) [23].

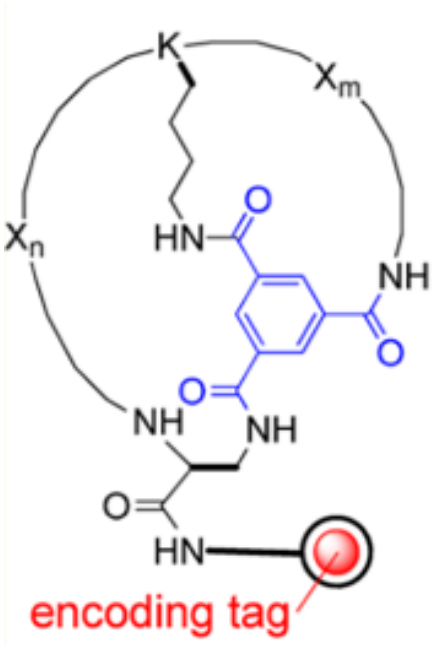


Figure 3.29: Generic bicyclic peptide formed using the linker trimesic acid which binds to the peptide via amide bonds [23].

4

Results and Discussion

In this section is discussed the design, properties and synthesis of four linkers TCPA, BTBE, TBCA and NTCA.

4.1 Design and properties of linkers

4.1.1 General design

Each of the trivalent linkers were designed with a 3-fold rotational symmetry, which will lead to the formation of a single product isomer when conjugated to a linear peptide and prevent polymorphism. The linker (see **Figure** 4.1) is designed containing either a central core molecule or an atom. Attached to the central core are three side arms containing various chemical groups and substituents and at the end of each side arm is a halide leaving group.

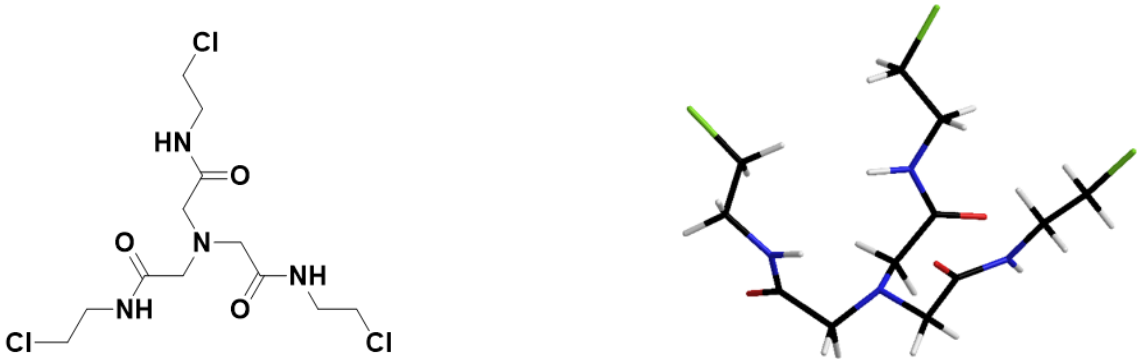
Figure 4.1: Schematic representation of a trivalent thiol reactive linker.

4.1.2 Design and properties of synthesized linkers

a) TCPA **5**

b) BTBE **9**

c) TBCA **16**



d) NTCA **21**

*Element corresponding to the colours in the 3D structure: black-carbon, blue-nitrogen, red-oxygen, green-chlorine, white-hydrogen, pink-bromine.

Figure 4.2: Structures and three-dimensional structures of the synthesized linkers a) TCPA **5**, b) BTBE **9**, c) TBCA **16** and d) NTCA **21** illustrating the different spatial arrangements of the reactive linker arms. The three-dimensional conformations are based on MM2 energy minimization calculations.

4.1.2.1 Central core of the linker

TCPA and NTCA (see **Figure** 4.2a and d) contain a central nitrogen atom as the core of the linker. The central nitrogen atom contains a lone pair of electrons and could potentially form hydrogen bonds with the peptide. This can further aid in conformational stabilization of the bicyclic peptide (see **Figure**1.2). The nitrogen atom at the core is sterically hindered by the side chain substituents and is less susceptible to form H-bonds with the peptide loops. However, the nitrogen core can increase the polarity of the molecule and leading to higher solubility in aqueous solvent as opposed to a hydrophobic core.

BTBE and TBCA (see **Figure** 4.2b and c) contain a benzene central core with the three side arms connected at 1-, 3- and 5-positions of the benzene core. The central core is planar and might impart the linker a slightly planar structure. This might influence the bicyclic peptide to have a more planar geometry as compared to a globular one [23].

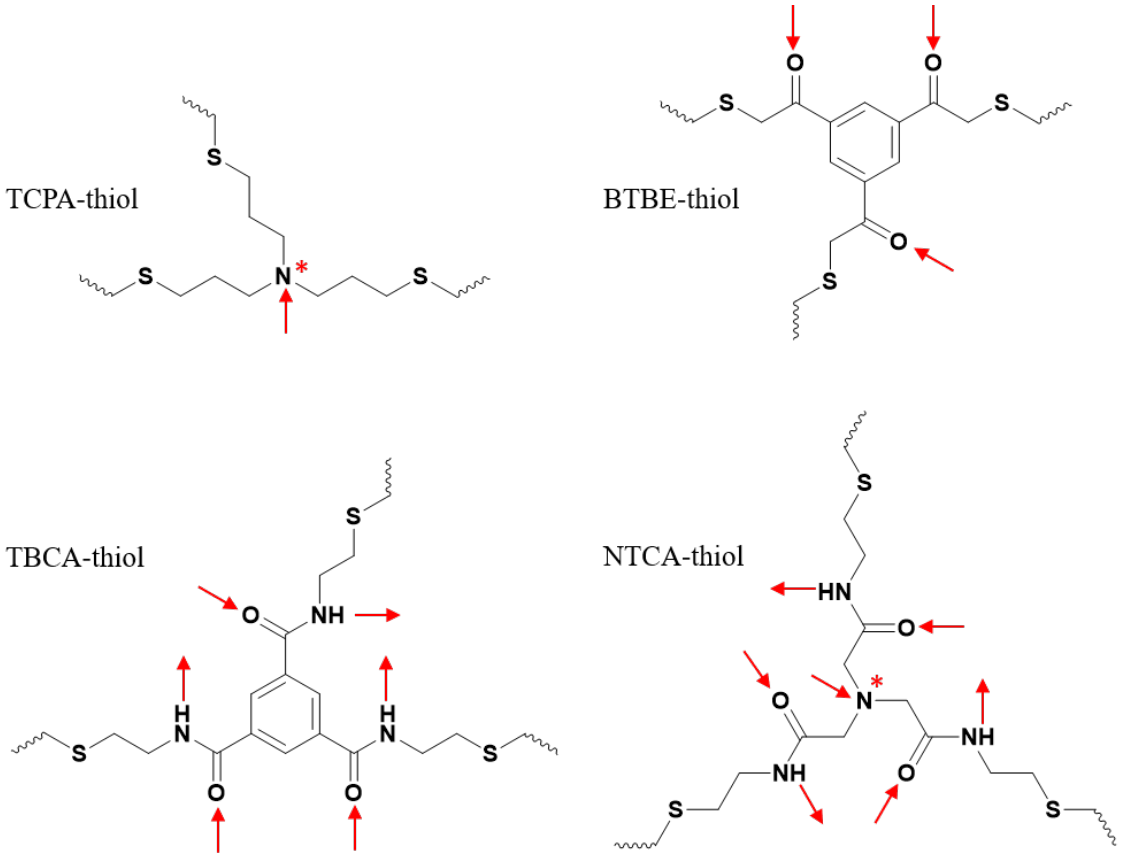
4.1.2.2 Side chain substituents

TCPA contains of three *n*-propyl groups, while BTBE contains three bromoacetyl groups and both TBCA and NTCA contain three electrophilic chloroethyl acetamide groups in the side arms. Out of the four linkers BTBE has the shortest side arm while NTCA has the longest side arm. Since bulky substituents on the side arm can potentially slow down the S_N2 reaction between the linker and the peptide [7], BTBE might react faster with a peptide as compared to NTCA.

4.1.2.3 Leaving groups

TCPA, TBCA and NTCA have chloride while BTBE has bromide as leaving group in their side arms. Both chloride and bromide are good leaving groups but due to the larger size of bromide atoms, they are substituted faster during S_N2 reaction between the linker and the peptide. These halide leaving groups were shown to react with free thiol groups of cysteine residues in the peptide preferentially but when in excess it was shown to also react with the primary amine of the peptide.

4.1.2.4 Non-covalent interactions



*** Depending on the pH, the central nitrogen atom, N* of TCPA and NTCA can get protonated and become a H-bond donor.**

Figure 4.3: Structure of the linkers after reacting with thiol, after substitution of halide groups by sulfur to form thioether bonds. Potential H-bond donors and acceptors are indicated by red arrows.

The side arms of TCPA are hydrophobic and do not take part in H-bonding. BTBE has one while both TBCA and NTCA have two potential H-bond donor/acceptor groups per side arm respectively as shown in **Figure** 4.3. The bicyclic peptide becomes more constrained when there are larger number of non-covalent interactions between the linker and the peptide chain. Also, these contacts could potentially help stabilize a three-dimensional structure in solution, which in turn leads to lower entropy loss when binding to targets resulting in higher binding affinities. [6][5]

The nitrogen core of TCPA and NTCA can also take part in H-bonding but more importantly increase the hydrophilicity of the linker. The presence of polar groups in the side chains of linkers BTBE (carbonyl), TBCA (amide) and NTCA (amide) increases the polarity and hydrophilicity of the linker as compared to a hydrophobic side chain in TCPA. This leads to a higher solubility in aqueous solution which can aid in aqueous phase reactions with peptides. Linkers can play a major role in

influencing the properties of bicyclic peptides. From previously reported literature [30], hydrophilic linkers can increase the aqueous solubility of the bicyclic peptides which aids in their purification and also crucial for biological activity characteristics. To understand the hydrophilicity of the linkers synthesized, calculated logP values were determined and shown in **Table** 4.1

Calculated logP values of linkers

Linker	c logP
TCPA	3.02 ± 0.27
BTBE	3.09 ± 0.85
TBCA	0.76 ± 0.55
NTCA	0.48 ± 0.67

Table 4.1: Calculated logP (clogP) values for the linkers.

Lower clogP values indicate the preference of the molecule to stay in the aqueous phase while higher values indicate their preference to stay in the hydrophobic organic phase (octane). These clogP values are estimations based on similar structures and might not depict the exact hydrophilicity/hydrophobicity of the linker but can be a good approximation to follow.

The linkers are hydrophilic (see **Table** 4.1) and the order of hydrophilicity of the linkers are: NTCA > TBCA > TCPA > BTBE.

Linkers TCPA and BTBE have similar clopP values around 3. While TCPA contains a hydrophilic central core (nitrogen atom) and hydrophobic side chains, BTBE has a hydrophobic central core (benzene ring) and hydrophilic side chains. It could then be inferred that either the central nitrogen core of TCPA, plays a major role in determining the hydrophilicity of the linker or the benzene core of BTBE imparts a significant hydrophobicity character to the linker that overcomes the hydrophilic side chain character.

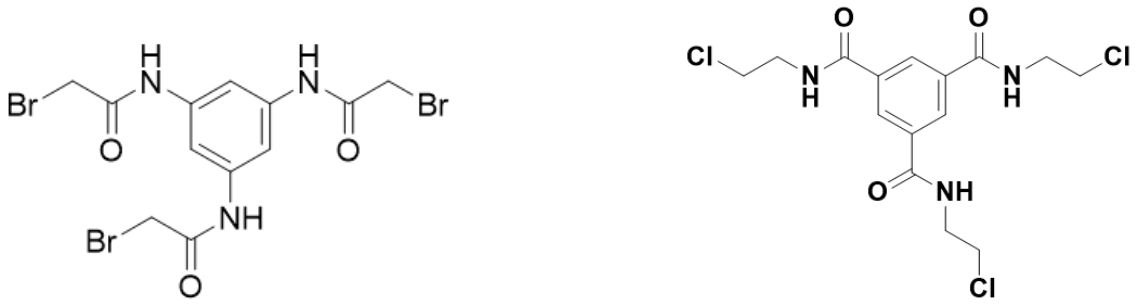
The amide functional group should impart a higher hydrophilicity to the linker as compared to a carbonyl functional group. This can be corroborated by comparing the clogP values of TBCA and BTBE, the former has ≈4 times lower value. Hence, TBCA is more hydrophilic than BTBE since, both BTBE and TBCA have the same benzene core, and differ only in their side chains.
NTCA and TBCA both have the same side chain substituents but differ in their central core, nitrogen atom and benzene ring respectively. From **Table** 4.1, NTCA has ≈ 1.6 times lower clogP value as compared to TBCA. Hence, the nitrogen core also plays a crucial role in affecting the hydrophilicity of the linker.

Overall, based on these clogP values, the linker NTCA is the most hydrophilic. Future designs of linkers can incorporate nitrogen in the core of the linker and/or have amide groups in the side chain to increase the hydrophilicity of the linker.

Polar molecules have been showed to be most suitable and superior as linkers for bicyclic peptides due to their ability to form non-covalent interactions with the peptide loops [5]. Hence, NTCA which has the highest number of potential H-donor and acceptor groups could form more non-covalent interactions with the peptide as compared to other linker molecules. Although, this needs to be tested with peptides to determine if they do form non-covalent interactions with loops of the bicyclic peptides.

4.1.2.5 Comparison between TBCA and TBAB

TBCA has a similar structure to TBAB (see **Figure** 4.4) which has been previously reported by Shiyu Chen *et al.* [7]. TBCA differs from TBAB on the basis of leaving group (Cl vs Br), number of carbon atoms in the side arms (3 C's vs 2 C's) and in the location of the carbonyl carbon of the amide bond with respect to the benzene ring. In TBAB the carbonyl carbon is attached to the benzene ring via a nitrogen by an amide bond, while in TBCA, it is attached directly to the benzene carbon. The effect of the leaving group and size of the side arm could make the reaction of TBCA and peptide slower as compared to TBAB.



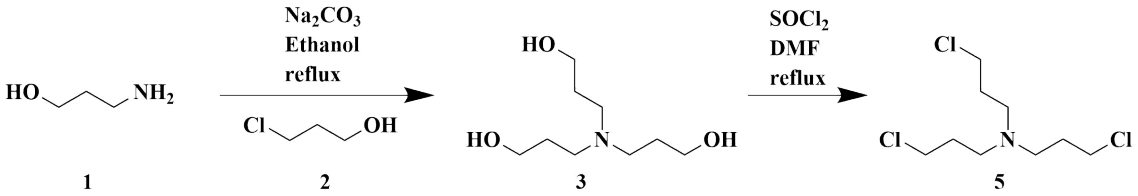
TBAB TBCA

Figure 4.4: Structure of TBAB [5][30] and TBCA **15**. TBCA has a similar structure to previously reported [5] [30] linker TBAB.

4.2 Synthesis of trivalent thiol reactive linkers

The general chemical synthesis, various challenges and workarounds for the four linkers: TCPA **5**, BTBE **9**, TBCA **16** and NTCA **21** are discussed below.
Linker TCPA has been previously reported in literature [4] recently but has not been used for cyclization of peptides. Other linkers have not been described in literature to my knowledge.

4.2.1 Synthesis of TCPA: tris(3-chloropropyl)amine



Scheme 4.1: Synthesis of TCPA **5**.

The synthesis of compound TCPA (**5**) is shown in **Scheme** 4.1. TCPA was synthesized in two steps via an intermediate tertiary amine **3**. **3** was synthesized via a nucleophilic substitution reaction between primary amine, **1** and alkyl chloride, **2**. Sodium carbonate (a mild base) neutralizes HCl that is formed and according to Le Chatelier's principle, the equilibrium is pushed to the right of the arrow and more product is formed. This reaction yielded a mixture of the desired product, **3** and the bis substituted intermediate, **4** (see **Figure** 4.5).

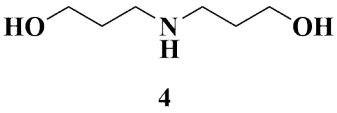
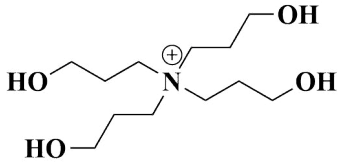


Figure 4.5: Bis intermediate, **4** was synthesized via a nucleophilic substitution reaction between **1** and **2**.

Compound **4** was the bis substituted intermediate where in 1 eq. of **2** had reacted with 1 eq. of **1**. These products as well as the starting material were readily soluble in water and an aqueous workup was not carried out.

The crude reaction mixture contained the desired product, **3**, the bis substituted product, **4** along with unreacted starting materials, **1** and **2**. Purification was carried out by distillation but required very low pressures (high vacuum) since compounds, **3** and **4** were very high boiling liquids. Distillation was carried out in a Koegler apparatus and the fractions obtained were **4** (175 °C/0.034 mbar) followed by **3** (200 °C/0.034 mbar). At higher temperatures, degradation products presumably via polymerization reactions were observed in proton NMR. Furthermore, the Koegler apparatus has a limitation during scale up (>2 g) due to the size of the apparatus. A normal vacuum distillation apparatus is even less efficient due to high reflux and increased time it takes for collection of distillate. This problem was overcome by performing the purification after the subsequent chlorination step with thionyl chloride.

Nucleophilic attack of the amine, **1** on the alkyl chloride, **2** can also theoretically result in **7** (see **Figure** 4.6), due to successive nucleophilic attacks until a quaternary molecule is formed. However, **7** which could form by the nucleophilic attack of **3** on another molecule of **2**, was not observed on NMR and LC-MS.

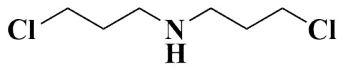


7

Figure 4.6: **7** is a byproduct which could theoretically be formed by the nucleophilic attack of **3** on **2**.

When only a slight excess of the alkyl chloride, **2** (0.3 eq. excess) was used, complete conversion from starting primary amine, **1** to tertiary amine, **3** was not achieved, instead the intermediate bis compound, **4** was favoured. Increasing the amount of the alkyl chloride, **2** (1 eq. excess) led to the major formation of the desired product, **3**.

Chlorination of the crude reaction mixture containing **3** with thionyl chloride and DMF as a catalyst resulted in the desired chlorinated product, **5** and by-products. A basic and acidic workup were performed to remove water soluble impurities. Crystallization in heptane as reported in literature failed [4]. Hence, purification was carried out by flash chromatography which resulted in two major fractions, **5** (the product, high Rf) and **6** (bis by-product, low Rf, see **Figure** 4.7) were isolated. The overall yield of the reaction was 80%.

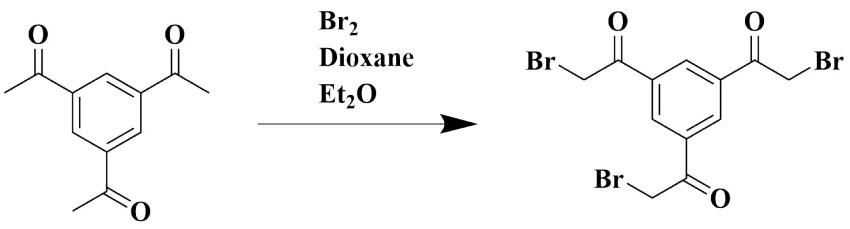


6

Figure 4.7: **6** is a by-product which forms upon chlorination of **4**.

4.2.2 Synthesis of BTBE: 1,1',1"-(benzene-1,3,5-triyl)tris (2-bromo-ethan-1-one)

4.2.2.1 Bromination using molecular bromine



8 9

Scheme 4.2: Synthesis of BTBE **9** using molecular bromine.

The synthesis of BTBE (9) is shown in **Scheme** 4.2 using molecular bromine as the brominating agent. A catalytic amount of 1,4-dioxane was used to form a complex, dioxane-dibromide *in situ* which is a mild brominating agent. The product, **9** is formed by electrophilic addition of bromine on the α-carbon of the three side chains of the starting material, **8**.

The reaction did not go to 100% completion and multiple spots were observed on a TLC plate, corresponding to unreacted starting material, **8** and other by-products. LC-MS and NMR of the sample showed that the reaction mixture consisted mostly of the product, **9** along with mono- and bis-brominated intermediates, **10** and **11**, respectively (see **Figure** 4.8).

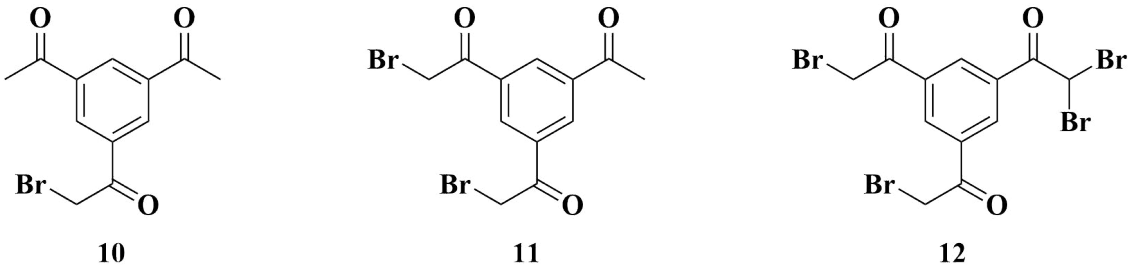
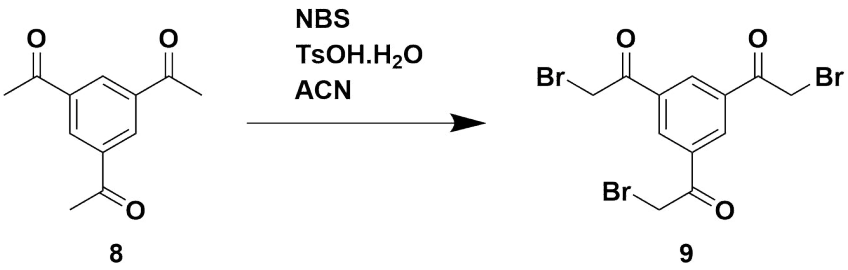


Figure 4.8: Intermediates **10**, **11** and side-product **12** from the bromination of **8**.

Other spots on the TLC plate were not identified. These compounds could be formed due to the high reactivity of bromine which leads to highly exothermic non-selective reactions. **12** (see **Figure** 4.8) could be a side product wherein the α-carbon gets brominated twice. The Rf's of **9**, **10** and **11** are also very similar and purification by a flash column resulted in co-elution of the compounds and pure product, **9** could not be isolated. The reaction could be enhanced by addition of an acid which promotes keto-enol tautomerization of **9** and aids the reaction since the enol is a better nucleophile.

However, due to the unfavorable properties of bromine – toxicity, corrosiveness, irritation - alternative brominating agents were considered. *N*-bromosuccinimide (NBS) was an attractive option since it is a white crystalline solid (ease of handling), a mild brominating agent which produces low constant concentrations of bromine during the course of the reaction and prevents the formation of exothermic side-reactions. Furthermore, since the bromine is attached to an electron withdrawing nitrogen in the succinimide, the bromine has a partial positive charge and hence is electrophilic. [15]

4.2.2.2 Bromination using *N*-bromosuccinimide (NBS)



Scheme 4.3: Synthesis of BTBE **9** using *N*-bromosuccinimide (NBS).

The synthesis of BTBE, **9** using NBS as the brominating agent and mediated by *p*-toluenesulfonic acid is shown in **Scheme** 4.3. The three α-carbon (methyl) of **8** are brominated using NBS and this follows an ionic pathway instead of a radical one.

Stoichiometric amounts of the reactants were used and refluxed using acetonitrile as solvent. TLC of the reaction mixture after 2 h showed the presence of 3 spots spaced close to each other and no starting material was observed. Complete conversion to product, **9** was not observed even after long reaction time (>21 h) and a crude NMR of the reaction mixture indicated the presence of product, **9** along with intermediate products, **10** and **11** (see **Figure** 4.8). Addition of excess NBS, converted all the mono-substituted intermediate, **10** to the bis-substituted intermediate, **11**. On TLC this also resulted in the appearance of an additional spot which could not be identified. Thus, large excess of NBS combined with longer reaction times lead to the formation of by-products.

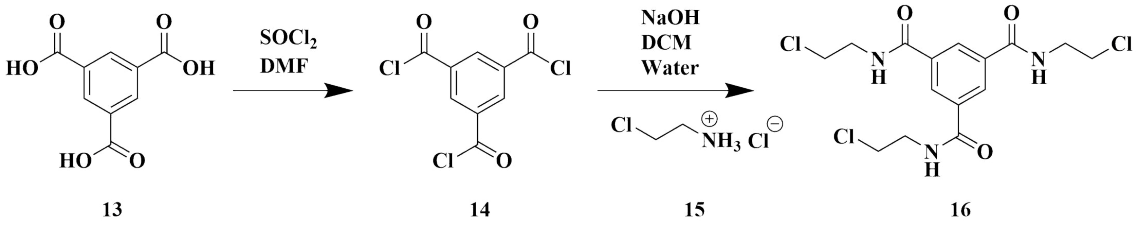
Therefore, addition of stoichiometric amounts of NBS along with shorter reaction time (2 h) could prevent the formation of by-products. Microwave assisted α-carbon bromination was reported in the literature to have better yield (>90%) and shorter reaction times (<30 min) [15]. This could be a better alternative to refluxing the reaction mixture, as well as increase the conversion and reduce the reaction time.

Purification was carried out on a silica column using a heptane/EtOAc gradient. The crude product mixture was not very soluble in heptane or EtOAc and needed to be loaded onto the column using DCM which affected the separation of the different compounds. Additionally, the product, **9** co-eluted with the intermediates (**10** and **11**). The by-products and the product have similar Rf's, which resulted in co-elution of the product leading to a reduced yield of 12%.

The first issue of using DCM to load the sample can be solved by dry loading the sample instead of wet loading. Whereas, the second issue of co-elution, can be mitigated by having a very gradual, less steep gradient. Another strategy could be to try other solvent mixtures for the mobile phase, that help widen the gap between

Rf's of the different compounds.

4.2.3 Synthesis of TBCA: N^1,N^3,N^5-tris(2-chloro-ethyl) benzene-1,3,5-tricarboxamide

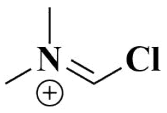


Scheme 4.4: Synthesis of TBCA **16**.

The synthesis of TBCA, **16** is shown in **Scheme** 4.4. TBCA was synthesized in two steps via an intermediate acyl chloride, **14**.

4.2.3.1 Chlorination of trimesic acid, 13

The acyl chloride, **14** was synthesized from trimesic acid, **13** using thionyl chloride and DMF as the catalyst. DMF initially reacts with the thionyl chloride to form an electrophilic chloro-iminium intermediate, **17** (Vilsmeier-Haack reagent) as shown in **Figure** 4.9. The iminium intermediate then reacts with the carboxylic acid of **13** to form the corresponding acid chloride, **14** and the catalyst, DMF gets regenerated in the process.

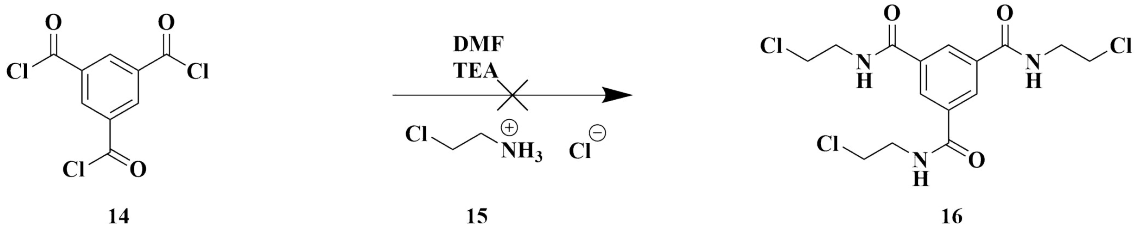


17

Figure 4.9: Chloro-iminium intermediate formed by the reaction between DMF and thionyl chloride, which acts as the electrophile for the chlorination of carboxylic acid.

No purification steps were required and this reaction resulted in **14** in a good yield of 92%. Small quantities of **14** started to hydrolyze back to the carboxylic acid, **13** even when stored under vacuum over 24 h at room temperature. Hence, it is crucial that the acid chloride is utilized for the next step without storing or stored at low temperatures to prevent hydrolysis.

4.2.3.2 Coupling reaction between 14 and 15



Scheme 4.5: Amide coupling reaction between **14** and **15**.

Amide coupling reaction between the acid chloride, **14** and the amine hydrochloride salt, **15** (see **Scheme** 4.5) using TEA as base and DMF as the solvent did not result in a clean reaction. Multiple peaks were observed in proton NMR and the various peaks could not be resolved. Furthermore, the amine hydrochloride salt, **15** was sparingly soluble in the solvent, DMF. This reduced the reaction rate, increased reaction time and did not lead to complete conversion to **16**. Other alternative reactions were tested which gave a cleaner reaction and involved less purification steps.

Amide coupling under Schotten-Baumaan conditions

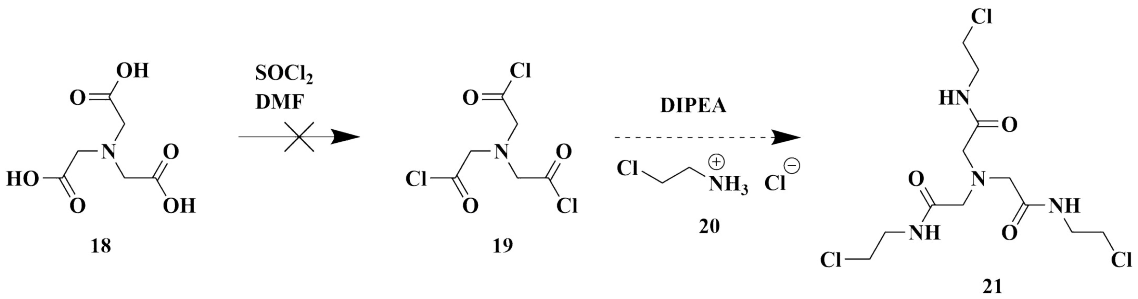
The desired product, **16** was formed by the amide coupling reaction between the acyl chloride, **14** and amine hydrochloride salt, **15** under Schotten-Baumaan (biphasic aqueous basic reaction) reaction conditions (see **Scheme** 4.4).

The product, **16** was sparingly soluble in the organic phase (DCM) and not soluble in the aqueous phase. Therefore purification of the crude product was carried out by re-precipitation to obtain the pure product with a yield of 60%. The overall yield for the two-step reaction was 55%.

4.2.4 Synthesis of NTCA: 2,2',2"-nitrilotris(N-(2-chloroethyl) acetamide)

4.2.4.1 Synthesis of NTCA, 21 in two steps via an intermediate acyl chloride

NTCA was initially envisioned to be synthesized from nitrilotriacetic acid, **18** in two steps via an acyl chloride, **19** as shown in **Scheme** 4.6. **19** was further intended to be coupled with free amine of **20** to form the amide **21** (NTCA). The problems faced with this reaction scheme are discussed below.

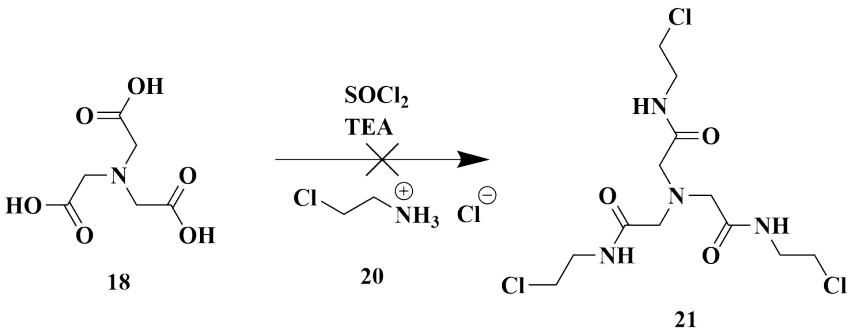


Scheme 4.6: Two-step reaction pathway initially envisioned to form NTCA, **21**.

Nitrilotriacetic acid, **18** was reacted with thionyl chloride under reflux with catalytic amount of DMF to increase the conversion. However, proton NMR of the crude product after workup showed multiple peaks (between 3 – 6 ppm). These could be due to degradation of product, elimination or side-reactions. These peaks could not be identified and the reaction had failed. Hence, other alternative synthesis pathways were considered.

A one-pot synthesis where the acid, **18** and the amine, **20** were coupled using thionyl chloride [20] as the activating agent in the presence of TEA is shown in **Scheme** 4.7.

4.2.4.2 One pot synthesis pathway to form NTCA

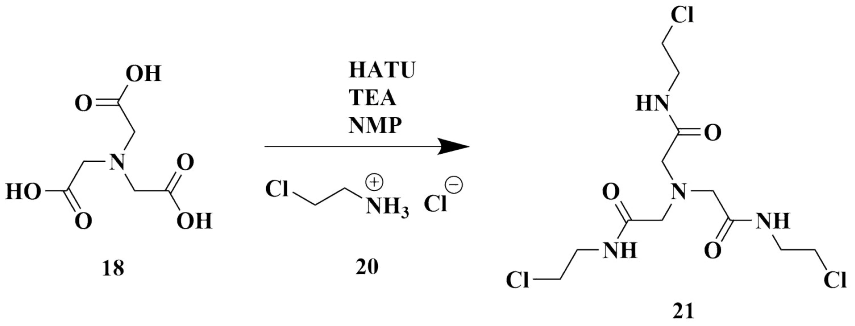


Scheme 4.7: One-pot synthesis pathway to form NTCA, **21**.

Thionyl chloride chlorinates the acid, **18** *in situ*, thus activating it for reaction with the free based amine of **20**. Proton NMR of the crude after workup, was found to contain a mixture of TEA, **20** and small quantities of the product, **21**. A black residue was formed during the course of the reaction which could be due to degradation or polymerization reactions. Due to the low amounts of the product formed other alternative coupling agents which result in a cleaner reaction were researched.

4.2.4.3 One-pot synthesis of NTCA using HATU as the coupling agent

A different amide coupling reagent, HATU was used in a one-pot synthesis by reacting the acid, **18** and the amine hydrochloride salt, **20** in the presence of base TEA, and NMP as the solvent is shown in **Scheme** 4.8.



Scheme 4.8: One-pot synthesis of NTCA, **21** using HATU as the coupling agent.

TEA was used to free base the amine hydrochloride salt, **20** and NMP was used as the solvent to enhance the solubility of **18** and the free amine.
The reactants were used in stoichiometric amounts and complete conversion of to **21** was not achieved. This was supported by a LC-MS analysis of the crude product after 20 h which showed the presence of the product, **21** along with a large fraction of bis-intermediate, **22** (see **Figure** 4.10).

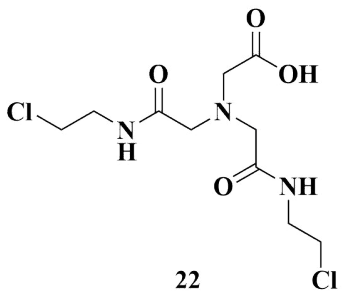


Figure 4.10: Bis-intermediate **22** formed during the HATU coupling reaction of **18** and **20**.

The yield for the reaction was low, 3%. To increase the yield of the reaction and conversion of intermediate, **22** to the product, **21** excess starting material, **20** and larger equivalents of the coupling reagent, HATU were used.

This reaction was performed again by adding 1 eq. excess of both **20** and HATU and a LC-MS analysis was carried out at different time intervals to check the progress of the reaction. At 3 h, the product **21** was observed along with a higher fraction of bis-intermediate, **22**. At 6 h, the ratio of **21** to **22** remained the same. Hence, an additional 2 equivalents of **20** was added but a LC-MS analysis after 18 h showed the presence of the di-amide intermediate, **22** (see **Figure** 4.10). There was no major change in conversion of **22** to **21** due to addition of **20**. However, addition

of two equivalents of HATU showed a reduction in the amount of intermediate, **22** and increase in the amount of product, **21**. The yield (31%) had also significantly improved compared to when using stoichiometric amount of reactants. Therefore, the addition of excess amounts of **20** and HATU led to a better conversion.

Another issue was the solubility of nitrilotriacetic acid, **18** and free based amine of **20**. **18** was not completely soluble in the solvent, NMP but the product, **21** was soluble. As the reaction progressed and as more **18** got converted to **20**, the undissolved **18** started to go into solution. Similarly, **20** was soluble in NMP but after free basing it with TEA, the amine had reduced solubility. Hence, the solubility of **18** and free based amine of **20** in NMP could have affected the rate and time of the reaction. No alternative solvents were tested but other polar solvents like THF, DMSO or ACN could be good alternatives.

Due to the low vapor pressure of NMP, it could not be removed by concentrating *in vacuo*. Although the majority of the solvent was removed during the aqueous workup, there was still some residual solvent present in the crude product. The sample also contained some triethyl amine (TEA) which along with the residual solvent altered the strength of the mobile phase and affected the elution of the compounds during column chromatography. This issue was dealt with by initially flushing the column with a low strength mobile phase to elute out NMP and TEA, after which a stronger mobile phase was used to elute the pure product, **21**.

5
Conclusion

In summary, four small molecule linkers with different substituents and chemical environment were successfully synthesized with low to good yield. Literature review resulted in a good understanding of the various linkers currently used for cyclization of peptides, their design and how they affect the structure and properties of the peptides. The linkers were designed to be C_3 symmetric and contained three side chains attached to a central core (atom/molecule). The side chains were either hydrophobic or contained multiple H-bond donor and acceptor groups. Similarly, the central core was either hydrophobic or polar. The side chains were terminated by a halide, which would act as the leaving group upon substitution reaction between the linker and free thiol groups of peptides.

The central core and substituents on the side chains contribute to the hydrophilicity of the linker. The four synthesized linkers were hydrophilic and NTCA was found to be the most polar. In general, polar linkers contribute to the solubility of the bicyclic peptides in aqueous solution which facilitates purification and affects the peptide's biological activity. Furthermore, polar linkers confer superior stablity to the structure of the bicyclic peptide compared to hydrophobic linkers due to their ability to form non-covalent interactions with the amino acids of the peptide. [30] This leads to a higher structural constraint in the bicyclic peptide thus reducing the entropy loss upon binding to targets and increasing the binding affinity. [5][30]

5.1 Future Outlook

The compounds were synthesized to see if they are viable candidates as linkers for bicyclization of peptides and the yield was not the main priority. However, these reactions can be optimized to achieve higher conversion and better yields.

The synthesized compounds have not been reported for use in bicyclization of peptides. However, they are promising candidates as linkers and have potential H-bond donor and acceptor groups. These linkers can therefore form H-bonds with the peptide loops adding additional constraint to bicyclic peptide which is favourable. Hence, these linkers need to be tested with a model peptide to see if they can bicyclize the peptide, better understand what properties they impart to the bicyclic peptide and see if they preferentially react only with thiol groups.

Bibliography

[1] Shahrzad Ahangarzadeh, Mohammad M. Kanafi, Simzar Hosseinzadeh, Ahad Mokhtarzadeh, Mahmood Barati, Javad Ranjbari, and Lobat Tayebi. Bicyclic peptides: types, synthesis and applications. *Drug Discovery Today*, 24(6):1311–1319, 2019.

[2] Alessandro Angelini, Laura Cendron, Shiyu Chen, Jeremy Touati, Greg Winter, Giuseppe Zanotti, and Christian Heinis. Bicyclic peptide inhibitor reveals large contact interface with a protease target. *ACS Chemical Biology*, 7(5):817–821, 2012.

[3] Neelam Azad and Yon Rojanasakul. Macromolecular drug delivery. *Biopharmaceutical Drug Design and Development*, pages 293–323, 2008.

[4] Pritha Bagchi, M. Thomas Morgan, John Bacsa, and Christoph J. Fahrni. Robust affinity standards for Cu(I) biochemistry. *Journal of the American Chemical Society*, 135(49):18549–18559, 12 2013.

[5] Shiyu Chen, Davide Bertoldo, Alessandro Angelini, Florence Pojer, and Christian Heinis. Peptide ligands stabilized by small molecules. *Angewandte Chemie - International Edition*, 53(6):1602–1606, 2014.

[6] Shiyu Chen and Christian Heinis. CHAPTER 9. Phage Selection of Mono- and Bicyclic Peptide Ligands. In Andrew J McKnigh Lyn Jones, editor, *Biotherapeutics: Recent Developments using Chemical and Molecular Biology*, number 36, pages 241–262. Royal Society of Chemistry, 2013.

[7] Shiyu Chen, Julia Morales-Sanfrutos, Alessandro Angelini, Brian Cutting, and Christian Heinis. Structurally Diverse Cyclisation Linkers Impose Different Backbone Conformations in Bicyclic Peptides. *ChemBioChem*, 13(7):1032–1038, 2012.

[8] Moo J. Cho and Rudy Juliano. Macromolecular versus small-molecule therapeutics: Drug discovery, development and clinical considerations. *Trends in Biotechnology*, 14(5):153–158, 1996.

[9] Maaike de Loos, Jan H. van Esch, Richard M Kellogg, and Ben L Feringa. C3-Symmetric, amino acid based organogelators and thickeners: a systematic study of structure-property relations. *Tetrahedron*, 63(31):7285–7301, 2007.

[10] Philippe Diderich, Davide Bertoldo, Pierre Dessen, Maola M. Khan, Irene Pizzitola, Werner Held, Joerg Huelsken, and Christian Heinis. Phage Selection of Chemically Stabilized α-Helical Peptide Ligands. *ACS Chemical Biology*, 11(5):1422–1427, 2016.

[11] Martin Empting. Peptides. In Wilfred A van der Donk Jesko Koehnke, James Naismith, editor, *Cyclic Peptides: From Bioorganic Synthesis to Applications*, number 6, pages 1–14. Royal Society of Chemistry, 2018.

[12] Christoph Ernst, Johannes Heidrich, Catharina Sessler, Julia Sindlinger, Dirk Schwarzer, Pierre Koch, and Frank M. Boeckler. Switching between bicyclic and linear peptides - the sulfhydryl-specific linker TPSMB enables reversible cyclization of peptides. *Frontiers in Chemistry*, 6(OCT):1–8, 2018.

[13] Christoph Ernst, Julia Sindlinger, Dirk Schwarzer, Pierre Koch, and Frank M. Boeckler. The Symmetric Tetravalent Sulfhydryl-Specific Linker NATBA Facilitates a Combinatorial "tool Kit" Strategy for Phage Display-Based Selection of Functionalized Bicyclic Peptides. *ACS Omega*, 3(10):12361–12368, 2018.

[14] Thaddeus S Franczyk, Kenneth R Czerwinski, and Kenneth N Raymond. Stereognostic Coordination Chemistry. 1. The Design and Synthesis of Chelators for the Uranyl Ion. *Journal of the American Chemical Society*, 114(21):8138–8146, 1992.

[15] Xiao-yu Guan. Efficient and selective a -bromination of carbonyl compounds with N -bromosuccinimide under microwave. *Arabian Journal of Chemistry*, 8(6):892–896, 2015.

[16] Marcel Günther, Jonas Lategahn, Michael Juchum, Eva Döring, Marina Keul, Julian Engel, Hannah L Tumbrink, Daniel Rauh, and Stefan Laufer. Trisubstituted Pyridinylimidazoles as Potent Inhibitors of the Clinically Resistant L858R/T790M/C797S EGFR Mutant: Targeting of Both Hydrophobic Regions and the Phosphate Binding Site. *Journal of Medicinal Chemistry*, 60(13):5613–5637, 2017.

[17] Christian Heinis, Trevor Rutherford, Stephan Freund, and Greg Winter. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nature Chemical Biology*, 5(7):502–507, 2009.

[18] Sang Hoon Joo. Cyclic peptides as therapeutic agents and biochemical tools, 1 2012.

[19] Yu Heng Lau, Peterson De Andrade, Yuteng Wu, and David R. Spring. Peptide stapling techniques based on different macrocyclisation chemistries. *Chemical Society Reviews*, 44(1):91–102, 2015.

[20] A. Leggio, E. L. Belsito, G. De Luca, M. L. Di Gioia, V. Leotta, E. Romio, C. Siciliano, and A. Liguori. One-pot synthesis of amides from carboxylic acids activated using thionyl chloride. *RSC Advances*, 6(41):34468–34475, 4 2016.

[21] Marcello Leopoldo, Francesco Berardi, Nicola A. Colabufo, Paola De Giorgio, Enza Lacivita, Roberto Perrone, and Vincenzo Tortorella. Structure-affinity relationship study on N-[4-(4-arylpiperazin-1-yl)butyl]arylcarboxamides as potent and selective dopamine D3 receptor ligands. *Journal of Medicinal Chemistry*, 45(26):5727–5735, 12 2002.

[22] Meiqing Li, Ximing Shao, Chunlei Wu, Danyi Lu, Ke Liu, Wei Wang, Jiakai Liu, Hongchang Li, Wu Su, and Lijing Fang. Chlorotoxin-derived bicyclic peptides for targeted imaging of glioblastomas. *Chemical Communications*, 56(66):9537–9540, 2020.

[23] Wenlong Lian, Punit Upadhyaya, Curran A. Rhodes, Yusen Liu, and Dehua Pei. Screening bicyclic peptide libraries for protein-protein interaction inhibitors: Discovery of a tumor necrosis factor-α Antagonist. *Journal of the American Chemical Society*, 135(32):11990–11995, 2013.

[24] Minami Nakajima, Yukiko Oda, Takamasa Wada, Ryo Minamikawa, Kenji Shirokane, Takaaki Sato, and Noritaka Chida. Chemoselective reductive nucleophilic addition to tertiary amides, secondary amides, and N-methoxyamides. *Chemistry - A European Journal*, 20(52):17565–17571, 12 2014.

[25] Leonard T. Nguyen, Johnny K. Chau, Nicole A. Perry, Leonie de Boer, Sebastian A.J. Zaat, and Hans J. Vogel. Serum stabilities of short tryptophan- and arginine-rich antimicrobial peptide analogs. *PLoS ONE*, 5(9):1–8, 2010.

[26] Fabien Renaud, Carole Decurnex, and Gérard Hopfgartner. Flexible covalent tripods in nonadentate podands: synthesis of tris [3-(6-diethylcarbamoylpyridine-2-carboxamido) propyl] amine and its complexing properties with trivalent lanthanides. *Journal of the Chemical Society, Dalton Transactions*, (12):1863–1871, 2001.

[27] Curran A. Rhodes and Dehua Pei. Bicyclic Peptides as Next-Generation Therapeutics. *Chemistry - A European Journal*, 23(52):12690–12703, 2017.

[28] Chinwon Rim, Lauren J. Lahey, Vijita G. Patel, Hongming Zhang, and David Y. Son. Thiol-ene reactions of 1,3,5-triacryloylhexahydro-1,3,5-triazine (TAT): facile access to functional tripodal thioethers. *Tetrahedron Letters*, 50(7):745–747, 2009.

[29] Peter Timmerman, Joris Beld, Wouter C. Puijk, and Rob H. Meloen. Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *ChemBioChem*, 6(5):821–824, 2005.

[30] Helmus van de Langemheen, Valerijs Korotkovs, Joachim Bijl, Claire Wilson, Sangram S. Kale, Christian Heinis, and Rob M.J. Liskamp. Polar Hinges as Functionalized Conformational Constraints in (Bi)cyclic Peptides. *ChemBioChem*, 18(4):387–395, 2017.

[31] Hong Bo Wang, James A Wisner, and Michael C Jennings. Anion receptors containing thiazine-1,1-dioxide heterocycles as hydrogen bond donors. *Beilstein Journal of Organic Chemistry*, 6(50):1–6, 2010.

A

Appendix A - Experimental section

All chemicals were purchased from commercial suppliers and used without further purification unless stated otherwise.

Solvents were evaporated using BÜCHI Vacuum Rotavapor R-200 with BÜCHI Vacuum controller V-850 as vacuum pump controller. Reactions were monitored using Thin Layer Chromatography (TLC) analysis using Merck TLC Silica gel 60 F254, 2.5x7.5 cm glass plates under UV-light (254 nm) or staining with potassium permanganate stain for visualization. Column chromatography was performed using Isolera™ Spektra One with either Biotage SNAP®-10 g, 25 g or 40 g KP-sil columns.

^{1}H and ^{13}C NMR were measured in $CDCl_3$ or DMSO-d6 and recorded on Bruker Avance-400, Bruker Avance-500 or Oxford NMR AS500 spectrometers at 25°C. The chemical shifts are reported as parts per million (ppm) and the coupling constants of the multiplicities are reported in Hertz (Hz), with the residual solvent signal used as a reference. NMR abbreviations used are as follows: br = broad, s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, m = multiplet. NMR peaks were assigned using MestReNova 14.1.1.

LCMS analysis was performed on Waters Acquity Ultra Performance LC using BEH C18 column (50 mm x 2.1 mm, 1.7 μm particles). GCMS (EI) analysis was performed on Agilent Technologies 7890A with 5975C inert MSD mass system equipped with an 19091S-433L (30 m x 250 μm x 0.25 μm) capillary column and 7683A series injector.

All structures and schemes were drawn in ChemDraw v.16.0.1.4.

A.1 Synthesis of linkers

A.1.1 Procedure for the synthesis of TCPA: tris(3-chloro propyl)amine, 5

The procedure was performed according to previously reported procedures but the purification strategy was modified. [14][4][26]

A mixture of 3-aminopropan-1-ol, **1** (2.00 g, 26.63 mmol, 1 eq.), 3-chloropropan-1-ol, **2** (7.55 g , 79.88 mmol, 3 eq.) and sodium carbonate (5.64 g, 53.25 mmol, 2 eq.) were dissolved in ethanol (15 mL) and refluxed for 23 h. The reaction was cooled to room temperature, dichloromethane (80 mL) was added, filtered to remove salts and concentrated *in vacuo* and dried under vacuum to obtain a clear, light yellow viscous oil. The resulting yellowish viscous oil was taken up in acetonitrile (15 mL) and a solution of thionyl chloride (31.7 g, 266.30 mmol, 10 eq.) in DCM (10 mL) was added slowly. Catalytic amount of DMF (0.047 g, 0.65 mmol, 0.024 eq.) was added to the rapidly stirred mixture and refluxed for 3 h. The solution was allowed to cool to room temperature, methanol was added to quench the reaction, concentrated *in vacuo* and added to a 100 mL solution of 1:1 mixture of EtOAc and heptane. This was then washed with 3.8 M NaOH solution (50 mL) and the organic layer was extracted using 1.9 M HCl solution (3 × 30 mL). The resulting aqueous layers were combined, neutralized with 3.8 M NaOH and adjusted to pH 12. The oily layer formed was extracted with heptane (5 × 40 mL), dried over anhydrous Na_2SO_4, filtered and concentrated *in vacuo*. Flash chromatography through a silica column using a gradient of 0 - 20% EtOAc in heptane (see **Figure** B.15 in Appendix B) resulted in a clear, colourless oil, **5** (5.27 g, 80%). The product, **5** that was obtained was not a colourless crystalline solid as reported in literature [4] but was instead a colourless liquid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 3.62 (t, J = 6.2 Hz, 6H), 2.54 (t, J = 6.5 Hz, 6H), 1.90 (p, J = 6.4 Hz, 6H).

^{13}C NMR (126 MHz, $CDCl_3$) δ 50.59, 43.10, 30.17.

A.1.2 Procedure for the synthesis of BTBE: 1,1',1"-(benzene-1,3,5-triyl)tris(2-bromoethan-1-one),9

A.1.2.1 Using molecular bromine as the brominating agent

The procedure followed was according to an in-house AstraZeneca procedure.

1,1',1"-(Benzene-1,3,5-triyl)tris(ethan-1-one), **8** (500 mg, 2.45 mmol, 1 eq.) and 1,4-dioxane (10.47 µL, 0.12 mmol, 0.049 eq.) were dissolved in diethyl ether (10 mL) to form a light brown solution. Bromine (59 mg, 0.367 mmol, 5% of the total moles of Br_2) was added to the solution at rt to form a light brownish yellow solution. When the colour had disappeared (approx. 30 min), the remaining amount of bromine

(1113.5 mg, 6.97 mmol, 95% of the total moles of Br_2) was added dropwise and the solution was stirred at room temperature for 4 h. An additional amount of bromine (352 mg, 2.21 mmol) was added and the solution stirred for 1 h. The reaction was quenched with 100 mL of sodium thiosulfate solution (10% (w/w)) and the aqueous layer was extracted with Et_2O (3 × 50 mL). The organic layers were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography through a silica column using a gradient of 0 - 30% EtOAc in heptane as the mobile phase (see **Figure** B.16 in Appendix B). The product co-eluted with by-products and no pure product was isolated.

A.1.2.2 Using NBS as the brominating agent

The procedure was performed according to previously reported procedures for acetophenone bromination. [16][31]

NBS (1307 mg, 7.34 mmol, 3 eq) was added to a solution of 1,1',1"-(benzene-1,3,5-triyl)tris(ethan-1-one), **8** (500 mg, 2.45 mmol, 1 eq.) and 4-methylbenzenesulfonic acid hydrate (2096 mg, 11.02 mmol, 4.5 eq.) in acetonitrile (20 mL). The solution was refluxed and the reaction followed by LC-MS and TLC. After 21 h, additional NBS (435.5 mg, 2.45 mmol, 1 eq.) was added to the stirred solution and refluxed. After 3 h, additional NBS (435.5 mg, 2.45 mmol, 1 eq.) was added to the stirred solution and refluxed for 2 h. The filtrate was concentrated *in vacuo* to remove the solvent. DCM (40 mL) was added, washed with DI water (3 × 20 mL) and the organic phase was concentrated *in vacuo* to give the crude product as a yellow oil. The crude product was purified by flash chromatography through a silica column using a gradient of 0 - 30% EtOAc in heptane as the mobile phase (see **Figure** B.16 in Appendix B) which resulted in a white fluffy solid, **9** (133 mg, 12.32%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.79 (s, 3H), 4.51 (s, 6H).

^{13}C NMR (126 MHz, $CDCl_3$) δ 189.56, 135.24, 133.48, 29.84.

A.1.3 Procedure for the synthesis of TBCA, 16

A.1.3.1 Step 1: Synthesis of intermediate benzene-1,3,5-tri-carbonyl trichloride, 14

The procedure was performed according to previously reported procedure. [9]

A dry round botton flask was charged with benzene-1,3,5-tricarboxylic acid, **13** (1 g, 4.76 mmol) and thionyl chloride (8.15 g, 68.51 mmol, 14 eq.) was added dropwise. A drop of *N,N*-dimethylformamide (0.037 ml, 0.48 mmol) was added and the reaction was refluxed for 3 h. Thionyl chloride was removed *in vacuo* and dried under vacuum to give the product, **14** (1.161 g, 92%), as a yellow oil.

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.08 (s, 3H).

^{13}C NMR (126 MHz, $CDCl_3$) δ 166.13, 138.16, 135.60.

A.1.3.2 Step 2: Synthesis of TBCA: *N¹,N³,N⁵*-tris(2-chloro-ethyl)benzene-1,3,5-tricarboxamide, 16

The procedure was performed according to previously reported procedure. [24] [21]

A solution of benzene-1,3,5-tricarbonyl trichloride, **14** (743.9 mg, 2.80 mmol, 1 eq.) in dichloromethane (12 mL) was added dropwise to a stirred solution of 2-chloroethan-1-aminium chloride, **15** (1300 mg, 11.21 mmol, 4 eq.) in 2.3% NaOH and stirred vigorously for 20 h. The pH of the solution was adjusted to 2 using 3.8 M HCl solution to protonate the excess amine, **15**. The white solid precipitate containing the product **16** was filtered and the reaction mixture was extracted with ethyl acetate (1 × 40 mL). The organic layer was concentrated *in vacuo* to yield a white solid which was combined with the precipitate initially filtered and re-precipitated in a mixture of ethyl acetate and DCM to yield a pure white solid, **16** (668 mg, 60%).

^{1}H NMR (500 MHz, DMSO) δ 9.00 (t, 3H), 8.46 (s, 3H), 3.77 (t, 6H), 3.63 (q, 6H).

^{13}C NMR (126 MHz, DMSO) δ 166.10, 135.16, 129.21, 43.50, 41.95.

A.1.4 Procedure for the synthesis of NTCA: 2,2',2"-nitrilotris(N-(2-chloro-ethyl)acetamide), 21

Solution A was prepared by adding HATU (3979 mg, 10.46 mmol, 4 eq.) to a stirred solution of 2,2',2"-nitrilotriacetic acid, **18** (500 mg, 2.62 mmol, 1 eq.) and triethylamine (1059 mg, 10.46 mmol, 4 eq.) in NMP (10 mL). Solution B was prepared by adding triethylamine (1059 mg, 10.46 mmol, 4 eq.) to a stirred solution containing 2-chloroethan-1-aminium chloride, **20** (1214 mg, 10.46 mmol, 4 eq.) in NMP (10 mL). Solution A was added dropwise to solution B over 5 minutes and stirred at room temperature. The reaction was monitored by LC-MS and after 6 h, an additional amount of 2-chloroethan-1-aminium chloride, **20** (607 mg, 5.23 mmol, 3 eq.) along with triethylamine (794 mg, 10.46 mmol, 3 eq.) in NMP (10 mL) was added and stirred for 18 h. HATU (1989 mg, 5.23 mmol, 2 eq.) was added to the reaction mixture and stirred for another 2 h. The reaction mixture was cooled, taken up in ethyl acetate (50 mL) and washed with sodium carbonate solution (3 × 30 mL). The combined aqueous layer was extracted with ethyl acetate (1 × 20 mL). The organic layers were combined, dried over anhydrous sodium sulfate, filtered and concentrated *in vacuo*. Flash chromatography through a silica column using a gradient of 0 - 20% methanol in EtOAc (see **Figure** B.17 in Appendix B) resulted in the product, **21** (374 mg, 31%) which was found to be a white solid.

^{1}H NMR (500 MHz, DMSO) δ 8.35 (t, J = 5.9 Hz, 3H), 3.63 (t, J = 6.2 Hz, 6H), 3.42 (q, J = 6.1 Hz, 6H), 3.25 (s, 6H).

B

Appendix B - Analytical data

B.1 NMR graphs

Herein are presented the NMR spectra along with peak assignment for the linkers synthesized. Also, included are the NMR spectra for the intermediate compounds and by-products (if any) formed.

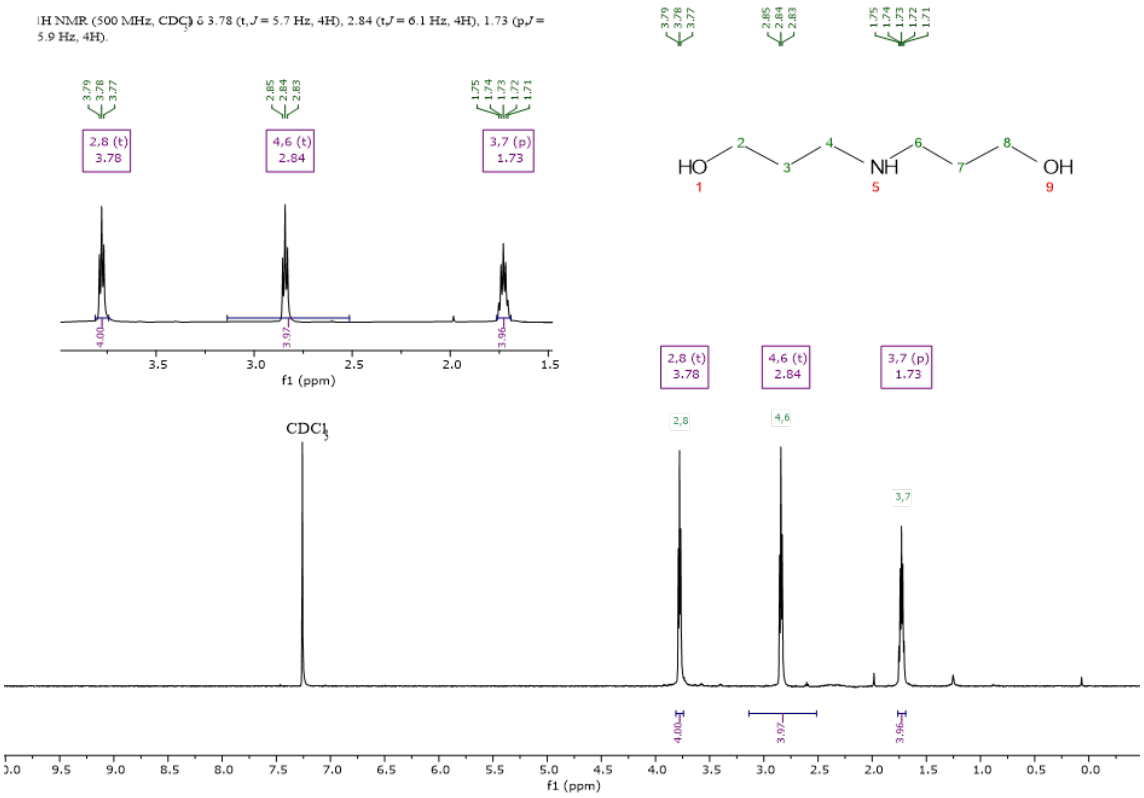


Figure B.1: ^{1}H NMR spectrum of 3,3'-azanediylbis(propan-1-ol), **4**.

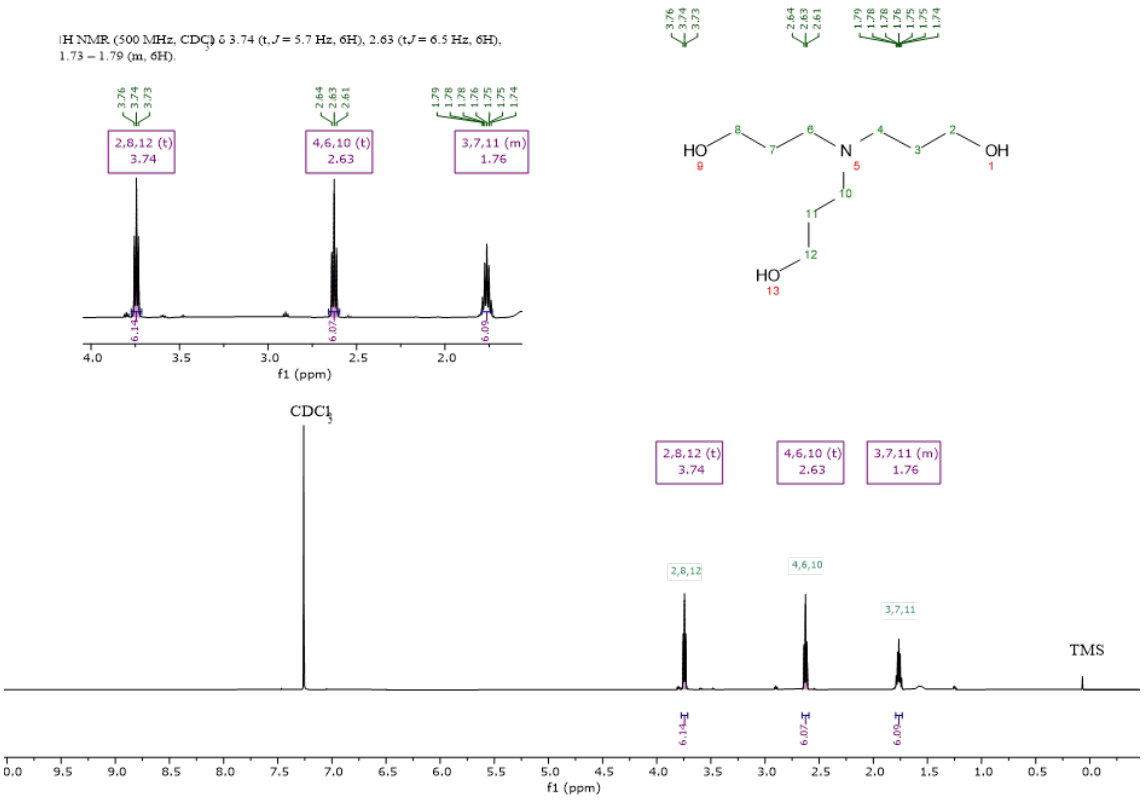


Figure B.2: ^{1}H NMR spectrum of 3,3',3"-nitrilotris(propan-1-ol), **3**.

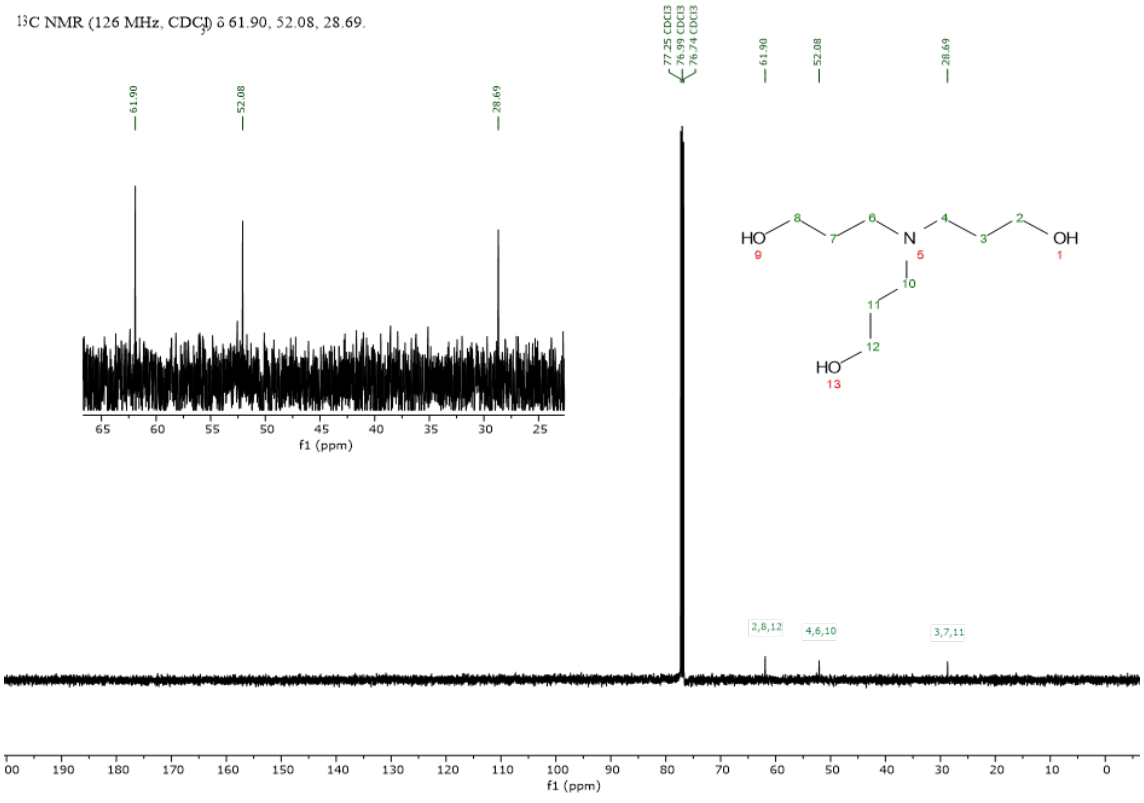


Figure B.3: ^{13}C NMR spectrum of 3,3',3"-nitrilotris(propan-1-ol), **3**.

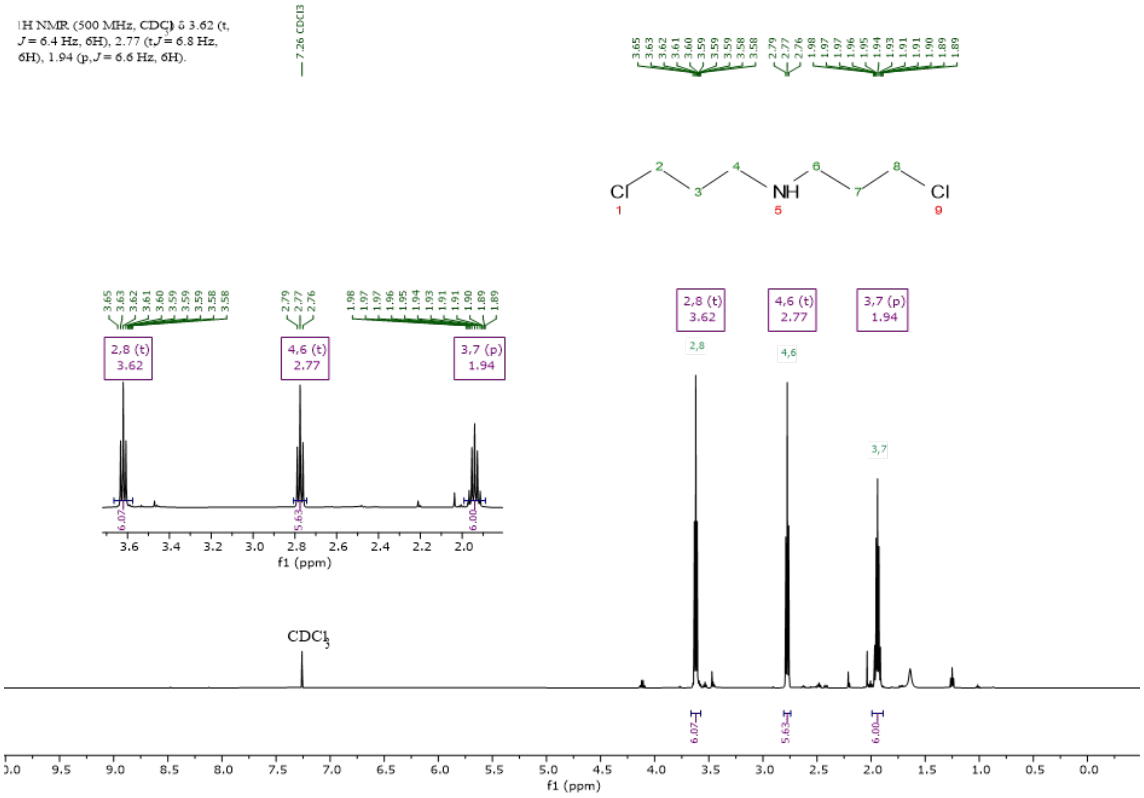


Figure B.4: ^{1}H NMR spectrum of bis(3-chloropropyl)amine, **6**.

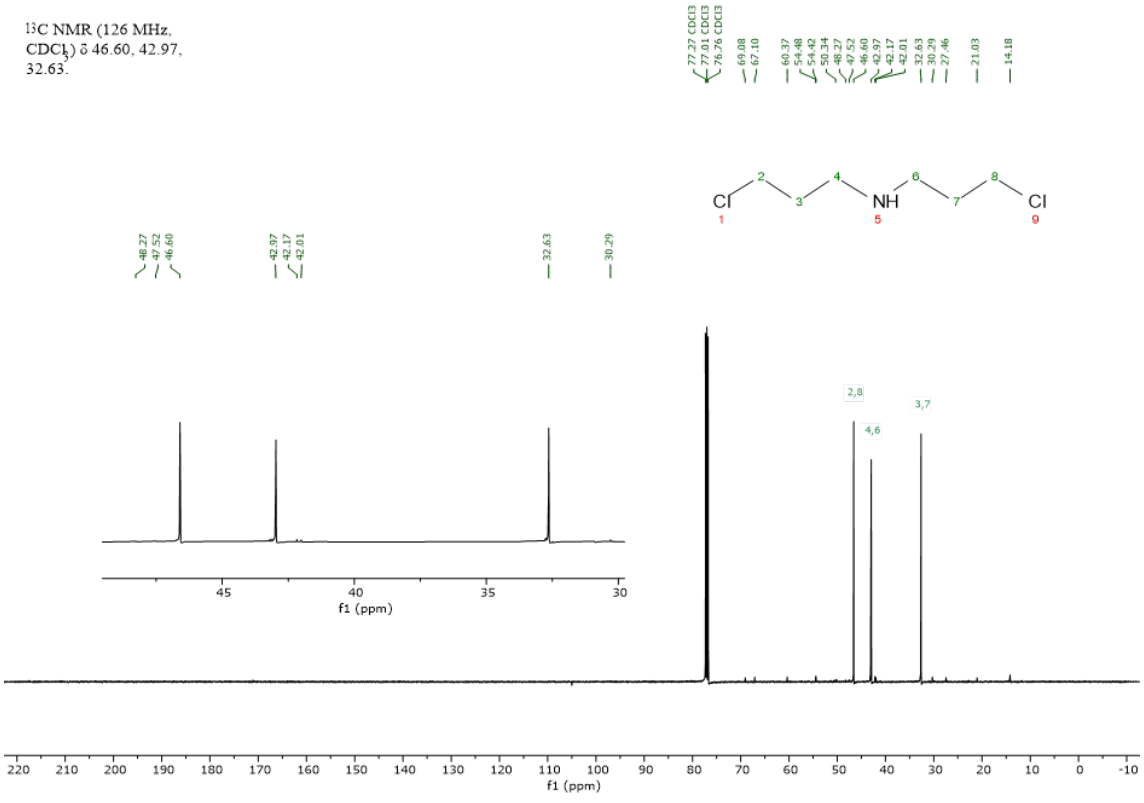


Figure B.5: ^{13}C NMR spectrum of bis(3-chloropropyl)amine, **6**.

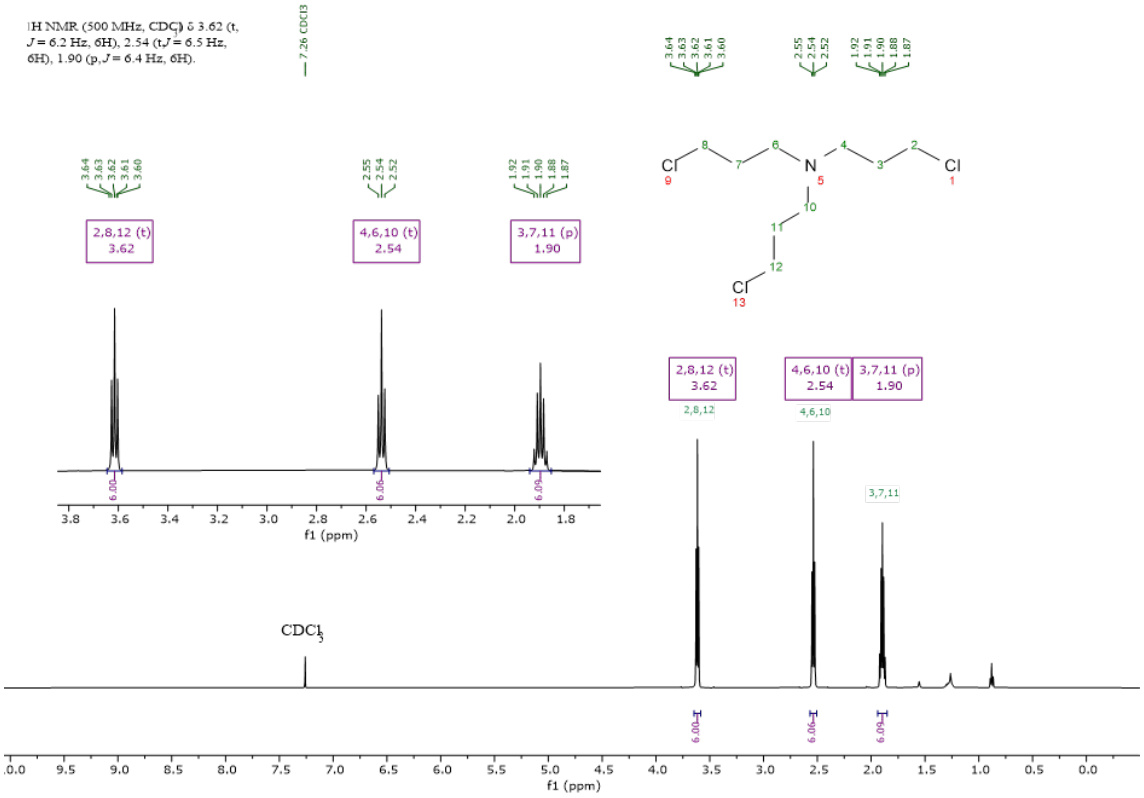


Figure B.6: ^{1}H NMR spectrum of tris(3-chloropropyl)amine, **5**.

Figure B.7: ^{13}C NMR spectrum of tris(3-chloropropyl)amine, **5**.

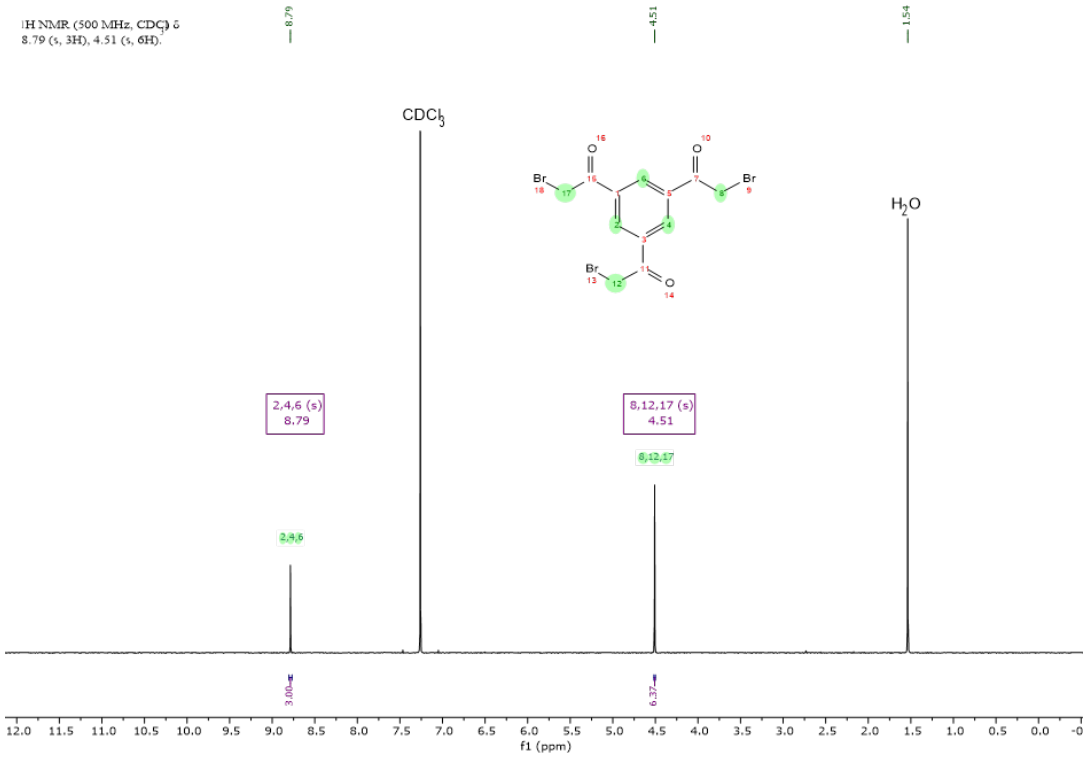


Figure B.8: ^{1}H NMR spectrum of 1,1',1"-(benzene-1,3,5-triyl)tris(2-bromoethan-1-one), **9**.

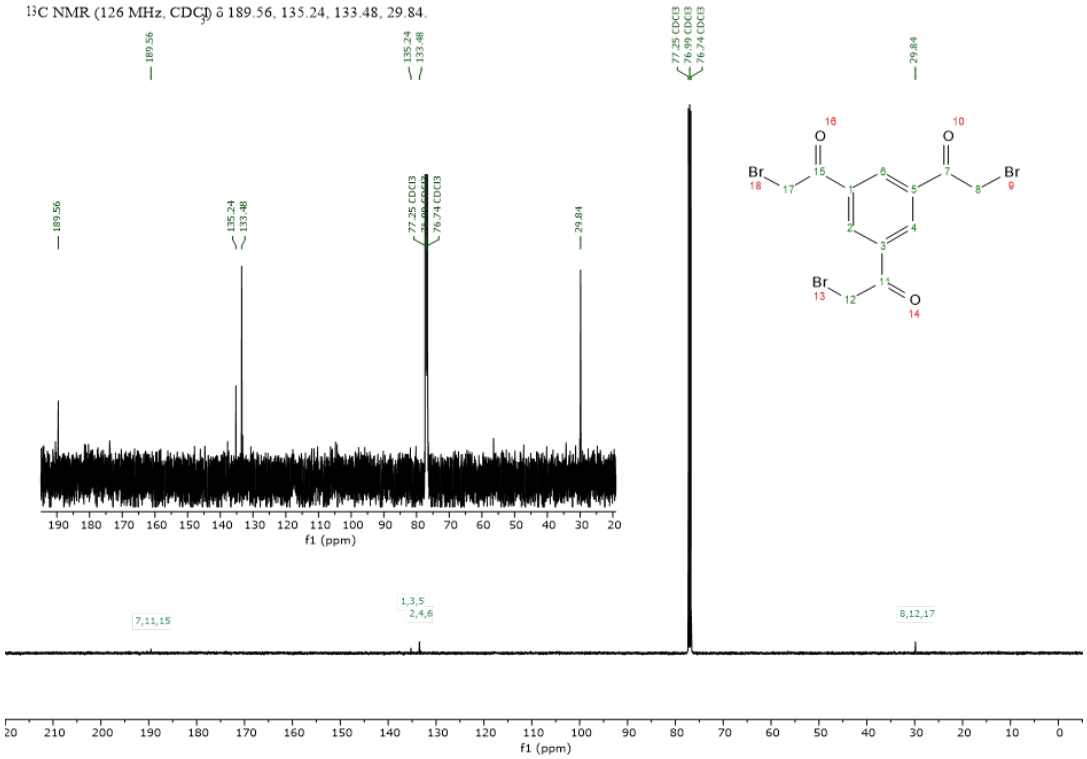


Figure B.9: ^{13}C NMR spectrum of 1,1',1"-(benzene-1,3,5-triyl)tris(2-bromoethan-1-one), **9**.

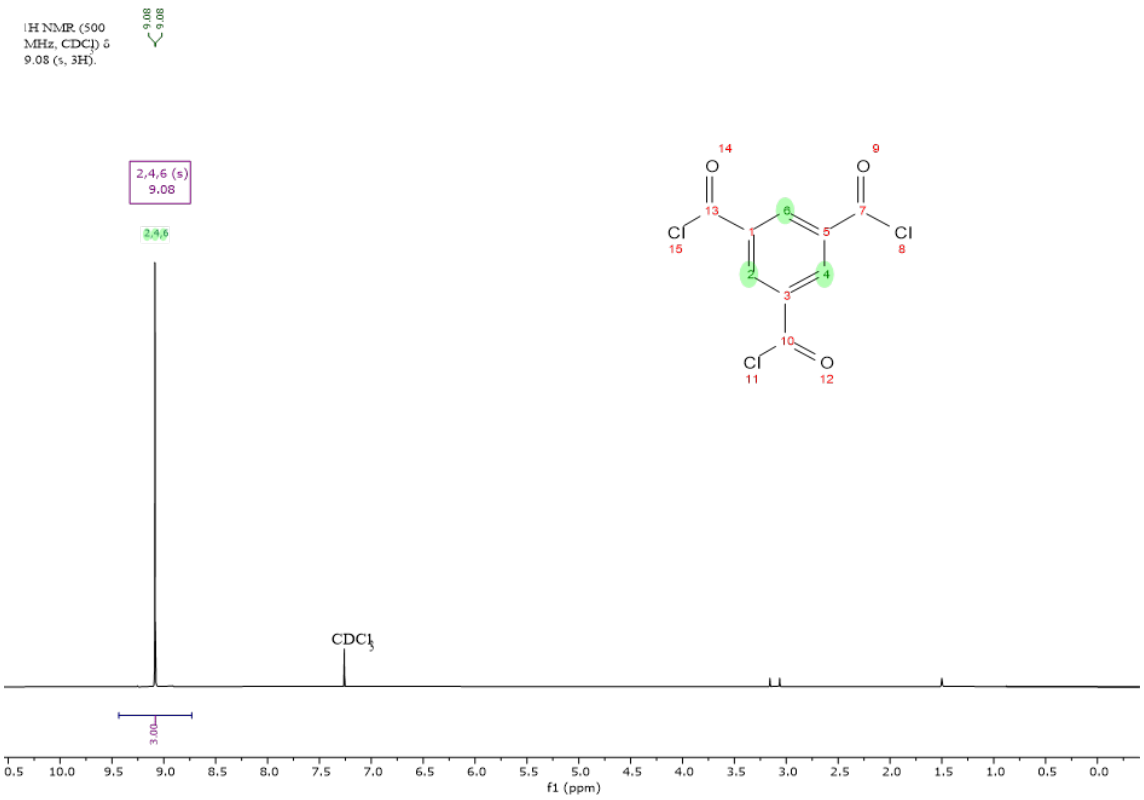


Figure B.10: ^{1}H NMR spectrum of benzene-1,3,5-tricarbonyl trichloride, **14**.

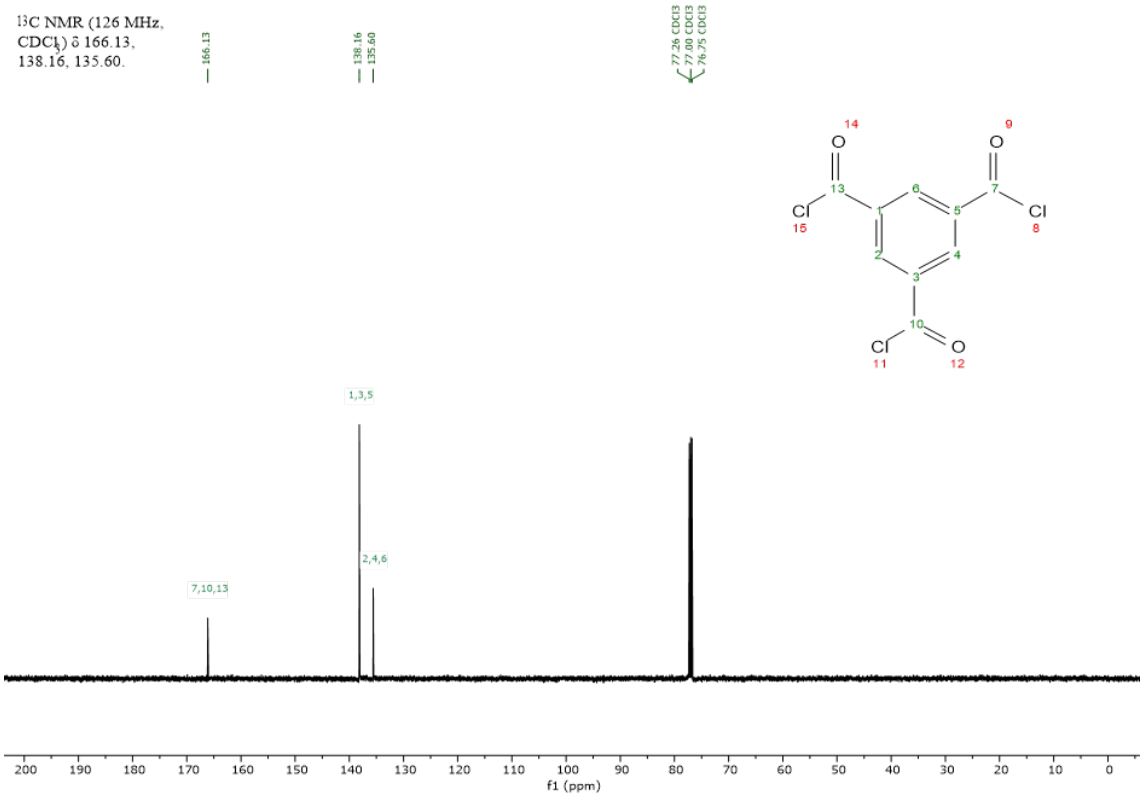


Figure B.11: ^{13}C NMR spectrum of benzene-1,3,5-tricarbonyl trichloride, **14**.

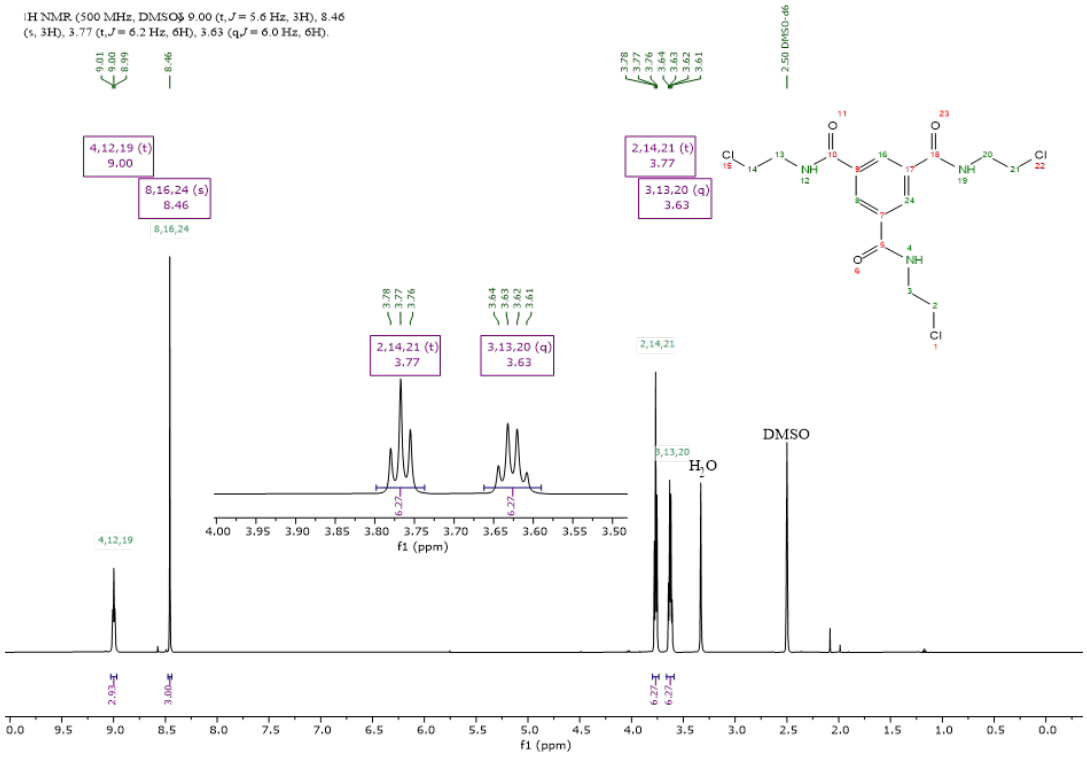


Figure B.12: ^{1}H NMR spectrum of N^{1},N^{3},N^{5}-tris(2-chloroethyl)-benzene-1,3,5-tricarboxamide, **16**.

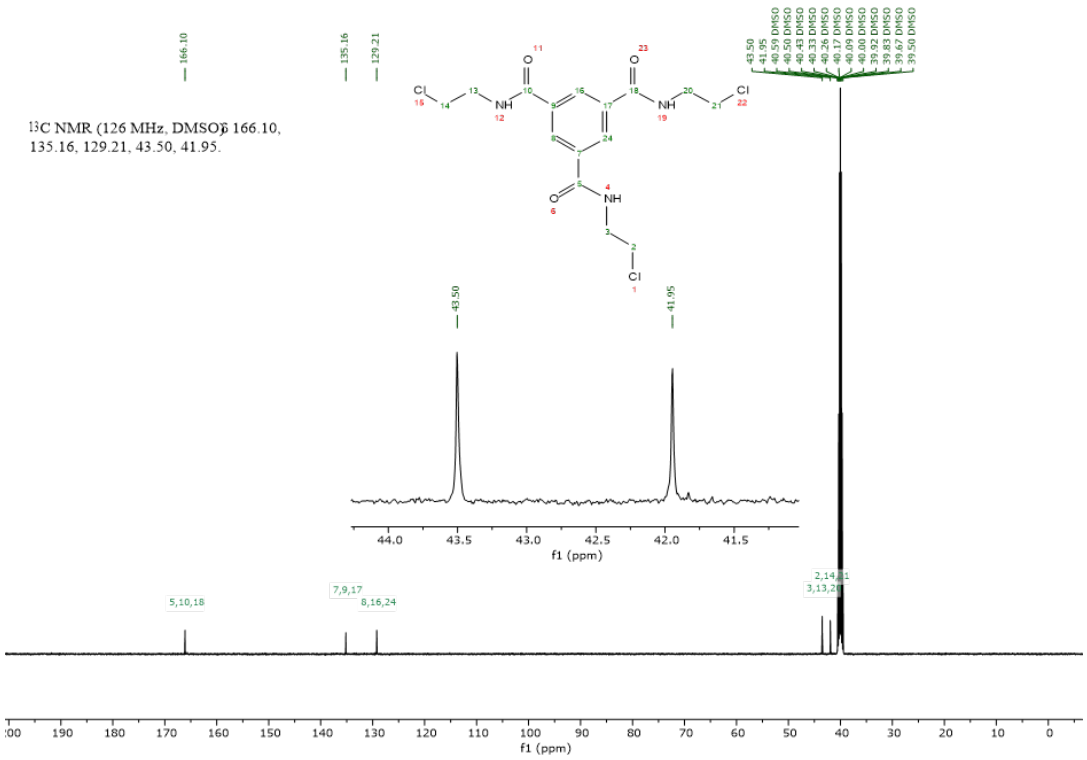


Figure B.13: ^{13}C NMR spectrum of N^{1},N^{3},N^{5}-tris(2-chloroethyl)benzene-1,3,5-tricarboxamide, **16**.

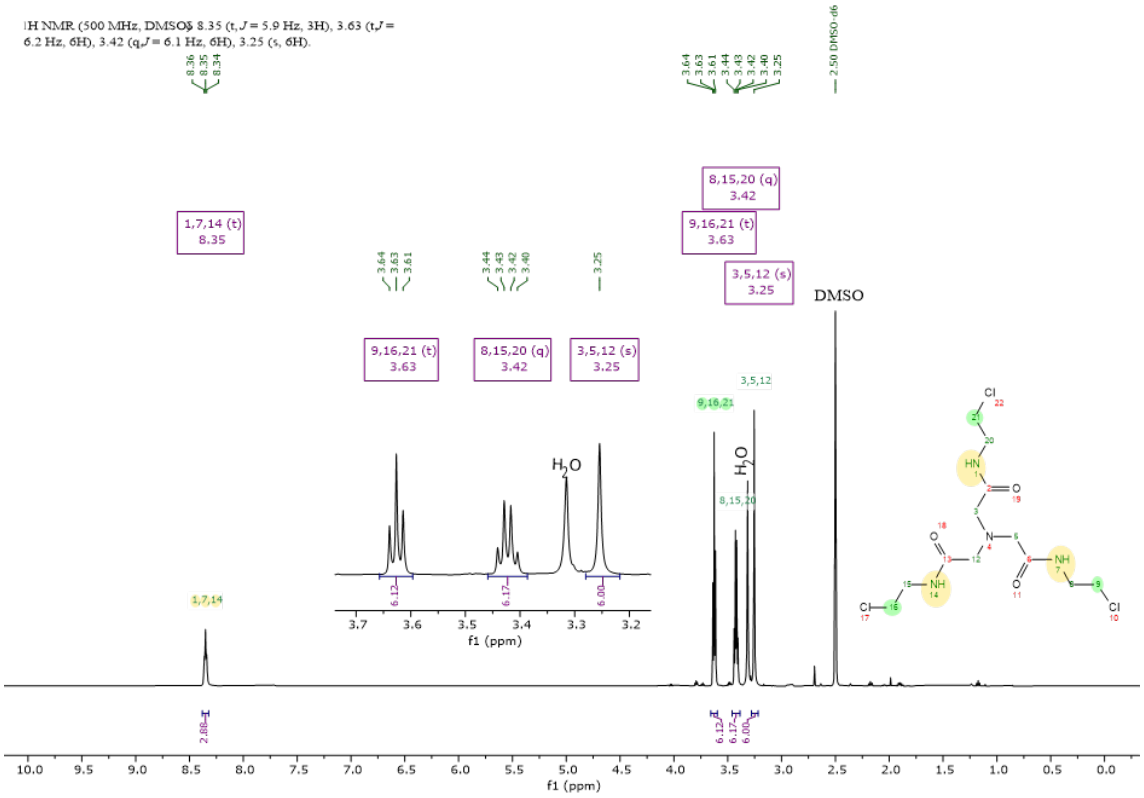


Figure B.14: ^{1}H NMR spectrum of 2,2',2"-nitrilotris(*N*-(2-chloroethyl)acetamide), **21**.

B.2 Flash column operating conditions

In this section are reported the mobile phase gradient and corresponding column volumes used for purification of the linkers during flash chromatography.

Step	Mobile phase: percentage of EtOAc in heptane, %		Column Volume (CV)
	Start	End	
Equilibriate	0	0	3
1	0	10	3
2	10	10	5
3	10	20	3

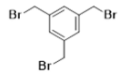
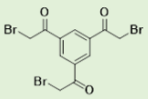
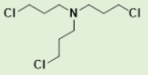
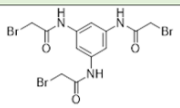
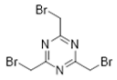
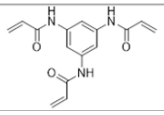
Figure B.15: Table showing the mobile phase gradient and corresponding column volumes used for purification of the linker, TCPA.

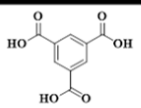
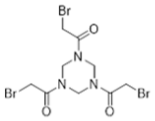
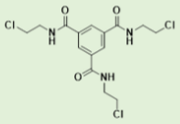
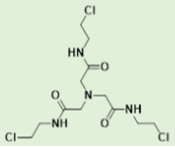
Step	Mobile phase: percentage of EtOAc in heptane, %		Column Volume (CV)
	Start	End	
Equilibriate	0	0	3
1	0	5	3
2	5	8	2
3	8	10	3
4	10	13	3
5	13	15	2
6	15	20	4
7	20	30	3
8	30	30	1

Figure B.16: Table showing the mobile phase gradient and corresponding column volumes used for purification of the linker, BTBE.

Step	Mobile phase: percentage of MeOH in EtOAc, %		Column Volume (CV)
	Start	End	
Equilibriate	0	0	3
1	0	1	1
2	1	3	1
3	3	5	3
4	5	10	3
5	10	20	3

Figure B.17: Table showing the mobile phase gradient and corresponding column volumes used for purification of the linker, NTCA.

Linker	Structure	clogP
TBMB		4.32 ± 0.27
BTBE		3.09 ± 0.85
TCPA		3.02 ± 0.27
TBAB		2.56 ± 0.45
TBMT		2.22 ± 0.53
TAAB		2.01 ± 0.42

Linker	Structure	clogP
Trimesic acid		1.51 ± 0.25
TATB		1.14 ± 0.79
TBCA		0.76 ± 0.55
NTCA		0.48 ± 0.67

*Synthesized linkers are highlighted in green.
*Values predicted on Chemscketch, ACD labs

Figure B.18: Calculated logP values of synthesized and reported linkers.