



**HAL**  
open science

## Design of bio-inspired catalysts based on a gamma-peptide foldamer architecture

Julie Aguesseau

► **To cite this version:**

Julie Aguesseau. Design of bio-inspired catalysts based on a gamma-peptide foldamer architecture. Other [q-bio.OT]. Université Montpellier, 2019. English. NNT : 2019MONT043 . tel-02403072

**HAL Id: tel-02403072**

**<https://theses.hal.science/tel-02403072>**

Submitted on 10 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Ingénierie Biomoléculaire

École doctorale Sciences Chimiques Balard (ED 459)

Unité de recherche Institut des Biomolécules Max Mousseron UMR 5247

Elaboration de catalyseurs bio-inspirés conçus autour  
d'une architecture  $\gamma$ -peptidique auto-structurée

*Design of bio-inspired catalysts based on a  $\gamma$ -peptide  
foldamer architecture*

Présentée par Julie AGUESSEAU-KONDROTAS

Le 18 Septembre 2019

Sous la direction de Ludovic MAILLARD

Devant le jury composé de

Gilles GUICHARD,  
Directeur de recherche CNRS, Université de Bordeaux  
Grégory CHAUME,  
Maître de conférences, Université de Cergy-Pontoise  
David AITKEN,  
Professeur, Université Paris-Sud  
Michael THORMANN,  
CSO, Origenis GmbH  
Renata MARCIA DE FIGUEIREDO,  
Chargée de recherche CNRS, Université de Montpellier

Rapporteur  
Rapporteur  
Président du Jury - Examineur  
Membre invité  
Membre invité



UNIVERSITÉ  
DE MONTPELLIER



*A mon mari, Mindaugas,  
A mes parents, Edith & Christophe,  
A mes sœurs, Manon & Clara,  
A mon frère, Mattéo,*

*je vous aime.*



## ACKNOWLEDGEMENTS

Premièrement, j'aimerais remercier mon directeur de thèse, Docteur Ludovic Maillard et ma co-encadrante, Docteur Renata Marcia De Figueiredo. Ludovic, tu as su être à mon écoute, me guider lorsqu'il le fallait, m'écouter et me faire confiance lorsque je souhaitais voler de mes propres ailes. Pour ta confiance en moi, ta disponibilité et tes précieux conseils, je te suis grandement reconnaissante. Renata, toujours disponible et volontaire, je te remercie d'avoir accepté d'encadrer mon travail et de m'avoir accueillie occasionnellement dans votre laboratoire. Tous deux, vous m'avez permis d'acquérir de nouvelles connaissances, techniques et scientifiques, et vos qualités humaines n'ont fait que renforcer mon enthousiasme et mon envie de développer ce projet. Merci.

Je tiens sincèrement à remercier les Docteurs Gilles Guichard et Grégory Chaume qui me font l'honneur d'être les rapporteurs de mon jury de thèse ainsi que le Professeur David Aitken qui a gentiment accepté d'examiner mon travail.

I would like to address my special thanks to Doctor Michael Thormann who accepted to evaluate my work and who has always encouraged me through the years. I am looking forward to working with you again!

Je remercie le Professeur Jean-Louis Bantignières et David Maurin d'avoir pris le temps de m'accueillir au sein de leur laboratoire et m'avoir formée pour les acquisitions en spectroscopie infrarouge. J'adresse mes remerciements au Docteur Baptiste Legrand pour son implication dans les analyses structurales de nos oligomères.

Mes remerciements vont également à toutes les personnes de l'équipe F9 de l'IBMM. J'en profite pour remercier chaleureusement le Docteur Matthieu Simon pour nos échanges sur ce projet mais également pour sa bonne humeur et ses encouragements. Viennent ensuite les remerciements aux doctorants, Christophe, mon fidèle voisin de bureau et de labo, tu as été le premier à m'accueillir dans ces locaux, grâce à toi, je me suis sentie un peu moins perdue dans ce dédale de labos. Barth, je pourrai faire un paragraphe juste pour toi mon ami ! Merci d'avoir été là, merci pour toutes les soirées qu'on a passé durant ces trois ans. Ces « Smash up » où j'ai enfin fini par gagner une partie, nos pas de danse déguisés devant la télé, ... et j'en passe ! Bien sûr, je n'oublie pas tous les autres doctorants, Alice, Julie, Julien, Kevin, Laurent, Laurine, Prisca, Titouan et Yohan ainsi que Manon et Caro ! Merci à vous d'avoir permis à ces trois années d'être festives !

Un grand merci à l'équipe AM2N et au Professeur Jean-Marc Campagne de m'avoir accueillie dans son laboratoire. Je suis reconnaissante envers Julien qui m'a toujours offert une place sur sa paillasse pour mes manip. Jimmy, je te remercie pour ta gentillesse, ta disponibilité et ta volonté de toujours faire au mieux pour tout le monde, c'est aussi grâce à toi que mes passages chez vous étaient si réussis !

Enfin, j'aimerais remercier mes meilleures amies, Marie et Marjorie, toujours disponibles pour un coup de fil malgré la distance qui nous sépare. Merci pour votre soutien les filles !

Bien entendu, je n'aurai jamais accompli cela sans ma famille : mes parents Edith et Christophe, mes sœurs Manon et Clara, et mon frère Mattéo. Vous n'avez pas toujours compris en quoi consistaient mes études et mon travail, je pense que vous allez encore moins comprendre ma présentation le jour de la soutenance, mais vous avez toujours été à mes côtés et je vous en remercie !

Last but not least, my husband, Min, thank you. Thank you for being by my side every day, providing me with this wonderful support. Of course, I cannot finish this paragraph without thanking you for proofreading this manuscript, you are amazing!

# TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS .....</b>	<b>5</b>
<b>TABLE OF CONTENTS .....</b>	<b>7</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>11</b>
<b>NOMENCLATURE.....</b>	<b>15</b>
<b>GENERAL INTRODUCTION .....</b>	<b>19</b>
<b>CHAPTER 1: BIBLIOGRAPHIC INTRODUCTION .....</b>	<b>23</b>
I.    SECONDARY STRUCTURES ASSOCIATED WITH PROTEINS DIHEDRAL ANGLES .....	27
1. <i>The helical secondary structures</i> .....	29
2. <i>The <math>\beta</math>-sheets</i> .....	31
3. <i>The turns</i> .....	33
i. $\beta$ -turns .....	34
ii. $\gamma$ -turns.....	36
4. <i>Secondary structure predictions for the natural amino acids in protein conformation</i> .....	37
II.   FOLDAMERS AND CATALYSIS .....	39
1. <i>Definition foldamer</i> .....	39
i.    Considerations from the dynamics of $\alpha$ -peptides and peptoids .....	42
ii. $\beta$ -peptides .....	44
ii.a.   14-helix .....	45
ii.b.   12-helix .....	47
ii.c.   12/10-helix .....	48
ii.d.   10-helix .....	49
ii.e.   8-helix .....	49
iii. $\gamma$ -peptides .....	51
iii.a.   9- and 14-helices: two structures accessible to acyclic $\gamma$ -peptides.....	52
iii.b.   Structures accessible to cyclic aliphatic $\gamma$ -peptides .....	56
iii.c.   ATC-based oligomers as prototype of aromatic $\gamma$ -peptide foldamers.....	56
iii.d.   Oligoureas peptidomimetics .....	59
iii.e.   Conclusion .....	61
2. <i>Catalytic foldamers</i> .....	62
i.    Peptoids for enantioselective catalysis .....	63
ii. $\beta$ -Peptides as catalysts .....	64
ii.a.   Juliá-Colonna asymmetric epoxidation of enones .....	64
ii.b.   Retroaldol cleavage of $\beta$ -hydroxyketones .....	64
ii.c.   Hydrolysis of trisulfonate esters.....	65
ii.d.   Bifunctional catalysis .....	66
iii.   Aib-oligomers and oligoureas for the C-C bond formation and control of chirality.....	68

III.	THESIS OBJECTIVES .....	71
<b>CHAPTER 2: INTRODUCTION OF A CATALYTIC FUNCTION ON AN ATC-OLIGOMER .....</b>		<b>73</b>
I.	ENAMINE CATALYSIS .....	75
1.	<i>History: proline as a prototype for enamine catalysis</i> .....	75
i.	Asymmetric aldol reaction .....	77
ii.	Asymmetric three-component Mannich reaction.....	79
iii.	Michael addition reaction .....	82
2.	<i>Peptides in enamine catalysis</i> .....	84
i.	H-Pro-Pro-Asp-NH <sub>2</sub> as a prototype catalyst for aldol reaction .....	84
ii.	Scope extension: from aldol reaction to nitro-Michael addition .....	86
iii.	Mechanistic studies of the nitro-Michael addition catalysed by H-D-Pro-Pro-Asp-NH <sub>2</sub> .....	87
iv.	Structural modulation of the tripeptide catalyst .....	89
iv.a.	Effect of ring-size analogues of H-D-Pro-Pro-Glu-NH <sub>2</sub> .....	89
iv.b.	Effect of β <sup>3</sup> -amino acids analogues of H-D-Pro-Pro-Glu-NH <sub>2</sub> .....	91
II.	DESIGN AND SYNTHESIS OF COMPETENT ATCS FOR ENAMINE-TYPE CATALYSIS.....	92
1.	<i>ATC design</i> .....	92
2.	<i>ATC synthesis</i> .....	92
3.	<i>Evaluation of the catalytic properties of the bifunctional ATC-monomer (46)</i> .....	96
4.	<i>Optimization of the nitro-Michael addition reaction catalysed by (46)</i> .....	97
<b>CHAPTER 3: CATALYTIC PROPERTIES ACCORDING TO THE OLIGOMER LENGTHS AND THE POSITION OF THE COMPETENT ATC .....</b>		<b>99</b>
I.	NOMENCLATURE USED FOR THE ATC-OLIGOMERS .....	101
II.	DESIGN AND SYNTHESIS OF FOLDED ATC-BASED CATALYSTS.....	102
1.	<i>Synthesis of (47)</i> .....	103
2.	<i>Synthesis of the ATC-trimers ATC<sub>3</sub>-2-NHiPr, ATC<sub>3</sub>-2*-NHiPr and ATC<sub>3</sub>-2(OMe)-NHiPr</i> .....	103
3.	<i>Synthesis of the ATC-pentamer ATC<sub>5</sub>-3-NHiPr</i> .....	104
4.	<i>Synthesis of ATC-pentamer ATC<sub>5</sub>-3-NH<sub>2</sub></i> .....	105
III.	COMPARISON OF THE CATALYTIC PROPERTIES WITH THE ATC NUMBER UNITS AND THE C-TER EXTREMITY .....	107
IV.	EVALUATION OF THE OVERALL FOLDING OF THE CATALYSTS .....	108
1.	<i>Characteristic structural markers for ATC-based helical foldamers</i> .....	108
i.	Circular dichroism .....	108
ii.	Nuclear Magnetic Resonance structural markers .....	109
iii.	Mid infrared spectroscopy .....	112
2.	<i>Structural markers associated to the ATC-catalysts</i> .....	113
i.	Global folding of (47) .....	113
ii.	Global folding of ATC <sub>3</sub> -2-NHiPr, ATC <sub>5</sub> -3-NHiPr and ATC <sub>5</sub> -3-NH <sub>2</sub> .....	114
ii.a.	Circular dichroism signatures .....	114
ii.b.	<sup>1</sup> H NMR indicators .....	115

ii.c. Fourier transform infrared indicators.....	117
V. MODULATION OF THE POSITION OF THE CATALYTIC CENTRE .....	119
1. <i>Synthesis of ATC-pentamers ATC<sub>5</sub>-P-NH<sub>2</sub> (1 ≤ P ≤ 5)</i> .....	119
2. <i>Comparison of the catalytic properties with the position of the competent ATC moiety</i> .....	121
VI. CONCLUSION .....	122
<b>CHAPTER 4: MODULATION OF THE MICROENVIRONMENT AROUND THE CATALYTIC CENTRE.....</b>	<b>123</b>
I. EFFECT OF THE DISTANCE BETWEEN THE PYRROLIDINE AMINE AND THE CARBOXYLIC ACID.....	125
II. EFFECT OF THE REMOTE LATERAL CHAINS ON THE CATALYSIS .....	128
1. <i>Local modifications of the remote lateral chains</i> .....	128
2. <i>Removal of the steric hindrance induced by the benzyl lateral chains</i> .....	131
III. EXPLORATION OF COOPERATIVE EFFECT .....	135
IV. CONCLUSION .....	138
<b>CHAPTER 5: STRUCTURE DETERMINATION AT THE ATOMIC LEVEL OF THE ATC-OLIGOMERS .....</b>	<b>141</b>
I. GENERALITIES ON NMR REFINEMENT .....	143
II. 3D-MODELLING UNDER NMR RESTRAINTS .....	145
1. <i>Method 1: Simulated annealing protocol and limitation of the NMR refinements</i> .....	145
2. <i>Variation of the NOE restraints</i> .....	148
3. <i>Setting the force field</i> .....	151
i. Method 2: Creation of the AMBER parameter files for individual ATC monomers using GAFF .....	153
i.a. Generation of ATC monomer 3D-structures.....	153
i.b. Application of the GAFF force field on the ATC monomers and generation of the AMBER file types ....	154
ii. Method 3: Optimization of the ATC-force field by DFT calculation .....	156
iii. Building an oligomer sequence and generation of the related Amber parameter files.....	159
4. <i>Validation of the ATC-optimized force field</i> .....	159
i. Comparative performances of Methods 1, 2 and 3 .....	159
ii. Validation of the calculated 3D models by comparison with the XRD structure, CD and FTIR data .....	162
III. APPLICATION OF THE OPTIMIZED SIMULATED ANNEALING PROTOCOLE TO THE STRUCTURAL DETERMINATION OF A SET OF CATALYTIC ATC-OLIGOMERS.....	164
1. <i>Structures according to the position of the catalytic residue</i> .....	164
2. <i>Limitations of the method</i> .....	164
3. <i>Structural validation of other M(X)-Mutant oligomers</i> .....	166
IV. CONCLUSION .....	167
<b>CONCLUSION AND OUTLOOK .....</b>	<b>171</b>
<b>RESUME EN FRANÇAIS .....</b>	<b>175</b>
<b>SUPPORTING INFORMATION.....</b>	<b>197 - S1</b>
<b>REFERENCES .....</b>	<b>385</b>



## LIST OF ABBREVIATIONS

<b>Å</b>	Angström
<b>aa</b>	Amino acid
<b>Ac</b>	Acetyl
<b>ABHC</b>	2-amino-6,6-dimethyl-bicyclo[3.1.1]-heptane-3-carboxylic acid
<b>ACBC</b>	Aminocyclobutanecarboxylic acid
<b>ACHC</b>	Aminocyclohexanecarboxylic acid
<b>ACN</b>	Acetonitrile
<b>ACPC</b>	Aminocyclopentanecarboxylic acid
<b>Aib</b>	Aminoisobutyric acid
<b>AMBER</b>	Assisted Model Building with Energy Refinement
<b>ATC</b>	4-amino-(methyl)-1,3-thiazole-5-carboxylic acid
<b>Aze</b>	Azetidine carboxylic acid
<b>Bn</b>	Benzyl
<b>Boc</b>	<i>tert</i> -Butyloxycarbonyl protecting group
<b>Bu</b>	Butyl
<b>C-ter</b>	Carboxyl terminus
<b>Caa</b>	Carbo-amino acid
<b>Calcd.</b>	Calculated
<b>Cat.</b>	Catalyst
<b>CD</b>	Circular Dichroism
<b>CDCl<sub>3</sub></b>	Deuterated Chloroform
<b>CDI</b>	Carbonyldiimidazole
<b>CH<sub>3</sub>CN</b>	Acetonitrile
<b>CHCl<sub>3</sub></b>	Chloroform
<b>COSY</b>	COrrrelation SpectroscopY
<b>C<sub>x</sub></b>	x-membered hydrogen bond ring
<b>d</b>	Doublet
<b>DCM</b>	Dichloromethane
<b>de</b>	Diastereomeric Excess
<b>DFT</b>	Density Functional Theory
<b>DIC</b>	<i>N,N'</i> -Diisopropylcarbodiimide
<b>DIEA</b>	<i>N,N</i> -Diisopropylethylamine
<b>DMAP</b>	4-Dimethylaminopyridine
<b>DMF</b>	Dimethylformamide
<b>DMSO</b>	Dimethylsulfoxide

---

<b>dr</b>	Diastereomeric Ratio
<b>e. g.</b>	<i>exempli gratiā</i>
<b>EDCI</b>	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
<b>ee</b>	Enantiomeric Excess
<b>equiv.</b>	Equivalent
<b>er</b>	Enantiomeric Ratio
<b>ESI</b>	Electrospray ionization
<b>Et<sub>3</sub>N</b>	Triethylamine
<b>EtOAc</b>	Ethyl Acetate
<b>EtOH</b>	Ethanol
<b>EWG</b>	Electron Withdrawing Group
<b>Fmoc</b>	Fluorenylmethyloxycarbonyl
<b>FTIR</b>	Fourier-Transform InfraRed spectroscopy
<b>GABA</b>	$\gamma$ -aminobutyric acid
<b>GAFF</b>	General Amber Force Field
<b>Gpn</b>	Gabapentin
<b>h</b>	Hour
<b>Hex</b>	Hexane
<b>HMBC</b>	Heteronuclear Multiple-Bond Correlation spectroscopy
<b>HOBt</b>	Hydroxybenzotriazole
<b>HPLC</b>	High-Performance Liquid Chromatography
<b>HSQC</b>	Heteronuclear Single-Quantum Correlation spectroscopy
<b>Hz</b>	Hertz
<b><i>i</i></b>	iso
<b>i. e.</b>	<i>id est</i>
<b><i>i</i>BCF</b>	<i>isobutylchloroformate</i>
<b><i>i</i>Pr</b>	<i>isopropyl</i>
<b><i>J</i></b>	Coupling constant
<b>K</b>	Kelvin
<b>LCMS</b>	Liquid Chromatography Mass Spectrometry
<b>LiHMDS</b>	Lithium bis(trimethylsilyl)amide
<b>(M)</b>	Minus (left-handed helix)
<b>M</b>	Molar = mol/L
<b>m</b>	Multiplet
<b>MD</b>	Molecular Dynamic
<b>MeOH</b>	Methanol
<b>MO</b>	Molecular Orbital

<b>mre</b>	Molar ellipticity per residue
<b>MS</b>	Mass Spectrometry
<b>N-ter</b>	Amino terminus
<b>na</b>	Not assigned
<b>nd</b>	Not detected
<b>NBS</b>	<i>N</i> -Bromosuccinimide
<b>NMM</b>	<i>N</i> -Methylmorpholine
<b>NMP</b>	<i>N</i> -Methyl-2-pyrrolidone
<b>NMR</b>	Nuclear Magnetic Resonance
<b>NOE</b>	Nuclear Overhauser Effect
<b>o</b>	Overlap
<b>(P)</b>	Plus (right-handed helix)
<b>PG</b>	Protecting Group
<b>Ph</b>	Phenyl
<b>Pip</b>	Piperidine
<b>ppm</b>	Parts Per Million
<b>PrOH</b>	Propanol
<b>quant.</b>	Quantitative
<b>RMSD</b>	Root Mean Square Deviation
<b>ROESY</b>	Rotating frame nuclear Overhauser Effect Spectroscopy
<b>rt</b>	Retention Time
<b>RT</b>	Room Temperature
<b>s</b>	Singlet
<b>SA</b>	Simulated Annealing
<b>SPPS</b>	Solid Phase Peptide Synthesis
<b><i>t</i></b>	tert
<b>t</b>	Triplet
<b>T</b>	Temperature
<b>TEMPO</b>	2,2,6,6-tetramethylpiperidine-1-oxyl
<b>TFA</b>	Trifluoroacetic acid
<b>THF</b>	Tetrahydrofuran
<b>TIS</b>	Tris(propan-2-yl)silyl
<b>TOCSY</b>	Total Correlation Spectroscopy
<b>UPLC</b>	Ultra Performance Liquid Chromatography
<b>XRC</b>	X-Ray Crystallography



## NOMENCLATURE

Amino acids will be designated by their full names, their 3- or 1-letter codes without difference.

Glycine (Gly, G)	Alanine (Ala, A)	Valine (Val, V)	Leucine (Leu, L)	Isoleucine (Ile, I)
Methionine (Met, M)	Proline (Pro, P)	Phenylalanine (Phe, F)	Tyrosine (Tyr, Y)	Tryptophan (Trp, W)
Asparagine (Asn, N)	Glutamine (Gln, Q)	Threonine (Thr, T)	Serine (Ser, S)	Cysteine (Cys, C)
Aspartic Acid (Asp, D)	Glutamic Acid (Glu, E)	Lysine (Lys, K)	Arginine (Arg, R)	Histidine (His, H)

Figure 1: The twenty amino acids found in the standard genetic code

ATC monomers will be designated in this document as following: **PG-NH-{R $\tau$ ;R $\gamma$ }-ATC-OH**  
 Chemset numbering is related to the lateral chains **R $\tau$** , **R $\gamma$** , where **R $\tau$**  corresponds to the substituent bore by the thiazole ring in position 2 and **R $\gamma$**  to the  $\gamma$  lateral chain. **OH** means that the acidic function on the C-terminal extremity is unprotected and **PG-NH** corresponds to the amine function on the N-terminal side, with its protecting group.

For the 3D-modelling studies, a nomenclature based on three characters was necessary; it can be found below each structure. For further information, the two first letters correspond to the amino acid used as a starting block for the design of the monomer (e.g., LF means L-Phenylalanine, LA: L-Alanine and etc). The digit refers to the chain carried by the thiazole ring, 1 for a methyl, 2 for (S)-pyrrolidine and so on.

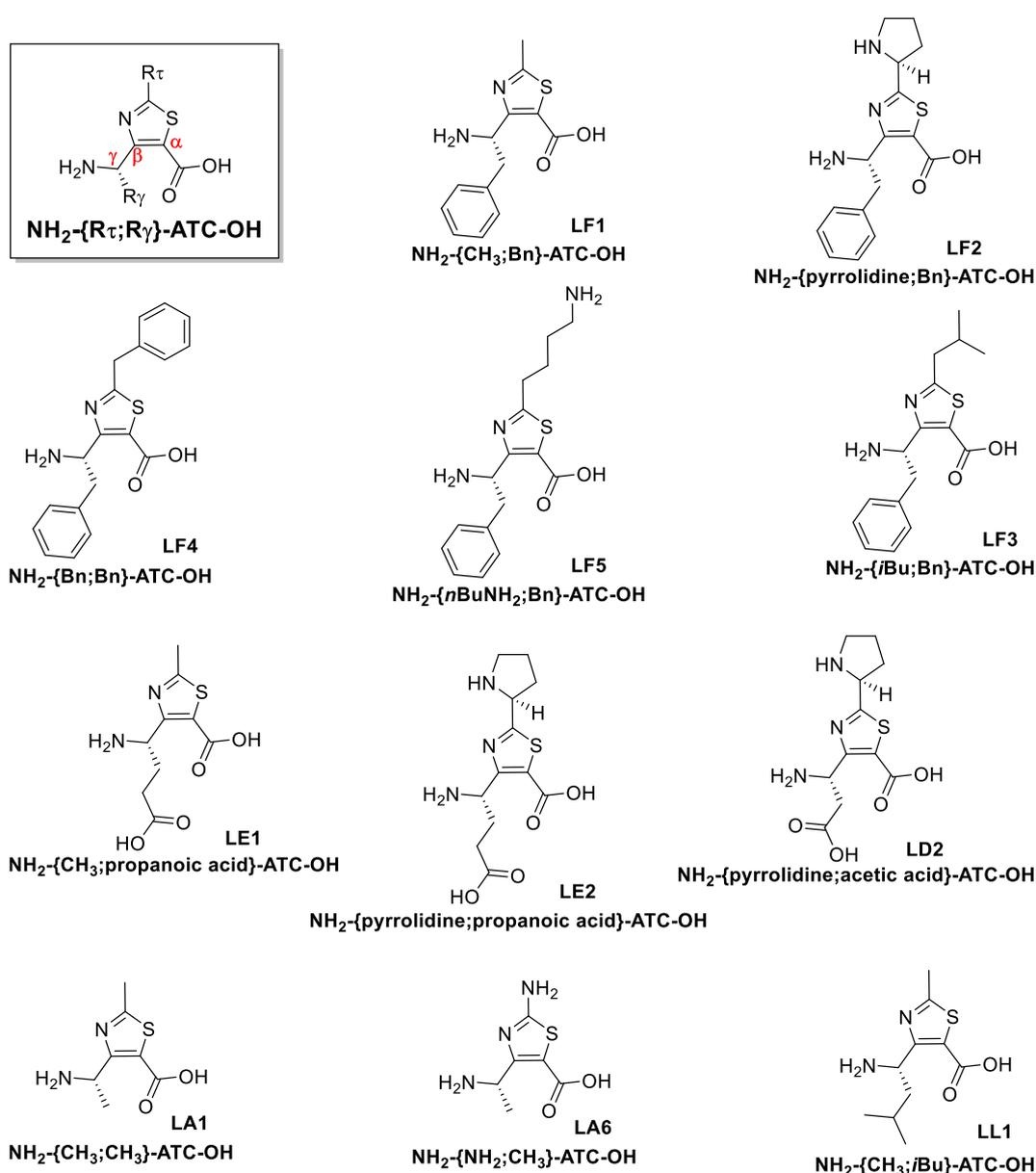


Figure 2: Nomenclature of ATC-monomers

The nomenclature used for the ATC-oligomers in this document is defined as following: when ATC-oligomers contain **Fmoc-NH-{Pyrrolidine;propanoic acid}-ATC-OH (45i)** and **Fmoc-NH-{CH<sub>3</sub>;Bn}-ATC-OH (45a)** residues, these are defined as **ATC<sub>L</sub>-P-X**, with **L** the length of the oligomer (3 for a trimer, 5 for a pentamer and etc), **P** the position along the chain of the ATC unit bearing the pyrrolidine group, from the N-Ter to the C-terminal extremity and **X** the C-terminal extremity (for example NH<sub>i</sub>Pr or NH<sub>2</sub>).

The pyrrolidine group in the ATC-oligomers is usually set in the (*S*) configuration, however, when the (*R*) configuration is used, a star (\*) is added after the position number (**P**), such as **ATC<sub>L</sub>-P\*-X** represents an ATC-oligomer bearing a pyrrolidine group of (*R*) configuration.

Finally, if the ATC-oligomer contains one residue that differs from (45a) and/or (45i) or (45j), it will be designated as a mutant and thus will be called **M(n)-ATC<sub>L</sub>-P-X** with **n** a number in order of appearance.

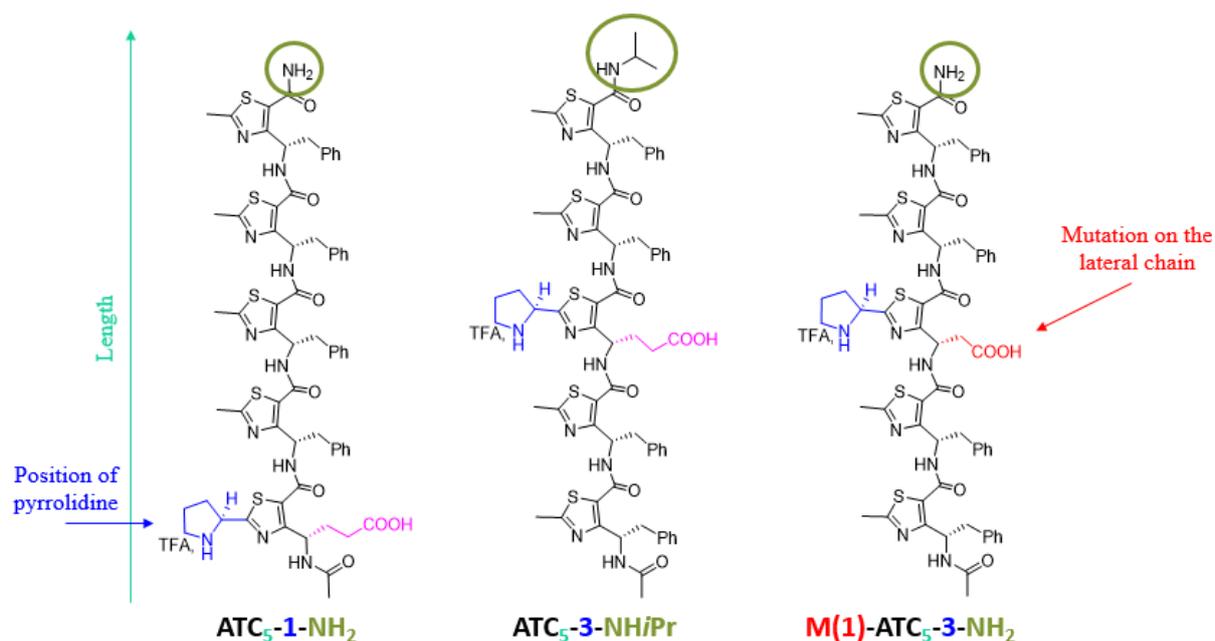


Figure 3: Nomenclature of ATC-oligomers



# GENERAL INTRODUCTION



Current environmental, economic, and societal restrictions are forcing the change in production technologies in favour of renewable resources. Most of the catalysts used in the pharmaceutical industries are small molecules that might be poorly selective or even toxic. In contrast to small molecule catalysts, biocatalysts offer an attractive alternative to conventional chemical methods. Enzymes are powerful catalysts being typically non-toxic and environmentally friendly while capable of performing remarkably difficult chemical transformations with relative ease, unmatched substrate, and product selectivity. In this context, research to design peptides and peptide-based molecules to mimick enzymatic systems seems auspicious. However, for a better efficiency, the enzyme-like catalysts should have a well defined structure, which is hard to control in peptides.<sup>[1]</sup>

Recently, new synthetic systems named “foldamers”, mimicking the protein secondary structures (helices, sheets and ribbons) were developed. They are defined as artificial oligomers with high conformational stability and structural predictabilities.<sup>[2]</sup> Substantial results were obtained and were patented for material and biomedical applications. Nevertheless, their potential for modular catalyst design has been studied little.

The work in this manuscript is devoted to the development of bio-inspired foldamer catalysts and how to adapt a small foldamer to achieve catalytic functions. After reviewing the different levels of protein structures, we will give an overview of some available structures obtained by oligomerization of  $\beta$ - and  $\gamma$ -amino acids. We will focus on the driving forces for the genesis of these structures. It will be shown from the literature that some of those foldamers could potentially be used as catalysts for different reactions. Thereafter, we will concentrate on enamine-type catalysis and how proline-containing peptides inspired the design of a catalytically competent 4-amino-(methyl)-1,3-thiazole-5-carboxylic acid (ATC) monomer.

The next two chapters will be devoted to the introduction of the ATC monomer, which bears the catalytic function, into a  $\gamma$ -peptide foldamer. A systematic exploration of the connection between foldamer sizes and/or shapes, and, both, reactivity and enantioselectivity will be detailed. Depiction of the global folding of ATC-based foldamers using the study of NMR, FTIR and circular dichroism experiments will be elucidated. Eventually, the development of a new script for NMR-based 3D-modelling, used for the structure determination at the atomic level of the synthesized ATC-oligomers, will be detailed.

---

<sup>1</sup> Nanda, V. *et al.*, *Nat. Chem.*, **2**, 15–24 (2010).

<sup>2</sup> Gellman, S. H., *Acc. Chem. Res.*, **31**, 173–80 (1998).



CHAPTER 1:  
BIBLIOGRAPHIC INTRODUCTION



Proteins are defined by four levels of structure, the simplest one being the primary structure, corresponding to the amino acid sequence in a polypeptide chain. The next level is the secondary structure; it refers to the folding of structures that form within a polypeptide to become a  $\alpha$ -helix or a  $\beta$ -pleated-sheet. Both structures are held in shape by hydrogen bonds, formed between the carbonyl oxygen of one amino acid and the amino hydrogen of another one. The tertiary structure is the overall three-dimensional structure of a polypeptide and is mostly due to the interactions between the lateral groups of the amino acids forming the protein. Various bonds among which are hydrogen bonds, electrovalent bonds or disulfide bonds, maintain it. Finally, the quaternary structure corresponds to two or more polypeptide chains that can bond in a non-covalent way and thus create a bigger structure.

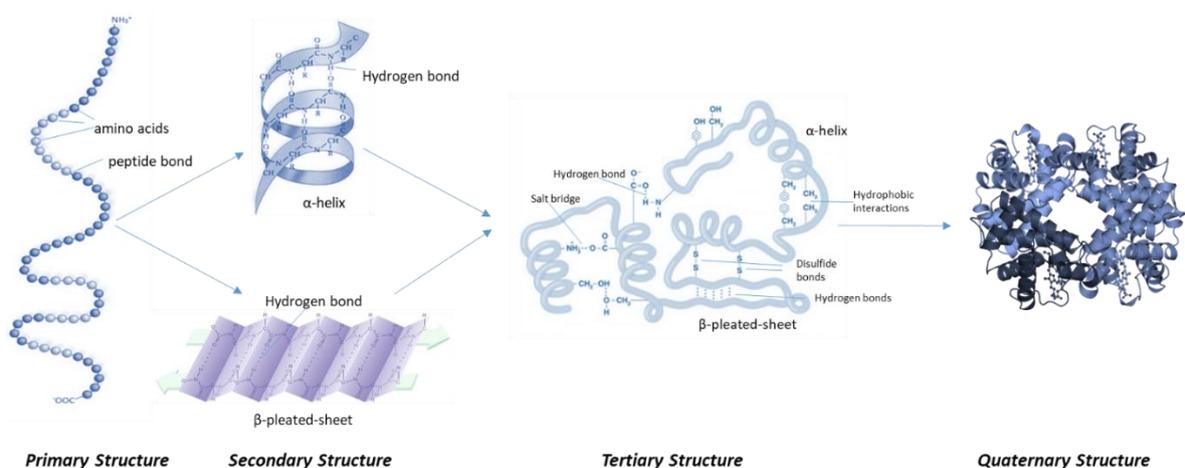


Figure 4: The four levels of protein structure

It has now been widely accepted that the 3D structure of a protein determines its function.<sup>[3]</sup> The correlation between the low conformational stability of many enzymes within temperature, pH or solvent conditions, and, the loss of catalytic activity is evidence correlating with the said fact. The catalytic site of an enzyme is composed solely of a small part of the protein in which only two to four amino acids, constituting the catalytic centre, are directly involved in the chemical transformation. The other residues in the catalytic site are useful for their orientations and dynamics as well as for the substrate recognition. In such a context, one major area to address is understanding the role of the peripheral residues, meaning those that do not constitute the active site. Indeed, by introducing catalytic centres into small peptides, the catalytic enzyme properties could be mimicked and, the molecular complexity of the catalyst could theoretically be reduced.

<sup>3</sup> Berg, J. *et al.*, in *Biochem. 5th Ed.* (New York, NY: W H Freeman, 2002).

So far, many attempts have been made to recreate all the desired catalytic properties of the enzymes starting from short peptide sequences. However, outside a protein context, small to medium-size peptides are often highly flexible, meaning that they do not adopt a single conformation but more likely a dynamic conformational ensemble. In the proteins, the conformation of a small peptide sequence is directly related to its local environment, which is to the presence of other elements of the secondary structure close to each other. Therefore, mimicking the enzyme activity with a peptide is challenging as the folding of this peptide needs to be controlled to a certain extent.

Some progress was made in the field and the desire to control the structural properties of peptides is very prominent in research nowadays.<sup>[4,5]</sup> Hence, the concept of restricting the conformation of peptides in order to mimic the activity of another natural peptide or protein was adopted in the design of molecules called peptidomimetics.<sup>[6]</sup> In order to reduce the flexibility of the peptide and control its shape, different approaches were considered. The most common approach is the incorporation of disulfide bridges, whereas, another possibility is to tie together several side chains, for example a glutamate and a lysine in an amide bridge in order to create a “multiple side chain knotting”.<sup>[7]</sup> Additionally, one can envisage using chemical elements, different from peptides able to mimic typical secondary structures, such as the  $\alpha$ -helices,  $\beta$ -strands or turns, found in proteins. The structuration properties within an artificial oligomer, associated with the properties carried by its internal specificities would permit to access new functions, like catalysis, specific binding, or controlled crystallization of inorganic phases. This would not be conceivable if the molecule was only a covalent succession of monomers. Such molecules exhibiting those properties are called “foldamers”.<sup>[2]</sup>

After a short presentation of the secondary structures found in proteins, how it is possible to design fold artificial pseudo-peptides will be reviewed. Finally, the end of this section will be dedicated to the insertion of a catalytic function within such foldamers. An attempt to highlight the relationships existing between the structural behaviours and the pseudo-enzymatic activities will be made.

---

<sup>4</sup> Grauer, A. *et al.*, *Eur. J. Org. Chem.*, **2009**, 5099–5111 (2009).

<sup>5</sup> Qvit, N. *et al.*, *Drug Discov. Today*, **22**, 454–62 (2017).

<sup>6</sup> Gante, J., *Angew. Chem. Int. Ed. Engl.*, **33**, 1699–1720 (1994).

<sup>7</sup> Liskamp, R. M. J. *et al.*, *ChemBioChem*, **12**, 1626–53 (2011).

## I. SECONDARY STRUCTURES ASSOCIATED WITH PROTEINS

### DIHEDRAL ANGLES

On a polypeptide chain, the amino acids are connected by an amide bond (in red, Figure 5). The geometry of this link is partially responsible for the structure of the proteins. Torsion angles are dihedral angles defined by four points in space; here  $\omega$  is defined as the dihedral angle  $C\alpha_i$ -CO-NH- $C\alpha_{i+1}$ . Two conformations for this specific bond are possible: the *cis* conformation, when  $\omega = 0^\circ$  and the *trans* conformation, when  $\omega = 180^\circ$  (Figure 5). Usually, since the lateral chains on the carbon  $\alpha$  are sources of steric hindrance, the *trans* configuration is preferred.

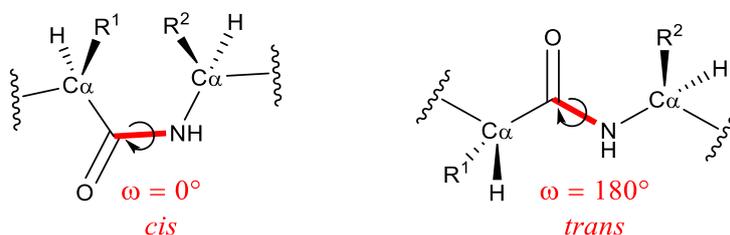


Figure 5: Peptide bond, *cis* and *trans* conformations

In addition, there is a partial double bond delocalized between the N-C-O bonds due to the resonance stabilization, leading to a planar and rigid geometry of the amide bond. Consequently, the length of this bond is shorter than a simple C-N bond (1.33 Å against 1.47 Å) but longer than a C=N bond (1.28 Å).<sup>[8]</sup> In proteins, two torsion angles,  $\phi$  and  $\psi$ , are noteworthy. These describe the rotation of the polypeptide chain around the two bonds on both sides of the carbon  $\alpha$ , as represented in Figure 6, where  $\phi$  defines the torsion angle created by  $C_{i-1}$ - $N_i$ - $C\alpha_i$ - $C_i$  and  $\psi$  by  $N_i$ - $C\alpha_i$ - $C_i$ - $N_{i+1}$ .

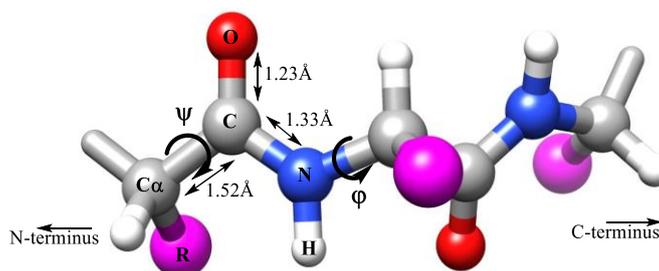


Figure 6: Ball and stick model of a polypeptide chain indicating the distances  $C\alpha$ -C, C=O and C-N and the rotations defined by the torsion angles of  $\phi$  and  $\psi$ . Figure created with UCSF Chimera (University of California)<sup>[9]</sup>

<sup>8</sup> Allen, F. H. *et al.*, *J Chem Soc Perkin Trans 2*, 19 (1987).

<sup>9</sup> Pettersen, E. F. *et al.*, *J. Comput. Chem.*, **25**, 1605–12 (2004).

Because of the rigidity of the amide bond, it is impossible for the angles  $\phi$  and  $\psi$  to take every single value between 0 and 180°. As part of this manuscript, we will only focus on the most common secondary structures allowed by the combination of these dihedral angles. Such possibilities are represented on the Ramachandran plot Figure 7.<sup>[10,11]</sup>

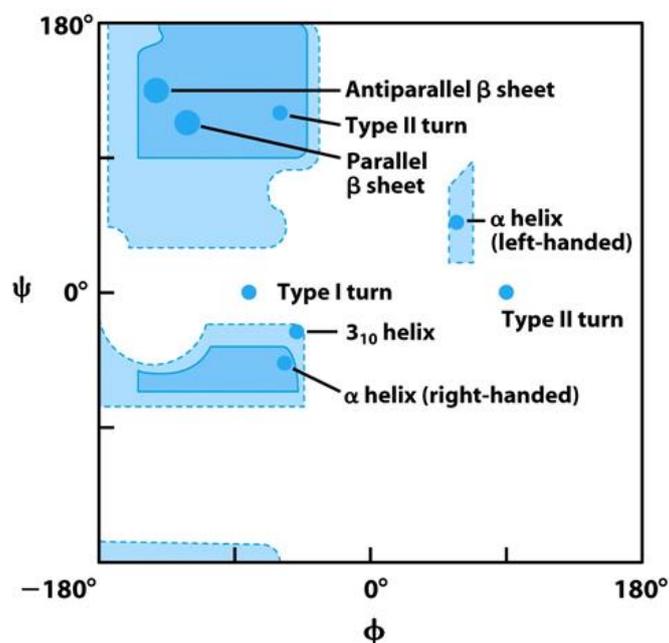


Figure 7: Ramachandran plot<sup>[11]</sup>

The representation in Figure 7 displays recurrent structures such as turns,  $\beta$ -sheets and helices are detected in proteins. Furthermore, several types of helices such as  $\alpha$ -,  $3_{10}$ -,  $\pi$ - or even polyproline helices can be observed in nature, the most common one being the  $\alpha$ -helix (32% of secondary structure of globular proteins).<sup>[12]</sup> The  $3_{10}$ -helices represent only 3.4% of the secondary structure, comparable with polyproline II helices (2 to 4%) while the occurrence of  $\pi$  helices is even lower (less than 1%).<sup>[13,14]</sup>

<sup>10</sup> Hollingsworth, S. A. *et al.*, *Biomol. Concepts*, **1**, 271–83 (2010).

<sup>11</sup> Horton, H. R., *Pearson Prentice Hall* (2003).

<sup>12</sup> Barlow, D. J. *et al.*, *J. Mol. Biol.*, **201**, 601–19 (1988).

<sup>13</sup> Fodje, M. N. *et al.*, *Protein Eng. Des. Sel.*, **15**, 353–58 (2002).

<sup>14</sup> Adzhubei, A. A. *et al.*, *J. Mol. Biol.*, **425**, 2100–2132 (2013).

### 1. *The helical secondary structures*

In 1931, Astbury *et al.* performed X-ray studies of hair and wool structures. The data obtained showed strong differences in the diffraction of hair fibres or moist wool depending on whether it was stretched or unstretched. This study demonstrated that unstretched fibres take the shape of a helix with a vertical distance between two consecutive turns (pitch) of 5.1 Å, that Astbury called the  $\alpha$ -form. Also, the existence of another form was shown, the  $\beta$ -strand, when the fibres are stretched.<sup>[15]</sup> Afterwards, the exploration of intramolecular fold was pursued with the investigation of  $\alpha$ -keratin and  $\alpha$ -myosin.<sup>[16]</sup> Even though all the details provided by those articles were not entirely accurate as Astbury's models involved clashes of atoms,<sup>[17]</sup> the elements of secondary structures, being the  $\alpha$ -helix and the  $\beta$ -strand, were touched upon and gave Pauling, Corey and Branson the opportunity to characterize the  $\alpha$ -helix in 1951.<sup>[18]</sup>

In theory,  $\alpha$ -helices can be either a right-handed screw or a left-handed screw but are usually right-handed.<sup>[19,20]</sup> These are the most abundant structure inherent to proteins, the pitch of the helix is equal to 5.4 Å, each turn is composed of 3.6 amino acids and the rise per residue is of 1.52 Å. A pseudocycle of 13 atoms forms within the helix between the carbonyl oxygen, eleven backbone atoms and the amide hydrogen that closes the loop and gives one complete turn (Figure 8, A). In Bragg's nomenclature, the  $\alpha$ -helix is called a  $3.6_{13}$ -helix, based on its average number of residues per helical turn and the hydrogen-bonded loop size.<sup>[21]</sup> The dihedral angles  $\phi$  and  $\psi$  of the  $\alpha$ -helix are  $-57^\circ$  and  $-47^\circ$  respectively.<sup>[12,22]</sup>

Along the axis of the helix, each carbonyl oxygen (residue  $i$ ) of the polypeptide backbone is hydrogen-bonded to the backbone amide hydrogen, four residues away towards the C-terminus (residue  $i+4$ ). This way, the hydrogen bonds stabilize the structure by being almost parallel to the axis (Figure 8, A). The side chains are located around the cylinder of this helix, pointing outward. The length of a  $\alpha$ -helix in a protein can range from about 4 or 5 residues to more than 40, but the average is  $14 \pm 5$  amino acids.<sup>[23]</sup> Each peptide bond within the helix has a dipole

<sup>15</sup> Astbury, W. T. *et al.*, *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, **230**, 75–101 (1932).

<sup>16</sup> Astbury, W. T. *et al.*, *Nature*, **147**, 696–99 (1941).

<sup>17</sup> Neurath, H., *J. Phys. Chem.*, **44**, 296–305 (1940).

<sup>18</sup> Pauling, L. *et al.*, *Proc. Natl. Acad. Sci.*, **37**, 205–11 (1951).

<sup>19</sup> Kabsch, W. *et al.*, *Biopolymers*, **22**, 2577–2637 (1983).

<sup>20</sup> Haimov, B. *et al.*, *Sci. Rep.*, **6**, 38341 (2016).

<sup>21</sup> Bragg, L. *et al.*, *Proc. R. Soc. Lond. Ser. Math. Phys. Sci.*, **203**, 321–57 (1950).

<sup>22</sup> Arnott, S. *et al.*, *J. Mol. Biol.*, **21**, 371–83 (1966).

<sup>23</sup> Kumar, S. *et al.*, *Biophys. J.*, **75**, 1935–44 (1998).

positioned from the oxygen to the nitrogen. Owing to that, the entire helix is a macrodipole oriented from the  $\delta^-$  C-terminus towards the  $\delta^+$  N-terminus (Figure 8, C).

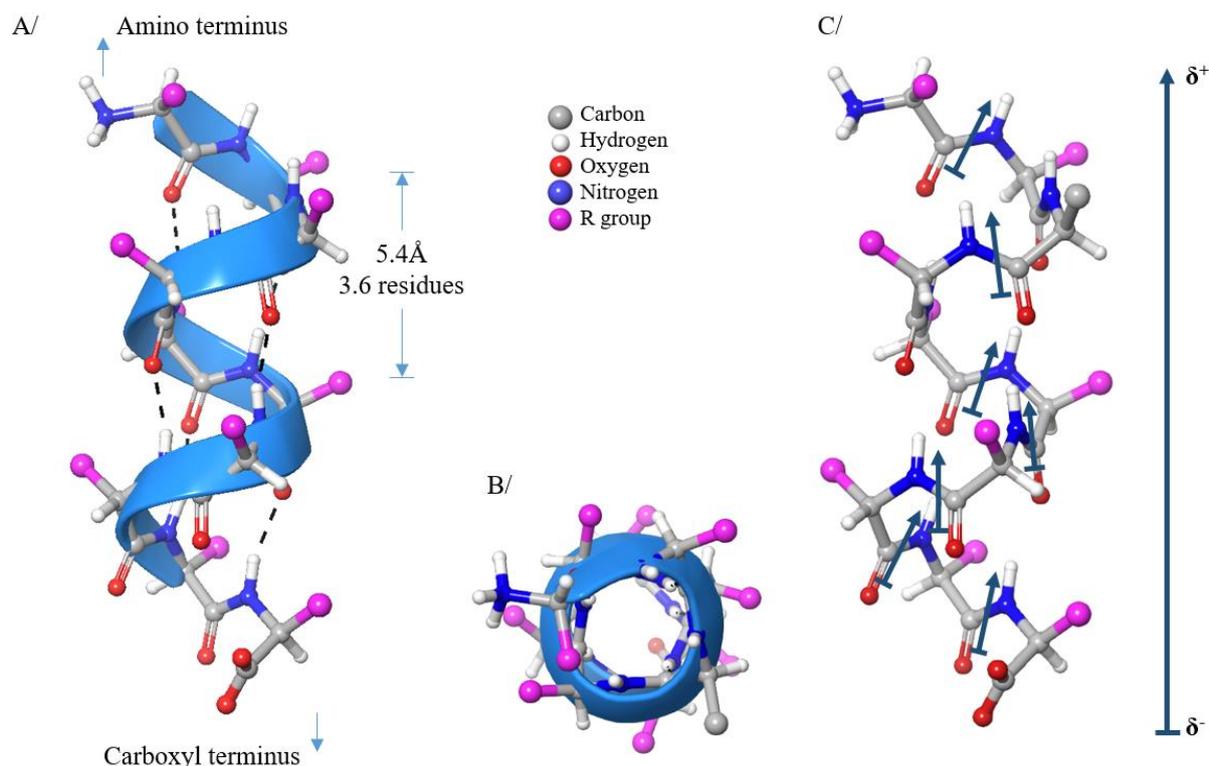


Figure 8: A/Hydrogen bonds stabilizing a right-handed  $\alpha$ -helix. B/An end view showing the coiled backbone as the inside of the helix and the side chains projecting outward. C/Macrodipole oriented to the  $\delta^-$  C-terminus towards the  $\delta^+$  N-terminus. Figures created with Maestro, Schrödinger software<sup>[24]</sup>

As previously reported, left-handed  $\alpha$ -helices ( $\phi$  and  $\psi$  are positive, opposite of the right-handed helix) also exist. However, they are less observed due to the dihedral angle  $\psi$  causing a higher steric hindrance between the carbonyl and the lateral chain.

Regarding other types of helices found in proteins, the  $3_{10}$ -helix differs from the  $\alpha$ -helix as 10 atoms compose a pseudocycle, with a hydrogen bond between  $\text{CO}_i$  and  $\text{NH}_{i+3}$ . The  $3_{10}$ -helix is also less regular, hence its dihedral angle values,  $\phi$  and  $\psi$ , fluctuate. The observed values in proteins are respectively of  $-71^\circ$  and  $-18^\circ$ , giving a very compact helical structure with a small diameter.<sup>[25]</sup>

Less frequent, but nonetheless present, are the  $\pi$ -helices. They are stabilized by a network of hydrogen bonds between  $\text{CO}_i$  and  $\text{NH}_{i+5}$ . These long-distance interactions impose 4.4 amino acids per turn and a pitch of 5.0 Å. This is the largest helix in terms of diameter and is usually

<sup>24</sup> Maestro, Schrödinger, LLC, *Schrödinger Release 2019-1 New-York NY* (2019).

<sup>25</sup> Toniolo, C. *et al.*, *Trends Biochem. Sci.*, **16**, 350–53 (1991).

found within  $\alpha$ -helices. Its dihedral angles  $\phi$  and  $\psi$  are fixed at  $-57^\circ$  and  $-70^\circ$ , respectively. On average, such helices are composed of seven amino acids, usually aromatic amino acids (tryptophan and phenylalanine) for which the lateral chains stabilize the helix by  $\pi$ - $\pi$  stacking.<sup>[26]</sup>

Finally, polyproline helices occur in two different conformation: polyproline-I (PPI) and polyproline-II (PPII). The PPI helix is a right-handed helix composed of all peptide bonds in *cis* conformation, while the PPII helix has all its peptide bonds in *trans* conformation. Only the latter is found in globular proteins.<sup>[27]</sup>

## 2. The $\beta$ -sheets

A  $\beta$ -strand is a long and extended polypeptide chain, while, as previously discussed, the  $\alpha$ -helix is tightly coiled. The distance between adjacent amino acids along the  $\beta$ -strand is of approximately  $3.5 \text{ \AA}$  and the side chains project from the sheet alternating up and down directions.  $\beta$ -strands are usually drawn as wide arrows with the tip of the arrowhead representing the C-terminal end of the polypeptide chain (Figure 9). In proteins, it is quite rare to find single  $\beta$ -strands, as their structure is not very stable, and can be compared to a random coil in terms of their stability.

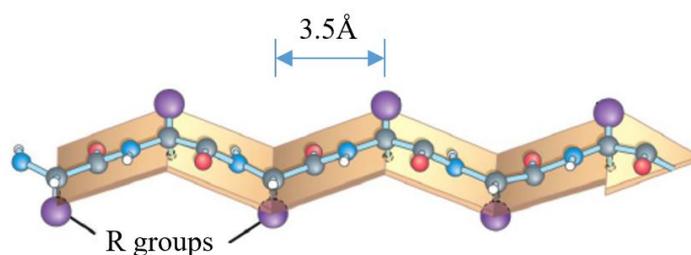


Figure 9: Representative structure of a  $\beta$ -strand<sup>[28]</sup>

However, when two adjacent  $\beta$ -strands line up, they can form hydrogen bonds creating a very stable structure known as a  $\beta$ -sheet. Connected by turns, many strands can assemble in  $\beta$ -sheets. Adjacent chains in a  $\beta$ -sheet can run in opposite directions (antiparallel  $\beta$ -sheet) or in the same direction (parallel  $\beta$ -sheet) (Figure 10).

<sup>26</sup> Cooley, R. B. *et al.*, *J. Mol. Biol.*, **404**, 232–46 (2010).

<sup>27</sup> Kakinoki, S. *et al.*, *Polym. Bull.*, **53**, 109–15 (2005).

<sup>28</sup> Nelson, D. L. *et al.*, 5th Revised edition (New York, NY: W.H.Freeman & Co Ltd, 2008).

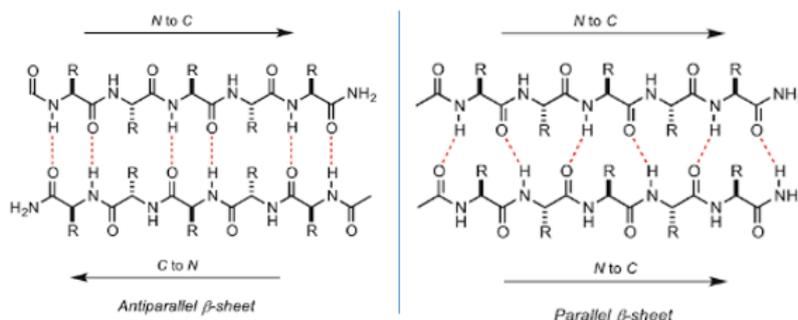


Figure 10: Antiparallel  $\beta$ -sheet versus parallel  $\beta$ -sheet

In the antiparallel arrangement (Figure 11), the NH group and the CO group of each amino acid are respectively hydrogen bonded to the CO and the NH of a partner on the adjacent chain. The dihedral angles of the amino acids in this type of sheet are  $\varphi = -139^\circ$  and  $\psi = +135^\circ$ .<sup>[29]</sup>

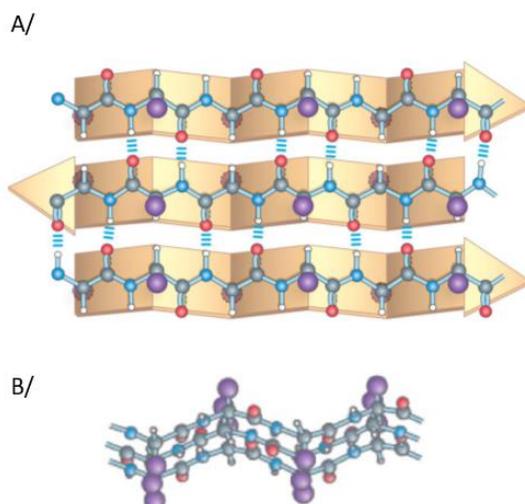


Figure 11: Representative structure of an antiparallel  $\beta$ -sheet. A/Top view. B/Side view<sup>[28]</sup>

In the parallel arrangement (Figure 12), the hydrogen-bonding scheme is slightly more complicated. For each amino acid, their hydrogen amide is bonded to the carbonyl oxygen of one amino acid on the adjacent strand, whereas their carbonyl oxygen is bonded to the hydrogen amide of the amino acid two residues further along the chain of the adjacent strand. Usually, two strands of such structure are connected by a long non-structured polypeptide chain or, sometimes, a  $\alpha$ -helix. The  $\varphi$  and  $\psi$  dihedral angles are then  $-119^\circ$  and  $+113^\circ$  respectively.<sup>[29]</sup>

<sup>29</sup> Lubell, W. D., Topics in Heterocyclic Chemistry (Springer, 2016).

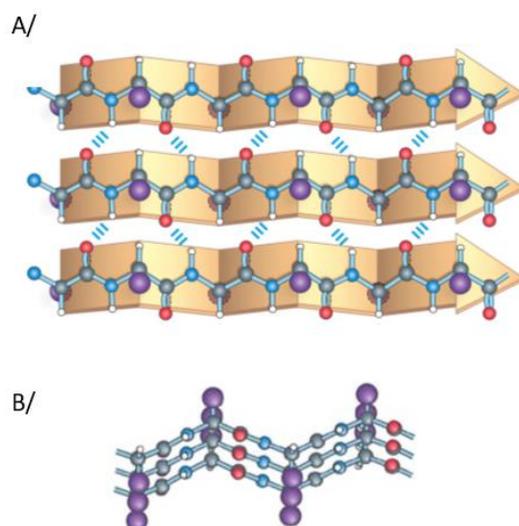


Figure 12: Representative structure of a parallel  $\beta$ -sheet. A/Top view. B/Side view<sup>[28]</sup>

Due to the optimum alignment of the donors and acceptors along the chains and the resulting parallelism given by the antiparallel conformation, it is the most common  $\beta$ -sheet conformation in proteins. The antiparallel  $\beta$ -sheet can be found at the surface of the proteins while, usually, the parallel arrangement tends to be found inside the core structure of proteins where the hydrogen bonds between strands are protected from competition with water molecules.

### 3. The turns

In addition to  $\alpha$ -helices and  $\beta$ -strands, a folded polypeptide chain can contain two other types of secondary structure called loops and turns. They are the simplest and most irregular secondary structures. Loops are not well defined, but generally hold hydrophilic residues found on the surface of the protein. Initially, turns were described by Venkatachalam in 1968, they were designated as sites where the polypeptide chain folds back on itself.<sup>[30]</sup> Nowadays, a widely accepted definition of turns is when loops possess an internal hydrogen bond between the main-chain carbonyl group from the first residue of the loop and the main-chain amide group from the last residue of this loop.<sup>[31]</sup> Several types of turns exist (Table 1), they differ by the number of residues separating the pair connected via the hydrogen bond.<sup>[32]</sup> The  $\delta$  and  $\epsilon$  turns contrast from the other families as they are characterized by a hydrogen bond between the main-chain amide group of the first residue and the main-chain carbonyl group of the last

<sup>30</sup> Venkatachalam, C. M., *Biopolymers*, **6**, 1425–36 (1968).

<sup>31</sup> Chou, K.-C., *Anal. Biochem.*, **286**, 1–16 (2000).

<sup>32</sup> Koch, O. *et al.*, *Proteins Struct. Funct. Bioinforma.*, **74**, 353–67 (2009).

residue (opposite to the  $\gamma$ ,  $\beta$ ,  $\alpha$  and  $\pi$  turns). Therefore, in these conformations, a “reverse” hydrogen bond is observed.

Table 1: Different turn types: number of residues and hydrogen bond pattern

Turn name	Number of residues	H bond type
$\delta$	2	$\text{NH}_i - \text{CO}_{i+1}$
$\varepsilon$	3	$\text{NH}_i - \text{CO}_{i+2}$
$\gamma$	3	$\text{CO}_i - \text{NH}_{i+2}$
$\beta$	4	$\text{CO}_i - \text{NH}_{i+3}$
$\alpha$	5	$\text{CO}_i - \text{NH}_{i+4}$
$\pi$	6	$\text{CO}_i - \text{NH}_{i+5}$

We will focus, in the following paragraphs, on the most common turns, the  $\beta$ -turns and  $\gamma$ -turns.

### i. $\beta$ -turns

The  $\beta$ -turns are composed of four amino acids, their stabilization is based on the establishment of an intramolecular hydrogen bond between the carbonyl oxygen of the residue  $i$  and the amine hydrogen of the residue  $i+3$  (Figure 13). They are called  $\beta$ -turns because they usually connect adjacent  $\beta$ -strands in an antiparallel  $\beta$ -sheet. They generally cause a change in direction of the polypeptide chain allowing it to fold back on itself to create a more compact structure.

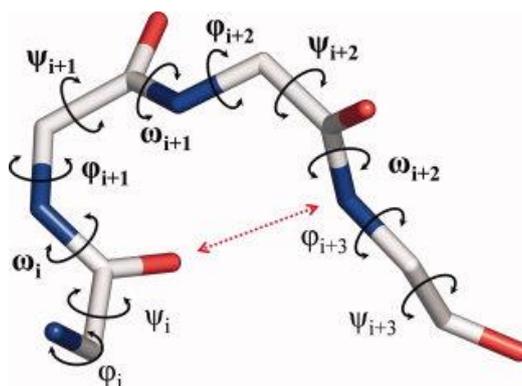


Figure 13:  $\beta$ -turn connected by hydrogen bond between residue  $i$  and  $i+3$

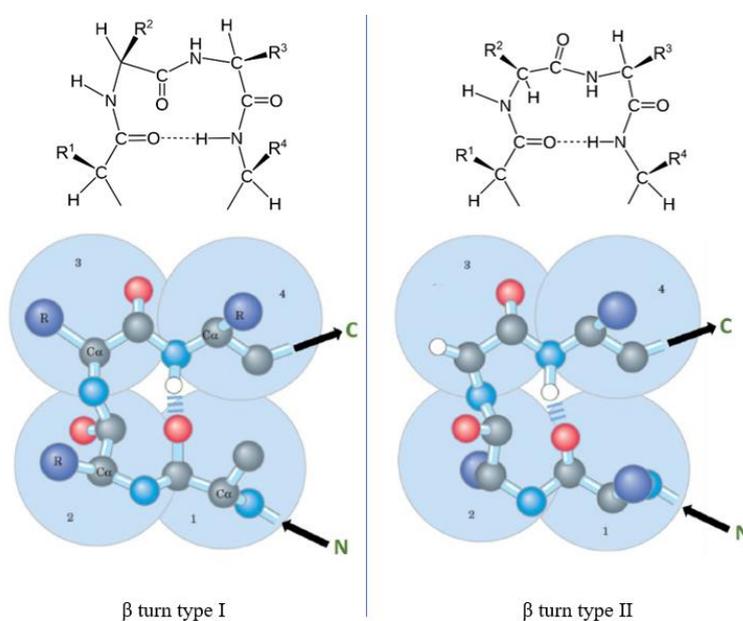
Several families of  $\beta$ -turns exist and were described in 1994 by Hutchinson and Thornton.<sup>[33]</sup> Their classification relies on the dihedral angles of the central residues ( $i+1$  and  $i+2$ ) in the turn. The main types of  $\beta$ -turns are summarized in the following table.

<sup>33</sup> Hutchinson, E. G. *et al.*, *Protein Sci.*, 2207–16 (1994).

Table 2: Dihedral angles of  $\beta$ -turn types

Turn type	Dihedral angles ( $^{\circ}$ )			
	$\Phi_{i+1}$	$\Psi_{i+1}$	$\Phi_{i+2}$	$\Psi_{i+2}$
I	-60	-30	-90	0
I'	+60	+30	+90	0
II	-60	+120	+80	0
II'	+60	-120	-80	0
VIa ( <i>cis</i> )	-60	+120	-90	0
VIb ( <i>cis</i> )	-120	+120	-60	0

Types I and II (Figure 14) are the most common turns found in proteins, types I' and II' are their mirror images and are significantly rarer.

Figure 14:  $\beta$ -turns type I and II

In these types of  $\beta$ -turns, it is quite typical to find proline in the position 2 and glycine as residue number 3. Indeed, in such configuration, the proline adds a high restraint in the structure, which is unconstrained by the following glycine.

## ii. $\gamma$ -turns

As reported in Table 1,  $\gamma$ -turns possess three amino acids and are stabilized by a hydrogen bond between the main chain  $\text{CO}_i$  and  $\text{NH}_{i+2}$ .

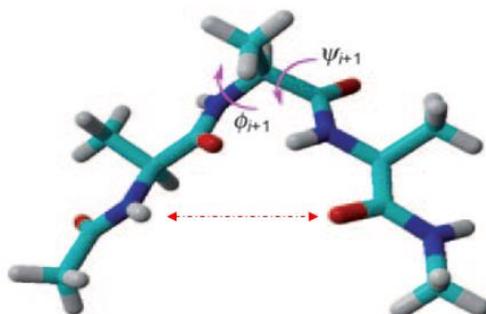


Figure 15:  $\gamma$ -turn involving a hydrogen bond between  $i$  and  $i+2$ <sup>[7]</sup>

In 1972, Matthews described the first example of a  $\gamma$ -turn in a protein at the end of a  $\beta$ -hairpin in thermolysin.<sup>[34,35]</sup> Afterwards, this type of turn was also reported in the early 80s by several groups.<sup>[36,37,38]</sup> These occur as two types: inverse  $\gamma$ -turn and classic  $\gamma$ -turn. However, it is less common to find classic  $\gamma$ -turns in proteins.<sup>[39]</sup> The values of dihedral angles for these two types of  $\gamma$ -turns are reported in the following table.

Table 3: Dihedral angles of  $\gamma$ -turns

	Dihedral angles ( $^{\circ}$ )	
	$\phi_{i+1}$	$\psi_{i+1}$
Classic	+70 to +85	-60 to -70
Inverse	-70 to -85	+60 to +70

Finally, it is quite frequent to find proline in the central position of the turn (Figure 16, A). It is to be noted that the intramolecular hydrogen bond is only possible if the proline adopts a *cis* conformation as represented in Figure 16, B.

<sup>34</sup> Matthews, B. W., *Macromolecules*, **5**, 818–19 (1972).

<sup>35</sup> Némethy, G. *et al.*, *Macromolecules*, **5**, 755–58 (1972).

<sup>36</sup> Smith, J. A. *et al.*, *Crit. Rev. Biochem.*, **8**, 315–99 (1980).

<sup>37</sup> Valle, G. *et al.*, *J Chem Soc Perkin Trans 2*, 885–89 (1986).

<sup>38</sup> Rose, G. D. *et al.*, *Adv. Protein Chem.*, **37**, 1–109 (1985).

<sup>39</sup> Milner-White, E. J. *et al.*, *J. Mol. Biol.*, **204**, 777–82 (1988).

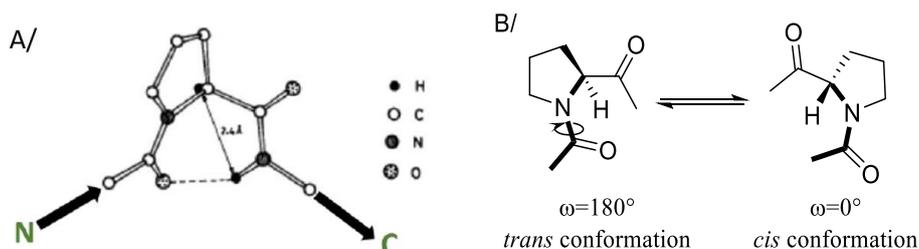


Figure 16: A/Isomerism of the proline's amide link. B/Representative structure of a  $\gamma$ -turn with proline in position 2

#### 4. Secondary structure predictions for the natural amino acids in protein conformation

As previously presented, proteins can adopt different secondary structures such as helices,  $\beta$ -sheets or turns. These structures are amino acid dependent, meaning that each amino acid tends to direct towards one preferential structure. A statistic rank was proposed by Chou and Fasman in 1974 for the twenty natural amino acids, in regards of the three most common secondary structures ( $\alpha$ -helices,  $\beta$ -sheets and turns).<sup>[40]</sup> In this study, fifteen proteins with known conformations were crystallized and the theoretical data was compared with the crystalline data. Using this procedure, the overall accuracy of prediction of the secondary structure of globular proteins was about 80%.<sup>[41]</sup> In 1987, Williams *et al.* proposed a more precise model that is currently used as a reference in terms of structure prediction.<sup>[42]</sup> These statistical data are presented in the Table 4. If the preferential score is higher than one, then this amino acid will likely be found within the corresponding structure.

Table 4: Conformational preferences of the amino acids

Amino acid	Glu	Ala	Leu	Met	Gln	Lys	Arg	His
$\alpha$ -helix	<b>1.59</b>	<b>1.41</b>	<b>1.34</b>	<b>1.30</b>	<b>1.27</b>	<b>1.23</b>	<b>1.21</b>	<b>1.05</b>
$\beta$ -sheet	0.52	0.72	1.22	1.14	0.98	0.69	0.84	0.80
turn	1.01	0.82	0.57	0.52	0.84	1.07	0.90	0.81
Amino acid	Val	Ile	Tyr	Cys	Trp	Phe	Thr	
$\alpha$ -helix	0.98	1.09	0.74	0.66	1.02	1.16	0.76	
$\beta$ -sheet	<b>1.87</b>	<b>1.67</b>	<b>1.45</b>	<b>1.40</b>	<b>1.35</b>	<b>1.33</b>	<b>1.17</b>	
turn	0.41	0.47	0.76	0.54	0.65	0.59	0.90	

<sup>40</sup> Chou, P. Y. *et al.*, *Biochemistry*, **13**, 211–22 (1974).

<sup>41</sup> Chou, P. Y. *et al.*, *Biochemistry*, **13**, 222–45 (1974).

<sup>42</sup> Williams, R. W. *et al.*, *Biochim. Biophys. Acta*, **916**, 200–204 (1987).

Amino acid	Gly	Asn	Pro	Ser	Asp
$\alpha$ -helix	0.43	0.76	0.34	0.57	0.99
$\beta$ -sheet	0.58	0.48	0.31	0.96	0.39
turn	<b>1.77</b>	<b>1.34</b>	<b>1.32</b>	<b>1.22</b>	<b>1.24</b>

Among the natural amino acids, two of them are identified as “helix breakers”: glycine and proline. The first is highly flexible and, therefore, disrupts the regularity of the  $\alpha$ -helical backbone conformation, while the pyrrolidine ring of proline is too rigid and causes an important steric strain. Conversely, those two amino acids are commonly found in turns.<sup>[43]</sup> Some others, such as methionine, alanine, leucine, glutamate and lysine, prefer to adopt helical conformations. Various amino acids like isoleucine, valine, threonine or large aromatic residues (tryptophan, tyrosine and phenylalanine), presumptively adopt  $\beta$ -strand conformations. Therefore, it appears that the structuration properties depend on the primary structure of the polypeptide chains, due to specific values on the dihedral angles  $\phi$  and  $\psi$  and interactions between lateral chains along the main-chains. Nonetheless, even if the sequence itself provides an indication of which amino acid could disrupt a structure or induce another one, these are not entirely reliable in predicting with exactitude the secondary structure.

In conclusion, structural complexity holds small, simple and randomly structured peptides as much as turns, helices and sheets. Structures with assemblies of helices, sheets and turns are also quite common in nature. The number of possibilities for the creation of proteins with these twenty amino acids is immense. Moreover, each of the amino acids carries some specific properties due to their side chain, it can then be established that the structural diversity will lead to an extensive variety in biological function. In this respect, it can be alleged that the biochemical functions carried out by the proteins must rely on their capacity to structure themselves.

<sup>43</sup> Imai, K. *et al.*, *Biophysics*, **1**, 55–65 (2005).

## II. FOLDAMERS AND CATALYSIS

### 1. *Definition foldamer*

The term “foldamer” was introduced for the first time by Samuel Gellman in 1996. He described it as “polymer with a strong tendency to adopt a specific compact conformation”.<sup>[2]</sup> In 2001, Moore gave a more specific definition of a foldamer as being “any oligomer that folds into a conformationally ordered state in solution, the structures of which are stabilized by a collection of non-covalent interactions between non-adjacent monomer units”.<sup>[44]</sup> In other words, a foldamer is an artificial oligomer, peptidic or not, for which, in solution, the backbone adopts a well-defined secondary structure, such as turns, helices or sheets. It is important to point out that these molecular objects must adopt only one conformational state.

Moreover, the interactions that provide the foldamer its structure ought not to be contained within adjacent repeating units, they must be immutable with the notion of secondary structure mimic. The conformational stability is required to be the result of the combination of steric hindrance, intrinsic rigidity of the backbone and non-covalent interaction between non-adjacent monomers.<sup>[45]</sup> Hydrogen bonds, ionic interactions or hydrophobic effect are governing the stability of the structure, in the same manner as they stabilize the structure of proteins.

Foldamers can be sorted in two classes: the single-stranded foldamers and the multiple-stranded foldamers (Figure 17). As the second class tends to access molecules presenting a level of structuration of the tertiary structure, we will focus on the following paragraphs on the single-stranded foldamers. Indeed, it is of great interest to create such foldamers as they are designed to organize themselves into defined secondary structures found in nature (helices, turns and sheets).

---

<sup>44</sup> Hill, D. J. *et al.*, *Chem. Rev.*, **101**, 3893–4012 (2001).

<sup>45</sup> Hecht, S. *et al.* (John Wiley & Sons, 2007).

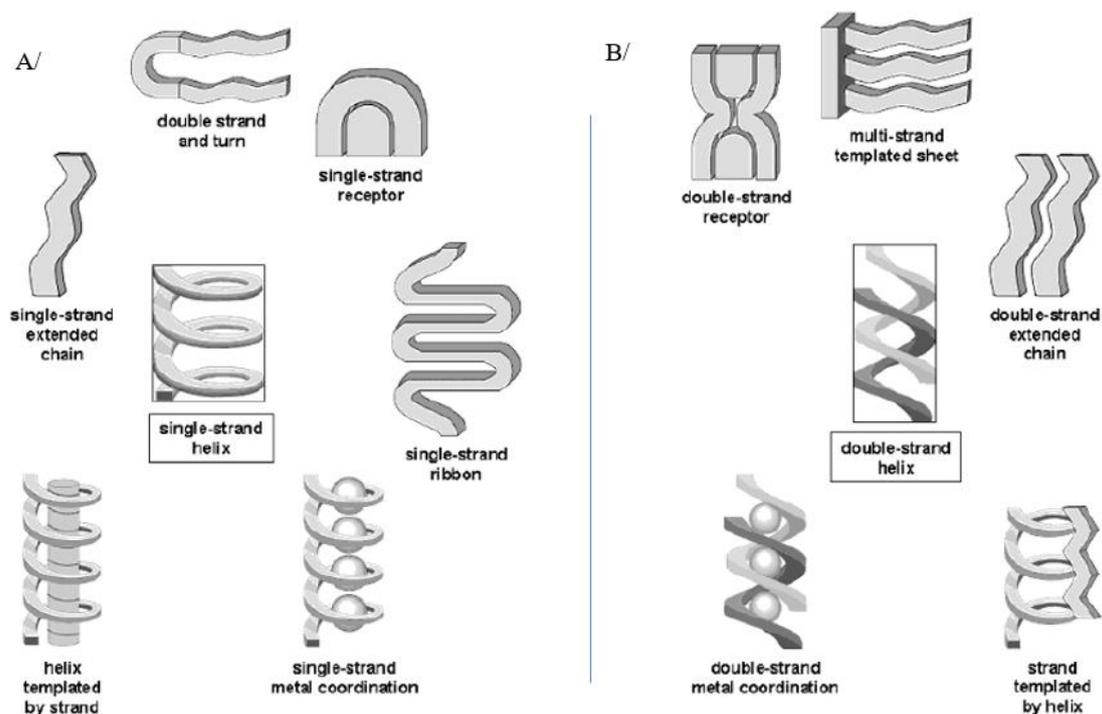


Figure 17: Illustrations of the different types of foldamer secondary structures in A/single-stranded foldamer or in B/multiple-stranded foldamers<sup>[44]</sup>

Foldamer-based chemical systems can be used for various applications if they are functionalized in a suitable manner. For instance, they could be adapted for different biological purposes such as protein-protein interaction inhibitors<sup>[46,47]</sup> or antimicrobial activity<sup>[48,49]</sup>, their exploitation as polyfunctional platforms for synthetic glycocluster construction<sup>[50]</sup> was investigated too. They also present possibilities to be used for further diverse applications like nanoarchitecture<sup>[51]</sup> or catalysis.<sup>[52,53]</sup>

Single-stranded foldamers are composed of various numbers of different foldamer-based molecules.<sup>[44,54]</sup> In this document, we will focus on peptides and pseudopeptides foldamers that can be classified into four separated groups. The artificial oligomers composing these clusters are sorted by the number of atoms separating the CO from the NH within a monomer (Figure 18). Hence, the four families are constituted either with  $\alpha$ -,  $\beta$ -,  $\gamma$ - or  $\delta$ -peptides. For example, the  $\alpha$ -peptides are oligomers created with  $\alpha$ -amino acids connected by an amide bond;  $\beta$ - and

<sup>46</sup> Bartus, É. *et al.*, *ChemistryOpen*, **6**, 236–41 (2017).

<sup>47</sup> Mándity, I. M. *et al.*, *Expert Opin. Drug Discov.*, **10**, 1163–77 (2015).

<sup>48</sup> Schmitt, M. A. *et al.*, *J. Am. Chem. Soc.*, **126**, 6848–49 (2004).

<sup>49</sup> Schmitt, M. A. *et al.*, *J. Am. Chem. Soc.*, **129**, 417–28 (2007).

<sup>50</sup> Simon, M. *et al.*, *Chem. - Eur. J.*, **24**, 11426–32 (2018).

<sup>51</sup> Li, C. *et al.*, *Chem. - Eur. J.*, **13**, 9990–98 (2007).

<sup>52</sup> Maayan, G. *et al.*, *Proc. Natl. Acad. Sci.*, **106**, 13679–84 (2009).

<sup>53</sup> Bécart, D. *et al.*, *J. Am. Chem. Soc.*, **139**, 12524–32 (2017).

<sup>54</sup> Guichard, G. *et al.*, *Chem. Commun.*, **47**, 5933 (2011).

$\gamma$ -peptides are obtained by oligomerization of  $\beta$ - and  $\gamma$ -amino acids, which are homologated proteinogenic amino acid residues, with respectively one and two extra carbons in comparison to the  $\alpha$ -amino acids. Also represented in the family tree of foldamers (Figure 18), we can distinguish subunits, sorted by the functional group classes such as aminoxy acids, ureas, vinylogous peptides and so on.

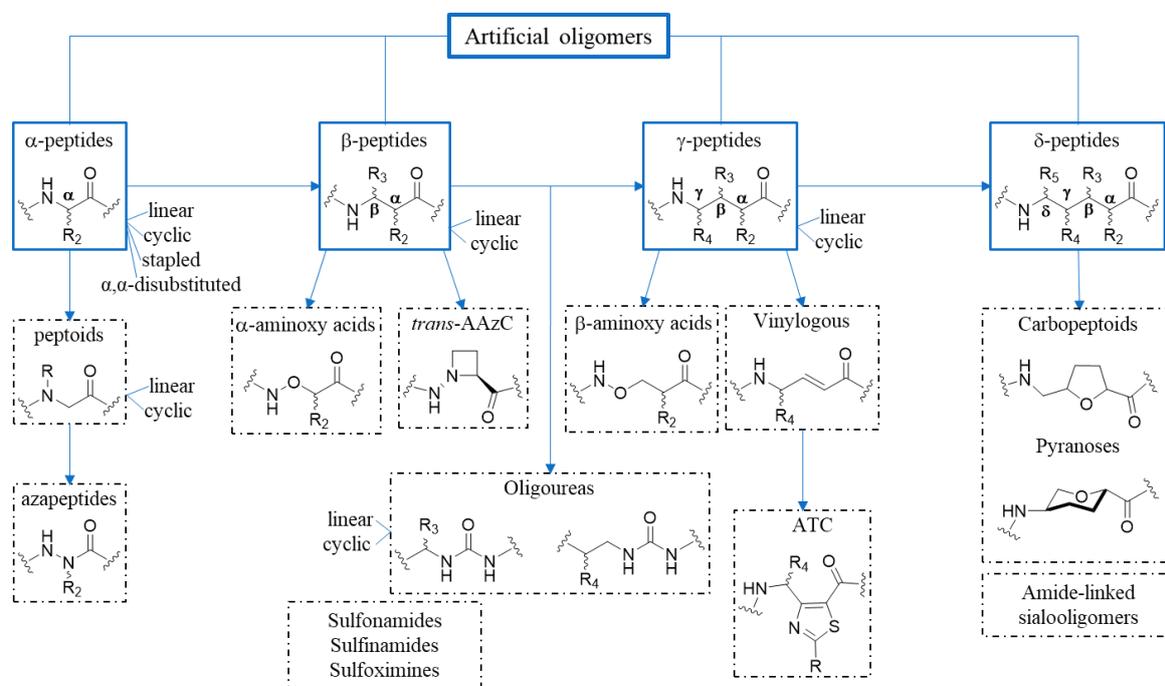


Figure 18: Structural lineage of existing artificial oligomers<sup>[44]</sup>

Historically, the first synthesized non-natural oligomers consisted of the same category of monomer in the succession along the backbone, therefore, they were qualified as homofoldamers.<sup>[2,55]</sup> Afterwards, new combinations between different types of monomers were attempted in order to reach new conformations, this category is known as heterofoldamers.<sup>[56,57]</sup>

In this manuscript, the objective is not to give a complete description of all the possible foldamers and their associated structures. The point here is to give an overview of the most common structures found on the  $\alpha$ -,  $\beta$ - and  $\gamma$ -peptides. Eventually, as the goal of this thesis is the design of bio-inspired catalysts based on a  $\gamma$ -peptide foldamer architecture, we will centre our interest on foldamers bearing specific functions that could lead to a catalytic activity.

<sup>55</sup> Seebach, D. *et al.*, *Helv. Chim. Acta*, **79**, 913–41 (1996).

<sup>56</sup> De Pol, S. *et al.*, *Angew. Chem. Int. Ed.*, **43**, 511–14 (2004).

<sup>57</sup> Hayen, A. *et al.*, *Angew. Chem. Int. Ed.*, **43**, 505–10 (2004).

### i. Considerations from the dynamics of $\alpha$ -peptides and peptoids

As previously discussed,  $\alpha$ -amino acids are proteinogenic amino acid residues. In proteins, peptidic helices mainly correspond to the  $\alpha$ -helices ( $3.6_{13}$ -helices) or the  $3_{10}$ -helices. For most amino acids, the  $\alpha$ -helical geometry ( $\varphi = -57^\circ$ ,  $\psi = -47^\circ$ ) is of lower energy than the  $3_{10}$  geometry ( $\varphi = -71^\circ$ ,  $\psi = -18^\circ$ ).<sup>[12]</sup> However there is no energetic barrier between the two conformations in the Ramachandran plot, and a peptide can therefore drive an easy transition from a  $\alpha$ -helix to a  $3_{10}$ -helix.<sup>[25,58]</sup> Interestingly, when a  $C_{\alpha,\alpha}$ -disubstituted amino acid, such as  $\alpha$ -aminoisobutyric acid (Aib), is incorporated inside the peptide, the formation of a  $3_{10}$  helix tends to be favoured, demonstrating that a slight modification in a peptide sequence may have a great impact on the conformational behaviour.<sup>[59]</sup> When no intramolecular hydrogen bonding network can be formed, for example in sequences consisting solely of proline, the peptide may adopt a new helical conformation called polyproline that could exist as both left- (PPII ( $\varphi$ ,  $\psi$ ,  $\omega$ ) =  $(-75^\circ, +145^\circ, +180^\circ)$ ) and right-handed (PPI ( $\varphi$ ,  $\psi$ ,  $\omega$ ) =  $(-75^\circ, +160^\circ, 0^\circ)$ ) helices.<sup>[60]</sup> Polyproline type II peptides are constituted with only *trans*-configuration and are observed in aqueous conditions. By contrast, the all-*cis* polyproline I is favoured in hydrophobic solvents but has not been found in protein structures. Interestingly, a polyproline II helix can propagate through adjacent non-proline residues. Theoretical and experimental studies demonstrated that water molecules form carbonyl-water-carbonyl H-bond within the chain, which seems to be the driving force to favour the *trans*-configuration.<sup>[61,14]</sup>

By moving the side chain by one atom, from the  $\alpha$ -carbon to the amide nitrogen, the resulting N-substituted glycine becomes a new monomer for the synthesis of  $\alpha$ -peptoids (Figure 19).

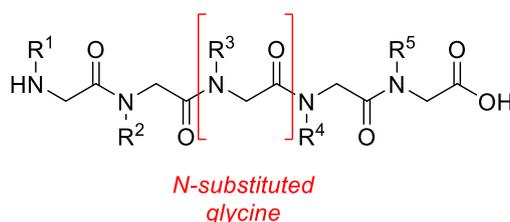


Figure 19: Structure of a generic peptoid oligomer

<sup>58</sup> Bolin, K. A. *et al.*, *Acc. Chem. Res.*, **32**, 1027–33 (1999).

<sup>59</sup> Karle, I. L. *et al.*, *Biochemistry*, **29**, 6747–56 (1990).

<sup>60</sup> Chiang, Y.-C. *et al.*, *Protein Sci. Publ. Protein Soc.*, **18**, 1967–77 (2009).

<sup>61</sup> Sreerama, N. *et al.*, *Proteins Struct. Funct. Genet.*, **36**, 400–406 (1999).

Due to the lack of amide backbone protons, peptoids are expected to be more flexible than their peptide counterparts as it is the inter- and intra-chain hydrogen bonding involving the amide proton in peptides that enables the formation of  $\alpha$ -helices,  $\beta$ -sheets and so on. However, despite a low conformational stability of peptoids, many types of secondary structures such as turns (Figure 20, C), sheets (Figure 20, E) or even ribbons (Figure 20, F) were obtained.<sup>[62]</sup> Helices are probably the most interesting type of folding. The monomers constituting peptoids do no longer possess a chiral centre on the  $\alpha$ -carbon nor do they still have the hydrogen bond donor hold by the NH on a proteinogenic amino acid making the  $\alpha$ - and  $3_{10}$ -helices inaccessible. The only accessible helical structure is analogous to the polyproline type I like helices (Figure 20 A, B and D)<sup>[62]</sup> for which the electronic interactions backbone to side chains ( $n \rightarrow \pi^*$ ) and the side chains bulkiness are the major driving force governing the folding.

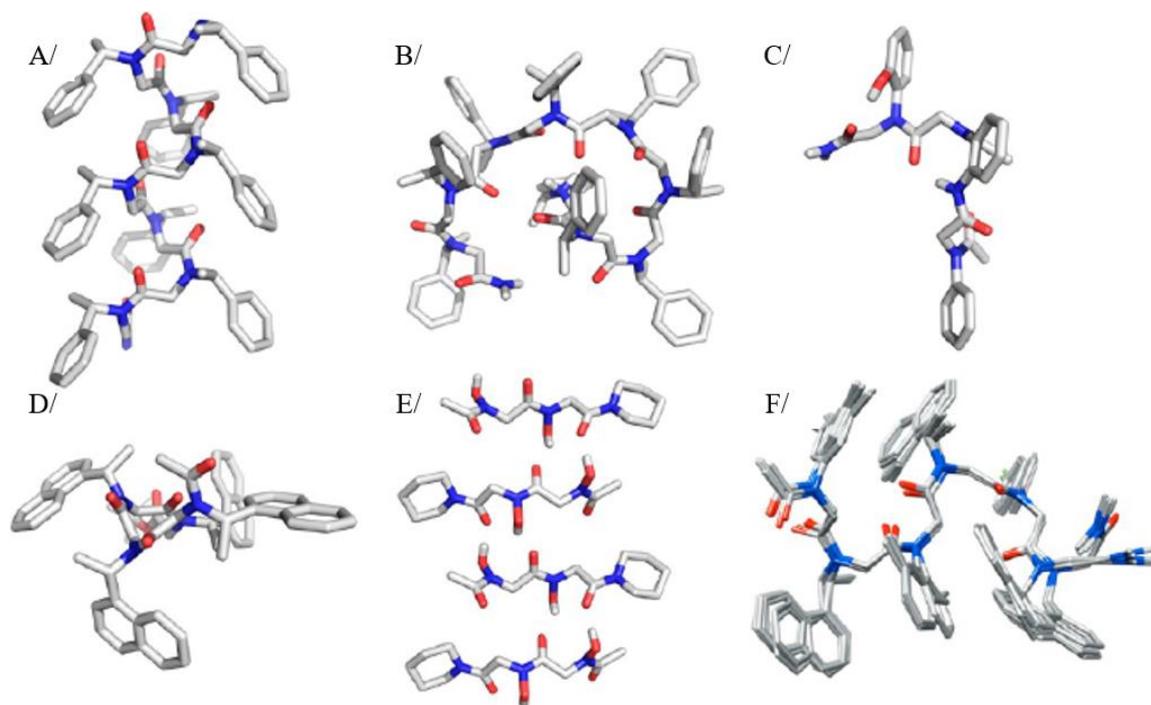


Figure 20: Peptoid helix from a homopolymer of *N*-(*S*)-(1-phenylethyl)glycine A/predicted by molecular mechanics, B/determined by solution-phase 2D NMR. C/*N*-aryl peptoid reverse turn. D/*N*-(*S*)-(1-naphthylethyl)glycine based peptoid helix looking down along the helix axis. E/*N*-hydroxyamide sheet-like structures determined by X-ray crystallography. F/alternating *N*-aryl/*N*-1-naphthylethyl peptoid ribbon structure<sup>[62]</sup>

All these examples are highlighting that the folding of an oligomer into a specific conformation is a multi-parametric problem. The free energy landscape is the result of repeating long-range intramolecular forces such as regular H-bond network or electronic interactions, local

<sup>62</sup> Sun, J. *et al.*, *ACS Nano*, **7**, 4715–32 (2013).

conformational restriction, as in Aib-containing peptides, and in some cases of intermolecular interactions with the environment. Over the last two decades, chemists have tended to combine all these approaches to control the folding behaviour of abiotic oligomers.

## ii. $\beta$ -peptides

$\beta$ -peptides have probably been the most extensively studied foldamer series. They result from the oligomerization of  $\beta$ -amino acids. Those residues are homologues of proteinogenic amino acids. This adds to the molecule an extra torsion angle  $\theta_i$  between the  $\alpha$ - and  $\beta$ -carbons.<sup>[63]</sup> Due to this new angle, the angle  $\varphi_i$  is found between the  $C_\beta$  and the NH, while  $\psi_i$  is the corresponding angle for the  $C_\alpha$ -CO bond (Figure 21).

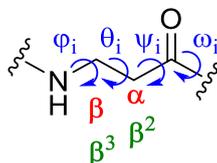


Figure 21: Nomenclature of carbons and dihedral angles on a  $\beta$ -peptide

Noticeably, the  $\alpha$ - and  $\beta$ -carbons can be the starting point for functionalization. When the substituent is branched on the  $\alpha$ -carbon relatively to the acid group, the residue is typically called  $\beta^2$ -amino acid and the resulting oligomer is a  $\beta^2$ -peptide. On the other hand, for  $\beta^3$ -peptides, the lateral chain is on the  $\beta$ -carbon relative to the acid group. Di-substitutions are also possible (Figure 22).

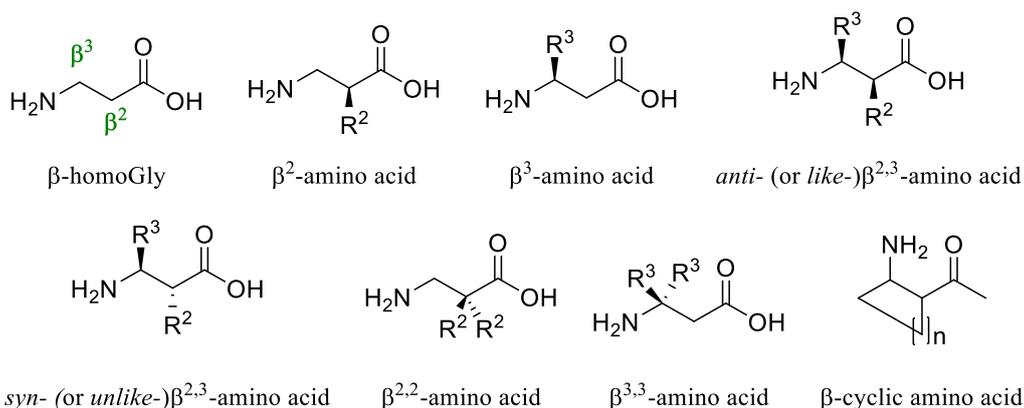


Figure 22: Possible substitutions for the  $\beta$ -amino acids

<sup>63</sup> Banerjee, A. *et al.*, *Curr. Sci.*, **73**, 1067–77 (1997).

In the  $\beta$ -peptide family, the most common structures found are helical. Many groups have been working to describe them, resulting in various nomenclature to characterize the structures.<sup>[55,64,65,66]</sup> In this manuscript, we will use Gellman's nomenclature, which clearly indicates the number of atoms involved in a hydrogen-bonded pseudocycle, from the carbonyls towards the NH amide groups, in the C-terminal direction.<sup>[66]</sup> For instance, if the pseudocycle is composed of 14 atoms, the helix is called "14-helix", "helix 14" or "C<sub>14</sub>-helix". This nomenclature is close to Bragg's nomenclature in which the number of atoms composing the hydrogen-bonded ring is indicated after the number of residues per turn, it would be "3<sub>14</sub>" for a pseudocycle of 14 atoms, as observed for 3<sub>6</sub> and 3<sub>10</sub>-helices in proteins.<sup>[21]</sup> To Gellman's nomenclature, we will take the liberty to add, for clarity, the sense of rotation of the helix (right or left (respectively (P) or (M) in Seebach's nomenclature)).

The prevalent structures accessible to  $\beta$ -peptides are the helices 14, 12, 12/10, 10 and 8 (Figure 23).

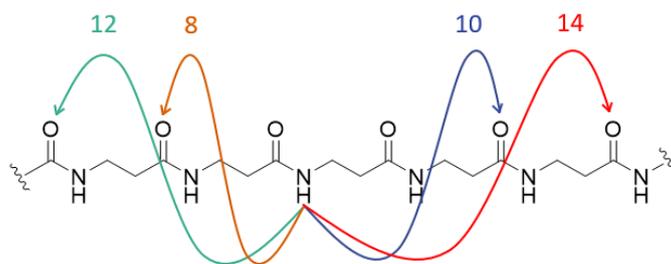


Figure 23: Most common hydrogen bonds in a  $\beta$ -peptide<sup>[67]</sup>

#### ii.a. 14-helix

The first helical structure of  $\beta$ -peptides was simultaneously described by Gellman and Seebach in 1996. Seebach's team was working on the study of the hexa- $\beta$ -peptide H-( $\beta_3$ -HVal- $\beta_3$ -HAla- $\beta_3$ -HLeu)<sub>2</sub>-OH (**1**) (Figure 24, A).<sup>[68]</sup> By performing <sup>1</sup>H NMR studies in deuterated pyridine and methanol, they were able to confirm that it would adopt a 14-helix structure with dihedral angles  $\phi = -134.3^\circ$ ,  $\theta = +60^\circ$ ,  $\psi = -139.9^\circ$  and  $\omega = +102^\circ$  (Figure 24, B).<sup>[67]</sup> This structure is characterized by the establishment of intramolecular hydrogen bonds between the NH in position  $i$  and the CO located on the residue  $i+2$ , leading to a 14 atoms pseudocycle. Due to the alignment of the hydrogen bonds along the main-chain, the net dipole of the helix is established

<sup>64</sup> Bode, K. A. *et al.*, *Macromolecules*, **30**, 2144–50 (1997).

<sup>65</sup> Bella, J. *et al.*, *Macromolecules*, **25**, 5225–30 (1992).

<sup>66</sup> Appella, D. H. *et al.*, *J. Am. Chem. Soc.*, **118**, 13071–72 (1996).

<sup>67</sup> Cheng, R. P. *et al.*, *Chem. Rev.*, **101**, 3219–32 (2001).

<sup>68</sup> Seebach, D. *et al.*, *Helv. Chim. Acta*, **79**, 2043–66 (1996).

in an opposite direction to that of the  $\alpha$ -helix. As represented in Figure 24 C, three residues are necessary to complete a full turn and the pitch is of 4.8 Å. The side chains are distributed around the helix every 120°, positioning them directly above one another every third residue along one face of the helix, giving three distinct sides on the foldamer (Figure 24, D). Only four  $\beta^3$ -amino acids are necessary to lead to such structure but the 14-helix conformation is generally much less stable in water than in organic solvents.

The increase of the 14-helix stability was finally obtained by using the 6-member cyclic *trans*-2-aminocyclohexanecarboxylic acid, so called *trans*-ACHC  $\beta^{2,3}$ -amino acid (**2**) (Figure 24, E), for which the constrained dihedral angles prefigure the oligomer folding.<sup>[69]</sup>

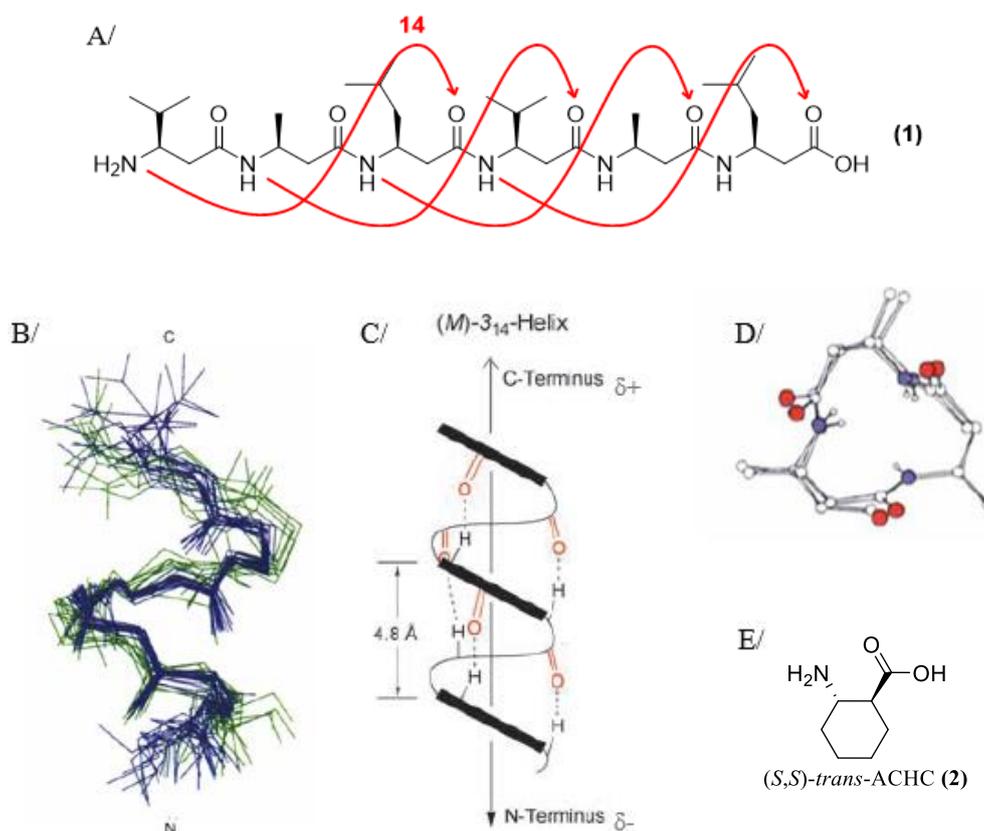


Figure 24: A/Hexa- $\beta$ -peptide (**1**) stabilized by hydrogen bonds. B/Superimposition of the conformations of (**1**) in methanol (green) and in pyridine (blue). C/Representative structure of the 14-helix. D/View N $\rightarrow$ C of the left-handed 14-helix of (**1**) in pyridine solution.<sup>[68,70]</sup> E/Monomer of *trans*-ACHC (**2**)

Afterwards, several other  $\beta$ -peptides folding into 14-helix were described and depending on the stereochemistry of the  $\beta$ -amino acids, the resulting helix is either right or left-handed. Usually,

<sup>69</sup> Appella, D. H. *et al.*, *Nature*, **387**, 381–84 (1997).

<sup>70</sup> Seebach, D. *et al.*, *Helv. Chim. Acta*, **88**, 1969–82 (2005).

when synthesized with  $\beta^3$ -amino acid derived from naturally occurring L-amino acids, the oligomer adopts a left-handed (M) helix.<sup>[67]</sup>

### ii.b. 12-helix

Another type of stable secondary structure, the 12-helix, was discovered in 1997 by Gellman *et al.* who used a new pre-organized cyclic  $\beta^{2,3}$ -amino acid (**3**) amino acid named *trans*-2-aminocyclopentanecarboxylic acid or *trans*-ACPC (Figure 25, A). The cyclopentane ring restrains the value of  $\theta$  at about  $-95^\circ$ , appearing as favouring the 12-helix. The other dihedral angle values correspond to  $\varphi = +100^\circ$ ,  $\psi = +104^\circ$  and  $\omega = -182^\circ$ .<sup>[71]</sup> This result is of prime importance since, for the first time, it demonstrated the possibility of changing the conformational behaviour of  $\beta$ -peptides by altering the cyclic constraint.

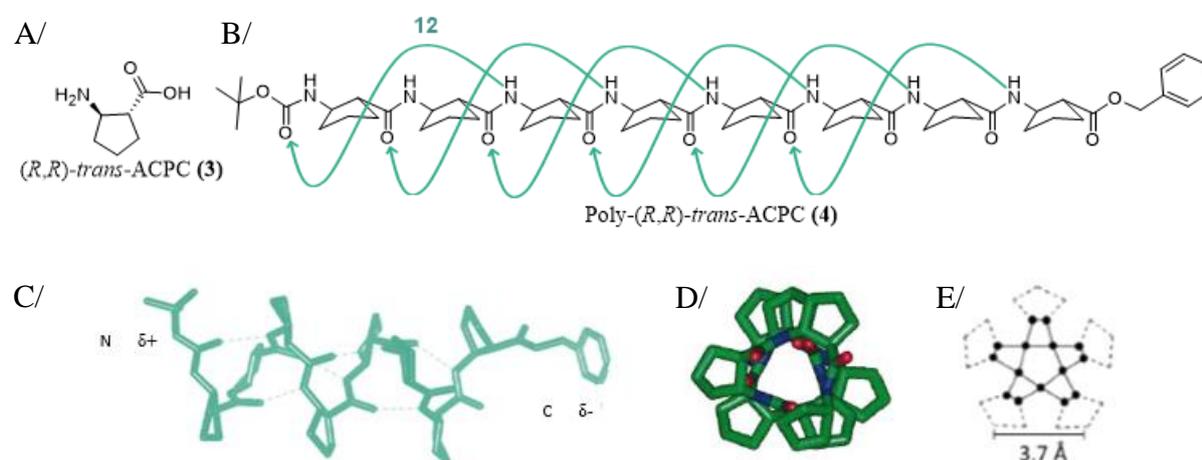


Figure 25: A/Monomer of *trans*-ACPC (**3**) B/ $\beta$ -peptide of (*R,R*)-*trans*-ACPC (**4**) stabilized by hydrogen bonds  $i, i+3$  C=O  $\cdots$  H-N, C/Crystallographic structure of (**4**). D/Representation of the monomers' repartition around the helical axis. E/Axial view N $\rightarrow$ C of (**4**)<sup>[67,72]</sup>

The 12-helix is a left-handed (M) helix stabilized by interwoven twelve-membered-ring hydrogen bonds between the CO of the residue  $i$  and NH of the residue  $i+3$  (Figure 25, B).<sup>[67,72]</sup> It is characterized by a pitch of  $5.6 \text{ \AA}$  and 2.5 residues per turn. The orientation of the dipole is the same as for the  $\alpha$ -helix (Figure 25, C, D and E).

Even though most of the publications found in the literature were describing the 12-helix structure via sequences synthesized using five-membered-ring residues, other groups have been able to reproduce the 12-helical shape using other  $\beta$ -cyclic amino acids. For instance, Aitken *et*

<sup>71</sup> DeGrado, W. F. *et al.*, *J. Pept. Res.*, **54**, 206–17 (2008).

<sup>72</sup> Appella, D. H. *et al.*, *J. Am. Chem. Soc.*, **121**, 7574–81 (1999).

*al.* used cyclobutane residues (**5**) named (*R,R*)-*trans*-ACBC in which the amine and acid groups are in *trans* configuration.<sup>[73]</sup> The angle  $\theta$  is  $-95.7^\circ$  which is close to the one of the canonical 12-helix. Following a similar approach, Martinek *et al.* also obtained such 12-helix employing bicyclic monomers (**6**) (Figure 26).<sup>[74]</sup>

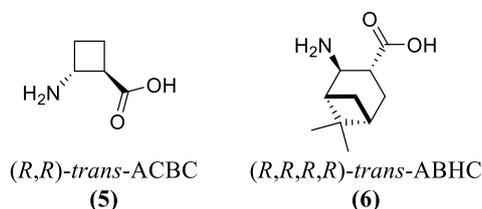


Figure 26: Other  $\beta$ -amino acids folding into 12-helix shape after oligomerization

### ii.c. 12/10-helix

Early after beginning to study the conformational dynamic of  $\beta$ -peptides, Seebach *et al.* demonstrated that unique structures may be obtained by alternating  $\beta^2$ - and  $\beta^3$ -amino acids.<sup>[75]</sup> For instance, heterochiral hybrid  $\beta^3$ -HAxx- $\beta^2$ -HAxx peptides fold forming a 12/10-helix.<sup>[45,76]</sup> The structure results from intramolecular hydrogen bonds alternating in the forward and backward directions along the sequence. The twelve-membered-ring occurs between the NH of the residue  $i$  from a  $\beta^2$ -amino acid and the  $\text{CO}_{i-3}$  of a  $\beta^3$ -amino acid while the ten-membered-ring exists between the  $\beta^2$ -NH $_i$  and the  $\beta^3$ -CO $_{i+1}$  (Figure 27).<sup>[77]</sup> The sense of helicity is right-handed (P) with a pitch of 5.7 Å and 2.7 residues per turn.<sup>[76]</sup> Finally, since the amide C=O bonds point alternately up and down the helix axis, the resulting dipole is strongly reduced in comparison to a 14-helix.<sup>[77]</sup>

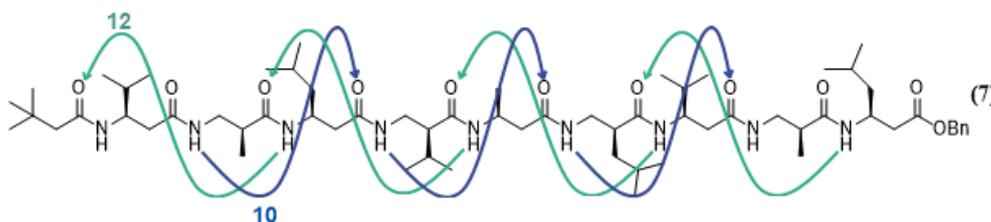


Figure 27:  $\beta$ -peptide of Boc-( $\beta^3$ -HAxx- $\beta^2$ -HAxx) $_n$ -OBn (**7**) stabilized by hydrogen bonds  $i-3,i$  C=O  $\cdots$  H-N and  $i,i+1$  N-H  $\cdots$  C=O<sup>[77]</sup>

<sup>73</sup> Fernandes, C. *et al.*, *Org. Lett.*, **12**, 3606–9 (2010).

<sup>74</sup> Hetényi, A. *et al.*, *Chem Commun*, 177–79 (2009).

<sup>75</sup> Seebach, D. *et al.*, *Helv. Chim. Acta*, **80**, 2033–38 (1997).

<sup>76</sup> Wu, Y.-D. *et al.*, *Acc. Chem. Res.*, **41**, 1418–27 (2008).

<sup>77</sup> Rueping, M. *et al.*, *Helv. Chim. Acta*, **85**, 2577–93 (2002).

## ii.d. 10-helix

In 2001, Fleet *et al.* described the structure of an hexamer of an oxetin  $\beta$ -amino acid derivative (Figure 28).<sup>[78]</sup> By comparison with the (*R,R*)-*trans*-ACBC (**5**), another four-member ring containing  $\beta$ -amino acid, the amine and acid groups in the oxetin residues are in *cis* configuration. Thus, the angle  $\theta$  shifts from  $-95.7^\circ$  to  $-12.1^\circ$  driving the folding towards a new 10-helical conformation in organic solvents.<sup>[79]</sup> The latest value seems somewhat far from the model predicted by Hofmann for an idealized left-handed 10-helix hexapeptide ( $\theta = -59^\circ$ ), but the two other dihedral angles ( $\varphi$  and  $\psi$ ) compensate for this deviation (respectively  $-104.7^\circ$  and  $-101.2^\circ$  in solution versus  $-64^\circ$  and  $-75^\circ$  for the idealized model).<sup>[80]</sup>

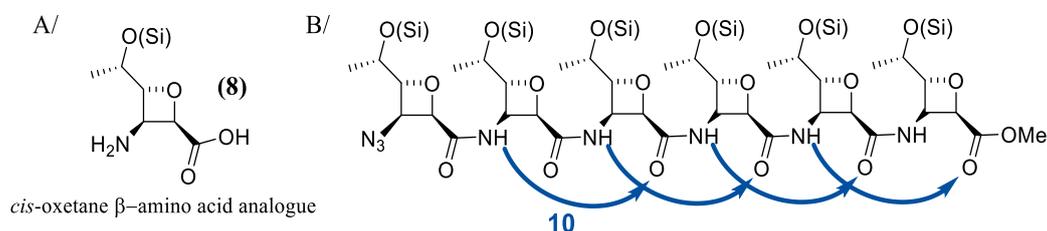


Figure 28: A/Analogue of *cis*-oxetane  $\beta$ -amino acid (**8**). B/ $\beta$ -peptide of the (**8**) stabilized by H-bond  $i, i+1$  N-H  $\cdots$  C=O

## ii.e. 8-helix

Eventually, Seebach *et al.* described, in 2003, the first synthesis of a right-handed 8-helix foldamer.<sup>[81]</sup> The structure was studied in methanol and possessed a pitch of 4.7–4.9 Å with about 2 residues per turn. The eight-membered-ring hydrogen bonds were found to be between the NH group of the amino acid  $i$  and the CO group of the amino acid  $i-2$ . This structure was the result of the oligomerization of *syn*- (or *unlike*)- $\beta^{2,3}$ -amino acids, such as the (*2R,3S*)-3-amino-2-hydroxy acid residues (**9**) (Figure 29). However, if a methyl group replaces the -OH group in position  $\beta^2$  of the building blocks, the 8-helical structure is lost in favour of a pleated-sheet. This can be explained by the presence of a second hydrogen-bonding pattern between hydroxyl groups, borne by the  $\beta^2$ -carbons, and the adjacent carbonyls.<sup>[81]</sup>

<sup>78</sup> Claridge, T. D. W. *et al.*, *Tetrahedron Lett.*, **5** (2001).

<sup>79</sup> Ragab, S. S. *et al.*, *Chem. Commun.*, **54**, 1968–71 (2018).

<sup>80</sup> Günther, R. *et al.*, *J. Phys. Chem. B*, **105**, 5559–67 (2001).

<sup>81</sup> Gademann, K. *et al.*, *Angew. Chem. Int. Ed.*, **42**, 1534–37 (2003).

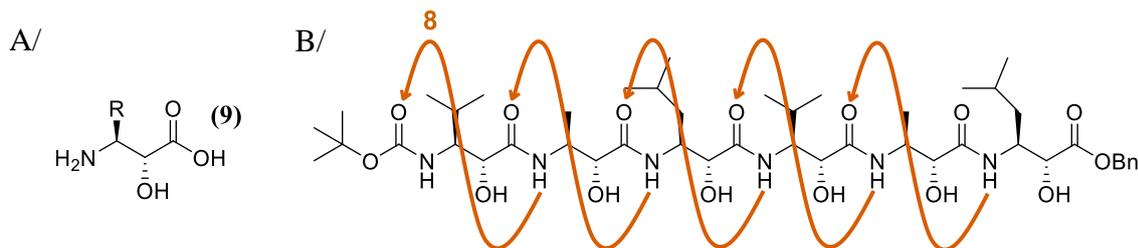


Figure 29: A/(2*R*,3*S*)-3-amino-2-hydroxy acid residue (**9**), B/ $\beta$ -peptide of (**9**) stabilized by hydrogen bonds  $i-2,i$  C=O $\cdots$ H-N

As reported previously, Aitken *et al.* were able to induce 12-helix structure using *trans*-ACBC monomers. This building block can be modified to display a strong tendency to form hydrazine turn by only changing one atom from the ring to create a *N*-aminoazetidone-2-carboxylic acid (AAzC) moiety.<sup>[82]</sup> The hydrazine turn is a split H-bonded 8/5-ring that can behave as an inducer of an 8-helix into a peptide sequence. The first studies were performed by replacing the *trans*-ACBC residue of the N-terminal extremity by a *trans*-AAzC (**10**) (Figure 30, A) to induce the 8-helix, but after six residues, the 12-helix would take over and the 8-helix would not be found.<sup>[83]</sup> This team did further research and, even if the result might be solvent dependant, they eventually depicted the *trans*-AAzC monomer as an 8-helix promoter over the entire length of the peptide, even when placed in the middle of a polypeptide. The octamer being the longest sequence to prove it yet (Figure 30, B).<sup>[84]</sup> Interestingly, the dihedral angles  $\varphi$  ( $\approx +86^\circ$ ) and  $\theta$  ( $\approx -95^\circ$ ) of the *trans*-ACBC residues were similar for the 12- and 8-helices though the dihedral angle  $\psi$  was found to be significantly altered ( $+95^\circ$  for the 12-helix versus  $+35^\circ$  for the 8-helix).<sup>[83]</sup> Finally, the AAzC residue displays a slightly different set of backbone dihedral angles ( $\varphi, \theta, \psi$ )  $\approx (+109^\circ, -98^\circ, +9^\circ)$  without interfering with the 8-helix conformation along the chain.<sup>[84]</sup>

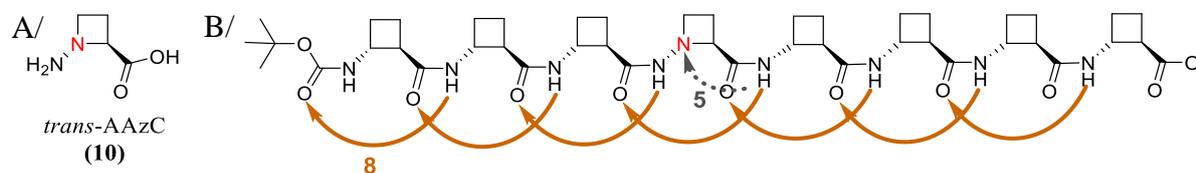


Figure 30: A/*trans*-AAzC residue (**10**), B/Hydrogen bonding network in an octamer possessing one AAzC residue in its centre

<sup>82</sup> Altmayer-Henzi, A. *et al.*, *J. Org. Chem.*, **78**, 6031–39 (2013).

<sup>83</sup> Altmayer-Henzi, A. *et al.*, *Angew. Chem. Int. Ed.*, **54**, 10807–10 (2015).

<sup>84</sup> Declerck, V. *et al.*, *J. Org. Chem.*, **83**, 8793–8800 (2018).

In conclusion, despite a higher degree of freedom than in  $\alpha$ -peptides, oligomers of  $\beta$ -amino acids may adopt stable 3D-structures (Figure 31). Due to numerous NMR, CD and X-ray investigations, reinforced by crystallography, different helical structures were observed in  $\beta$ -peptide series for which topologies are mainly driven either by establishment of regular H-bond networks and by dihedral angle restrictions. In the next sections, we will see that similar concepts are involved in the folding of higher order foldamers such as  $\gamma$ -peptides and related oligomers, including oligoureas.

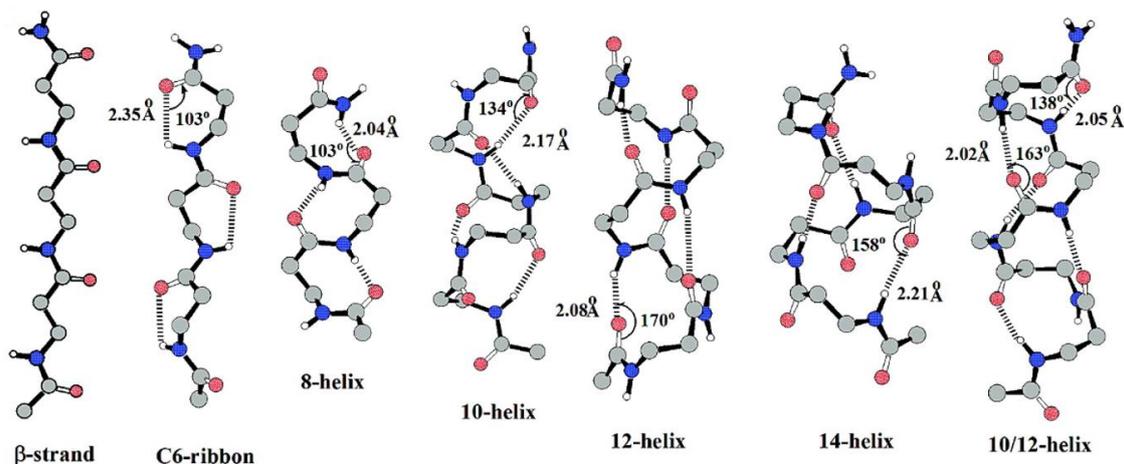


Figure 31: Recap of the possible helix for  $\beta$ -peptide<sup>[76]</sup>

### iii. $\gamma$ -peptides

$\gamma$ -peptides are constituted of  $\gamma$ -amino acids, homologues of  $\beta$ -amino acids, giving a total of three methylene groups between the CO and NH of each monomer. Each of these carbons becomes then a possible starting point for functionalization. Depending on the nomenclature chosen, the position of the substituent is called  $\alpha$ ,  $\beta$  and  $\gamma$ , or  $\gamma^2$ ,  $\gamma^3$  and  $\gamma^4$ . The corresponding dihedral angles are  $\varphi_i$ ,  $\theta_i$ ,  $\zeta_i$  and  $\psi_i$  as represented in Figure 32.

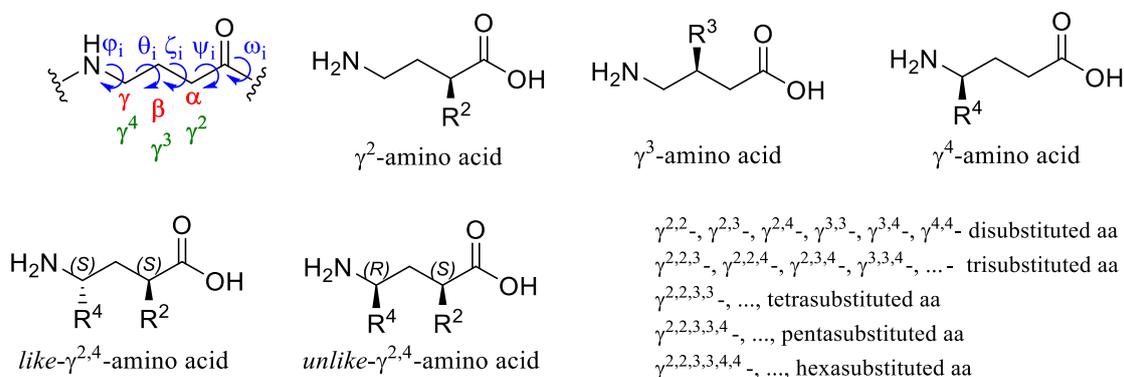


Figure 32: Nomenclature of carbons and dihedral angles on a  $\gamma$ -peptide and examples of its possible substitutions

Because of this third carbon on the main-chain, the  $\gamma$ -peptides are allegedly more flexible than the  $\alpha$ - or  $\beta$ -peptides. Furthermore, the number of hydrogen-bond donors and acceptors, essential for the structuration of the helix, is lower for a  $\gamma$ -peptide than a  $\beta$ - or  $\alpha$ -peptide of the same number of atoms. These two parameters seem unfavourable regarding a possible structuration, but it does not mean that it is impossible, as proved by many researchers who designed folded  $\gamma$ -peptides.

### iii.a. 9- and 14-helices: two structures accessible to acyclic $\gamma$ -peptides

Systematic prediction of the most stable periodic structures in unconstrained  $\gamma$ -peptides was first postulated by Hofmann *et al.* who used an *ab initio* molecular orbital (MO) method.<sup>[85]</sup> They computationally generated oligomers of  $\gamma$ -aminobutyric acid (GABA) by variations of the torsion angles  $\varphi$ ,  $\theta$ ,  $\zeta$  and  $\psi$  of the backbone. This generated 36,000 different conformations which were optimized at the *ab initio* Hartree-Fock MO level of theory (HF/6-31G\*). This study provided various possible structures for the  $\gamma$ -peptides, all of them being helices or even flat ribbons with H-bonded pseudo-rings with atoms ranging from seven up to 24 in number. From these calculations, the 9- and 14-membered-ring helices were recognized as the most stable conformations (Figure 33).

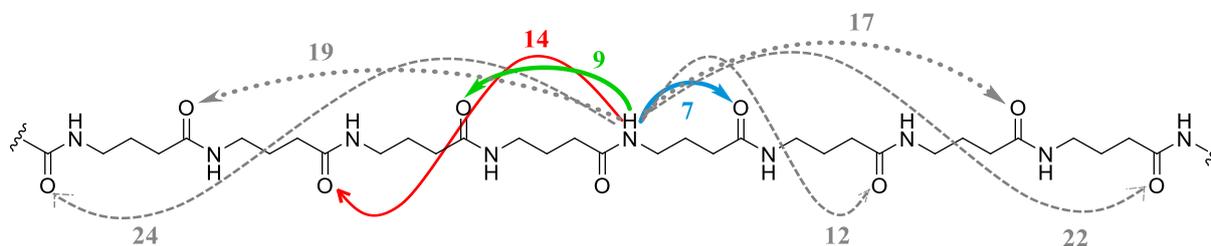


Figure 33: Most common hydrogen bonds in a  $\gamma$ -peptide

The 7-membered-ring structure was also considered as accessible but corresponds to a ribbon due to its small pitch, giving a flat shape of the structure as represented in Figure 34.

<sup>85</sup> Baldauf, C. *et al.*, *Helv. Chim. Acta*, **86**, 2573–88 (2003).

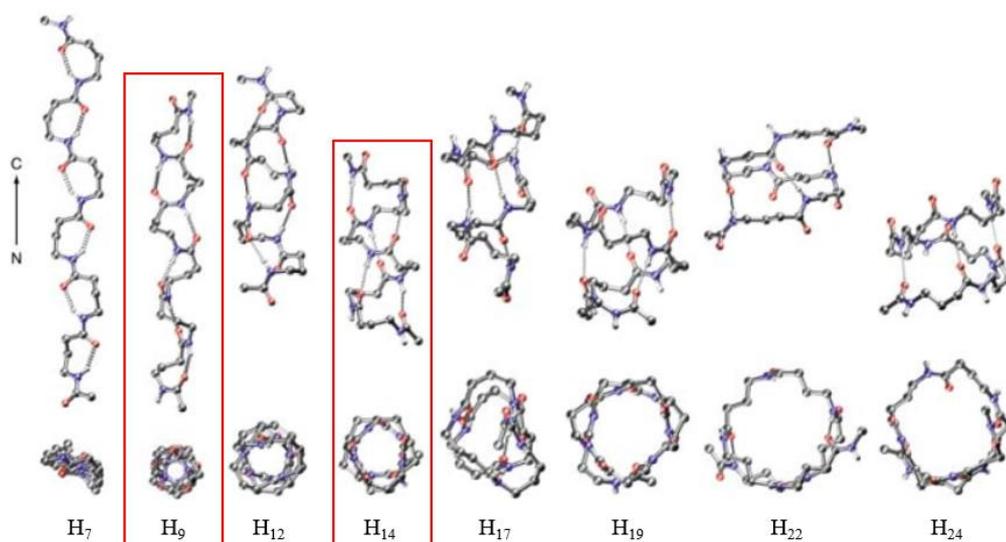


Figure 34: Most stable periodic structures of GABA hexamers obtained after MO theory optimization ( $H_x$  = number of atoms to complete the pseudocycle)<sup>[85]</sup>

In fact, Hanessian *et al.* and Seebach *et al.* independently discovered, in 1998, the 14-helix by oligomerization of  $\gamma^4$ -amino acids.<sup>[86,87]</sup> The folding is characterized by a right-handed helix connecting fourteen atoms by a hydrogen bond between  $NH_i$  and  $CO_{i-3}$ , with a pitch of 5 Å and 2.6 residues per turn; this is why it is also called the 2.6<sub>14</sub>-helix (Figure 35).

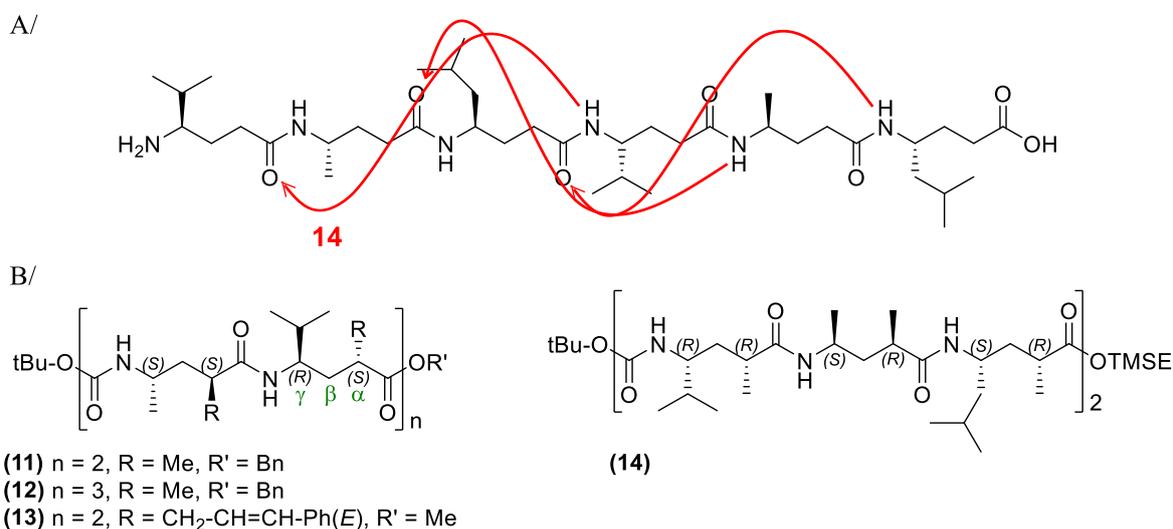


Figure 35: A/H-bonded pattern in Seebach's hexamer<sup>[87]</sup> B/Chemical structure of ( $\alpha S$ )-methylated  $\gamma$ -peptide (11) and (12), ( $\alpha S$ )-cinnamyl  $\gamma$ -peptide (13) and ( $\alpha R$ )-methylated  $\gamma$ -peptide (14)<sup>[86,88]</sup>

Based on NMR experiments performed in pyridine- $d_5$ , NOE correlations were extracted from the ROESY spectra. These data provided distance restraints for 3D-modelling using NMR

<sup>86</sup> Hanessian, S. *et al.*, *J. Am. Chem. Soc.*, **120**, 8569–70 (1998).

<sup>87</sup> Hintermann, T. *et al.*, *Helv. Chim. Acta*, **81**, 983–1002 (1998).

refinement by simulated annealing. The 3D-models obtained showed that, in addition to the H-bond network, hydrophobic interactions between lateral chains projected perpendicular to the helical axis may stabilize the helix.

Disubstituted  $\gamma^{2,4}$ -amino acids were also used for  $\gamma$ -peptide elaboration. This additional substitution reduces the number of accessible conformations for the backbone. Thus, Hanessian *et al.* compared the folding potential of sequences **(11)** to **(14)** having opposite relative configuration.<sup>[86,88]</sup> Only peptides **(11)**–**(13)** adopt a 14-helix formation, ( $\alpha R$ )-methylated analogue **(14)** do not fold. When the 2,4 relative configuration is compatible with the 14-helix secondary structure, a supplementary substitution in position 3 could be supported without disturbing the overall folding.<sup>[89,90]</sup>

Interestingly, for the 14-helix, the macrodipole is orientated as for the  $\alpha$ -helix, from the  $\delta^-$  C-terminus to the  $\delta^+$  N-terminus. Thus, the structure of the  $\alpha$ -helix (13-membered-ring) and of the 14-helices of  $\beta$ - and  $\gamma$ -peptides are quite similar. However, each homologation, from  $\alpha$ - to  $\beta$ - and then from  $\beta$ - to  $\gamma$ -amino acid leads to an inversion of the macrodipole of the helix (Figure 36).<sup>[91]</sup>

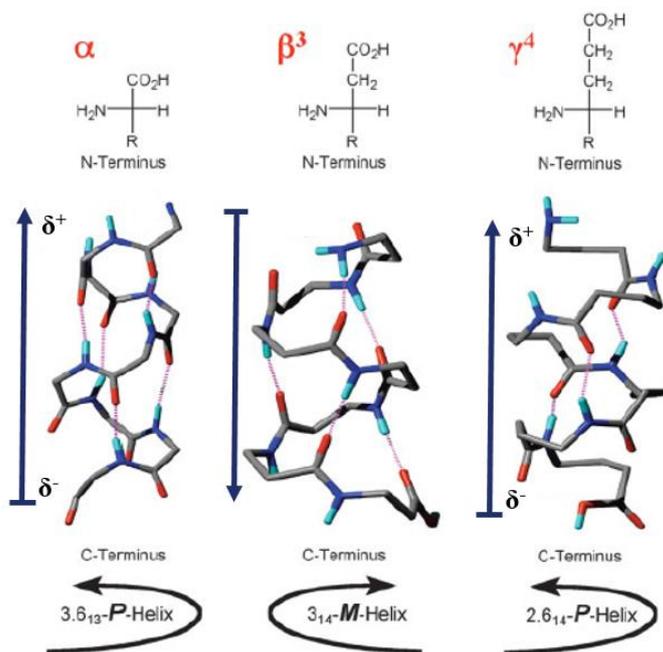


Figure 36: Comparison of  $\alpha$  3.6<sub>13</sub>-,  $\beta$  3<sub>14</sub>- and  $\gamma$  2.6<sub>14</sub>-helices, all amino acids are in L-configuration<sup>[92]</sup>

<sup>88</sup> Hanessian, S. *et al.*, *Tetrahedron Lett.*, **40**, 4925–29 (1999).

<sup>89</sup> Seebach, D. *et al.*, *Chem. Commun.*, 207–8 (2001).

<sup>90</sup> Seebach, D. *et al.*, *Chem. – Eur. J.*, **8**, 573–84 (2002).

<sup>91</sup> Seebach, D. *et al.*, *Chem. Biodivers.*, **1**, 1111–1239 (2004).

<sup>92</sup> Seebach, D. *et al.*, *Pept. Sci.*, **84**, 23–37 (2005).

The second most stable helical structure given by Hofmann's calculations is the 9-helix which was characterized in 2005 by Balaram *et al.*, starting from an oligomer of gabapentin, the Boc-Gpn-Gpn-Gpn-Gpn-NHMe (**15**) (Figure 37).<sup>[93]</sup> The disubstitution of the central  $\beta$ -carbon (or  $\gamma^3$ -) limits the rotation around the  $C_\alpha$ - $C_\beta$  (angle  $\theta$ ) and  $C_\beta$ - $C_\gamma$  (angle  $\zeta$ ) to two *gauche* conformations, where  $\theta \approx \zeta \approx \pm 60^\circ$ . The H-bond pseudocycle is formed between  $CO_{i-2}$  and  $NH_i$ .

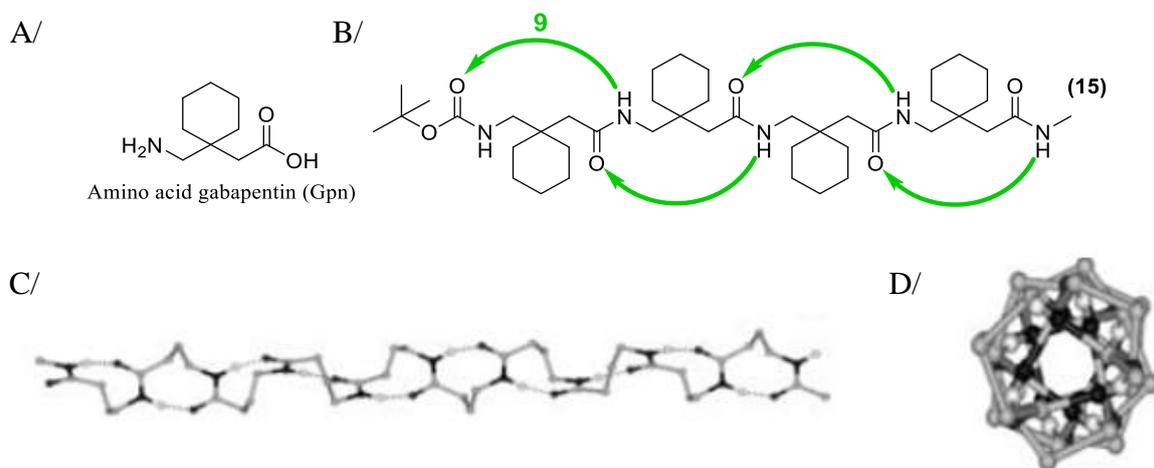


Figure 37: A/Amino acid gabapentin, B/Oligomer (**15**) stabilized by hydrogen bonds  $i-2, i C=O \cdots H-N$ , C/9-helix formed by (**15**), D/view of (**15**) down the axis<sup>[93]</sup>

After this description, Kunwar *et al.* also reported  $\gamma$ -peptides alternating carbo- $\gamma^4$ -amino acid ( $\gamma$ -Caa) (**16**) and  $\gamma$ -aminobutyric acid (GABA) as able to fold into 9-helices.<sup>[94]</sup> The 9-helix is found after only four residues and each turn is completed with 2.3 monomers (Figure 38).

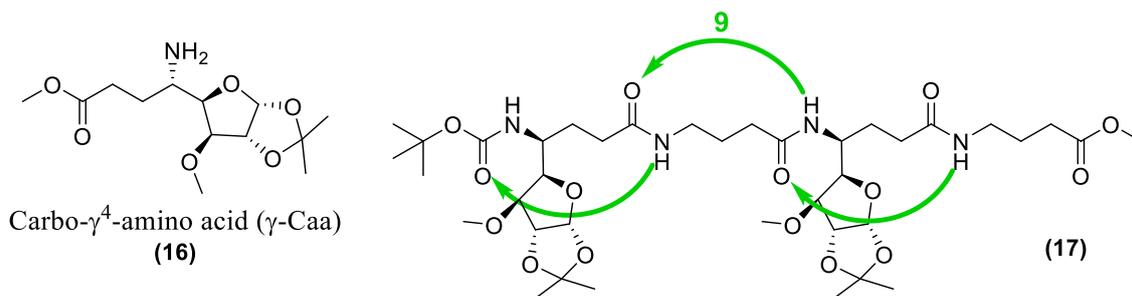


Figure 38: Structure of (**16**) and tetramer (**17**) stabilized by hydrogen bonds  $i-2, i C=O \cdots H-N$

<sup>93</sup> Vasudev, P. G. *et al.*, *Angew. Chem. Int. Ed.*, **44**, 4972–75 (2005).

<sup>94</sup> Sharma, G. V. M. *et al.*, *Angew. Chem. Int. Ed.*, **45**, 2944–47 (2006).

iii.b. Structures accessible to cyclic aliphatic  $\gamma$ -peptides

As in  $\beta$ -peptide series, a cyclic conformational restriction may improve the stability of a specific folding or even favour the formation of a new secondary structure. Thus, Gellman *et al.* used a cyclohexyl-containing  $\gamma$ -amino acid<sup>[95]</sup>. The ring constraint and stereochemistry promote a *gauche*<sup>+</sup>, *gauche*<sup>+</sup> torsion-angle around the C $_{\alpha}$ -C $_{\beta}$  and C $_{\beta}$ -C $_{\gamma}$  bonds favouring the 14-helix (Figure 39, A). In a similar manner, a C $_9$ -bend ribbon was obtained from oligomers of *cis*- $\gamma$ -amino-L-proline (Figure 39, B).<sup>[96]</sup> A C $_7$ -bend-ribbon was also observed for homochiral and heterochiral tartrate-derived peptides in benzene (Figure 39, C).<sup>[97]</sup> Finally, cyclopropane-containing  $\gamma$ -peptides were also studied by Smith.<sup>[98]</sup> Such oligomers adopt a unique infinite parallel sheet structure at both the solid state and in chloroform solution.<sup>[99]</sup>

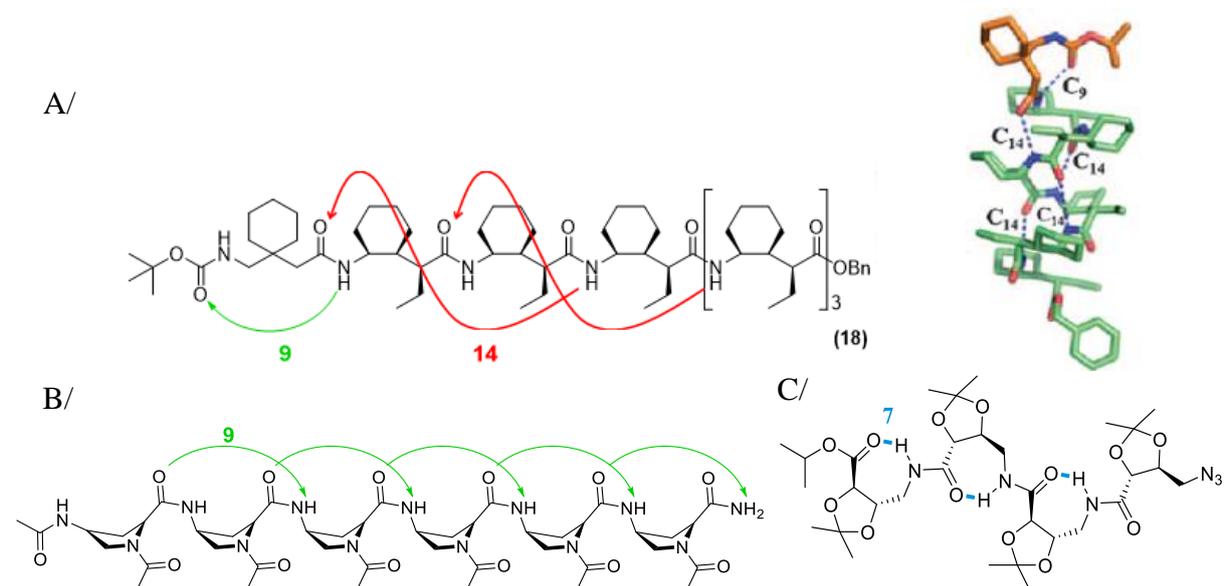


Figure 39: A/Structure of cyclohexyl-containing  $\gamma$ -peptide (18) and its crystal structure,<sup>[95]</sup> B/ Chemical structure of the *cis*- $\gamma$ -aminoproline peptide,<sup>[96]</sup> C/Observed bend-ribbon solution conformation of tartrate-derived peptides<sup>[97]</sup>

iii.c. ATC-based oligomers as prototype of aromatic  $\gamma$ -peptide foldamers

Over the last few years, our laboratory has been working on the study of aromatic  $\gamma$ -amino acid building blocks and the structural behaviour of their oligomers. This research program was inspired by the work of Baldauf and Hofmann, who explored, by *ab initio* calculations, the folding properties of vinylogous  $\gamma$ -peptides. They suggested that the introduction of a

<sup>95</sup> Guo, L. *et al.*, *Angew. Chem. Int. Ed.*, **50**, 5843–46 (2011).

<sup>96</sup> Farrera-Sinfreu, J. *et al.*, *J. Am. Chem. Soc.*, **126**, 6048–57 (2004).

<sup>97</sup> Kothari, A. *et al.*, *Chem. Commun.*, **0**, 2814–16 (2007).

<sup>98</sup> Smith, E. M. *et al.*, *J. Org. Chem.*, **62**, 7906–7 (1997).

<sup>99</sup> Jones, C. R. *et al.*, *Angew. Chem. Int. Ed Engl.*, **47**, 7099–7102 (2008).

(*E*)-double bond between the C<sub>α</sub> and C<sub>β</sub> would lead to helical structures starting from 14- to up to 27-membered pseudocycles but, would not conduct to smaller helices.<sup>[85]</sup> On the other hand, the (*Z*)-configuration, with a  $\zeta$  dihedral angle at 0°, would favour peptidic nearest-neighbour interactions heading to helices with 7- and 9-membered pseudocycles, with unique dihedral values.<sup>[100]</sup> This hypothesis was corroborated by Grison *et al.* who observed an intramolecular C<sub>9</sub>-pseudocycle in a hybrid tripeptide that contained a *cis*-vinylogous derivative of proline (Figure 40).<sup>[101]</sup> A similar C<sub>9</sub>-hydrogen bond was also described by Mann and Kessler in a oxazole-based peptidomimetic.<sup>[102]</sup> As in a (*Z*)-vinylogous  $\gamma$ -amino acid, the heterocycle imposes that the  $\zeta$  dihedral angle is 0°.

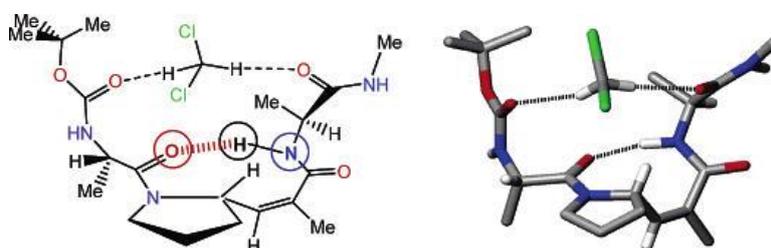


Figure 40: Schematic and stereoscopic view of the crystal molecular structure of the *cis*-vinylogous tripeptide Boc-Ala-‘Pro<sup>cis</sup>’-Ala-NH-Me with a DCM molecule joining the N- and C-terminal extremities<sup>[101]</sup>

In 2013, our group published the first structure of a heterocyclic  $\gamma$ -peptide. The monomer called 4-Amino-(methyl)-1,3-Thiazole-5-Carboxylic acid (ATC) (**20**) was built around a thiazole ring, providing a conformational limitation at 0° around the C<sub>α</sub> and C<sub>β</sub> (Figure 41).<sup>[103]</sup>

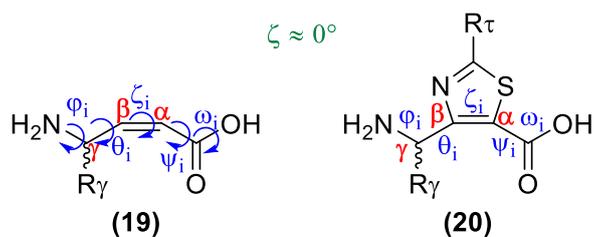


Figure 41: (*Z*)-vinylogous  $\gamma$ -amino acid (**19**) and ATC (**20**) pattern

NMR studies performed in CDCl<sub>3</sub> and water showed well-dispersed signals associated with high downfield chemical shifts of the amide protons, higher than 9.10 ppm. Later, this high deshielding effect was unambiguously associated with the formation of a C<sub>9</sub> intermolecular

<sup>100</sup> Baldauf, C. *et al.*, *J. Org. Chem.*, **70**, 5351–61 (2005).

<sup>101</sup> Grison, C. *et al.*, *J. Org. Chem.*, **70**, 10753–64 (2005).

<sup>102</sup> Mann, E. *et al.*, *Org. Lett.*, **5**, 4567–70 (2003).

<sup>103</sup> Mathieu, L. *et al.*, *Angew. Chem. Int. Ed.*, **52**, 6006–10 (2013).

hydrogen-bond.<sup>[104]</sup> The structure of the ATC oligomers was finally solved either by simulated annealing protocol (AMBER 10<sup>[105]</sup>) starting from NMR constraints or by XRD. The folding corresponds to a right-handed helix different from those of Balaram,<sup>[93]</sup> stabilized by hydrogen bonds between the main-chain carbonyl at position  $i$  and the amide proton at position  $i+2$ , forming a series of interwove 9-membered pseudocycles (Figure 42, B). The crystal structure revealed the same H-bonding pattern in one direction, from the N- to the C-terminus (Figure 42, C). Each turn of the 9-helix is composed of three residues with a pitch of 11.8 Å. The rise-per-residue is of 3.9 Å and each turn exhibits six substitution patterns, distributed around a 60° angle along the axis (Figure 42, C and D). Interestingly, only two residues are needed to induce the C<sub>9</sub> folding.

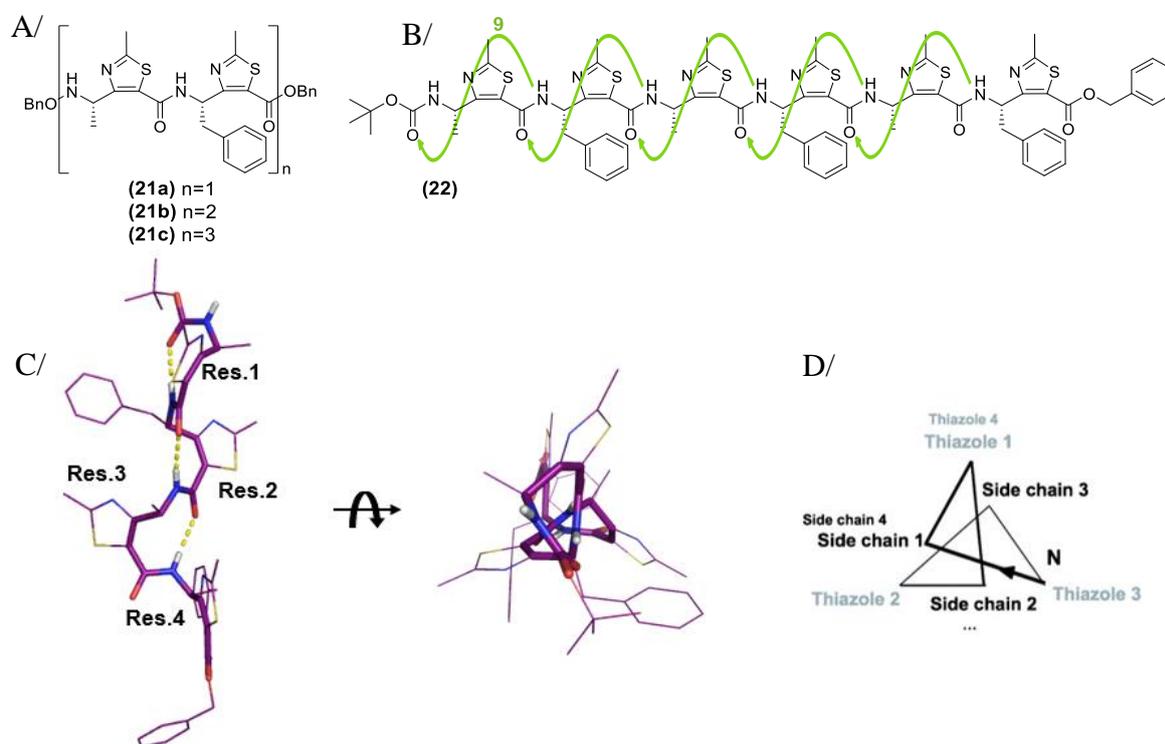


Figure 42 A/ATC-based  $\gamma$ -peptides (**21**) synthesized in solution, B/Representation of the intrahelical H-bonding pattern on hexamer (**22**), C/X-ray structure of (**21b**), D/Axial projection of the ATC-oligomers

The structure of the ATC oligomers is perfectly consistent with the 9-helix predicted by Baldauf and Hofmann for oligomers of (*Z*)-vinyllogous  $\gamma$ -amino acids (Table 5 for comparison of dihedral angles).<sup>[100,105,106]</sup> However, it is not clear why ATC oligomers never adopt the 7-helical conformation which was calculated by the same authors as similar in energy as the 9-

<sup>104</sup> Bonnel, C. *et al.*, *Org. Biomol. Chem.*, **14**, 8664–69 (2016).

<sup>105</sup> Case, D. A. *et al.* (University of California, San Francisco, 2008).

helix. A partial answer can be found in recent studies on hybrid oligomers alternating either ATCs and  $\alpha$ -amino acids, or (Z)-vinylogous  $\gamma$ - and  $\alpha$ -amino acids which display different conformational behaviours. While the latter adopt a 12-helical shape,<sup>[106]</sup> The ATC-containing oligomers fold into C<sub>9/12</sub>-ribbons.<sup>[107]</sup> In ATC-based molecules, the carbonyls and double bonds are coplanar while for the 12-helix of (Z)-vinylogous  $\gamma$ -amino acids they are not. Whatever the sequences studied, the ATC carbonyl group point towards the sulfur atom of the thiazole ring suggesting a dipole-dipole interaction between the C=O and the heterocycle. Thus, conformational behaviour of the ATC oligomers seems to result from the establishment of a regular H-bonding network oriented by both a strict limitation of the  $\zeta$  dihedral angle at 0° and a possible electronic interaction between the aromatic ring and the amide bond.

Table 5: Average backbone torsion angles for (Z)-vinylogous  $\gamma$ -oligomers and ATC-based foldamers (**21a,b,c**) and (**22**). The  $\psi$  angles of the last residue of the NMR and XRD structures were omitted.

Oligomers	$\phi$	$\theta$	$\zeta$	$\psi$
Predicted (Z)-vinylogous $\gamma$ -oligomers	-79.8°	122.8°	0.1°	-46.5°
ATC-oligomers ( <b>21a</b> ), ( <b>21b</b> ), ( <b>21c</b> )	-76 ± 16°	128 ± 12°	-6 ± 5°	-28 ± 6°
XRD structure of ( <b>22</b> )	-78 ± 3°	127 ± 14°	0 ± 3°	-41 ± 4°

As highlighted by circular dichroism, the 9-helical folding of ATC-based  $\gamma$ -peptides is, surprisingly, very stable in a wide range of solvents such as nonpolar chloroform or polar methanol or even more protic water from pH 3 to 11. In addition, the ATC-monomers present the advantage of an easy modulation of their side chains. Notably, the folding appeared of low dependency of the lateral chains flanking the backbone. For instance, we recently reported ATC-based foldamer as a template for the design of glycoclusters.<sup>[50]</sup> The main purpose of this project was to flank sugar moieties on an original glycol-peptide scaffold in order to target lectins. The obtained results demonstrated that the overall shape of the ATC-oligomer was not affected by the mannosyl substituents.

### iii.d. *Oligourea peptidomimetics*

Although there are many peptidomimetics in which the peptide amide bond was replaced by another moiety, the exploration towards oligomeric synthesis, and study of the resulting properties of these larger peptidomimetics containing more than three residues was performed on only a very small subset of them. Among these are the urea peptidomimetics. In such

<sup>106</sup> Ganesh Kumar, M. *et al.*, *Angew. Chem. Int. Ed.*, **55**, 7847–51 (2016).

<sup>107</sup> Bonnel, C. *et al.*, *Chem. - Eur. J.*, **23**, 17584–91 (2017).

oligomers, the number of atoms between two sequential carbonyl groups is similar to those found in  $\gamma$ -peptides. However, because of an additional nitrogen atom within each residue, a new H-bond pattern may be formed resulting in a different folding behaviour.

The first studies of oligoureas were performed by Nowick *et al.* in 1992, to mimic turns and  $\beta$ -sheets.<sup>[108,109]</sup> Thanks to the NMR and IR studies, they were able to prove that di- and triureas form intramolecular hydrogen bonds between their  $\text{CO}_i$  and  $\text{NH}_{i+3}$  so that their scaffold could be used for mimicking  $\beta$ -sheet-like structures. Even though it was demonstrated in 1995 that the synthesis of oligoureas is possible with solid phase peptide synthesis (SPPS)<sup>[110,111,112]</sup> their structural preferences had not been investigated much until the Guichard group started in 2002.<sup>[113,114]</sup> At first, they postulated that the substitution of the  $\alpha$ -carbon by a NH group within a  $\gamma$ -amino acid would be compatible with a folding into a 14-helix, similar to that found in the  $\gamma^4$ -peptides (Figure 43, A).

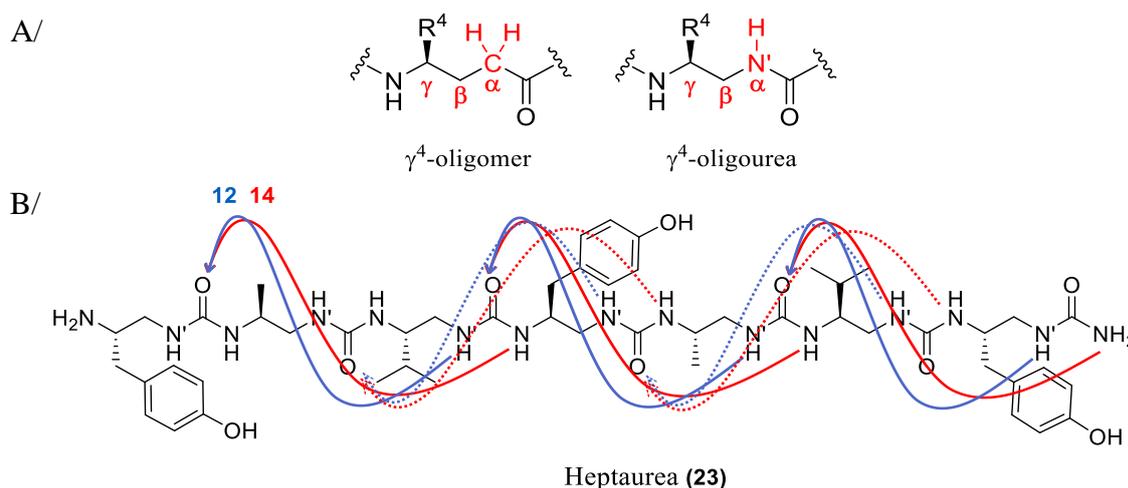


Figure 43: A/ $\gamma^4$ -oligomer versus  $\gamma^4$ -oligourea, B/Structure of heptaurea (**23**) stabilized by hydrogen bonds  $i,i+3 \text{ C=O} \cdots \text{H-N}$  and  $i,i+2 \text{ C=O} \cdots \text{H-N}'$

In order to verify this hypothesis, Guichard *et al.* prepared the heptaurea (**23**) by solid phase synthesis (Figure 43, B). NMR studies of (**23**) in pyridine- $d_5$  showed that this oligomer adopts a well-defined right-handed 12/14-helix stabilized by hydrogen bonds between  $\text{CO}_i \cdots \text{HN}_{i+3}$  for the 14-membered-ring (in red, Figure 43, B) and  $\text{CO}_i \cdots \text{HN}'_{i+2}$  regarding the pseudocycle

<sup>108</sup> Nowick, J. S. *et al.*, *J. Org. Chem.*, **57**, 3763–65 (1992).

<sup>109</sup> Nowick, J. S. *et al.*, *J. Am. Chem. Soc.*, **117**, 89–99 (1995).

<sup>110</sup> Burgess, K. *et al.*, *Angew. Chem. Int. Ed. Engl.*, **34**, 907–9 (1995).

<sup>111</sup> Burgess, K. *et al.*, *J. Am. Chem. Soc.*, **119**, 1556–64 (1997).

<sup>112</sup> Kim, J.-M. *et al.*, *Tetrahedron Lett.*, **37**, 5305–8 (1996).

<sup>113</sup> Semetey, V. *et al.*, *Angew. Chem. Int. Ed.*, **41**, 1893–95 (2002).

<sup>114</sup> Hemmerlin, C. *et al.*, *Helv. Chim. Acta*, **85**, 3692–3711 (2002).

of 12 atoms (in blue, Figure 43, B). The pitch of this helix is of 5.1 Å and 2.5 residues are needed to complete a turn, therefore, this helix is also called the  $2.5_{12/14}$ -helix. These helical parameters are quite close to the  $\gamma^4$ -peptides corresponding 14-helix (5 Å and 2.6 residues per turn). Circular dichroism experiments were performed on this  $N,N'$ -linked oligoureas and it seemed that there is a specific signature for such  $2.5_{12/14}$ -helix (Figure 44). Indeed, in methanol, a maximum band is found at 203 nm while  $\gamma^4$ -peptides folding into 14-helices do not provide any specific CD signature. Further studies led to the conclusion that five residues were enough to stabilize the  $2.5_{12/14}$ -helices.<sup>[115,116]</sup> Finally, one can notice that the first two urea-residues, located on the N-terminal extremity of the oligoureas chain were not involved in the intramolecular hydrogen-bonding pattern.<sup>[53]</sup>

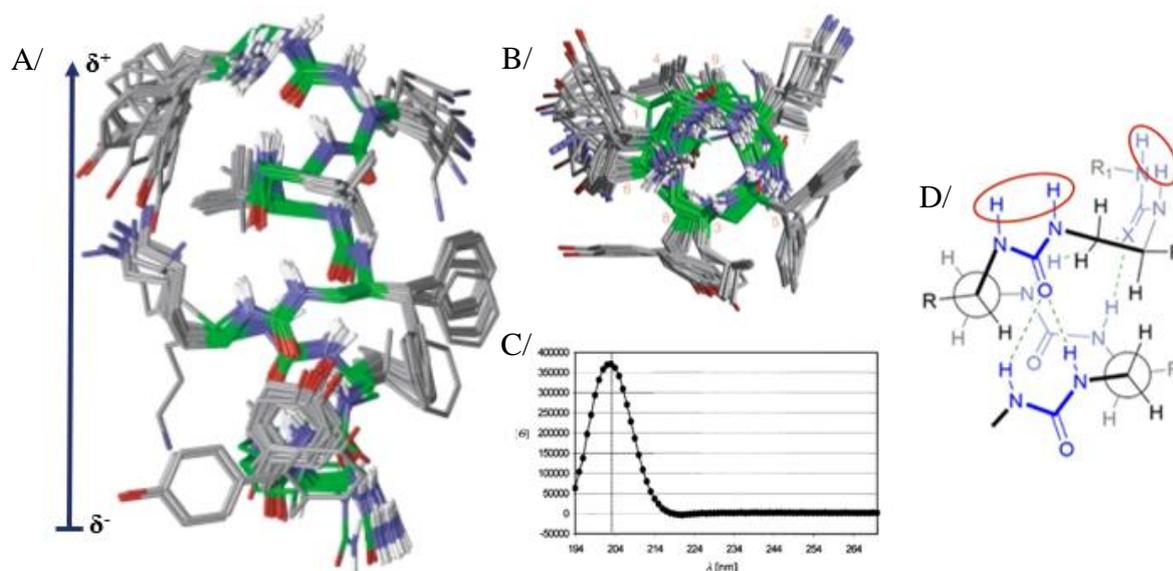


Figure 44: NMR Structure of oligourea (**23**) in pyridine- $d_5$ : A/stereo view along the helix axis and B/top view. C/CD signature of (**23**) in MeOH at room temperature.<sup>[114]</sup> D/Representation of an  $N,N'$ -linked oligoureas and its folded behaviour illustrating the two accessible H-bond donor sites near the positive pole of the helix macrodipole<sup>[53]</sup>

### iii.e. Conclusion

Over the last two decades, many folded architectures mimicking protein secondary structures have been described. Earlier studies mainly focused on their structural properties. Advances in synthetic folded molecules, coupled with the discovery that oligomeric backbones can retain folding in water, opened the way for foldamers to interfere with biological functions. In this context, initial work focused on the design of polycationic amphipathic sequences to tailor the

<sup>115</sup> Violette, A. *et al.*, *J. Am. Chem. Soc.*, **127**, 2156–64 (2005).

<sup>116</sup> Violette, A. *et al.*, *Chem. – Eur. J.*, **14**, 3874–82 (2008).

physicochemical characteristics of antimicrobial peptides (AMPs).<sup>[117]</sup> Peptidomimetic antimicrobials were made using  $\beta$ -peptides,<sup>[118,119,120,121,122]</sup> peptoids<sup>[123]</sup> and oligoureas<sup>[124,125]</sup> as scaffold. Unlike the  $\alpha$ -peptide backbone of natural AMPs, they offer the additional advantage of being resistant to proteases. A more challenging application of peptidomimetic foldamers concerns their use for mimicking protein segments, particularly  $\alpha$ -helix, involved in protein-protein interactions. Prototype works were employed to target the intracellular p53/hDM2<sup>[126,127]</sup> and Bcl-2/Bcl-x<sub>L</sub>,<sup>[128]</sup> interactions demonstrating the potential for foldamers to bind undruggable flat protein surfaces.

Although peptidomimetic foldamers were extensively explored for biomedical applications, their potential as asymmetric catalysts was little explored. One major question to be addressed is to what extent can a small foldamer sequence be adapted to achieve catalytic functions. The next section will review this literature.

## 2. Catalytic foldamers

The discovery of short peptide sequences that function as asymmetric catalysts for a variety of reactions is now well documented. Both combinatorial and *de novo* methods of catalyst design were demonstrated successfully to perform a variety of reactions including kinetic resolutions of alcohols, aldol reactions, phosphorylation, Michael addition, and Morita–Baylis–Hillman reactions.<sup>[129,130,131,132]</sup> Despite both combinatorial and rational design methods proved fruitful, it remains a need for robust systems that permit rational design of versatile peptide-based catalysts. In many cases, secondary structural elements showed to be crucial for catalytic activity and chemical, regio-, or enantio-selectivity. However, the state of a short peptide in solution is generally characterized not by one configuration but by an ensemble of

<sup>117</sup> Tew, G. N. *et al.*, *Acc. Chem. Res.*, **43**, 30–39 (2010).

<sup>118</sup> Hamuro, Y. *et al.*, *J. Am. Chem. Soc.*, **121**, 12200–201 (1999).

<sup>119</sup> Liu, D. *et al.*, *J. Am. Chem. Soc.*, **123**, 7553–59 (2001).

<sup>120</sup> Porter, E. A. *et al.*, *Nature*, **404**, 565 (2000).

<sup>121</sup> Porter, E. A. *et al.*, *J. Am. Chem. Soc.*, **124**, 7324–30 (2002).

<sup>122</sup> Arvidsson, P. I. *et al.*, *Chem. Biodivers.*, **2**, 401–20 (2005).

<sup>123</sup> Patch, J. A. *et al.*, *J. Am. Chem. Soc.*, **125**, 12092–93 (2003).

<sup>124</sup> Violette, A. *et al.*, *Chem. Biol.*, **13**, 531–38 (2006).

<sup>125</sup> Claudon, P. *et al.*, *Angew. Chem. Int. Ed Engl.*, **49**, 333–36 (2010).

<sup>126</sup> Kritzer, J. A. *et al.*, *J. Am. Chem. Soc.*, **127**, 167–78 (2005).

<sup>127</sup> Bautista, A. D. *et al.*, *J. Am. Chem. Soc.*, **132**, 2904–6 (2010).

<sup>128</sup> Sadovsky, J. D. *et al.*, *J. Am. Chem. Soc.*, **127**, 11966–68 (2005).

<sup>129</sup> Davie, E. A. C. *et al.*, *Chem. Rev.*, **107**, 5759–5812 (2007).

<sup>130</sup> Miller, S. J., *Acc. Chem. Res.*, **37**, 601–10 (2004).

<sup>131</sup> Mogharabi, M. *et al.*, *Trends Pept. Protein Sci.*, **1**, 89–98 (2017).

<sup>132</sup> Zozulia, O. *et al.*, *Chem. Soc. Rev.* (2018).

structures limiting the ability to re-create all the desirable characteristics for an enzyme-like catalyst. Foldamer science may provide solutions to these problems.

### i. Peptoids for enantioselective catalysis

In this context, a chiral environment associated with a competent catalytic functional group is the key elements for catalyst design. Therefore, one can imagine positioning reactive sites on the foldamer side chains. This was achieved by Maayan *et al.* who described peptoids substituted on their lateral chains with a 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) group (Figure 45). The oligomers were explored for oxidation of a racemic mixture of 1-phenylethanol and their activity were compared to those of the corresponding peptides.<sup>[52]</sup>

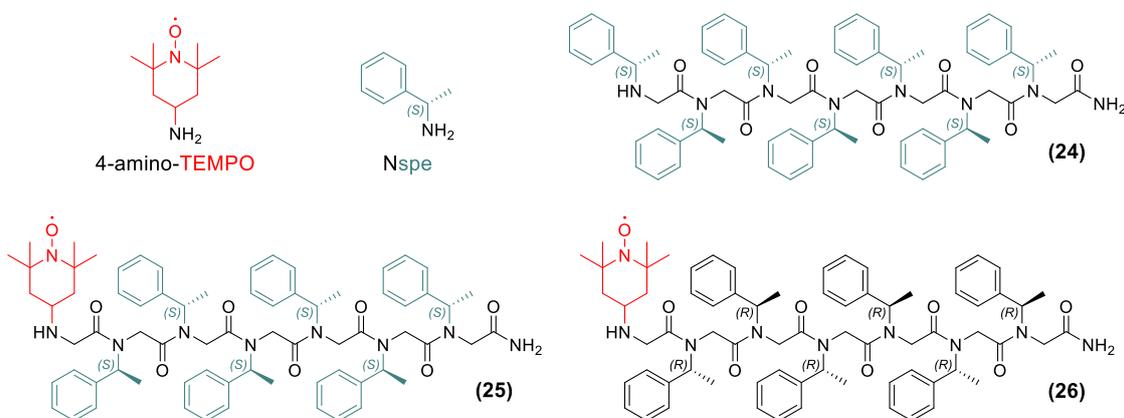
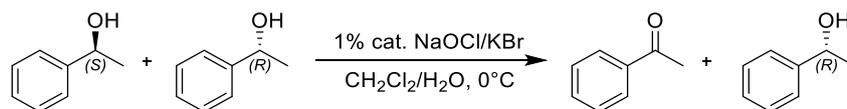


Figure 45: 4-amino-TEMPO, Nspe monomer and peptoid sequences (24), (25) and (26)

While (S)-phenylalanine monomer (spe) mixed with 4-amino-TEMPO led to a low conversion with no selectivity for one enantiomer (Table 6, entry 1), the peptoid counterpart TEMPO-N(spe)<sub>6</sub> (25) showed excellent conversion with a selectivity towards the (S) enantiomer (Table 6, entry 3). Inversion of the chirality of the oligomer (compound (26)) led to a complete change of the enantioselectivity (Table 6, entry 4). Interestingly, no asymmetric induction was observed when the TEMPO group was not covalently linked to the peptoid backbone (Table 6, entry 2).

Table 6: Conversions, enantioselectivities, and enantiomeric excess for the catalytic oxidation of 1-phenylethanol by peptoids<sup>[52]</sup>

Entry	Catalytic system	Conv. <sup>a</sup> (%)	Selectivity <sup>b</sup> (%)	ee (%)
1	4-amino-TEMPO + Nspe	22	None	None
2	4-amino-TEMPO + <b>(24)</b>	86	None	None
3	<b>(25)</b> (right-handed helix)	84	60 ( <i>S</i> )	>99 ( <i>R</i> )
4	<b>(26)</b> (left-handed helix)	85	59 ( <i>R</i> )	>99 ( <i>S</i> )

<sup>a</sup>After 2h reaction, <sup>b</sup>Selectivity at the quoted conversion value is defined as (% preferred enantiomer/% conversion)

## ii. $\beta$ -Peptides as catalysts

### ii.a. Juliá–Colonna asymmetric epoxidation of enones

The first example of a catalytic pseudo-peptide foldamer was probably reported by Smith *et al.*<sup>[133]</sup> By analogy to the work of Juliá and Colonna, who used poly- $\alpha$ -leucines as catalytic species, they described poly- $\beta^3$ -leucines as able to catalyse the epoxidation of (*E*)- $\alpha,\beta$ -enones by H<sub>2</sub>O<sub>2</sub> with significant enantioselectivity. The exact role of the  $\alpha$ - and  $\beta$ -peptides in the reaction mechanism is not clear. Since the reaction occurred in a heterogeneous medium, it should be hypothesized that the substrate is adsorbed on the oligomer surface that induce a chiral environment for the action of H<sub>2</sub>O<sub>2</sub>. In this article, the authors did not study the conformational preference of the oligomers nor demonstrated any superiority of the  $\beta$ -peptides over the  $\alpha$ -ones.

### ii.b. Retroaldol cleavage of $\beta$ -hydroxyketones

In 2009, Hilvert *et al.* used  $\beta$ -peptides as platforms to create a foldamer that has an aldolase activity in aqueous buffer.<sup>[134]</sup> It was previously known that lysine rich amphiphilic  $\alpha$ -helices can catalyse retroaldol cleavage of  $\beta$ -hydroxyketones through the formation of imine or enamine intermediates. Thus, a  $\beta$ -peptide sequence that contains multiple ACHC residues along with  $\beta^3$ -homolysine ( $\beta^3$ -hLys) residues in an *i, i+3, i+6* array (catalyst **(27)**) was designed to

<sup>133</sup> Coffey, P. E. *et al.*, *Chem. Commun. Camb. Engl.*, 2330–31 (2001).

<sup>134</sup> Müller, M. M. *et al.*, *Angew. Chem.*, **121**, 940–43 (2009).

lead to alignment of the amine-containing side chains along one side of a 14-helix (Figure 46). The N-terminus was capped by a  $\beta^3$ -hTyr residue to facilitate concentration determination and with an heptanoyl moiety to promote self-assembly of the oligomer. The  $\beta$ -peptide (**27**) was tested for its ability to promote the retroaldol cleavage of the 4-phenyl-4-hydroxy-2-oxobutyrate, an anionic  $\beta$ -hydroxyketone expected to bind to the cationic catalyst. While (**27**) showed high catalytic activity, a scrambled sequence that does not match the three-residue repeat ACHC-ACHC- $\beta^3$ -hLys of the 14-helix and, an analogue of (**27**) in which all of the preorganized ACHC residues are replaced by flexible  $\beta^3$ -hVal residues, were relatively poor catalysts under identical reaction conditions. The authors also demonstrated a favourable effect of the  $\beta$ -peptide clustering for the aldolase activity. As highlighted in the article, this favourable effect may reflect a reduction in the  $pK_a$  value of the  $\beta^3$ -hLys side-chain ammonium group. This could be due to an increase in the hydrophobicity of the environment of catalytic amines or a high local positive-charge density. Finally, another effect that could actually operate jointly with the previous one is such that the interface between helical  $\beta$ -peptides would be acting as a primitive substrate-binding pocket.

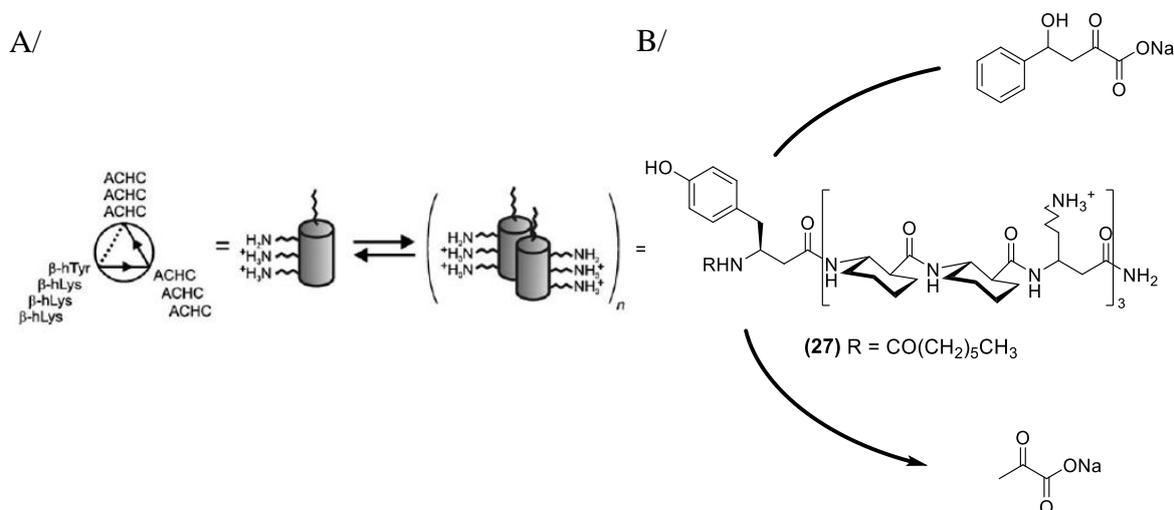


Figure 46: A/Helical wheel diagram of  $\beta$ -peptide (**27**) (left) and a cartoon showing  $\beta$ -peptide self-assembly (right; the partial protonation state shown is purely hypothetical) B/Retroaldol reaction of sodium 4-phenyl-4-hydroxy-2-oxobutyrate catalysed by (**27**)

### ii.c. Hydrolysis of trisulfonate esters

Independently to Hilvert and Gellman, Schepartz group rationally designed a  $\beta^3$ -peptide bundle catalyst for hydrolysis of 8-acetoxypyrene-1,3,6-trisulfonate.<sup>[135]</sup> They modified the Zwitter-YYYK sequence, known to fold in a bundle octamer, and introduced  $\beta^3$ -hArg (=  $\beta^3$ R, Figure

<sup>135</sup> Wang, P. S. P. *et al.*, *J. Am. Chem. Soc.*, **136**, 6810–13 (2014).

47, A) at key positions to electrostatically orientate the recognition of the substrate in close proximity to a catalytic  $\alpha$ -His residue. The helical stability of the architecture was essential to guide the design of a  $\beta^3$ -peptide bundle. Despite the low propensity to self-assemble, the first generation of synthesized peptides, at 10 mol%, was increasing the background reaction rate by a factor of 20 to 30 while free histidine at 10 mol% enhanced the reaction rate by less than 5-fold. More importantly, steady-state measurements revealed that hydrolysis was following Michaelis-Menten kinetics. To improve bundle stability, they synthesized a covalent dimer containing two  $\beta$ Est-2 monomers joined with a tetra- $\beta$ -hGly linker (Figure 47, B). The  $\beta$ Est-2-sequence showed a propensity to structure in an octamer and displayed  $K_{\text{cat}}/K_{\text{M}} = 5102 \text{ M}^{-1} \text{ min}^{-1}$ , almost two orders of magnitude greater than that for the  $\beta$ Est-2 monomer. Based on XRD structure of the bundle catalyst, twenty potential active sites per assembly were recognized, this could explain the differences in activity between bundle forming and monomeric  $\beta^3$ -peptide catalysts.

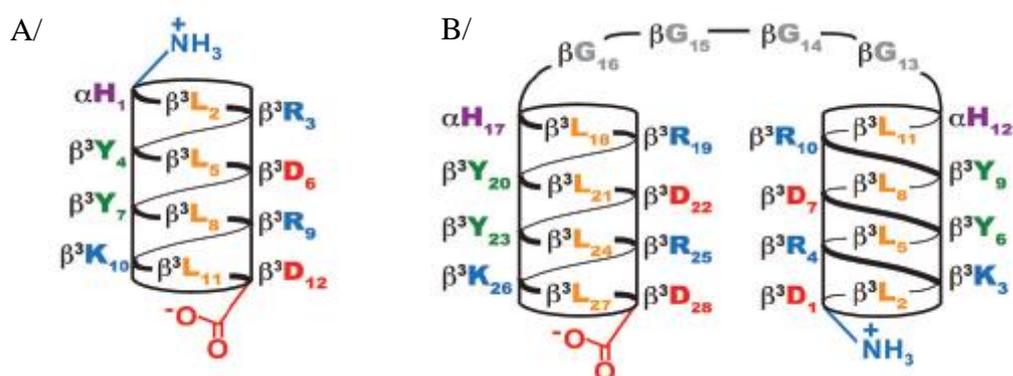


Figure 47: Helical schematic of A/ $\beta$ Est-2, B/  $\beta$ Est-2 monomers joined with a tetra- $\beta$ -hGly linker<sup>[136]</sup>

#### ii.d. Bifunctional catalysis

A few other types of folded catalysts were investigated, one of the latest, developed by Gellman *et al.*, relies on a  $\beta$ -peptide helix bearing pyrrolidine groups for cross aldol reactions.<sup>[137]</sup> As previously settled, this group showed the propensity for ACPC-oligomers to fold into 12-helices (*Chapter 1, II, 1, ii.b*), based on this evidence, they developed pyrrolidine derivative (APC) for introduction into oligomers (Figure 48, A). A geometric study of APC-containing foldamers was performed in order to get the best spatial arrangement to provide a favourable secondary amine diad for bifunctional catalysis. Hence,  $\beta$ -peptides of ACPC/APC and mixed  $\alpha/\beta$ -peptides

<sup>136</sup> Wang, P. S. P. *et al.*, *Chem. Commun.*, **52**, 7420–32 (2016).

<sup>137</sup> Girvin, Z. C. *et al.*, *J. Am. Chem. Soc.*, **140**, 12476–83 (2018).

of  $\alpha$ -residues Aib and ACPC/APC  $\beta$ -monomers were synthesized (Figure 48, B). It appeared that the 1:2  $\alpha/\beta$ -oligomer (**31**) provided a propitious geometry of the secondary amine diad, for bifunctional catalysis as this helix is stabilized by hydrogen bonds between the  $\text{CO}_i$  and the  $\text{NH}_{i+3}$ . This helix contains three residues per turn, thus bringing the two APC residues of (**31**) above one another and exactly one turn apart.

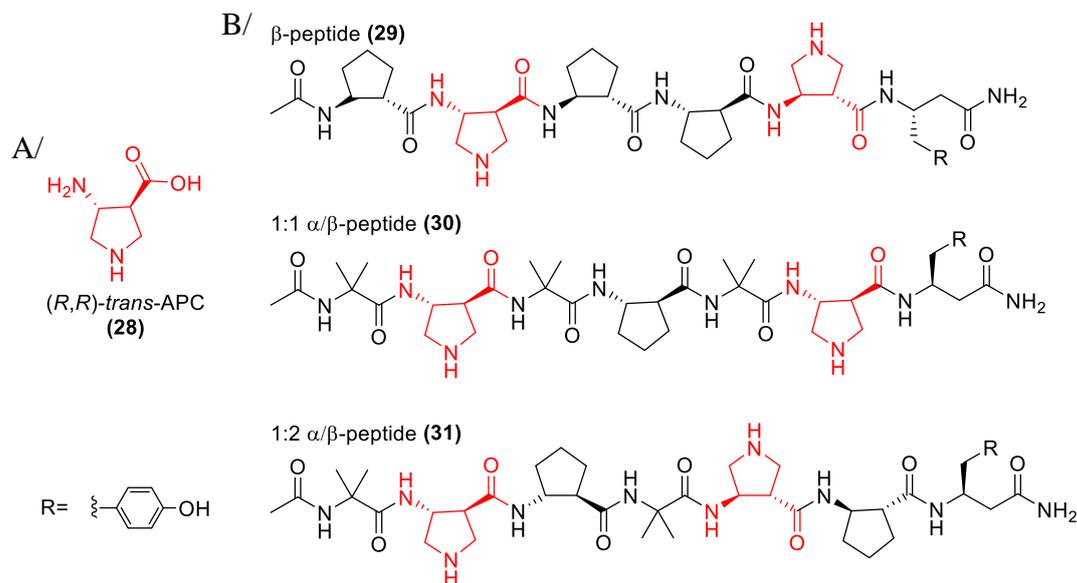


Figure 48: A/Cyclic  $\beta$ -amino acid pyrrolidine derivative APC, B/Foldamers to evaluate different APC diad geometries

The  $\alpha/\beta$ -peptide (**31**) was tested in the cross-aldol reaction of formaldehyde to hydrocinnamaldehyde. By monitoring this reaction by UPLC, it appeared that this catalysis was indeed based on the principle of intramolecular dual covalent catalysis, with a coordinate action of the two secondary amine moieties, as depicted in the proposed mechanism, Figure 49. To start, one of the bis-APC foldamer secondary amine is condensed with the hydrocinnamaldehyde while the nitrogen of the other moiety reacts with the formaldehyde to generate an enamine-iminium intermediate (**A**). Then, an iminium (**B**) is formed by intramolecular attack of the enamine on the iminium, creating a cross-linked intermediate. (**C**) is generated by tautomerism of (**B**), followed by N-protonation and scission of the C-N bond (**D**). The catalyst is then recovered by hydrolysis of the iminium (**D**) which releases the expected cross-aldol product.

The authors highlight with this paper the importance of the spatial organization of reactive moieties and the possibility to access new scaffolds by modulating the position or the nature of the monomers to promote bifunctional catalysis.

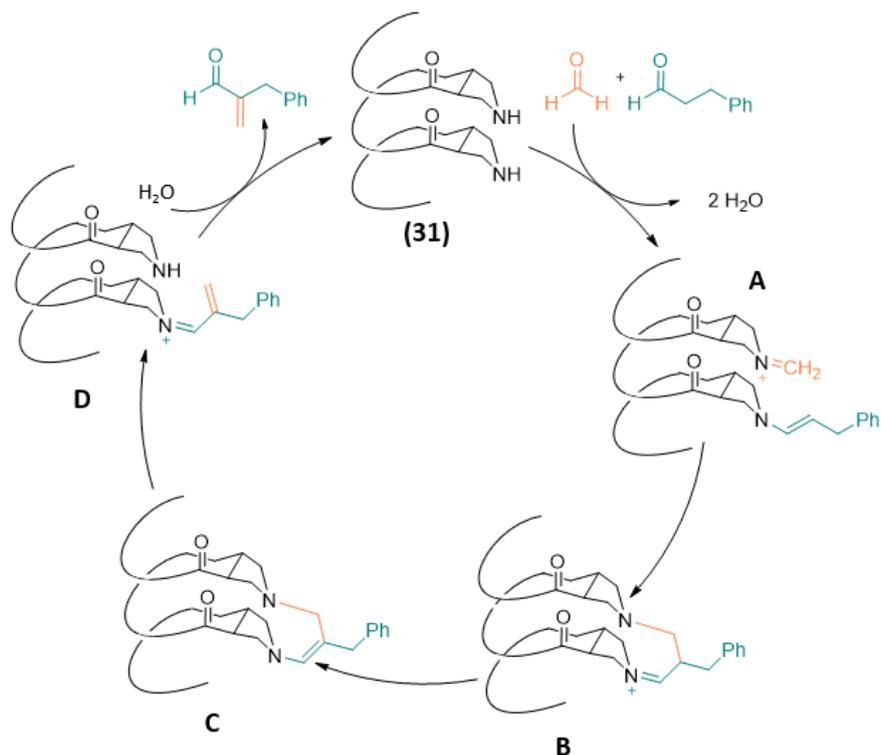


Figure 49: Proposed catalytic cycle for the foldamer-catalysed crossed aldol reaction<sup>[137]</sup>

### iii. Aib-oligomers and oligoureas for the C-C bond formation and control of chirality

Over the last decade, Clayden's group explored the concepts of screw-sense preference, that is the thermodynamic preference for a helix to adopt a specific left- or right-handedness, by building synthetic helices made from achiral monomers.<sup>[138]</sup> They demonstrated that if an Aib-oligomer is connected to a chiral terminus extremity, its two screw-sense conformers might be desymmetrised and thus drive to an imbalance in the population of the two interconverting screw-sense conformers. To induce a screw-sense preference, the terminal residue must be connected to the organization of the final  $\beta$ -turn of the  $3_{10}$ -helix (as Aib-oligomers of more than three residues are usually  $3_{10}$ -helical (*Chapter 1, II, 1*)). This ability was used to control the chiral environment of an achiral aminothiurea moiety placed at the second extremity of the oligomer. The remotely inducible chiral thiurea centre was able to catalyse the nitro-Michael addition reaction of diethylmalonate to  $\beta$ -trans-nitrostyrene. The enantiomeric ratio is reflecting the screw-sense preference of the Aib-oligomer (Figure 50).

<sup>138</sup> Le Bailly, B. A. F. *et al.*, *Angew. Chem. Int. Ed Engl.*, **55**, 2132–36 (2016).

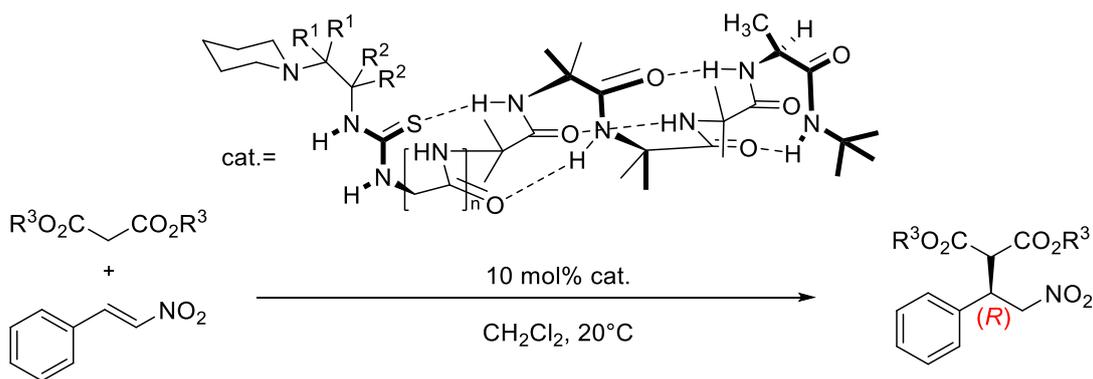


Figure 50: Catalytic foldamer with a remotely inducible catalytic site<sup>[138]</sup>

Concerning the  $\gamma$ -peptides, Guichard *et al.* reported helical oligopeptides foldamers as catalysts for C-C bonding formation.<sup>[53]</sup> For a reminder, they demonstrated that their structure owns two hydrogen-bond donors sites on the N-terminal extremity (circled in red on the hexamer (32a), Figure 51, A). Therefore, they focused their efforts on a catalysis based on the principle of H-bond donor catalysts. Moreover, as demonstrated by Clayden and explained in the previous paragraph, the internally organized amine/thiourea catalyst was proved to be an efficient catalyst for asymmetric induction. Hence, the objective of this study was to demonstrate that this well-defined 2.5-helix of oligopeptide could act as chiral catalyst capable of forming extended intramolecular hydrogen bonds. This would then increase the H-bond donor ability of the thiourea and leave space for the substrate-catalyst coordination around this moiety.

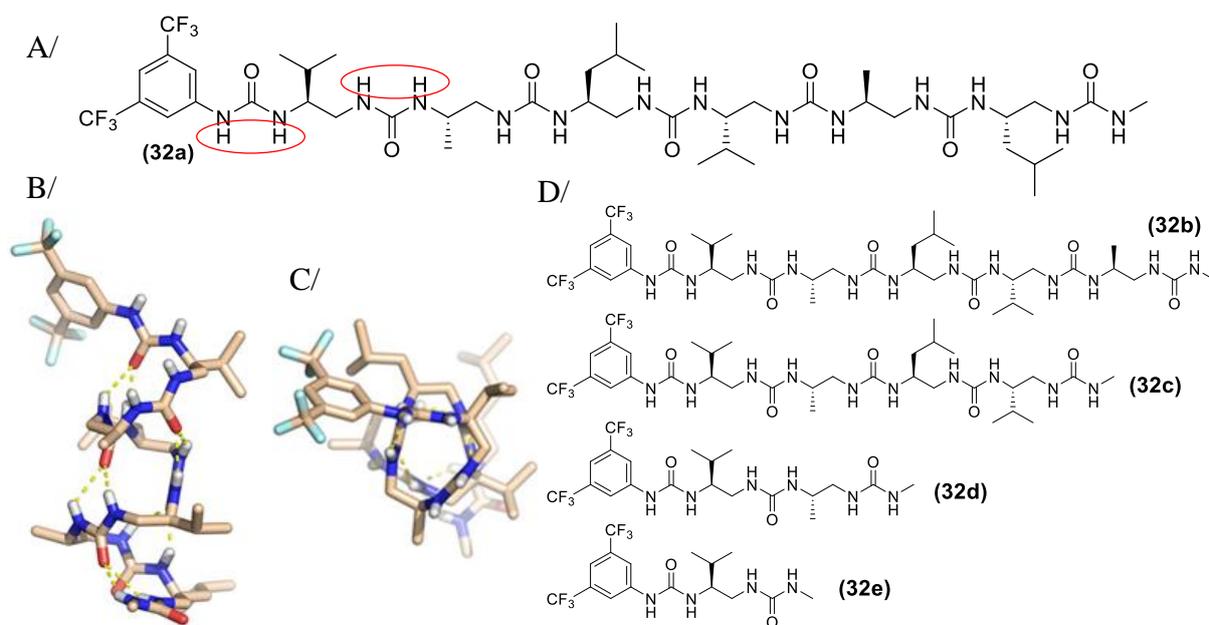
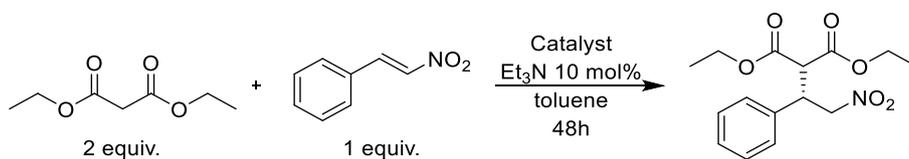


Figure 51: A/Oligopeptide foldamers (32a) the free H of the hexamer are circled in red, B/X-ray structure of (32a) along the axis, C/down the axis, D/Oligopeptide foldamers (32b-e)

The study established that only 0.01 mol% of foldamer (**32a**), in association with an external Brønsted base, can catalyse the Michael addition reaction between enolizable carbonyl compounds and nitroolefins (Table 7, entries 1-3). The reaction is performed in toluene, various parameters were tested, such as the loading of the catalyst, the influence of the temperature and the length of the foldamer (Table 7).

With this work, the authors were able to promote an efficient aliphatic N,N'-linked oligourea as chiral catalyst. Moreover, they demonstrated the importance of the folding in such oligomers as almost no enantiocontrol was observed when the reaction took place in presence of monomers and dimers of ureas but the highest reactivity and selectivity were found for the penta- and hexamers (Table 7).

Table 7: Impact of the oligourea catalyst loading, chain length and temperature



Entry	Catalyst	Cat. loading (mol%)	Temp. (°C)	Yield <sup>a</sup> (%)	ee <sup>b</sup> (%)
1	<b>(32a)</b>	0.1	-20	86	95
2	<b>(32a)</b>	0.1	rt	75	89
3	<b>(32a)</b>	0.01	-20	82	91
4	<b>(32b)</b>	0.1	rt	83	87
5	<b>(32c)</b>	0.1	rt	47	36
6	<b>(32d)</b>	0.1	rt	44	6
7	<b>(32e)</b>	0.1	rt	55	5

<sup>a</sup>Yield of isolated product after chromatography, <sup>b</sup>Enantiomeric excess determined by chiral HPLC.<sup>[53]</sup>

As we have seen in this paragraph, exploring the propensity for folded unnatural oligomers as a scaffold for catalysis remains a challenging topic. Research in this area is ongoing and many breakthroughs remain to be achieved.

### III. THESIS OBJECTIVES

As explored in this chapter, over the past several decades, peptides and peptide-based molecules have emerged as promising minimal enzymatic systems. Combinatorial methods and de novo designs have led to real successes for a large variety of reactions, such as aldolisation, phosphorylation, Michael addition or Morita-Baylis-Hillman reactions.<sup>[129]</sup> However, the difficulty to control their three-dimensional structure limits their use to re-create all the desirable characteristics for an enzyme-like catalyst.<sup>[1]</sup> As we have reviewed, recently it was proposed to build bio-inspired catalytic systems from artificial oligomers, called “foldamers”, with high conformational stability and structural predictabilities, mimicking the protein secondary structures (helices, sheets and ribbons).<sup>[139,134,52]</sup> Potentially, such catalysts could present numerous advantages and could be an original solution to answer both the economic and environmental issues faced in organic chemistry. Chemically accessible, they can, unlike the enzyme, be used in various organic solvents and within a large temperature and pressure scale. Furthermore, the activity and specificity of the catalyst can be easily modulated by changing the nature of its catalytic centre and/or by modulating the topology of the folded backbone and the microenvironment around the catalytic centre.

In this context, my work was devoted to the development of bio-inspired foldamer catalysts based on a heterocyclic  $\gamma$ -peptide scaffold and one major question to be addressed was to what extent a small foldamer can be adapted to achieve catalytic functions. The general aim was to use the ATC-based foldamers as a template for the development of modular enamine-type catalysts. In this context, the next chapter of this manuscript will provide a short overview of the use of peptides in enamine catalysis. Particularly, we will see how the proline-containing peptides reported by the group of H. Wennemers inspired us for the design of a catalytically competent ATC monomer. The next section will be devoted to the introduction of the ATC monomer bearing the catalytic function into a set of ATC-based oligomers. One central feature was to explore systematically the connection between foldamer sizes and/or shapes, and, both, reactivity and enantioselectivity. Consequently, a steady study of the catalytic properties of the synthesized oligomers will be performed. We will also focus on the reliable structural markers to assess the folding properties of the catalytic oligomers by combining NMR, FTIR and circular dichroism experiments. Finally, I will describe the development of a script for NMR-

---

<sup>139</sup> Smaldone, R. A. *et al.*, *Chem. – Eur. J.*, **14**, 2650–57 (2008).

based 3D-modelling that was used to calculate the folding of the ATC oligomers on an atomic level.

In this document, we will use the following nomenclature for the ATC monomers: **PG-NH-{R $\tau$ ;R $\gamma$ }-ATC-OH**. Chemset numbering is related to the lateral chains **R $\tau$** , **R $\gamma$** , where **R $\tau$**  corresponds to the substituent bore by the thiazole ring in position 2 and **R $\gamma$**  to the  $\gamma$  lateral chain. **OH** means that the acidic function on the C-terminal extremity is unprotected and **PG-NH** corresponds to the amine function on the N-terminal side, with its protecting group.

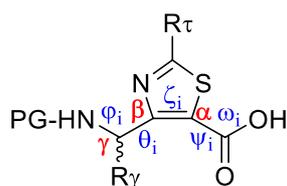


Figure 52: General structure of an ATC-monomer

CHAPTER 2:  
INTRODUCTION OF A CATALYTIC  
FUNCTION ON AN ATC-OLIGOMER

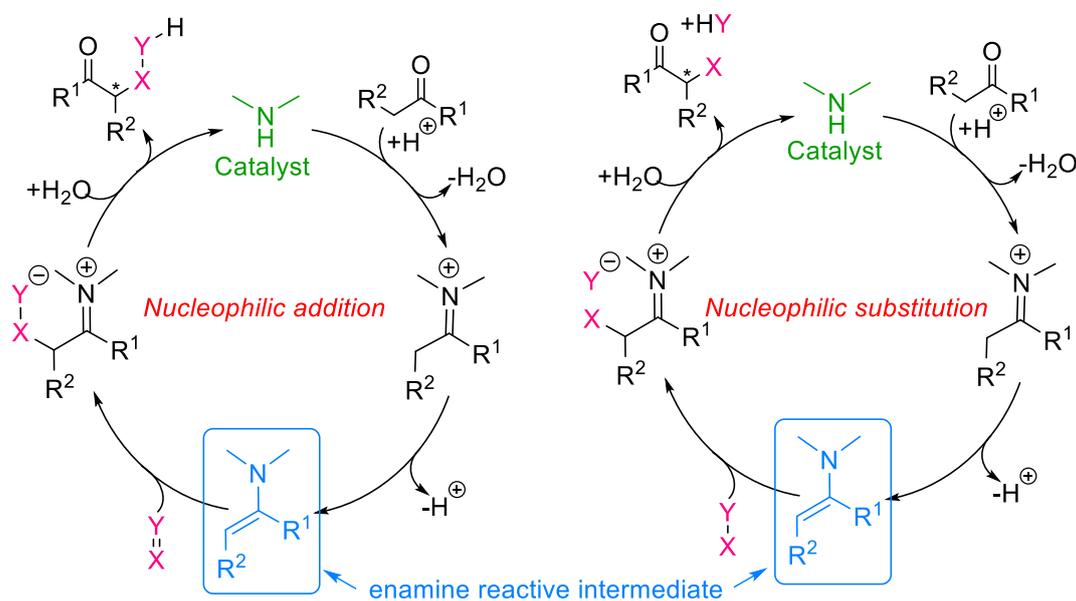


Before reviewing the work that has been accomplished over these three years of PhD, I would like to finish my introduction by presenting a short review focused on the enamine-type catalysis.

## I. ENAMINE CATALYSIS

### 1. History: proline as a prototype for enamine catalysis

Enamine catalysis is defined as being the catalysis of electrophilic substitution reactions in the  $\alpha$ -position of a carbonyl compound, by a primary or secondary amine via an enamine intermediate. This type of catalysis was subject to much discussion as it provides, in terms of the number of steps, an economical way for enantioselective synthesis and one of the most powerful methods for the stereoselective  $\alpha$ -functionalization of aldehydes and ketones. It involves catalytically generated enamine intermediates formed via deprotonation of an iminium ion that can react with various electrophiles. There are two modes of enamine catalysis, determined by the class of electrophile used (Scheme 1).<sup>[140,141]</sup>



Scheme 1: Enamine catalysis, arrows may be considered equilibria

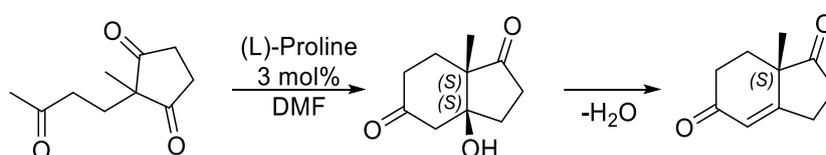
A primary or secondary amine reacts with a ketone or an aldehyde to create an iminium ion that spontaneously converts into an enamine through a tautomerism equilibrium. The resulting

<sup>140</sup> Mukherjee, S. *et al.*, *Chem. Rev.*, **107**, 5471–5569 (2007).

<sup>141</sup> List, B., *Chem. Commun.*, 819–24 (2006).

enamine then reacts in one of the two possible ways. In a first potential chemical route, a nucleophilic *addition* of the enamine, occurs when the electrophile contains double bonds, like aldehydes, imines or Michael acceptors and so on. The second possible pathway results from the nucleophilic *substitution* by the enamine of a single bond containing electrophiles such as alkyl halides.

The independent discoveries in 1971, of Hajos and Parrish's group and Eder, Sauer and Wiechert's group, of L-proline enantioselectively catalysing 6-enolendo intramolecular aldol reactions of di- and triketones, is probably the fundamental to enamine catalysis (Scheme 2).<sup>[142,143,144]</sup>



Scheme 2: Hajos-Parrish-Eder-Sauer-Wiechert reaction<sup>[145]</sup>

Hajos and Parrish initially proposed two possible mechanisms for this reaction. The first hypothesis being the formation of a protonated enamine and an oxazolidone ring (Figure 53, A). The second one is based on the addition of the L-proline's zwitterionic form on one of the two carbonyls of the cyclopentanedione (Figure 53, B).

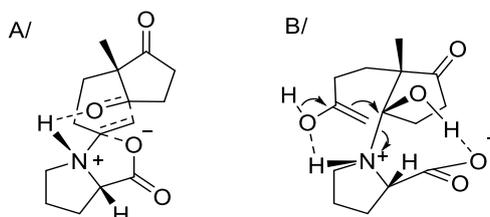


Figure 53: A/Formation of a protonated enamine; B/Addition of the L-proline's zwitterionic form

Initially considered as true by Hajos,<sup>[146]</sup> the second hypothesis was rejected by List *et al.* In fact, when performed in <sup>18</sup>O-labelled water, the latter obtained more than 90% of labelled oxygen incorporation in the final product, which is not compatible with the putative

<sup>142</sup> Hajos, Z. G. *et al.*, *Ger. Pat.* DE2102623 (1971).

<sup>143</sup> Eder, U. *et al.*, *Ger. Pat.* DE2014757 (1971).

<sup>144</sup> Eder, U. *et al.*, *Angew. Chem. Int. Ed. Engl.*, **10**, 496–97 (1971).

<sup>145</sup> Clemente, F. R. *et al.*, *Angew. Chem. Int. Ed.*, **43**, 5766–68 (2004).

<sup>146</sup> Hajos, Z. G. *et al.*, *J. Org. Chem.*, **39**, 1615–21 (1974).

mechanism.<sup>[147]</sup> In 1976, Jung proposed a third mechanism involving the formation of a C-C bond, together with the transfer of the carboxylic acid's proton on the carbonyl group (Figure 54, A).<sup>[148]</sup> This model was challenged in the 80s by Agami's group who proposed a mechanism based on the hydrogen transfer guided by a second molecule of proline (Figure 54, B).<sup>[149]</sup>

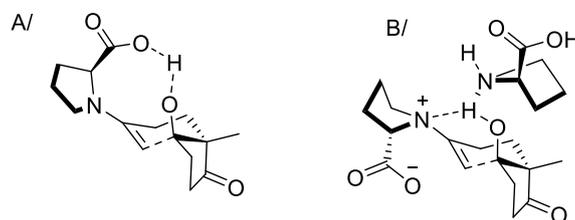
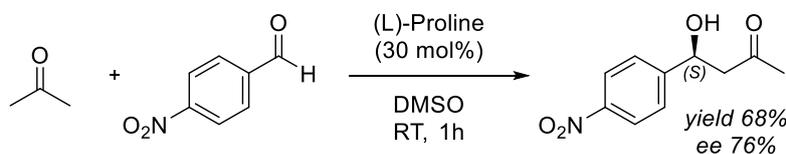


Figure 54: Reaction intermediates proposed by A/Jung, B/Agami

However, because of the absence of dilution effect on the enantioselectivity and, secondly, due to the kinetics obtained during this reaction, List, Houk *et al.* showed, in 2003, that only one proline would be involved in the mechanism.<sup>[150]</sup> Subsequently, Jung's model became the most probable one. Eventually, helped with the Density Functional Theory (DFT), Houk brought new elements supporting the formation of an enamine intermediate.<sup>[145]</sup>

### i. Asymmetric aldol reaction

As proline can catalyse intramolecular aldol reactions, many groups were interested in whether proline would catalyse intermolecular aldol reactions or not. The first case of such reaction was reported by List *et al.* in 2000, in which L-proline catalysed the aldol condensation of acetone to *p*-nitrobenzaldehyde giving the aldol product with an enantiomeric excess of 76% (Scheme 3).<sup>[151]</sup>



Scheme 3: The first proline-catalysed asymmetric intermolecular aldol reaction<sup>[151]</sup>

<sup>147</sup> List, B. *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 5839–42 (2004).

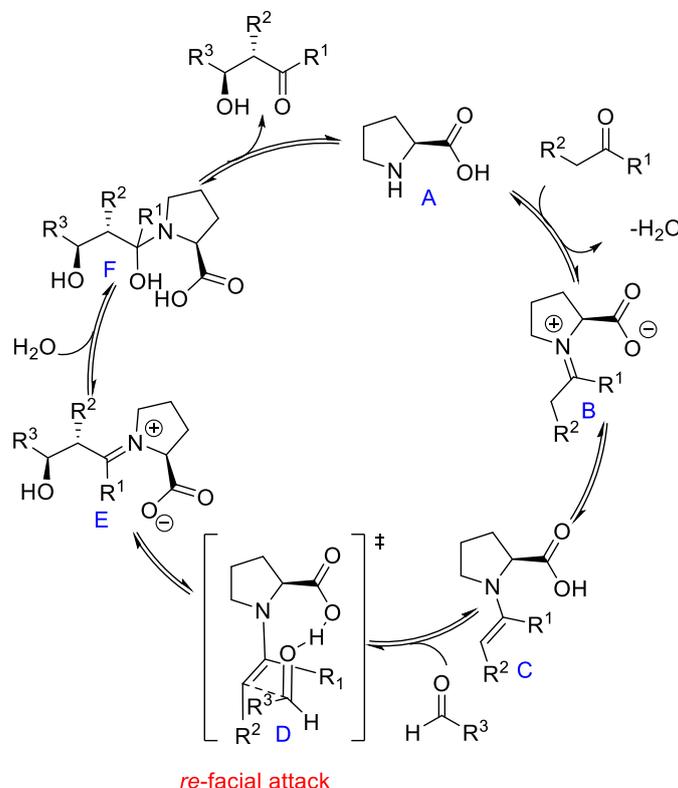
<sup>148</sup> Jung, M. E., *Tetrahedron*, **32**, 3–31 (1976).

<sup>149</sup> Agami, C., *Bull. Société Chim. Fr.*, 499–507 (1988).

<sup>150</sup> Hoang, L. *et al.*, *J. Am. Chem. Soc.*, **125**, 16–17 (2003).

<sup>151</sup> List, B. *et al.*, *J. Am. Chem. Soc.*, **122**, 2395–96 (2000).

The resulting catalytic mechanism involving the proline would be, from then onwards, described as an enamine type reaction, very similar to the accepted mechanism of class I aldolases (Scheme 4).<sup>[140,152]</sup>



Scheme 4: Catalytic cycle of proline-catalysed aldolizations

Proline **A** reacts with a ketone or aldehyde to form the iminium ion **B** which evolves to the nucleophilic enamine **C** by imine-enamine tautomerism. In the transition state of the carbon-carbon formation (**D**), the electrophile aldehyde  $R^3\text{CHO}$  approaches in a way that permits a *re*-facial attack of the enamine to the carbonyl, leading to the *anti*-product. This transition state was widely studied, and it is suggested that the carboxylic acid acts as a Brønsted cocatalyst, replacing the several acid/base functional groups involved in the aldolase mechanism, making proline a “bifunctional catalyst”. Finally, the iminium ion **E** is hydrolysed into a carbinolamine **F** to release the aldol product and the proline catalyst.<sup>[153]</sup> Supporting this mechanism, Marquez and Metzger isolated and characterized the intermediates **C**, **E** and **F** with an ESI-MS study.<sup>[154]</sup> With the success of the first proline-catalysed intermolecular reaction by List *et al.* and the non-

<sup>152</sup> Dalko, P. I. *et al.*, *Angew. Chem. Int. Ed.*, **43**, 5138–75 (2004).

<sup>153</sup> List, B. *et al.*, *Org. Lett.*, **3**, 573–75 (2000).

<sup>154</sup> Marquez, C. *et al.*, *Chem. Commun.*, 1539 (2006).

existent steric constraints applied to proline compared to enzymes, a large scope of electrophiles were tested. As highlighted by some examples found on the literature, the yields and enantiomeric excesses depend on the aldehyde used for the reaction (Table 8).<sup>[155]</sup>

Table 8: Examples of proline-catalysed aldol reactions<sup>[155]</sup>

Entry	R <sup>1</sup>	R <sup>2</sup>	Yield (%)	<i>anti</i> : <i>syn</i>	<i>anti</i> ee <sup>c</sup> (%)
1 <sup>a</sup>	OH	<i>i</i> Pr	63	>20:1	>99
2 <sup>a</sup>	OH	Ph	83	1:1	80
3 <sup>b</sup>	Me	Et	80	4:1	99
4 <sup>b</sup>	Me	<i>i</i> Bu	88	3:1	97
5 <sup>b</sup>	<i>n</i> Bu	<i>i</i> Pr	80	24:1	98
6 <sup>b</sup>	OBn	CH <sub>2</sub> OCH <sub>2</sub> Ph	73	4:1	99
7 <sup>b</sup>	Me	CH <sub>2</sub> OTIPS	75	4:1	99

<sup>a</sup>Performed in DMSO with 20 mol% of L-Proline, <sup>b</sup>Performed in DMF with 10 mol% of L-Proline, <sup>c</sup>Enantiomeric excess determined by chiral HPLC

## ii. Asymmetric three-component Mannich reaction

The Mannich reaction represents a highly useful extension of aldol-type approaches. The most widely recognized form of this reaction consists of three components, those are two carbonyl compounds and an amine, condensing to form a β-amino-carbonyl product.<sup>[156]</sup> The concomitant formation of a carbon-carbon bond and a carbon-nitrogen bond makes the Mannich reaction an extremely useful synthetic transformation. This reaction can be either *direct*, when using unmodified ketone donors or *indirect*, with preformed enolates generated stoichiometrically either *in situ* or by prior synthesis (Scheme 5).

<sup>155</sup> Tanaka, F. *et al.*, in *Enantioselective Organocatalysis*, ed. by Peter I. Dalko, pp. 19–55 (Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2007).

<sup>156</sup> Kleinman, E. F., in *Compr. Org. Synth.*, by Barry M. Trost *et al.*, **2**, 893–951 (Elsevier Ltd, 1991).



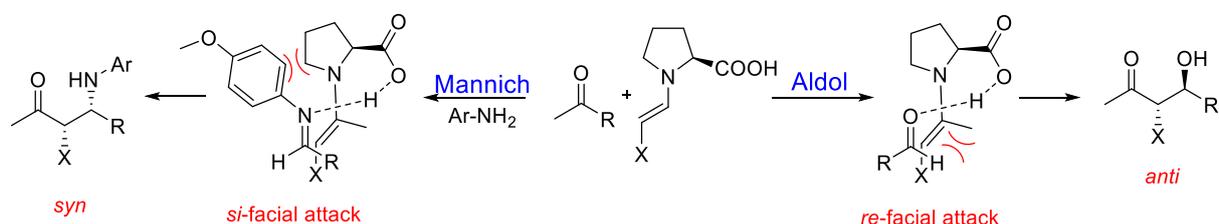
Table 9: Three-component direct Mannich reaction with different ketones and aldehydes

Entry	(ketone) R <sup>1</sup>	(aldehyde) R <sup>2</sup>	Product	Yield (%)	ee <sup>a</sup> (%)
1	H			50	94
2	CH <sub>3</sub>			96 (ratio 2.5:1)	99 (de>95%) 94%
3	OCH <sub>3</sub>			93	98
4	OH			92	>99
5	H			35	96
6	H			90	93
7	H			74	73
8	H			58	70

<sup>a</sup>Enantiomeric excess determined by chiral HPLC

Noticeably, the products from proline-catalysed Mannich reactions have the opposite absolute configuration to those of the aldol products, meaning that Mannich and aldol reactions have opposite enantiofacial selectivities (*si* and *re* respectively). It is assumed that the proline imine and enamine intermediates have a fixed (*E*)-configuration leading to a selective attack of the enamine on the *si*-face of the imine. The access on the *re*-face is limited by unfavourable steric interactions between the pyrrolidine of the proline and the *N*-substituent of the imine. This does

not exist in the aldol reaction, in which steric repulsion between aldehyde and enamine carbon substituents is dominant, leading to a *re*-facial attack (Scheme 7).<sup>[160]</sup>



Scheme 7: Opposite enantiofacial selectivity in Mannich (left) and aldol (right) transition states

### iii. Michael addition reaction

The Michael addition reaction is another efficient pathway to form carbon-carbon bonds. As the demand for optically active compounds grew, this reaction was widely studied to provide 1,4-addition adducts as valuable synthetic intermediates, such as  $\gamma$ -nitro ketones. Thus, the diversity of Michael acceptors used was broadly extended from malonic acid esters and  $\beta$ -ketoesters to nitroolefins. Regarding the Michael donors, ketones are usually used once activated into a more reactive species like enols or enamines even though this is not optimum in terms of atom economy. Therefore, a direct addition of unmodified carbonyl donors to Michael acceptors would make this reaction more advantageous. The first result of an enamine-catalytic asymmetric intermolecular Michael reaction was obtained by List *et al.* in 2001.<sup>[161]</sup> They first performed a proline-catalysed Michael addition between acetone and  $\beta$ -*trans*-nitrostyrene (Table 10, entry 1). Thereafter, they disclosed a variety of Michael addition between unmodified ketones and nitroolefins, even if this first enantioselectivity was low (7% ee). In general, it appeared that proline-catalysed Michael reactions are less enantioselective than the aldol or Mannich reactions, but its high efficiency and simple process are noteworthy for enamine catalysis. Finally, a high *syn* selectivity was observed, explained by an acyclic synclinal transition state based on a *re, re*-approach (Figure 55).<sup>[162]</sup> In such model, the reaction is under kinetic control.<sup>[163,164]</sup>

<sup>160</sup> List, B. *et al.*, *J. Am. Chem. Soc.*, **124**, 827–33 (2002).

<sup>161</sup> List, B. *et al.*, *Org. Lett.*, **3**, 2423–25 (2001).

<sup>162</sup> Seebach, D. *et al.*, *Helv. Chim. Acta*, **64**, 1413–23 (1981).

<sup>163</sup> Enders, D. *et al.*, *Synlett*, **2002**, 0026–0028 (2002).

<sup>164</sup> Sharma, A. K. *et al.*, *J. Org. Chem.*, **77**, 10516–24 (2012).

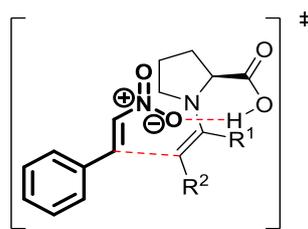


Figure 55: Proposed transition state for the diastereoselective Michael-addition of (*E*)-enamines to (*E*)-nitroolefins

The reaction is general as it could be extended to many ketones and nitroolefins. However, the chiral controls appeared much less favourable than in the aldol-type reaction (Table 10).

Table 10: Michael reaction between unmodified carbonyl donors and nitroolefins<sup>[161]</sup>

Entry	Product	Yield (%)	ee <sup>a</sup> (%) ( <i>syn</i> )	dr <sup>b</sup> ( <i>syn/anti</i> )
1		97	7	-
2		85	10	3:1
3		94	23	>20:1
4		92	10	>20:1
5		95	19	10:1

<sup>a</sup>Enantiomeric excess determined by chiral HPLC, <sup>b</sup>Determined by <sup>1</sup>H NMR of crude mixture.

In conclusion, following the pioneer work of Hajos, Parrish, Eder, Sauer and Wiechert, List's group opened the possibility of proline-based enamine catalysis. They demonstrated its potential for aldol-type condensation and nitro-Michael addition reactions. However, despite a worthy diastereoselectivity and regioselectivity in the addition of ketones to nitroolefins, the level of enantioselectivity was disappointing. Thus, different groups imagined using proline-containing peptides as an alternative to the low complexity of proline

catalysts.<sup>[165,166,167,168,169,170]</sup> They assumed that the peptide folding could drive a highly chiral local microenvironment thus, enhancing the chiral transfer. Considering both activity and selectivity, di-proline-containing tripeptides discovered by H. Wennemers are probably the most efficient catalysts for diastereo- and enantioselective aldol-type reactions and Michael additions of aldehydes to  $\beta$ -substituted nitroolefins.

## 2. Peptides in enamine catalysis

### i. H-Pro-Pro-Asp-NH<sub>2</sub> as a prototype catalyst for aldol reaction

In 2005, Wennemers's team identified two tripeptides, H-L-Pro-D-Ala-D-Asp-NH<sub>2</sub> (**33**) and H-L-Pro-L-Pro-L-Asp-NH<sub>2</sub> (**34**), as highly efficient catalysts on the aldol reaction between the nitro-benzaldehyde and the acetone (Figure 56).<sup>[171]</sup>

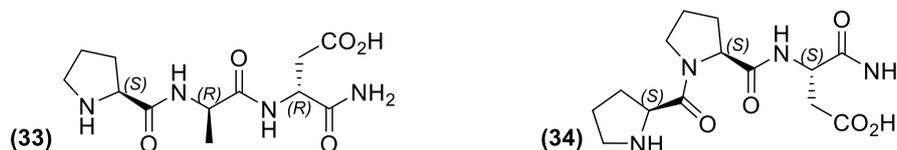


Figure 56: Wennemers's peptides (**33**) and (**34**)

While the use of 30 mol% of proline is needed to catalyse this transformation in four hours at room temperature ( $y = 68\%$ ,  $ee = 76\%$ ), only 10 mol% of (**33**) are necessary to lead to 73% yield, after 72 hours, and an enantiomeric excess up to 70%. Moreover, only 1 mol% of peptide (**34**) gave, in 4 hours, 99% of the aldol product in 80% ee. These results were explained by a better aptitude for this second peptide to adopt a compact structure, in comparison with the first one. In fact, based on molecular dynamic studies, the authors highlighted that the peptides (**33**) and (**34**) hold different conformations, a left-handed turn for the first one and a right-handed turn for the second. It is also interesting to note that the two peptides lead to two different enantiomers, even if the chirality of the catalytic N-terminus prolines are identical (Table 11). This result is of prime importance as it demonstrates that the chirality of the transformation

<sup>165</sup> Kofoed, J. *et al.*, *Bioorg. Med. Chem. Lett.*, **13**, 2445–47 (2003).

<sup>166</sup> Martin, T. J. *et al.*, *Synlett*, **12**, 1091–1902 (2003).

<sup>167</sup> Jarvo, E. R. *et al.*, *Tetrahedron*, **58**, 2481–95 (2002).

<sup>168</sup> Andrae, M. R. M. *et al.*, *Tetrahedron Asymmetry*, **16**, 2487–92 (2005).

<sup>169</sup> Luppi, G. *et al.*, *J. Org. Chem.*, **70**, 7418–21 (2005).

<sup>170</sup> Tang, Z. *et al.*, *J. Am. Chem. Soc.*, **127**, 9285–89 (2005).

<sup>171</sup> Krattiger, P. *et al.*, *Org. Lett.*, **7**, 1101–3 (2005).

might be driven not only by the catalytic centre but also by the overall conformational behaviour of the peptide catalyst.

Table 11: Aldol reactions catalysed by Wennemers's peptides<sup>[171,151]</sup>

Entry	R	10 mol% ( <b>33</b> )		1 mol% ( <b>34</b> )		30 mol% Proline	
		yield (%)	ee <sup>a</sup> (%)	yield (%)	ee <sup>a</sup> (%)	yield (%)	ee <sup>a</sup> (%)
1	4-NO <sub>2</sub> Ph	73	70 ( <i>R</i> )	99	80 ( <i>S</i> )	68	76 ( <i>R</i> )
2	Ph	58	66 ( <i>R</i> )	69	78 ( <i>S</i> )	62	60 ( <i>R</i> )
3	<i>c</i> -Hex	56	83 ( <i>R</i> )	66	82 ( <i>S</i> )	63	84 ( <i>R</i> )
4	<i>i</i> -Pr	75	91 ( <i>R</i> )	79	79 ( <i>S</i> )	97	96 ( <i>R</i> )
5	<i>neo</i> -Pent	24	70 ( <i>S</i> )	28	73 ( <i>R</i> )	22	36 ( <i>S</i> )

<sup>a</sup>Enantiomeric excess determined by chiral HPLC or GC analysis, (*R*) and (*S*) indicate the absolute configuration of the aldol product<sup>[171]</sup>

On a mechanistic aspect, the amine and acid functions are essential. If the secondary amine is either acetylated (Ac-Pro-Pro-Asp-NH<sub>2</sub> (**35**)) or methylated (Me-Pro-Pro-Asp-NH<sub>2</sub> (**36**)), or if the carboxylic acid is replaced by an amide (H-Pro-Pro-Asn-NH<sub>2</sub> (**37**)) or a sulfonic acid (H-Pro-Pro-Cys(SO<sub>3</sub>H)-NH<sub>2</sub> (**38**)), the catalytic activity is drastically reduced (Figure 57 and Table 12).<sup>[172]</sup>

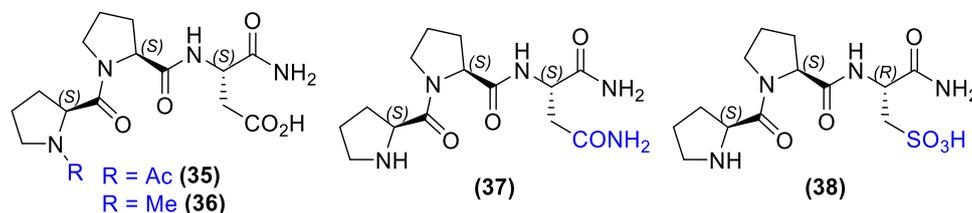


Figure 57: Peptides (**35**) to (**38**)

Table 12: Aldol reaction between *p*-nitrobenzaldehyde and acetone catalysed by peptides (**34**) to (**38**)

Entry	Catalyst	mol %	Time (h)	Yield (%)	ee <sup>a</sup> (%)
1	( <b>34</b> )	1	4	99	80
2 <sup>b</sup>	( <b>34</b> )	5	24	67	77
3	( <b>35</b> )	10	72	≤5	nd
4	( <b>36</b> )	10	72	≤5	nd
5	( <b>37</b> )	10	24	39	54
6	( <b>38</b> )	10	24	26	53

<sup>a</sup>Enantiomeric excess determined by chiral HPLC, <sup>b</sup>Reaction was performed in the absence of NMM

<sup>172</sup> Revell, J. D. *et al.*, *Tetrahedron*, **63**, 8420–24 (2007).

## ii. Scope extension: from aldol reaction to nitro-Michael addition

An article from 2008 suggested that the distance between the N-terminal secondary amine and the carboxylic acid might play a determining role in the reaction control.<sup>[173]</sup> While such a distance is of 3.7 Å for proline, it reaches 7.3 Å in the tripeptide (**34**) (Figure 58, A and B).

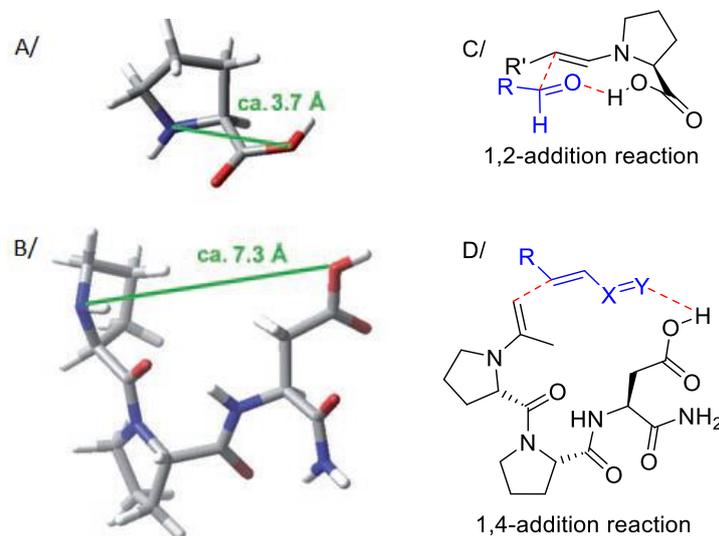


Figure 58: A/Lowest-energy conformation for the proline as calculated by MacroModel 8.0<sup>[173]</sup> B/Lowest-energy conformation of the (**34**) as calculated by MacroModel 8.0<sup>[171,173]</sup> C/Transition state of aldol reactions catalysed by proline proposed by Houk and List<sup>[145,147]</sup> D/Transition state for the 1,4-addition catalysed by Wennemers's tripeptide, X and Y represent heteroatoms or carbon atoms<sup>[173]</sup>

Thus, it was hypothesized that the tripeptide may be a better support for substrate recognition of 1,4-additions than proline. Tested on the nitro-Michael addition of *n*-butanal and  $\beta$ -*trans*-nitrostyrene, the tripeptide H-L-Pro-Pro-Asp-NH<sub>2</sub> (**34**) and its diastereomeric version (H-D-Pro-Pro-Asp-NH<sub>2</sub> (**39**)) led to 96 and 93% yields with a high preference for the *syn* product. The enantiomeric excesses (85 and 95%) were much higher than with proline. However, since the two catalysts led to products with opposite configurations, it appeared that the chirality of the transformation was driven by the absolute configuration of the N-terminal proline. This observation implies two different transition states resulting from an opposite approach of the nitroolefin to the enamine intermediate (Figure 59).

<sup>173</sup> Wiesner, M. *et al.*, *Angew. Chem. Int. Ed.*, **47**, 1871–74 (2008).

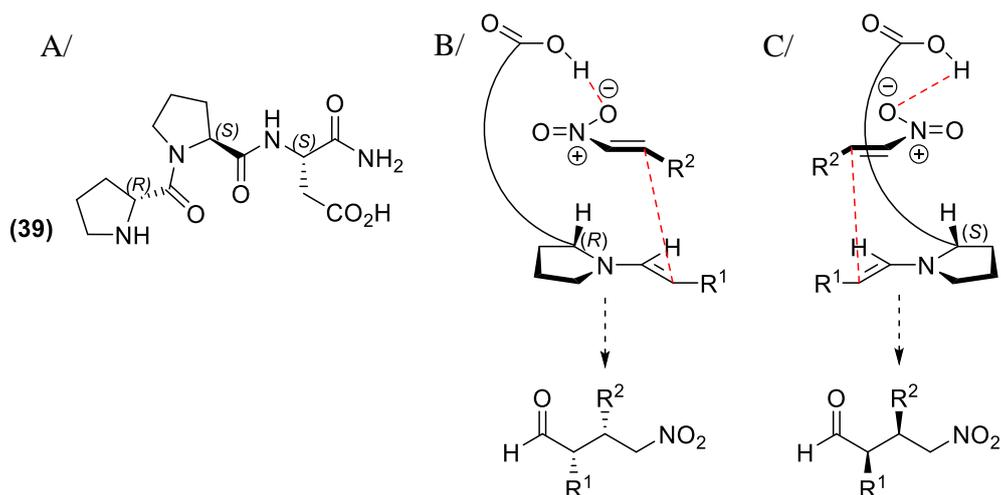


Figure 59: A/Wennemers's peptides (39). Proposed transition-state structures for B/(34) and C/(39) leading to enantiomeric syn products

### iii. Mechanistic studies of the nitro-Michael addition catalysed by H-D-Pro-Pro-Asp-NH<sub>2</sub>

Wennemers *et al.* attempted to describe the mechanism of the reaction and especially its rate determining step. ESI-MS back-reaction screening experiments supported an enamine mechanism since enamine and iminium were detected as intermediates.<sup>[174]</sup> Considering such pathway, the catalytic cycle should be divided into four main steps:

- 1- enamine formation
- 2- reaction with the nitroolefin to yield a new C-C bond
- 3- protonation of the nitronate intermediate
- 4- hydrolysis of the iminium

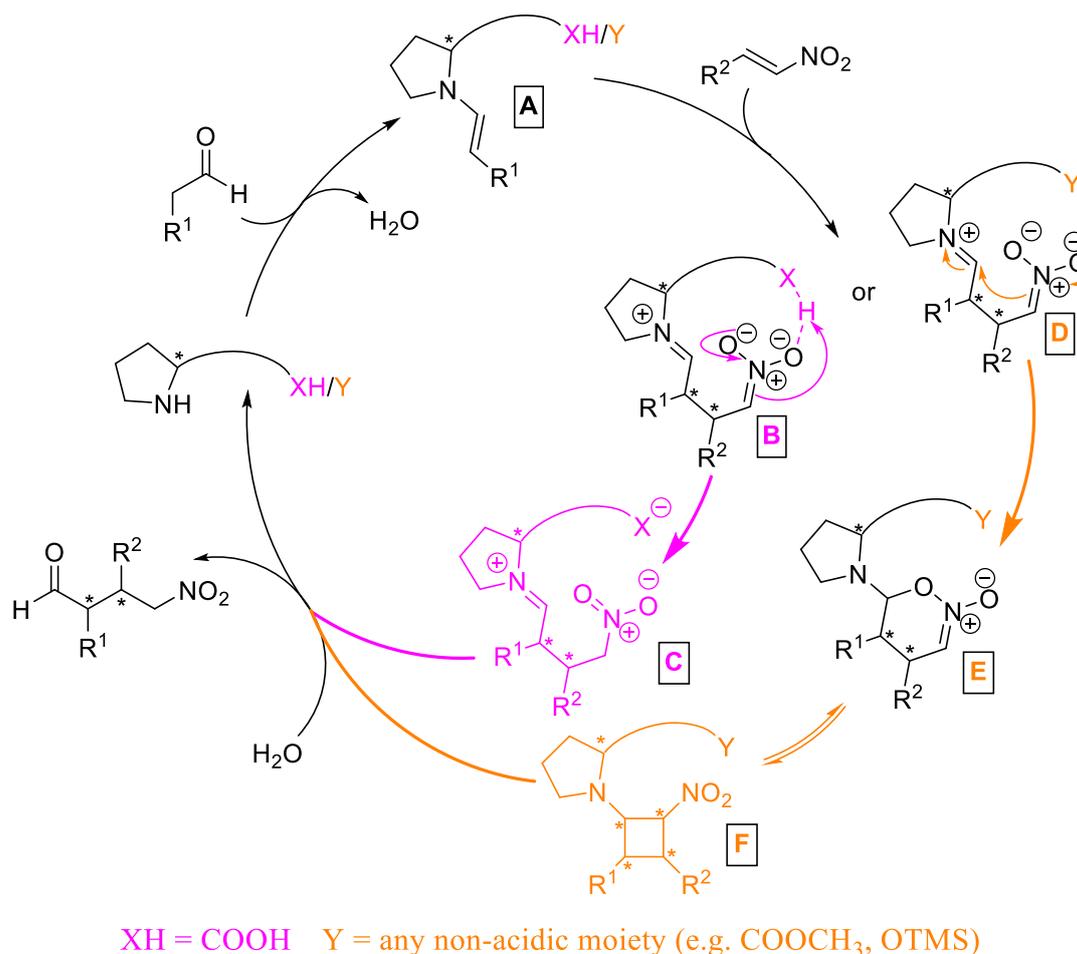
Kinetic studies provided insights that the reaction of the enamine (A, Scheme 8) with the electrophile is rate limiting and highlights a double role of the acidic group. It first orientates the reactivity and stereoselectivity and secondly, it improves the reaction rate by promoting protonation of the iminium nitronate (B, Scheme 8).<sup>[175,176]</sup> Although the acid group probably coordinates the nitronate function, its exact role as well as the protonation states along the catalytic pathway are not fully understood. Especially, its presence is not indispensable since a powerful catalytic tripeptide without acidic group was reported for the addition of nitro-Michael

<sup>174</sup> Bächle, F. *et al.*, *Angew. Chem. Int. Ed.*, **52**, 12619–23 (2013).

<sup>175</sup> Duschmalé, J. *et al.*, *Chem. Sci.*, **4**, 1312–18 (2013).

<sup>176</sup> Duschmalé, J. *et al.*, *Chem. – Eur. J.*, **18**, 1111–20 (2012).

between aldehydes and  $\beta,\beta$ -disubstituted nitro-olefins.<sup>[177]</sup> When the catalyst bears a carboxylic acid, the iminium nitronate **B** is intramolecularly protonated to form the iminium ion **C** which is then hydrolysed to release the expected product and the catalyst (Scheme 8, pink pathway). This intramolecular protonation is fast and does not limit the reaction rate, suggesting again that the limiting step is the C-C bond formation, which depends on the concentration of the nitroolefin and the enamine but not on the external proton sources.



Scheme 8: Catalytic cycle proposed for 1,4-additions of aldehydes to nitroolefins using catalyst with (pink) or without (orange) a proton-donor<sup>[175]</sup>

Alternative mechanisms were proposed for catalysts which do not own acidic hydrogen. For example, with chiral prolinol ether derivatives, cyclic intermediates such as dihydrooxazine and cyclobutane (respectively **E** and **F** in Scheme 8) were identified as resting states of

<sup>177</sup> Kastl, R. *et al.*, *Angew. Chem. Int. Ed.*, **52**, 7228–32 (2013).

catalysts<sup>[178,179,180,181,182]</sup> and the rate-determining step of the reaction was suggested to be the protonation of the dihydrooxazine oxide species,<sup>[183]</sup> instead of the iminium or enamine nitronates. For example, an analogous cyclobutane intermediate was found in the reaction of butanal and nitrostyrene with H-D-Pro-Pro-Asp(OMe)-NH<sub>2</sub>.<sup>[175]</sup> On the other hand, even though they are not populated to a significant extent, dihydrooxazine and cyclobutane intermediates could not exclude H-D-Pro-Pro-Asp-NH<sub>2</sub> (**39**) as catalyst.<sup>[184]</sup>

Further studies demonstrated that the related sequence H-D-Pro-Pro-Glu-NH<sub>2</sub> (**40**) is also able to catalyse the nitro-Michael addition reaction between aldehydes and the nitroethylene as Michael acceptor.<sup>[185]</sup> The best result (yield  $\geq$  95%, ee  $\geq$  95%) was obtained after 15 hours at room temperature with 1 mol% of (**40**), 1 mol% of NMM, 1.5 equivalent of aldehyde and nitroethylene in chloroform. Recently, Rigling *et al.* investigated the conformational preferences of H-D-Pro-Pro-Glu-NH<sub>2</sub> (**40**) and its enamine intermediate (**A**) obtained by reaction with phenylacetaldehyde in a solution of CDCl<sub>3</sub>/CD<sub>3</sub>OH (9:1).<sup>[186]</sup> NMR data support the formation of a  $\beta$ -turn conformation with a H-bond between the carbonyl group of the D-Proline in position 1 and one amide proton of the C-terminal group. The side chain of the glutamic acid points towards the D-Pro-Pro moiety. Moreover, it was found that the enamine intermediate **A** also adopts a  $\beta$ -turn conformation, even though its overall folding appeared more flexible than H-D-Pro-Pro-Glu-NH<sub>2</sub> (**40**). The higher flexibility of the glutamic side chain and of the backbone in **A** was suggested to be essential to handle the incoming nitroolefin, which would allow a very stereoselective C-C bond formation.

#### iv. Structural modulation of the tripeptide catalyst

##### iv.a. Effect of ring-size analogues of H-D-Pro-Pro-Glu-NH<sub>2</sub>

In 2017, Wennemers and her team accessed structurally related catalysts of (**40**) with different *trans/cis* amide bond ratio, by synthesizing analogues bearing the four- and six-membered ring, respectively azetidine carboxylic acid (Aze) and piperidine (Figure 60).<sup>[187]</sup> This study was

<sup>178</sup> Patora-Komisarska, K. *et al.*, *Helv. Chim. Acta*, **94**, 719–45 (2011).

<sup>179</sup> Seebach, D. *et al.*, *Helv. Chim. Acta*, **95**, 1064–78 (2012).

<sup>180</sup> Burés, J. *et al.*, *J. Am. Chem. Soc.*, **133**, 8822–25 (2011).

<sup>181</sup> Burés, J. *et al.*, *J. Am. Chem. Soc.*, **134**, 6741–50 (2012).

<sup>182</sup> Burés, J. *et al.*, *J. Am. Chem. Soc.*, **134**, 14264–14264 (2012).

<sup>183</sup> Sahoo, G. *et al.*, *Angew. Chem. Int. Ed.*, **51**, 13144–48 (2012).

<sup>184</sup> Maillard, L. T. *et al.*, *ACS Omega*, **4**, 8862–73 (2019).

<sup>185</sup> Wiesner, M. *et al.*, *J. Am. Chem. Soc.*, **130**, 5610–11 (2008).

<sup>186</sup> Rigling, C. *et al.*, *J. Am. Chem. Soc.*, **140**, 10829–38 (2018).

<sup>187</sup> Schnitzer, T. *et al.*, *J. Am. Chem. Soc.*, **139**, 15356–62 (2017).

conducted to observe the influence of the *trans/cis* amide bond ratio on stereoselectivity. Based on NMR, they demonstrated that all the peptides adopt as similar  $\beta$ -turn with a hydrogen bond between the C=O of  $D$ -Pro and the C-terminal amide. However, whatever the considered solvent, the *trans/cis* amide bond ratio increased with the size of the central azacycle (Table 13, 3<sup>rd</sup> column).

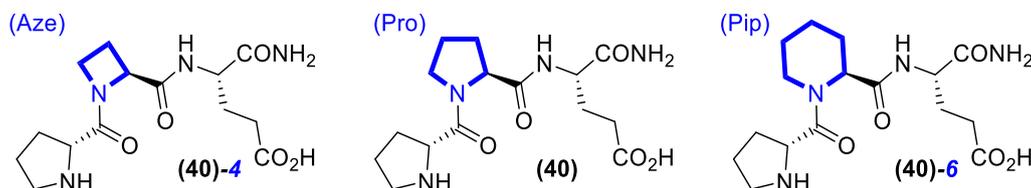


Figure 60:  $H$ - $D$ -Pro-Yaa-Glu- $NH_2$  catalysts

Table 13: Conjugate addition reaction of butanal to  $\beta$ -*trans*-nitrostyrene catalysed by  $H$ - $D$ -Pro-Yaa-Glu- $NH_2$  peptides<sup>[187]</sup>

Entry	Catalyst	$K_{t/c}$ <sup>a</sup>	Conv. <sup>b</sup> (%)	dr <sup>c</sup> (%)	ee <sup>c</sup> (%)
1	(40)-4	10	>95	26:1	92
2	(40)	46	>95	35:1	97
3	(40)-6	71	>95	50:1	99

<sup>a</sup>*Trans/cis* conformer ratio in  $CDCl_3/CD_3OH$  9:1, <sup>b</sup>Conversion determined by  $^1H$  NMR of the crude material  
<sup>c</sup>Determined by chiral stationary phase

Evaluation of the catalytic properties demonstrated a strong dependency of the diastereo- and enantioselectivities to the *trans/cis* ratio: the more the *trans* isomer is favoured, the better is the selectivity. In addition, since the *trans/cis* ratio is mainly related to the solvent (Figure 61), it is not surprising to find strong solvent effects on the reaction selectivity.<sup>[187]</sup>

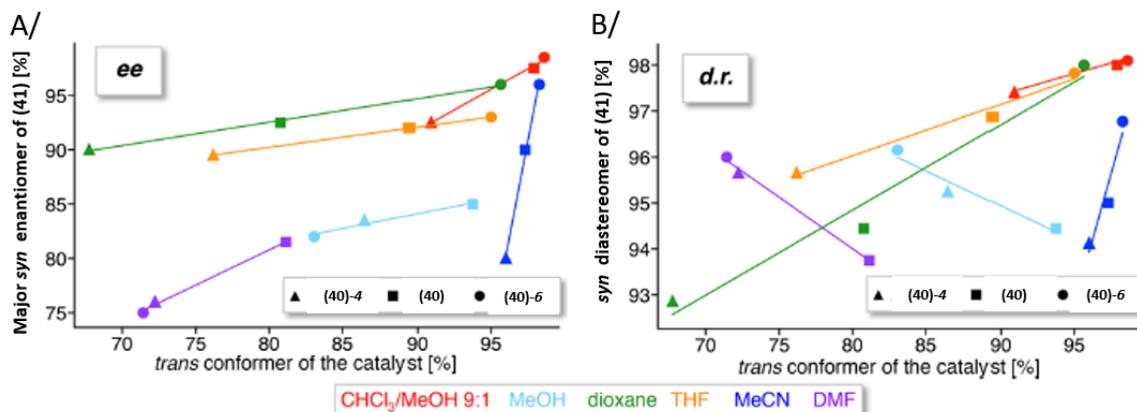
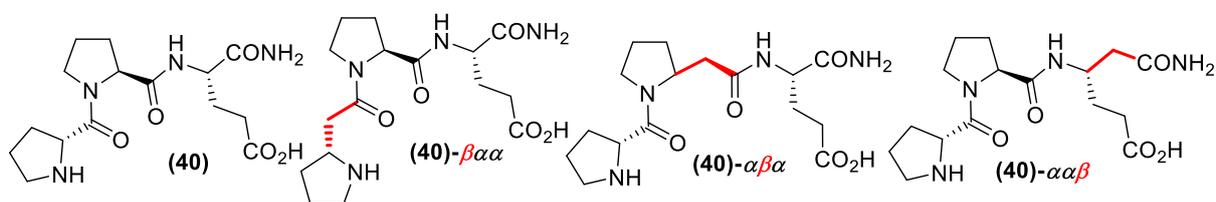


Figure 61: Correlation between the amount of *trans* conformer of peptides (40)-4, (40) and (40)-6 with their A/enantioselectivity and B/diastereoselectivity as examined in different solvents<sup>[187]</sup>

iv.b. Effect of  $\beta^3$ -amino acids analogues of H-D-Pro-Pro-Glu-NH<sub>2</sub>

Finally, in a recent publication, Wennemers and Schnitzer explored the effects of conformational flexibility on catalytic performances by synthesizing  $\beta^3$ -amino acid peptides analogues of **(40)** (Figure 62).<sup>[188]</sup> Based on computational analysis of the enamines derived from these four peptides, they observed that tripeptides **(40)- $\alpha\beta\alpha$**  and **(40)- $\alpha\alpha\beta$**  tend to have a similar orientation of their enamine and carboxylic acid groups when compared to the all- $\alpha$ -amino acids tripeptide **(40)**. Tested on the conjugate addition reaction of butanal to  $\beta$ -trans-nitrostyrene, these two analogues showed a lower reactivity and stereoselectivity than H-D-Pro-Pro-Glu-NH<sub>2</sub> **(40)** but the effect was not critical (Table 14, entries 1, 3 and 4). However, when the tripeptide bears an extra methylene group on its N-terminal side (**(40)- $\beta\alpha\alpha$** ), a rotation of the  $\beta^3$ -D-Pro residue occurred. This homologation leads to an important distance between the enamine and the carboxylic acid. In that case, the outcome is not too surprising; an increase of flexibility on the N-terminal proline can drive the nitroolefin to approach from both faces, hence it directs to a low enantioselectivity of the reaction (Table 14, entry 2). In conclusion, the catalytic properties of these tripeptides seem closely related to the balance between conformational flexibility and rigidity.

Figure 62: Analogues of H-D-Pro-Pro-Glu-NH<sub>2</sub> (**40**) bearing  $\beta^3$ -amino acids<sup>[188]</sup>Table 14: Peptide catalysed conjugate addition reaction of butanal to  $\beta$ -trans-nitrostyrene<sup>[188]</sup>

Entry	Catalyst	Conv. <sup>a</sup> (%)	dr <sup>b</sup> (%)	ee <sup>b</sup> (%)
1	<b>(40)</b>	quant.	50:1	97
2	<b>(40)-<math>\beta\alpha\alpha</math></b>	81	6:1	-13
3	<b>(40)-<math>\alpha\beta\alpha</math></b>	81	17:1	87
4	<b>(40)-<math>\alpha\alpha\beta</math></b>	94	24:1	90

<sup>a</sup>Determined by <sup>1</sup>H NMR of the crude material <sup>b</sup>Determined by chiral stationary phase. A minus indicates (2R,3S) enantiomer predominance

<sup>188</sup> Schnitzer, T. *et al.*, *Helv. Chim. Acta*, hlca.201900070 (2019).

## II. DESIGN AND SYNTHESIS OF COMPETENT ATCS FOR ENAMINE-TYPE CATALYSIS

### 1. ATC design

These insights, from the literature, provide a general guide for designing an enamine-type catalyst. It appeared that an ideal catalyst should combine two functionalities, *i.e.* a nucleophilic amine such as a pyrrolidine and a Brønsted acid. The average distance of the two key functions is of prime importance to establish the catalysed transformation. Finally, the chemo- and enantioselectivity of the reaction is drastically related to the dynamics of the folding and could be modulated by an appropriate choice of solvent.

Thus, the first objective of my thesis was to develop an ATC-monomer bearing a carboxylic acid on its  $\gamma$ -lateral chain and an azacycle on the position 2 of the thiazole ring. The general structure is presented in Figure 63. The chemical route ought to guarantee a strict control of the stereochemistry of the two asymmetric carbons.

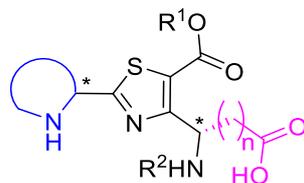


Figure 63: General structure of the catalytic ATC monomer

### 2. ATC synthesis

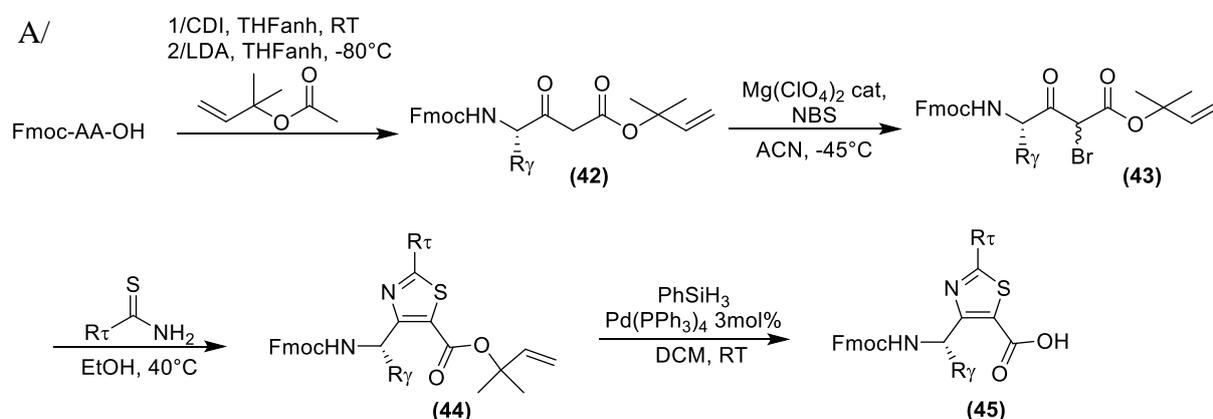
In 2016, Dr. L. Mathieu and Dr. C. Bonnel reported a robust ATC synthesis pathway providing an easy way to modulate the lateral chains, either on the  $\gamma$ -carbon atom or on the thiazole core of the amino acids.<sup>[189]</sup>

The synthetic strategy starts by a Cross-Claisen condensation between the *N*-Fmoc-amino acid and the 1,1-dimethylallyl acetate to access an *N*-Fmoc-amino- $\beta$ -keto ester intermediate (**42**). The use of a sterically hindered acetate is needed to preserve the orthogonality of the protections and to avoid any side reactions due to acetate autocondensation. Actually, the best conditions consist in the use of CDI (1.1 equiv.) and DMAP (3 mol%) to activate the *N*-protected-amino acid as an imidazolide intermediate, which is not isolated. The reaction is usually performed at

<sup>189</sup> Mathieu, L. *et al.*, *Eur. J. Org. Chem.*, 2262–70 (2015).

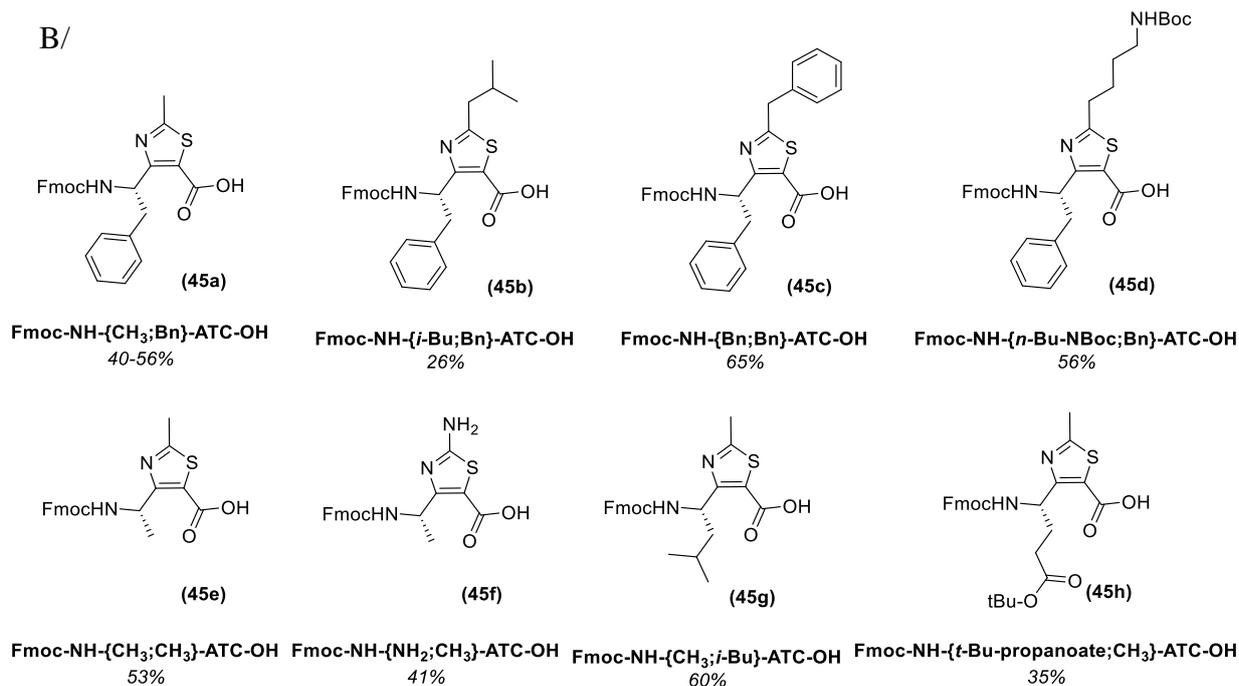
room temperature, in dry THF. In parallel, the enolate is generated from the 1,1-dimethylallyl acetate using LiHMDS as base. Then, the activated amino acid is added dropwise to the enolate. A strict temperature control, at  $-80^{\circ}\text{C}$  is necessary to prevent any Fmoc deprotection. The *N*-Fmoc-amino  $\beta$ -keto esters are classically obtained in ten grams scales with overall yields of 34–80%, depending on the selected *N*-Fmoc-amino acid starting material.<sup>[189]</sup> A monohalogenation at the malonic position of compound (**42**) is then performed to give the  $\alpha$ -bromo- $\beta$ -keto ester intermediate (**43**) (Scheme 9, A). The bromination is carried out under Yang *et al.* conditions,<sup>[190]</sup> in presence of a catalytic quantity of  $\text{Mg}(\text{ClO}_4)_2$  behaving as a Lewis acid, and *N*-bromosuccinimide (NBS), at  $-45^{\circ}\text{C}$ . A Hantzsch cyclisation using a thioamide reactant leads to the formation of the fully protected ATC (**44**). The bromine substitution, the cyclisation and the spontaneous aromatisation are performed in ethanol, at  $35^{\circ}\text{C}$ . A chemoselective deprotection of the acid function is finally done using a catalytic amount of palladium tetrakis triphenylphosphine and phenylsilane as nucleophile.

The *N*-Fmoc-ATC-monomer (**45**) is orthogonally protected on the amino group by the baso-labile Fmoc and on its lateral chain by an acido-labile group. Following this chemical route (Scheme 9), several monomers were synthesized with 26 to 65% yield. These monomers were used for the synthesis of a variety of foldamers described later in this manuscript.



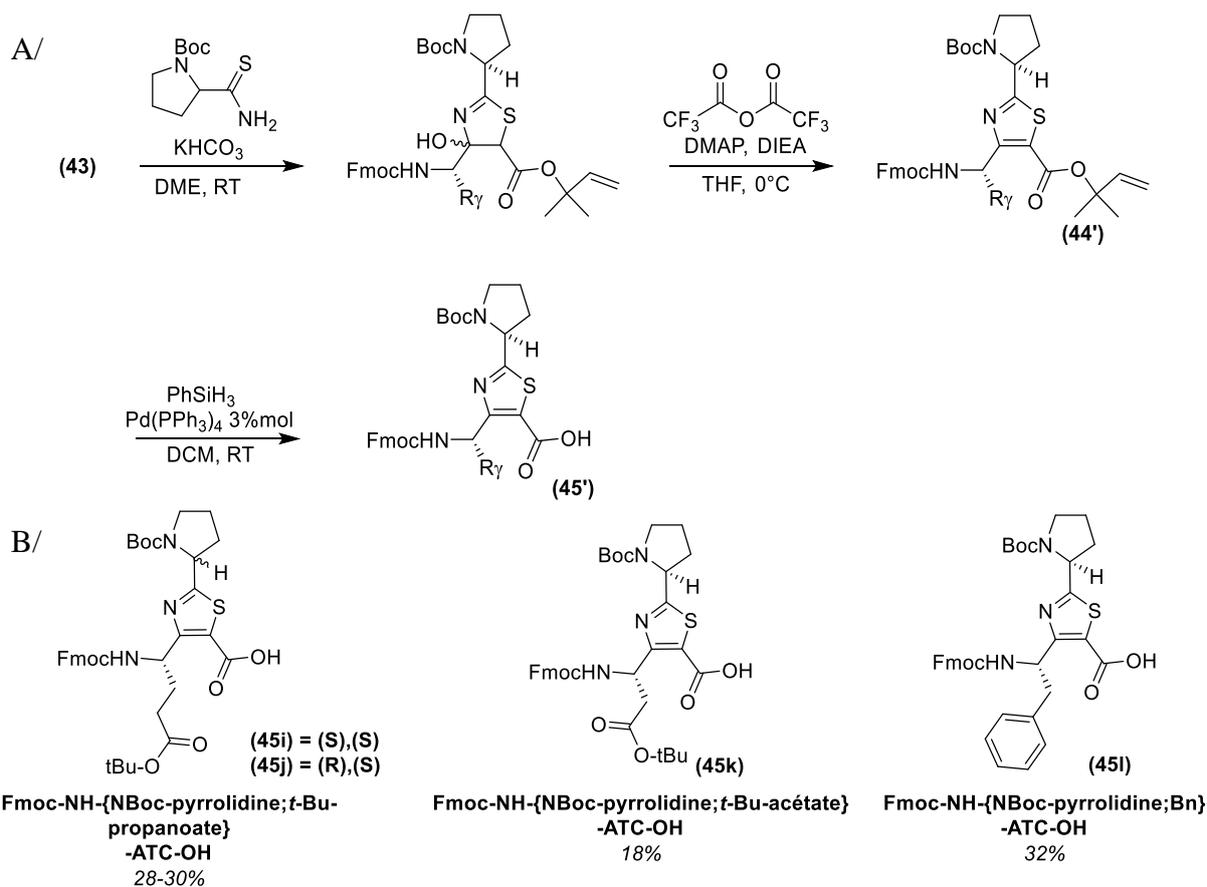
<sup>190</sup> Yang, D. *et al.*, *J. Org. Chem.*, **67**, 7429–31 (2002).

B/



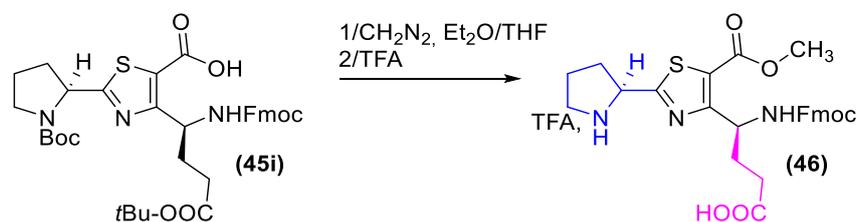
Scheme 9: A/Short route for the synthesis of ATC-monomers. B/Synthesized ATC-monomers, yields are calculated from both bromination and Hantzsch cyclisation steps

However, as represented here, none of the synthesized monomers bear an azacycle on the position 2 of the thiazole ring. Because of poor yields, we had to modify the Hantzsch conditions to introduce a pyrrolidine ring (Scheme 10). *N*-Boc-(*L* or *D*)-proline thioamide was prepared in two steps from *N*-Boc-(*L* or *D*)-proline. The thioamide reacted then with the  $\alpha$ -bromo- $\beta$ -keto ester (**43**) in presence of KHCO<sub>3</sub> as base. Under basic conditions, no aromatisation occurred and the resulting hydroxythiazoline had to be dehydrated using trifluoroacetic anhydride and DIEA at 0°C, leading to ATC (**44'**). The acid deprotection was then conducted under Tsuji-Trost conditions. Following this procedure, we synthesized the four ATCs (**45i-i**). NMR and chiral HPLC analyses were performed on ATC (**45i**) and (**45j**) in order to confirm the absence of epimerisation of the pyrrolidine ring (See SI, *section II*).



Scheme 10: A/Hantzsch cyclisation in two steps. B/Synthesized ATC-monomers via cyclisation then aromatisation, yields are calculated from both bromination and Hantzsch cyclisation steps

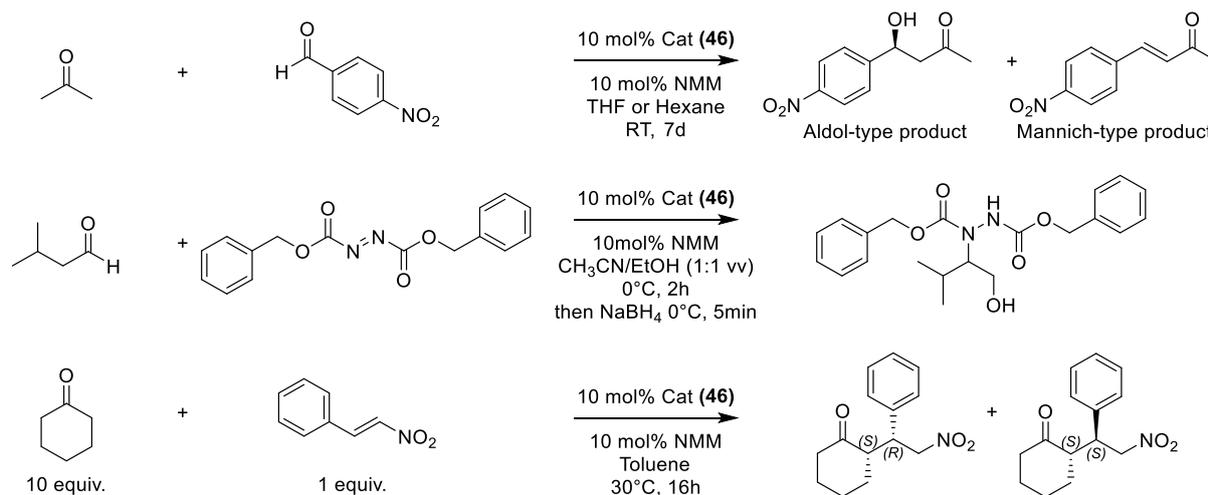
Finally, the *N*-Fmoc-ATC-OH (**45i**) was treated with diazomethane (Scheme 11). Deprotection of the pyrrolidine and  $\gamma$ -lateral chains by TFA led to **Fmoc-NH-{Pyrrolidine;propanoic acid}-ATC-OCH<sub>3</sub> (46)** which corresponds to the general structure expected for a bifunctional catalytic ATC.



Scheme 11: Access to Fmoc-NH-{Pyrrolidine;propanoic acid}-ATC-OCH<sub>3</sub> (**46**)

### 3. Evaluation of the catalytic properties of the bifunctional ATC-monomer (46)

**Fmoc-NH-{Pyrrolidine;propanoic acid}-ATC-OCH<sub>3</sub> (46)** was first evaluated at 10 mol% in model enamine reactions: the cross condensation of acetone with *p*-nitrobenzaldehyde<sup>[153]</sup>, the amination of isovaleraldehyde with dibenzylazodicarboxylate<sup>[191]</sup> and the Michael addition reaction of cyclohexanone to  $\beta$ -*trans*-nitrostyrene (Scheme 12).<sup>[161]</sup>



Scheme 12: Explored model reactions using (46) as catalyst

In the first case, using 10 mol% catalyst, less than 28% *p*-nitrobenzaldehyde was converted after 6 hours (66% after 7 days), and the reaction led to the Mannich-type product as a major species instead of the desired aldol-type compound (See SI, section IV). With the addition of isovaleraldehyde to dibenzylazodicarboxylate, the conversion rate was 67% after 2 hours (94% after 3.5 days) but no asymmetric induction was observed at 10 mol% catalyst (See SI, section IV). The effectiveness of ATC (46) was finally established in the nitro-Michael addition reaction of cyclohexanone to  $\beta$ -*trans*-nitrostyrene (Scheme 12). Initially performed in toluene at 30°C (Table 15, entry 1), the reaction led to a yield of 54% after 16 hours, with a good diastereoselectivity towards the *syn* product (*syn/anti* ratio 95/5) and a modest, but significant, enantiomeric excess of 27% in favour of the product with a ((*S*),(*R*)) configuration, determined by HPLC and optical rotation. This reaction has the advantage of being widely studied in organocatalysis since the resulting  $\gamma$ -nitroketones possess two stereogenic centres and can be

<sup>191</sup> Kumaragurubaran, N. *et al.*, *J. Am. Chem. Soc.*, **124**, 6254–55 (2002).

used for the synthesis of more complex molecules, such as chiral  $\gamma$ -amino acids,  $\gamma$ -butyrolactams or polyketide natural products.<sup>[176,192,193,194]</sup>

#### 4. Optimization of the nitro-Michael addition reaction catalysed by (46)

Having established the effectiveness of **Fmoc-NH-{Pyrrolidine;propanoic acid}-ATC-OCH<sub>3</sub> (46)** in the nitro-Michael addition reaction of cyclohexanone to  $\beta$ -*trans*-nitrostyrene, we then explored the reaction conditions. We have first tested the influence of the temperature and the solvent on the yield, the *syn/anti* diastereoselectivity and the enantioselectivity (Table 15). The reaction was performed using 10 equivalents of cyclohexanone, 1.0 equivalent of  $\beta$ -*trans*-nitrostyrene and 10 mol% of catalyst (46) as TFA salt and 10 mol% *N*-methylmorpholine as base. The yields were **estimated** by HPLC at 214nm.

Table 15: Optimization of nitro-Michael addition reaction, using catalyst (46)

Entry <sup>a</sup>	Solvent	T (°C)	Yield <sup>b</sup> (%)	<i>syn/anti</i>	ee <sup>c</sup> (%)
1	Toluene	30	54	95/5	27
2	DMSO	30	82	87/13	29
3	CHCl <sub>3</sub>	30	81	93/7	27
4	THF	30	80	94/6	28
5	CH <sub>3</sub> CN	30	80	92/8	22
6	Hexane	30	88	94/6	31
7	MeOH	30	18	95/5	nd
8	EtOH	30	98	92/8	34
9	<i>i</i> PrOH	30	92	96/4	41
10	<i>i</i> PrOH	4	98	96/4	45

<sup>a</sup>Reaction performed over 16 hours with 0.1mmol nitroolefin, 1mmol cyclohexanone, 0.01mmol NMM and 0.01mmol catalyst in 400  $\mu$ L solvent; <sup>b</sup>Determined via RP-HPLC at 214 nm; <sup>c</sup>Determined via chiral HPLC (CHIRALPAK IC column; Hex/*i*PrOH 8:2 v/v, *d*=1 ml/min,  $\lambda$ =214 nm). Selectivity towards (2*S*)-2-[(1*R*)-2-Nitro-1-phenylethyl]cyclohexanone determined by optical rotation and comparison with reported data in literature<sup>[195]</sup>.

As highlighted in Table 15 (entries 2 to 6), when the reaction was performed in dimethyl sulfoxide, chloroform, tetrahydrofuran, acetonitrile or else hexane, the yield raised up to more

<sup>192</sup> Belot, S. *et al.*, *Adv. Synth. Catal.*, **352**, 667–95 (2010).

<sup>193</sup> Roca-Lopez, D. *et al.*, *Tetrahedron Asymmetry*, **21**, 2561–2601 (2010).

<sup>194</sup> Marcia de Figueiredo, R. *et al.*, *Eur. J. Org. Chem.*, **2007**, 2575–2600 (2007).

<sup>195</sup> Xu, Y. *et al.*, *Adv. Synth. Catal.*, **348**, 418–24 (2006).

than 80%. However, the asymmetric induction was not significantly increased, reaching a limit of 31%. Methanol (Table 15, entry 7) was the solvent for which the conversion rate obtained was the lowest. This was quite an unexpected result, especially since under the same reaction conditions, ethanol led to a yield of 98% with an enantiomeric excess of 34%. Supported by the latest result, *isopropanol* was tested (Table 15, entries 9–10), giving the best yield with an ee of 41% at 30°C. By decreasing the temperature to 4°C, the asymmetric induction was increased, reaching 45% ee with a yield of 98%.

To conclude, we designed a bifunctional ATC, which, as proline or Wennemer's tripeptides, bears a nucleophilic pyrrolidine and a Brønsted acid lateral chain. After a slight modification of the chemical procedure, we were able to prepare the catalyst on a multigram scale. We established its effectiveness on the nitro-Michael addition reaction of cyclohexanone and  $\beta$ -*trans*-nitrostyrene and optimized the reaction conditions. At 10 mol% of **Fmoc-NH-{Pyrrolidine;propanoic acid}-ATC-OCH<sub>3</sub> (46)** in *i*PrOH at 4°C, the 2-[2-Nitro-1-phenylethyl]cyclohexanone was obtained in 98% yield with a selectivity towards the *syn* product (dr > 95:5) and an enantiomeric excess of 45%. The next step of the project was to introduce the catalytic monomer in ATC oligomers to explore how a folded  $\gamma$ -peptide environment may modulate the catalytic properties.

CHAPTER 3:  
CATALYTIC PROPERTIES ACCORDING TO  
THE OLIGOMER LENGTHS AND THE  
POSITION OF THE COMPETENT ATC



## I. NOMENCLATURE USED FOR THE ATC-OLIGOMERS

Most of the oligomers described in this section alternate two ATC monomers, consisting on the **Fmoc-NH-{Pyrrolidine;propanoic acid}-ATC-OH (45i)** and **Fmoc-NH-{CH<sub>3</sub>;Bn}-ATC-OH (45a)**. The sequences are named based on the following nomenclature: **ATC<sub>L</sub>-P-X**, where **L** is related to the length of the oligomer (3 for a trimer, 5 for a pentamer, and so on), **P** corresponds to the position along the chain of the ATC unit bearing the pyrrolidine group, from the N-Ter to the C-terminal extremity. **X** pertains to the C-terminal extremity (for example NH<sub>i</sub>Pr or NH<sub>2</sub>).

The pyrrolidine group in the ATC-oligomers is usually set in the (*S*) configuration, however, when the (*R*) configuration is used, a star (\*) is added after the position number (**P**), such as **ATC<sub>L</sub>-P\*-X** represents an ATC-oligomer bearing a pyrrolidine group of (*R*) configuration.

Finally, if the ATC-oligomer contains one residue that differs from **(45a)** and/or **(45i)** (eventually **(45j)**) when the pyrrolidine is in the (*R*) configuration), it will be designated as a mutant, thus will be called **M(n)-ATC<sub>L</sub>-P-X** where **n** is a number in order of appearance in the manuscript.

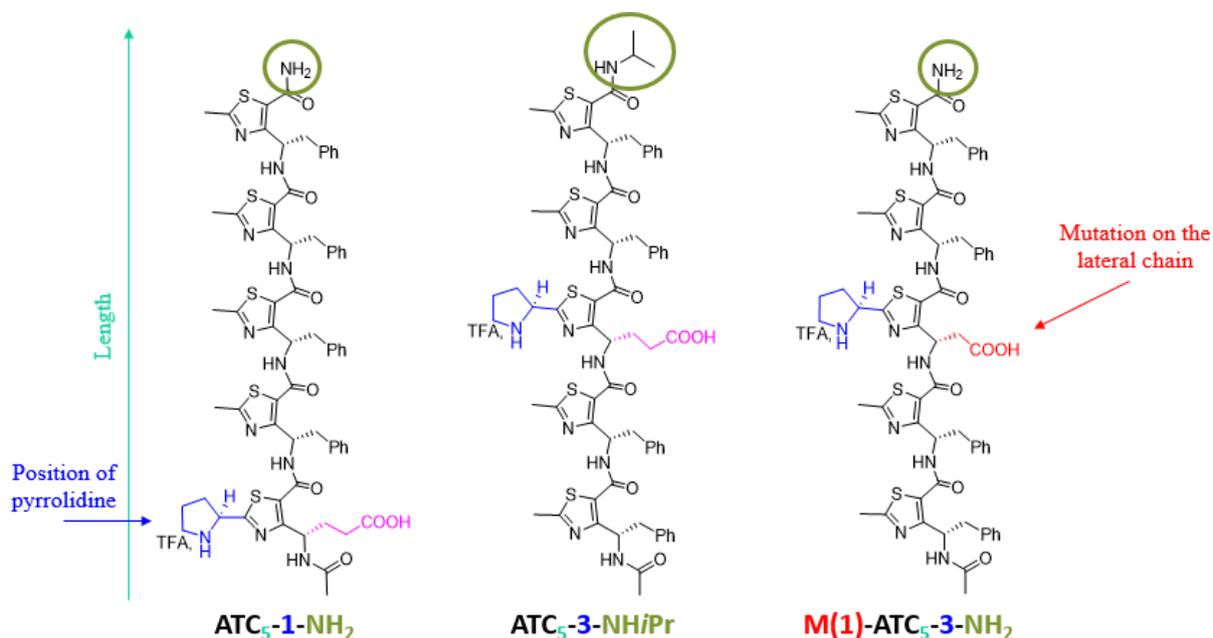


Figure 64: Illustration of this nomenclature

## II. DESIGN AND SYNTHESIS OF FOLDED ATC-BASED CATALYSTS

We initially explored the relationship existing between the oligomer length and the catalytic properties. Therefore, we designed five ATC-based catalysts with an increasing number of ATC units. We first attempted to rigidify the monomeric catalyst (**46**) by introducing an isopropyl amide group to its C-terminal extremity and acetyl to its N-terminal extremity (Figure 65, compound (**47**)). Based on a previous study conducted within our laboratory,<sup>[104]</sup> we expected a stiffening of the ATC-scaffold resulting from a C<sub>9</sub> intramolecular H-bond between the acetyl carbonyl and *i*PrNH. (**47**) was used as a reference catalyst. Then, we inserted the catalytic motif inside an ATC-trimer called **ATC<sub>3</sub>-2-NHiPr** and an ATC-pentamer named **ATC<sub>5</sub>-3-NHiPr**. We believed that whatever the sequence, the oligomers would still conserve a C<sub>9</sub>-helical folding.

We also designed the trimer **ATC<sub>3</sub>-2\*-NHiPr**, for which the pyrrolidine chiral centre was in the opposite configuration than in **ATC<sub>3</sub>-2-NHiPr** (Figure 65). The aim was to dissociate the effect of the folding to those from the pyrrolidine configuration on the asymmetric induction. To test the role of the acidic lateral chain, we also prepared the methyl ester **ATC<sub>3</sub>-2(OMe)-NHiPr**. Finally, we planned the synthesis of the pentamer **ATC<sub>5</sub>-3-NH<sub>2</sub>** differing from **ATC<sub>5</sub>-3-NHiPr** at the C-terminal extremity. The main goal was to explore the impact of the C-terminal extremity on both the folding and the catalytic properties of the oligomers.

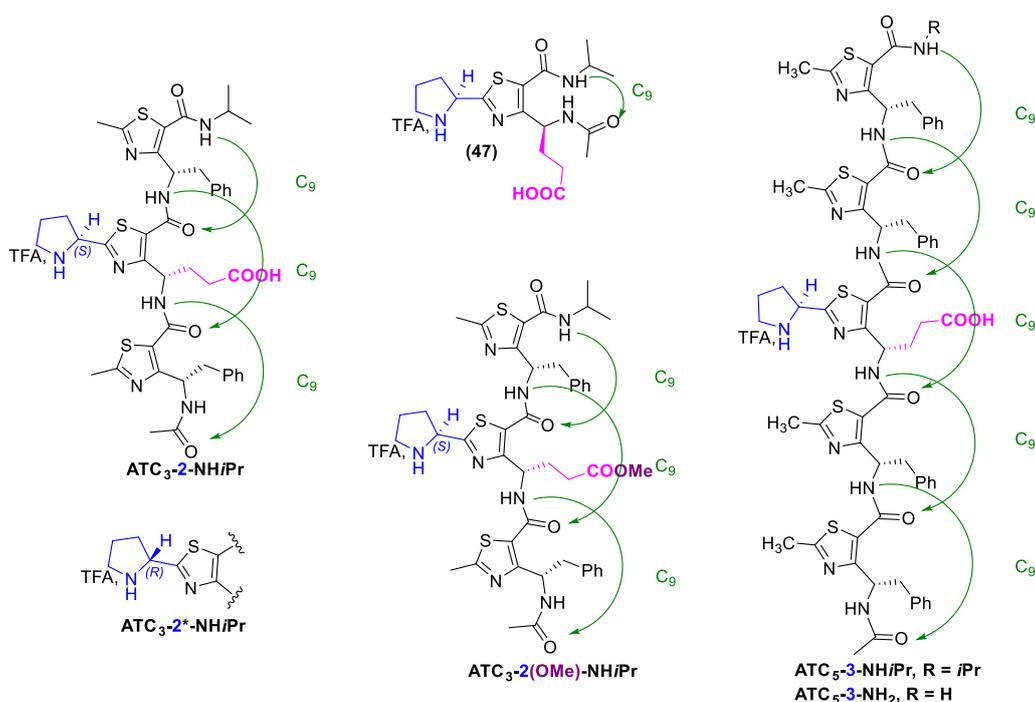
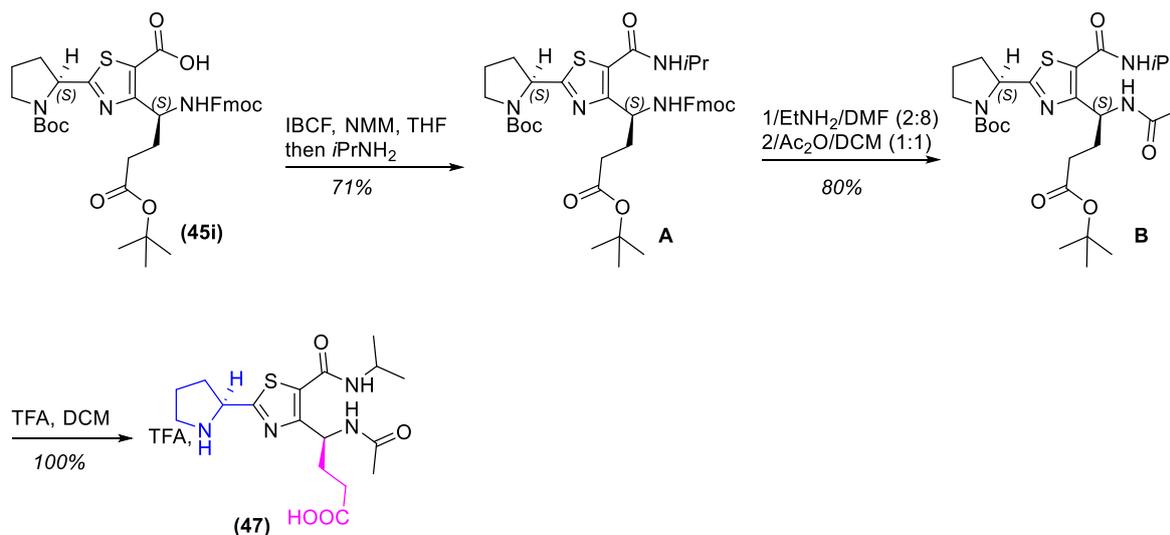


Figure 65: Structured ATC-monomer (**47**), trimers **ATC<sub>3</sub>-2-NHiPr**, **ATC<sub>3</sub>-2\*-NHiPr** and **ATC<sub>3</sub>-2(OMe)-NHiPr** and pentamers **ATC<sub>5</sub>-3-NHiPr** and **ATC<sub>5</sub>-3-NH<sub>2</sub>**

### 1. Synthesis of (47)

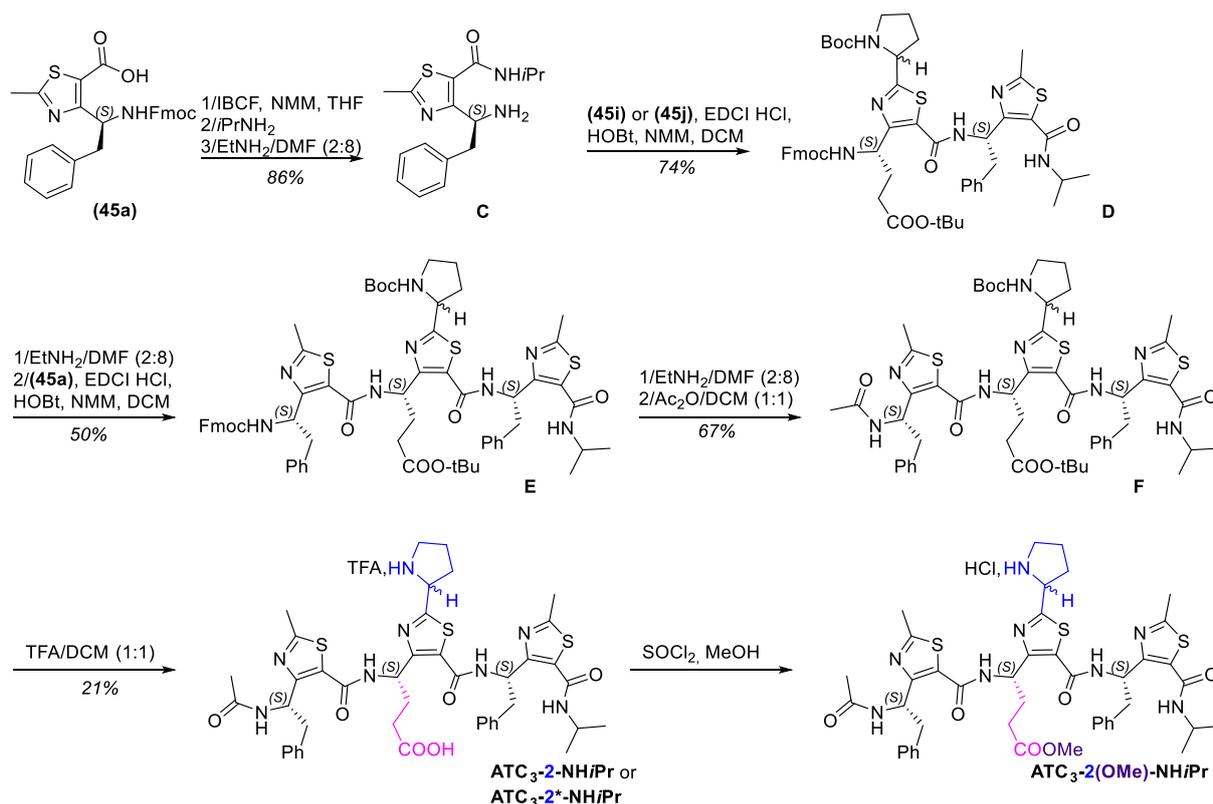
The **Ac-NH-{Pyrrolidine;propanoic acid}-ATC-NHiPr (47)** was prepared from **(45i)** as depicted in Scheme 13. After activation by *isobutylchloroformate* (*i*BCF) and DIEA, the acid function was converted into *isopropylamide*. Treatment with diethylamine in DMF led to terminal amine deprotection which was then acetylated with acetic anhydride in DCM. Finally, the pyrrolidine and the propanoic acid lateral chains were deprotected by TFA (Scheme 13).



Scheme 13: Access to **Ac-NH-{Pyrrolidine;propanoic acid}-ATC-NHiPr (47)**

### 2. Synthesis of the ATC-trimers **ATC<sub>3</sub>-2-NHiPr**, **ATC<sub>3</sub>-2\*-NHiPr** and **ATC<sub>3</sub>-2(OMe)-NHiPr**

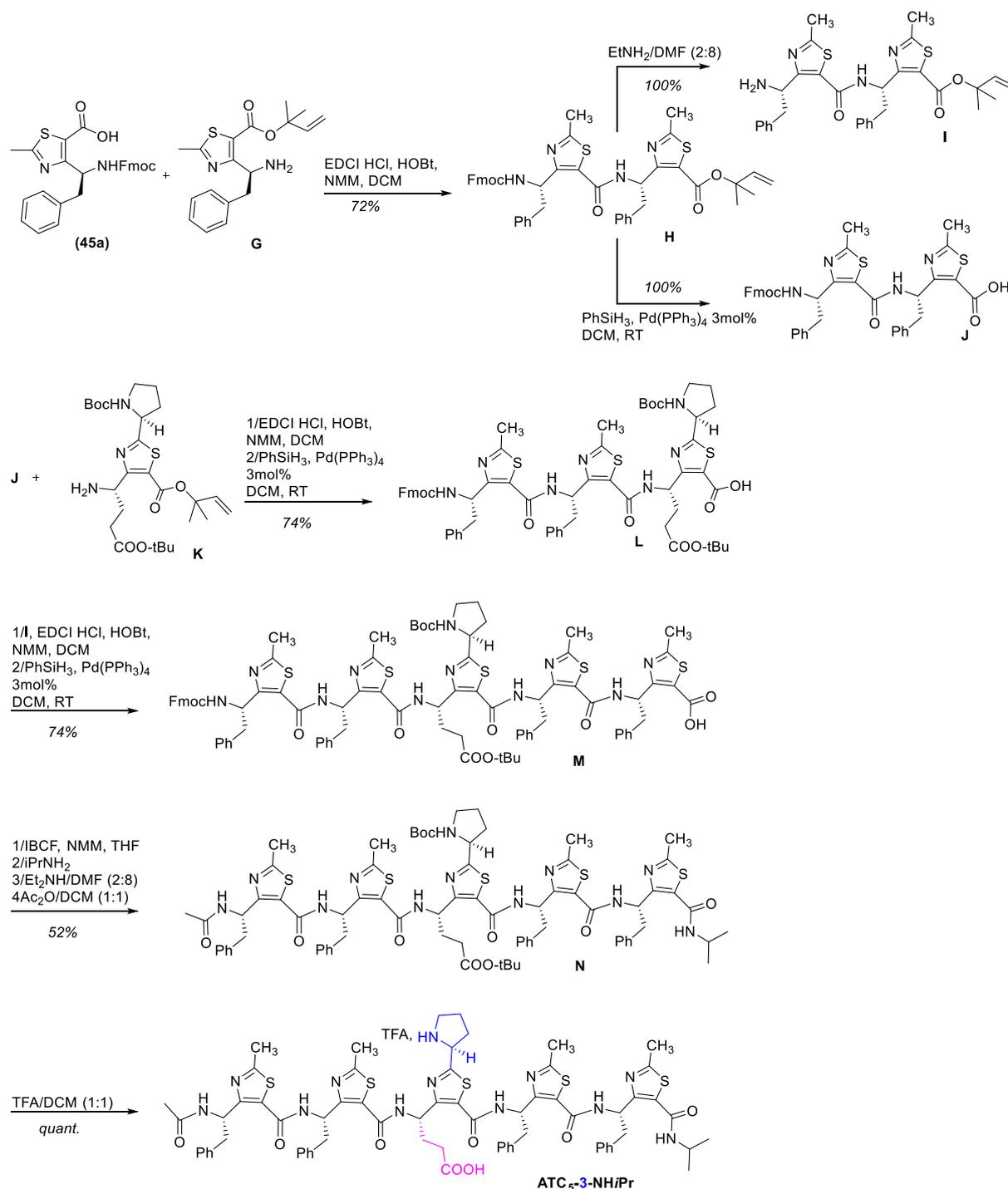
The two diastereomers **ATC<sub>3</sub>-2-NHiPr** and **ATC<sub>3</sub>-2\*-NHiPr** were synthesized in solution, following a step-by-step strategy, using a mix of EDCI and HOBt (1:1) as coupling agents. During the procedure, the Fmoc groups were removed using a solution of diethylamine in DMF (2:8 v/v) (Scheme 14). The trimers, **ATC<sub>3</sub>-2-NHiPr** and **ATC<sub>3</sub>-2\*-NHiPr**, were obtained with an overall yield of respectively 34% and 21%. Access to the methyl ester analogue **ATC<sub>3</sub>-2(OMe)-NHiPr** was finally performed by treating **ATC<sub>3</sub>-2-NHiPr** with SOCl<sub>2</sub> in methanol.



Scheme 14: Synthesis of  $ATC_3\text{-2-NHiPr}$ ,  $ATC_3\text{-2}^*\text{-NHiPr}$  and  $ATC_3\text{-2(OMe)-NHiPr}$

### 3. Synthesis of the ATC-pentamer $ATC_5\text{-3-NHiPr}$

In contrast, the pentamer  $ATC_5\text{-3-NHiPr}$  was prepared following a convergent fragment approach (Scheme 15). As the first step, the monomer (**44a**) was quantitatively deprotected from its Fmoc group with a solution of diethylamine in DMF (2:8 vv) to obtain compound **G**. Thereafter, a general coupling between monomers (**45a**) and **G** was performed using a mix of EDCI and HOBT (1:1) as coupling agents to lead to dimer **H**. Half of the amount of **H** was treated with  $Et_2NH/DMF$  (2:8 vv) to remove the Fmoc group (compound **I**, 100% yield), the other half was deprotected on its acid group (compound **J**, 100% yield) under Tsuji-Trost conditions. **J** was coupled with the free amine group of **K** using EDCI and HOBT (1:1). The trimer **L** was isolated with 74% yield. After removal of the dimethylallyl group, it was coupled with the dimer fragment **I**. Deprotection of the acid function lead to pentamer **M** (yield = 74%). This oligomer **M** was then activated by *isobutylchloroformate* (*iBCF*) and DIEA, to convert the acid function into *isopropylamide*. Treatment with diethylamine in DMF led to terminal amine deprotection which was then acetylated with acetic anhydride in DCM. Finally, the pyrrolidine and the propanoic acid lateral chains were deprotected by TFA.  $ATC_5\text{-3-NHiPr}$  was isolated with 22% overall yield.

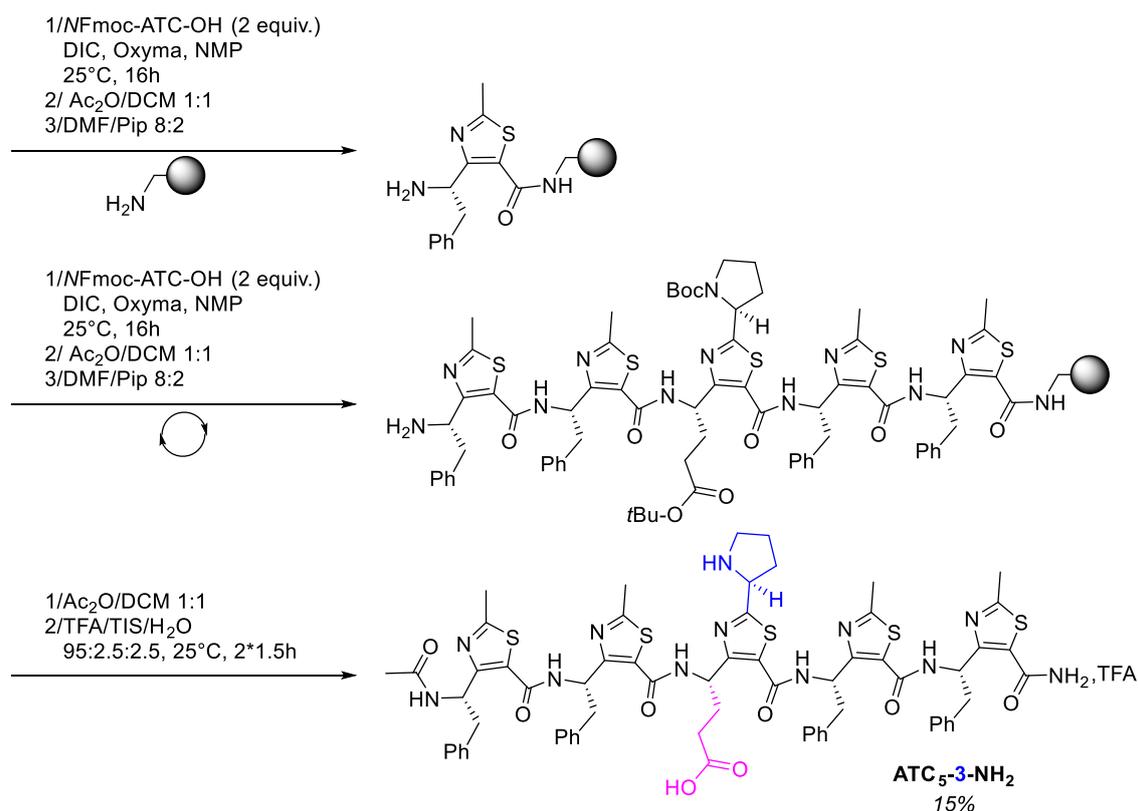


Scheme 15: Solution synthesis of **ATC<sub>5</sub>-3-NHiPr**

#### 4. Synthesis of ATC-pentamer **ATC<sub>5</sub>-3-NH<sub>2</sub>**

The synthesis of **ATC<sub>5</sub>-3-NH<sub>2</sub>** was guided by two main objectives. As previously underlined, the first one was to explore the impact of the C-terminal extremity on both the folding and the catalytic properties of the oligomers. The second objective was to test the feasibility of a Fmoc/*t*Bu Solid Phase Peptide Synthesis (SPPS) approach for the preparation of catalytic

oligomers. In fact, because of the simplified procedures after coupling and deprotection steps, SPPS permits the parallel synthesis of multiple ATC-oligomers in a more convenient way than a solution phase strategy. Thus, **ATC<sub>5</sub>-3-NH<sub>2</sub>** was synthesized step-by-step on a Rink amide ChemMatrix resin, following the procedure optimized by Dr. Bonnel three years ago.<sup>[196]</sup> Each coupling is effective after one night, in NMP (20 mL per gram of resin), using two equivalents of Fmoc-ATC-OH monomer, 2.0 equiv. of DIC and 2.0 equiv. of OxymaPure®, in NMP (20 mL per gram of resin). Other coupling reagents, such as BOP, PyBOP, HBTU, HATU and T3P had previously been evaluated. All had resulted in an incomplete reaction (around 70% yield). In the case of **ATC<sub>5</sub>-3-NH<sub>2</sub>**, each coupling was monitored by a Kaiser test to detect the presence of free amines and was followed by an acetylation step. Deprotections of the Fmoc groups were performed in three times ten minutes in a 20% piperidine solution in DMF. Cleavages were achieved, after acetylation of the N-terminal extremity by acetic anhydride, using a solution of TFA/TIS/H<sub>2</sub>O 95/2.5/2.5 vvv, for three hours (Scheme 16). The peptide was finally purified by RP-HPLC, with a yield of 15%.



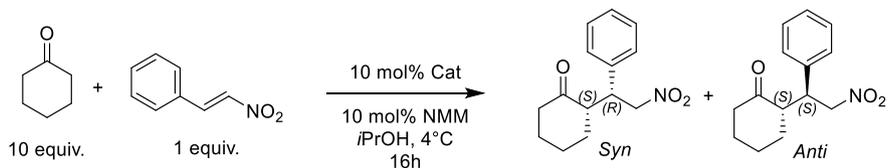
Scheme 16: Solid Phase Peptide Synthesis of **ATC<sub>5</sub>-3-NH<sub>2</sub>**

<sup>196</sup> Bonnel, C., *Thesis*, Oligopeptides Construits autour du  $\gamma$ -Aminoacide ATC : Syntheses, Analyses Structurales et Evaluation Biologique (2016).

### III. COMPARISON OF THE CATALYTIC PROPERTIES WITH THE ATC NUMBER UNITS AND THE C-TER EXTREMITY

The catalytic properties of each molecule were determined on the nitro-Michael addition reaction of cyclohexanone to  $\beta$ -*trans*-nitrostyrene, using the conditions that we had previously determined, *i.e.* 10 mol% catalyst, *i*PrOH as solvent at 4°C over 16 hours. While the **Ac-NH-{Pyrrolidine;propanoic acid}-ATC-NH*i*Pr (47)** seems less bulky than the **Fmoc-NH-{Pyrrolidine;propanoic acid}-ATC-OCH<sub>3</sub> (46)**, the asymmetric induction was slightly increased (ee = 45% and 52% for **(46)** and **(47)** respectively). Nonetheless, even though the conversion rate was excellent (> 99%), the selectivity towards the compound *syn* was a bit lower with **(47)** (*syn/anti*= 96/4 and 93/7 for **(46)** and **(47)** respectively) (Table 16, entries 1-2). The trimer **ATC<sub>3</sub>-2-NH*i*Pr** and pentamer **ATC<sub>5</sub>-3-NH*i*Pr** retained a good catalytic activity (yields = 98% and 93% respectively) associated with an excellent *syn/anti* selectivity (Table 16, entries 3 and 4). The asymmetric induction was significantly increased to 74% for the trimer **ATC<sub>3</sub>-2-NH*i*Pr** and 77% for the pentamer **ATC<sub>5</sub>-3-NH*i*Pr**. Finally, it can be noted that **ATC<sub>5</sub>-3-NH*i*Pr** and **ATC<sub>5</sub>-3-NH<sub>2</sub>** produced similar results, demonstrating no real impact of the C-ter capping (H or *i*Pr) on the catalysis (Table 16, entries 4-5).

Table 16: Comparison of different ATC-catalysts in the nitro-Michael addition reaction between cyclohexanone and  $\beta$ -*trans*-nitrostyrene



Entry <sup>a</sup>	Catalyst	Yield <sup>b</sup> (%)	<i>syn/anti</i> ratio <sup>c</sup>	ee <sup>d</sup> (%)
1	<b>(46)</b>	92	96/4	45
2	<b>(47)</b>	94	93/7	52
3	<b>ATC<sub>3</sub>-2-NH<i>i</i>Pr</b>	91	>99/1	74
4	<b>ATC<sub>5</sub>-3-NH<i>i</i>Pr</b>	>99	>99/1	77
5	<b>ATC<sub>5</sub>-3-NH<sub>2</sub></b>	>99	>99/1	79
6	<b>ATC<sub>3</sub>-2*-NH<i>i</i>Pr</b>	93	98/2	65 <sup>e</sup>
7	<b>ATC<sub>3</sub>-2(OMe)-NH<i>i</i>Pr</b>	<5	nd	nd

<sup>a</sup>Reaction performed over 16 hours with 0.1mmol nitroolefin, 1mmol cyclohexanone, 0.01mmol NMM and 0.01mmol catalyst in 400  $\mu$ L solvent; <sup>b</sup>Determined via RP-HPLC at 214 nm; <sup>c</sup>Determined via chiral HPLC (CHRALPAK IC column; Hex/*i*PrOH 8:2 v/v,  $d=1$  ml/min,  $\lambda=214$  nm). <sup>d</sup>((R),(S)) enantiomer was obtained as major compound.

As observed by Wennemers' group, we also found by comparing the trimers **ATC<sub>3</sub>-2-NH<sub>i</sub>Pr** and **ATC<sub>3</sub>-2\*-NH<sub>i</sub>Pr** (Table 16, entries 3 and 6), that the pyrrolidine configuration mainly orients the chiral transfer, with the latter catalyst, the ((*R*),(*S*)) enantiomer was obtained as major compound. Finally, a total loss of reactivity was observed with **ATC<sub>3</sub>-2(OMe)-NH<sub>i</sub>Pr**, demonstrating the crucial role of the acid group in the catalytic activity (Table 16, entry 7).

These results showed a dependency of the diastereo- and enantioselectivity of the reaction within the oligomer length. This also suggested an impact of the secondary structure on the catalytic properties which remained to be demonstrated by exploring the conformational behaviour of the molecules.

#### IV. EVALUATION OF THE OVERALL FOLDING OF THE CATALYSTS

The structure elucidation of the oligomers described within this document is based on the interpretation of several spectroscopic methods. Among them, the nuclear magnetic resonance (NMR) and the X-ray crystallography (XRC), which enable to determine the structure on an atomic level, respectively in solution- and solid-state. This will be the object of the last chapter of this manuscript. However, some NMR observables and other complementary methods, such as circular dichroism (CD) or Fourier-transform infrared spectroscopy (FTIR), can be used to give qualitative information on the global folding of the molecules.

##### *1. Characteristic structural markers for ATC-based helical foldamers*

The aim of this paragraph is to give to the reader an overview of the structural markers related to the C<sub>9</sub>-helical folding of the ATC oligomers.

##### **i. Circular dichroism**

Studying macromolecules by circular dichroism allows obtaining information about their folding. Circular dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light, which arise due to structural asymmetry. Dichroic spectra may be difficult to interpret because the dichroic signature depends not only on the structure of the object, but also on the nature of the monomeric units of which it is composed. Thus, CD is probably the most rapid method for evaluating the folding of non-natural oligomers, including ATC-based  $\gamma$ -peptides, provided that the spectrum pattern must be carefully established based on high resolution structural data.

The CD signature for the C<sub>9</sub>-helical folding of the ATC oligomers in solvents such as methanol and water was described in 2013.<sup>[103]</sup> Whatever the solvent considered, the CD spectra displayed two minima, at around 195 and 230 nm, and a strong maximum at 263 nm. The two minima were assigned respectively to the  $\pi$ - $\pi^*$  and  $n$ - $\pi^*$  transitions of the amide chromophore, while the maximum could be attributed to the thiazole core. We generally observe a blue shift for the  $n$ - $\pi^*$  transition (226 nm) in aqueous solution certainly resulting from a stronger solute-solvent hydrogen bond interaction and/or to the higher water polarizability. A molar ellipticity value per residue (mre) around 80 000 deg.cm<sup>2</sup>.dmol<sup>-1</sup> is characteristic of the 9-helix (Figure 66).

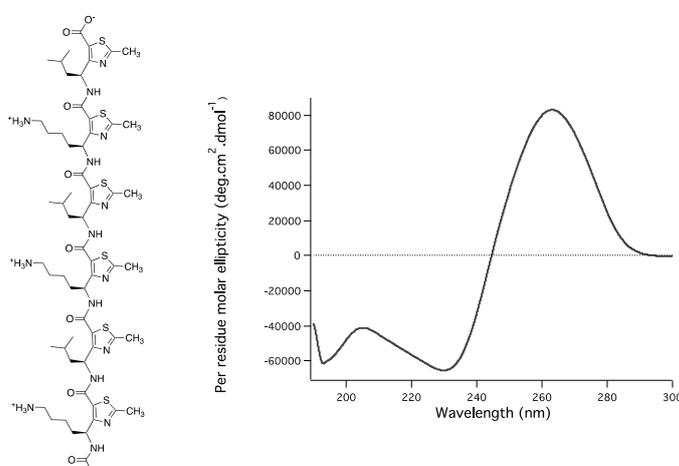


Figure 66: Representative CD spectrum for an ATC-oligomer folding in 9-helix<sup>[103]</sup>

## ii. Nuclear Magnetic Resonance structural markers

Nuclear Magnetic Resonance (NMR) spectroscopy is the second method used to explore the folding properties of ATC-oligomers.

Sequential attribution of protons in ATC polypeptides is performed by the combination of multi-dimensional experiments providing homonuclear <sup>1</sup>H/<sup>1</sup>H <sup>n</sup>J-couplings and heteronuclear <sup>13</sup>C/<sup>1</sup>H <sup>n</sup>J-couplings (Table 17).<sup>[197,198]</sup>

<sup>197</sup> Diehl, B., in *NMR Spectrosc. Pharm. Anal.*, pp. 1–41 (Elsevier, 2008).

<sup>198</sup> Wagner, G. *et al.*, *Eur. J. Biochem.*, **114**, 375–84 (1981).

Table 17: Most common scalar and bidimensional experiments

<b>Homonuclear through-bond correlation methods</b>			
<i>Name</i>	<i>Nuclei 1</i>	<i>Coupling</i>	<i>Nuclei 2</i>
COSY	<sup>1</sup> H	<sup>2</sup> J and <sup>3</sup> J (between geminal and vicinal protons)	<sup>1</sup> H
TOCSY	<sup>1</sup> H	<i>J, J, J...</i> (multiple transfer, up to 3 bonds)	<sup>1</sup> H
NOESY/ROESY	<sup>1</sup> H	NOE (dipolar interactions between H up to 5 Å)	<sup>1</sup> H
<b>Heteronuclear through-bond correlation methods</b>			
<i>Name</i>	<i>Nuclei 1</i>	<i>Coupling</i>	<i>Nuclei 2</i>
HSQC	<sup>13</sup> C or <sup>15</sup> N	<sup>1</sup> J <sub>CH</sub> and <sup>1</sup> J <sub>NH</sub>	<sup>1</sup> H
HMQC	<sup>13</sup> C	<sup>1</sup> J <sub>CH</sub>	<sup>1</sup> H
HMBC	<sup>13</sup> C	<sup>2</sup> J <sub>CH</sub> and <sup>3</sup> J <sub>CH</sub>	<sup>1</sup> H

In 2016, our laboratory published a study related to NMR structural markers for ATC-containing molecules.<sup>[104]</sup> Among them, it was recognized that a high deshielding effect ( $\delta > 9.0$  ppm) constitutes a confident <sup>1</sup>H NMR observable, representative of an ATC NH engaged in a C<sub>9</sub>-intramolecular hydrogen bond. In the ATC C<sub>9</sub>-helix, only the N-terminal NH is not engaged in a H-bond and appears at  $\delta = 8.0$  to 8.5 ppm.

The <sup>3</sup>J(NH,<sup>γ</sup>CH) values constitute a second indicator of the C<sub>9</sub>-helical folding. The coupling constants can be derived from the <sup>1</sup>H (or from the 2D COSY<sup>[199]</sup>). Vicinal coupling constants <sup>3</sup>J(NH,<sup>γ</sup>CH) are used for the deduction of the possible values for the torsion angle  $\theta$  (H-N-C<sub>γ</sub>-H<sub>γ</sub>), using the Karplus equation:

$$^3J = A \cdot \cos^2(\theta) - B \cdot \cos(\theta) + C$$

where A, B and C are empiric values (which differ depending on the authors<sup>[200,201,202]</sup>). Usually, the ATC-oligomer folding is preferably described by the torsion angle  $\varphi$  (H-N-C<sub>γ</sub>-C<sub>β</sub>) instead of  $\theta$  (H-N-C<sub>γ</sub>-H<sub>γ</sub>). Though, the two torsion angles are linked by the equation  $\theta = |\varphi - 60|$ . The graph relative to the Karplus equation for the  $\varphi$  (H-N-C<sub>γ</sub>-C<sub>β</sub>) torsion angle is drawn Figure 67, using the values given by Pardi *et al* (A = 6.4, B = 1.4 and C = 1.9).<sup>[203]</sup> When ATC-oligomers

<sup>199</sup> Marion, D. *et al.*, *Biochem. Biophys. Res. Commun.*, **113**, 967–74 (1983).

<sup>200</sup> Cung, M. T. *et al.*, *Macromolecules*, **7**, 606–13 (1974).

<sup>201</sup> Ramachandran, G. N. *et al.*, *Biopolymers*, **10**, 2113–31 (1971).

<sup>202</sup> Bystrov, V. F., *Prog. Nucl. Magn. Reson. Spectrosc.*, **10**, 41–82 (1976).

<sup>203</sup> Pardi, A. *et al.*, *J. Mol. Biol.*, **180**, 741–51 (1984).

adopt the C<sub>9</sub>-helical folding, the φ angles are around -78°, making the <sup>3</sup>J(NH,γCH) values in the range of 6.0 to 7.0 Hz.

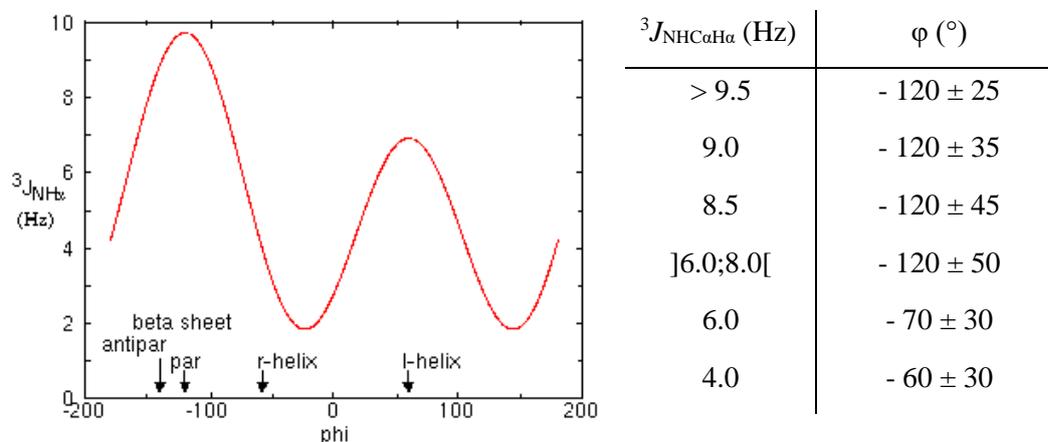


Figure 67: The empirically-derived Karplus relationship between the vicinal three-bond coupling constant <sup>3</sup>J(NH,γCH) and φ

Finally, NOE correlations are expected if the corresponding interproton distances are lower than 5.5 Å. The integration of NOE cross-peaks enables to define interatomic distances, which can be used afterwards to provide 3D-models of the studied molecule. Based on the previously reported crystal structure of ATC oligomers, seven sequential NOEs were recognized for the C<sub>9</sub>-helix conformational preference (Figure 68). Even if not all the correlations are always observed on the spectra, three recurrent NOEs are generally enough to lead to a convergent structure during computational calculations.

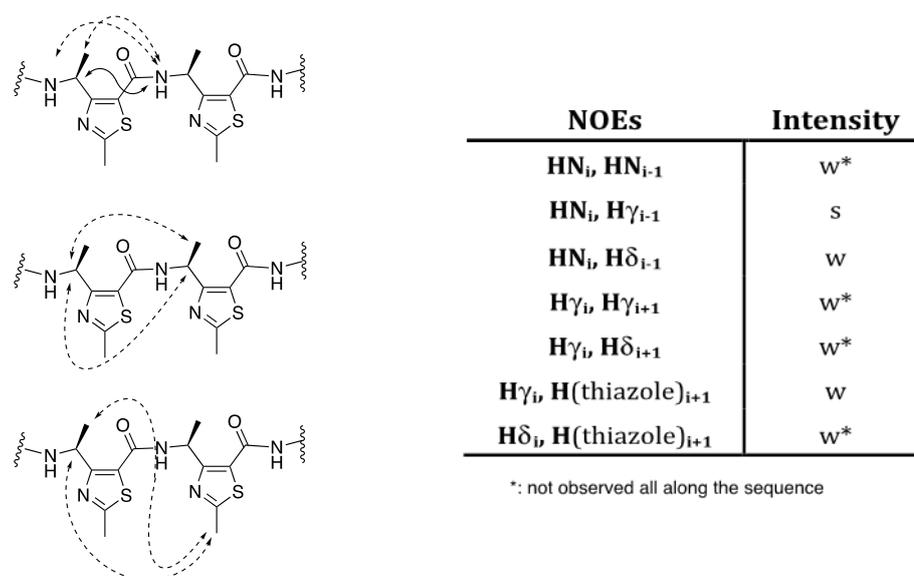


Figure 68: Recurrent NOE cross-peaks in an ATC-oligomer

### iii. Mid infrared spectroscopy

Mid infrared spectroscopy is one of the most effective methods to establish the folding as well as the amide hydrogen bonding state in peptides and pseudo-peptide molecules.<sup>[204,205]</sup> The amide groups of polypeptides and proteins possess nine characteristic vibrational modes or group frequencies, named A, B, and from I to VII.<sup>[206]</sup> Three of them, Amide A, I and II are the most explored in the infrared spectra of proteins because these arise from the amide bonds that link the amino acids.<sup>[207]</sup> Frequencies of Amide A (mainly  $\nu(\text{NH})$ ), Amide I (dominated by  $\nu(\text{CO})$ ) and Amide II (mix of NH bending and CN stretching) vibrations, which are highly influenced by the state (free or bound) and the strength of hydrogen bonds involving amide groups, may provide unambiguous structural assignments of the H-bonding states of the oligomers.

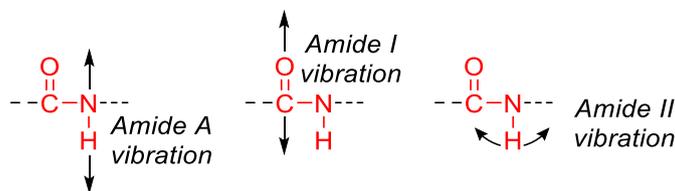


Figure 69: The vibrations responsible for the Amide A, I and II bands in the infrared spectra of proteins

In peptides, as well as in ATC-oligomers, the Amide A band corresponding to the NH stretching is located between  $3450$  and  $3200\text{ cm}^{-1}$ , making experiments in water or methanol usually quite hard to interpret as they strongly absorb in this spectral region. Nonetheless, the frequency associated with this band typically depends on whether the NH is involved in the establishment of an intramolecular hydrogen bond (absorption band around  $3450\text{--}3400\text{ cm}^{-1}$ ) or not (around  $3300\text{--}3250\text{ cm}^{-1}$ ). For example, a change in hydrogen bonding distance of only  $0.002\text{ \AA}$  shifts the wavenumber of the Amide A band of a polypeptide by approximately  $1\text{ cm}^{-1}$ .<sup>[208]</sup> However, it was demonstrated, from the ATC-oligomers, a strong deviation of the molar absorption coefficient between the free and bound Amide A bands. Consequently, Amides A are not considered as quantitative markers of the H-bonding state of ATC-oligomers.

The absorption associated with the Amide I band is due to the stretching vibrations of the C=O bond of the amide, coupled with in-plane NH bending and CN stretching modes (Figure 69).

<sup>204</sup> Byler, D. M. *et al.*, *Biopolymers*, **25**, 469–87 (1986).

<sup>205</sup> Haris, P. I. *et al.*, *Biopolymers*, **37**, 251–63 (1995).

<sup>206</sup> Surewicz, W. K. *et al.*, *Biochim. Biophys. Acta*, **952**, 115–30 (1988).

<sup>207</sup> Vass, E. *et al.*, *Chem. Rev.*, **103**, 1917–54 (2003).

<sup>208</sup> Braiman, M. S. *et al.*, *Annu. Rev. Biophys. Biophys. Chem.*, **17**, 541–70 (1988).

This band is located between 1800 and 1600  $\text{cm}^{-1}$ . For instance,  $\nu(\text{CO})$  and  $\nu(\text{NH})$  are downshifted when hydrogen bonds are reinforced. Based on a careful comparison of ATC-oligomers and other 1,3-thiazole-containing molecules, it was established that a bounded ATC carbonyl leads to a  $\nu(\text{CO}) \approx 1618\text{-}1630 \text{ cm}^{-1}$ , while a free carbonyl gives a band at  $\nu(\text{CO}) > 1645 \text{ cm}^{-1}$ . By contrast to Amide A, the Amide I bands in ATC-molecules are particularly relevant as the molar absorption coefficients are similar for the free and the bound C=O. Decomposition of the spectrum into individual Gaussian functions and quantitative analysis of each contribution enables to estimate the number of carbonyls involved in a hydrogen-bonding pattern (Figure 70).

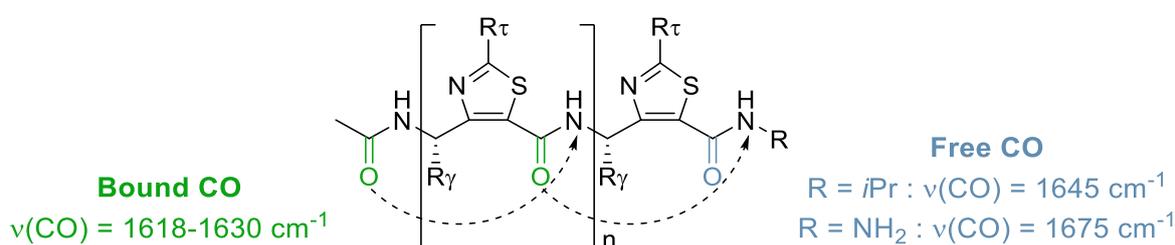


Figure 70: Representation of the carbonyls involved or not in a hydrogen-bonding pattern for an ATC-oligomer

Finally, the Amide II band, which corresponds primarily to the NH bending and CN stretching vibrations (Figure 69), is positioned around 1575-1480  $\text{cm}^{-1}$ . If the band is shifted towards the highest wavenumber of the region, it implies that the proton of NH is involved in a hydrogen bond. However, since the Amide II bands overlap the stretching bands of the thiazole rings, they are not suitable to the study of the folding of ATC-based peptides.

Fourier Transform InfraRed (FTIR) brings structural indications but the resolution with such method is not always optimum, especially in the Amide I region, as that can lead to misinterpretation due to a superposition of the absorption bands. Consequently, FTIR experiments must be correlated with other spectroscopic techniques to provide a bundle of irrefutable structural markers.

## 2. Structural markers associated to the ATC-catalysts

### i. Global folding of (47)

We started to explore the folding behaviour of the *N*-acetyl-ATC-carboxamide (**47**). We had hypothesized that the higher enantiomeric excess observed with the *N*-acetyl-ATC-carboxamide (**47**) compared to the ATC catalyst (**46**) might be a result of a stiffening of the

ATC scaffold. Such increase in the rigidity of the scaffold could be explained by a C<sub>9</sub> intramolecular H-bond between the acetyl carbonyl and the *i*PrNH in (47). The intramolecular hydrogen bond was supported by the strong chemical shift ( $\delta = 9.64$  ppm) of the *i*PrNH extremity.

## ii. Global folding of ATC<sub>3</sub>-2-NHiPr, ATC<sub>5</sub>-3-NHiPr and ATC<sub>5</sub>-3-NH<sub>2</sub>

The folding properties, in the solution state, of oligomers ATC<sub>3</sub>-2-NHiPr, ATC<sub>5</sub>-3-NHiPr and ATC<sub>5</sub>-3-NH<sub>2</sub> were then explored by combining circular dichroism, NMR and FTIR. Isopropanol was chosen as solvent because it was considered as the best one for the catalysis.

### ii.a. Circular dichroism signatures

All compounds shared comparable CD signatures with a strong positive band centred at 263 nm and a negative maximum at 232 nm. These signatures indicated that all sequences had adopted a similar conformational preference, consistent with the above-described right 9-helix structure for ATC oligomers.<sup>[104,209]</sup>

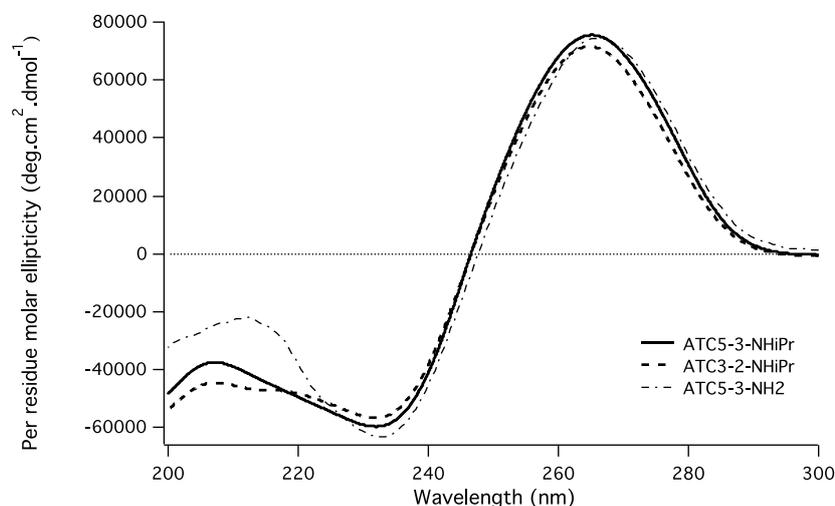


Figure 71: Overlay of the far-UV circular dichroism spectra of compounds ATC<sub>3</sub>-2-NHiPr, ATC<sub>5</sub>-3-NHiPr and ATC<sub>5</sub>-3-NH<sub>2</sub>. All experiments were performed at 4°C with a 50  $\mu$ M peptide solution in *i*PrOH

<sup>209</sup> Aguesseau-Kondrotas, J. *et al.*, *Chem. – Eur. J.*, **25**, 7396–7401 (2019).

ii.b.  $^1\text{H}$  NMR indicators

The  $^1\text{H}$  NMR spectra of **ATC<sub>3</sub>-2-NHiPr**, **ATC<sub>5</sub>-3-NHiPr** and **ATC<sub>5</sub>-3-NH<sub>2</sub>** were recorded in *d*<sub>7</sub>-*i*PrOH and nearly all  $^1\text{H}$  resonances were assigned, combining COSY, TOCSY and ROESY spectra.

As clearly visible on the  $^1\text{H}$  NMR spectrum of **ATC<sub>5</sub>-3-NHiPr** (Figure 72, A), we observed strong NH deshielding ( $> 9.8$  ppm) as well as  $^3J(\text{NH},\gamma\text{CH})$  values  $\approx 6.5$  Hz (Figure 72) previously recognized as a structural marker related to the formation of the poly-ATC C<sub>9</sub>-helix pattern.<sup>[104,107]</sup> The less deshielded proton was attributed to the N-terminal ATC NH, an indication that this hydrogen was not engaged in such hydrogen bond network. Similar NH behaviours were observed for the trimer **ATC<sub>3</sub>-2-NHiPr**.

On the ROESY spectrum of **ATC<sub>5</sub>-3-NHiPr** (Figure 72, B), canonical correlations between the H $\gamma$  of the residues *i* and the NH of the residues *i*+1 were detected. Some other correlations between H $\gamma$  of the residues *i* and the protons of the lateral chains worn by the thiazole rings of the residues *i*+1 were also observed (Figure 72, C). Once again, these features, also appearing on the ROESY spectrum of **ATC<sub>3</sub>-2-NHiPr**, are characteristic of a C<sub>9</sub>-helical structure.

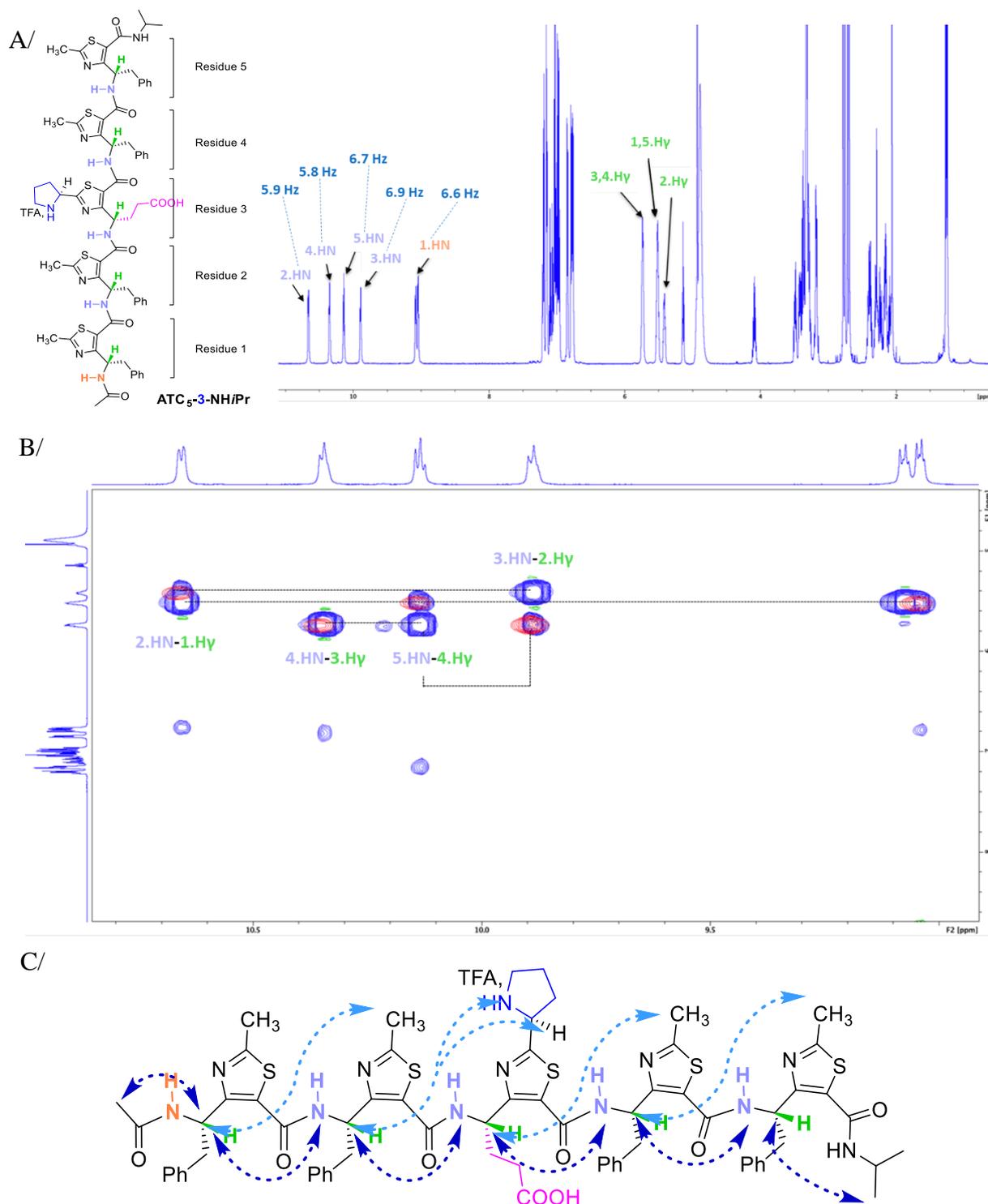


Figure 72: A/ $^1\text{H}$  NMR spectrum of ATC<sub>5</sub>-3-NHiPr in CD<sub>3</sub>OH at 293K, with the coupling constants  $^3\text{J}(\text{NH}, \gamma\text{CH})$  in purple. B/Superimposition of ROESY (blue) and TOCSY (red) spectra of ATC<sub>5</sub>-3-NHiPr in CD<sub>3</sub>OH at 293K. C/Representation of the observable NOE cross-peaks.

In the case of ATC<sub>5</sub>-3-NH<sub>2</sub>, we observed one major conformer and two minor species contributing to 82 %, 16% and 2%, suggesting that ATC<sub>5</sub>-3-NH<sub>2</sub> is more dynamic than ATC<sub>5</sub>-3-NHiPr. However, as for the other oligomers, most of the amide protons were highly deshielded which was consistent with the previously set hypothesis of intramolecular hydrogen

bonds forming C<sub>9</sub> pseudocycles (Figure 73). We also detected weak NH<sub>*i*</sub>/δCH<sub>*i*-1</sub> and γCH<sub>*i*</sub>/δCH<sub>*i*+1</sub> sequential NOE connectivities characteristic of the ATC helix, all along the backbone.

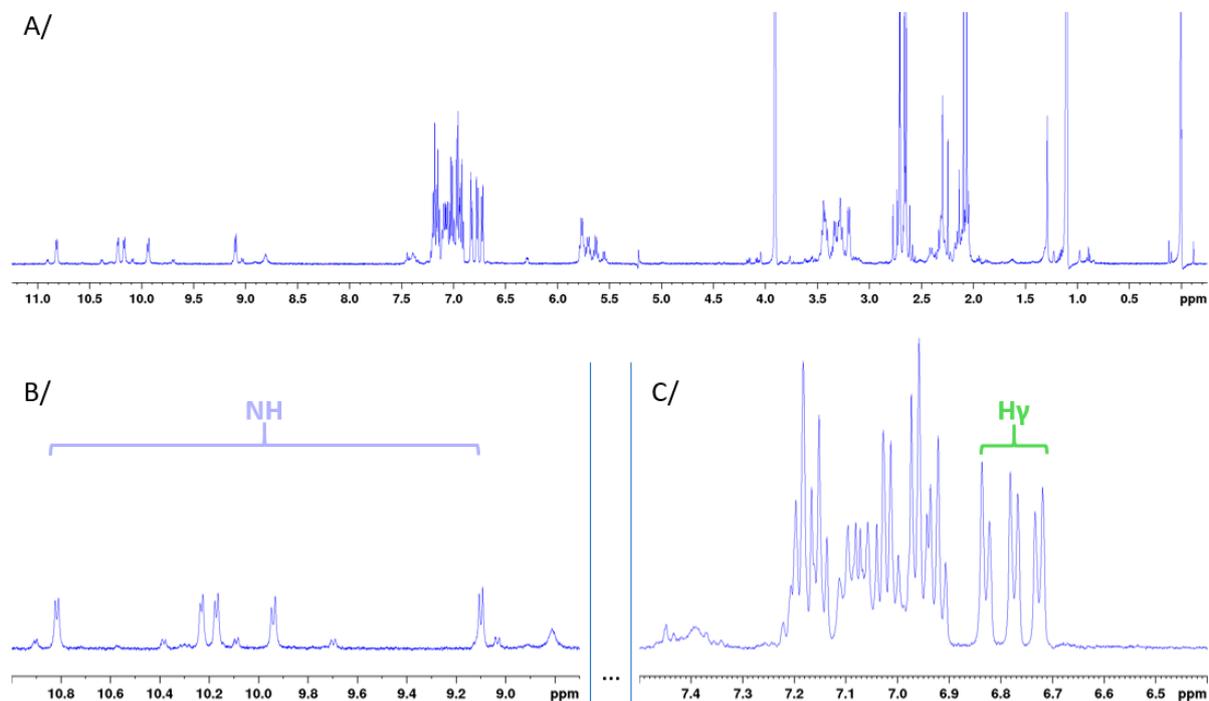


Figure 73: <sup>1</sup>H NMR of ATC<sub>5</sub>-3-NH<sub>2</sub>: A/full spectrum, B/NH zone: zoom between 8.0 and 11.0 ppm, C/Hγ zone: zoom between 6.4 and 7.5 ppm

### ii.c. Fourier transform infrared indicators

It was decided to conduct FTIR experiments on ATC-oligomers in *isopropanol* at room temperature. Considering that the Amide I frequencies (1600-1800 cm<sup>-1</sup>) were formerly recognized as structural markers of the C<sub>9</sub> H-bond network for ATC-containing oligomers in a previous work performed in the laboratory,<sup>[104,107]</sup> thus, we focused our efforts on the interpretation of signals in this area.

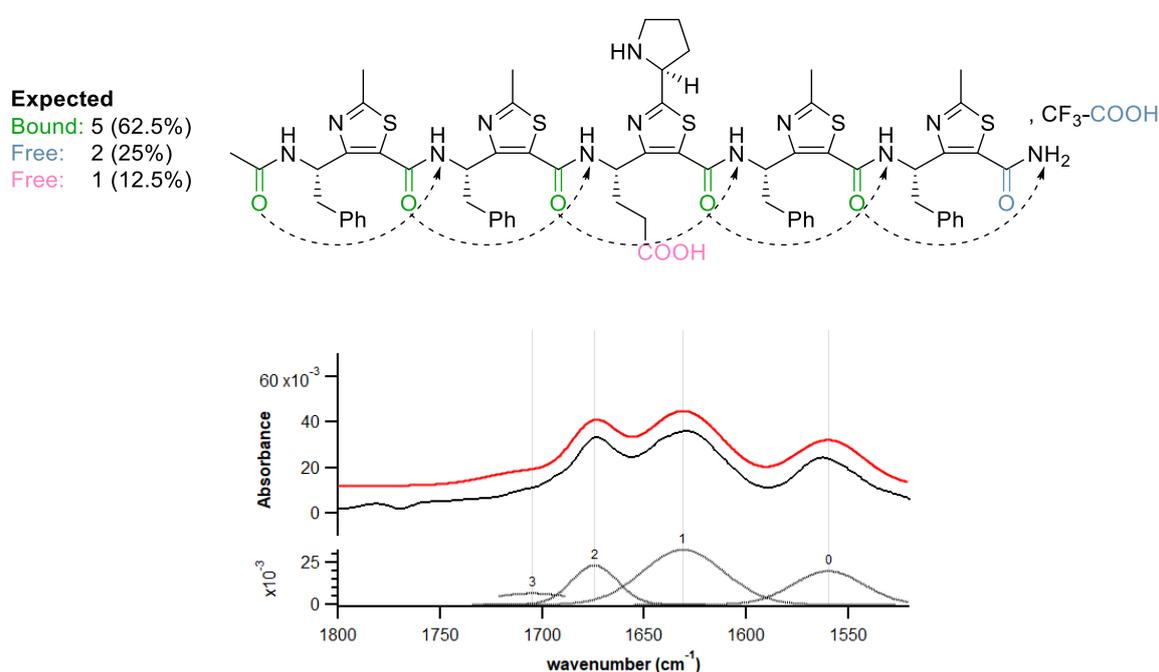
In the 9-helix model, oligomer ATC<sub>5</sub>-3-NH<sub>2</sub> should display five bounded CO corresponding to the acetyl group and to ATCs **1-4** and two free carbonyls resulting from the ATC **5** and the acid lateral chain COOH. In addition, since ATC<sub>5</sub>-3-NH<sub>2</sub> was obtained as a salt, a band corresponding to the carbonyl of the TFA was also expected.

The FTIR spectrum exhibited three absorption components in ν(CO) region (Table 18). The shoulder contribution at 1705-1710 cm<sup>-1</sup> is tied to the free acidic lateral chain as reported in the literature.<sup>[210]</sup> Based on the C<sub>9</sub>-helix model and on our previously reported studies on ATC

<sup>210</sup> Max, J.-J. *et al.*, *J. Phys. Chem. A*, **108**, 3324–37 (2004).

oligomers,<sup>[104,107]</sup> the band at 1674 cm<sup>-1</sup> should correspond to a combination of the C-terminal carboxamide (ATC 5) and of the TFA counter ions.<sup>[211]</sup> Finally, the low frequency band at 1631 cm<sup>-1</sup> was attributed to five bounded CO, *i.e.* the acetyl group at the N-terminus of the sequence, as well as to the four carbonyls of ATCs **1-4**, which are engaged in the H-bond network. After deconvolution of the spectrum into individual Gaussian functions (Table 18), the experimental ratio of free and H-bonded NHs was 3/5 in accordance with the C<sub>9</sub>-helix model.

Table 18: FTIR spectrum of ATC<sub>5</sub>-3-NH<sub>2</sub> at 2 mM in iPrOH. Band positions, intensities, widths, and areas in the region of 1800-1500cm<sup>-1</sup> are reported in the table



Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1631	0.0322	27.7	1.580	62 %
<b>Contribution 2</b> (free CO)	1674	0.0251	16.0	0.712	28 %
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1705	0.0068	20.0	0.240	10 %

<sup>211</sup> Roux, S. *et al.*, *J. Pept. Sci.*, **14**, 354–59 (2008).

In conclusion, all the CD, NMR and FTIR data were consistent with the folding model. Despite that, based on  $^1\text{H}$  NMR spectra, **ATC<sub>5</sub>-3-NH<sub>2</sub>** appeared more dynamic than the *isopropylamide* terminated counterpart **ATC<sub>5</sub>-3-NH*i*Pr**, but should likely adopt the expected C<sub>9</sub>-topology. In addition, as previously underlined, the capping of the C-terminal extremity did not influence the reactivity nor the enantiomeric excess in the nitro-Michael addition reaction. Consequently, the following studies were performed on oligomers with a primary amide, being more readily available than the NH*i*Pr homologues.

## V. MODULATION OF THE POSITION OF THE CATALYTIC CENTRE

### 1. *Synthesis of ATC-pentamers ATC<sub>5</sub>-P-NH<sub>2</sub> (1 ≤ P ≤ 5)*

We then explored the impact of the position of the catalytic **NH<sub>2</sub>-{Pyrrolidine;propanoic acid}-ATC-OH** residue along the sequence. The length was fixed at five ATC units. Thus, the oligomers, so called **ATC<sub>5</sub>-1-NH<sub>2</sub>**, **ATC<sub>5</sub>-2-NH<sub>2</sub>**, **ATC<sub>5</sub>-3-NH<sub>2</sub>**, **ATC<sub>5</sub>-4-NH<sub>2</sub>** and **ATC<sub>5</sub>-5-NH<sub>2</sub>**, have a CO-NH<sub>2</sub> extremity and can be prepared by SPPS, as previously presented (Scheme 16). After elongation and cleavage, each oligomer was purified using semi-preparative HPLC. After lyophilisation, the five foldamers were isolated as TFA salts.

The modest yields obtained (around 15%) were explained afterwards by some difficulties encountered during the cleavage phase. In fact, we found out later that, after a cleavage of three hours in a solution of TFA/TIS/H<sub>2</sub>O 95/2.5/2.5 vvv, the resin was only partially cleaved. Another cleavage was then performed on the resin, giving the possibility to recover as much product as obtained from the first cleavage; then yields were doubled.

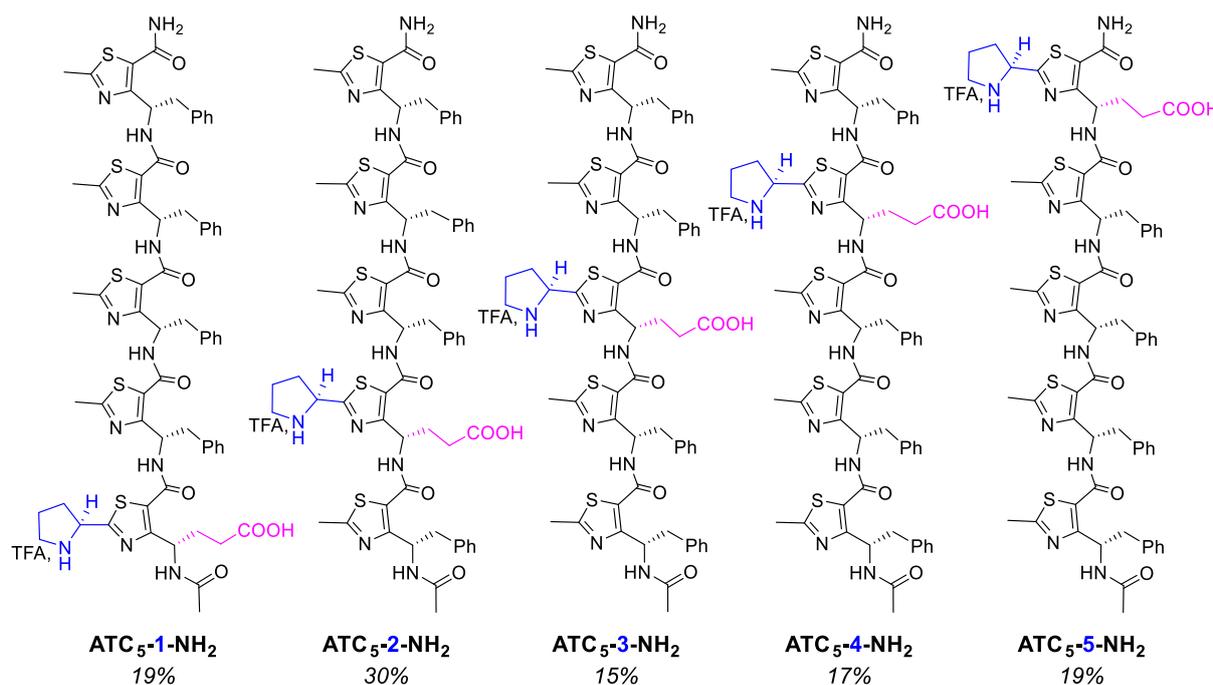


Figure 74: Foldamers  $ATC_5-P-NH_2$  with  $P = 1$  to 5, yields are in italics

All these oligomers shared comparable CD signatures with a strong positive band centred at 265 nm and a negative maximum at 232 nm (Figure 75). These signatures indicate once again that all sequences adopt a similar conformational preference, consistent with the straight helix structure. The folding properties were also explored by NMR and FTIR spectrometry. Since the data are similar to those obtained for  $ATC_5-3-NH_2$  (see the supporting information, section VI, 6 to 9), they are not considered here.

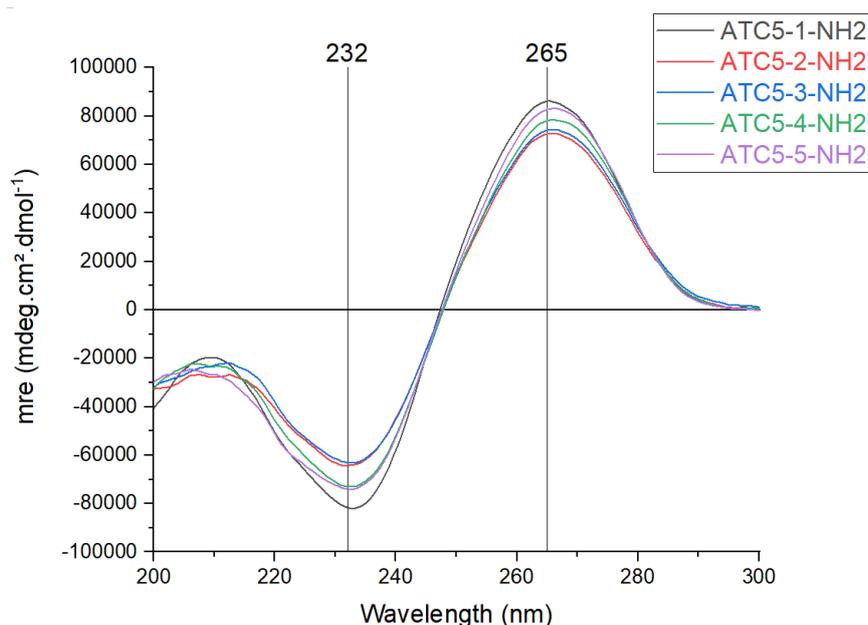


Figure 75: Overlay of the far-UV circular dichroism spectra of compounds  $ATC_5-P-NH_2$  with  $1 \leq P \leq 5$ . All experiments were performed at 4°C with a 50  $\mu$ M peptide solution in *i*PrOH

## 2. *Comparison of the catalytic properties with the position of the competent ATC moiety*

The catalytic properties of all the oligomers **ATC<sub>5</sub>-P-NH<sub>2</sub>** (with  $1 \leq P \leq 5$ ) were evaluated on the nitro-Michael addition reaction, using three different substrates: *β*-*trans*-nitrostyrene, *trans*-4-methyl-*β*-nitrostyrene and *trans*-4-methoxy-*β*-nitrostyrene. Based on the previous optimizations, all the reactions were performed at 4°C in *isopropanol*. Whatever the substrate considered, oligomer **ATC<sub>5</sub>-3-NH<sub>2</sub>**, in which the catalytic centre is in the middle of the sequence, led to the best yield after 16 h reaction (> 95%, Figure 76). We observed a slight decrease in the conversion speed when using the electron rich *trans*-4-methoxy-*β*-nitrostyrene as reactant, and especially for the ATC-oligomers bearing the competent moiety on another position other than the middle of sequence.

Finally, with the exception of the peptide **ATC<sub>5</sub>-1-NH<sub>2</sub>**, in which the catalytic ATC was positioned at the N-terminus, the asymmetric induction was almost similar for all the peptides (ee = 77-82%) demonstrating that the location of the active site along the backbone had little effect on the stereocontrol (Figure 76).

These results are indicative of a low dependency of the **NH<sub>2</sub>-{Pyrrolidine;propanoic acid}-ATC-OH** residue's position on the conversion and enantioselectivity of the reaction. Nevertheless, when the competent moiety is located on the N-terminal or C-terminal extremity, the enantioselectivity ratio drops.

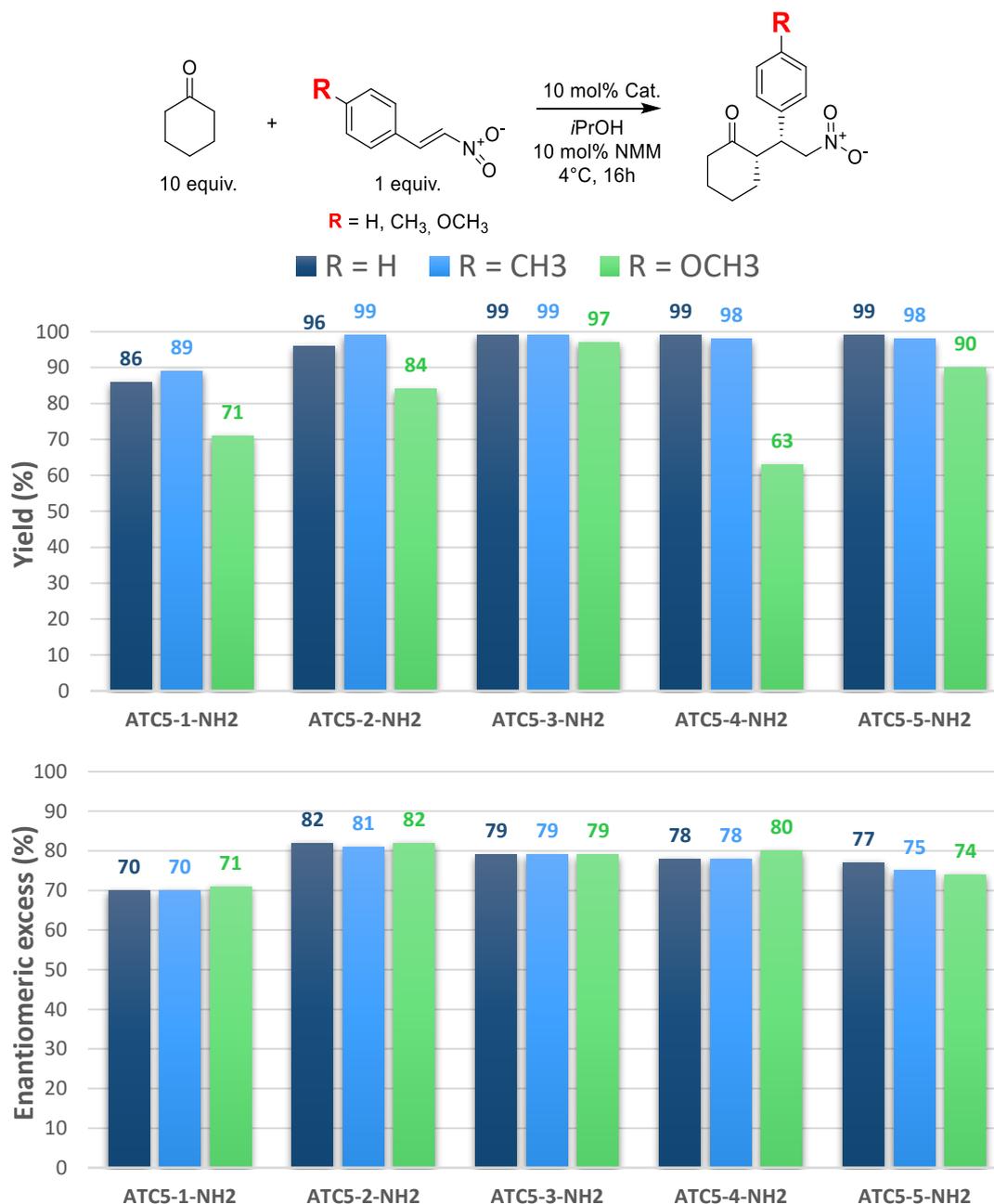


Figure 76: Evaluation of the nitro-Michael addition reaction using  $\beta$ -trans-nitrostyrene ( $R=H$ ), trans-4-methyl- ( $R=CH_3$ ) and trans-4-methoxy- $\beta$ -nitrostyrene ( $R=OCH_3$ ) and oligomers ATC5-*P*-NH<sub>2</sub> as catalysts

## VI. CONCLUSION

As emphasised by the Wennemers' group, the catalytic activity is mainly guided by the average distance between the pyrrolidine and the Brønsted acid functions. Thus, the next objective was to study the impact of the relative position of the lateral carboxylic chain with respect to the pyrrolidine on the catalysis properties. Therefore, the next chapter will be dedicated to this exploration and of the effects driven by the modification of the microenvironment around the catalytic centre.

CHAPTER 4:  
MODULATION OF THE  
MICROENVIRONMENT AROUND THE  
CATALYTIC CENTRE



## I. EFFECT OF THE DISTANCE BETWEEN THE PYRROLIDINE AMINE AND THE CARBOXYLIC ACID AND THE CARBOXYLIC ACID

The first modification we applied to the sequence was done to evaluate the importance of the distance between the pyrrolidine amine and the carboxylic acid. Thus, three ATC-oligomers were synthesised using SPPS as previously reported (Scheme 16), using the ATC-monomers presented in *Chapter 3*. The modification of the length of the lateral chain carrying the carboxylic acid is symbolized by the foldamer **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>** (Figure 77). The separation of the pyrrolidine and the carboxylic acid on two different ATC-monomers is represented by foldamers **M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>** and **M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>**.

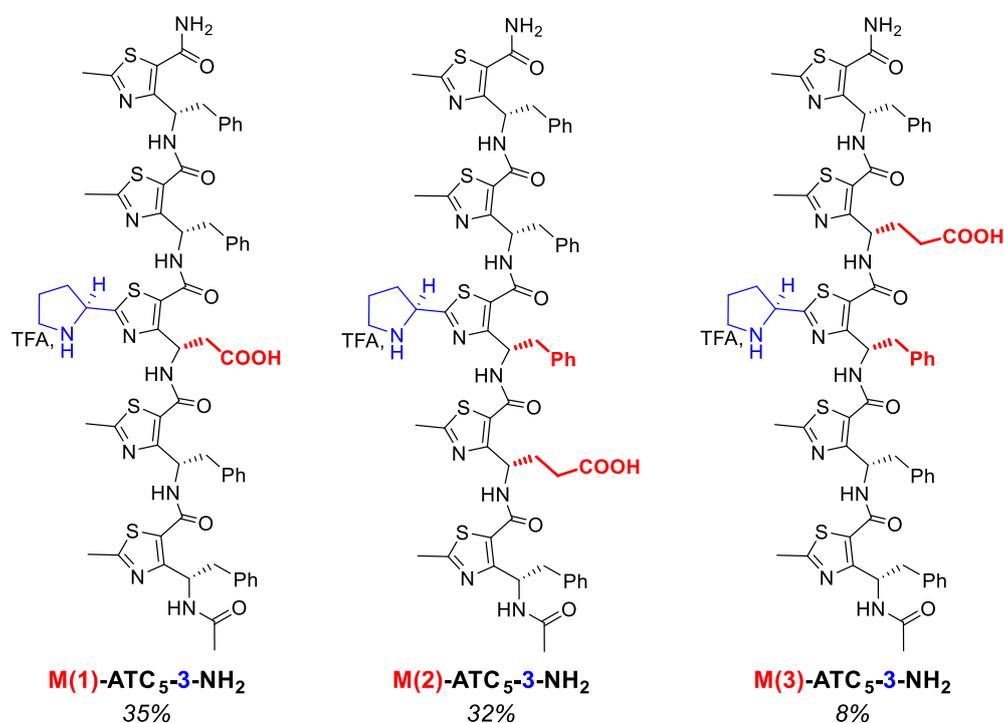


Figure 77: Modulation of the position and the nature of the lateral chain carrying the carboxylic acid, yields are in italics

All these oligomers shared the same CD signature as the previous synthesized oligomers, with a strong positive band centred at 265 nm and a negative maximum at 232 nm (Figure 78). The NMR NOE data sets as well as the FTIR spectra were similar to those described previously for **ATC<sub>5</sub>-3-NH<sub>2</sub>**. These are all presented in the supporting information (*section VI, 10 to 12*). Consequently, we concluded that the C<sub>9</sub>-helical structure was conserved, regardless of the modifications applied.

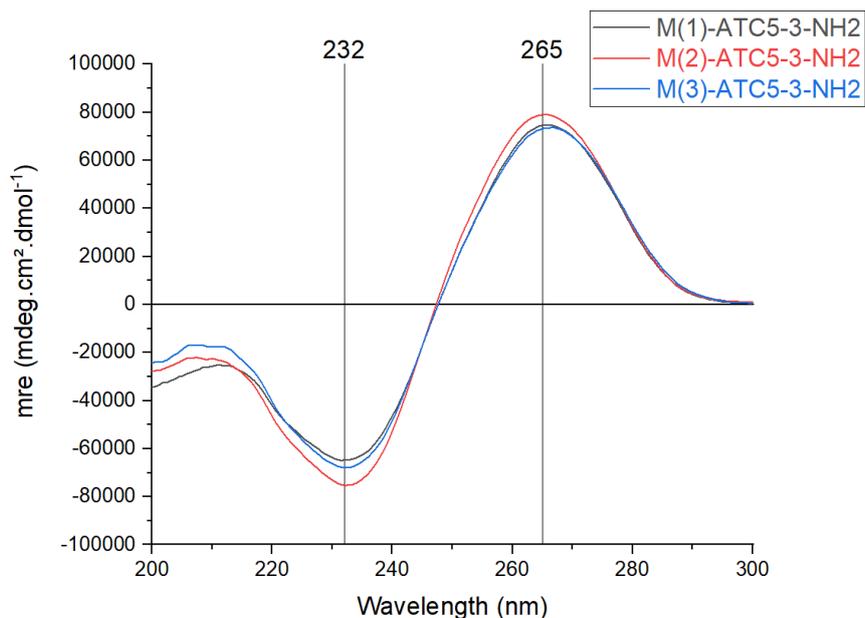


Figure 78: Overlay of the far-UV circular dichroism spectra of compounds **M(1)**- to **M(3)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub>. All experiments were performed at 4°C with a 50 μM peptide solution in iPrOH

The catalytic activities of the three pseudo-peptides were compared to those of **ATC<sub>5</sub>-3-NH<sub>2</sub>**. After a 16 hour reaction, **M(1)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub> led to only 45% yield with β-*trans*-nitrostyrene as substrate. This result was quite surprising since with *p*-methyl-nitrostyrene or *p*-methoxy-nitrostyrene, the conversions were almost quantitative. Nevertheless, in all cases the asymmetric induction remained lower than with **ATC<sub>5</sub>-3-NH<sub>2</sub>** (Figure 79).

The separation of the acidic chain from the pyrrolidine provided two types of profiles. Whatever the olefin considered, the yields obtained with **M(2)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub> were around 94% and the enantiomeric excess varied slightly, from 59 to 63%. On the other hand, when the acidic chain was closer to the C-terminus (**M(3)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub>), the yields were lower, (70 to 90%), but the asymmetric induction was higher, around 75% (Figure 79). Nonetheless, **ATC<sub>5</sub>-3-NH<sub>2</sub>** still produced the best results in terms of conversions and enantiomeric excesses.

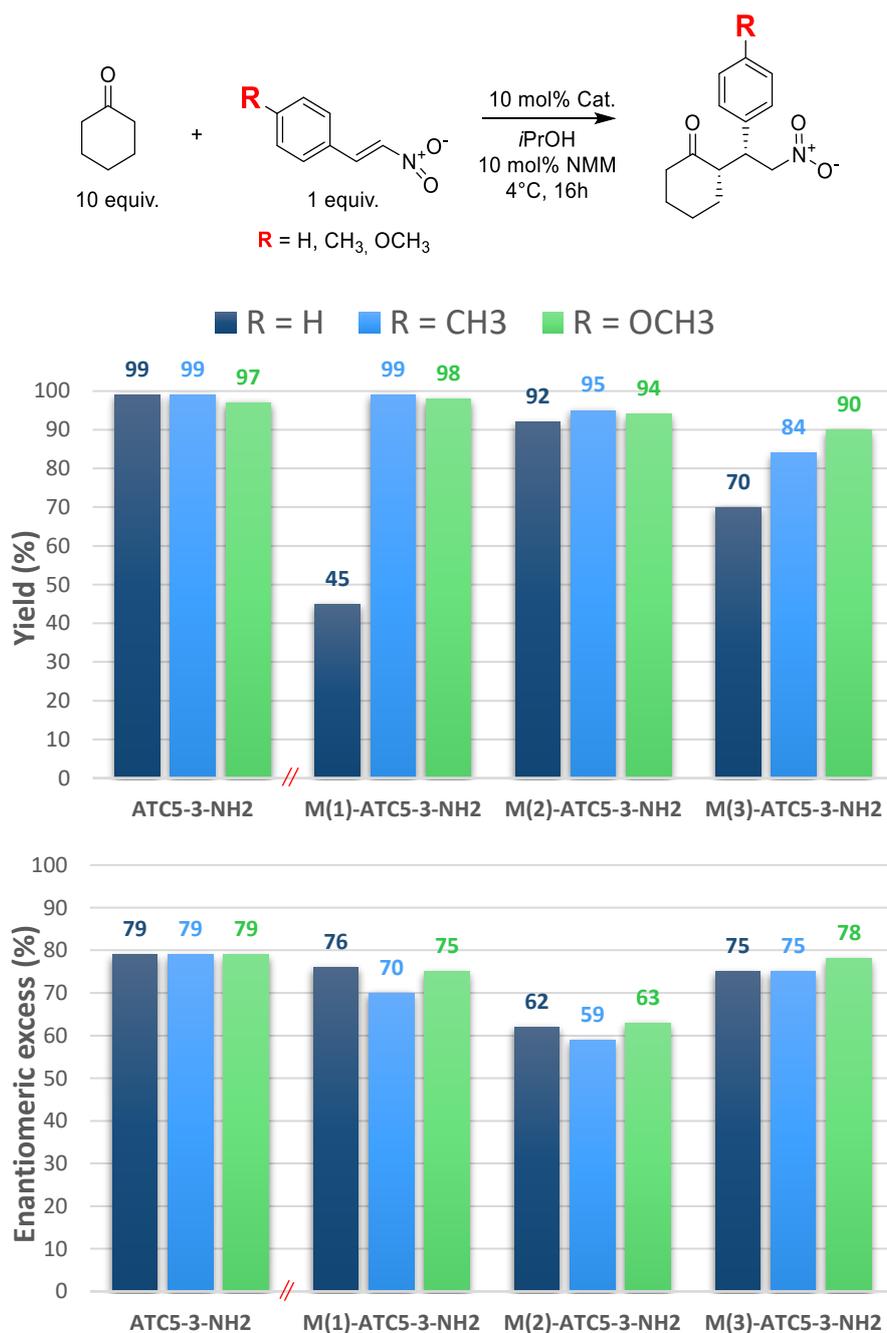


Figure 79: Evaluation of the nitro-Michael addition reaction using  $\beta$ -trans-nitrostyrene ( $R=H$ ), trans-4-methyl- ( $R=CH_3$ ) and trans-4-methoxy- $\beta$ -nitrostyrene ( $R=OCH_3$ ) and oligomers ATC<sub>5</sub>-3-NH<sub>2</sub> and M(1)- to M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub> as catalysts

## II. EFFECT OF THE REMOTE LATERAL CHAINS ON THE CATALYSIS

### 1. Local modifications of the remote lateral chains

In order to modulate the steric hindrance around the catalytic site, we planned to modify the nature of the side chains that are distant from the catalytic residue. Starting from the canonical ATC-helix, we built a 3D-model of the oligomer **ATC<sub>5</sub>-3-NH<sub>2</sub>**. This provided us with an insight about the lateral chains that share the same face of the helix as the catalytic functions (in red, Figure 80). Two of those chains are carried by the thiazole ring of the residues 1 and 5, the third one corresponds to the  $\gamma$ -lateral chain of the residue 2.

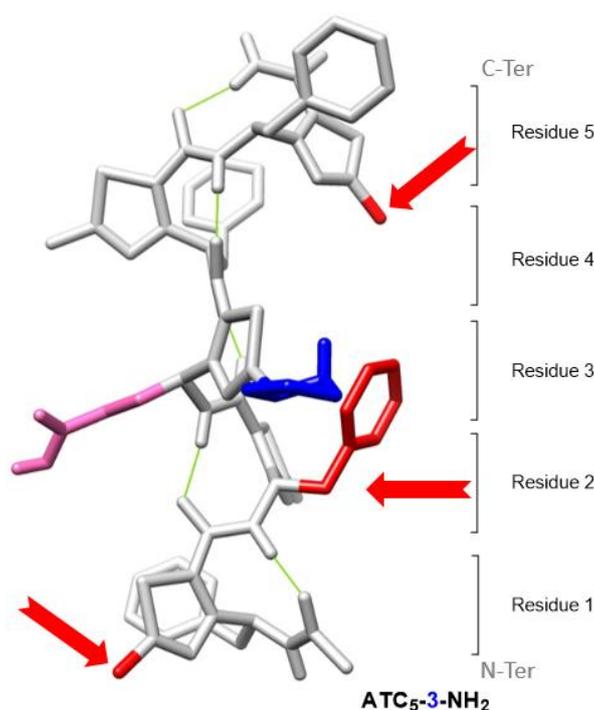


Figure 80: 3D-model of **ATC<sub>5</sub>-3-NH<sub>2</sub>** with the most likely interaction zones between the catalytic centre and its environment represented in red

Therefore, a series of mutated **ATC<sub>5</sub>-3-NH<sub>2</sub>** oligomers was designed (Figure 81). The methyl groups of the thiazole rings of residues **1** and **5** were replaced by a branched aliphatic chain such as *isobutyl*, an aromatic substituent like *benzyl*, or a polar function like an amine or a butyl amine group. The *benzyl* group on the  $\gamma$ -chain of the ATCs **1** and **2** were also modified by and *isobutyl* or a methyl. The ATC-monomers used for those syntheses are the ones discussed at the beginning of this manuscript (*Chapter II, 2*). All the oligomers were synthesized following the SPPS approach previously described.

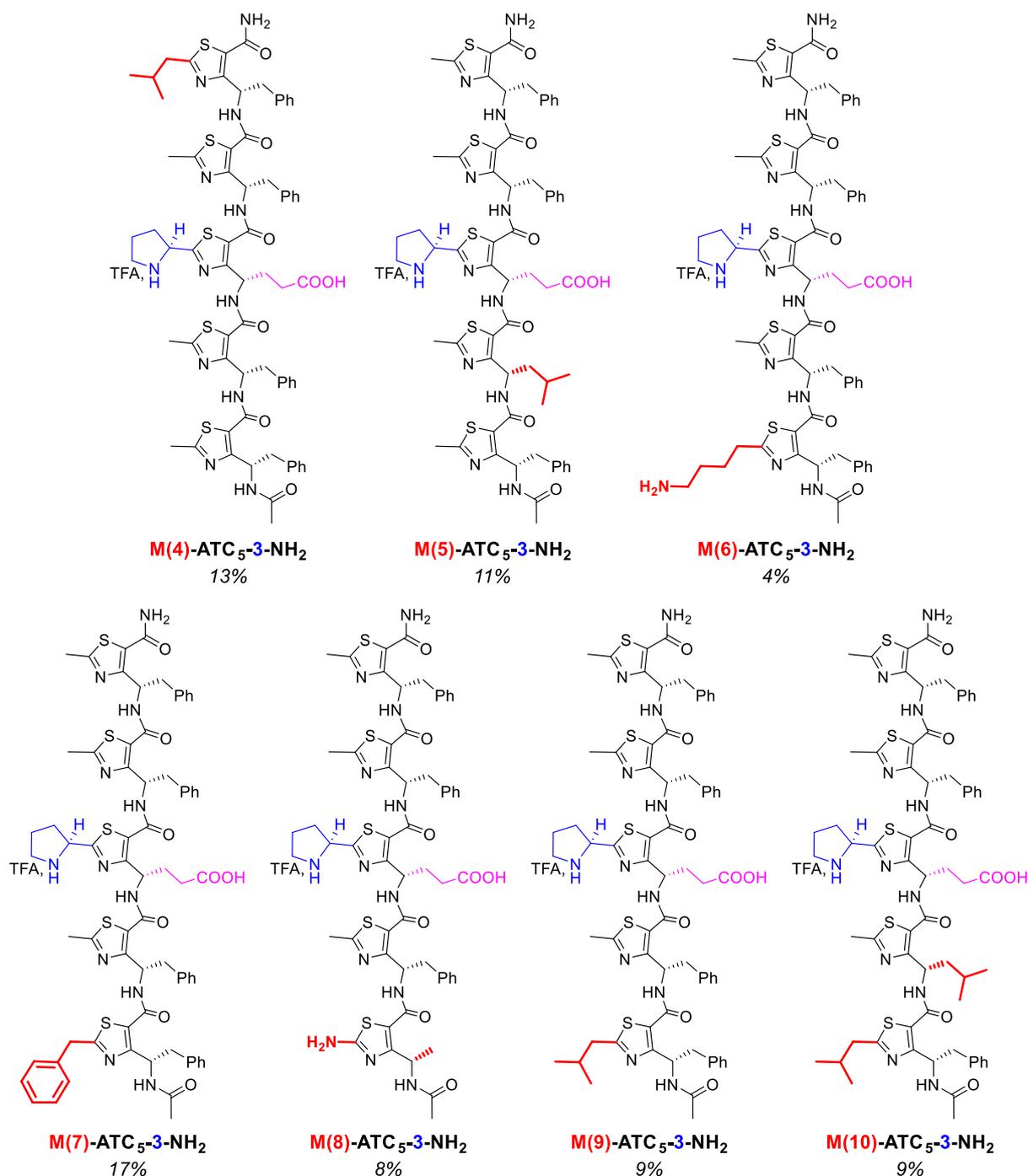


Figure 81: Modulation of the microenvironment around the catalytic centre on a series of foldamers, yields are in italics

All the oligomers displayed the same CD signature and the same set of NOEs as the parent compound **ATC<sub>5</sub>-3-NH<sub>2</sub>** except for **M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub>** for which the intensities were somewhat lower than for the other oligomers and the maxima and minima slightly shifted to 267 and 231nm, respectively (Figure 82).

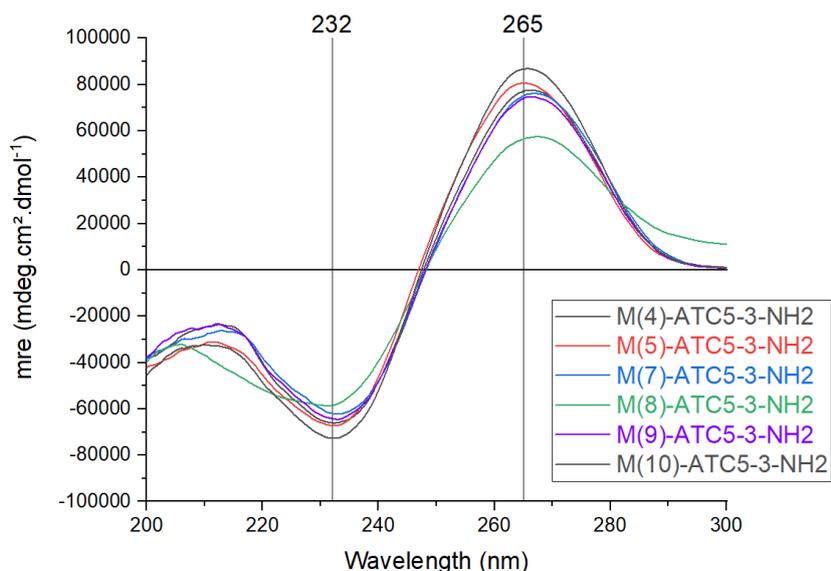


Figure 82: Overlay of the far-UV circular dichroism spectra of compounds **M(4)-** to **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>**. All experiments were performed at 4°C with a 50  $\mu$ M peptide solution in *i*PrOH

When tested on the nitro-Michael addition reaction, notable results were observed. Most of the modifications did not change the catalytic properties significantly though, as the estimated yields were higher than 80% after a 16 hour reaction at 4°C and the enantiomeric excesses were around 75 to 80%, depending on the catalyst and olefin tested. The best asymmetric induction ( $ee = 83\%$ ) was obtained using the oligomer **M(7)-ATC<sub>5</sub>-3-NH<sub>2</sub>** and the *p*-methoxy-nitrostyrene. However, compared to other oligomers, **M(9)-ATC<sub>5</sub>-3-NH<sub>2</sub>** and **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>**, having an *isobutyl* group on the thiazole of the N-terminus residue, showed a marked preference for the conversion of *trans*-4-methoxy- $\beta$ -nitrostyrene (yield > 76%) over the two other olefins (11 to 31% after 16 hours at 4°C) (Figure 83). Currently, this result is not totally understood but clearly demonstrates the possibility of modulating the substrate specificity.

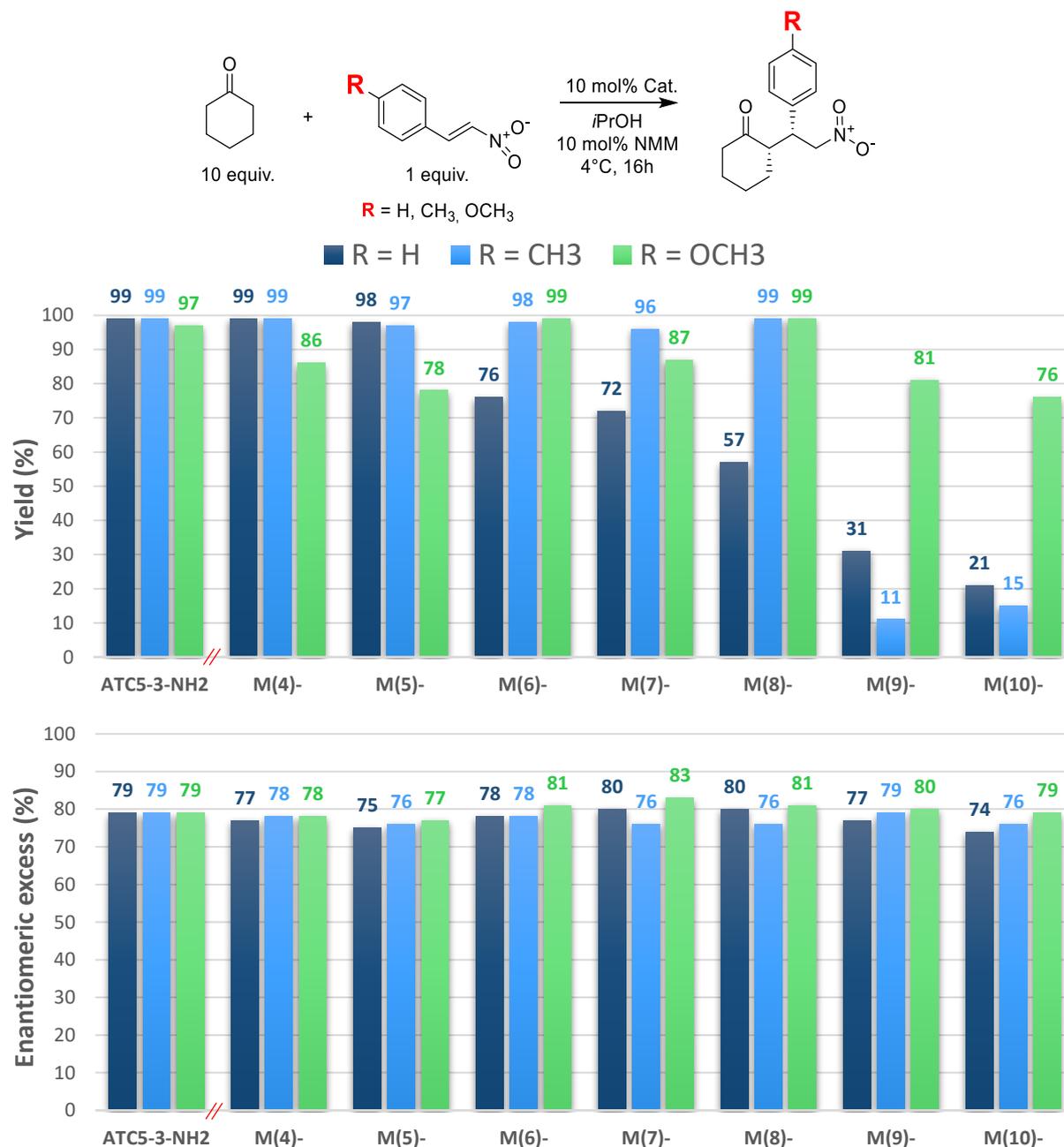


Figure 83: Evaluation of the nitro-Michael addition reaction using  $\beta$ -trans-nitrostyrene ( $R=H$ ), trans-4-methyl- ( $R=CH_3$ ) and trans-4-methoxy- $\beta$ -nitrostyrene ( $R=OCH_3$ ) and oligomers ATC<sub>5</sub>-3-NH<sub>2</sub> and M(4)- to M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub> as catalysts

## 2. Removal of the steric hindrance induced by the benzyl lateral chains

Finally, we sought to remove all the possible steric interactions that could occur between the lateral chains. Thus, all the benzyl lateral chains were replaced by methyl groups. Following the same reasoning as the benzyl series, five oligomers with four identical ATC-monomers NH<sub>2</sub>-{CH<sub>3</sub>;CH<sub>3</sub>}-ATC-OH, and one catalytic centre NH<sub>2</sub>-{Pyrrolidine;propanoic acid}-ATC-OH were synthesized by SPPS, with the position of the catalytic centre varying along the

pentameric structure (Figure 84). Eventually, we synthesized a trimer and a heptamer of this series, with the catalytic centre in the middle of the chain (Figure 84).

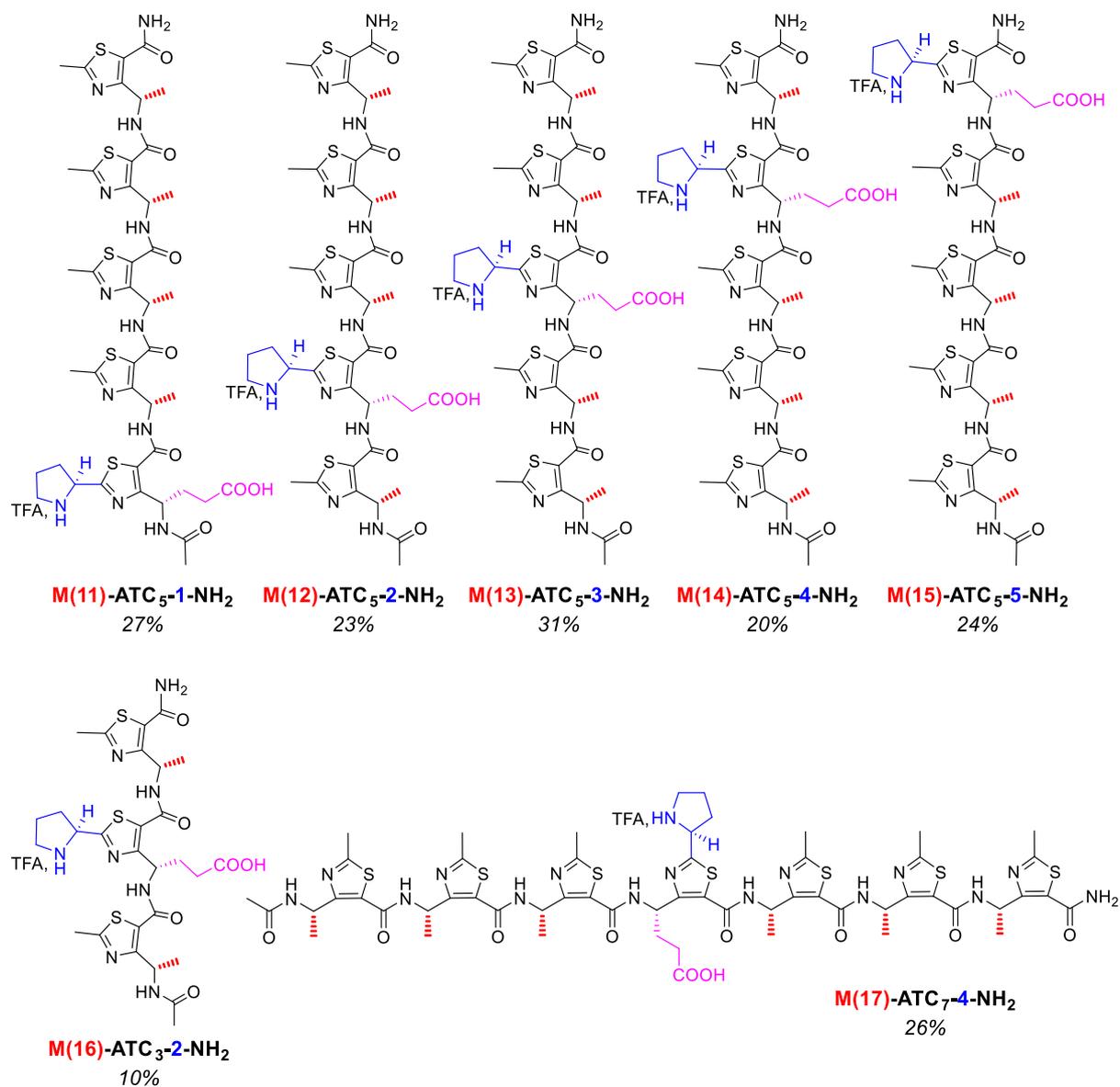


Figure 84: Foldamers **M(11)**- to **M(17)**-ATC<sub>L</sub>-**P**-NH<sub>2</sub>, with **P** from 1 to 5 and *L* = 3, 5 or 7, yields are in italics

The five pentamers presented the same kind of signature in circular dichroism, only with the minima and maxima slightly shifted, respectively from 232 to 231 nm and 265 to 263 nm (Figure 85). The trimer **M(16)**-ATC<sub>3</sub>-2-NH<sub>2</sub> and heptamer **M(17)**-ATC<sub>7</sub>-4-NH<sub>2</sub> also displayed the same kind of signature as the previous five oligomers but the intensities were quite different for each oligomer. It appears here that the length of the oligomer had an impact on the intensity of the signature, something we had not seen before.

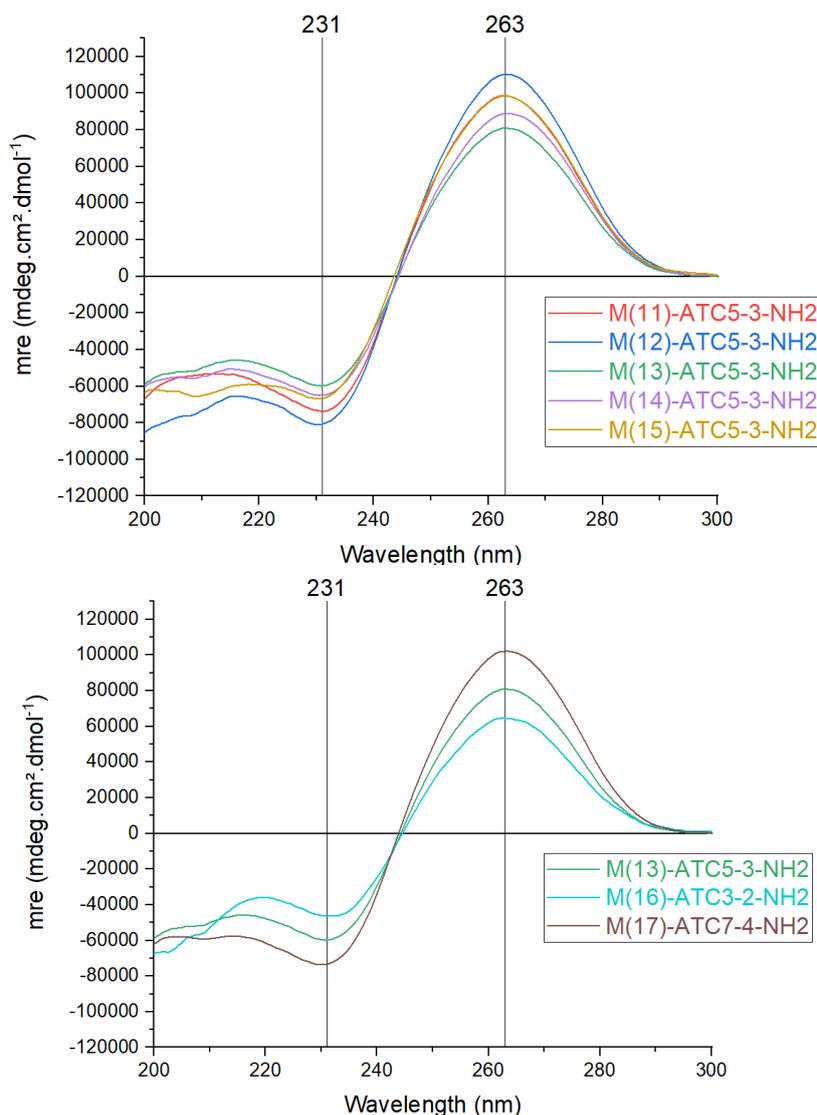


Figure 85: Overlay of the far-UV circular dichroism spectra of compounds **M(11)**- to **M(15)**-ATC<sub>5</sub>-3-NH<sub>2</sub>. All experiments were performed at 4°C with a 50 μM peptide solution in iPrOH

The catalytic properties of the series of oligomers are presented in Figure 86. In contrast to the benzylated series, the reaction yields were inversely correlated to the oligomer length: while the trimer **M(16)**-ATC<sub>5</sub>-3-NH<sub>2</sub> provided good results regarding the yield (>99%), the pentamer **M(13)**-ATC<sub>5</sub>-3-NH<sub>2</sub> led from 9% to 60%, depending on the evaluated substrate. The nitroolefin conversions were even worse with the heptamer **M(17)**-ATC<sub>7</sub>-4-NH<sub>2</sub> (yields = from 5% to 19%). In all the cases, the enantiomeric excesses dropped by almost 8% to 9% in comparison with ATC<sub>5</sub>-3-NH<sub>2</sub>.

Regarding the position of the catalytic residue along the sequence, it appeared that, with the exception of **M(13)**-ATC<sub>5</sub>-3-NH<sub>2</sub>, all the pentamers retained good catalytic properties with

yields greater than 93%, whatever the olefin considered. Conversely, for this series, enantiomeric excesses remained several points lower than those of **ATC<sub>5</sub>-3-NH<sub>2</sub>**.

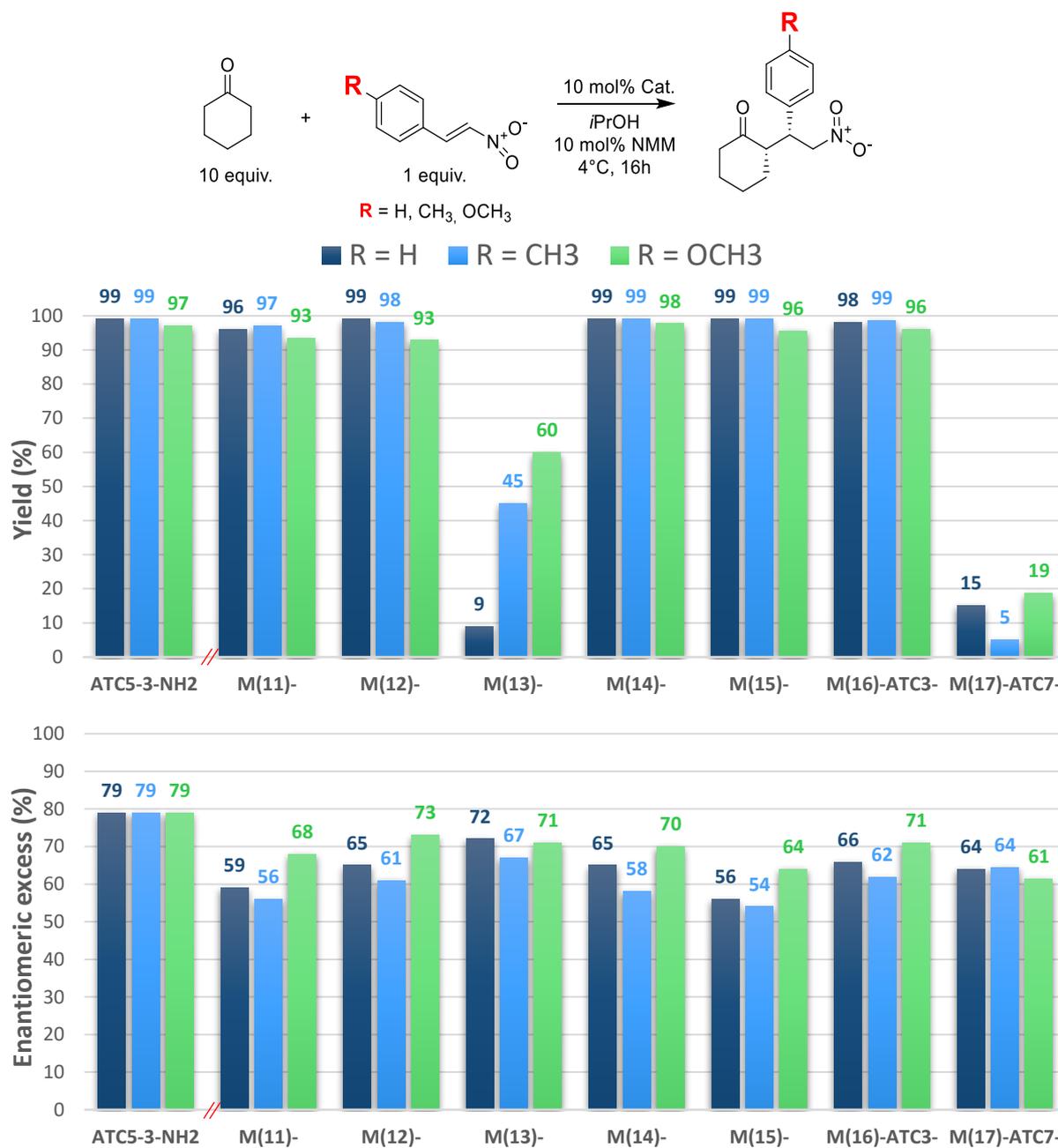


Figure 86: Evaluation of the nitro-Michael addition reaction using  $\beta$ -trans-nitrostyrene ( $R=H$ ), trans-4-methyl- ( $R=CH_3$ ) and trans-4-methoxy- $\beta$ -nitrostyrene ( $R=OCH_3$ ) and oligomers **ATC<sub>5</sub>-3-NH<sub>2</sub>** and **M(10)-** to **M(17)-ATC<sub>L</sub>-P-NH<sub>2</sub>** as catalysts

### III. EXPLORATION OF COOPERATIVE EFFECT

Finally, we explored the possibility of cooperative effects on the catalysis by introducing several catalytic ATC-monomers along a sequence. Hence, four structures were designed. The homotrimer **M(18)**-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub> and homopentamer **M(19)**-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub> were formed by the repetition of the catalytic residue NH<sub>2</sub>-{Pyrrolidine;propanoic acid}-ATC-OH. The two sequences **M(20)**-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub> and **M(21)**-ATC<sub>5</sub>-2,4-NH<sub>2</sub> alternate the NH<sub>2</sub>-{CH<sub>3</sub>;Bn}-ATC-OH monomer and the catalytic residue (Figure 87).

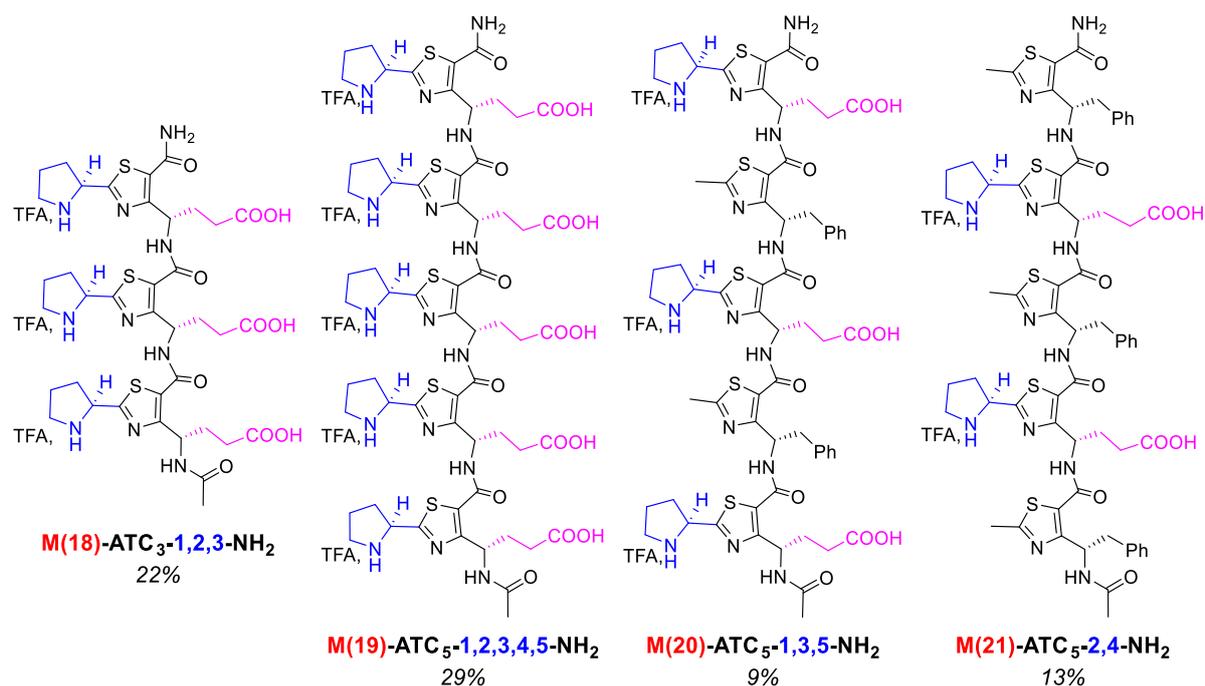


Figure 87: Foldamers **M(18)**-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub>, **M(19)**-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub>, **M(20)**-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub> and **M(21)**-ATC<sub>5</sub>-2,4-NH<sub>2</sub> bearing several catalytic centres, yields are in italics

After circular dichroism, the shape of the signature remained somewhat the same as for other ATC-oligomers, with a minimum at 235 nm and a maximum around 265 nm. Once again, the trimer **M(18)**-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub> and its pentameric equivalent **M(19)**-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub> presented an important shift in intensity. Moreover, the two other pentamers seemed less structured as their signals were a lot less intense. On top of that, we can see that their maxima were not at the same wavelength, 269 nm for **M(20)**-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub> and 267 nm for **M(21)**-ATC<sub>5</sub>-2,4-NH<sub>2</sub>. As circular dichroism is a qualitative tool, it provides clues on the global folding of peptide-like molecules, but further studies, such as NMR, must be performed to obtain the exact structure of the compound.

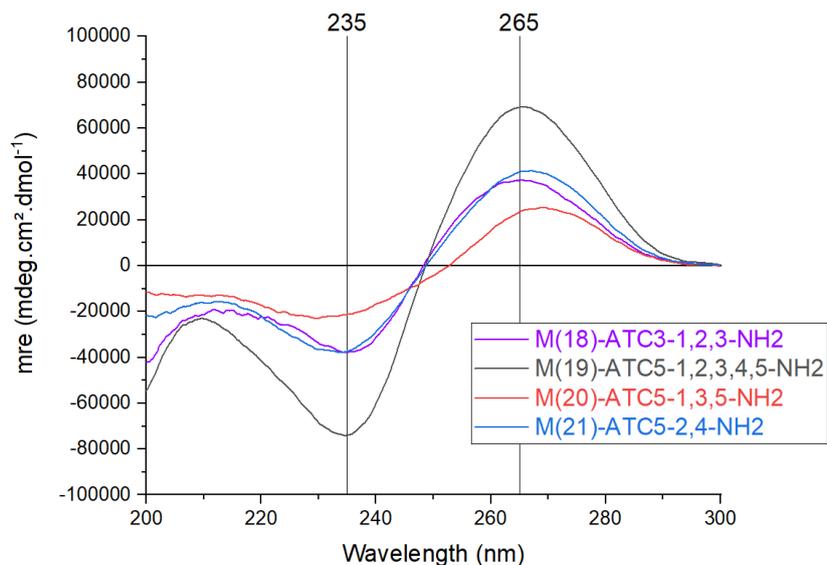


Figure 88: Overlay of the far-UV circular dichroism spectra of compounds **M(18)-** to **M(21)-ATCL-P-NH<sub>2</sub>**. All experiments were performed at 4°C with a 100  $\mu$ M peptide solution in *i*PrOH

The nitro-Michael addition reaction between cyclohexanone and the three defined nitroolefins was performed using these four latter oligomers as catalysts. At first, even though the number of catalytic moieties was higher than **ATC<sub>5</sub>-3-NH<sub>2</sub>**, it was decided to realize the reaction under the same conditions, with 10 mol% of oligomers. The reactivity for all four oligomers was, as expected, exceptional (Figure 89). On the other hand, the enantioselectivity did not rise with the number of catalytic moieties. Actually, it seems that the trimer and pentamer bearing solely the catalytic moiety obtained the worst results in terms of enantiomeric excesses (41 to 65%). Regarding the oligomer bearing three catalytic residues, **M(20)-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub>**, the enantiomeric excess was of about 71% (Figure 89). In contrast, the oligomer **M(21)-ATC<sub>5</sub>-2,4-NH<sub>2</sub>**, bearing only two catalytic moieties, presented a similar enantioselectivity than **ATC<sub>5</sub>-3-NH<sub>2</sub>** (ee = 79%) (Figure 89).

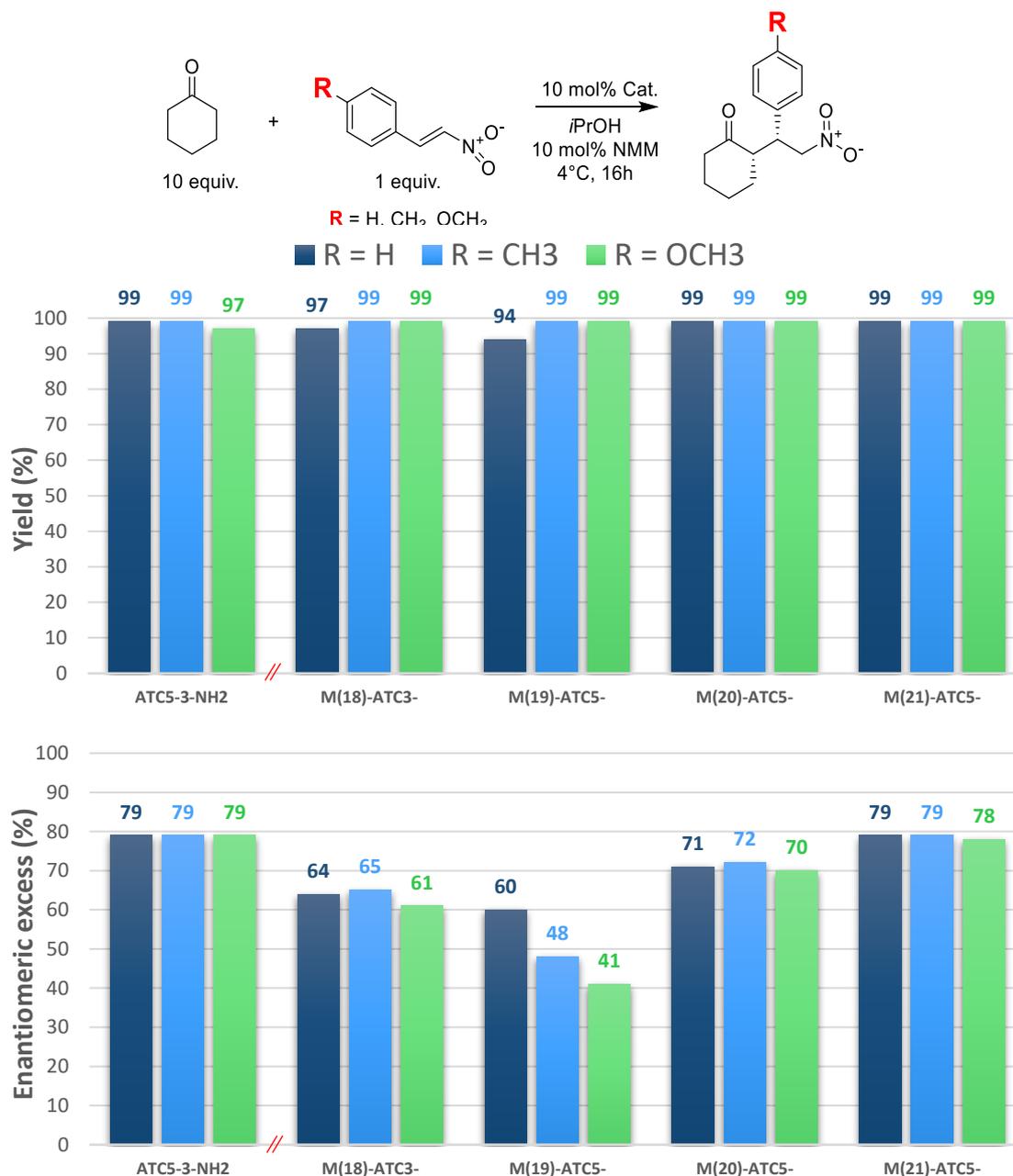


Figure 89: Evaluation of the nitro-Michael addition reaction using  $\beta$ -*trans*-nitrostyrene ( $R=H$ ), *trans*-4-methyl- ( $R=CH_3$ ) and *trans*-4-methoxy- $\beta$ -nitrostyrene ( $R=OCH_3$ ) and oligomers **ATC<sub>5</sub>-3-NH<sub>2</sub>**, **M(18)-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub>**, **M(19)-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub>**, **M(20)-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub>** and **M(21)-ATC<sub>5</sub>-2,4-NH<sub>2</sub>** as catalysts

Finally, to evaluate the possibility of a cooperation, a kinetic study was performed using the pentamer **M(19)-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub>** at 10 mol% and at 2 mol%, which corresponds to a 5/10 a 1/10 ratio of ATC units with respect to  $\beta$ -*trans*-nitrostyrene. The kinetics were compared to the reaction catalysed by the pentamer **ATC<sub>5</sub>-3-NH<sub>2</sub>** at 10 mol% (catalytic ratio ATC/ $\beta$ -*trans*-nitrostyrene = 1/10) (Figure 90).

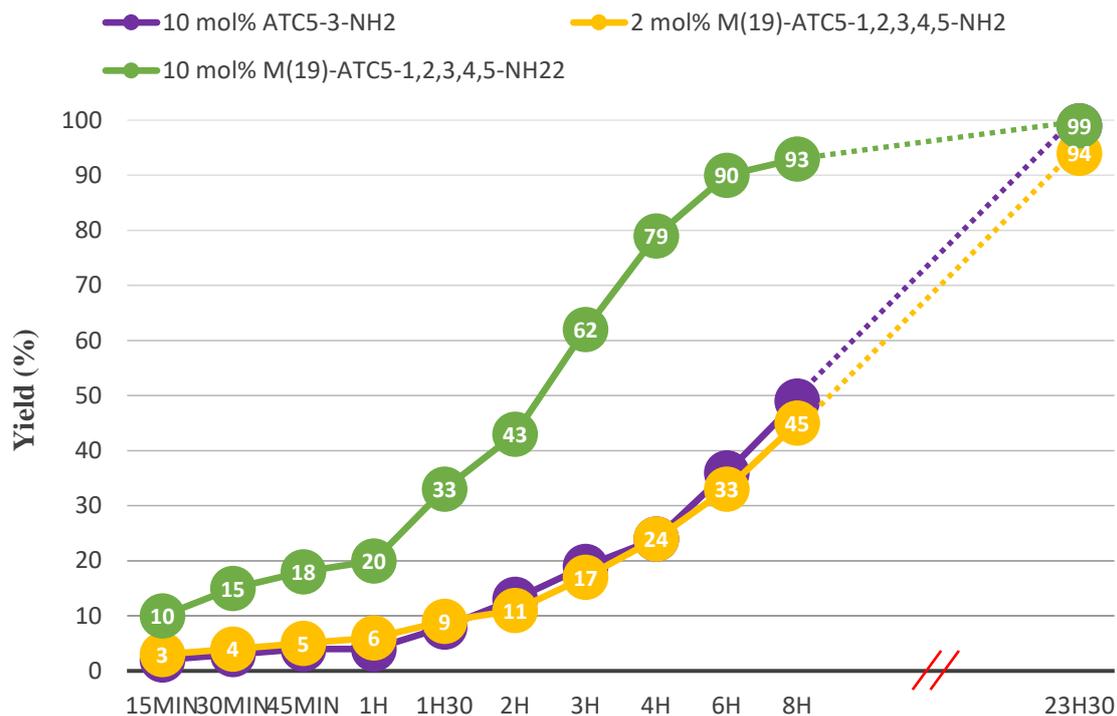


Figure 90: Evolution of the yield in the nitro-Michael addition reaction between cyclohexanone and  $\beta$ -trans-nitrostyrene in function of amount of catalyst used.

As represented in the figure above, when the catalyst bearing only competent moieties was used at 10 mol% (green), the reaction was nearly completed after 8 hours (yield = 93%). On the contrary, when 2 mol% were used (yellow), after 8 hours, the yield reached only 45%. This latest result was similar to using 10 mol% of ATC5-3-NH<sub>2</sub> (purple), thus, catalytic moieties do not appear to act in a cooperative way, the activity is only correlated with the amount of catalytic function used for the reaction.

#### IV. CONCLUSION

On the one hand, the systematic evaluation of the twenty-six synthesized ATC-oligomers as catalysts for the nitro-Michael addition reaction between cyclohexanone and three different olefins permitted to demonstrate their activity in an enamine-type reaction.

Contrary to the expectations, it seems that the correlation between the sequence and the asymmetric induction is quite weak, as the enantiomeric excesses remained mostly greater than 75% but reached their maximum around 80%. On the other hand, the catalytic activity appears highly interrelated with the remote lateral chains, especially when the lateral chains are modulated on the N-terminal extremity. Even if these effects are not totally understood, this is

probably the biggest breakthroughs of this work. It clearly demonstrates the possibility of modulating the substrate specificity which is an inherent property of the enzyme catalysts.

Finally, low resolution circular dichroism and FTIR spectroscopies provided clues for the overall folding of the oligomers. Some NMR data have also corroborated the helical topology generally associated with ATC-oligomers. Nevertheless, structures at the atomic level have not been discussed yet. Therefore, the next chapter will be devoted to the description of the simulated annealing protocol that we used to generate 3D-models from NMR restraints. Some of the obtained structures in the solution state were then compared to crystallographic data.



CHAPTER 5:  
STRUCTURE DETERMINATION AT THE  
ATOMIC LEVEL OF THE ATC-OLIGOMERS



Due to the NMR studies, we imposed experimental restraints and, using the software AMBER16<sup>[212]</sup>, performed NMR refinements by simulated annealing. A part of my thesis consisted of improving the method of 3D-model generation by NMR refinement. The way to computationally build oligomers was modified as well as different force fields were applied to the models. Details of this research and enhancement will be explained in the following paragraphs.

## I. GENERALITIES ON NMR REFINEMENT

It is now well established that three-dimensional structures of small proteins can be determined using a combined approach of multidimensional NMR spectroscopy and computational techniques.<sup>[213,214,215]</sup> Developments in NMR methodology make it possible to extract a large number of structural information, including estimations of interproton distances and vicinal scalar coupling constants. Most internuclear distance constraints come from NOESY data, while scalar coupling constants are more useful for confirming secondary structural features.

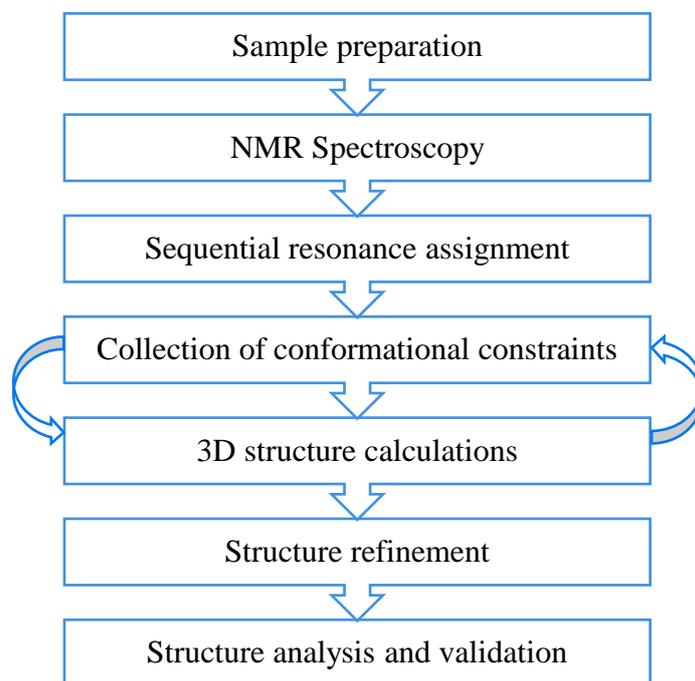


Figure 91: Typical scheme for NMR structure determination

<sup>212</sup> Case, D. A. *et al.* (University of California, San Francisco, 2016).

<sup>213</sup> Wüthrich, K., Wiley-Interscience (New-York, 1986).

<sup>214</sup> Clore, G. *et al.*, *Science*, **252**, 1390–99 (1991).

<sup>215</sup> Gronenborn, A. M. *et al.*, *Proteins Struct. Funct. Genet.*, **19**, 273–76 (1994).

A variety of algorithms were developed to use NOE cross-peaks for structural determination. Molecular dynamic (MD) simulations provide the means to solve the equation of motion of the particles to obtain trajectories for the molecular system. The Newton equation of motion for a molecule consisting of  $N$  atoms is:

$$ma \frac{d^2 r_i}{dt^2} = \frac{\partial}{\partial r_i} U[t|\sigma(r_1, \dots, r_N)]$$

where  $U[t|\sigma(r_1, \dots, r_N)]$  is the potential energy at the moment  $t$ , for the complete set of  $3N$  atomic coordinates. The basic functional form of potential energy includes bonded terms for interactions of atoms that are linked by covalent bonds, and noncovalent terms, that describe the long-range electrostatic and van der Waals forces. Because each interacting force depends on the atoms involved and the environment, the potential must be adapted locally. The force fields define a set of parameters for different types of atoms, chemical bonds, dihedral angles, and so on. The parameter sets are usually empirical but can also be derived from ab-initio calculations.

In MD, the presence of kinetic energy allows the molecule conformations to cross potential surface barriers. In structural determination of peptides and pseudo-peptides, MD is usually combined with simulated annealing (SA) protocols. The kinetic energy provided in terms of temperature defines the maximum height of energy barrier that can be overcome in MD simulations. During the SA calculations, temperature is varied along the MD simulation to illustrate a broad conformational space of the molecule. In a second step, the molecule reaches its minimum energy configuration by slow cooling. To consider the experimental data, a hybrid energy function is defined that incorporates the potential energy function combined with an appropriate force field and NMR structural restraints as pseudopotential energy terms.

One could conclude that there is no single recipe to the "best NMR refinement" since it is highly correlated to quality and interpretation of the NOE data set, the force field optimisation and of common sense, trials and errors.

Eventually, based on the ROESY spectra, NOE cross-peaks were integrated and assigned within the NMRView software.<sup>[216]</sup> The signal intensity ( $I_{ij}$ ) depends on the distance  $r$  between two nuclei  $i$  and  $j$ . Therefore, distances  $r_{ij}$  by  $r^{-6}$  averaging were obtained using these integration

---

<sup>216</sup> Johnson, B. A. *et al.*, *J. Biomol. NMR*, **4**, 603–14 (1994).

values. Comparison was made with an integrated known reference distance  $r_{ref}$  using the following equation:

$$r_{ij} = r_{ref} * \left( \frac{I_{ref}}{I_{ij}} \right)^{1/6}$$

Distances obtained were grouped into three classes and each of them is defined by an upper and a lower bound of distance.

## II. 3D-MODELLING UNDER NMR RESTRAINTS

### *1. Method 1: Simulated annealing protocol and limitation of the NMR refinements*

The first methodology used to generate 3D-models of our oligomers was based on previous 3D modelling studies performed before my arrival within our team. The ATC-oligomers were built using Maestro<sup>[24]</sup> and UCSF Chimera<sup>[9]</sup> and the parameter files were generated using the General AMBER Force Field, so called GAFF. The reference for the volume of a NOE was taken between methylene pair protons (1.8 Å). The lower bound for all restraints was fixed at 1.8 Å and upper bounds at 2.7, 3.3, and 5.0 Å, for strong, medium, and weak correlations, respectively. Structure calculations were performed with AMBER 16<sup>[212]</sup> in two stages: cooking and simulated annealing (SA) in vacuum. The cooking stage was performed at 1000 K to generate 100 initial random structures. SA calculations were carried during 20 ps (20,000 steps, 1 fs long). First, the temperature was raised quickly and was maintained at 1000 K for the first 5000 steps, then the system was cooled gradually from 1000 K to 100 K from steps 5001 to 18,000. Finally, the temperature was brought to 0 K over the 2000 remaining steps. For the 3000 first steps, the force constant of the distance restraints was increased gradually from 2.0 kcal.mol<sup>-1</sup>.Å<sup>-2</sup> to 20 kcal.mol<sup>-1</sup>.Å<sup>-2</sup>. For the rest of the simulation (step 3001 to 20,000), the force constant was kept at 20 kcal.mol<sup>-1</sup>.Å<sup>-2</sup>. The calculations were launched on the Hawkins, Cramer, Truhlar (HCT) form of pairwise generalized Born model for solvation.<sup>[217]</sup> The 20 lowest energy structures with no violations greater than 0.3 Å were considered as representative of the

---

<sup>217</sup> Tsui, V. *et al.*, *Biopolymers*, **56**, 275–91 (2000).

oligomer structure. The representation and quantitative analysis were carried out using MOLMOL<sup>[218]</sup> and UCSF Chimera.

Using the integration of the observed NOE cross-peaks, three major restraints between  $\text{NH}_i$  and  $\text{NH}_{i+1}$ ,  $\text{H}\gamma_i$  and  $\text{NH}_{i+1}$ , and,  $\text{H}\gamma_i$  and  $\text{H}\tau_{i+1}$  were set as weak, strong and weak, respectively. In addition to this basic set, we also included, when visible on the NOE spectrum, some weak restraints between the  $\gamma$ -lateral chains of the residue  $i$  and the lateral chains born by the thiazole ring of the residue  $i+1$ . In the case of **ATC<sub>5</sub>-3-NH<sub>2</sub>**, the 20 structures of lowest energy converged towards the expected C<sub>9</sub>-helix with a Root-Mean-Square Deviation (RMSD) of 1.175 Å (Figure 92). The dihedral angle values were consistent with those expected for the canonical 9-helix. However, we can observe that some dihedral angles were somewhat shifted from the expected canonic C<sub>9</sub>-helix values (Table 19).

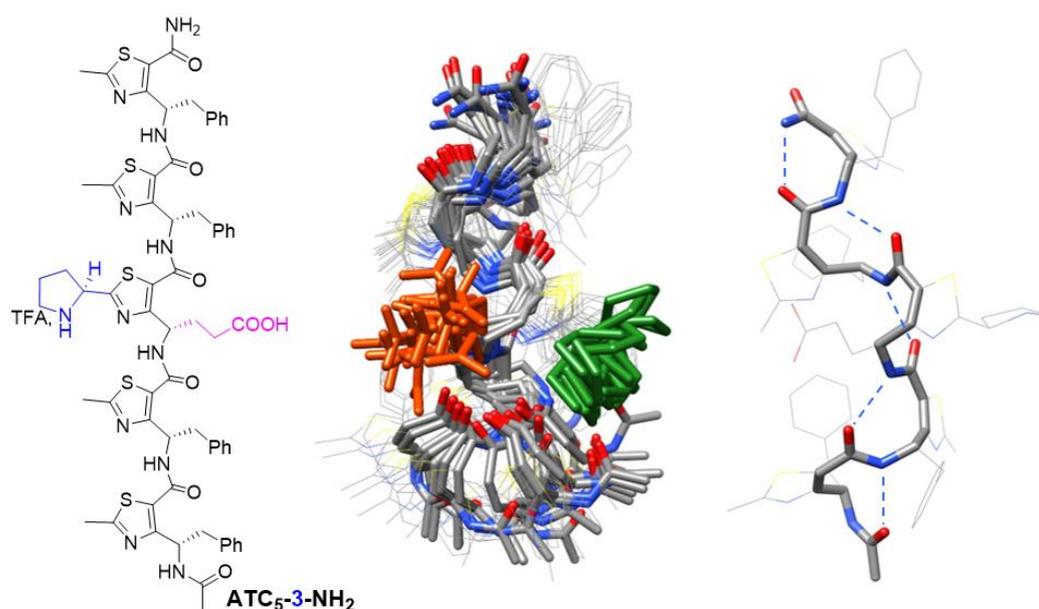


Figure 92: Superimposition of the 20 lowest energy structures and isolated structure to observe the H-bonds of **ATC<sub>5</sub>-3-NH<sub>2</sub>**, RMSD (backbone without capping) = 1.175 Å

<sup>218</sup> Koradi, R. *et al.*, *J. Mol. Graph.*, **14**, 51–55 (1996).

Table 19: Average dihedral angle values observed for the 20 best structures obtained after NMR refinement for foldamers ATC<sub>5</sub>-3-NH<sub>2</sub>

Canonic C <sub>9</sub> -helix		-78	+127	0	-41	
Oligomer	Res. <sup>a</sup>	φ (°)	θ (°)	ζ (°)	ψ (°)	H-bond <sup>b</sup> (Å)
ATC <sub>5</sub> -3-NH <sub>2</sub>	1	-71.3 ± 14.9	+122.2 ± 6.1	-5.1 ± 1.7	-28.5 ± 10.8	1.92
	2	-121.9 ± 34.7	+53.7 ± 87.4	-1.6 ± 3.0	-6.8 ± 16.4	1.94
	3	-108.0 ± 29.1	+107.1 ± 13.3	-2.9 ± 2.8	-19.8 ± 7.8	1.98
	4	-77.1 ± 23.8	+120.4 ± 9.6	-4.6 ± 1.6	-25.4 ± 3.7	1.94
	5	-111.5 ± 22.1	+115.6 ± 18.5	-0.6 ± 3.3	-3.2 ± 15.4	2.21 <sup>c</sup>

<sup>a</sup>Residues are ordered from N to C-terminal extremities, <sup>b</sup>Average length of H-bond formed between CO<sub>i</sub> and NH<sub>i+2</sub>, <sup>c</sup>Average length between O of CO(4) and the barycentre of the two H from NH<sub>2</sub> on the C-terminus

However, starting from the same set of NOEs, some oligomers did not converge towards the helical shape. As a representative example, **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>** led to a rotation around two dihedral angles (Figure 93), hence the well-established C<sub>9</sub> hydrogen-bonding pattern was unbalanced (Table 20).

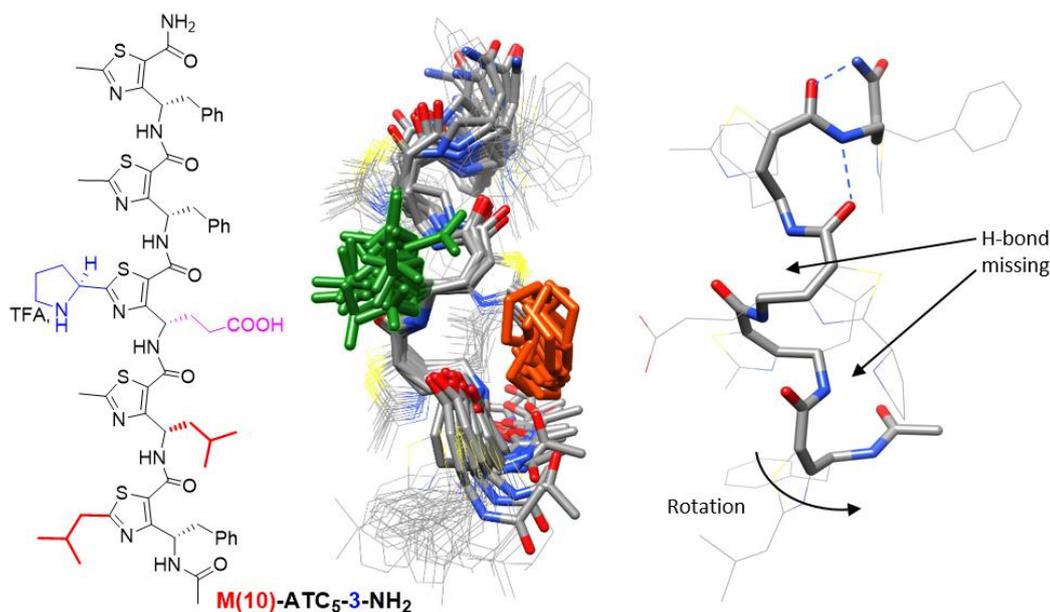


Figure 93: Superimposition of the 20 lowest energy structures and isolated structure to observe the H-bonds of **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>**, RMSD (backbone without capping) = 0.818 Å

Table 20: Average dihedral angle values observed for the 20 best structures obtained after NMR refinement for foldamers **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

Canonic C <sub>9</sub> -helix		-78	+127	0	-41	
Oligomer	Res. <sup>a</sup>	φ (°)	θ (°)	ζ (°)	ψ (°)	H-bond <sup>b</sup> (Å)
<b>M(10)-ATC<sub>5</sub>- 3-NH<sub>2</sub></b>	1	-115.9 ± 49.5	+110.1 ± 32.8	-3.4 ± 3.6	-14.3 ± 6.2	2.85 <sup>c</sup>
	2	-67.2 ± 6.8	+126.4 ± 10.6	-2.6 ± 7.3	-12.0 ± 16.0	1.91
	3	-137.0 ± 14.6	+94.6 ± 16.4	-3.5 ± 3.6	-13.5 ± 5.9	2.50 <sup>c</sup>
	4	-97.0 ± 24.6	+114.1 ± 13.6	-6.8 ± 2.8	-20.0 ± 5.3	1.90
	5	-118.3 ± 18.8	+112.1 ± 15.6	-1.9 ± 6.7	-7.4 ± 19.6	2.20 <sup>d</sup>

<sup>a</sup>Residues are ordered from N to C-terminal extremities, <sup>b</sup>Average distance of H-bond formed between CO<sub>i</sub> and NH<sub>i+2</sub>, <sup>c</sup>Average distance between O of CO<sub>i</sub> and H of NH<sub>i+2</sub>, <sup>d</sup>Average length between O of CO(4) and the barycentre of the two H from NH<sub>2</sub> on the C-terminus

The first issue we addressed was whether this result disproved our well-defined C<sub>9</sub>-helical structure hypothesis, or that it resulted from a bias due to the simulated annealing procedure. According to the first hypothesis, we should have observed more NOE cross-peaks, especially between ATC-lateral chains around the flip, though this was not the case. Consequently, two other options were considered:

1. A possible over-estimation of some NOE cross peaks, leading to a misfolding of the oligomers.
2. An incorrect setting of the applied force field.

## 2. Variation of the NOE restraints

As previously mentioned, while NMR refinements usually work fine, the quality of the results is highly dependent on interpretation of the NMR data. One major limitation with NOESY-NMR is the fact that only an average distance over time is measured meaning that experimental observations are thus driven by population statistics. Significant fluctuation of internuclear distances between non-sequential <sup>1</sup>H nuclei on the NOESY timescale is likely to be found within peptide foldamers. Consequently, calculation of distances from NOE cross-peaks can give the wrong impression of a fixed distance but still might be used to describe the conformational behavior of the molecule in solution. A cautious analysis is thus required to select the relevant NOE cross-peaks associated with a single conformational cluster. In the worst case, a set of NOE cross-peaks observed in an unstructured foldamer may map to a single incorrect structural interpretation, resulting in spurious conclusions.

Therefore, to explain the differences in the calculation outcomes, we hypothesized that some of the restraints could be irrelevant and/or overvalued. Therefore, we decided to categorize restraints in two sets: the “mandatory” and the “optional” restraints.

From the past observations, we noticed that the NOE cross-peaks between  $H\gamma_i$  and  $NH_{i+1}$  and the ones corresponding to  $NH_i-NH_{i+1}$  were necessary for the establishment of the helical structure so these two were classified as “mandatory” restraints. Usually, the distances between  $H\gamma_i$  and  $NH_{i+1}$  are below 2.7 Å and the ones corresponding to  $NH_i-NH_{i+1}$  are greater than 3.3 Å, this can be verified on the NOESY experiments. Furthermore, NOE cross-peaks between  $H\gamma_i$  and  $H\gamma_{i+1}$  are not always visible so this restraint was categorized as “optional” and is set above 3.3 Å (Figure 94A). In some cases, other correlations could be found, between  $H\gamma_i$  and  $H\delta_{i+1}$  for instance, or between  $H\gamma_i$  and the protons carried by the lateral chain of the thiazole-ring of the residue  $i+1$ . Cross-peaks can also be integrated between protons in position  $\delta$  or further along the gamma lateral chain of the residue  $i$  and the protons carried by the lateral chain of the thiazole-ring of the residue  $i+1$ . All these restraints were categorized as “optional” (Figure 94B).

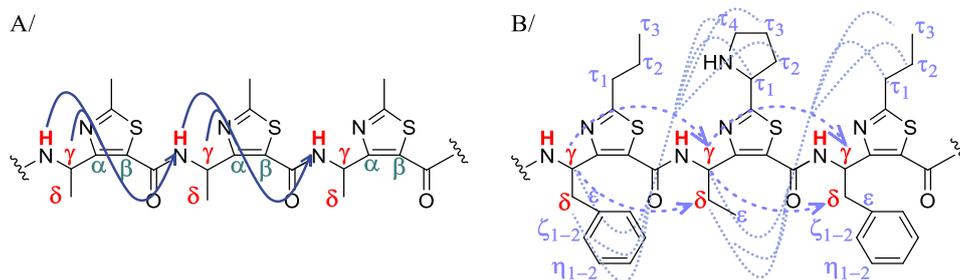
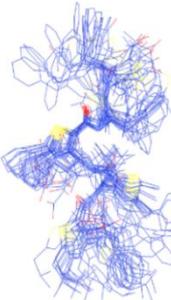
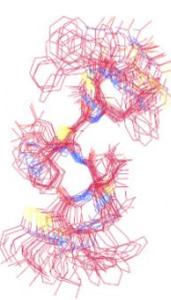
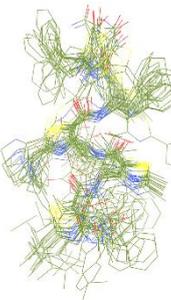
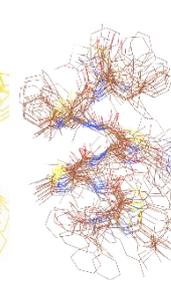


Figure 94: Distance restraints: A/Mandatory, B/ Optional

Eventually, a systematic generation of the foldamer 3D-models was performed with one parameter being changed after another, as represented in the Table 21 for the foldamer **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>**. Values filling this table correspond to the values set in the script for the generation of the structure. “o” indicates “off”, where no restraint were imposed between the two atoms, “w” is for “weak”, the value is set at 5.0 Å, “m” for “medium is fixed at 3.3 Å and “s” for “strong” at 2.7 Å. Some restraints had a penalty of “+1” applied to them because several protons were in competition (*e.g.* Ac-1.NH: “s+1” as there are 3 hydrogens on the acetyl group and if we had restrained the distance at “s” only, it would have locked an equal distance for the three protons while they are not equidistant to the amide hydrogen and the restraint would then have been too restrictive). Each number in *italics* corresponds to the residue numbers along the chain, from the N- to the C-terminal extremity. In entry 1 all the **observed** NOE cross-peaks

have been considered. In entry 2, we added some optional restraints between lateral chains. In entries 3 to 6, we kept open the mandatory restraints and closed each of the optional ones. These changes of parameterization have led to new structures (Table 21).

Table 21: Variation of the parameters for **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>** and superimposition of the 20 structures of lowest energy for each entry

Entry	1	2	3	4	5	6
1.NH-2.NH	w	w	w	w	w	w
2.NH-3.NH	w	w	w	w	w	w
3.NH-4.NH	w	w	w	w	w	w
4.NH-5.NH	w	w	w	w	w	w
5.NH-NH <sub>2</sub>	o	o	o	o	o	o
Ac-1.NH	o	o	o	o	o	o
1.Hγ-2.NH	s	s	s	s	s	s
2.Hγ-3.NH	s	s	s	s	s	s
3.Hγ-4.NH	s	s	s	s	s	s
4.Hγ-5.NH	s	s	s	s	s	s
5.Hγ-NH <sub>2</sub>	s/s	s/s	s/s	s/s	s/s	s/s
1.Hγ-2.Hγ	o	o	o	o	o	o
2.Hγ-3.Hγ	w	w	w	o	o	o
3.Hγ-4.Hγ	w	w	w	o	o	o
4.Hγ-5.Hγ	w	w	w	o	o	o
1.Hγ-2.Hτ	w+1	w+1	w+1	w+1	w+1	o
2.Hγ-3.Hτ	w+1	w+1	w+1	w+1	w+1	o
3.Hγ-4.Hτ	w+1	w+1	w+1	w+1	w+1	o
4.Hγ-5.Hτ	w+1	w+1	w+1	w+1	w+1	o
1.Hγ-2.Hδ	w+1	w+1	o	w+1	o	o
2.Hγ-3.Hδ	w+1	w+1	o	w+1	o	o
3.Hγ-4.Hδ	w+1	w+1	o	w+1	o	o
4.Hγ-5.Hδ	w+1	w+1	o	w+1	o	o
1.Hζ-2.Hτ	m+2	m+2	o	o	o	o
2.Hζ-3.Hτ	o	o	o	o	o	o
3.Hζ-4.Hτ	o	o	o	o	o	o
4.Hζ-5.Hτ	m+2	m+2	o	o	o	o
1.Hζ-1.Hτ	m+2	m+2	o	o	o	o
2.Hζ-3.Hγ	o	w+1	o	o	o	o
2.Hζ-4.Hζ	o	w+1	o	o	o	o
3D-Structure						
RMSD (Å) <sup>a</sup>	0.818	0.653	0.959	0.823	0.765	1.040

<sup>a</sup>backbone without capping; w: weak = 5.0 Å; m: medium = 3.3 Å; s: strong = 2.7 Å; o: off = no restraint applied

One can see that by adding restraints between lateral chains, the structure seems to converge more towards the 9-helix (Table 21, *entry 2*). On the other hand, when removing restraints observed in the ROESY spectrum, the structures became less well organized (Table 21, *entries 3 to 5*). Also, when only “mandatory” restraints were taken into account (Table 21, *entry 6*), the structure became a lot more compact and the canonic C<sub>9</sub>-helical structure was not observed. Therefore, even though the structure tends toward a C<sub>9</sub>-helix, the “mandatory” set is not enough on its own to provide the canonic C<sub>9</sub>-helix, the “optional” restraints must be taken into consideration as well. It seems that all the **observed** NOE cross-peaks are relevant and therefore should be taken into consideration during the generation of the 3D-sequences.

As the hypothesis of a misinterpretation of the NOE cross-peaks was lifted, we decided to focus on the applied force field. Indeed, as we were working on pseudopeptide oligomers, one could argue that the applied force field was not relevant for such structure. Indeed, a significant challenge in the use of computation to study pseudo-peptide systems is force-field accuracy.

### 3. *Setting the force field*

Force fields are composed of a molecular-mechanics energy function with empirical parameters, which are typically obtained from fitting to experimental or high-level quantum mechanical data. These approximations may lead to inaccuracies in calculated kinetic and/or thermodynamic properties. AMBER force fields such as ff99SB have a good reputation in the study of proteins and nucleic acids. However, they are not suitable for other organic molecules. A general set of parameters to allow simulations for small organic molecules is provided by the General AMBER Force Field (GAFF). Compared to the traditional AMBER force field, atom properties in GAFF include element type, hybridization, aromaticity and chemical environment for most of the organic chemical space.<sup>[219]</sup> Parameters exist for almost all the molecules made of C, N, O, H, S, P, F, Cl, Br, and I, and are compatible with the AMBER functional forms. The partial atomic charges are derived by fitting the gas phase electrostatic potential calculated at the Hartree-Fock 6-31G\* level. GAFF also includes parameters for bond lengths, bond angles and torsional angles. Classically, recognition of the atom and bond types as well as assignment of the atomic equivalence and of the force field parameters is automatically performed by Antechamber, a set of auxiliary programs for molecular mechanic studies. However, this General Amber Force Field is known to be efficient for molecules up to 100 or 200 atoms. The problem for bigger molecules is that a quantum calculation is needed to obtain partial charges

---

<sup>219</sup> Wang, J. *et al.*, *J. Comput. Chem.*, **25**, 1157–74 (2004).

and one does not know what sort of performance are to be expected with large systems. Since Antechamber has to infer the bonding characteristics from the geometry, the number of possibilities grows with the size of the molecule. Therefore, we decided to develop a homemade script to generate the ATC-oligomer sequences from individually parametrized ATC-monomers (Method 2).

In contrast to the previously explained method (Method 1), the atomic charges were calculated independently for each monomer using GAFF. We also intended to optimize the force field for the ATC monomers, using high-level quantum mechanical method for charge assignments (Method 3). The three methods used to generate the AMBER parameter files is depicted in Figure 95.

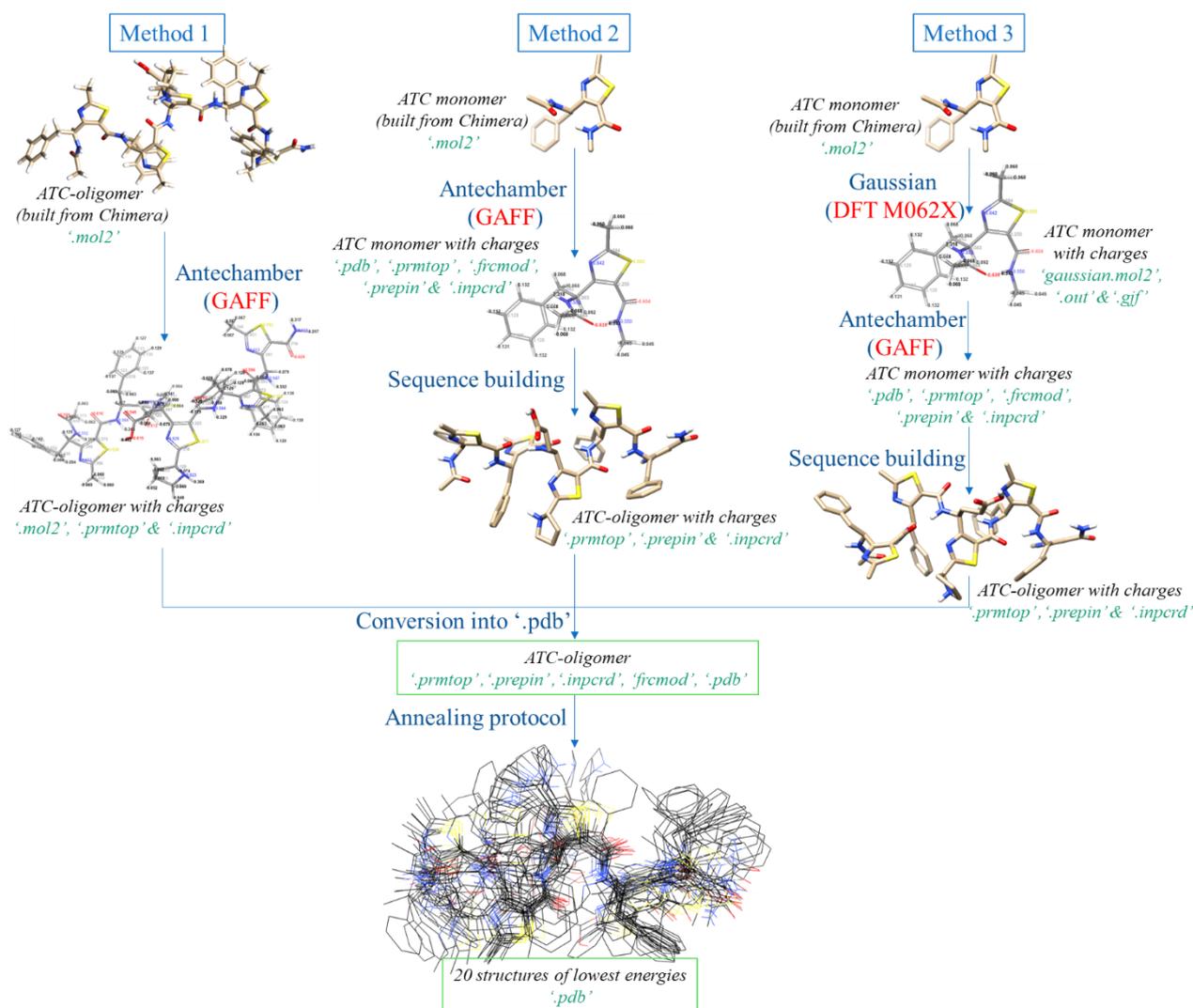


Figure 95: Comparison of the 3 methods employed to generate the AMBER parameter files of ATC sequences

## i. Method 2: Creation of the AMBER parameter files for individual ATC monomers using GAFF

### i.a. Generation of ATC monomer 3D-structures

The 3D-structures of ATC-monomers were generated using the USCF Chimera software. The dihedral angle values were defined from the ATC oligomer crystal structure CCDC 922304 (Table 22 and supporting information, *section IX*).<sup>[103]</sup> Each ATC was capped using an acetyl group (called Ace) at the N-terminus and an *N*-methylamide (Nme) at the C-terminus, resulting in Ace-ATC-Nme. This was done to mimic the environment of an oligomer backbone for an ATC-monomer.

Table 22: Average values of dihedral angles found in the RX structures of CCDC 922304

$\varphi$	$\theta$	$\zeta$	$\psi$
$-78 \pm 3^\circ$	$+127 \pm 14^\circ$	$0 \pm 3^\circ$	$-41 \pm 4^\circ$

For simplicity in the construction and analysis of the sequences, we applied the same nomenclature code for the atoms of each monomer backbone. The amide extremity was identified as following: N for the nitrogen, HN for the amide proton. The carbons  $\delta$ ,  $\gamma$ ,  $\beta$  and  $\alpha$  were named respectively CD, CG, CB and CA while the carbon between the nitrogen (NT) and the sulfur (S) of the thiazole ring was called CT1. The first carbon on the chain held by the thiazole was named CT. The atoms of the acetyl group were named CX, OX, CX2, HH31, HH32 and HH33. Atoms constituting the NH-CH<sub>3</sub> group on the C-terminal extremity were called NX, HNX, CX3, HH34, HH35 and HH36 (Figure 96).

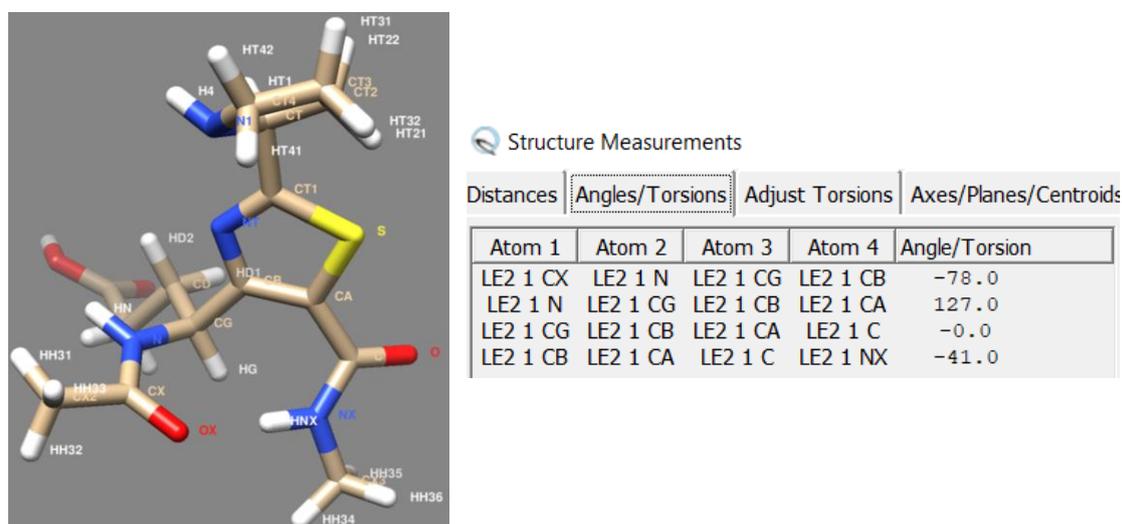


Figure 96: Creation of an ATC-monomer on UCSF Chimera by forcing dihedral angle values and implementing a nomenclature on the backbone atoms

*i.b. Application of the GAFF force field on the ATC monomers and generation of the AMBER file types*

To generate the Amber model parameter file types such as '.prepin', '.inpcrd', '.frcmod', '.pdb' and '.prmtop', we used the LEaP program. LEaP is an Antechamber module that collects parameters in order to make a complete description of a molecule. Each of the parameter files generated by LEaP contained pieces of information needed for the construction of a simulation, whether for energy minimization or for molecular dynamics.

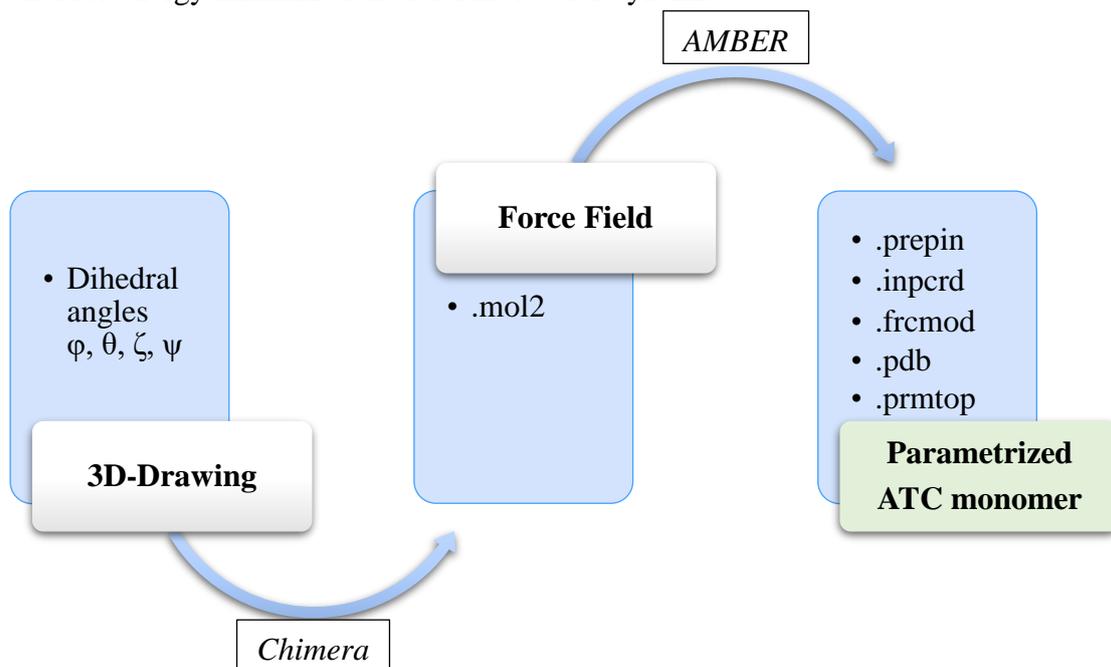


Figure 97: Generation of the parametrized 3D-ATC monomers with the input/output files and the software used for each step

For instance, the '.inpcrd' and '.prepin' files hold initial coordinates and periodic box size. Datasets of new parameters assigned for the ATC building-blocks are consigned in '.frcmod'. The '.prmtop' regroups all the informations as it contains the molecular topology, force field, atom and residue names. Table 23 illustrates the information contained in the different file types.

Table 23: Information contained in different files, Y means yes, and D means that it will be discarded when converted to force field files.

Information	mol2	prepin	inpcrd	frcmmod	pdb	prmtop
Residue name	Y	Y			Y	Y
Atom name	Y	Y			Y	Y
Atom type	Y	Y		Y		Y
Charges	Y	Y				Y
Connectivity	Y	Y			Y	Y
Coordinates	Y	D	Y		Y	
Atom mass				Y		Y
Bonded parameters				Y		Y
Non-bonded parameters				Y		Y

I developed a short script to generate the Amber model parameter file types ‘.prepin’, ‘.inpcrd’, ‘.frcmmod’, ‘.pdb’ and ‘.prmtop’. The structures drawn with UCSF Chimera were saved as a ‘.mol2’ files and used as input files. In lines 5 to 14, the force field GAFF ff99SBnmr was applied using antechamber, generating four out of the five parameter and coordinate output files needed for the rest of the process. Line 16 permitted to generate the ‘.pdb’ file from the ‘.prmtop’ (Figure 98).

```

1 #Antechamber for 1 ATC-monomer
2
3 #Creation of the parameters and coordinates files from a mol2 file (Chimera)
4
5 antechamber -i ./init/name_of_residue.mol2 -fi mol2 -o ./prep/name_of_residue.prepin -fo prepi -c bcc -s 2 > ./prep/antechamber.out
6 parmchk -i ./prep/name_of_residue.prepin -f prepi -o ./prep/name_of_residue.frcmmod
7 cat > leap.in <<EOF
8 source leaprc.gaff
9 source leaprc.ff99SBnmr
10 loadamberprep ./prep/name_of_residue.prepin
11 loadamberparams ./prep/name_of_residue.frcmmod
12 saveamberparm name_of_residue ./prep/name_of_residue.prmtop ./prep/name_of_residue.inpcrd
13 quit
14 EOF
15 tleap -f leap.in > ./prep/leap.out
16 ambpdb -a atm -p ./prep/name_of_residue.prmtop < ./prep/name_of_residue.inpcrd > ./prep/name_of_residue.pdb

```

Figure 98: Script of the parameters and coordinates files generation using Antechamber

A second script was then used to remove the acetyl and amine capping groups from the parametrized Ace-ATC-Nme monomer (Figure 102). The output files could then be used as an ATC library to generate an oligomer sequence.

```

27 #Removal of extremities
28
29 cat > leap.in <<EOF
30 source leaprc.gaff
31 source leaprc.ff99SBnmr
32
33 #Here, loading of the .prepin and .frcmmod files of the ATC for which the extremities have to be removed.
34 #Files are in folder ./final created after the antechamber step
35
36 loadamberprep ./final/name_of_residue.prepin
37 loadamberparams ./final/name_of_residue.frcmmod
38
39 #Identification of the atoms to remove and of the connectors
40 #Head et Tail must be identified even if one of them won't be used in the sequence.
41
42 #Nter
43 remove name_of_residue name_of_residue.1.CX
44 remove name_of_residue name_of_residue.1.CX2
45 remove name_of_residue name_of_residue.1.HH31
46 remove name_of_residue name_of_residue.1.HH32
47 remove name_of_residue name_of_residue.1.HH33
48 remove name_of_residue name_of_residue.1.OX
49
50 #Cter
51 remove name_of_residue name_of_residue.1.CX3
52 remove name_of_residue name_of_residue.1.HH34
53 remove name_of_residue name_of_residue.1.HH35
54 remove name_of_residue name_of_residue.1.HH36
55 remove name_of_residue name_of_residue.1.NX
56 remove name_of_residue name_of_residue.1.HNX
57
58
59 set name_of_residue head name_of_residue.1.N
60 set name_of_residue tail name_of_residue.1.C
61 set name_of_residue.1 connect0 name_of_residue.1.N
62 set name_of_residue.1 connect1 name_of_residue.1.C
63
64 #Creation of the residue without its extremities and saving the corresponding .prepin/.prmtop/.inpcrd files
65
66 name_of_residue = sequence { name_of_residue }
67 saveamberprep name_of_residue ./Librairie_ATC_sans_cap/name_of_residue.prepin
68 saveamberparm name_of_residue ./Librairie_ATC_sans_cap/name_of_residue.prmtop ./Librairie_ATC_sans_cap/name_of_residue.inpcrd
69 quit
70 EOF
71 tleap -f leap.in > ./final/leap.out
72
73 parmchk -i ./final/name_of_residue.prepin -f prepi -o ./Librairie_ATC_sans_cap/name_of_residue.frcmmod
74
75 ambpdb -aatm -p ./Librairie_ATC_sans_cap/name_of_residue.prmtop < ./Librairie_ATC_sans_cap/name_of_residue.inpcrd > ./Librairie_ATC_sans_cap/name_of_residue.pdb

```

Figure 99: Script permitting the removal of the Ace and Nme extremities

## ii. Method 3: Optimization of the ATC-force field by DFT calculation

For the method 3, it was decided to perform an intermediate step where the electrostatic potential was taken into consideration to generate atom charges on ATC-monomers (Figure 100). Force field libraries for ATC fragments were generated using the restrained electrostatic potential (RESP) method.<sup>[220]</sup> After having locally optimized the geometry, the atom charges were derived from the electrostatic potential obtained by DFT calculations with Gaussian 09<sup>[221]</sup> at the M062X/6311++g(d) level of theory in vacuum (See SI, *section VI, 31*). The M06-2X is a global hybrid functional, meaning a class of approximations to the exchange–correlation energy functional in density functional theory, that has shown good performance in predicting noncovalent interactions of small molecules and structures and relative stabilities of biological compounds such as peptides.<sup>[222]</sup>

<sup>220</sup> Wang, J. *et al.*, *J. Comput. Chem.*, **21**, 26.

<sup>221</sup> Frisch, M. J. *et al.*, *Gaussian Inc Wallingford CT* (2016).

<sup>222</sup> Zhao, Y. *et al.*, *Chem. Phys. Lett.*, **502**, 1–13 (2011).

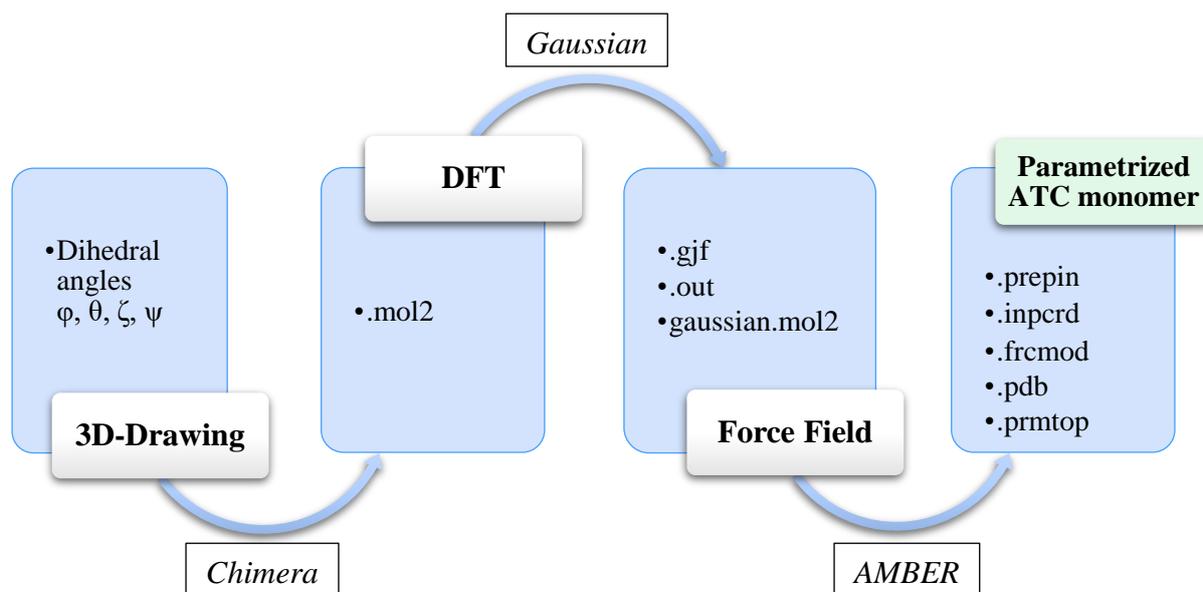


Figure 100: Generation of the parametrized 3D-ATC monomers with the input/output files and the software used for each step

The basic structure of a Gaussian input file includes several different sections:

```
# M062X/6-311++g(d) geom=connectivity
```

*Method and basis set that specifies the model chemistry used.*

*Calculated molecular properties:*

```
Pop=MK iop(6/33=2) iop(6/42=6) iop(6/50=1)
```

*Pop refers to the atomic charges while MK refers to the Merz-Singh-Kollman method.<sup>[223]</sup>*

*iop values define the resolution of the charge distribution calculated around the molecule.*

```
0 1
S -1.93170000 -1.57420000 -0.80170000
N -1.57800000 0.95090000 -0.53170000
C -0.32110000 0.38980000 -0.57730000
C -2.50890000 0.04870000 -0.65160000
C -0.30810000 -0.97090000 -0.73570000
```

*Molecule specification including the atomic coordinates and connectivities.*

```
...
1 4 1.0 5 1.0
2 3 1.5 4 2.0
3 5 2.0 19 1.0
4 6 1.0
5 39 1.0
...
```

<sup>223</sup> Singh, U. C. *et al.*, *J. Comput. Chem.*, **5**, 129–45 (1984).

At the end of the calculation, Gaussian generated an output '.LOG' or '.out' file depending on the software version. It contained all the calculated molecular properties, including the electrostatic surface potential (ESP) values. It is important to notice that Gaussian modified the nomenclature code for the atoms so a new 'gaussian.mol2' file had to be saved. The script that we developed is depicted in Figure 101. The Gaussian '.out' file that contains the charge distribution information as ESP values, the '.mol2' and 'gaussian.mol2' files were used as inputs. Derivation of RESP charges (Atomic charges) from ESP values was done in Lines 6 and 7 of the script. In Line 13, the match\_atom command used the '.mol2' and 'gaussian.mol2' files as input and permitted to return to the initial atom code nomenclature. Finally, lines 15 to 25 provided the Amber model parameter files for the Ace-ATC-Nme molecule.

```

1 #Entry file: name_of_residue.LOG -> change .LOG par .out
2 #Entry file: name_of_residue_g.mol2 obtained via Gaussian
3 #Entry file: name_of_residue.mol2 obtained via Chimera
4 #replace in the template name_of_residue by the corresponding three letter code
5
6 espgen -i ./init/name_of_residue_g.out -o ./prep/name_of_residue.esp
7 respgen -i ./init/name_of_residue_g.mol2 -o ./prep/name_of_residue.respin -f resp1
8
9 #replace name_of_residue by the corresponding three letter code
10 #-nc represents the net charge of the molecule: 0 by default
11
12 antechamber -i ./init/name_of_residue_g.out -fi gout -o ./prep/name_of_residue.prepin -fo prepi -c resp -s 2 -rn name_of_residue -at amber -nc 0
13 match_atomname -i ./prep/name_of_residue.prepin -fi prepi -r ./init/name_of_residue.mol2 -fr mol2 -o ./final/name_of_residue.prepin -h 1
14
15 parmchk -i ./final/name_of_residue.prepin -f prepi -o ./final/name_of_residue.frcmod
16 cat > leap.in <<EOF
17 source leaprc.gaff
18 source leaprc.ff99SBnmr
19 loadamberprep ./final/name_of_residue.prepin
20 loadamberparams ./final/name_of_residue.frcmod
21 saveamberparm name_of_residue ./final/name_of_residue.prmtop ./final/name_of_residue.inpcrd
22 quit
23 EOF
24 tleap -f leap.in > ./final/leap.out
25 ambpdb -aatm -p ./final/name_of_residue.prmtop < ./final/name_of_residue.inpcrd > ./final/name_of_residue.pdb

```

Figure 101: Renaming ATC-monomer's atoms and generation of the parameter's files

The script, which had been edited to remove the extremities of the monomer, was then used. In the generated ".pdb" file, verification could be performed to ensure that the extremities were removed on the ATC-monomer and that the nomenclature of the atoms had remained the same as the one that had been implemented at the beginning of the process (Figure 102).

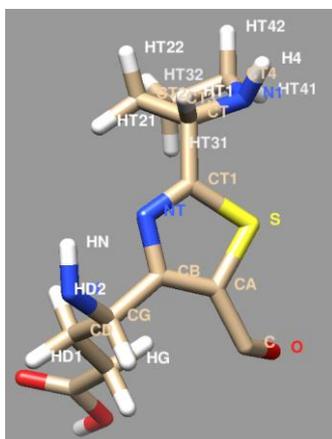


Figure 102: Example of an ATC-monomer without its N-ter and C-ter extremities after application of the ATC-optimized GAFF using Gaussian

### iii. Building an oligomer sequence and generation of the related Amber parameter files

Whatever the methods 2 or 3 considered, ATC-sequences were generated by the assembly of ATC-residues (Figure 102). Again, we developed a special script to link the ATC monomers (Figure 103). The parameter files for each ATC residue were entered in the lines 9 to 24 of the script. The oligomer sequence defined in line 29 was constructed using LEaP. Finally, the related ‘.prepin’, ‘.prmtop’ and ‘.inpcrd’ AMBER files that were to be used in the simulated annealing protocol were generated.

```

1 #Replace "name_of_sequence" by the appropriate name
2 parmchk -i ./ATC/rep/lambda.prepin -f prepi -o ./Sequences/rep/name_of_sequence1.frcmod
3
4 #Association of residues from the ATC-library without capping then minimisation of the obtained structure
5 cat > leap.in <<EOF
6 source leaprc.gaff
7 source leaprc.ff99SBnmr
8
9 #load all the prepin and frcmod files needed (every useful residue)
10 #as well as the ACE and NHE extremities modified to match the same nomenclature as ours (ACE_1 and NHE_1)
11 loadamberprep ./ATC/Librairie_ATC_sans_cap/LF1.prepin
12 loadamberprep ./ATC/Librairie_ATC_sans_cap/LF2.prepin
13
14 ...
15
16 loadamberprep ./ATC/Librairie_ATC_sans_cap/ACE_1.prepin
17 loadamberprep ./ATC/Librairie_ATC_sans_cap/NHE_1.prepin
18 loadamberparams ./ATC/Librairie_ATC_sans_cap/LF1.frcmod
19 loadamberparams ./ATC/Librairie_ATC_sans_cap/LF2.frcmod
20
21 ...
22
23 loadamberparams ./ATC/Librairie_ATC_sans_cap/ACE_1.frcmod
24 loadamberparams ./ATC/Librairie_ATC_sans_cap/NHE_1.frcmod
25
26 #change the name of the sequence and the order and name of the specified residues
27 #As many sequences as we want can be created at the same time,
28 #it is just needed to add new lines and double the following command lines to save the new sequences files
29 name_of_sequence1 = sequence { ACE LF1 LF1 LE2 LF1 LF1 NHE }
30
31 saveamberprep name_of_sequence1 ./Sequences/rep/name_of_sequence1.prepin
32 saveamberparm name_of_sequence1 ./Sequences/rep/name_of_sequence1.prmtop ./Sequences/rep/name_of_sequence1.inpcrd
33 quit
34 EOF

```

Figure 103: Generation of an ATC-sequence

## 4. Validation of the ATC-optimized force field

### i. Comparative performances of Methods 1, 2 and 3

A systematic evaluation of the three methods was performed for several oligomers. The 3D-models of ATC-oligomers were computed based on all the observed NOE cross-peaks (mandatory and optional), these are enumerated for each studied ATC-oligomer on the SI, *section VI*, 2-30. In the Figure 104, the superimposition of the lowest energy structures generated using methods 1 (pink), 2 (beige) and 3 (blue) are represented for **ATC<sub>5</sub>-3-NH<sub>2</sub>**, **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>** and **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>**. As we can observe, the last two methods lead

towards a conserved helical shape, especially for **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>** which did not converge to the C<sub>9</sub>-helix with method 1. However, by looking at the structures, one could not conclude regarding the differences and improvements between methods 2 and 3.

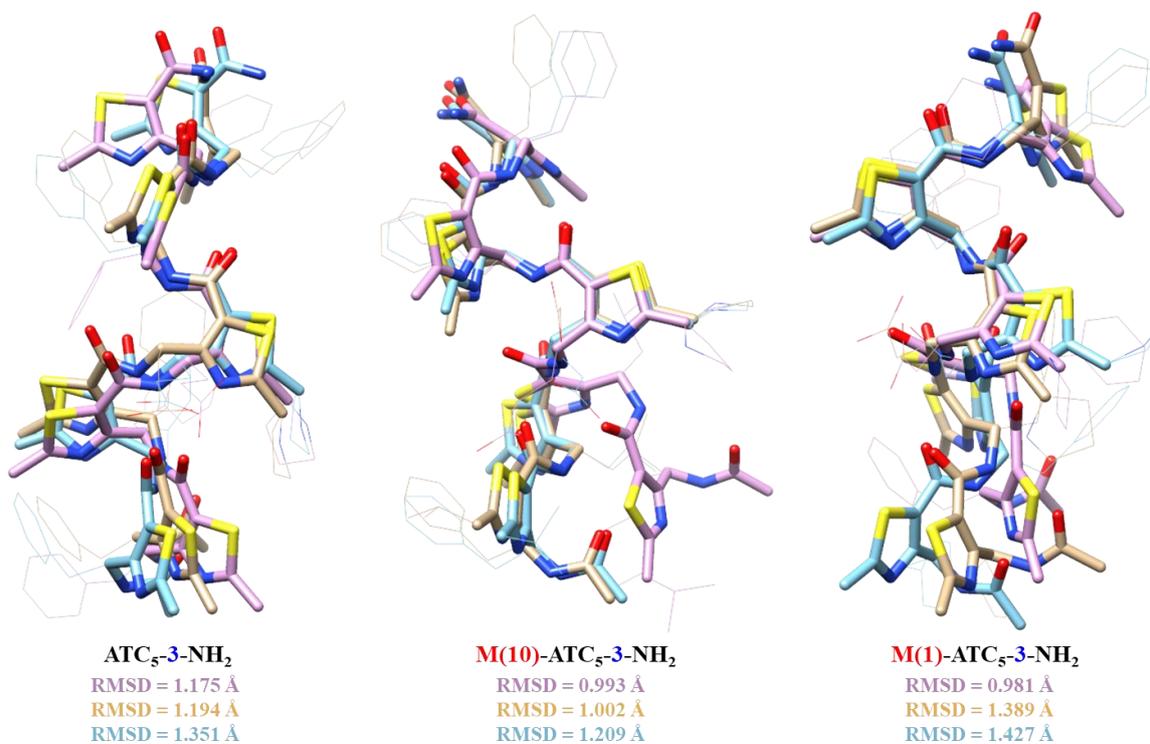


Figure 104: Superimposition of the lowest energy structures generated using methods 1 (pink), 2 (beige) and 3 (blue), for ATC<sub>5</sub>-3-NH<sub>2</sub>, M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub> and M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>

RMSD values were lower using Method 2, meaning a higher convergence in the calculated structures than with Method 3. However, using Method 2 we observed a significant deviation from the canonic C<sub>9</sub>-helix dihedral angles (values in red). In contrast to Method 2, for Method 3, the average dihedral angles were completely consistent with the establishment of a 9-helix structure (Table 24).

Table 24: Comparison of dihedral angle values obtained for the 20 best structures obtained after NMR refinement for foldamers ATC<sub>5</sub>-3-NH<sub>2</sub>, M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub> and M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub> built with monomers created with or without ATC-optimized GAFF

Canonic C <sub>9</sub> -helix			-78°	+127°	0°	-41°	H-bond <sup>c</sup> (Å)
Oligomer	Method <sup>a</sup>	Res. <sup>b</sup>	φ (°)	θ (°)	ζ (°)	ψ (°)	
ATC <sub>5</sub> -3-NH <sub>2</sub>	Method 2 GAFF	1	-73.5 ± 36.1	120.8 ± 12.6	-9.6 ± 3.4	-23.8 ± 10.1	1.91
		2	-89.9 ± 24.4	122.3 ± 9.9	-6.8 ± 5.2	-18.3 ± 14.3	1.90
		3	-101.9 ± 32.1	111.8 ± 14.3	-6.8 ± 3.9	-22.2 ± 6.4	1.85
		4	-77.1 ± 14.2	124.6 ± 5.9	-8.6 ± 2.2	-22.8 ± 2.5	1.93
		5	-108.5 ± 20.9	125.9 ± 14.2	-1.3 ± 7.0	-2.7 ± 14.8	1.93

<b>ATC<sub>5</sub>-3-NH<sub>2</sub></b>	<b>Method 3 ATC- optimized GAFF</b>	1	<b>-43.1 ± 42.2</b>	110.6 ± 18.8	<b>-5.0 ± 1.6</b>	-36.6 ± 10.9	1.88
		2	-78.1 ± 27.6	126.9 ± 8.7	-1.9 ± 3.6	-21.6 ± 34.0	1.92
		3	-75.8 ± 59.0	107.4 ± 18.1	-4.5 ± 2.3	-31.0 ± 12.4	1.85
		4	-74.3 ± 26.8	122.4 ± 9.1	-4.0 ± 0.6	-41.7 ± 5.4	1.93
		5	-77.6 ± 14.1	121.2 ± 7.6	-1.1 ± 4.3	<b>-7.1 ± 25.4</b>	1.92
<b>M(10)- ATC<sub>5</sub>-3- NH<sub>2</sub></b>	<b>Method 2 GAFF</b>	1	-66.9 ± 33.5	122.3 ± 12.9	<b>-10.3 ± 2.2</b>	<b>-26.7 ± 5.1</b>	1.99
		2	-85.4 ± 20.6	115.8 ± 5.9	<b>-6.4 ± 2.6</b>	<b>-19.6 ± 5.0</b>	1.94
		3	<b>-111.6 ± 33.8</b>	105.7 ± 18.8	<b>-6.1 ± 5.0</b>	<b>-24.3 ± 8.8</b>	2.05
		4	-76.6 ± 9.6	134.2 ± 6.5	<b>-6.5 ± 2.1</b>	<b>-23.3 ± 1.9</b>	2.20
		5	<b>-136.1 ± 6.5</b>	139.7 ± 7.2	<b>6.4 ± 2.0</b>	<b>9.2 ± 2.9</b>	-
<b>M(10)- ATC<sub>5</sub>-3- NH<sub>2</sub></b>	<b>Method 3 ATC- optimized GAFF</b>	1	-52.4 ± 34.2	116.9 ± 16.1	<b>-5.2 ± 0.9</b>	-39.7 ± 8.0	1.93
		2	-67.6 ± 26.2	114.6 ± 7.7	-4.4 ± 1.6	-31.1 ± 7.8	1.93
		3	-90.8 ± 42.3	114.4 ± 11.5	<b>-5.0 ± 1.8</b>	-44.9 ± 8.3	1.90
		4	-71.6 ± 7.3	130.5 ± 15.1	-1.5 ± 1.5	-37.6 ± 6.0	2.12
		5	<b>-120.2 ± 19.2</b>	135.5 ± 14.0	3.5 ± 3.0	<b>6.4 ± 17.7</b>	2.14
<b>M(1)- ATC<sub>5</sub>-3- NH<sub>2</sub></b>	<b>Method 2 GAFF</b>	1	<b>-40.7 ± 56.8</b>	107.2 ± 21.8	<b>-7.9 ± 2.5</b>	<b>-21.2 ± 4.9</b>	1.90
		2	<b>-115.2 ± 30.0</b>	120.2 ± 19.5	-4.3 ± 6.1	<b>-11.6 ± 17.0</b>	1.92
		3	<b>-112.5 ± 42.6</b>	102.6 ± 17.5	<b>-6.2 ± 4.1</b>	<b>-20.0 ± 8.0</b>	1.85
		4	-81.2 ± 20.6	124.2 ± 7.5	<b>-8.0 ± 2.0</b>	<b>-23.5 ± 2.1</b>	1.88
		5	-99.9 ± 26.6	129.2 ± 10.0	1.9 ± 6.8	<b>-4.2 ± 15.0</b>	1.90
<b>M(1)- ATC<sub>5</sub>-3- NH<sub>2</sub></b>	<b>Method 3 ATC- optimized GAFF</b>	1	-53.8 ± 34.1	98.4 ± 53.5	-4.5 ± 1.9	-36.0 ± 13.7	1.90
		2	-77.2 ± 29.3	116.6 ± 7.1	-3.3 ± 2.1	-32.7 ± 17.6	1.85
		3	-59.8 ± 57.3	113.9 ± 12.5	-4.3 ± 1.6	-37.6 ± 8.2	1.86
		4	-71.3 ± 26.1	120.0 ± 6.2	-4.3 ± 0.9	-39.9 ± 4.3	1.91
		5	-66.6 ± 10.0	125.0 ± 10.8	-1.2 ± 3.8	<b>-6.6 ± 26.4</b>	1.91

<sup>a</sup>Method applied during the generation of ATC-monomers, <sup>b</sup>Residues are ordered from N to C-terminal extremities, <sup>c</sup>Average length of H-bond formed between CO<sub>i</sub> and NH<sub>i+2</sub>

In summary, the procedure could be divided into 8 steps:

- Design of ATC-monomers using dihedral angles found on an ATC-monomer crystal (*Software UCSF Chimera*)
- Assignment of atomic charges by DFT calculations at the M062X/6311++g(d) level of theory (*Gaussian*)
- Generation by Antechamber of the AMBER parameter files for each ATC monomer by combining the atomic charges obtained by DFT and the other atom properties as defined in GAFF ff99SBnmr (*AMBER16*)
- Construction of the oligomer sequences using LEaP (*AMBER16*)
- Introduction of the distance restraints resulting from the study of the NMR spectra and NOE cross-peaks (*TopSpin and NMRView*)
- Cooking and simulated annealing in vacuum of each oligomer (*AMBER16*)

- Selection of the 20 lowest energy structures without the violations being greater than 0.3 Å (*MOLMOL*)
- Exploitation and observation of the structures (*UCSF Chimera*)

## ii. Validation of the calculated 3D models by comparison with the XRD structure, CD and FTIR data

Finally, the foldamer **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>** crystallized by slow evaporation of an isopropanol solution (crystallographic data of this oligomer can now be obtained free of charge from The Cambridge Crystallographic Data Centre CCDC 1894172). Therefore, a careful comparison of the XRD structure permitted to validate the calculated models. This ATC-oligomer self-assembled into infinite columns linked together by hydrogen bonds between the acid and pyrrolidine side chains (Figure 105 and SI, *section IX*). It cannot be discounted that such intermolecular association could have occurred in solution, thus explaining the minor species observed in NMR spectra as previously discussed.

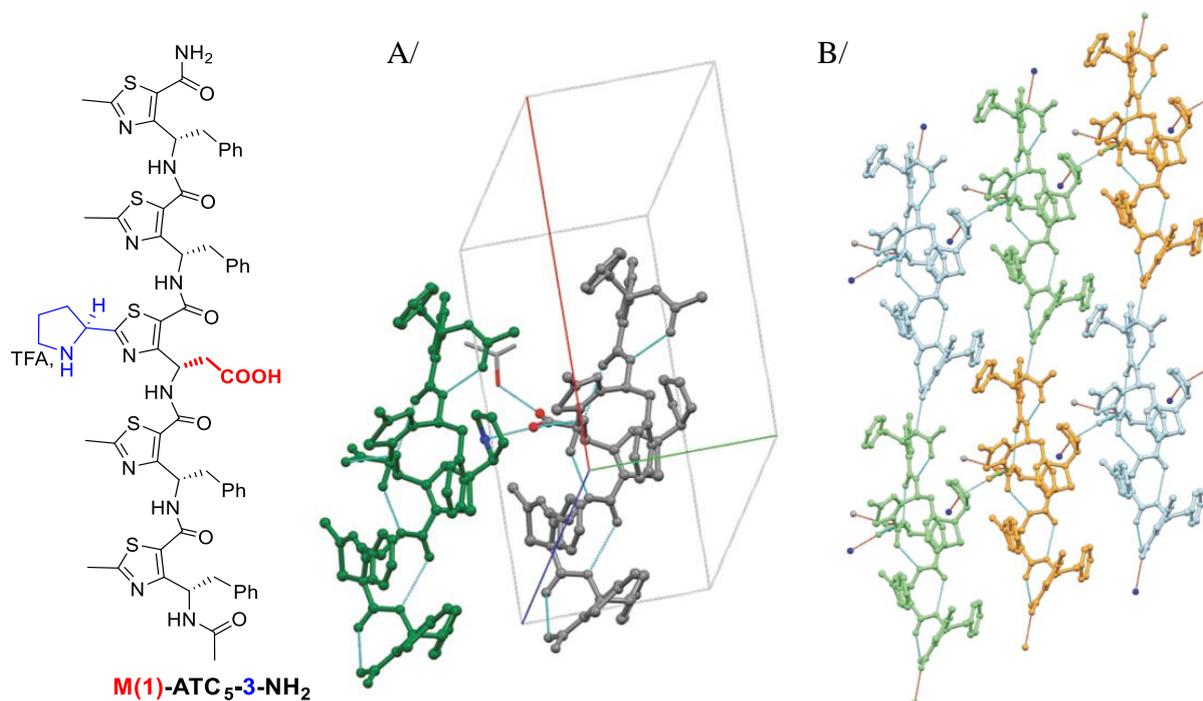


Figure 105: Part of the crystal packing of **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>**. A/The view highlights the asymmetric unit of **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>** that consists of two independent ATC oligomers (grey and green) and one *iPrOH* (wireframe). Hydrogen atoms are not indicated for clarity. B/ATC oligomers self-assemble in infinite columns linked together by hydrogen bonds between the acid and pyrrolidine side chains.

Moreover, it could be observed that the ATC  $\gamma$ -peptide **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>** adopts the same overall helical structure in solution and in a solid state (Figure 106, Table 25). Its folding is defined by an intramolecular hydrogen-bonding pattern found between the carbonyl of the residue  $i$  and the proton amide of the residue  $i+2$ , the C<sub>9</sub>-helix seems preserved in both states. As previously described (*Chapter 4, I and SI, VII and VIII*), the folding is perfectly consistent with the data obtained by CD and FTIR.

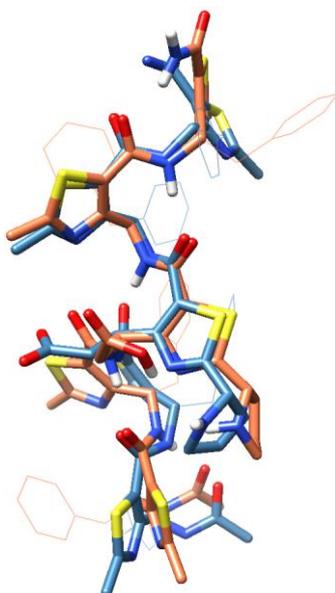


Figure 106: Superimposition of the XRD (blue) and lowest energy NMR structures of **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>** generated by the Method 3 (coral)

Table 25: Average backbone torsion angles for 9-helical fold  $\gamma$ -peptides **ATC<sub>5</sub>-3-NH<sub>i</sub>Pr**, **ATC<sub>5</sub>-3-NH<sub>2</sub>** and **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

	$\phi$	$\theta$	$\zeta$	$\psi$
Average NMR values				
<b>ATC<sub>5</sub>-3-NH<sub>i</sub>Pr</b>	$-71 \pm 17^\circ$	$119 \pm 4^\circ$	$-4 \pm 1^\circ$	$-35 \pm 6^\circ$
<b>ATC<sub>5</sub>-3-NH<sub>2</sub></b>	$-70 \pm 15^\circ$	$118 \pm 8^\circ$	$-3 \pm 2^\circ$	$-28 \pm 13^\circ$
<b>M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	$-90 \pm 30^\circ$	$116 \pm 11^\circ$	$-5 \pm 4^\circ$	$-17 \pm 7^\circ$
XRD structure				
<b>M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	$-83 \pm 11^\circ$	$117 \pm 8^\circ$	$-1 \pm 2^\circ$	$-36 \pm 15^\circ$

### III. APPLICATION OF THE OPTIMIZED SIMULATED ANNEALING PROTOCOLE TO THE STRUCTURAL DETERMINATION OF A SET OF CATALYTIC ATC-OLIGOMERS

Ultimately, Method 3 was used for the structural determination of other catalytic ATC-oligomers as it had been recognized as being the most relevant method for their 3D-modelling.

#### 1. Structures according to the position of the catalytic residue

We first solved the structures of **ATC<sub>5</sub>-2-NH<sub>2</sub>**, **ATC<sub>5</sub>-3-NH<sub>2</sub>**, **ATC<sub>5</sub>-4-NH<sub>2</sub>** and **ATC<sub>5</sub>-5-NH<sub>2</sub>**. Whatever the position of the catalytic residue along the backbone, the oligomers displayed the same 9-helical structure (Figure 107). Once again, these results were in accordance with the FTIR and CD data. In the case of **ATC<sub>5</sub>-1-NH<sub>2</sub>**, the structure was not solved since it precipitated into *i*PrOH, making the NMR study impossible. Nevertheless, based on FTIR and CD experiments, it was assumed that it adopted the same overall folding as the other oligomers.

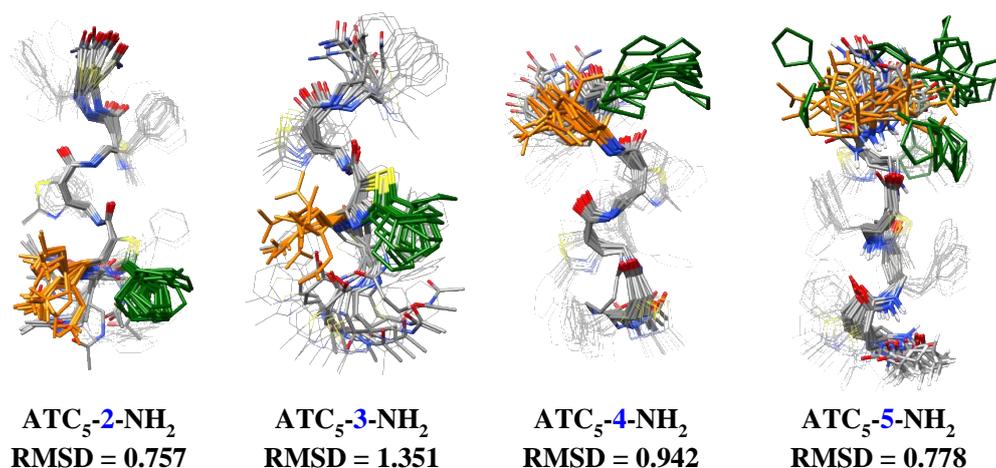


Figure 107: Superimposition of the 20 lowest energy NMR solution structures of **ATC<sub>5</sub>-2-NH<sub>2</sub>**, **ATC<sub>5</sub>-3-NH<sub>2</sub>**, **ATC<sub>5</sub>-4-NH<sub>2</sub>** and **ATC<sub>5</sub>-5-NH<sub>2</sub>** in *i*PrOH-*d*<sub>7</sub>. RMSD (Å) corresponds to the root mean square deviation value calculated on the heavy backbone atoms.

#### 2. Limitations of the method

As presented previously, **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>** displayed the same C<sub>9</sub>-helical folding in both the solution and the solid state. The two oligomers, **M(2)-** and **M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>**, in which the catalytic pyrrolidine and the carboxylic acid are born by two different ATC monomers, also folded into a C<sub>9</sub>-helix. However, the backbones curved around the 3<sup>rd</sup> residue (Figure 108). A careful analysis of the dihedral angles of **M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>**, presented in the supporting information (*section VI, 34*), showed that  $\varphi$  of residue 3 was equal to  $-12.8 \pm 30^\circ$ . This value is

far from the expected  $-78^\circ$  for a canonic  $C_9$ -helix. However, a variation around the 3<sup>rd</sup> residue was confirmed by looking at the coupling constant  $^3J(\text{NH}, \gamma\text{CH})$ , which was of 4.7 Hz, instead of 6.0-7.0 Hz in other oligomers (SI, *section VI, 11*). In addition, analysis of the 20 lowest energy structures calculated from NMR data suggests some labilities of the  $C_9$ -H bond around ATC 3. This is in agreement with the FTIR data (SI, *section VIII*), in which the contribution of bound CO (50% of all the C=O) is lower than those expected (62.5% of all the C=O). It is actually unclear whether the folding dynamic around ATC 3 could explain the observed decrease of the enantiomeric excess ( $ee < 70\%$ ) in the nitro-Michael addition reaction catalysed by **M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>**.

In **M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>**, the dihedral angle  $\varphi$  of ATC 3 also differed from the expected  $-78^\circ$  ( $-34.5 \pm 27.4^\circ$ ). The coupling constant  $^3J(\text{NH}, \gamma\text{CH})$  was measured at 7.6 Hz, meaning that this angle should be around  $-87^\circ$ . In the same manner, **M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub>** had its ATC 3 dihedral angle  $\varphi$  at  $-34.7 \pm 72.6^\circ$ , moreover, its  $^3J(\text{NH}, \gamma\text{CH})$  was measured at 7.8 Hz. In those cases, the comparison between the coupling constant of this residue 3, associated with its dihedral angle  $\varphi$ , is not relevant for the expected models. We believe that we might have reached the limitations of the method and that the observed coupling constant would not reflect a single conformational state but more a dynamic one around the ATC 3.

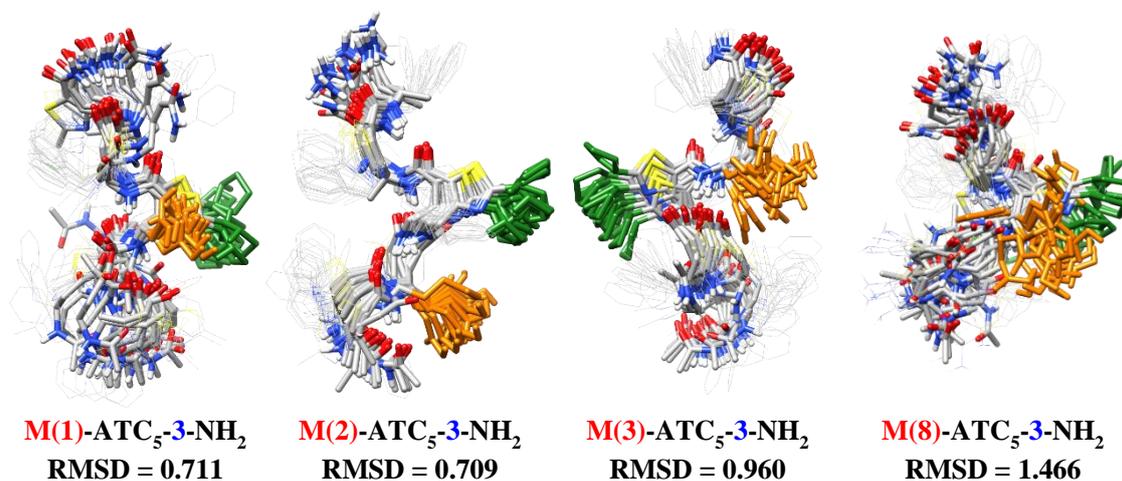


Figure 108: Superimposition of the 20 lowest energy NMR solution structures of **M(1)-**, **M(2)-**, **M(3)-** and **M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>*. RMSD (Å) corresponds to the root mean square deviation value calculated on the heavy backbone atoms.

### 3. Structural validation of other $M(X)$ -Mutant oligomers

Eventually, other ATC-oligomers were studied using 3D-modelling and  $M(5)$ -,  $M(9)$ - and  $M(10)$ -ATC<sub>5</sub>-3-NH<sub>2</sub> were fully consistent with the canonic C<sub>9</sub>-helix (Figure 109).

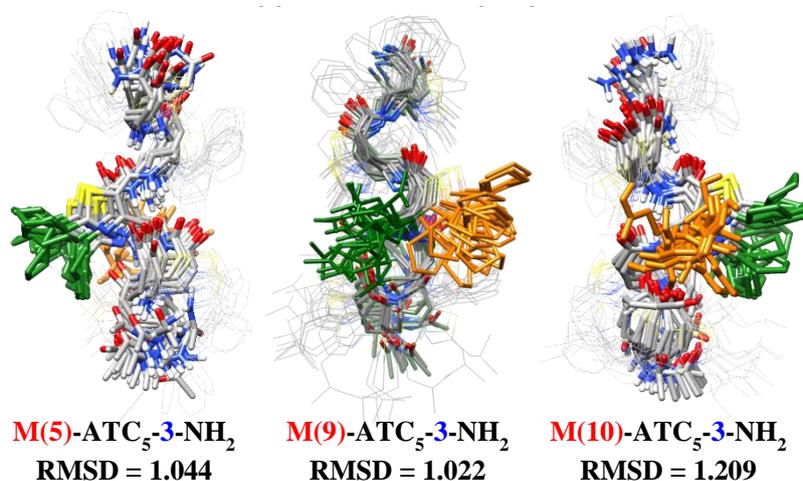


Figure 109: Superimposition of the 20 lowest energy NMR solution structures of  $M(5)$ -,  $M(9)$ - and  $M(10)$ -ATC<sub>5</sub>-3-NH<sub>2</sub> in *i*PrOH-*d*<sub>7</sub>. RMSD (Å) corresponds to the root mean square deviation value calculated on the heavy backbone atoms.

Furthermore, trimer  $M(16)$ -ATC<sub>3</sub>-2-NH<sub>2</sub>, pentamer  $M(13)$ -ATC<sub>5</sub>-3-NH<sub>2</sub> and heptamer  $M(17)$ -ATC<sub>7</sub>-4-NH<sub>2</sub> adopted the same 9-helical structure, regardless of the length of the oligomer (Figure 110). In the figure it might be challenging to visualize perfectly the shape of this helix on the trimer, however, it really folds into a well-defined 9-helical structure even though, in such a case, the helix is only one turn long.

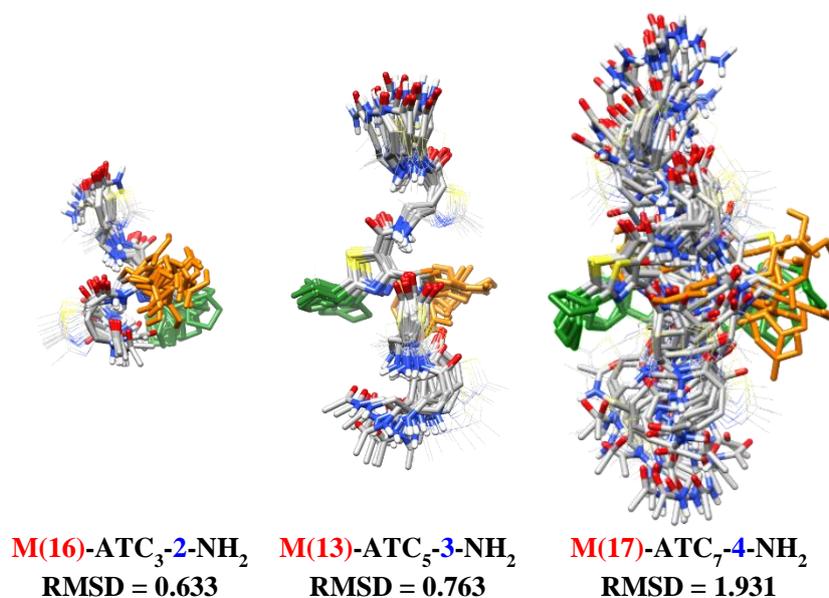


Figure 110: Superimposition of the 20 lowest energy NMR solution structures of trimer  $M(16)$ -ATC<sub>3</sub>-2-NH<sub>2</sub>, pentamer  $M(13)$ -ATC<sub>5</sub>-3-NH<sub>2</sub> and heptamer  $M(17)$ -ATC<sub>7</sub>-4-NH<sub>2</sub> in *i*PrOH-*d*<sub>7</sub>. RMSD (Å) corresponds to the root mean square deviation value calculated on the heavy backbone atoms.

Finally, the last four oligomers (**M(18)**- to **M(21)**-), which bear several catalytic ATC residues did not entirely converge towards the C<sub>9</sub>-helix (Figure 111). In fact, the definition of the ROESY spectra of these four oligomers was not sufficient to provide substantial NOE cross-peaks and most of them overlapped, making them almost impossible to integrate. Hence, these models probably cannot fully reflect the structures of these molecules.

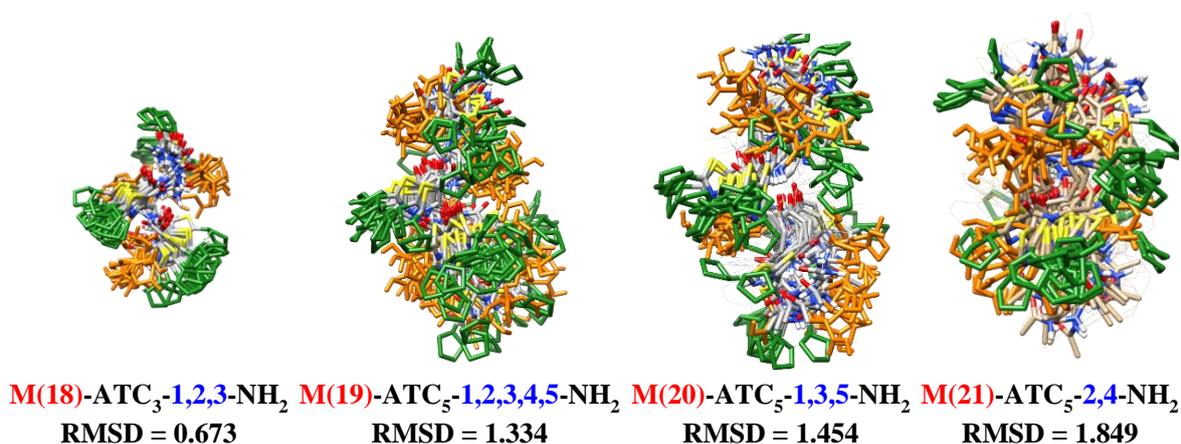


Figure 111: Superimposition of the 20 lowest energy NMR solution structures of **M(18)**-ATC<sub>3</sub>-**1,2,3**-NH<sub>2</sub>, **M(19)**-ATC<sub>5</sub>-**1,2,3,4,5**-NH<sub>2</sub>, **M(20)**-ATC<sub>5</sub>-**1,3,5**-NH<sub>2</sub> and **M(21)**-ATC<sub>5</sub>-**2,4**-NH<sub>2</sub> in *i*PrOH-*d*<sub>7</sub>. RMSD (Å) corresponds to the root mean square deviation value calculated on the heavy backbone atoms.

#### IV. CONCLUSION

During this whole project, it was quite challenging to investigate the relationship between the structure and the catalytic activity. However, in this last section, I will try to provide some elements to enable the correlation of the obtained results to some extent.

In fact, the X-ray structure of oligomer **M(1)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub> provided information regarding the distance between the pyrrolidine NH and the carboxylic acid, being of about 7.1 Å (Figure 112). Furthermore, if we compare **M(1)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub> to ATC<sub>5</sub>-**3**-NH<sub>2</sub>, **M(2)**- and **M(3)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub>, the only difference is the length of the lateral chain bearing the carboxylic acid function. The average distance between the two key functions for ATC<sub>5</sub>-**3**-NH<sub>2</sub>, **M(2)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub> and **M(3)**-ATC<sub>5</sub>-**3**-NH<sub>2</sub> (measured from 3D-models) were of 7.9 Å, 8.1 Å and 11.9 Å, respectively.

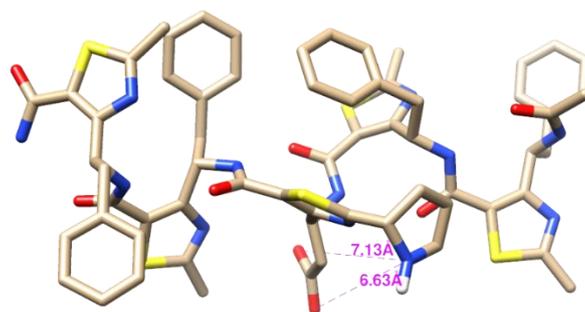


Figure 112: Representation of the distance between pyrrolidine NH and carboxylic acid in crystal structures of **M(1)-ATC5-3-NH<sub>2</sub>** (CCDC 1894172)

In contrast, it seems that Wennemers tripeptide H-D-Pro-Pro-Glu-NH<sub>2</sub> is quite flexible. In fact, from the crystallographic data provided by Wennemers *et al.*, it appeared that the distance between the nitrogen of the 1<sup>st</sup> proline and the oxygen of the carboxylic acid was of 9.9 Å (structure in **green**, Figure 113, A).

On the other hand, a conformational exploration of the peptide along the reactional pathway lead by DFT calculations<sup>[184]</sup> permitted to observe that when the tripeptide was in presence of the substrates, a different conformation was adopted. Figure 113 displays the conformations adopted at a resting state (B, in **beige**), during the enamine formation (C, in **blue**) and for the C-C bond formation (D, in **pink**). Therefore, the distances between the key functions are quite different when looking at the crystal structure or at the catalytic process.

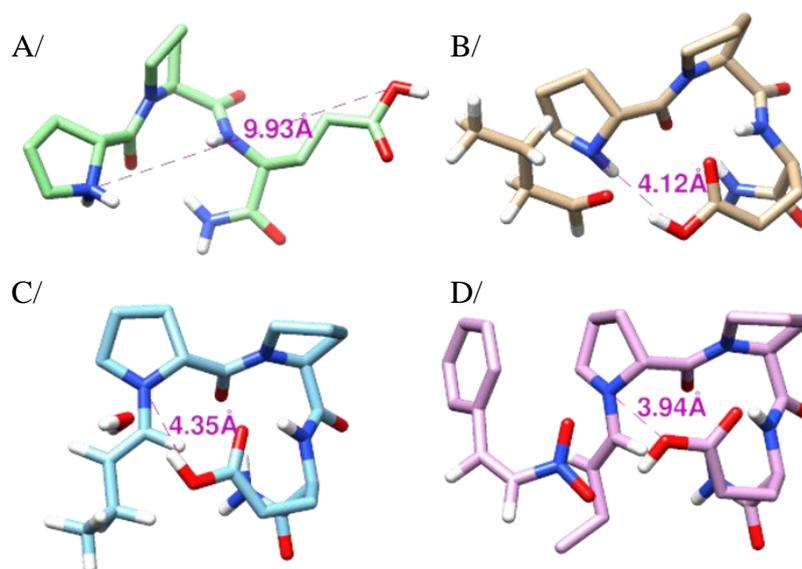


Figure 113: A/Representation of the distance between pyrrolidine NH and carboxylic acid in crystal structures of H-D-Pro-Pro-Glu-NH<sub>2</sub> (CCDC 1556800)<sup>[187]</sup>, Plausible pathways for the H-D-Pro-Pro-Glu-NH<sub>2</sub>-catalyzed conjugate addition reaction of butanal to  $\beta$ -nitrostyrene into (2S,3R)-2-ethyl-4-nitro-3-phenylbutanal in chloroform B/Resting state, C/Enamine formation, D/C-C bond formation<sup>[184]</sup>

Regarding the ATC-oligomers, there was not an obvious correlation between the pyrrolidine-carboxylic acid distance and the observed enantioselectivity in the nitro-Michael reaction (Table 26).

Table 26: Enantiomeric excesses obtained for the nitro-Michael addition reaction of cyclohexanone to *trans*-, *p*-methyl- and *p*-methoxy-nitrosytrene in respects to the distance between pyrrolidine NH and carboxylic acid for  $ATC_5-3-NH_2$ ,  $M(1)-ATC_5-3-NH_2$ ,  $M(2)-ATC_5-3-NH_2$  and  $M(3)-ATC_5-3-NH_2$ .

Oligomer	Distance <sup>a</sup>	Enantiomeric excess (%)		
		R=H	R=CH <sub>3</sub>	R=OCH <sub>3</sub>
$ATC_5-3-NH_2$	7.9 Å	79	79	79
$M(1)-ATC_5-3-NH_2$	7.1 Å <sup>b</sup>	76	70	75
$M(2)-ATC_5-3-NH_2$	8.1 Å	62	59	63
$M(3)-ATC_5-3-NH_2$	11.9 Å	75	75	78

<sup>a</sup>Average distance between pyrrolidine NH and carboxylic acid from the 20 lowest energy structures, <sup>b</sup>distance taken from crystal structure

Nevertheless, we can notice that for  $M(2)-ATC_5-3-NH_2$ , the enantiomeric excesses were lower for any considered substrate than for the other oligomers. This could have been due to a lack of flexibility between the chains and therefore the substrates might not have had a proper access to the catalytic site. If we have a close look at the 3D-models (Figure 114), we can observe that for  $M(2)-ATC_5-3-NH_2$ , the key functions were completely above one another, while for the other oligomers, these functions were not located on the same face of the oligomer. This could be a part of the explanation for the drop in enantioselectivity.

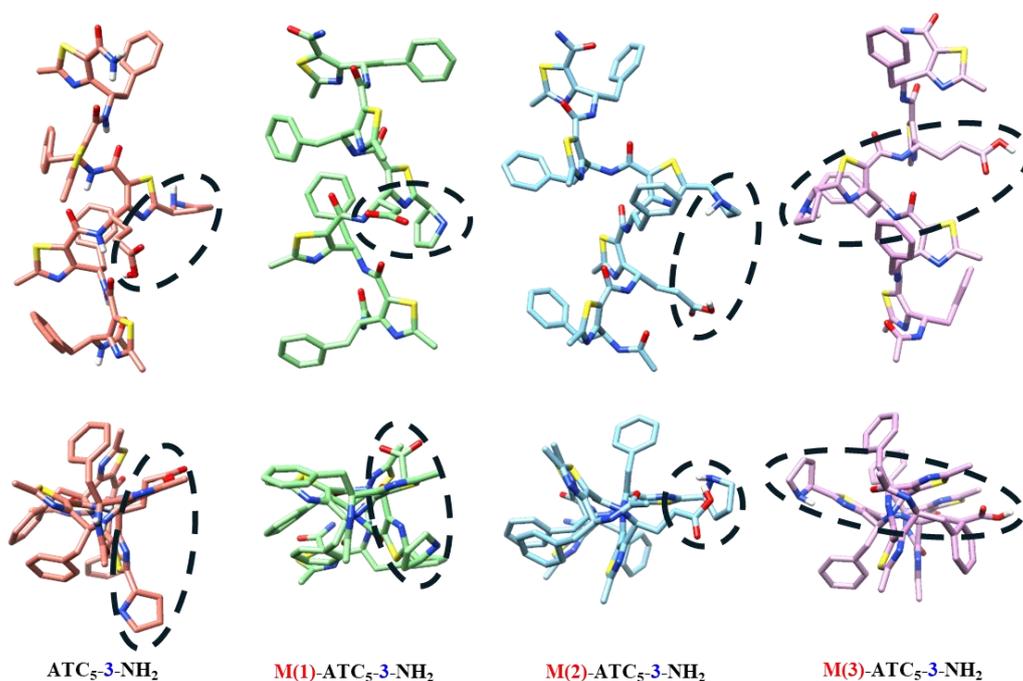


Figure 114: Position of the two key functions on the oligomers  $ATC_5-3-NH_2$  and  $M(1)-$  to  $M(3)-ATC_5-3-NH_2$ , along the axis (up) and above the axis (down), with the “catalytic site” circled in black.

In conclusion, it appears that the ATC-oligomers have less flexibility regarding their catalytic centre than the tripeptides proposed by Wennemers. This lack of adaptability could explain, to some extent, why the ATC catalysts were less enantioselective than their peptidic counterpart. Hence, the distance and flexibility between the key functions appear to be of prime importance.

## CONCLUSION AND OUTLOOK



In 2016, our team reported a robust ATC synthesis pathway providing an easy way to modulate the lateral chains, either on the  $\gamma$ -carbon atom or on the thiazole core of the amino acids. These monomers were designed to exhibit a conformational limitation around the  $C_\alpha$  and  $C_\beta$ , blocking the angle  $\zeta$  at about  $0^\circ$  to reduce the flexibility of the residue. Moreover, it was established that ATC monomers bearing a pyrrolidine on their thiazole ring and a carboxylic acid on their  $\gamma$ -lateral chain were able to catalyse enamine-type reactions and especially the nitro-Michael addition reaction. We were first able to prove that a rise of the structuration on an ATC monomer bearing catalytic functions was increasing the asymmetric induction in the reaction of cyclohexanone to  $\beta$ -*trans*-nitrostyrene. Therefore, the main objective of this project was to develop foldamers bearing a catalytic activity and observe the influence of their size and/or shape, on both, reactivity and enantioselectivity.

A large number of ATC-monomers was synthesized in solution, they differed in their lateral chains, either on their  $\gamma$ -lateral chain or at the position 2 of their thiazole ring. We then obtained, by solid phase peptide synthesis, twenty-six different ATC-oligomers which were tested on the nitro-Michael addition reaction of cyclohexanone to  $\beta$ -*trans*-nitrostyrene, *p*-methyl-nitrostyrene and *p*-methoxy-nitrostyrene. They showed a potential as catalysts. Yields were generally around 95% after 16 hour reaction, and the selectivity was towards the *syn* compound, regardless of the pyrrolidine configuration. For the pentamers, the enantiomeric excesses were around 79%. The modulation of the remote lateral chains displayed a chemoselectivity that we had not expected at the beginning of this project but demonstrated a substrate selectivity for some oligomers, mimicking some enzymes properties.

A study combining NMR, FTIR and circular dichroism permitted to evaluate the structural markers characteristic of the ATC-oligomers. This enabled to show that ATC-oligomers fold into a  $C_9$ -helix, stabilized by hydrogen bonds between  $CO_i$  and  $NH_{i+2}$ . NMR studies demonstrated that a high unshield of the amide protons ( $> 9$  ppm) is associated with this structure, and FTIR experiments showed that the study of Amide I band was relevant to quantify the bounded and free carbonyls of the sequence. A study of the circular dichroism signatures confirmed to us the conformational stability of these oligomers.

Finally, the creation of a script pipeline for molecular dynamic associated with simulated annealing, based on the integration of NOE cross-peaks allowed to determine the structure of the foldamers on an atomic level. Moreover, an X-ray study of an ATC-oligomer proved that the  $C_9$ -helix adopted by ATC-oligomers was conserved in a solid state.

Eventually, it was established that not only the length of the oligomer and the microenvironment around the catalytic centre influence the catalysis, but also the distance between the carboxylic acid and the pyrrolidine seems to play a role in the activity and the enantiomeric excess. It appears that a balance between flexibility and rigidity must be found within the catalytic functions.

Since ATCs are built around a thiazole ring, the scaffold can be modified for the construction of thiazolium containing oligomers that could be of value to design a new family of NHC-organocatalysts. The use of a coenzyme chimera in the form of an ATC building block to mimic thiamine diphosphate dependent enzymes could be an interesting outlook for this project. In fact, the development of such a prosthetic strategy would offer a straightforward way to implement catalytic functionality into foldamer constructs. Thiazolium ATC analogues, also named ATzCs could be included in reverse-turns or twist-turns folded  $\alpha$ - $\gamma$ - $\alpha$ - $\alpha$  hybrid tetrapeptides. Their catalytic activities could, for instance, be evaluated on benzoin condensation and Stetter reaction.

Finally, it could be possible to study ATC-based systems that adopt a tertiary structure mimicking proteins, not only in terms of their structure but also in size. Therefore, a design of a true substrate-binding pocket represents the next challenge for the conception of an enzyme-like catalyst.

## RESUME EN FRANÇAIS



## I. INTRODUCTION ET OBJECTIFS DE LA THESE

Les contraintes environnementales, économiques et sociologiques actuelles invitent nos sociétés à repenser les structures et technologies de production afin de favoriser l'utilisation de ressources renouvelables, de proposer des procédés économes en énergie et de limiter la production de déchets. Dans un tel contexte, ces nouvelles exigences passent en particulier par la mise au point et l'exploitation de nouveaux systèmes catalytiques pour la chimie. L'une des approches pour développer ces catalyseurs verts consiste à s'inspirer de la nature pour inventer, en mariant biologie et chimie, des biocatalyseurs et des systèmes de catalyse dits bio-mimétiques ou bio-inspirés.<sup>[1]</sup>

Ainsi, au cours des dernières décennies la détermination structurale d'un nombre toujours croissant de protéines a mis en exergue l'importance du contrôle conformationnel pour expliquer l'activité et la sélectivité des enzymes. En résumé, la structure est le support de l'activité. Actuellement, on compte plus d'une centaine de procédés industriels faisant intervenir des biocatalyseurs. Cependant, dans de nombreux cas, l'instabilité enzymatique limite leur utilisation et les enzymes doivent alors être remplacées par des catalyseurs artificiels, souvent moins sélectifs et parfois toxiques pour l'environnement. Aspirant à imiter l'efficacité enzymatique, un des objectifs actuels de la chimie vise à concevoir des catalyseurs artificiels reproduisant certaines propriétés de structure et de fonction des enzymes. Les peptides se sont révélés être des candidats prometteurs pour développer de telles plateformes de catalyse bio-inspirée. Les méthodes combinatoires et le design de novo ont conduit à de réels succès pour une large variété de réactions, comme les résolutions cinétiques d'alcools, réactions d'aldolisation, de phosphorylation, d'addition de Michael, réactions de Morita-Baylis-Hillman, etc.<sup>[129]</sup> Toutefois, l'activité catalytique et la sélectivité de la réaction sont souvent limitées par la difficulté à contrôler précisément la topologie du catalyseur peptidique. Plus récemment, il a été proposé de construire des systèmes de catalyse bio-inspirée à partir de plateformes hautement structurées, appelés foldamères, différents des peptides et oligonucléotides que la nature n'a pas explorés, ni sélectionnés.<sup>[134,139,52]</sup> Potentiellement, de tels catalyseurs peuvent présenter de nombreux avantages et pourraient être une solution originale pour répondre aux doubles enjeux économiques et environnementaux de la chimie de synthèse. Faciles à synthétiser et peu coûteux à produire à grande échelle, ils peuvent, contrairement aux enzymes dont ils s'inspirent, être utilisés dans de nombreux solvants organiques et dans une large gamme de températures ou de pressions. Par ailleurs, l'activité et la spécificité catalytique peuvent être

modulées facilement en changeant la nature du groupement catalytique et/ou en modulant la topologie du squelette foldamérique et le microenvironnement autour du centre catalytique.

C'est dans ce cadre que notre équipe s'est intéressée au développement d'architectures pseudo-peptidiques capables d'adopter en solution des états conformationnels stables et prévisibles. Mon projet de thèse a concerné le développement de catalyseurs artificiels reproduisant certaines propriétés de structure et de fonction des enzymes. Nous avons ainsi cherché à mimer les éléments de structures secondaires des protéines, i.e. feuillet- $\beta$  et hélice- $\alpha$  pour présenter de manière contrôlée en trois dimensions des éléments clés pour la catalyse.

Durant les années précédant mon arrivée, l'équipe que j'ai intégrée a décrit la synthèse et la structure d'une nouvelle famille de  $\gamma$ -aminoacides hétérocycliques contraints, des acides 4-amino(méthyl)-1,3-thiazole-5-carboxyliques, appelés ATCs (Figure 52). Ces monomères sont construits autour d'un thiazole afin de contraindre, aux alentours de  $0^\circ$ , la flexibilité de la liaison  $^{\alpha}\text{C}-^{\beta}\text{C}$ .<sup>[103]</sup> De plus, ces monomères présentent l'avantage de pouvoir être modifiés en deux positions, sur leur chaîne  $\gamma$  et en position 2 du thiazole.

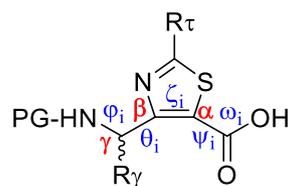


Figure 115: Structure générale d'un monomère d'ATC

Des travaux antérieurs ont permis de vérifier que, placé au cœur d'une séquence peptidique, un ATC induisait la formation d'un coude stabilisé par une liaison hydrogène intramoléculaire à 9 chaînons, permettant de structurer le peptide dans une conformation en épingle à cheveux. Dans le cas d'oligomères d'ATCs, il a été démontré que la conformation plane des angles de torsion  $^{\gamma}\text{C}-^{\beta}\text{C}-^{\alpha}\text{C}-\text{C}(\text{O})$  imposée par le thiazole tend à favoriser la formation d'une hélice droite résultant de l'établissement de liaisons hydrogènes intramoléculaires formant des pseudo-cycles en  $\text{C}_9$ . Ces liaisons s'établissent entre les carbonyles des résidus  $i$  et les NH des résidus  $i+2$ . Ainsi, l'utilisation de tels  $\gamma$ -aminoacides cycliques dans un oligomère permettrait de réduire la flexibilité du squelette carboné et d'améliorer la stabilité de la structure hélicoïdale. Seuls deux ATCs sont nécessaires pour induire l'hélice dont un tour est obtenu avec trois résidus. Les six chaînes latérales par tour d'hélice, réparties tous les  $60^\circ$ , permettent de moduler les propriétés physicochimiques des oligomères. Le pas de l'hélice est de  $11.8 \text{ \AA}$  et cette topologie existe tant dans les solvants organiques, chloroforme et méthanol que dans l'eau. Il est intéressant de noter

que la conformation en C<sub>9</sub> avait été prédite par modélisation moléculaire par Hofmann et al.<sup>[85]</sup> pour des oligomères de  $\gamma$ -aminoacides Z-vinylogues, c'est-à-dire présentant une insaturation entre les carbones  $\alpha$  et  $\beta$ . Une hélice 7 avait aussi été proposée mais elle n'a pas encore été mise en évidence pour les ATCs. Il est possible qu'une interaction dipolaire ou une délocalisation électronique  $n \rightarrow \sigma^*$  entre les carbonyles et les thiazoles (Figure 116, A) orientent le repliement vers la formation de l'hélice 9.

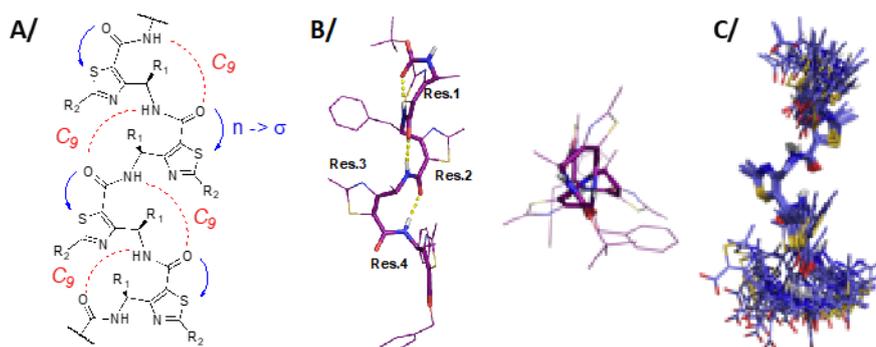


Figure 116 : A/ Réseau de liaisons hydrogènes au sein des oligomères d'ATCs. B/ Structure cristallographique d'un tétramère d'ATC. C/ Structure obtenue par RMN d'un hexamère d'ATCs dans l'eau.

Dans un tel contexte, le projet que je présente ici vise à introduire des propriétés de catalyse au sein des oligomères hélicoïdaux d'ATCs et à étudier l'impact de la conformation et du microenvironnement autour du centre catalytique sur la chimio-sélectivité et sur le transfert de chiralité. A terme, ces édifices pourront servir de plateformes pour la conception d'enzymes synthétiques.

## II. INTRODUCTION D'UNE FONCTION CATALYTIQUE SUR UN MONOMERE D'ATC

### 1. Design d'un monomère d'ATC portant une fonction de catalyse

Dans un premier temps, nous nous sommes intéressés à l'ajout d'une fonction catalytique sur un monomère d'ATC. En 1971, les groupes de Hajos et Parrish et d'Eder, Sauer et Wiechert ont démontré que la (*S*)-proline pouvait être utilisée comme catalyseur pour des réaction d'aldolisation.<sup>[142,143,144,146]</sup> Dans les années 1980, le groupe de C. Agami a démontré que ce type de mécanisme implique la formation d'un intermédiaire énamine.<sup>[149]</sup> Par la suite, de nombreuses études ont permis de démontrer que la L-Proline possédait la capacité de catalyser,

de manière asymétrique, des réactions telles que les  $\alpha$ -alkylations, réactions de Mannich, additions de Michael et  $\alpha$ -aminations de cétones.<sup>[155,157]</sup>

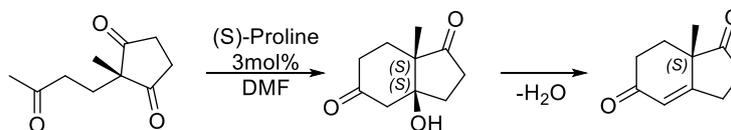


Schéma 1 : Réaction de Hajos-Parrish-Eder-Sauer-Wiechert<sup>[145]</sup>

Ce type de mécanisme a été exploré par Houk en 2004 pour des réactions de cyclisations intramoléculaires d'aldols catalysées par la proline<sup>[145]</sup> et également étudié, la même année, par l'équipe de List.<sup>[147]</sup>

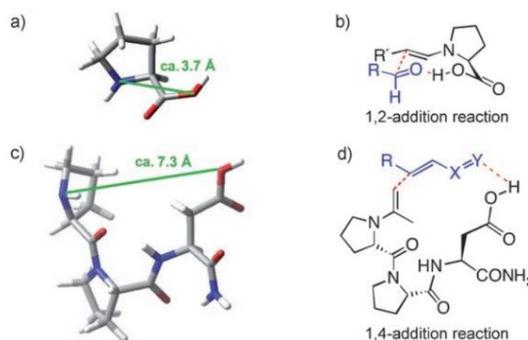


Figure 117: a) Conformation de plus basse énergie pour la proline<sup>[173]</sup>. b) Etat de transition de la proline proposé par Houk et List<sup>[145,147]</sup>. c) Conformation de plus basse énergie du tripeptide de Wennemers.<sup>[171,173]</sup> d) Etat de transition pour l'addition 1,4 catalysée par le tripeptide de Wennemers<sup>[173]</sup>.

En 2005, l'équipe de H. Wennemers a identifié deux tripeptides H-L-Pro-D-Ala-D-Asp-NH<sub>2</sub> et H-L-Pro-L-Pro-L-Asp-NH<sub>2</sub> comme catalyseurs hautement efficace de la réaction d'aldolisation entre le nitro-benzaldéhyde et l'acétone.<sup>[171]</sup> Il est intéressant de noter que les deux peptides conduisent à des énantiomères différents.

Dans un article de 2008, cette même équipe suggère que la distance entre la fonction acide carboxylique portée par l'aspartate et l'amine secondaire de la proline N-terminale puisse jouer un rôle majeur dans le contrôle réactionnel.<sup>[173]</sup> Dans le cas de la proline, cette distance est de 3,7Å tandis que pour le tripeptide, cette même distance est de 7.3Å (Figure 117) permettant d'envisager, outre des réactions d'additions de type 1,2 (réaction d'aldolisations), des additions de type 1,4.

Il a ainsi été démontré que les tripeptides H-D-Pro-Pro-Glu-NH<sub>2</sub> ou H-D-Pro-Pro-Asp-NH<sub>2</sub> sont capables de catalyser, de manière asymétrique, des réactions entre des aldéhydes et le nitro-éthylène, utilisé comme accepteur de Michael<sup>[185]</sup>. Le meilleur résultat, avec une conversion  $\geq$

95% et un ee > 95%, a été obtenu après 15h à température ambiante avec 1mol% de H-D-Pro-Pro-Glu-NH<sub>2</sub>, 1mol% de NMM, 1,5 équiv d'aldéhyde et de nitro-éthylène à 0,1M dans le chloroforme. Les deux fonctions amine et acide sont nécessaires à l'activité : lorsque l'amine de la proline est acétylée ou méthylée, l'activité catalytique est perdue. De plus, la réactivité est diminuée lorsque l'acide carboxylique du tripeptide est substitué par un carboxamide.

Des études cinétiques ont permis de démontrer que la réaction de l'énamine (**A**, Schéma 2) avec l'électrophile est limitante et de suggérer un double rôle de la fonction acide. Premièrement, elle orienterait la réactivité et la stéréosélectivité et, deuxièmement, la réaction serait accélérée du fait de la protonation de l'iminium nitronate (**B**, Schéma 2).<sup>[175,176]</sup> Un mécanisme alternatif a été proposé lorsque le catalyseur ne possède pas d'hydrogène acide (chemin orange, Schéma 2).

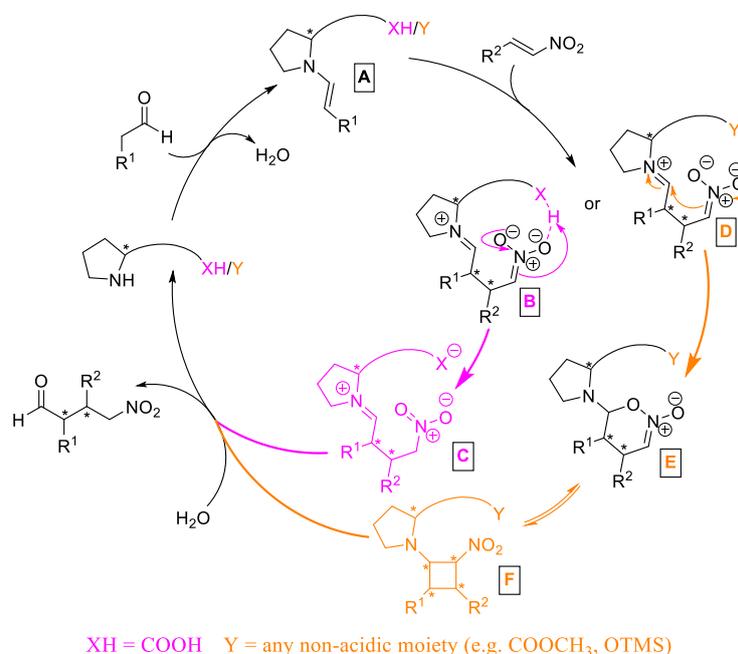


Schéma 2: Cycle catalytique proposé par l'équipe de Wennemers pour la réaction d'un aldéhyde sur une nitrooléfine<sup>[175]</sup>.

Comme nous l'avons vu dans cette succincte étude bibliographique, la proline mais aussi de petits peptides de séquence Pro-Pro-Glu ou Pro-Pro-Asp sont capables de catalyser des réactions de type énamine. Cette réactivité résulte à la fois de la présence d'une amine secondaire, sous forme de pyrrolidine et d'une fonction acide à proximité de l'amine. L'objectif de ma thèse étant d'introduire une activité catalytique sur un foldamère d'ATCs, c'est ainsi que nous avons cherché dans un premier temps à développer un monomère d'ATC présentant à la fois une pyrrolidine sur la position 2 du thiazole et une chaîne latérale en position  $\gamma$  comportant

une fonction acide carboxylique (Schéma 3). Nous avons ensuite vérifié ses capacités à catalyser une réaction modèle, la réaction d'addition de Michael de la cyclohexanone sur des nitrostyrènes. Cette réaction présente l'avantage d'avoir été largement étudiée en organocatalyse. Par ailleurs, les  $\gamma$ -nitroaldéhydes obtenus possèdent deux centres stéréogènes et peuvent servir à la construction de molécules plus complexes.<sup>[192,176]</sup>

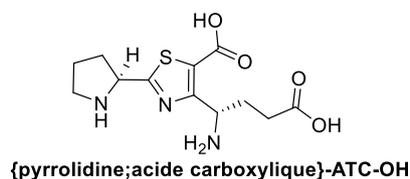


Schéma 3 : Structure générale du monomère catalytique

## 2. Synthèse des monomères d'ATC

La voie de synthèse très robuste développée par les Dr. L. Matthieu et C. Bonnel garantit l'accès à une large diversité d'ATC énantipures, permettant de facilement moduler la nature des chaînes latérales décorant les oligomères.<sup>[189]</sup>

La stratégie de synthèse repose sur l'utilisation d'un *N*-Fmoc-aminoacide comme précurseur d'un premier intermédiaire de type  $\beta$ -cétoester (**42**). La formation du  $\beta$ -cétoester (**42**) se fait selon un mécanisme de condensation croisée de Claisen entre le *N*-Fmoc-aminoacide et l'acétate de 1,1-diméthylallyle. L'utilisation d'un acétate encombré est nécessaire pour éviter des réactions secondaires d'auto-condensation de l'acétate. Le composé (**42**) subit ensuite une réaction de mono-bromation sur la position malonique selon les conditions de Yang et al. en présence de  $Mg(ClO_4)_2$ , en quantité catalytique agissant comme un acide de Lewis, et du *N*-bromosuccinimide, à  $-45^\circ C$ , pour donner le  $\alpha$ -bromo- $\beta$ -cétoester (**43**).<sup>[190]</sup> Une cyclisation de Hantzsch avec un thioamide permet la formation du thiazole **3** protégé de manière orthogonale sur les fonctions amine et acide ainsi que sur la chaîne latérale. Une déprotection chimiosélective de l'ester de 1,1-diméthylallyle en présence de phénylsilane conduit au monomère d'ATC (**44**) protégé par un groupement acido-labile sur les chaînes latérales. Selon cette voie de synthèse, plusieurs monomères ont été synthétisés avec des rendements variant de 26 à 65%. Ces monomères ont servi à la synthèse de foldamères qui seront présentés par la suite.

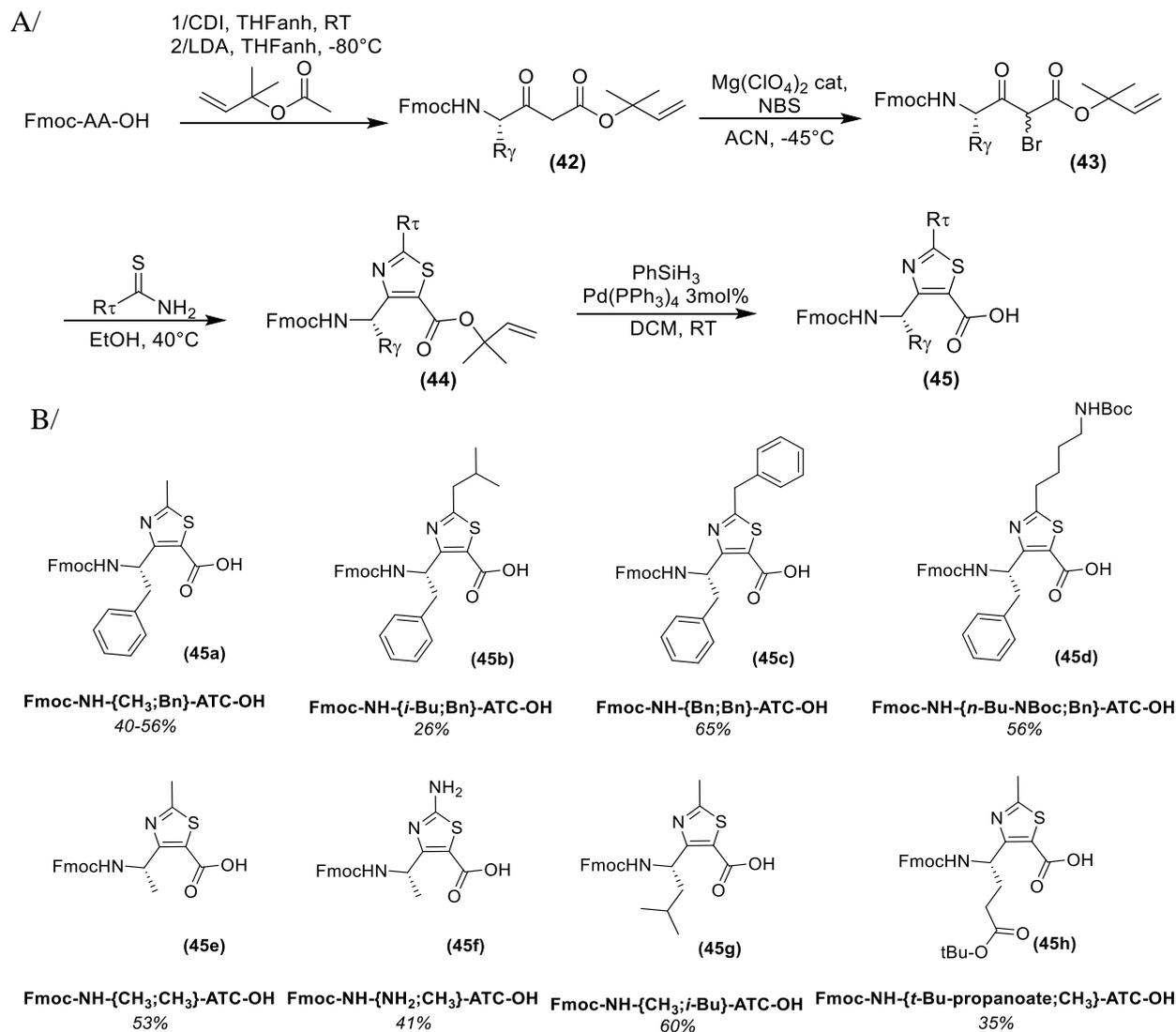


Schéma 4 : A/ Voie de synthèse des monomères d'ATCs. B/ Monomères d'ATCs synthétisés. Les rendements sont calculés sur les étapes de bromation et de cyclisation de Hantzsch.

Pour accéder aux ATCs comportant le motif pyrrolidine sur la position 2 du noyau thiazole, nous avons modifié le protocole (Schéma 5). Nous avons choisi de cycliser l' $\alpha$ -bromo- $\beta$ -cétostère afin de former les thiazolines en présence de Boc-Proline thioamide synthétisée au laboratoire (2 étapes, rdt 75%). Puis, une réaction d'aromatisation des thiazolines par l'anhydride trifluoroacétique en présence de DIEA à 0°C a été effectuée, conduisant au composé (44'). Une déprotection sélective selon les conditions de Tsuji-Trost a enfin été réalisée afin d'obtenir les monomères d'ATC (45i-I).

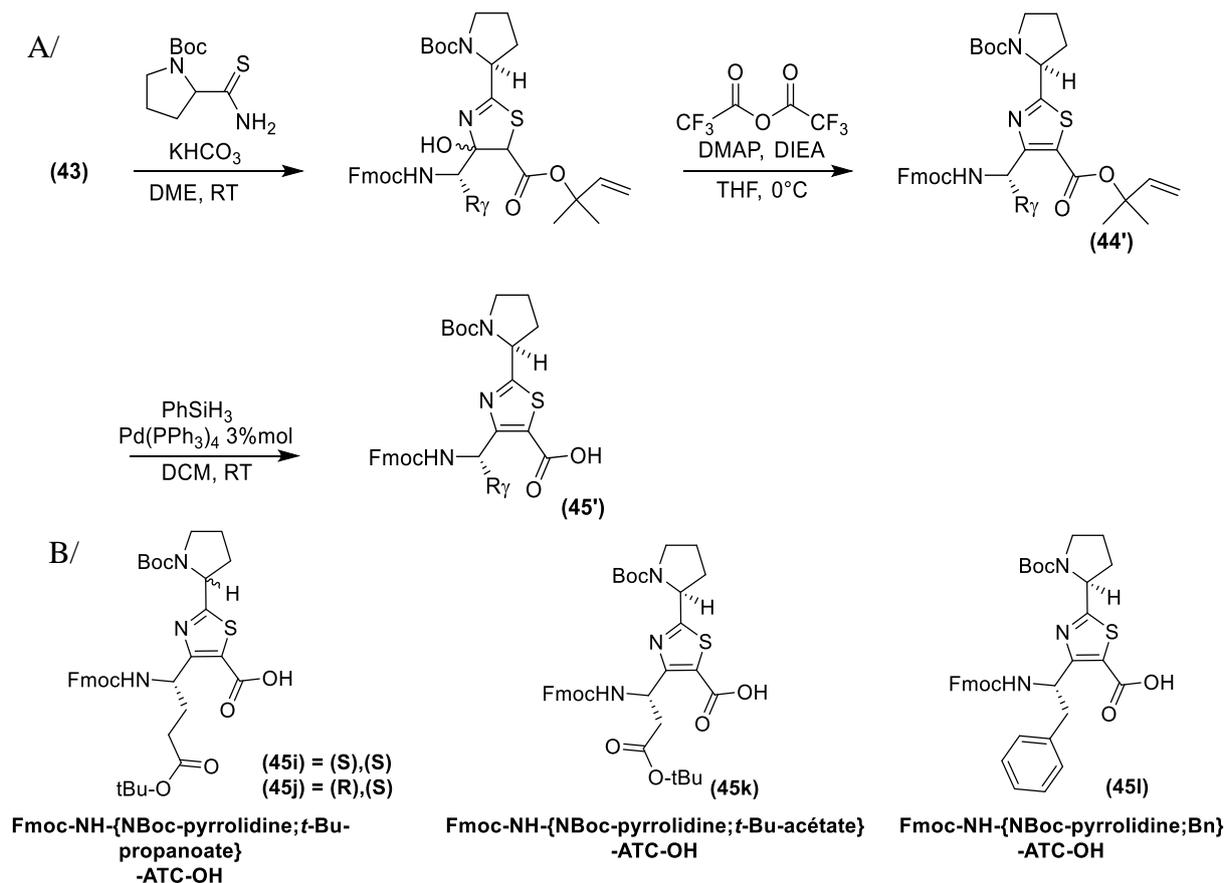


Schéma 5 : A/ Cyclisation de Hantzsch en 2 étapes. B/ Monomères d'ATCs synthétisés après cyclisation puis aromatisation. Les rendements sont calculés sur les étapes de bromation et de cyclisation de Hantzsch

### 3. Evaluation des propriétés de catalyse du monomère d'ATC (46)

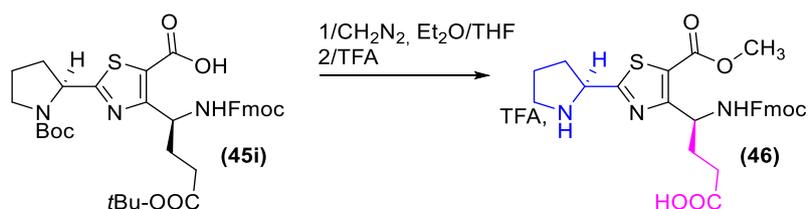


Schéma 6: NFmoc-{Pyrrolidine;acide propanoïque}-ATC-OCH<sub>3</sub>

Nous avons cherché dans un premier temps à explorer les propriétés de catalyse de l'ATC (46) dans la réaction de nitro-Michael. Celui-ci a été obtenu à partir du *N*-Fmoc-ATC-OH (45i) par traitement au diazométhane et déprotection au TFA des chaînes latérales (Schéma 6).

Premièrement, nous avons évalué l'influence de la température et du solvant sur la conversion, la diastéréosélectivité *syn/anti* et sur l'énantiosélectivité pour la réaction d'addition de la cyclohexanone sur le  $\beta$ -*trans*-nitrostyrène (Schéma 7). La réaction a été effectuée avec 10

équiv. de cyclohexanone, 1.0 équiv. de  $\beta$ -*trans*-nitrostyrène et 10 mol% de catalyseur. Les résultats obtenus après 16 h de réaction sont rapportés dans le Tableau 1.

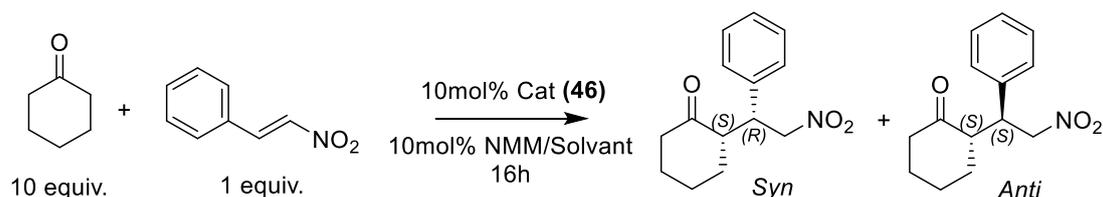


Tableau 1 : Optimisation de la réaction de nitro Michael avec le catalyseur (46) ;

Entrée <sup>a</sup>	Solvant	T (°C)	Conversion <sup>b</sup> (%)	Syn/Anti <sup>b</sup>	Excès énantiomérique <sup>c</sup> (%)
1	Toluène	30	54	95/5	27
2	DMSO	30	82	87/13	29
3	CHCl <sub>3</sub>	30	81	93/7	27
4	THF	30	80	94/6	28
5	CH <sub>3</sub> CN	30	80	92/8	22
6	Hexane	30	88	94/6	31
7	MeOH	30	18	95/5	nd
8	EtOH	30	98	92/8	34
9	<i>i</i> PrOH	30	92	96/4	41
10	<i>i</i> PrOH	4	98	96/4	45

<sup>a</sup>Réalisé dans 400  $\mu$ l de solvant ; <sup>b</sup>Déterminée par RP-HPLC à 214 nm ; <sup>c</sup>Déterminé par HPLC chirale (Colonne IC; Hex/*i*PrOH 8:2 v/v,  $d=1$  ml/min,  $\lambda=214$  nm). Sélectivité vers (2*S*)-2-[(1*R*)-2-Nitro-1-phényléthyl] cyclohexanone déterminée par rotation optique and comparaison avec les données reportées dans la littérature<sup>[195]</sup>

Dans le toluène à 30°C (entrée 1), la réaction a conduit à une conversion de 54%, une bonne diastéréosélectivité *syn/anti* de 95/5 et un excès énantiomérique, modeste mais significatif, de 27% en faveur du produit de configuration ((*S*),(*R*)). Afin d'améliorer la conversion ainsi que l'induction asymétrique, plusieurs solvants ont été testés. Le diméthylsulfoxyde, le chloroforme, le tétrahydrofurane, l'acétonitrile ou encore l'hexane (entrées 2 à 6) ont permis d'accroître le taux de conversion jusqu'à 88%. Cependant, l'induction asymétrique n'a pas été significativement augmentée, plafonnant à 31%. Le méthanol (entrée 7) est le solvant pour lequel le taux de conversion obtenu a été le plus faible. Ce résultat est plutôt inattendu, d'autant plus que l'éthanol a conduit à une conversion de 98% et un excès énantiomérique de 34%. L'étude des alcools comme solvant a donc été poussée et le solvant donnant les meilleurs résultats en termes d'induction asymétrique a été l'*isopropanol* (entrée 9) pour lequel, à 30°C,

l'excès énantiomérique obtenu est alors de 41%. A plus basse température (4°C), le rendement a été légèrement amélioré (98%) du fait de moins de produits de dégradation. Par ailleurs, l'induction asymétrique a été augmentée à 45%.

### III. INCORPORATION DES MONOMERES DANS DES OLIGOMERES

Le deuxième point abordé dans ce manuscrit a été l'étude de l'impact du contrôle conformationnel sur la chimio et l'énantiosélectivité.

Nous avons tout d'abord cherché à rigidifier le catalyseur en introduisant un groupement isopropylamide du côté C-terminal et un acétyle sur l'extrémité N-terminale (Figure 4). Dans une étude antérieure du laboratoire,<sup>[104]</sup> ces deux groupements s'étaient avérés compatibles avec l'établissement d'une liaison hydrogène intramoléculaire en C<sub>9</sub>. Alors que le composé **(47)** semble moins encombré que le catalyseur **(46)** protégé par un Fmoc, l'induction asymétrique a été légèrement améliorée (ee = 52%). Néanmoins, alors que la conversion était excellente > 99%, la sélectivité syn/anti s'est avérée légèrement moins bonne (93/7).

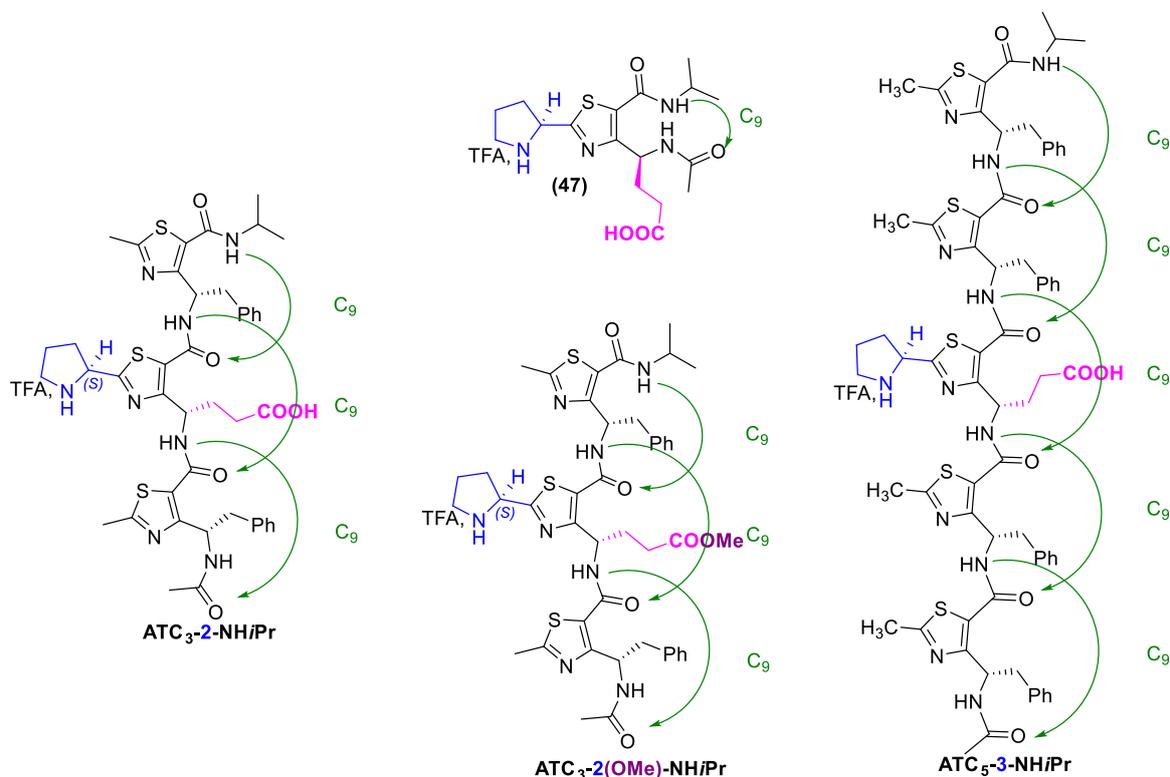


Figure 4 : ATC-monomer structuré **(47)**, trimères ATC<sub>3</sub>-2-NHiPr et ATC<sub>3</sub>-2(OMe)-NHiPr, et pentamère ATC<sub>5</sub>-3-NHiPr

Nous avons ensuite inclus le motif catalytique au sein d'un trimère **ATC<sub>3</sub>-2-NH<sub>i</sub>Pr** puis d'un pentamère **ATC<sub>5</sub>-3-NH<sub>i</sub>Pr**. Ces oligomères ont été synthétisés selon une stratégie pas-à-pas en solution en utilisant un mélange d'EDCI et d'HOBt comme agents de couplage. Au cours de la synthèse les groupements Fmoc ont été déprotégés par une solution de diéthylamine dans le DMF. Le trimère **ATC<sub>3</sub>-2-NH<sub>i</sub>Pr** a été obtenu avec un rendement global de 34% tandis que le pentamère **ATC<sub>5</sub>-3-NH<sub>i</sub>Pr** a été isolé avec un rendement de 22%.

Testés sur la réaction d'addition de nitro-Michael entre la cyclohexanone et le  $\beta$ -*trans*-nitrostyrène, ces deux oligomères ont conservés une bonne activité catalytique (conversions : 98 et 93 respectivement) associée à une excellente sélectivité *syn/anti* (99/1). L'induction asymétrique pour sa part a été améliorée de manière significative : 74% pour le trimère **ATC<sub>3</sub>-2-NH<sub>i</sub>Pr** et 77% pour le pentamère **ATC<sub>5</sub>-3-NH<sub>i</sub>Pr**, suggérant ainsi un impact de la structure secondaire sur les propriétés de catalyse. Finalement, pour vérifier l'importance de la chaîne propanoïque dans la catalyse, le trimère **ATC<sub>3</sub>-2-NH<sub>i</sub>Pr** a été estérifié sous forme d'ester méthylique. Dans ce cas, nous avons observé une perte de réactivité.

#### 4. Validation de la structure de l'oligomère

Après la synthèse du premier oligomère, il a été nécessaire de valider sa structure en solution. En effet, un foldamère est un oligomère artificiel qui doit présenter une structure définie et conservée dans différentes gammes de températures et de solvants. Ainsi, afin de confirmer notre hypothèse, le pentamère **ATC<sub>5</sub>-3-NH<sub>i</sub>Pr** a été soumis à des analyses RMN dans le méthanol d<sub>3</sub>.

Grâce à l'étude du spectre RMN <sup>1</sup>H de **ATC<sub>5</sub>-3-NH<sub>i</sub>Pr**, nous avons observé que les signaux des NH et H $\gamma$  étaient très dispersés et très déblindés pour des protons amides, autour de 10ppm. D'après des travaux antérieurs réalisés au laboratoire, cela signifie que ces protons sont engagés dans des liaisons hydrogènes intramoléculaires à 9 chainons.<sup>[104,107]</sup> Par ailleurs, nous avons pu remarquer que le proton NH du résidu 1 est moins déblindé (à 9ppm), signe que ce proton n'est pas engagé dans une liaison intramoléculaire en adéquation avec le modèle d'hélice C<sub>9</sub>. Afin de confirmer cette structure, nous avons procédé à l'étude des spectres ROESY. Nous avons ainsi pu détecter des corrélations canoniques entre les H $\gamma$  des résidus *i* et les NH des résidus *i*+1. Des corrélations secondaires entre les H $\gamma$  des résidus *i* et les protons des chaînes latérales portés par les thiazoles des résidus *i*+1, caractéristiques de l'hélice 9 ont également été mises en évidence (Figure 118).

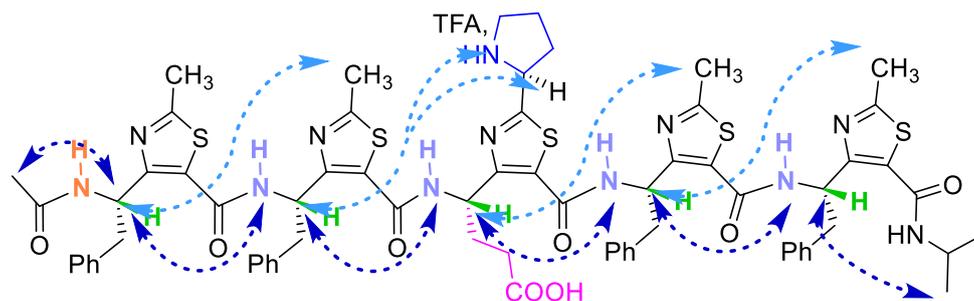


Figure 118 : Représentation des corrélations NOE observables.

Le chapitre 3 de cette thèse présente également l'analyse de cet oligomère par dichroïsme circulaire et infrarouge. Cela a permis d'établir l'existence d'une signature spécifique aux oligomères d'ATC. De plus, l'étude des fréquences Amides I ( $1600-1800\text{ cm}^{-1}$ ) en infrarouge a permis de déterminer le nombre de CO liés et de CO libres dans la séquence de l'oligomère.

## IV. MODULATION DE LA POSITION DU CENTRE CATALYTIQUE

### 1. Synthèse des oligomères

J'ai ensuite entrepris de modifier pas à pas la séquence du  $\gamma$ -peptide **ATC<sub>5</sub>-3-NH<sub>i</sub>Pr** afin d'en étudier l'impact sur la chimio et l'énantiosélectivité. Pour la première génération, il a été décidé de synthétiser des foldamères portant 4 monomères d'ATC identiques, **{CH<sub>3</sub>;Bn}-ATC-OH**, et un centre catalytique, **{Pyrrolidine;acide propanoïque}-ATC-OH**. Nous avons choisi de faire varier la position du centre catalytique le long de la chaîne pentamérique, comme présenté sur la figure suivante.

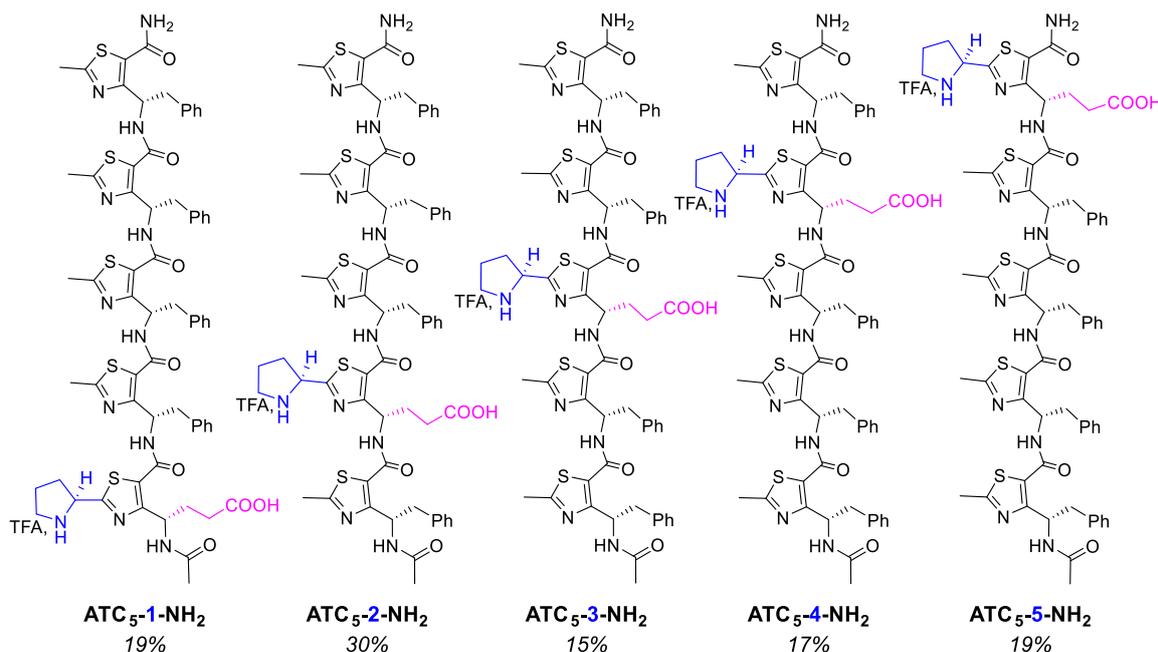


Figure 119: Foldamères  $ATC_5-P-NH_2$  avec  $P = 1$  to 5, les rendements sont en italique

Il a été choisi d'assembler les monomères d'ATC un à un, de manière linéaire, sur une résine ChemMatrix de type Rink amide selon une stratégie SPPS Fmoc/*t*Bu (Schéma 8). Par rapport au pentamère  $ATC_5-3-NH_iPr$ , nous avons choisi pour des raisons de facilité de synthèse, de ne pas introduire de groupement isopropyle sur l'extrémité C-terminale. Les couplages s'effectuent sur une nuit, dans la NMP à 20 mL/g de résine, en présence de 2 équivalents du monomère d'ATC choisi, de 2 équivalents de DIC et 2 équivalents d'Oxyma Pure. Chaque couplage a été contrôlé par un test de Kaiser et suivi d'une acétylation afin d'éviter la présence de peptides de délétion. Les déprotections des groupements Fmoc ont été faites en trois fois dix minutes dans une solution de pipéridine à 20% dans le DMF. Le clivage a été réalisé, après acétylation du groupement N-terminal par l'anhydride acétique, dans une solution de TFA/TIS/H<sub>2</sub>O 95/2,5/2,5 vvv.

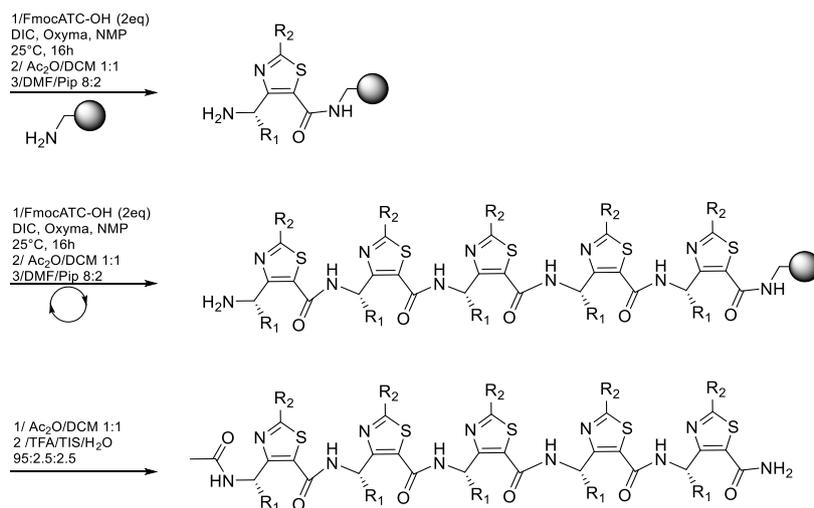


Schéma 8 : Synthèse sur phase solide

## V. MODULATION DU MICROENVIRONNEMENT AUTOUR DU CENTRE CATALYTIQUE

Nous avons également choisi de moduler le microenvironnement autour du centre catalytique lorsque celui-ci est au centre du foldamère. Le premier axe a consisté à étudier l'impact de la position et de la nature de la chaîne portant l'acide carboxylique sur les propriétés de catalyse (Figure 120). La modification de la longueur de la chaîne carbonée portant le groupement COOH est symbolisée par le foldamère **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>**. L'éloignement de la pyrrolidine de l'acide carboxylique est représenté par les foldamères **M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>** et **M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>**.

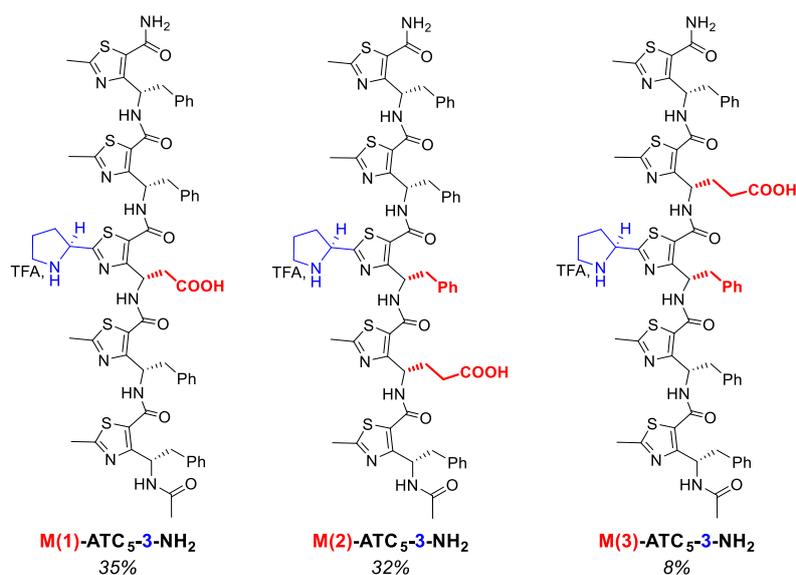


Figure 120: Modulation de la position et de la nature de la chaîne portant l'acide carboxylique, les rendements sont en italiques

Dans une seconde approche, nous avons envisagé de modifier la nature des chaînes latérales et de moduler les contraintes stériques autour du centre catalytique. Les monomères utilisés pour la synthèse de ces foldamères sont ceux qui ont été présentés en début de ce résumé. Tous ces oligomères ont été synthétisés comme précédemment selon une approche SPPS.

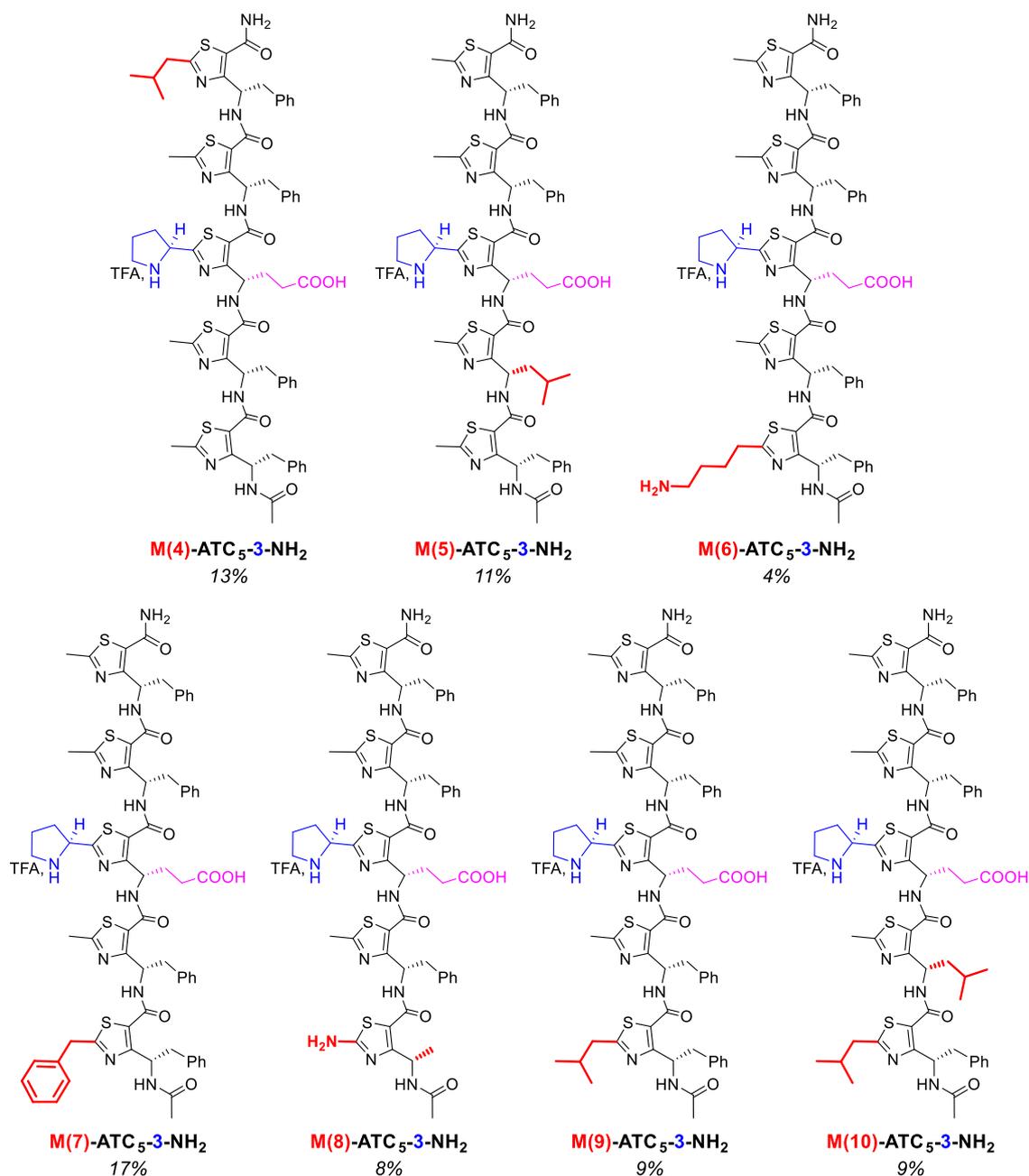


Figure 121 : Modulation du microenvironnement autour du centre catalytique sur une série de foldamères, les rendements sont en italiques

Dans le chapitre 4 de cette thèse, d'autres séquences pourront également être trouvées. Elles ont été synthétisées afin de lever toute gêne stérique entre les chaînes latérales ou pour explorer l'éventualité d'effets coopératifs.

Enfin, toutes ces séquences ont été testées sur la réaction de nitro-Michael. Celle-ci a été réalisée pour trois substrat différents. La première réaction étant l'addition de la cyclohexanone sur le  $\beta$ -*trans*-nitrostyrène, pour la deuxième réaction, ce dernier est remplacé par le *p*-méthyl-nitrostyrène. Finalement, le troisième substrat testé est le *p*-méthoxy-nitrostyrène. D'après l'optimisation des paramètres réalisée au préalable, la réaction est effectuée à 4°C dans 150 $\mu$ L d'*isopropanol* contenant 10mol% de NMM.

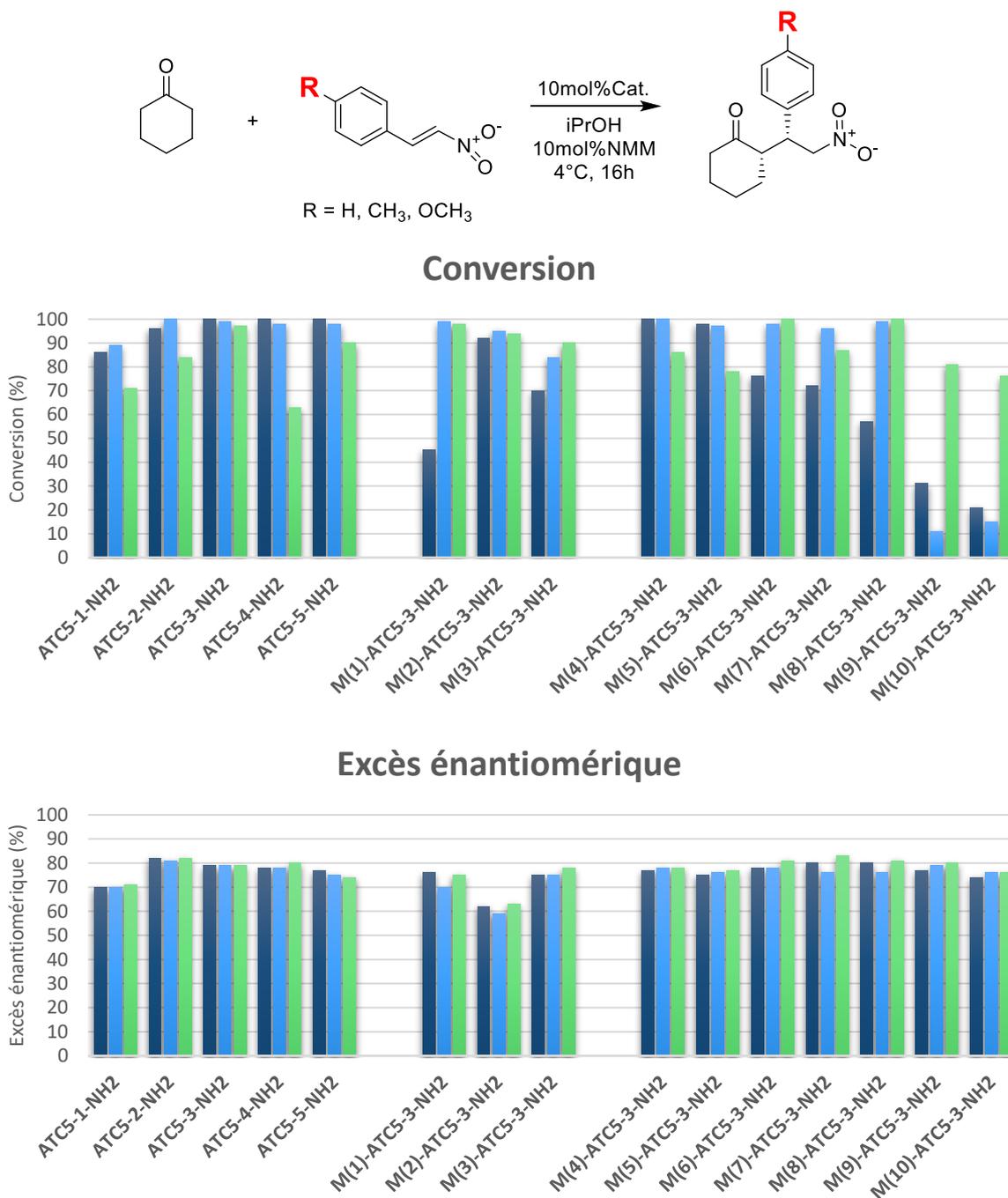


Figure 122 : Conversion et excès énantiomérique pour les réactions d'addition de la cyclohexanone sur le  $\beta$ -*trans*-nitrostyrène (bleu foncé), le *p*-méthyl-nitrostyrène (bleu) et le *p*-méthoxy-nitrostyrène (vert) pour les foldamères d'ATC

D'après les résultats obtenus lors des tests de catalyse (Figure 122), nous avons pu remarquer que la position du site catalytique le long de la chaîne (foldamères **ATC<sub>5</sub>-P-NH<sub>2</sub>**) n'avait que peu d'influence, la conversion étant aux alentours de 95-100% et l'induction asymétrique vers 75-80%. Nous avons tout de même noté que lorsque le centre catalytique était positionné du côté N-Terminal (**ATC<sub>5</sub>-1-NH<sub>2</sub>**), la conversion et l'induction asymétrique avaient tendance à être plus faibles, quel que soit le substrat engagé.

Lorsque la chaîne carbonée de l'acide carboxylique est raccourcie (**M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>**), nous avons observé un effet négatif sur la conversion du  $\beta$ -*trans*-nitrostyrène (chute à 45%), tandis que l'induction asymétrique restait autour de 79%. Cet effet n'a pas été retrouvé pour les deux autres substrats pour lesquels des conversions > 98% ont été obtenues.

La séparation de la chaîne acide de la pyrrolidine a également entraîné deux types de profils différents. Pour les trois substrats, quand la chaîne carboxylée est plus proche du côté N-Ter (**M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>**), la conversion variait entre 92 et 95% et l'excès énantiomérique entre 59 et 63%. Lorsque la chaîne carboxylée est placée côté C-Ter (**M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>**), le profil observé était inverse au précédent. La conversion était plus faible, entre 70 et 90%, mais l'induction asymétrique oscillait entre 75 et 78%.

Finalement, la 3<sup>ème</sup> catégorie de foldamères (**M(4)- à M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>**), a conduit à des résultats étonnants. En effet, nous avons remarqué que lorsque le thiazole du résidu 1 est décoré par un groupement benzyle (**M(7)-**), les excès énantiomériques obtenus pour ce foldamères étaient les plus élevés, jusqu'à 83%, notamment en comparaison avec la référence **ATC<sub>5</sub>-3-NH<sub>2</sub>** qui plafonnait à 79%. Par ailleurs, pour les foldamères **M(9)-** et **M(10)-** possédants, tous deux, un groupement isobutyle sur le thiazole du premier résidu du côté N-Ter., nous avons observé une sélectivité de substrat vis-à-vis du méthoxy-nitrostyrène (conversion > 76%). Pour le  $\beta$ -*trans*-nitrostyrène et le *p*-méthyl-nitrostyrène, dans des conditions expérimentales identiques les taux de conversion sont restés très limités (11 à 31% après 16h à 4°C). De tels résultats, fréquemment rencontrés dans des procédés de catalyse enzymatique, étaient inattendus avec des substrats de structure aussi similaire.

## VI. DETERMINATION STRUCTURALE DES OLIGOMERES D'ATC AU NIVEAU ATOMIQUE

Enfin, la dernière partie de ce manuscrit est dédiée à la détermination structurale des oligomères d'ATC synthétisés. Pour cela, trois méthodes de modélisation sous contraintes RMN ont été explorées et un nouveau protocole de définition des charges atomiques a été créé (Figure 123).

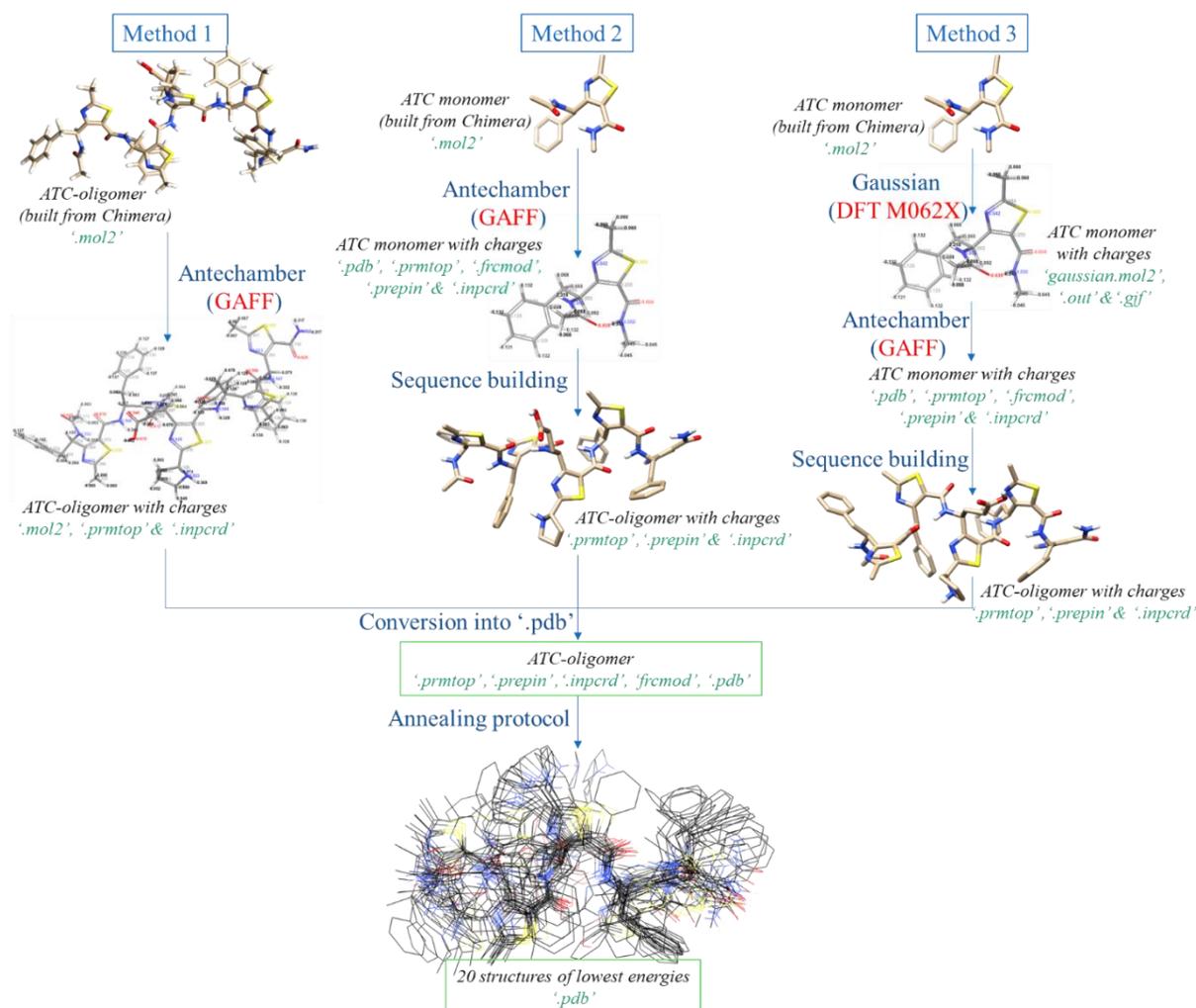


Figure 123 : Différentes méthodes explorées pour la modélisation d'oligomères d'ATC sous contraintes RMN

## VII. CONCLUSION

Il a été établi que des monomères d'ATC portant une pyrrolidine sur le thiazole et un acide carboxylique sur la chaîne  $\gamma$  étaient capables de catalyser des réactions de type énamine. La réaction de Nitro-Michael faisant partie de cette catégorie, l'influence d'une structuration sur l'énantiosélectivité a été une question majeure pour démarrer le projet. Ainsi, il a été prouvé qu'avec une augmentation de la structuration sur le monomère d'ATC puis en synthétisant un

trimère et enfin un pentamère d'ATC, l'induction asymétrique passait de 41% pour un monomère non structuré à 74% pour un pentamère.

Ainsi, nous avons été capable de synthétiser de nombreux monomères d'ATC comportant des modifications sur le thiazole et/ou la chaîne  $\gamma$ . Il a ensuite été possible d'assembler ces monomères par synthèse sur phase solide, ce qui a permis la synthèse de 26 foldamères différents. Ces derniers ont été testés sur la réaction de Nitro-Michael, par addition de la cyclohexanone sur trois substrats différents, le  $\beta$ -*trans*-nitrostyrène, le *p*-méthyl-nitrostyrène et le *p*-méthoxy-nitrostyrène. Ces foldamères ont montré un potentiel certain en tant que catalyseurs avec des taux de conversion atteignant pour certains les 99% en 16h, une sélectivité vers le composé *syn* et une induction asymétrique aux alentours de 80%. La modification du microenvironnement autour du centre catalytique a permis de montrer une chimiosélectivité que nous n'attendions pas forcément au début de ce projet. Ainsi, nous avons synthétisé des catalyseurs biomimétiques pouvant mimer certaines propriétés des enzymes en apportant une sélectivité de substrat.

Enfin, le développement d'un protocole original de modélisation sous contraintes RMN a permis de démontrer l'importance des charges partielles atomiques ainsi que du champ de force appliqué aux ATC. Une majorité des foldamères synthétisés a pu être modélisée par cette méthode et la structure en hélice  $C_9$  des oligomères d'ATC a été observée pour la plupart.



# SUPPORTING INFORMATION



## TABLE OF CONTENTS

I.	GENERAL EXPERIMENTAL CONDITIONS.....	S4
II.	SYNTHESIS OF ATC-MONOMERS.....	S5
III.	SYNTHESIS OF CATALYSTS <b>(46)</b> , <b>(47)</b> , <b>ATC<sub>3</sub>-2-NH/PR</b> , <b>ATC<sub>3</sub>-2*-NH/PR</b> AND <b>ATC<sub>5</sub>-3-NH/PR</b> .....	S18
IV.	EXPLORED MODEL REACTIONS USING <b>(46)</b> AS CATALYST.....	S41
V.	HPLC PROFILES OF THE NITRO-MICHAEL ADDITION REACTIONS CATALYSED BY <b>ATC<sub>5</sub>-3-NH<sub>2</sub></b> .....	S45
VI.	NMR AND MOLECULAR MODELLING STUDIES.....	S49
VII.	CIRCULAR DICHROISM.....	S160
VIII.	FTIR ANALYSIS.....	S161
IX.	XRD STRUCTURES.....	S175
X.	SPECTRA AND HPLC ANALYSIS.....	S179

## I. GENERAL EXPERIMENTAL CONDITIONS

Commercially available reagents and solvents were used without any further purification. Reactions were monitored by HPLC using an analytical Chromolith® Speed Rod RP-C18 185 Pm column (50 x 4.6 mm, 5 µm) using a flow rate of 5.0 ml/min, and gradients from 100/0 to 0/100 eluents A/B over 3 min, in which eluents A = H<sub>2</sub>O / TFA 0.1% and B = CH<sub>3</sub>CN / TFA 0.1%. Detection was done at 214 nm and 254 nm using a Photodiode Array Detector. Retention times were reported as follows: LC: rt = (min). LC-MS spectra (ESI) were recorded on an HPLC using an analytical Chromolith Speed Rod RP-C18 185 Pm column (50 x 4.6 mm, 5 µm) using a flow rate of 3.0 mL/min, and gradients of 100/0 to 0/100 eluents A/B over 2.5 min, in which eluent solvent A = H<sub>2</sub>O / HCOOH 0.1% and solvent B = CH<sub>3</sub>CN/HCOOH 0.1%. High-Resolution Mass Spectrometric analyses were performed with a time-of-flight (TOF) mass spectrometer fitted with an Electrospray Ionization Source (ESI). All measurements were performed in the positive ion mode. Melting points (mp) are uncorrected and were recorded on a capillary melting point apparatus. Enantiomeric excesses were determined by Chiral HPLC analysis using a Chiralpak® IC column (250 x 4.6 mm) with *n*-hexane/isopropanol (80/20) as eluent or Chiralpak® IA column (250 x 4.6 mm) with *n*-hexane/isopropanol (90/10) as eluent and a flow rate of 1 ml/min.

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded at room temperature in deuterated solvents. Chemical shifts (δ) are given, in this document, in parts per million, relative to TMS or relative to the solvent [<sup>1</sup>H: δ (CDCl<sub>3</sub>) = 7.24 ppm; <sup>13</sup>C: δ (CDCl<sub>3</sub>) = 77.2 ppm]. The following abbreviations are used to designate the signal multiplicities: s (singlet), d (doublet), dd (doublet doublet), t (triplet), q (quartet), m (multiplet), br (broad).

Circular dichroism (CD) experiments were performed using a Jasco J815 spectropolarimeter. The spectra were recorded at 50 µM compounds dissolved in *i*PrOH using a 1 mm path length CD cuvette at 4°C, over a wavelength range of 190-300 nm. Continuous scanning mode was used, with a response of 1.0 s with 0.5 nm steps and a bandwidth of 1 nm. The signal to noise ratio was improved by acquiring each spectrum over an average of three scans. Baseline was corrected by subtracting the background from the sample spectrum.

Middle-infrared experiments (400-5000 cm<sup>-1</sup>) were recorded in the transmission mode. The measurements were conducted on a Bruker Tensor 27 spectrometer equipped with a deuterated (*L*)-alanine doped triglycene sulphate (DLATGS) pyroelectric detector, a Globar source, and potassium bromide (KBr) beam splitter. The spectral resolution was 4 cm<sup>-1</sup>, and 128 scans were

co-added for each spectrum. The compounds were dissolved at 2 mM concentration in *i*PrOH transferred in a liquid cell equipped with CaF<sub>2</sub> windows separated by a teflon spacer (thickness: 50 μM). A spacer thickness of 50 μM provide exploitable signal in term of signal-to-noise ratio S/N (S/N (Peak to peak =  $4 \times 10^{-5}$ ), S/N (quadratic mean (RMS) =  $1 \times 10^{-5}$  between 2000 and 2100 cm<sup>-1</sup>). Bellow this value, significant compensation problems between solvent spectral bands used as reference and sample spectrum appear. The FT-IR spectra were not smoothed. Linear baseline subtraction and deconvolution were performed using IGOR Pro 7.0 software (WaveMetrics).

## II. SYNTHESIS OF ATC-MONOMERS

Catalysts (46), (47), ATC<sub>3</sub>-2-NHiPr, ATC<sub>3</sub>-2\*-NHiPr, ATC<sub>5</sub>-3-NHiPr and **M(n)**-ATC<sub>L</sub>-P-NH<sub>2</sub> were synthesized starting from the monomer units (45a-l)

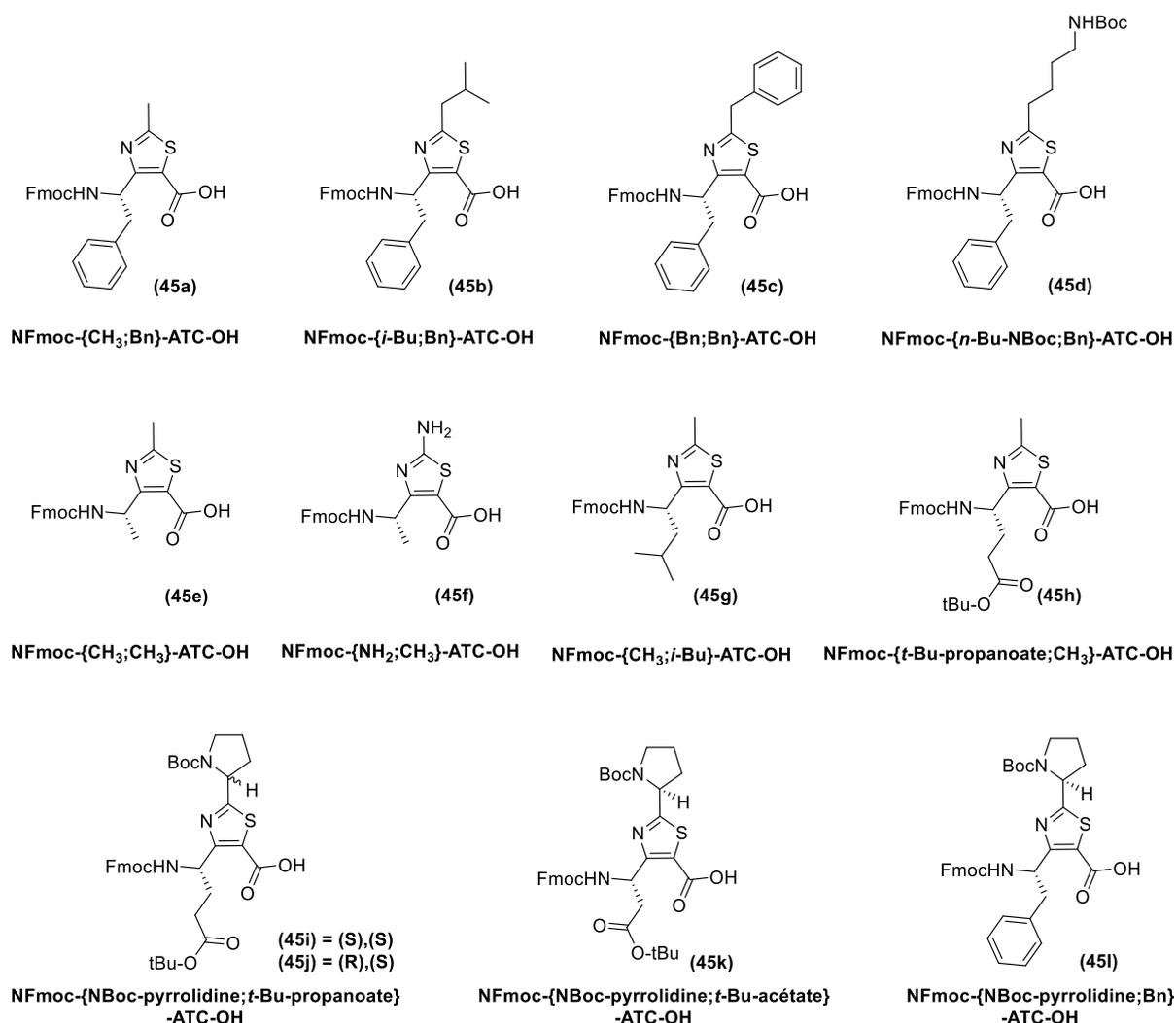


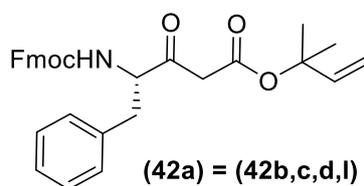
Figure S1: Fmoc protected ATC-monomers (45a-k)

### 1. General procedure for the synthesis of $\beta$ -keto-ester (42)

**Imidazolide formation:** In a 250 mL two-neck flask under a nitrogen atmosphere was dissolved the Fmoc-AA-OH (1.0 equiv., 27.14mmol) in dry THF. Then CDI (1.1 equiv., 29.85mmol, 4.84g) was added in three turns. A catalytic amount of DMAP (3mol%, 0.8142mmol, 99mg) was added 2-3 min after and the solution was stirred for 1h at room temperature.

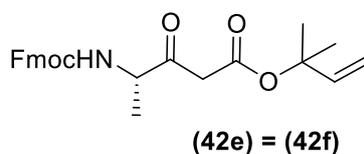
**Enolate formation:** Under a nitrogen atmosphere, a 500 mL three-neck flask was loaded with LiHMDS (22equiv, 597.6mmol, 100mL, 1M solution in THF) followed by dry THF (100mL). After cooling at  $-78^{\circ}\text{C}$  2-methylbut-3-en-2-yl (4.0 equiv, 108.6mmol, 13.91g) was added dropwise over 10 min. The solution was stirred at  $-78^{\circ}\text{C}$  for 10 min, then at r.t. for 10 min and finally at  $-78^{\circ}\text{C}$  for 20 min.

**Condensation:** Imidazolide solution was added dropwise to the enolate solution at  $-78^{\circ}\text{C}$  over 5-10 min. After stirring for 15 minutes (HPLC monitoring) the mixture was removed from the cold bath and poured onto a solution of 10% aqueous citric acid (200mL, pH 7). The crude was extracted with EtOAc (2 $\times$ 100mL). The combined organic layers were washed with saturated  $\text{NaHCO}_3$  solution (1 $\times$ 100mL) and brine (3 $\times$ 150mL), dried, filtered and evaporated under reduced pressure to yield crude product. Purification by flash column chromatography gave the pure  $\beta$ -ketoester.

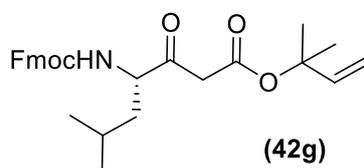


**(42a):** Pale yellow solid. Yield: 50-70 %. Mp: 101-102 $^{\circ}\text{C}$ .  $[\alpha]_{\text{D}}^{20^{\circ}\text{C}} = +0.5^{\circ}$  (c 1.00,  $\text{CHCl}_3$ ).  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz):  $\delta = 1.52$  (s, 6H); 3.01 (ABX,  $J = 14.0, 7.0$  Hz, 1H); 3.18 (ABX,  $J = 14.0, 7.0$  Hz, 1H); 3.40 (d,  $J = 5.3$  Hz, 2H); 4.17 (t,  $J = 6.9$  Hz, 1H); 4.40 (m, 2H); 4.67 (m, 1H); 5.10 (d,  $J = 10.8$  Hz, 1H); 5.19 (d,  $J = 17.5$  Hz, 1H); 5.34 (d,  $J = 7.6$  Hz, 1H); 6.07 (dd,  $J = 17.5, 10.8$  Hz, 1H); 7.14 (d,  $J = 7.0$  Hz, 1H); 7.23-7.33 (m, 6H); 7.40 (t,  $J = 7.4$  Hz, 2H); 7.53 (t,  $J = 7.4$  Hz, 2H); 7.77 (d,  $J = 7.4$  Hz, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 100 MHz):  $\delta = 26.4$  (2C); 37.1; 47.3; 48.0; 60.9; 67.1; 82.8; 113.5; 120.1 (2C); 125.2 (2C); 127.2 (2C), 127.3; 127.9 (2C); 128.9 (2C); 129.4 (2C); 135.9; 141.4 (2C); 141.8; 143.8 (2C); 155.8; 165.6; 201.7. LC rt = 2.04 min (major) et 2.19 min (minor) (keto-enol equilibrium). LC-MS: (ESI $^+$ ): m/z 498.2 ([M+H] $^+$ ) 100%, 520.2 ([M+Na] $^+$ ) 10%. HRMS (ESI) calcd. for  $\text{C}_{31}\text{H}_{31}\text{NO}_5\text{Na}^+$ : 520.2100, found

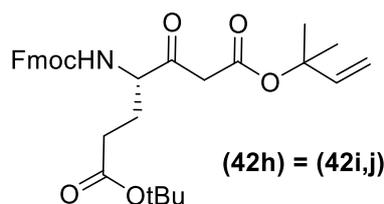
520.2098. FT-IR:  $\nu_{\max}$  (cm<sup>-1</sup>) 3313; 2984; 1738; 1713; 1692; 1534; 1323; 1254; 1124; 1031; 929; 836; 755; 741; 703.



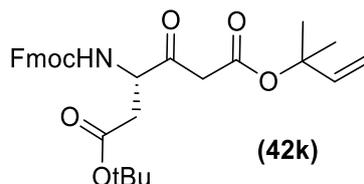
**(42e)**: Pale orange solid. Yield: 76-86%. Mp: 68-70°C.  $[\alpha]_{\text{D}}^{20^\circ\text{C}} = +6.6^\circ$  (c 2.00, CHCl<sub>3</sub>). <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz):  $\delta = 1.39$  (d,  $J = 7.1$  Hz, 3H); 1.54 (s, 6H); 3.45 (AB,  $J = 16.0$  Hz, 1H); 3.50 (AB,  $J = 16.0$  Hz, 1H); 4.22 (t,  $J = 6.9$  Hz, 1H); 4.38-4.50 (m, 3H); 5.10 (d,  $J = 11.0$  Hz, 1H); 5.20 (d,  $J = 17.5$  Hz, 1H); 5.47 (d,  $J = 7.2$  Hz, 1H); 6.08 (dd,  $J = 11.0, 17.5$  Hz, 1H); 7.32 (td,  $J = 0.7, 7.5$  Hz, 2H); 7.40 (t,  $J = 7.5$  Hz, 2H); 7.59 (d,  $J = 7.4$  Hz, 2H); 7.75 (d,  $J = 7.5$  Hz, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz):  $\delta = 17.4$ ; 26.4 (2C); 47.0; 47.3; 56.0; 67.1; 82.8; 113.6; 120.1 (2C); 125.1 (2C); 127.2 (2C); 127.9 (2C); 141.5; 141.8 (2C); 143.9 (2C); 155.8; 165.6; 202.2. LC rt = 2.25 min (major) et 2.48 min (minor) (keto-enol equilibrium). LC-MS (ESI<sup>+</sup>): m/z 422.3 ([M+H]<sup>+</sup>) 100 %, 444.2 ([M+Na]<sup>+</sup>) 70 %. HRMS (ESI) calcd. for C<sub>25</sub>H<sub>28</sub>NO<sub>5</sub><sup>+</sup>: 422.1972, found 422.1967. FT-IR:  $\nu_{\max}$  (cm<sup>-1</sup>) 3309; 2983; 2939; 1717; 1685; 1533; 1450; 1316; 1254; 1123; 1048; 923; 833; 757; 737.



**(42g)**: White solid. Yield: 73 %. Mp: 117-118°C.  $[\alpha]_{\text{D}}^{20^\circ\text{C}} = -10.2^\circ$  (c 1.00, CHCl<sub>3</sub>). <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz):  $\delta = 0.95$  (t,  $J = 6.2$  Hz, 6H); 1.41 (m, 2H); 1.53 (s, 6H); 1.68 (m, 1H); 3.41 (AB,  $J = 15.6$  Hz, 1H); 3.49 (AB,  $J = 15.6$  Hz, 1H); 4.22 (t,  $J = 6.6$  Hz, 1H); 4.42 (s, 1H); 4.45 (d,  $J = 6.6$  Hz, 2H); 5.09 (d,  $J = 10.8$  Hz, 1H); 5.19 (d,  $J = 17.4$  Hz, 1H); 5.19 (s, 1H); 6.08 (dd,  $J = 17.4, 10.9$  Hz, 1H); 7.32 (t,  $J = 7.4$  Hz, 2H); 7.40 (t,  $J = 7.4$  Hz, 2H); 7.59 (m, 2H); 7.78 (d,  $J = 7.4$  Hz, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz):  $\delta = 21.7$ ; 23.4; 25.0; 26.5 (2C); 40.3; 47.5 (2C); 58.8; 67.0; 82.7; 113.5; 120.2 (2C); 125.1 (2C); 127.2 (2C); 127.9 (2C); 141.5 (2C); 142.0; 143.9 (2C); 156.2; 165.7; 202.7. LC rt = 2.29 min (major) et 2.52 min (minor) (keto-enol equilibrium). LC-MS: (ESI<sup>+</sup>): m/z 464.3 ([M+H]<sup>+</sup>) 100%, 486.3 ([M+Na]<sup>+</sup>) 10%. HRMS (ESI) calcd. for C<sub>28</sub>H<sub>34</sub>NO<sub>5</sub><sup>+</sup>: 464.2437, found 464.2416. FT-IR:  $\nu_{\max}$  (cm<sup>-1</sup>) 3375; 3041; 2958; 1739; 1717; 1536; 1452; 1353; 1249; 1237; 1222; 1126; 1098; 1038; 942; 760; 737; 730.



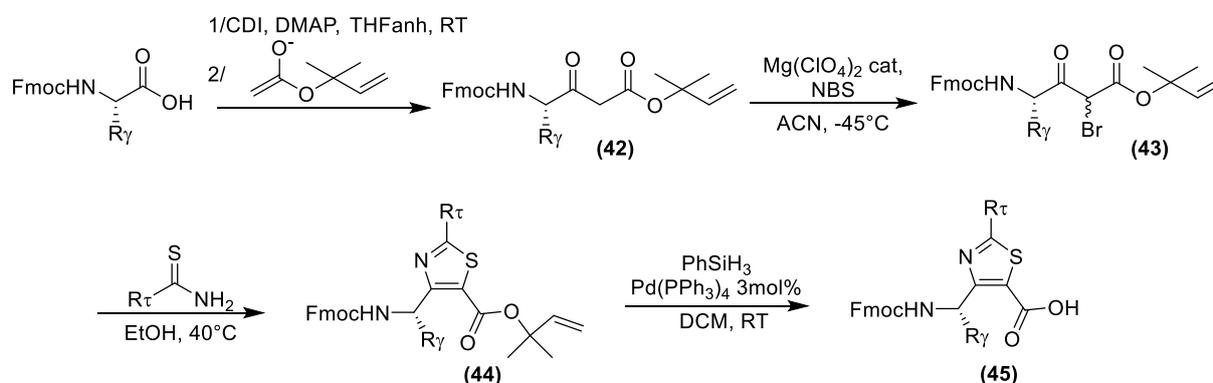
**(42h)**: Orange oil. Yield: 53%.  $[\alpha]_D^{20^\circ\text{C}} = +4.7^\circ$  (c 2.00,  $\text{CHCl}_3$ ).  $^1\text{H NMR}$  ( $\text{CDCl}_3$ , 400 MHz):  $\delta = 1.44$  (s, 9H), 1.53–1.57 (m, 6H), 1.82–1.90 (m, 1H), 2.17–2.38 (m, 3H), 3.47 (AB,  $J = 16.0$  Hz, 1H), 3.54 (AB,  $J = 16.0$  Hz, 1H), 4.21 (t,  $J = 6.6$  Hz, 1H), 4.38–4.47 (m, 3H), 5.09 (d,  $J = 11.3$  Hz, 1H), 5.18 (d,  $J = 17.4$  Hz, 1H), 5.60 (d,  $J = 7.4$  Hz, 1H), 6.08 (dd,  $J = 11.3, 17.4$  Hz, 1H), 7.32 (t,  $J = 7.4$  Hz, 2H), 7.40 (t,  $J = 7.4$  Hz, 2H), 7.59 (d,  $J = 7.4$  Hz, 2H), 7.76 (d,  $J = 7.4$  Hz, 2H) ppm.  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ , 100 MHz):  $\delta = 26.1, 26.4, 26.5, 28.2$  (3C), 31.2, 47.4 (2C), 59.7, 67.1, 81.1, 82.8, 113.5, 120.2 (2C), 125.2 (2C), 127.2 (2C), 127.9 (2C), 141.5, 141.9 (2C), 143.8, 143.9, 156.2, 165.6, 172.3, 201.7 ppm. LC rt = 2.54 min (major) et 2.73 min (minor) (keto-enol equilibrium). LC-MS: (ESI<sup>+</sup>): m/z 558.3 ( $[\text{M}+\text{Na}]^+$ ) 100%. HRMS (ESI) calcd. for  $\text{C}_{31}\text{H}_{38}\text{NO}_7^+$ : 536.2648, found 536.2648.



**(42j)**: Orange oil. Yield = 53%.  $[\alpha]_D^{20^\circ\text{C}} = -11.0^\circ$  (c 1.00,  $\text{CHCl}_3$ ).  $^1\text{H NMR}$  ( $\text{CDCl}_3$ , 400 MHz):  $\delta = 1.42$  (s, 9H), 1.51 (s, 6 H), 2.65–2.90 (dd,  $J = 4.6$  and 16.9 Hz, 2H), 3.47 (s, 2H), 4.21 (t,  $J = 6.6$  Hz, 1H), 4.38–4.43 (m, 1H), 4.46–4.48 (m, 1H), 4.55–4.57 (m, 1 H), 5.06 (d,  $J = 10.7$  Hz, 1H), 5.15 (d,  $J = 17.6$  Hz, 1H), 5.87 (d,  $J = 9.6$  Hz, 1H), 6.08 (dd,  $J = 10.7, 16.9$  Hz, 1H), 7.29 (t,  $J = 7.6$  Hz, 2H), 7.38 (t,  $J = 7.6$  Hz, 2H), 7.56 (d,  $J = 7.6$  Hz, 2H), 7.74 (d,  $J = 7.6$  Hz, 2H) ppm.  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ , 100 MHz):  $\delta = 26.4$  (2C), 28.2 (3C), 36.7, 47.4, 56.9, 67.3, 82.2, 113.4, 120.2 (2C), 125.2 (2C), 127.2 (2C), 127.9 (2C), 141.5, 142.0, 143.8, 156.3, 165.8, 201.7 ppm. LC rt = 2.06 min (major) and 2.27 min (minor; keto-enol equilibrium). LC-MS (ESI<sup>+</sup>) : m/z (%) = 544.3 (100)  $[\text{M}+\text{Na}]^+$ . HRMS (ESI): calcd. for  $\text{C}_{30}\text{H}_{36}\text{NO}_7$   $[\text{M}+\text{H}]^+$  522.2490; found 522.2492.

## 2. General procedure for ATC-monomer synthesis

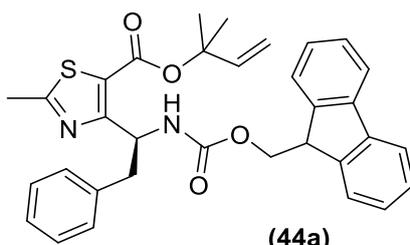
(45a) to (45g) were synthesized following Scheme S1.<sup>[189]</sup>



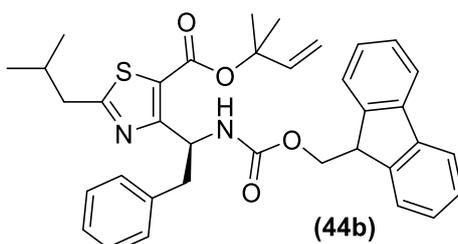
Scheme S1

**Synthesis of bromo- $\beta$ -ketoesters (43):** To a solution of the  $\beta$ -ketoester (42) (1.0 equiv., 17.5mmol) in acetonitrile (100mL) was added magnesium perchlorate (0.33 equiv., 5.77mmol, 1.29g). The solution was stirred at  $-45^{\circ}\text{C}$  for 10 min. A solution of NBS (1.05 equiv., 18.36mmol, 3.27g) in acetonitrile (160mL) was then added dropwise over 20 min. The reaction was completed after few minutes stirring at  $-45^{\circ}\text{C}$  (HPLC monitoring). The mixture was diluted with Et<sub>2</sub>O (150mL), washed with water (2 $\times$ 100mL) and brine (3 $\times$ 100mL). The organic layer was dried on MgSO<sub>4</sub>, filtered and the solvent was removed under vacuum to give the crude product which was used in the next step without further purification.

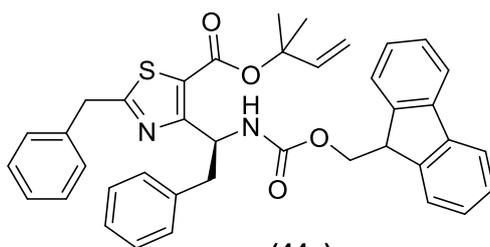
**Hantzsch cyclization to access (44):** To a solution of  $\alpha$ -mono-brominated  $\beta$ -ketoester (43) (1.0 equiv, 15.6mmol) in absolute ethanol was added a solution of the thioamide (3 equiv., 46.7mmol) dissolved in absolute ethanol. The solution was heated for 2 h at  $35^{\circ}\text{C}$  until completion of the reaction (HPLC and TLC monitoring). EtOH was evaporated and the yellowish solid was partitioned between EtOAc (50mL) and water (50mL). The organic layer was washed with water (3 $\times$ 50 mL) and brine (1 $\times$ 50mL). The combined organic layers were dried on magnesium sulfate, filtered and evaporated under vacuum to yield the crude product. Purification by column chromatography on silica gel afforded ATC (44).



**(44a)**: White solid. Yield: 40-56%. Mp: 50-51°C.  $[\alpha]_D^{20^\circ\text{C}} = +2.9^\circ$  (c 1.00,  $\text{CHCl}_3$ ).  $^1\text{H NMR}$  ( $\text{CDCl}_3$ , 500 MHz):  $\delta = 1.62$  (d,  $J = 5.7$  Hz, 6H); 2.67 (s, 3H); 3.13 (d,  $J = 6.8$  Hz, 2H); 4.21 (m, 2H); 4.39 (m, 1H); 5.14 (d,  $J = 11.0$  Hz, 1H); 5.23 (d,  $J = 17.5$  Hz, 1H); 5.98 (m, 1H); 6.11 (m, 2H); 7.09 (m, 2H); 7.18-7.23 (m, 3H); 7.29 (t,  $J = 7.4$  Hz, 2H); 7.39 (t,  $J = 7.4$  Hz, 2H); 7.55 (m, 2H); 7.76 (d,  $J = 7.4$  Hz, 2H).  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ , 100 MHz):  $\delta = 19.4$ ; 26.5; 26.6; 42.2; 47.2; 51.2; 66.8; 83.2; 113.4; 120.0 (2C); 124.4; 125.2; 125.3; 126.6; 127.0 (2C); 127.7 (2C); 128.2 (2C); 129.4; 129.6; 137.2; 141.3 (2C); 141.9; 144.1 (2C); 155.7; 160.0 (2C); 169.8. LC rt = 2.18 min. LC-MS: (ESI+):  $m/z$  553.3 ( $[\text{M}+\text{H}]^+$ ) 100%, 575.3 ( $[\text{M}+\text{Na}]^+$ ) 10%. HRMS (ESI) calcd. for  $\text{C}_{33}\text{H}_{33}\text{N}_2\text{O}_4\text{S}^+$ : 553.2161, found 553.2166. FT-IR:  $\nu_{\text{max}}$  ( $\text{cm}^{-1}$ ) 2925; 2310; 1703; 1501; 1449; 1319; 1245; 1131; 1083; 738; 699.

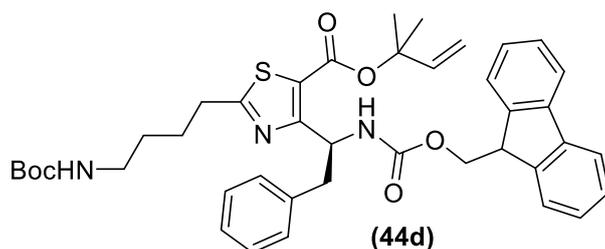


**(44b)**: Orange oil. Yield: 26%.  $^1\text{H NMR}$  ( $\text{CDCl}_3$ , 500 MHz):  $\delta = 0.97$  (dd, 7H); 1.58 (m, 6H); 2.77 (dd, 2H); 3.12 (m, 2H); 4.17 (t,  $J = 6.9$  Hz, 1H); 4.24 (t,  $J = 7.2$  Hz, 2H); 4.34 (dd,  $J = 7.6$  Hz and 2.9 Hz, 1H); 5.10 (d,  $J = 10.9$  Hz, 1H); 5.19 (d,  $J = 17.4$  Hz, 1H); 5.91 (m, 1H); 6.08 (dd,  $J = 10.9$  Hz and 17.4 Hz, 1H); 7.02 (m, 2H); 7.15 (m, 3H); 7.28 (m, 2H); 7.36 (m, 2H); 7.55 (m, 2H); 7.75 (d,  $J = 7.2$  Hz, 2H).  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ , 100 MHz):  $\delta = 22.2$ ; 23.5; 26.6 (2C); 29.7 (3C); 33.4; 42.3; 47.3; 51.6; 53.6; 59.7; 66.8; 82.7; 113.5; 120.1 (2C); 125.4 (2C); 126.5; 127.1 (2C); 127.8 (2C); 128.1 (2C); 129.8 (2C); 137.5; 141.4 (2C); 142.0; 144.2 (2C); 173.3 ppm. LC rt = 2.75min.



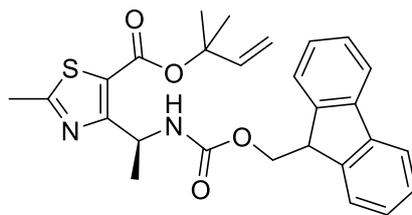
(44c)

(44c): White solid. Yield: 65%.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  = 1.57 (m, 6H); 3.15 (m, 2H); 4.20 (t,  $J$  = 6.9 Hz, 1H); 4.24 (s, 2H); 4.29 (dd,  $J$  = 6.9 Hz and 3.3 Hz, 1H); 4.39 (dd,  $J$  = 6.9 Hz and 3.3 Hz, 1H); 5.10 (d,  $J$  = 10.9 Hz, 1H); 5.18 (d,  $J$  = 17.4 Hz, 1H); 5.94 (m, 2H); 6.08 (dd,  $J$  = 10.9 Hz and 17.4 Hz, 1H); 7.06 (m, 2H); 7.15-7.21 (m, 3H); 7.28 (m, 2H); 7.30 (m, 3H); 7.35 (m, 2H); 7.40 (t,  $J$  = 7.5 Hz, 2H); 7.55 (d,  $J$  = 7.5 Hz, 1H); 7.58 (d,  $J$  = 7.5 Hz, 1H); 7.76 (d,  $J$  = 7.5 Hz, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 100 MHz):  $\delta$  = 26.9 (2C); 40.0; 42.3; 43.2; 47.5; 51.5; 66.8; 76.8; 83.2; 113.4; 120.1 (2C); 125.5 (2C); 126.6; 127.1 (2C); 127.6; 127.7 (2C); 128.2 (2C); 129.0 (2C); 129.1 (2C); 129.7 (2C); 136.7; 141.4 (2C); 141.9; 144.1 (2C); 155.5; 160.0; 169.3; 173.8 LC rt = 2.58min. HRMS (ESI) calcd. for  $\text{C}_{39}\text{H}_{37}\text{N}_2\text{O}_4\text{S}^+$ : 629.2472, found 629.2474.

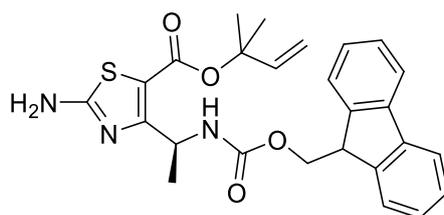


(44d)

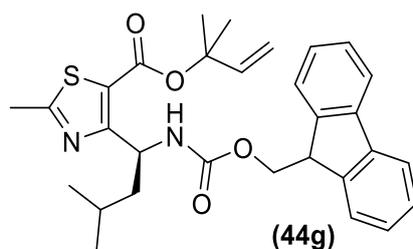
(44d): Orange oil. Yield: 56%.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  = 1.44 (s, 9H); 1.55 (m, 2H); 1.59 (m, 6H); 1.78 (m, 2H); 2.94 (t,  $J$  = 7.4 Hz, 2H); 3.12 (m, 2H); 3.16 (m, 2H) 4.19 (m, 1H); 4.31 (m, 2H); 4.60 (s, 2H); 5.12 (d,  $J$  = 10.6 Hz, 1H); 5.21 (d,  $J$  = 17.5 Hz, 1H); 5.9 (s, 1H); 6.10 (dd,  $J$  = 10.6 Hz and 17.5Hz, 1H); 7.01-7.12 (m, 2H); 7.13-7.22 (m, 3H); 7.29 (t,  $J$  = 6.8 Hz, 2H); 7.38 (t,  $J$  = 7.4 Hz, 2H); 7.56 (dd,  $J$  = 7.4 Hz and 6.8Hz, 2H); 7.76 (d,  $J$  = 7.4 Hz, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 100 MHz):  $\delta$  = 26.9 (2C); 27.1; 28.6 (3C); 29.6; 33.3; 40.3; 42.4; 47.4; 51.7; 66.9; 77.2; 78.8; 83.2; 113.5; 120.1 (2C); 125.4 (2C); 126.6; 127.2 (2C); 127.9 (2C); 128.6 (2C); 129.8 (2C); 135.9; 141.0 (2C); 142.0; 142.4; 144.0 (2C); 151.2; 155.9; 160.2; 174.1 HRMS (ESI) calcd. for  $\text{C}_{41}\text{H}_{48}\text{N}_3\text{O}_4\text{S}^+$ : 710.3247, found 710.3240.

**(44e)**

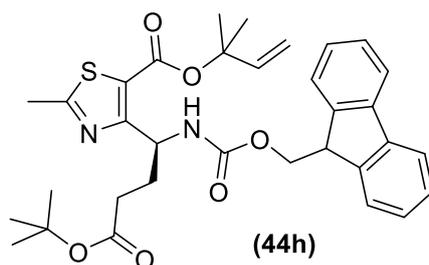
**(44e)**: White solid. Yield: 69 %. Mp: 51-53°C.  $[\alpha]^{D20}_{20} = +8.37$  (c 1.00, CHCl<sub>3</sub>). <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz):  $\delta$  = 1.48 (d,  $J$  = 6.5 Hz, 3H); 1.66 (s, 6H); 2.68 (s, 3H); 4.23 (t,  $J$  = 7.8 Hz, 1H); 4.36 (d,  $J$  = 6.5 Hz, 2H); 5.15 (d,  $J$  = 10.9 Hz, 1H); 5.27 (d,  $J$  = 17.4 Hz, 1H); 5.68 (m, 1H); 6.01 (d,  $J$  = 8.7 Hz, 1H); 6.18 (dd,  $J$  = 10.9, 17.4 Hz, 1H); 7.30 (t,  $J$  = 7.5 Hz, 2H); 7.39 (t,  $J$  = 7.5 Hz, 2H); 7.61 (d,  $J$  = 7.5 Hz, 2H); 7.76 (d,  $J$  = 7.5 Hz, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz):  $\delta$  = 19.5; 22.1; 26.7 (2C); 46.6; 47.4; 66.8; 83.2; 113.5; 120.0 (2C); 123.1; 125.3 (2C); 127.1 (2C); 127.7 (2C); 141.4; 142.0 (2C); 144.1; 144.2; 155.7; 160.3; 162.6; 169.8. LC rt = 2.49 min. LC-MS: (ESI+):  $m/z$  477.2 ([M+H]<sup>+</sup>) 100%. HRMS (ESI) calcd. for C<sub>27</sub>H<sub>29</sub>N<sub>2</sub>O<sub>4</sub>S<sup>+</sup>: 477.1844, found 477.1848. FTIR:  $\nu_{\max}$  (cm<sup>-1</sup>) 3317; 2980; 1705; 1499; 1448; 1366; 1315; 1233; 1192; 1132; 1082; 923; 830; 758.

**(44f)**

**(44f)**: White solid. Yield: 41%. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  = 1.44 (d,  $J$  = 6.7 Hz, 3H); 1.62 (m, 6H); 4.22 (t,  $J$  = 7.3 Hz, 1H); 4.36 (d,  $J$  = 7.5 Hz, 2H); 5.12 (ABX,  $J$  = 10.9 Hz, 17.5 Hz, 1H); 5.23 (ABX,  $J$  = 10.9 Hz, 17.5 Hz, 1H); 5.31 (s, 2H); 5.51 (quint,  $J$  = 7.5 Hz, 1H); 5.90 (d,  $J$  = 8.5, 1H); 6.16 (ABX,  $J$  = 10.9 Hz and 17.5 Hz, 1H); 7.30 (t,  $J$  = 7.5 Hz, 2H); 7.38 (t,  $J$  = 7.5 Hz, 2H); 7.60 (t,  $J$  = 7.5 Hz, 2H); 7.76 (d,  $J$  = 7.5 Hz, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz):  $\delta$  = 21.7; 26.7 (2C); 46.7; 47.3; 66.7; 82.5; 113.1; 119.9 (2C); 125.2 (2C); 127.0 (2C); 127.8 (2C); 128.6 (2C); 129.8 (2C); 141.0 (2C); 142.3; 144.0; 155.5; 169.8. HRMS (ESI) calcd. for C<sub>26</sub>H<sub>28</sub>N<sub>3</sub>O<sub>4</sub>S<sup>+</sup>: 478.1805, found 478.1801.

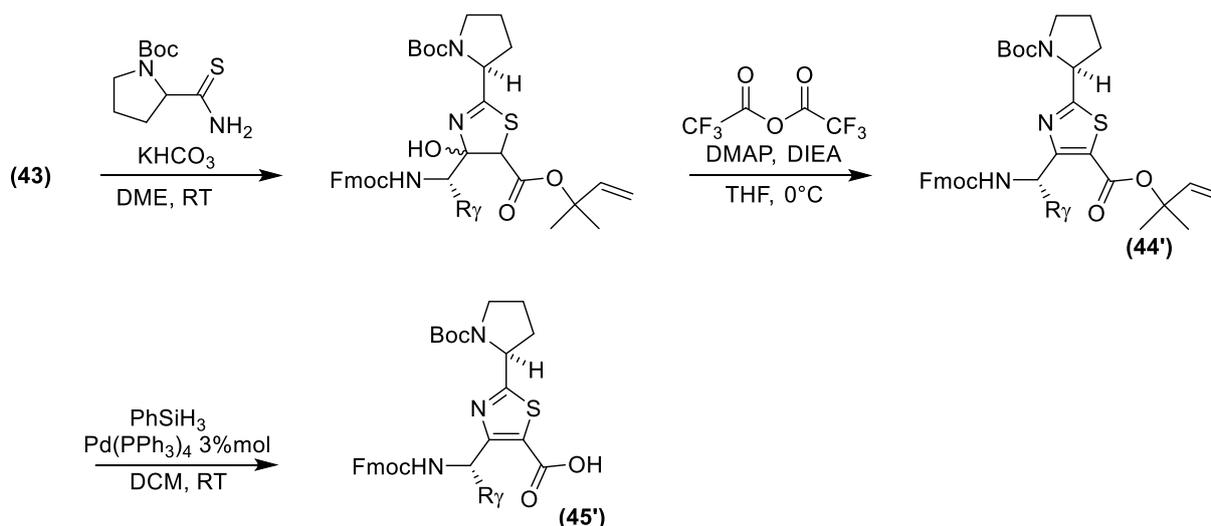


**(44g)**: White solid. Yield: 60%. Mp: 98-99°C.  $[\alpha]_{\text{D}}^{20} = -0.6$  (c 1.00, CHCl<sub>3</sub>). <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz)  $\delta$  = 0.95 (d, *J* = 6.4 Hz, 3H); 0.99 (d, *J* = 6.4 Hz, 3H); 1.56 (m, 2H); 1.66 (s, 6H); 1.75 (m, 1H); 2.67 (s, 3H); 4.21 (t, *J* = 7.4 Hz, 1H); 4.35 (m, 2H); 5.14 (d, *J* = 10.5 Hz, 1H); 5.26 (d, *J* = 17.5 Hz, 1H); 5.70 (m, 1H); 5.83 (m, 1H); 6.19 (dd, *J* = 17.5, 10.5 Hz, 1H); 7.29 (m, 2H); 7.39 (t, *J* = 7.4 Hz, 2H); 7.60 (m, 2H); 7.75 (d, *J* = 7.4 Hz, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz)  $\delta$  = 19.5; 22.3; 23.3; 25.1; 26.7; 26.8; 45.1; 47.4; 48.8; 66.7; 83.3; 113.5; 120.0 (2C); 123.8; 125.3 (2C); 127.1 (2C); 127.7 (2C); 141.4 (2C); 142.1; 144.1; 144.2; 155.9; 160.4; 162.2; 169.8. LC Rt = 2.45 min. LC-MS: (ESI+): *m/z* 519.3 ([M+H]<sup>+</sup>) 100%, 541.3 ([M+Na]<sup>+</sup>) 10%. HRMS (ESI) calcd. for C<sub>30</sub>H<sub>35</sub>N<sub>2</sub>O<sub>4</sub>S<sup>+</sup>: 519.2305, found 520.2318. FT-IR:  $\nu_{\text{max}}$  (cm<sup>-1</sup>) 3326; 2956; 1699; 1507; 1450; 1327; 1278; 1251; 1191; 1082; 1044; 923; 830; 759; 736



**(44h)**: White solid. Yield: 35%. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  = 1.42 (s, 9H); 1.65 (m, 6H); 2.07 (q, *J* = 7.5 Hz, 2H); 2.32 (m, 2H), 2.67 (s, 3H); 4.21 (t, *J* = 7.2 Hz, 1H); 4.35 (m, 2H); 5.15 (d, *J* = 10.9 Hz, 1H); 5.25 (d, *J* = 17.4 Hz, 1H); 5.59 (m, 1H); 5.97 (d, *J* = 9.3 Hz, 1H); 6.17 (dd, *J* = 10.9 Hz and 17.4 Hz, 1H); 7.30 (q, *J* = 7.4 Hz, 2H); 7.38 (t, *J* = 7.4 Hz, 2H); 7.60 (t, *J* = 7.4 Hz, 2H); 7.76 (d, *J* = 7.4 Hz, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz):  $\delta$  = 19.5; 26.7 (2C); 28.3 (3C); 31.1; 32.5; 47.5; 50.2; 66.9; 80.2; 83.3; 113.6; 120.1 (2C); 124.1; 125.4 (2C); 127.1 (2C); 127.8 (2C); 141.4; 142.0 (2C); 144.1 (2C); 155.5, 161.3; 169.9; 172.6. LC rt = 2.13min.

**(45i)**, **(45j)**, **(45k)** and **(45l)** were prepared following Scheme S2. The Hantzsch cyclization between **(43)** and N-Boc-proline thioamide was done in two steps in order to avoid any epimerization of the pyrrolidine moiety that might occur by HBr releasing.



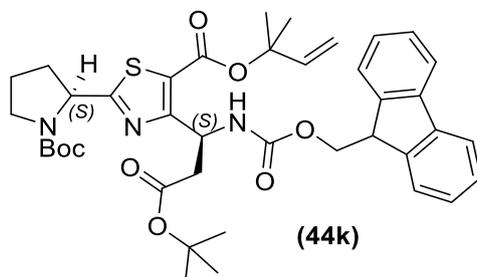
Scheme S2

**Cyclization and dehydration to access (44'):** In a three-neck flask under nitrogen atmosphere Boc-Proline-thioamide (9.29mmol, 1.1 equiv.) was dissolved in 160ml DME,  $\text{KHCO}_3$  (74.36mmol, 8 equiv.) was then added. After stirring for 15 min, the  $\alpha$ -monobrominated  $\beta$ -ketoester (43) (10.1mmol) dissolved in 100 ml DME was added drop-wise over 10 min. The reaction was completed after 18 h at room temperature (HPLC monitoring) leading the thiazolidine as two diastereoisomers. The organic layer was extracted with 60ml EtOAc then washed with  $3 \times 60\text{ml}$  brine. Dried, filtered and evaporated. The thiazoline was engaged in the next step without any more purification. Brown solid, yield 98%.

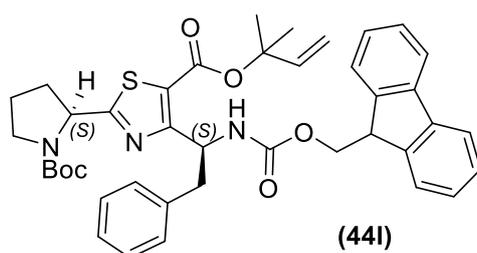
To a solution of the thiazoline (9.14mmol) dissolved in 100ml dry THF, DMAP (9.141mmol, 1 equiv.) and TFAA (36.56mmol, 4 equiv.) were added. After cooling at  $0^\circ\text{C}$ , DIEA (77.7mmol, 8.5 equiv.) dissolved in 50ml dry THF was added drop-wise. After stirring at  $0^\circ\text{C}$  for 2 hours, the reaction was completed (HPLC monitoring). The crude was extracted with 120ml DCM, washed with 100ml water and 100ml saturated  $\text{NaHCO}_3$  and  $3 \times 100\text{ml}$  brine. The organic layer was dried, filtered and evaporated. Purification by flash column chromatography gave pure (44'). Flash purification: Cyclohexane/Ethyl Acetate 80/20 for 1CV then gradient from 80/20 to 40/60 for 7CV.



127.2 (2C); 127.8 (2C); 141.5; 142.0 (2C); 144.1; 144.3; 156.0; 160.5; 161.3 (2C); 172.6; 179.3. LC (Column: Chromolith Speed Rod RP-C18, flow 3ml/min, Eluent: A - H<sub>2</sub>O/TFA 0,1%, B - ACN/TFA 0,1%, From 5:5 to 1:9 in 5 min): rt = 2.97 min. LC-MS: (ESI<sup>+</sup>): m/z 746.4 [M+H]<sup>+</sup> 100%. HRMS (ESI) calcd. for C<sub>41</sub>H<sub>52</sub>N<sub>3</sub>O<sub>8</sub>S<sup>+</sup>: 746.3475, found 746.3473.



**(44k)**: Yellowish solid. Yield: 33%. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ = 1.30-1.50 (m, 18 H); 1.66 (m, 6H); 1.94 (td, *J* = 4.3 Hz and 3.3 Hz, 2H); 2.23 (m, 2H); 2.78 (m, 2H); 3.52 (m, 2H); 4.21 (m, 1H); 4.33 (m, 2H); 5.04 (m, 1H); 5.15 (d, *J* = 10.9 Hz, 1H); 5.26 (d, *J* = 17.4 Hz, 1H); 5.94 (m, 1H); 6.05 (m, 1H); 6.19 (dd, *J* = 10.9 Hz and 17.4 Hz, 1H); 7.30 (t, *J* = 7.4 Hz, 2H); 7.38 (t, *J* = 7.4 Hz, 2H); 7.60 (t, *J* = 7.4 Hz, 2H); 7.76 (d, *J* = 7.4 Hz, 2H) ppm. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): δ = 23.7, 26.7 (2C), 28.3 (6C), 33.3, 41.5, 46.9, 47.3, 47.8, 50.4, 59.6, 67.0, 78.0, 80.5, 83.5, 113.7, 120.1 (2C), 124.0, 125.4 (2C), 127.1 (2C), 127.8 (2C), 141.4 (2C), 142.0, 144.2 (2C), 156.4, 161.3, 169.6, 171.3 and 179.6 ppm. LC (Column: Chromolith Speed Rod RP-C18, flow 3ml/min, Eluent: A - H<sub>2</sub>O/TFA 0,1%, B - ACN/TFA 0,1%, From 5:5 to 1:9 in 5 min): rt = 2.61 min. LC-MS: (ESI<sup>+</sup>): m/z 732.3 [M+H]<sup>+</sup> 100%. HRMS (ESI) calcd for C<sub>40</sub>H<sub>50</sub>N<sub>3</sub>O<sub>8</sub>S<sup>+</sup>: 732.3311, found 732.3319



**(44l)**: White solid. Yield: 39%. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ = 1.31-1.50 (m, 9H); 1.66 (m, 6H); 1.82-1.97 (m, 2H); 2.02-2.31 (m, 2H); 2.99-3.16 (m, 2H); 3.38-3.63 (m, 2H); 4.19 (m, 1H); 4.37 (m, 2H); 4.99 (m, 1H); 5.12 (m, 2H); 5.22 (m, 1H); 5.92 (m, 1H); 6.10 (m, 1H); 6.94 (m, 2H); 7.16 (m, 3H); 7.30 (t, *J* = 7.4 Hz, 2H); 7.39 (t, *J* = 7.4 Hz, 2H); 7.57 (m, 2H); 7.76 (d, *J* = 7.4 Hz, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): δ = 21.0; 22.7; 23.5; 26.6 (2C); 28.6 (3C); 33.4; 42.1; 46.8; 47.4; 51.3; 59.5; 66.8; 80.6; 83.2; 112.6; 113.5; 120.1 (2C); 125.4 (2C); 126.5; 127.1 (2C); 127.8 (2C); 128.1 (2C); 129.8 (2C); 137.5; 141.4 (2C); 142.0; 144.2 (2C); 161.3; 170.4; 171.4 ppm. LC rt = 2.57min. HRMS (ESI) calcd for C<sub>41</sub>H<sub>46</sub>N<sub>3</sub>O<sub>6</sub>S<sup>+</sup>: 708.3101, found 708.3107

**Removal of the dimethylallylester:** Under an argon atmosphere, Ac-ATC-O-dimethylallylester (**44**) (0.68 mmol) was dissolved in dry THF (10 mL). Tetrakis (triphenylphosphine)palladium<sup>(0)</sup> (3 mol %) was then added to the solution and the orange mixture was stirred for 5 min at r.t.. Then PhSiH<sub>3</sub> (1.2 equiv.) was added in one portion and the solution was stirred for 30 min (HPLC monitoring). The solvent was then evaporated under vacuum to give a brown foam of (**45**).

**(45a):** Quantitative yield. LC rt = 1.97 min. LC-MS: (ESI+): m/z 485.1 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.53 min. LC/MS: (ESI+): m/z 279.1 ([M+H]<sup>+</sup>) 100%.

**(45b):** Quantitative yield. LC rt = 2.24 min. LC-MS: (ESI+): m/z 527.0 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.52 min. LC/MS: (ESI+): m/z 279.1 ([M+H]<sup>+</sup>) 100%.

**(45c):** Quantitative yield. LC rt = 1.97 min. LC-MS: (ESI+): m/z 561.1 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.38 min. LC/MS: (ESI+): m/z 279.1 ([M+H]<sup>+</sup>) 100%.

**(45d):** Quantitative yield. LC rt = 1.95 min. LC-MS: (ESI+): m/z 642.1 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.42 min. LC/MS: (ESI+): m/z 279.0 ([M+H]<sup>+</sup>) 100%.

**(45e):** Quantitative yield. LC rt = 1.59 min. LC-MS: (ESI+): m/z 409.3 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.36 min. LC/MS: (ESI+): m/z 279.1 ([M+H]<sup>+</sup>) 100%.

**(45f):** Quantitative yield. LC rt = 1.41 min. LC-MS: (ESI+): m/z 410.0 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.41 min. LC/MS: (ESI+): m/z 279.0 ([M+H]<sup>+</sup>) 100%.

**(45g):** Quantitative yield. LC rt = 2.02 min. LC-MS: (ESI+): m/z 451.1 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.57 min. LC/MS: (ESI+): m/z 279.1 ([M+H]<sup>+</sup>) 100%.

**(45h):** Quantitative yield. LC rt = 2.03 min. LC-MS: (ESI+): m/z 467.1 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.55 min. LC/MS: (ESI+): m/z 279.1 ([M+H]<sup>+</sup>) 100%.

**(45i):** Quantitative yield. LC rt = 1.78 min. LC-MS: (ESI+): m/z 678.5 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.22 min. LC/MS: (ESI+): m/z 279.1 ([M+H]<sup>+</sup>) 100%.

**(45j):** Quantitative yield. LC rt = 1.78 min. LC-MS: (ESI+): m/z 678.5 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide : LC rt = 1.22 min. LC/MS: (ESI+): m/z 279.1 ([M+H]<sup>+</sup>) 100%.

**(45k):** Quantitative yield. LC rt = 1.97 min. LC-MS: (ESI+): m/z 664.6 ([M+H]<sup>+</sup>) 100%.  
Triphenylphosphine oxide: LC rt = 1.32 min. LC/MS: (ESI+): m/z 279.2 ([M+H]<sup>+</sup>) 100%.



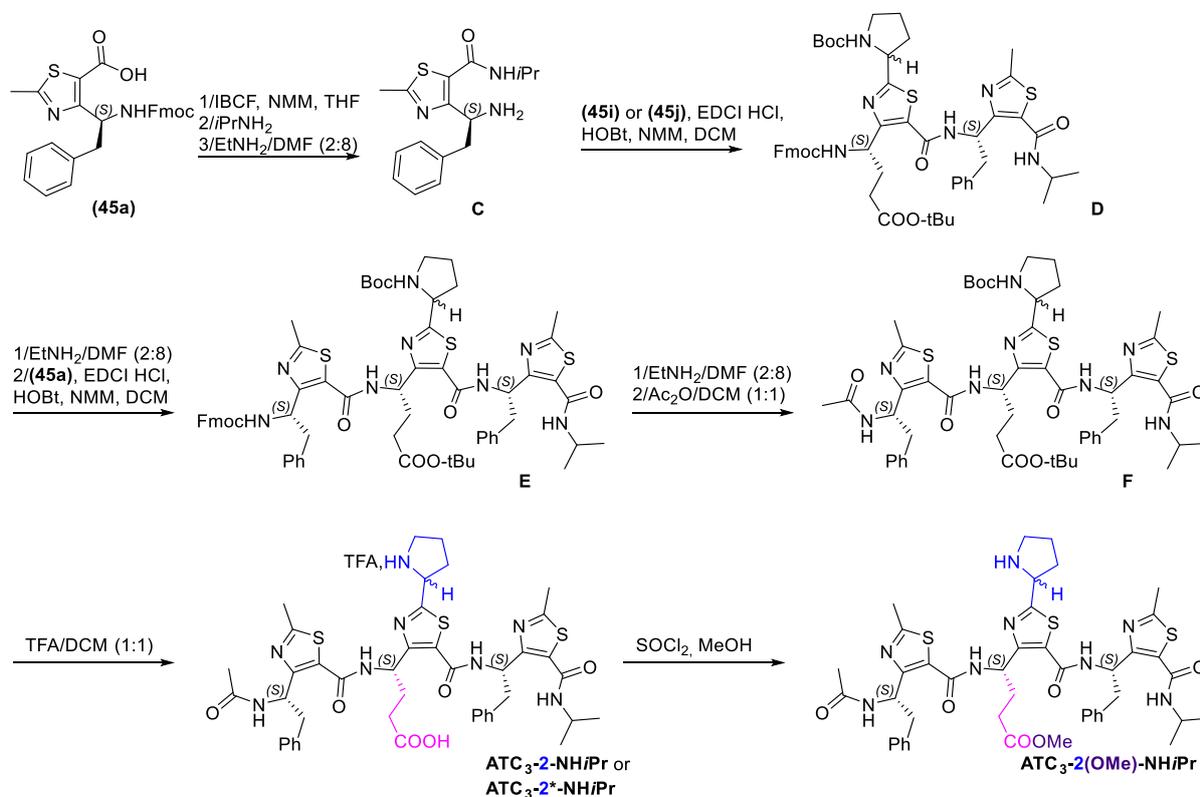
then the solution was stirred 10 min at room temperature (TLC monitoring: alumina, 100 % DCM). The solution was diluted using 10 ml DCM then washed twice with 1M aq KHSO<sub>4</sub> and brine. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. The crude was purified on silica (eluent: DCM). White solid. Yield: 71%. LC rt = 2.42 min. LC-MS: (ESI+): m/z 719.3 ([M+H]<sup>+</sup>) 100%. HRMS (ESI) calcd. for C<sub>39</sub>H<sub>51</sub>N<sub>4</sub>O<sub>7</sub>S<sup>+</sup>: 719.3478, found 719.3486.

**Synthesis of B:** **A** (2.96 mmol, 1.0 equiv.) was dissolved in 40 ml DMF/Et<sub>2</sub>NH (8:2 vv). The solution was stirred at room temperature then was evaporated under reduced pressure. The crude was dissolved in 20 ml DCM/Ac<sub>2</sub>O (1:1 vv) and stirred 15 min at RT. After solvent removal, **B** was purified by flash chromatography on silica gel (DCM/EtOAc 85/15 to 4/6). White solid. Yield 80%. <sup>1</sup>H NMR (DMSO-d<sub>6</sub>, 400 MHz): δ = 1-1.2 (m, 11H), 1.3-1.5 (d, J=24 Hz, 14H), 1.7-2.4 (m, 11H), 3.3-3.6 (m, 2H), 3.9-4.0 (m, 1H), 4.8-5.1 (m, 2H) ppm. <sup>13</sup>C NMR (DMSO-d<sub>6</sub>, 100 MHz): δ = 22.5 (3C), 23.81, 23.83 (3C), 28.18 (3C), 29.08, 31.42, 33.02, 41.90, 47.19, 47.96, 59.19, 79.86, 80.19, 131.15, 153.17, 153.78, 159.74, 171.30 (2C), 176.19. LC rt=2.85 min. LC-MS (ESI+): m/z (%) = 539.3 (100) [M+H]<sup>+</sup>, 561.3 (3) [M+Na]<sup>+</sup>. HRMS (ESI) calcd. for C<sub>26</sub>H<sub>43</sub>N<sub>4</sub>O<sub>6</sub>S<sup>+</sup>: 539.2903, found 539.2910.

**Synthesis of (47):** **B** was stirred 1h in 10 ml TFA/DCM (1:1 vv). <sup>1</sup>H NMR (DMSO-d<sub>6</sub>, 300 MHz): δ = 1.17 (d, 3H, J = 6.5 Hz); 1.20 (3H, J = 6.5 Hz); 1.89 (s, 3H); 1.97-2.21 (m, 7H); 2.40-2.48 (m, 1H); 3.3-3.4 (m, 2H), 3.9-4.1 (dd, J=15 Hz, 4H), 5.0-5.1 (m, 2H), 8.6-8.9 (d, 1H), 9-9.2 (br, 1H), 9.5-9.6 (d, 1H), 9.8-9.9 (br, 1H) ppm. <sup>13</sup>C NMR (DMSO-d<sub>6</sub>, 100 MHz): δ = 22.00 (2C) 23.02, 29.92, 30.44, 30.7, 30.98, 45.20, 47.59, 47.96, 58.9, 133.21, 153.15, 158.62, 164.96, 170.89, 173.52 ppm. LC tr = 0.82 min. LC-MS (ESI+): m/z (%) = 385.3 (100) [M+H]<sup>+</sup>, 407.3 [M+Na]<sup>+</sup>.

### 3. Synthesis of ATC<sub>3</sub>-2-NHiPr:

ATC<sub>3</sub>-2-NHiPr was prepared in solution following Scheme S4.



Scheme S4

**Synthesis of C:** IBCF (130  $\mu$ l, 0.98 mmol, 1.2 equiv.) and *N*-methylmorpholine (110  $\mu$ l, 0.98 mmol, 1.2 eq) were added to a solution of (45a) (400 mg, 0.82 mmol, 1.0 equiv.) in 2 ml anhydrous THF at 0°C. The suspension was stirred for 15 min at 0°C. *i*PrNH<sub>2</sub> (84  $\mu$ l, 1.2 equiv.) was then added to the mixture then the solution was stirred for 10 min at room temperature (TLC monitoring: alumina, 100 % DCM). The solution was diluted using 10 ml DCM then washed twice with 1M aq KHSO<sub>4</sub> and brine. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and evaporated under reduced pressure. White solid. Yield: 86%. LC-MS: (ESI+): *m/z* 526.2 ([M+H]<sup>+</sup>) 100%. The resulting material was then stirred for 1 hour at RT in 5 ml Et<sub>2</sub>NH / DMF (2:8 v/v). The solvent was removed under reduced pressure yielding **C** that was used without any purification.

**Synthesis of D:** To a solution of **C** (206 mg, 0.68 mmol, 1.0 equiv.) and (45i) (460 mg, 0.68 mmol, 1.0 equiv.) in 5 ml anhydrous DCM was added EDCI, HCl (143 mg, 1.1 equiv.), HOBT (100 mg, 1.1 equiv.) and NMM (82  $\mu$ l, 1.1 equiv.). The solution was stirred for 3 hours at RT to completion (HPLC monitoring). The solvent was removed under reduced pressure then the

crude was purified by flash chromatography on silica gel (DCM / AcOEt from 100/0 to 65/35). Yield: 74%. LC-MS: (ESI+):  $m/z$  963.5 ( $[M+H]^+$ ) 100%.  $^1H$  NMR ( $CDCl_3$ , 600 MHz, 278K):  $\delta$  ppm 1.29-1.46 (24H, m); 1.88-2.35 (8H, m); 2.71 (3H, s); 2.89 (1H, br); 3.14 (1H, br); 3.33-3.40 (1H, m); 3.44-3.63 (2H, m); 4.23 (2H, m); 4.38 (2H, m); 5.03 (0.6H, d,  $J=7.8$ Hz); 5.13 (0.4H, d,  $J=8.6$ Hz); 5.22 (1H, m); 5.91 (1H, t,  $J=10.4$ Hz); 7.17-7.21 (5H, m); 7.31 (2H, t,  $J=7.1$ Hz); 7.39 (2H, m); 7.56 (2H, dd,  $J=7.1$ Hz,  $J=3.3$ Hz); 7.78 (2H, d,  $J=7.1$  Hz); 8.97 (0.6H, br); 9.01 (0.4H, br); 9.31 (0.4H, br); 9.50 (0.6H, d,  $J=6.5$ Hz).  $^{13}C$  NMR ( $CDCl_3$ , 125 MHz, 278K):  $\delta$  ppm 19.1; 22.2; 22.8; 23.1 and 23.8 (0.6 and 0.4C); 28.1 (3C); 28.3 (3C); 29.8 and 32.7 (0.6 and 0.4C); 32.5; 22.9; 39.7; 42.3; 46.7 and 47.1 (0.4 and 0.6C); 47.1; 50.2; 51.7; 59.1; 67.4; 80.5; 80.8; 120.1-120.3 (2C); 125.1-125.2 (2C), 127.1-127.9-128.7-128.9 (10C); 132.5; 134.2; 137.1; 141.3 (2C); 143.6-143.8 (2C); 143.9; 154.1; 154.7; 157.0 (2C); 162.2 (2C); 172.3; 172.5; 177.8 and 178.5 (0.4 and 0.6C). Missing 1 carbon.

**Synthesis of E: D** (431 mg, 0.625 mmol, 1.0 equiv.) was stirred for 1 hour at RT in 5 ml  $Et_2NH$  / DMF (2:8 vv). The solvent was removed under reduced pressure. The crude was diluted in 5 ml anhydrous DCM containing (**45a**) (303 mg, 0.625 mmol, 1.0 equiv.), EDCI, HCl (132 mg, 1.1 equiv), HOBt (93 mg, 1.1 equiv.) and NMM (76  $\mu$ l, 1.1 equiv.). The solution was stirred 3h at rt to completion (HPLC monitoring). The solvent was removed under reduced pressure then the crude was purified by flash chromatography on silica gel (DCM / AcOEt from 100/0 to 6/4). Yield: 50%. LC-MS: (ESI+):  $m/z$  1207.5 ( $[M+H]^+$ ) 100%, 604.2 ( $[M+2H]^{2+}$ ) 22%.  $^1H$  NMR ( $CDCl_3$ , 600 MHz, 278K):  $\delta$  ppm 1.27 (9H, s); 1.28 (9H, s); 1.31 (3H, m); 1.39 (3H, m); 1.71-2.32 (8H, m); 2.69 (6H, s); 3.05 (1H, dd,  $J=14.5$ Hz,  $J=6.5$ Hz); 3.15-3.30 (2H, m); 3.35-3.47 (2H, m); 3.51 (1H, m); 4.14 (1H, t,  $J=7.2$ Hz); 4.18 (1H, m); 4.29 (1H, dd,  $J=10.5$ Hz,  $J=7.2$ Hz); 4.33-4.44 (1H, m); 5.01 (0.6H, d,  $J=7.9$ Hz); 5.11 (0.4H, d,  $J=7.0$ Hz); 5.41-5.51 (3H, m); 5.76 (1H, br); 6.96-7.32 (12H, m); 7.35-7.39 (2H, m); 7.47-7.56 (2H, m); 7.70-7.76 (2H, m); 9.07 (1H, br); 9.42 (0.4H, br); 9.62 (0.6H, br); 10.10 (0.4H, br); 10.26 (0.6H, br).  $^{13}C$  NMR ( $CDCl_3$ , 125 MHz, 278K):  $\delta$  ppm 19.2; 19.3; 22.2; 22.9; 28.0 (3C); 28.4 (3C); 28.5; 29.0; 32.5; 34.0; 39.7. 39.9; 42.2; 46.7; 47.0; 50.0; 50.1; 52.2; 59.0 (0.4C); 59.6 (0.6C); 67.5; 80.4; 80.6; 120.0 (2C); 125.2 (2C); 126.9-127.1-127.8-128.5-128.6-128.7-128.8-129.0 (14C); 132.1; 132.2; 134.8; 136.3; 137.4; 141.3 (2C); 143.5; 143.7; 154.2; 154.2-154.6 (3C); 157.2 (2C); 162.1 (2C); 171.7 (2C); 172.0; 178.6.

**Synthesis of F: E** (350 mg, 0.29 mmol) was stirred for 1 hour at RT in 5 ml  $Et_2NH$  / DMF (2:8 vv). The solvent was removed under reduced pressure. The crude was diluted in 4 ml DCM /  $Ac_2O$  (1:1 vv) and stirred for 15 min at room temperature. The solvent was removed under

reduced pressure then the crude was purified by flash chromatography on silica gel (DCM / AcOEt from 9/1 to 2/8). 201 mg. Yield: 67%. LC rt = 2.31 min. LC-MS: (ESI+): m/z 1027.6 ( $[M+H]^+$ ) 30%; 514 ( $[M+2H]^{2+}$ ) 12%; 486.0 ( $[M-tBu+2H]^{2+}$ ) 30%; 436.0 ( $[M-tBu-Boc+2H]^{2+}$ ) 100%. HRMS (ESI) calcd for  $C_{52}H_{67}N_8O_8S_3^+$ : 1027.4244, found 1027.4247.  $^1H$  NMR ( $d_6$ -DMSO, 400 MHz):  $\delta$  ppm 1.14-1.38 (24H, m); 1.84 (3H, s); 1.83-1.85 (4H, m); 2.05-2.35 (4H, m); 2.63 (6H, s); 2.90 (1H, m); 3.12-3.16 (3H, m); 3.33 (2H, m); 3.97-4.06 (1H, m); 4.87 (0.5H, m); 4.96 (0.5H, m); 5.31 (0.5H, m); 5.40 (0.5H, m); 5.54 (2H, m); 6.97-7.27 (10H, m); 8.68 (1H, br); 9.05 (1H, br); 9.92 (0.5H, br); 8.98 (0.5H, br); 10.47 (1H, br).  $^{13}C$  NMR ( $d_6$ -DMSO, 100 MHz):  $\delta$  ppm 19.5 (2C); 22.5; 22.7 (2C); 28.1 (6C); 28.3; 28.5; 29.1; 29.5; 32.0; 32.5; 33.5; 41.8; 46.7; 47.1; 49.3; 49.7; 50.2; 52.3; 79.8; 80.2; 126.7; 126.8; 128.6 (2C); 128.7 (2C); 129.0; 129.1; 129.5; 129.7 (2C); 137.9; 138.6; 154.5; 156.0; 158.5; 158.9; 159.9; 161.3; 161.5; 166.9; 168.4; 171.5; 171.6; 176.7.

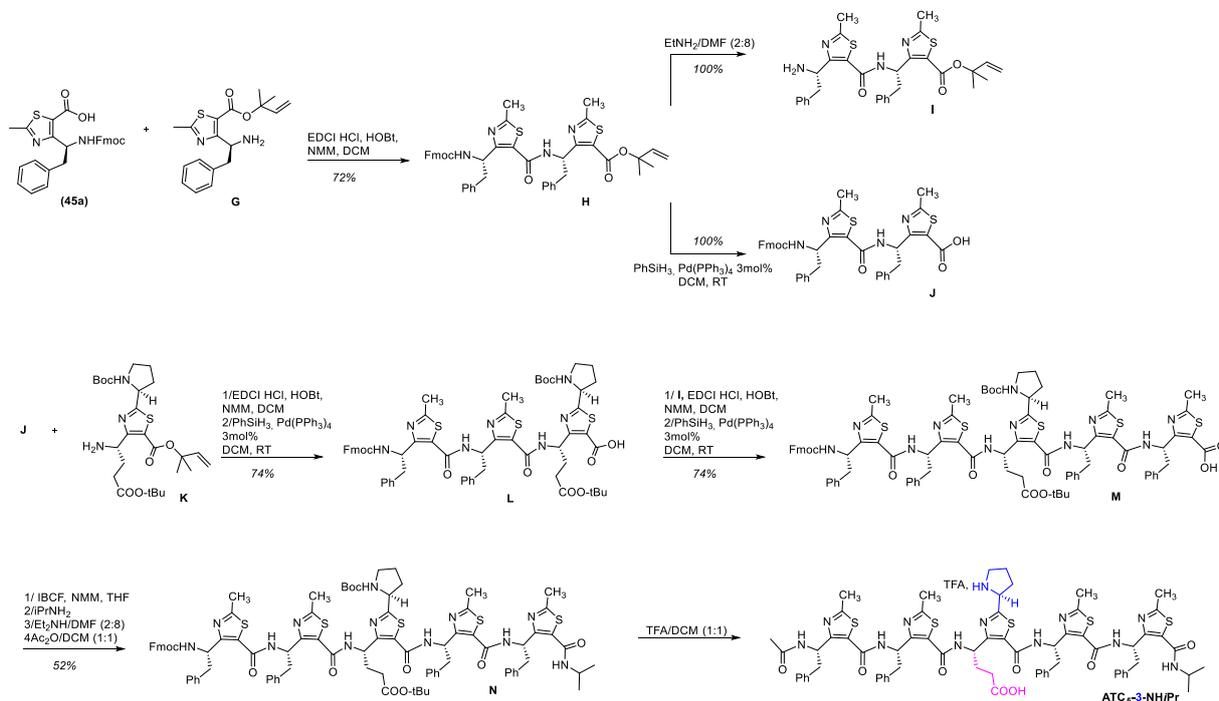
**Synthesis of ATC<sub>3</sub>-2-NHiPr:** **F** (195 mg, 0.19 mmol) was stirred for 1 hour at RT in 5 ml DCM/TFA (1:1 v/v). The solvent was removed under reduced pressure. The resulting material was then solubilized in ACN/Water (1:1 v/v) and lyophilized to yield **ATC<sub>3</sub>-2-NHiPr** as a white solid. Yield 100%.  $^1H$  NMR ( $CD_3OH$ , 600 MHz):  $\delta$  ppm 1.26 (3H, d,  $J = 6.5$  Hz); 1.30 (3H, d,  $J = 6.5$  Hz); 1.95 (3H, s); 2.05-2.45 (8H, m); 2.71 (3H, s); 2.73 (3H, s); 3.10 (1H, s); 3.11 (1H, s); 3.31 (2H, m); 3.41 (1H, m); 3.48 (1H, m); 4.07 (1H, m); 5.08 (1H, m); 5.31 (1H, m); 5.61 (1H, m); 5.63 (1H, m); 6.73 (2H, d,  $J = 7.5$  Hz); 7.14 (2H, m); 7.14-7.25 (6H, m); 8.85 (1H, d,  $J = 7.5$  Hz); 8.95 (1H, d,  $J = 7.5$  Hz); 9.90 (1H, d,  $J = 6.5$  Hz); 10.27 (1H, d,  $J = 6.5$  Hz). LC-MS: (ESI+): m/z 871.4 ( $[M+H]^+$ ) 35%, 436.2 ( $[M+2H]^{2+}$ ) 100%. HRMS (ESI) calcd. for  $C_{43}H_{51}N_8O_6S_3^+$ : 871.3094 found 871.3098. Overall yield 34%

**Synthesis of ATC<sub>3</sub>-2\*-NHiPr:** **ATC<sub>3</sub>-2\*-NHiPr** was prepared in solution following the same procedure as **ATC<sub>3</sub>-2-NHiPr**. LC-MS: (ESI+): m/z 871.4 ( $[M+H]^+$ ) 35%, 436.2 ( $[M+2H]^{2+}$ ) 100%. HRMS (ESI) calcd. for  $C_{43}H_{51}N_8O_6S_3^+$ : 871.3094 found 871.3098. Overall yield 21%

**ATC<sub>3</sub>-2(OMe)-NHiPr:** **ATC<sub>3</sub>-2(OMe)-NHiPr** was prepared by treating 1.0 equiv of **ATC<sub>3</sub>-2-NHiPr** with  $SOCl_2$  (8 mg) in methanol (1 mL). The solvent was removed under reduced pressure.

#### 4. Synthesis of ATC<sub>5</sub>-3-NHiPr

ATC<sub>5</sub>-3-NHiPr was prepared in solution following Scheme S5.



Scheme S5

**General procedure for Fmoc removal:** The Fmoc-protected compound was stirred for 1 hour at RT in Et<sub>2</sub>NH / DMF (2:8 vv). The solvent was removed under reduced pressure and the resulting product was used without any other purification.

**General procedure for removal of dimethylallyl protecting groups:** Under an argon atmosphere, the dimethylallylester (1 equiv.) was dissolved in dry THF (10 mL for 0.7 mmol). Tetrakis (triphenylphosphine)palladium(0) (3 mol %) was then added to the solution and the orange mixture was stirred for 5 min at RT. Then PhSiH<sub>3</sub> (1.2 eq) was added in one portion and the solution was stirred for 30 min (HPLC monitoring). The solvent was then evaporated under vacuum and the resulting product was used without any other purification.

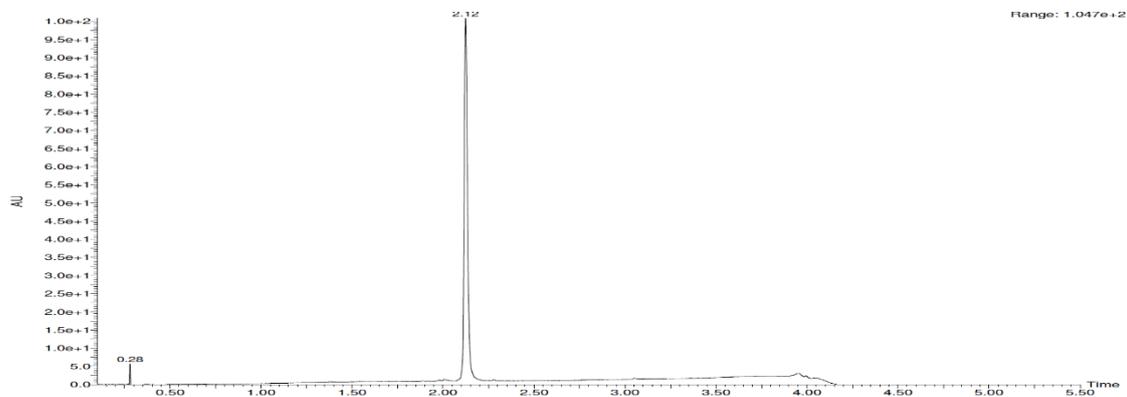
**General procedure for peptide coupling:** To a solution of acid (1.0 equiv.) and amine (1.0 equiv.) in anhydrous DCM (5 ml for 0.7 mmol) was added EDCI, HCl (1.1 equiv), HOBT (1.1 equiv.) and NMM (1.1 equiv.). The solution was stirred for 3 hours at RT to completion (HPLC monitoring). The solvent was removed under reduced pressure then the crude was purified by flash chromatography on silica gel.

### 1. Solid phase peptide synthesis:

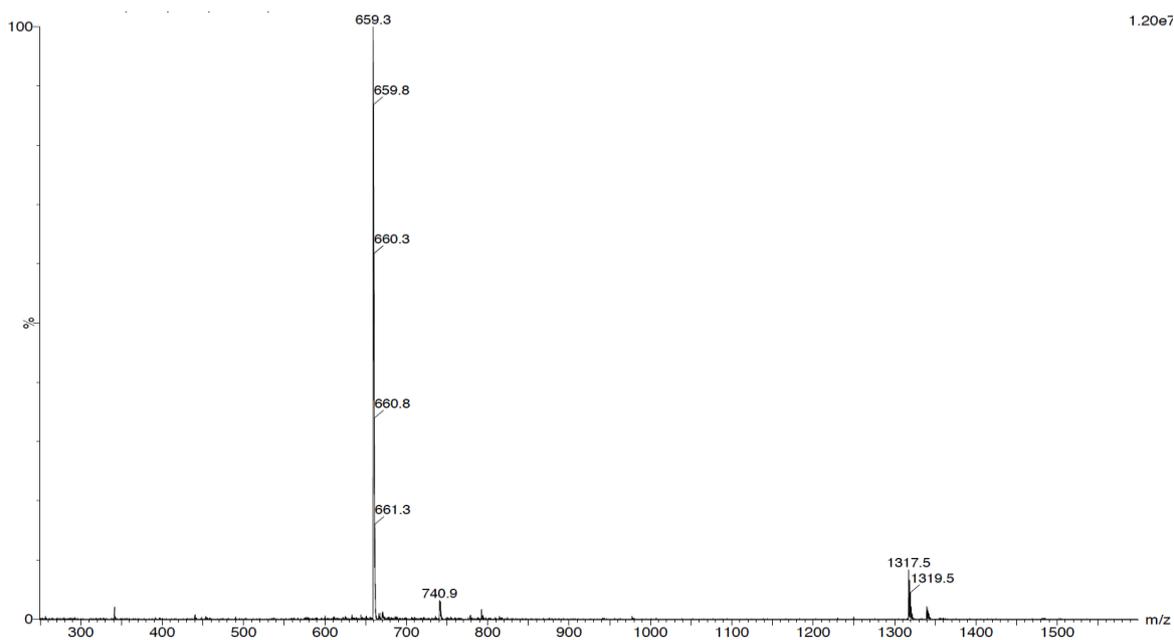
All reagents and solvents were obtained from commercial sources and used without further purification. Solid-phase synthesis was performed with a ChemMatrix® Rink Amide resin loaded at 0.49 mmol/g using Fmoc/*t*-Bu chemistry. Resin (0.5 mmol) was swollen in *N*-methylpyrrolidone (NMP) for 10 minutes and filtered. Fmoc-ATC-OH (2 equiv.), DIC (2 equiv.), Oxyma Pure (2 equiv.) and NMP (30 ml/g) were added in this order for each peptide coupling (overnight at r.t.). Resin was washed using the following procedure: 3 × DMF, 3 × MeOH, 3 × DCM. Each coupling was followed by a capping step with Ac<sub>2</sub>O/DCM (1:1 vv) during 5 min at r.t. The resin was then washed with 3 × DMF, 3 × MeOH, 3 × DCM. Fmoc deprotection at the *N*-terminus was performed using piperidine/DMF solution (2:8 vv) (3 × 10 min at r.t.) then the resin was washed with 3 × DMF, 3 × MeOH, 3 × DCM before the next coupling. Deprotection and coupling steps were monitored by Kaiser Test. After completion of the oligomerization process, the foldamer was capped using Ac<sub>2</sub>O/DCM (1:1 vv) (2 × 5 min at r.t.) and cleaved from the resin with a TFA/TIS/H<sub>2</sub>O (95:2.5:2.5 vvv) solution (2 × 90 min at r.t.). The resin was washed (1 × DCM and 1 × cleavage cocktail). The collected filtrate was evaporated under reduced pressure. The crude foldamer was precipitated with cold diethylether, centrifugated 10 min at 5°C and lyophilized. Purification was performed by preparative RP-HPLC on a Waters system controller equipped with a C<sub>18</sub> Waters Delta-Pack column (100 × 40 mm, 100 Å) at a flow of 50 mL/min; UV detection at 214 nm using a Waters 486 Tunable Absorbance Detector and a linear gradient of A = H<sub>2</sub>O (0.1 % TFA) and B = CH<sub>3</sub>CN (0.1 % TFA).

## 2. Compound ATC<sub>5</sub>-**1**-NH<sub>2</sub>

**HPLC analysis** - Chromolith® Speed Rod RP-C18 185 Pm column (50 x 4.6 mm, 5 μm) using a flow rate of 5.0 ml/min, and gradients from 100/0 to 0/100 eluents A/B over 3 min, in which eluents solvent A = H<sub>2</sub>O / TFA 0.1% and B = CH<sub>3</sub>CN / TFA 0.1%.

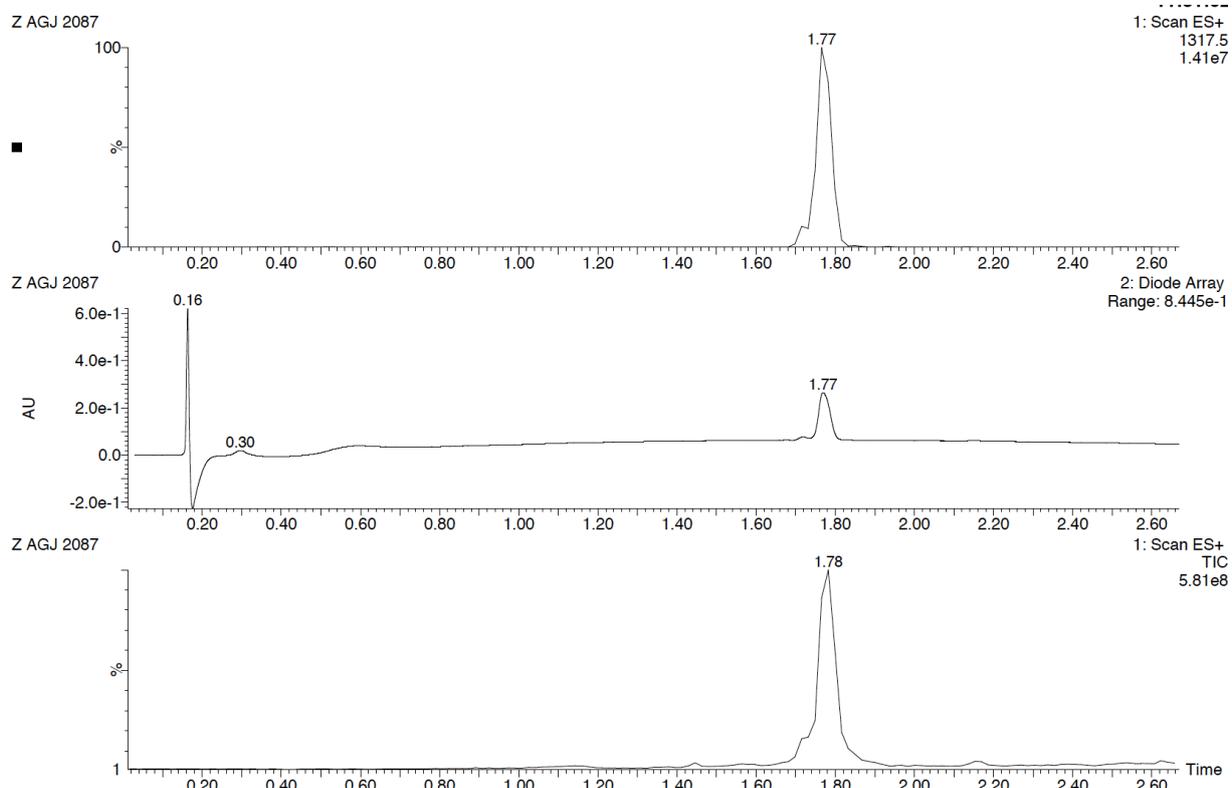


**MS spectrum (ESI):**  $m/z = 1317.5 [M+H]^+$ ,  $659.3 [M+2H]^{2+}$

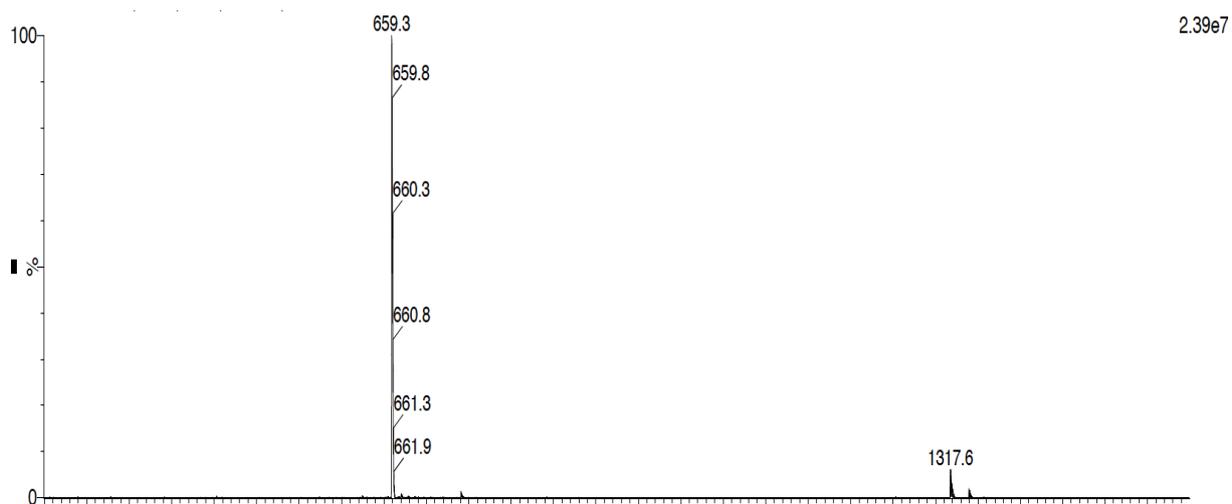


### 3. Compound ATC<sub>5</sub>-2-NH<sub>2</sub>

**LC-MS analysis** - Chromolith Speed Rod RP-C18 185 Pm column (50 x 4.6 mm, 5 μm) using a flow rate of 3.0 mL/min, and gradients of 100/0 to 0/100 eluents A/B over 2.5 min, in which eluents solvent A = H<sub>2</sub>O / HCOOH 0.1% and solvent B = CH<sub>3</sub>CN / HCOOH 0.1%.

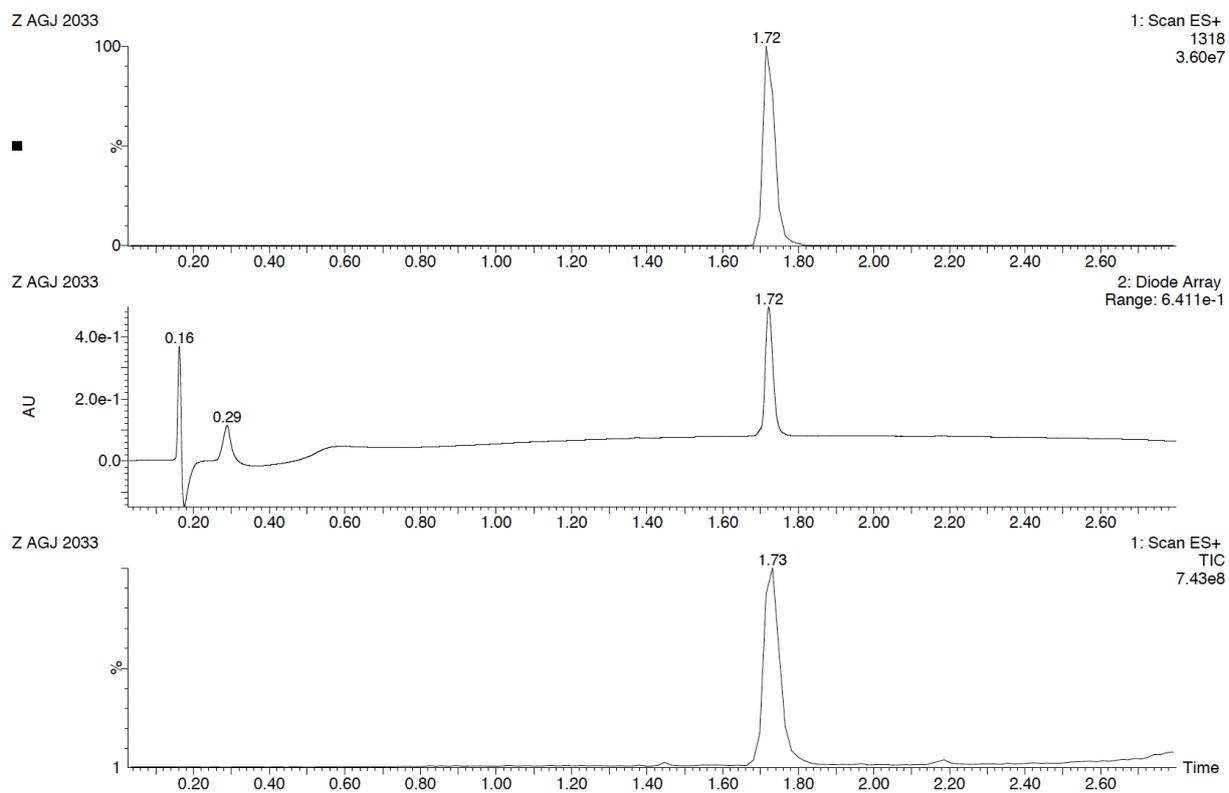


**MS spectrum (ESI):**  $m/z = 1317.5 [M+H]^+$ ,  $659.3 [M+2H]^{2+}$

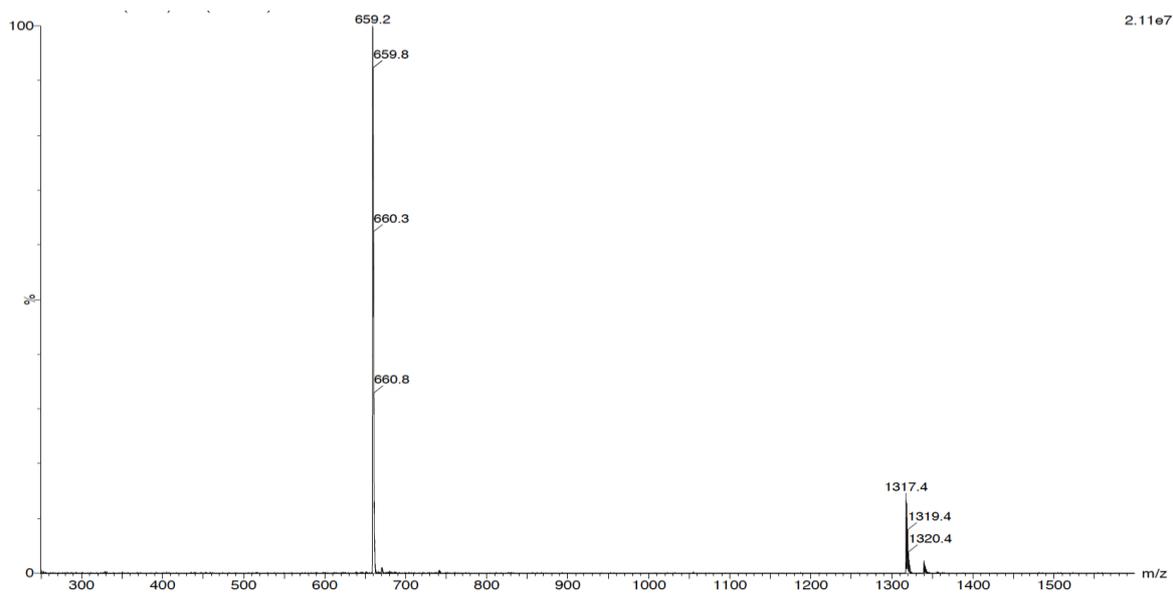


4. Compound ATC<sub>5</sub>-3-NH<sub>2</sub>

## LC-MS analysis

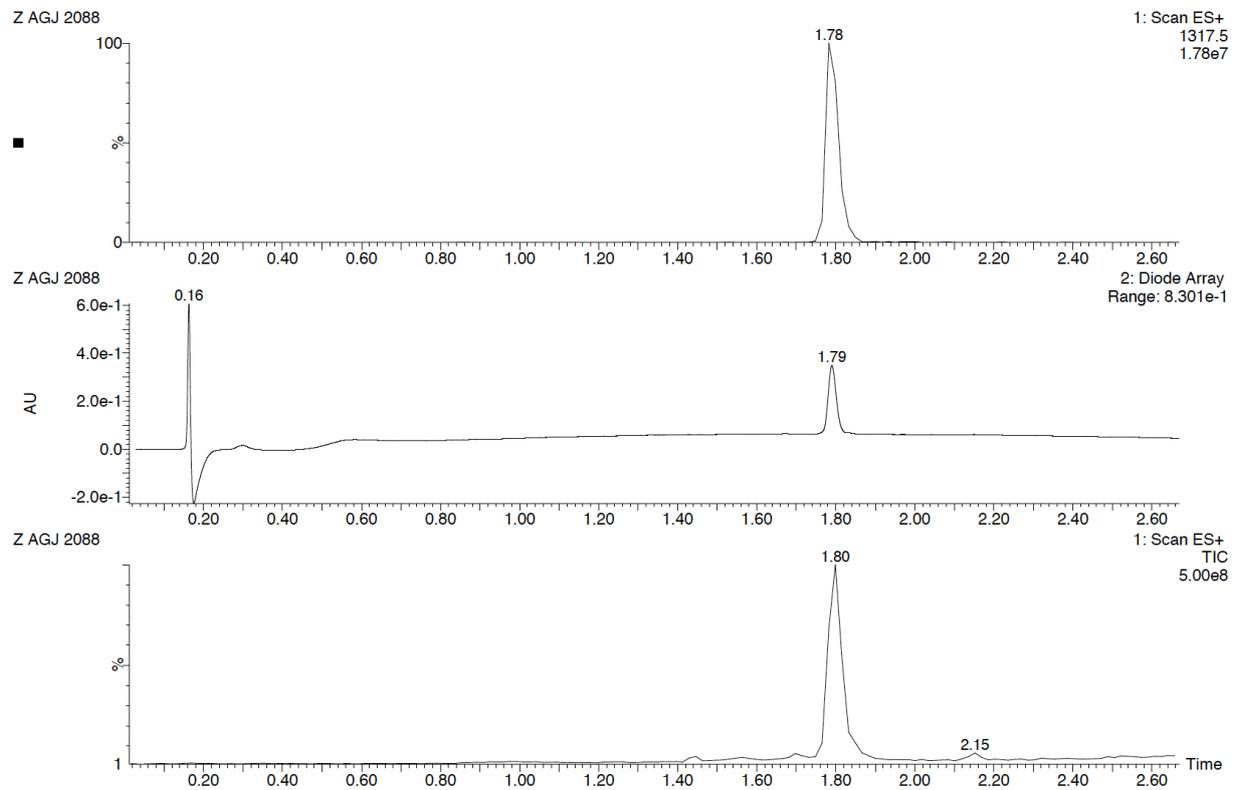
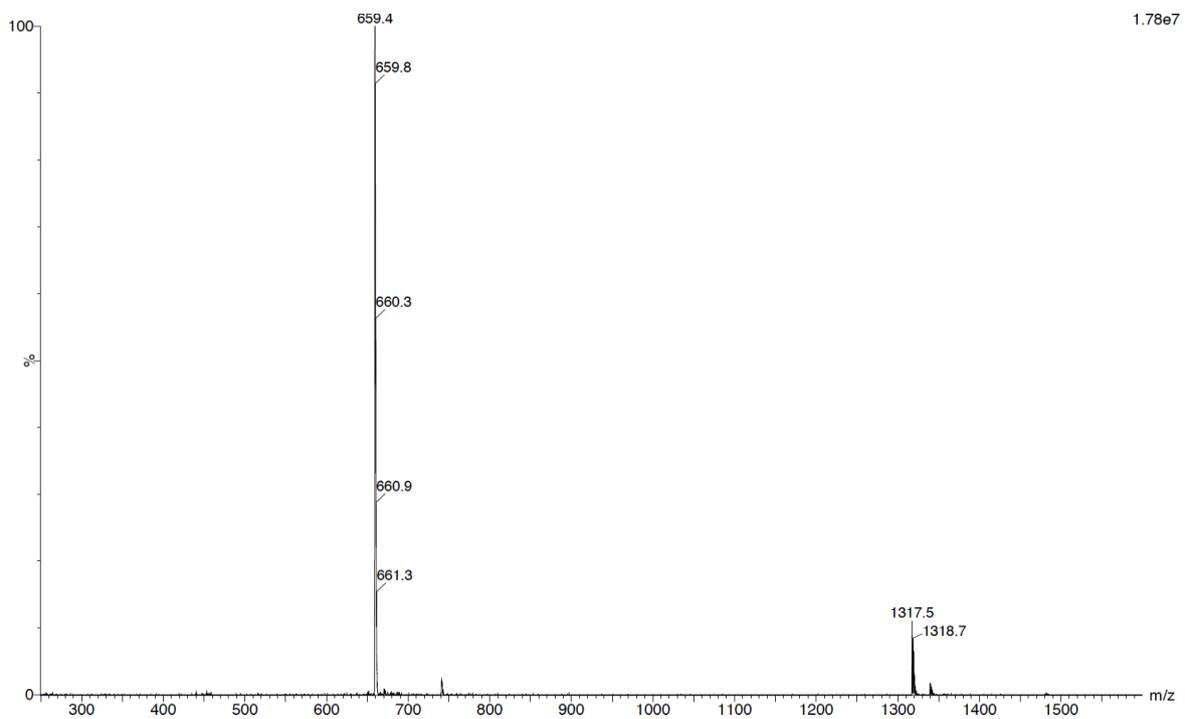


MS spectrum (ESI):  $m/z = 1317.5 [M+H]^+$ ,  $659.3 [M+2H]^{2+}$



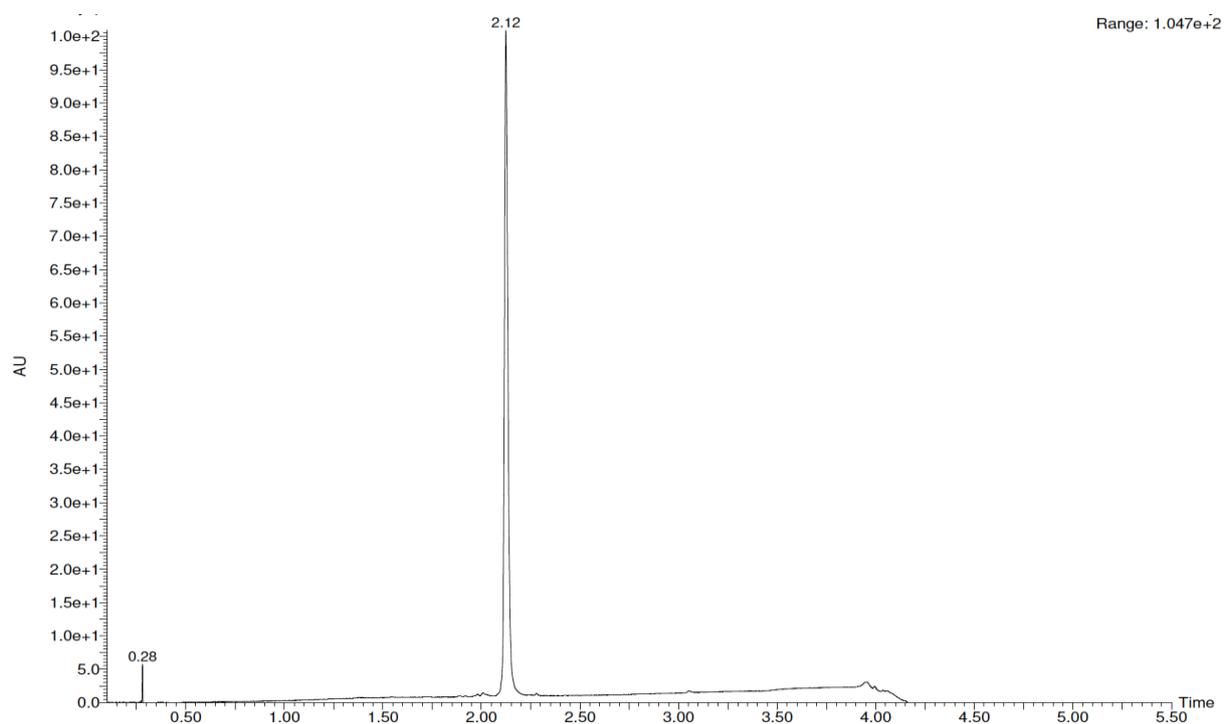
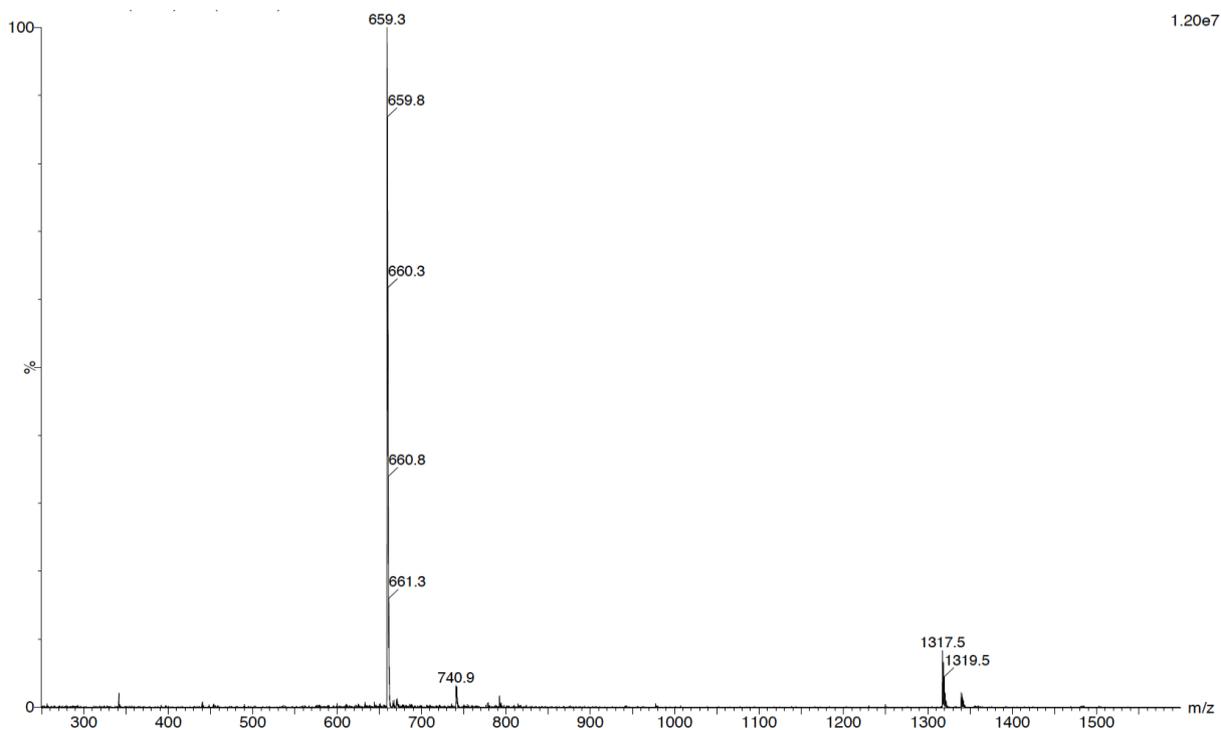
5. Compound ATC<sub>5</sub>-4-NH<sub>2</sub>

## LC-MS analysis

MS spectrum (ESI):  $m/z = 1317.5 [M+H]^+$ ,  $659.3 [M+2H]^{2+}$ 

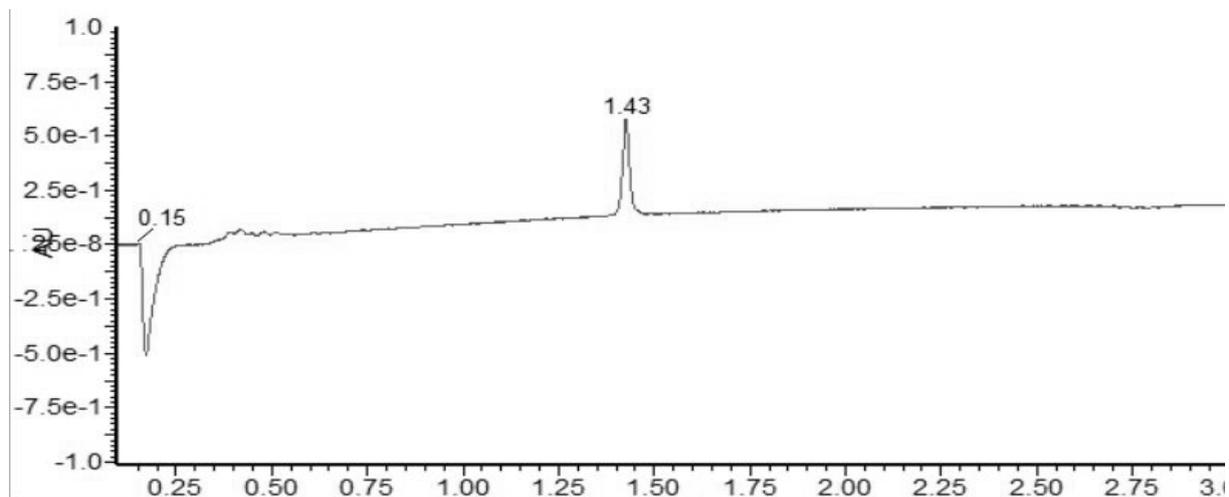
6. Compound ATC<sub>5</sub>-5-NH<sub>2</sub>

## HPLC analysis

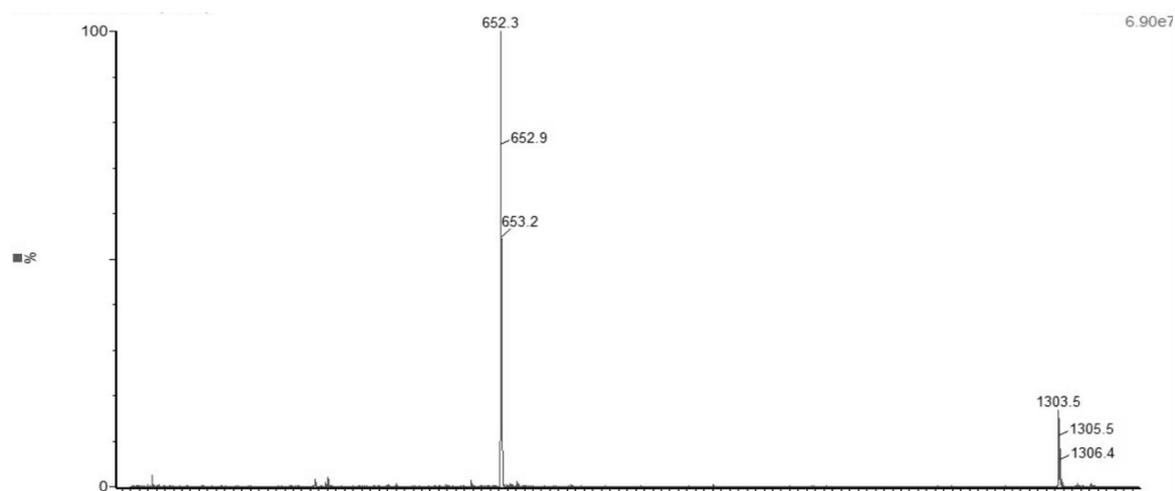
MS spectrum (ESI):  $m/z = 1317.5 [M+H]^+$ ,  $659.3 [M+2H]^{2+}$ 

7. Compound **M(1)**-ATC<sub>5</sub>-3-NH<sub>2</sub>

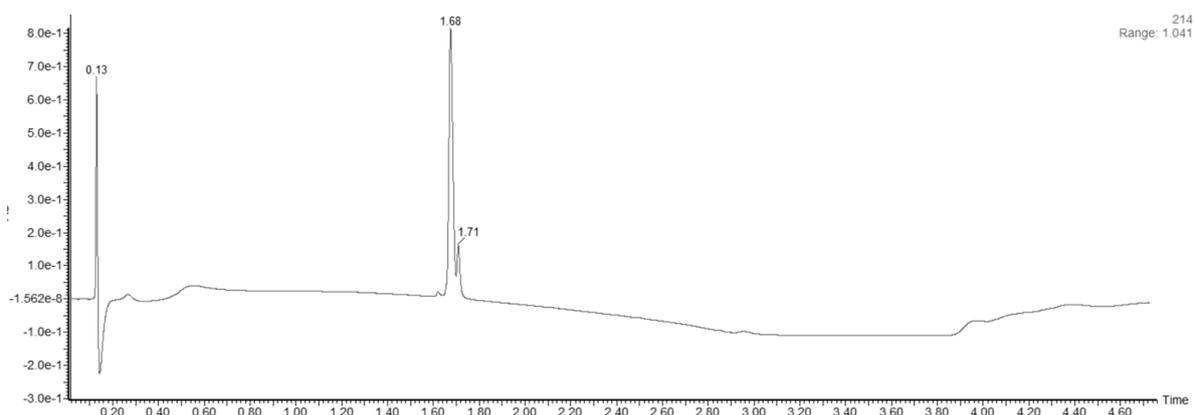
## HPLC analysis



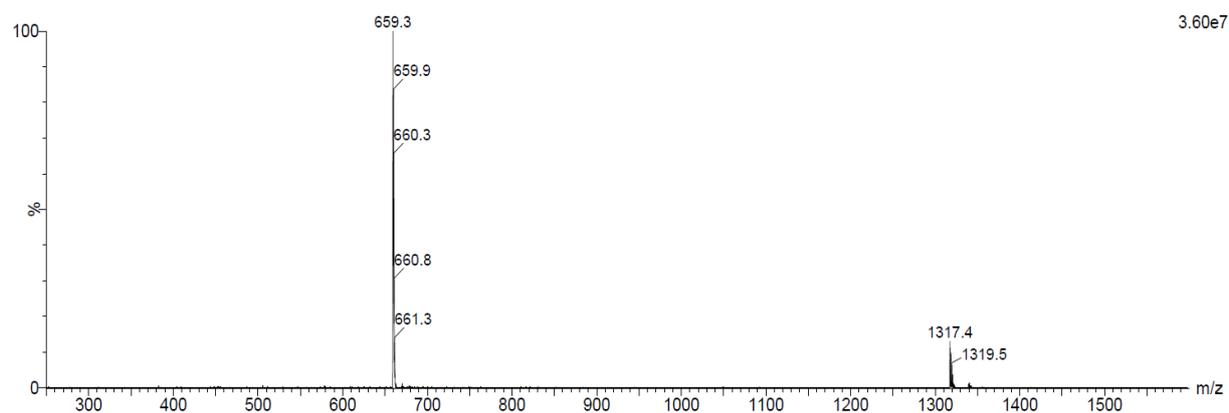
MS spectrum (ESI):  $m/z = 1303.5 [M+H]^+$ ,  $652.3 [M+2H]^{2+}$

8. Compound **M(2)**-ATC<sub>5</sub>-3-NH<sub>2</sub>

## LC-MS analysis

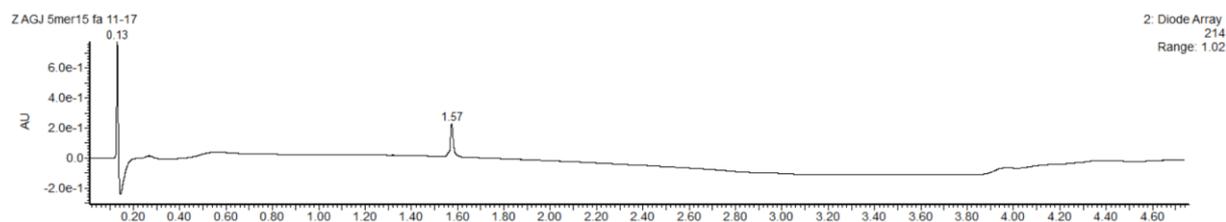


**MS spectrum (ESI):**  $m/z = 1317.4 [M+H]^+$ ,  $659.3 [M+2H]^{2+}$

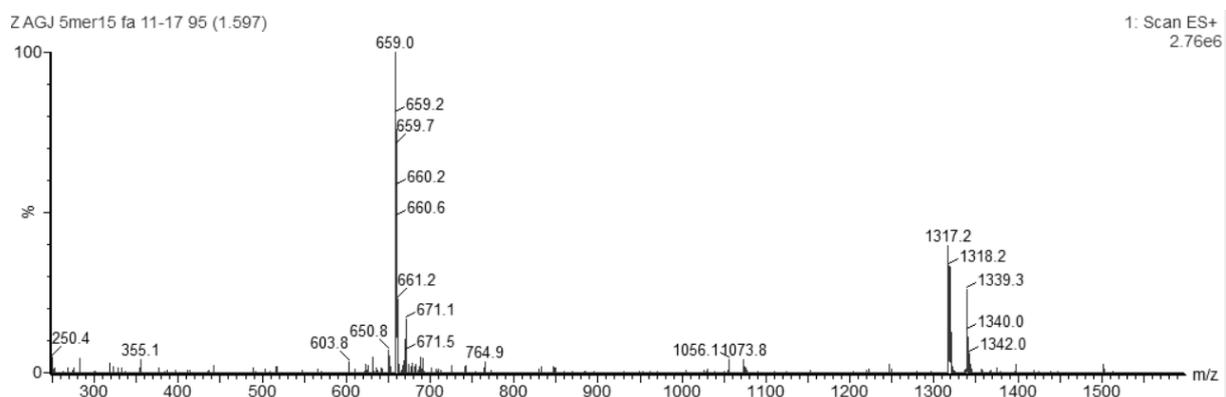


### 9. Compound **M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

#### LC-MS analysis

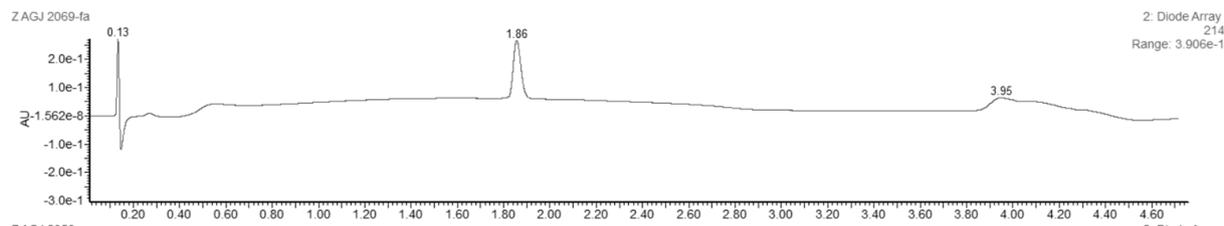


**MS spectrum (ESI):**  $m/z = 1317.4 [M+H]^+$ ,  $659.3 [M+2H]^{2+}$

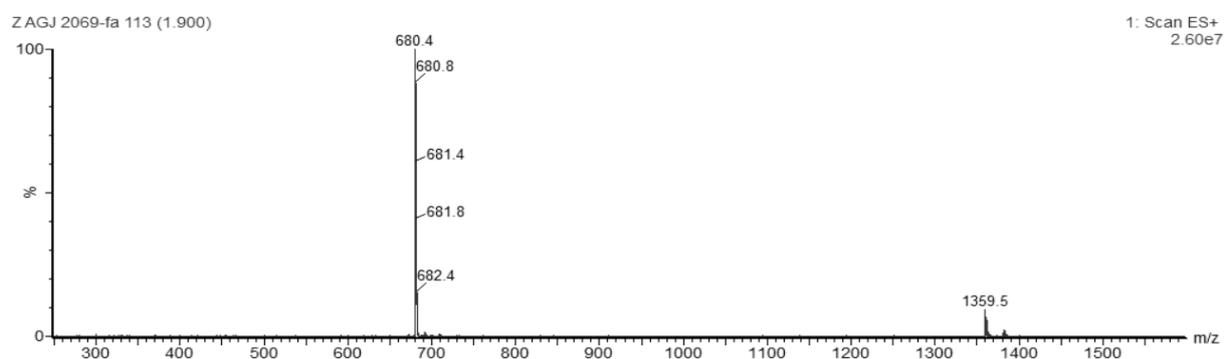


### 10. Compound **M(4)**-ATC<sub>5</sub>-3-NH<sub>2</sub>

#### LC-MS analysis

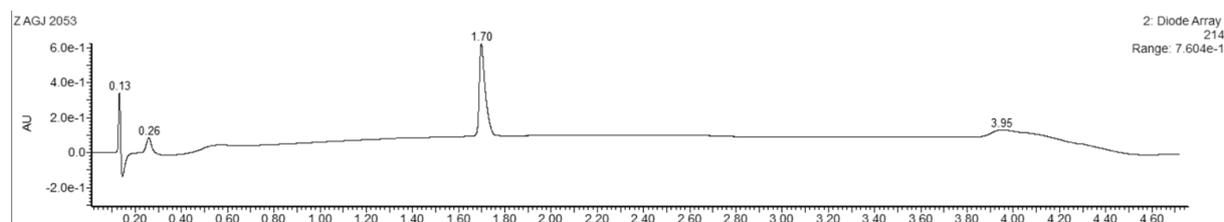


**MS spectrum (ESI):**  $m/z = 1359.5 [M+H]^+$ ,  $680.4 [M+2H]^{2+}$

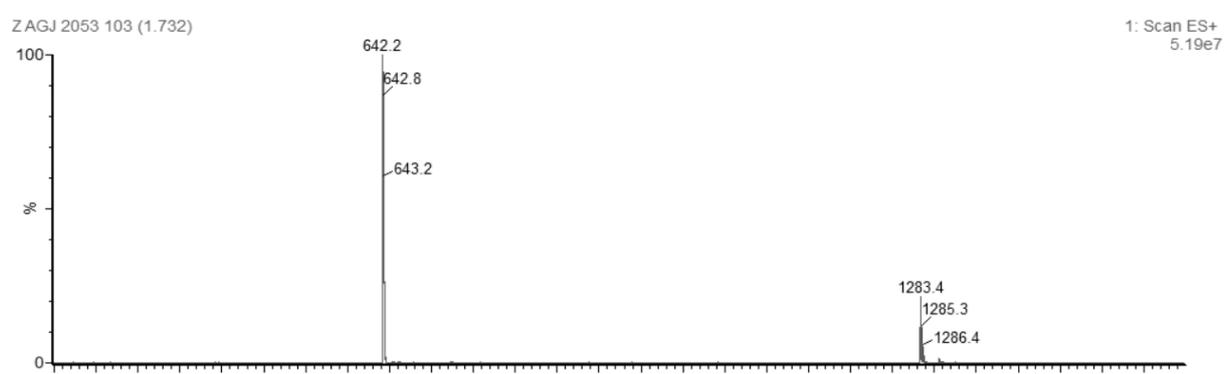


### 11. Compound **M(5)**-ATC<sub>5</sub>-3-NH<sub>2</sub>

#### LC-MS analysis

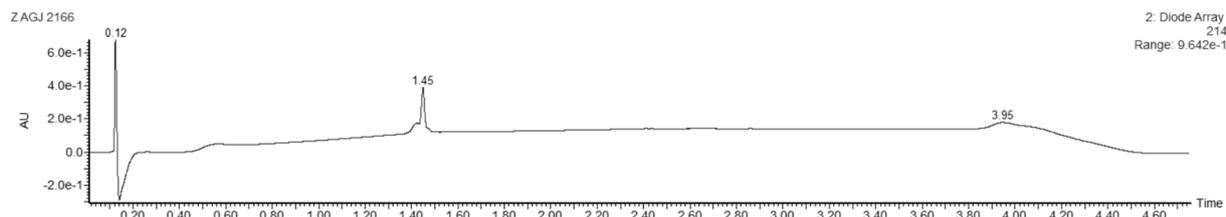


**MS spectrum (ESI):**  $m/z = 1283.4 [M+H]^+$ ,  $642.2 [M+2H]^{2+}$

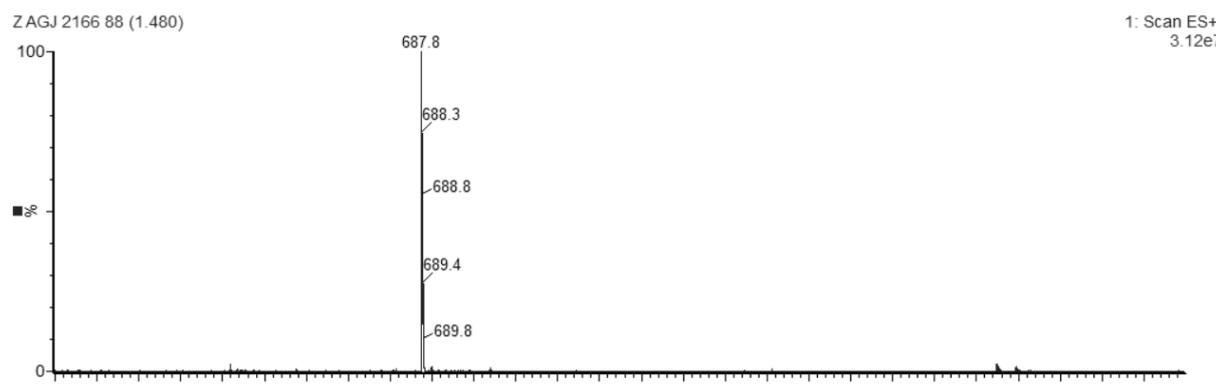


## 12. Compound **M(6)**-ATC<sub>5</sub>-3-NH<sub>2</sub>

### LC-MS analysis

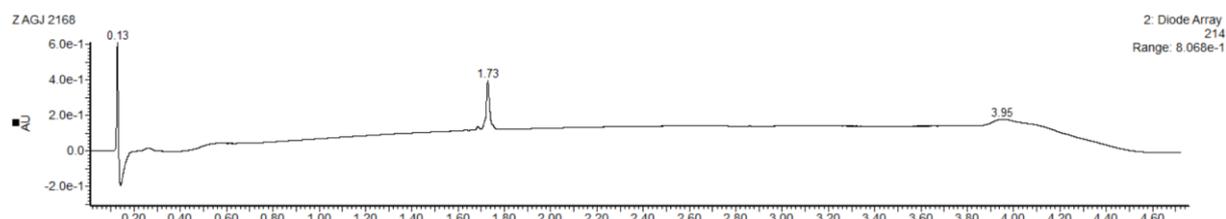


**MS spectrum (ESI):**  $m/z = 1374.2 [M+H]^+$ ,  $687.8 [M+2H]^{2+}$

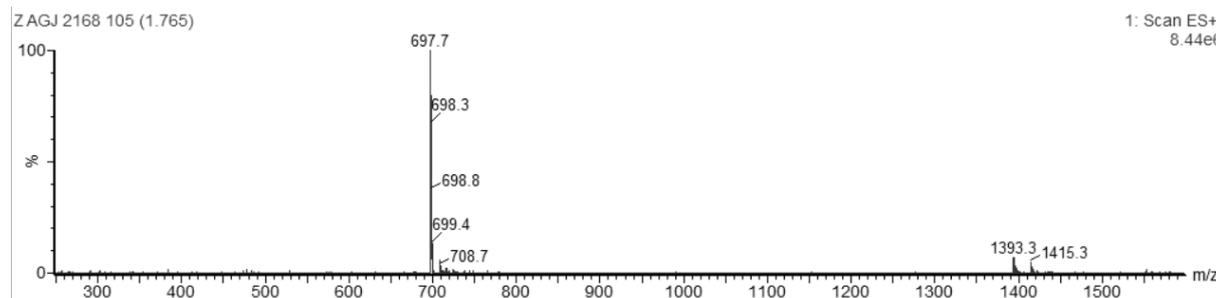


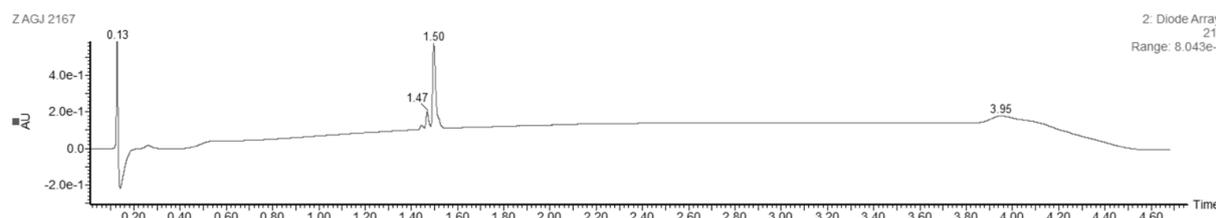
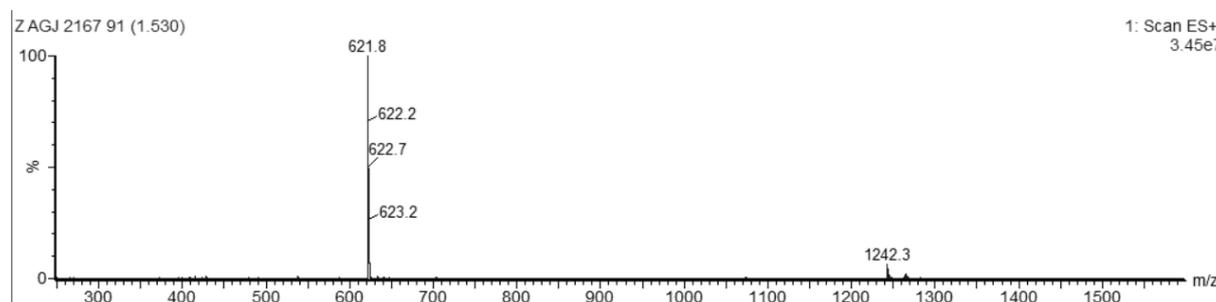
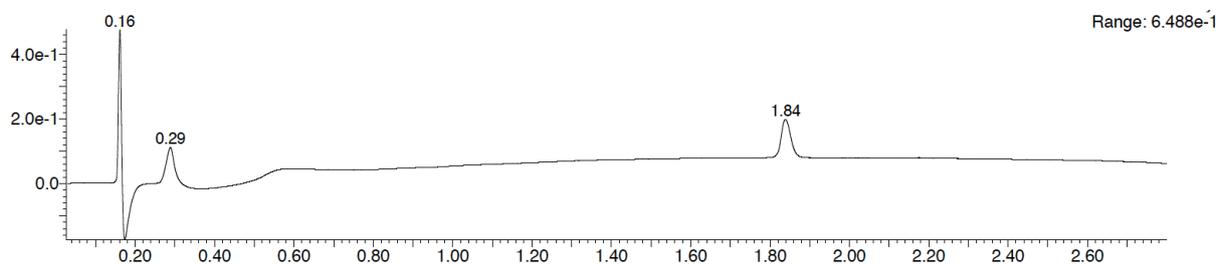
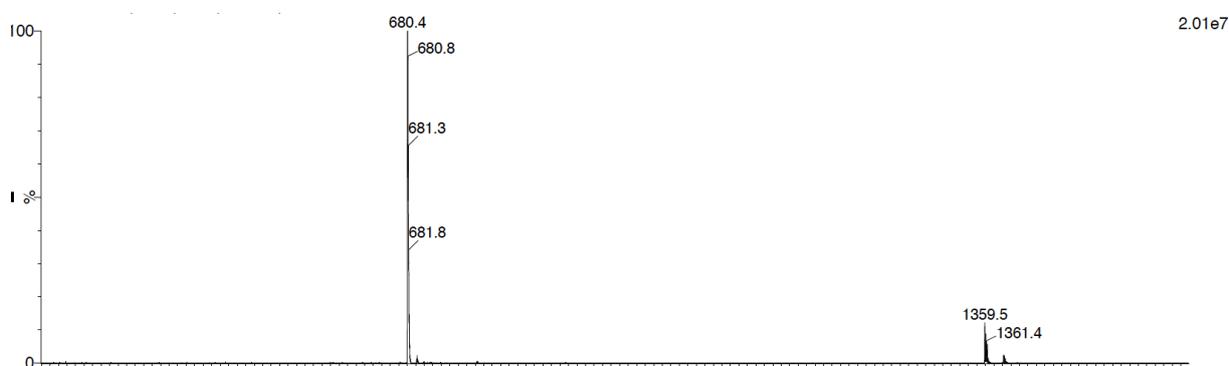
## 13. Compound **M(7)**-ATC<sub>5</sub>-3-NH<sub>2</sub>

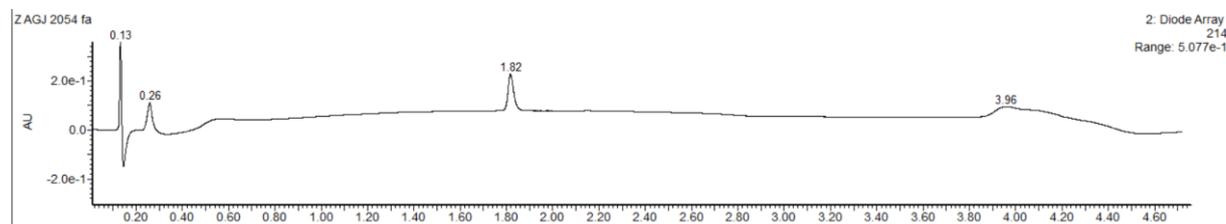
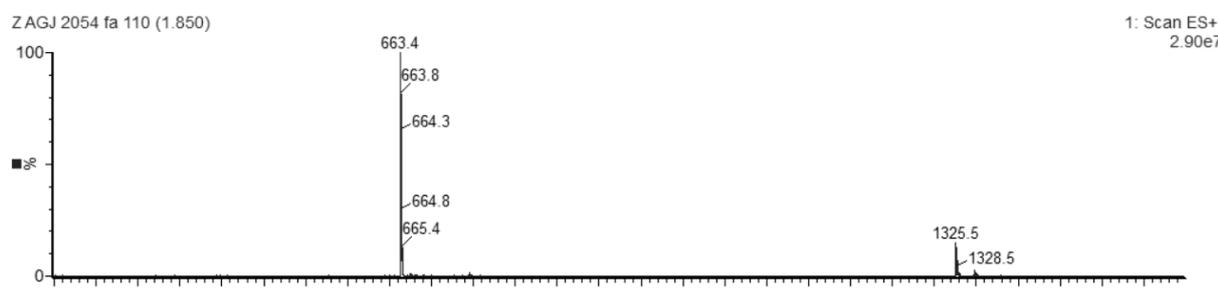
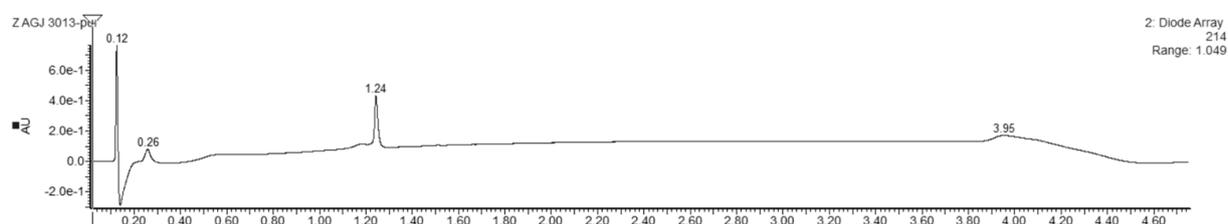
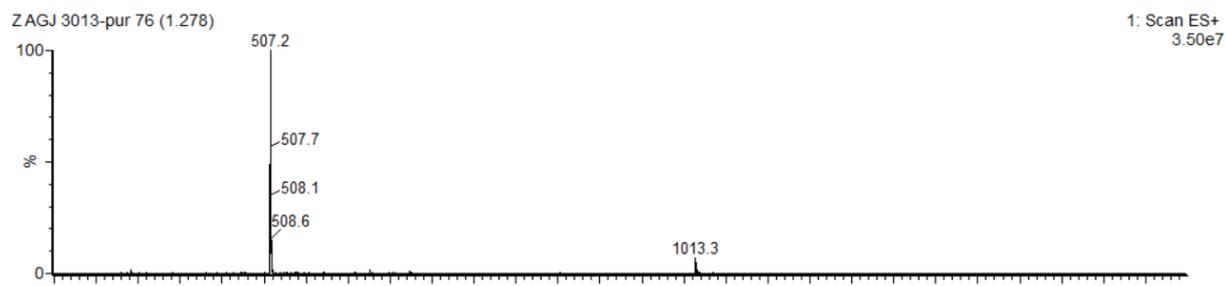
### LC-MS analysis

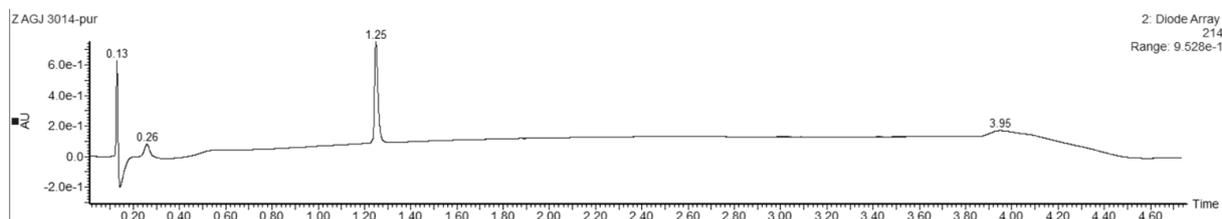
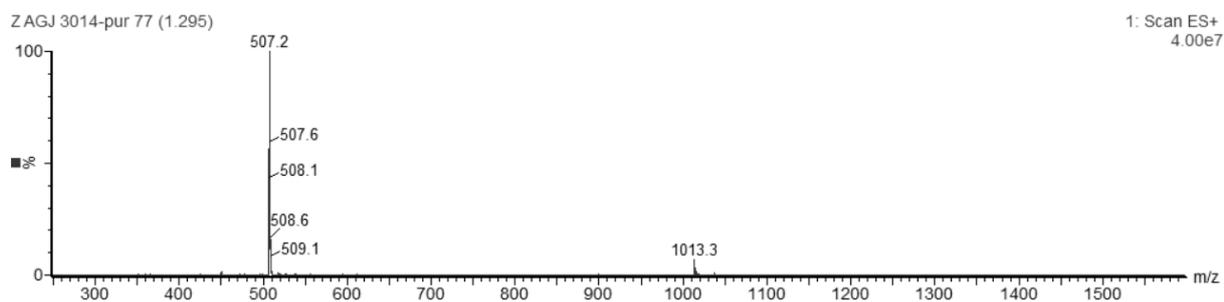
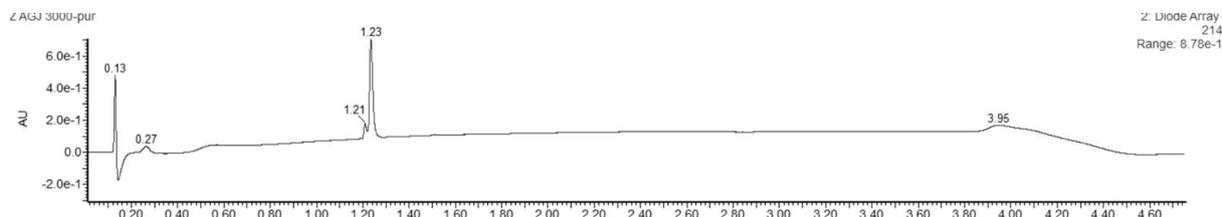
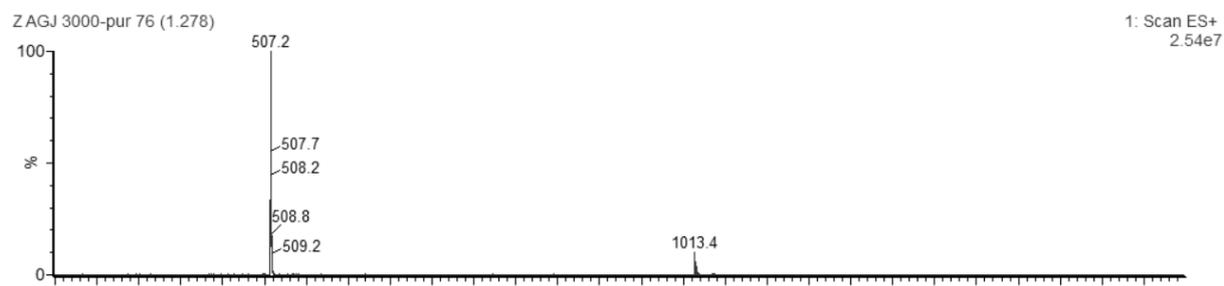


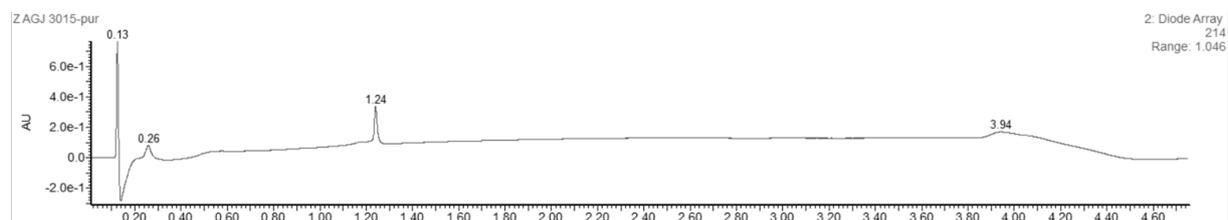
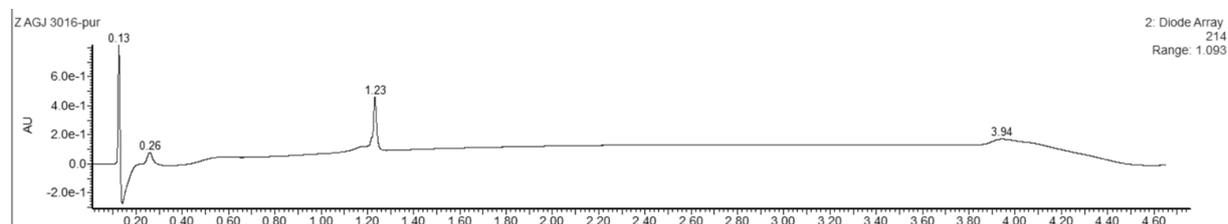
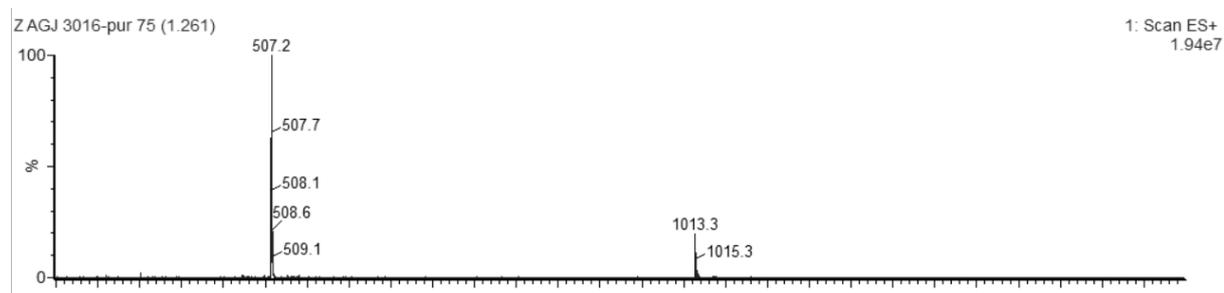
**MS spectrum (ESI):**  $m/z = 1393.3 [M+H]^+$ ,  $697.7 [M+2H]^{2+}$



**14. Compound M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1242.3 [M+H]^+$ ,  $621.8 [M+2H]^{2+}$ **15. Compound M(9)-ATC<sub>5</sub>-3-NH<sub>2</sub>****HPLC analysis****MS spectrum (ESI):**  $m/z = 1359.5 [M+H]^+$ ,  $680.4 [M+2H]^{2+}$ 

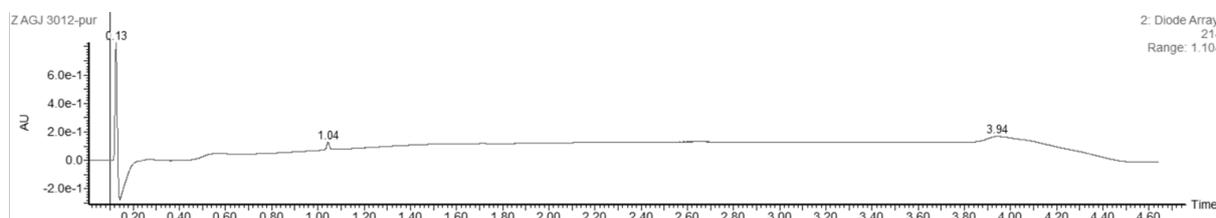
**16. Compound *M(10)*-ATC<sub>5</sub>-3-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1325.5 [M+H]^+$ ,  $663.4 [M+2H]^{2+}$ **17. Compound *M(11)*-ATC<sub>5</sub>-1-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1013.3 [M+H]^+$ ,  $507.2 [M+2H]^{2+}$ 

**18. Compound M(12)-ATC<sub>5</sub>-2-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1013.3 [M+H]^+$ ,  $507.2 [M+2H]^{2+}$ **19. Compound M(13)-ATC<sub>5</sub>-3-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1013.3 [M+H]^+$ ,  $507.2 [M+2H]^{2+}$ 

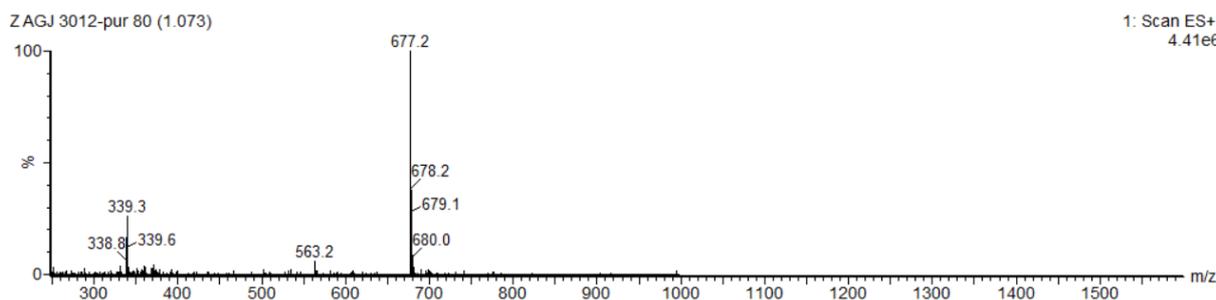
**20. Compound M(14)-ATC<sub>5</sub>-4-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1013.3 [M+H]^+$ ,  $507.2 [M+2H]^{2+}$ **21. Compound M(15)-ATC<sub>5</sub>-5-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1013.3 [M+H]^+$ ,  $507.2 [M+2H]^{2+}$ 

## 22. Compound **M(16)**-ATC<sub>3</sub>-2-NH<sub>2</sub>

### LC-MS analysis

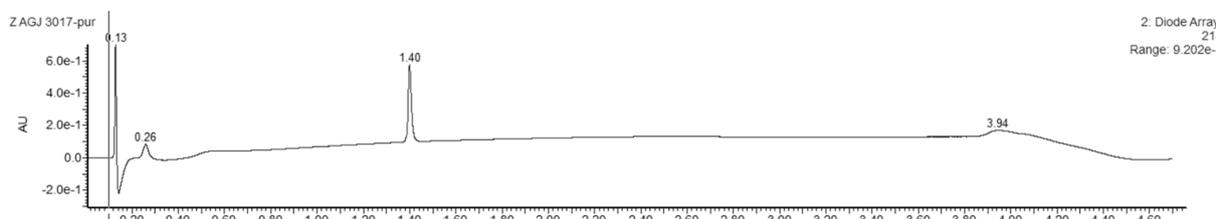


**MS spectrum (ESI):**  $m/z = 677.2 [M+H]^+$ ,  $339.3 [M+2H]^{2+}$



## 23. Compound **M(17)**-ATC<sub>7</sub>-4-NH<sub>2</sub>

### LC-MS analysis

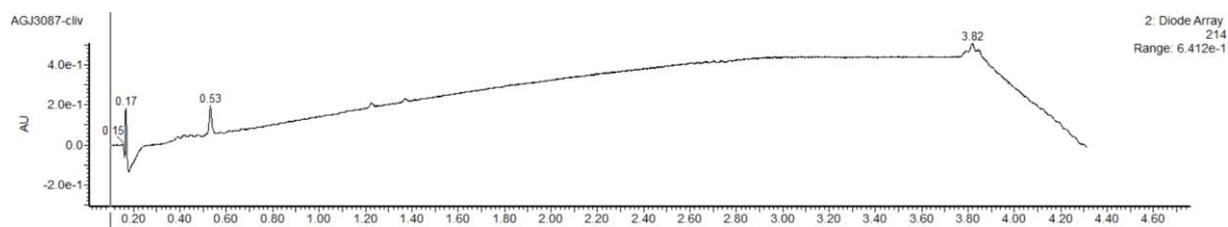


**MS spectrum (ESI):**  $m/z = 1349.4 [M+H]^+$ ,  $675.2 [M+2H]^{2+}$

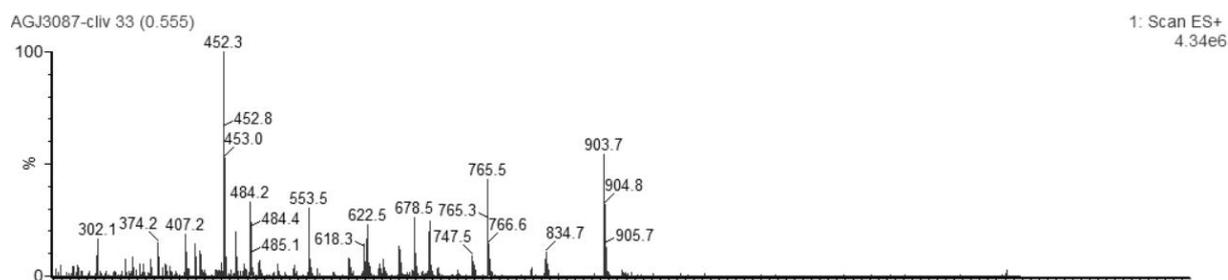


## 24. Compound *M(18)*-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub>

### LC-MS analysis

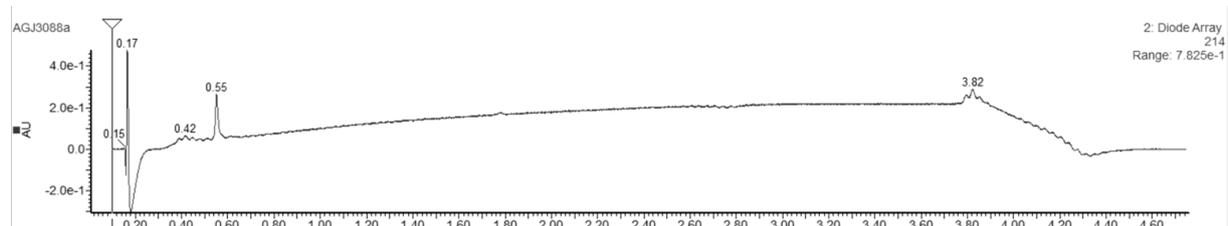


**MS spectrum (ESI):**  $m/z = 903.7 [M+H]^+$ ,  $452.3 [M+2H]^{2+}$

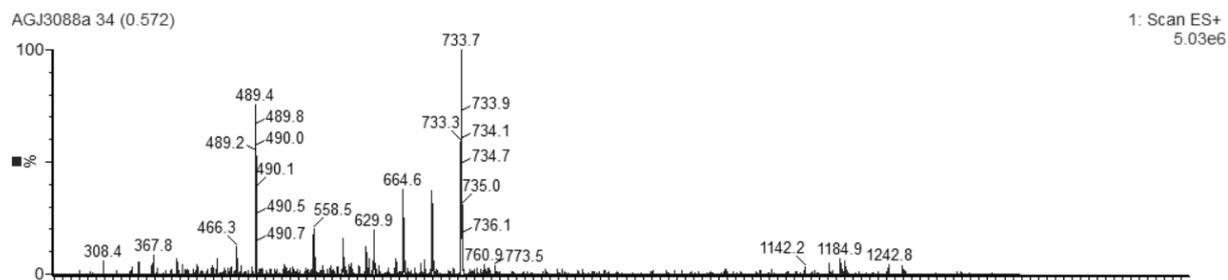


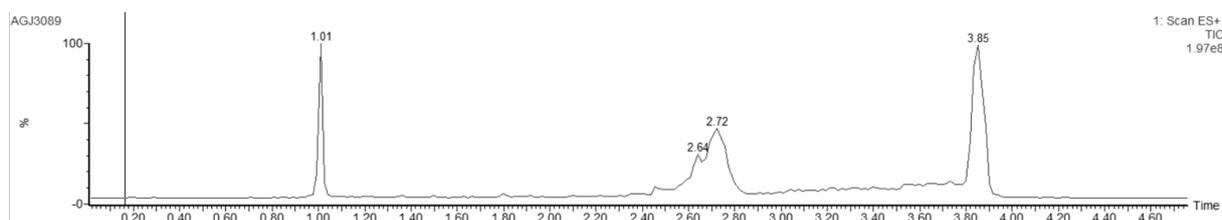
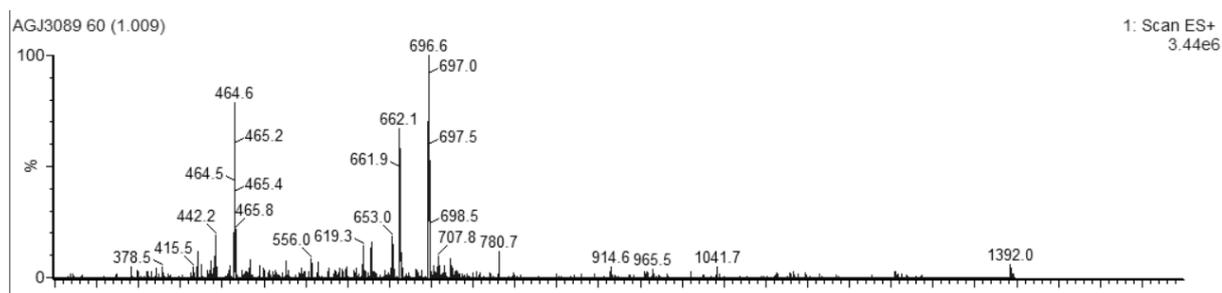
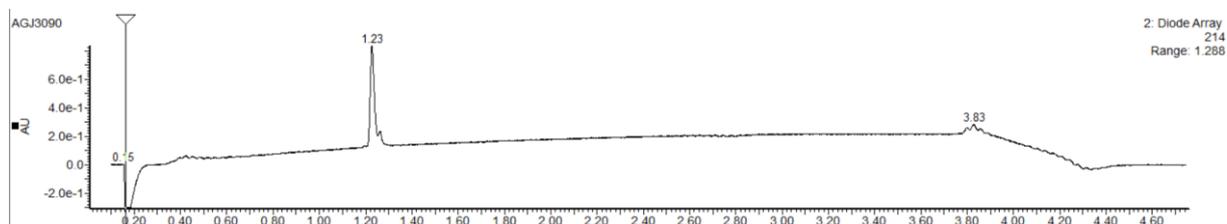
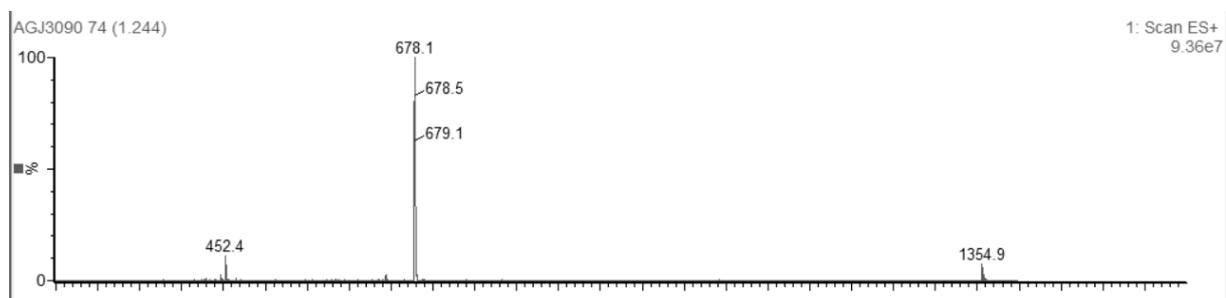
## 25. Compound *M(19)*-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub>

### LC-MS analysis



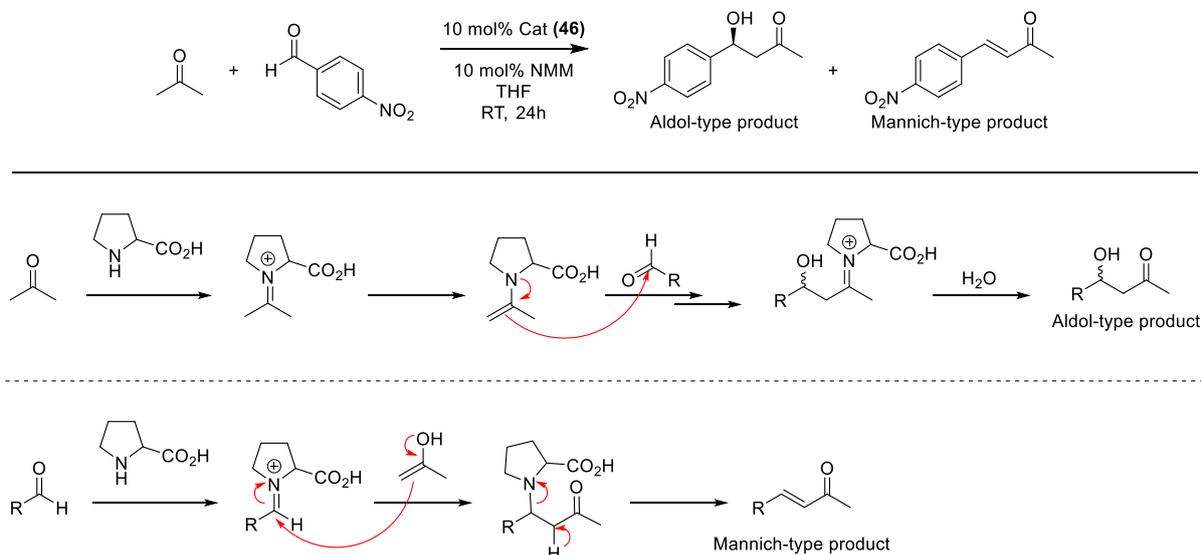
**MS spectrum (ESI):**  $m/z = 733.7 [M+2H]^{2+}$



**26. Compound M(20)-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1392.0 [M+H]^+$ ,  $696.6 [M+2H]^{2+}$ **27. Compound M(21)-ATC<sub>5</sub>-2,4-NH<sub>2</sub>****LC-MS analysis****MS spectrum (ESI):**  $m/z = 1354.9 [M+H]^+$ ,  $678.1 [M+2H]^{2+}$ 

## IV. EXPLORED MODEL REACTIONS USING (46) AS CATALYST

### 1. Cross condensation of acetone with *p*-nitrobenzaldehyde



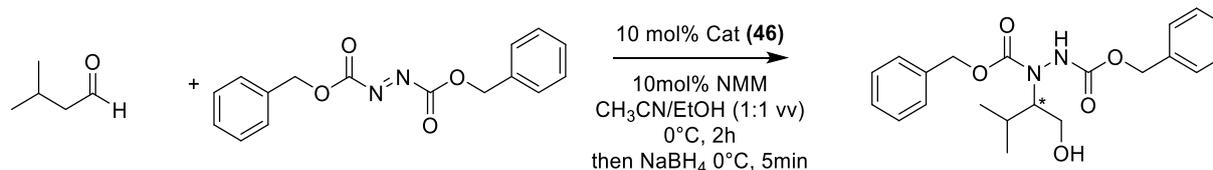
Scheme S6: Cross condensation of acetone with *p*-nitrobenzaldehyde: mechanisms for the obtention of the aldol-type and Mannich-type products

The catalyst (46) (10 mol%) was dissolved in 1mL of a solution of Acetone/THF (1:4). The *p*-nitrobenzaldehyde (15.1 mg 0.1 mmol) was then added and the reaction mixture was stirred at RT for 24 hours.

Monitoring was done by diluting 20 $\mu$ L of the reaction mixture in 400 $\mu$ L EtOH. 10 $\mu$ L of this solution were injected for the HPLC analysis performed C18 ONYX column using a flow rate of 3.0 mL/min, and gradients of 100/0 to 70/30 eluents A/B over 5 min, in which eluents solvent A = H<sub>2</sub>O / TFA 0.1% and solvent B = CH<sub>3</sub>CN / TFA 0.1%.

Estimated yield by HPLC at 214nm = 42%; Aldol-type product: *tr*=2.96 min 27%, Mannich-type product: *tr*=4.54 min 63%.

## 2. Amination of isovaleraldehyde with dibenzylazodicarboxylate



Scheme S7: Amination of isovaleraldehyde with dibenzylazodicarboxylate

Dibenzyl azodicarboxylate (90%, 330 mg, 1 mmol, 1eq) and cat (**46**) (53.62 mg, 0.1mmol, 10 mol%) in CH<sub>3</sub>CN (10 mL) were treated with an aldehyde (1.5 mmol, 1.5 eq) at 0°C. The mixture was stirred at RT for 3.5 days. After the mixture became colorless it was cooled to 0°C, treated with ethanol (10 mL) and NaBH<sub>4</sub> (40 mg) and was stirred for 5 min at 0°C. The reaction was worked up in the usual manner with half-concentrated aqueous ammonium chloride solution and ethyl acetate. The organic layers were dried (MgSO<sub>4</sub>), filtered, and concentrated. Yield 94%.

Monitoring was done by diluting 5μL of the reaction mixture in 400μL MeOH. 5 μL of this solution were injected for the HPLC analysis performed with Chromolith Speed Rod RP-C18 185 Pm column (50 x 4.6 mm, 5 μm) using a flow rate of 3.0 mL/min, and gradients of 100/0 to 0/100 eluents A/B over 3 min, in which eluents solvent A = H<sub>2</sub>O / TFA 0.1% and solvent B = CH<sub>3</sub>CN / TFA 0.1%.

Enantiomeric excess was determined by Chiral HPLC analysis using Chiralpak® OD-H column (250 x 4.6 mm) with *n*-hexane/isopropanol (90/10) as eluent at a flow rate of 1 ml/min.

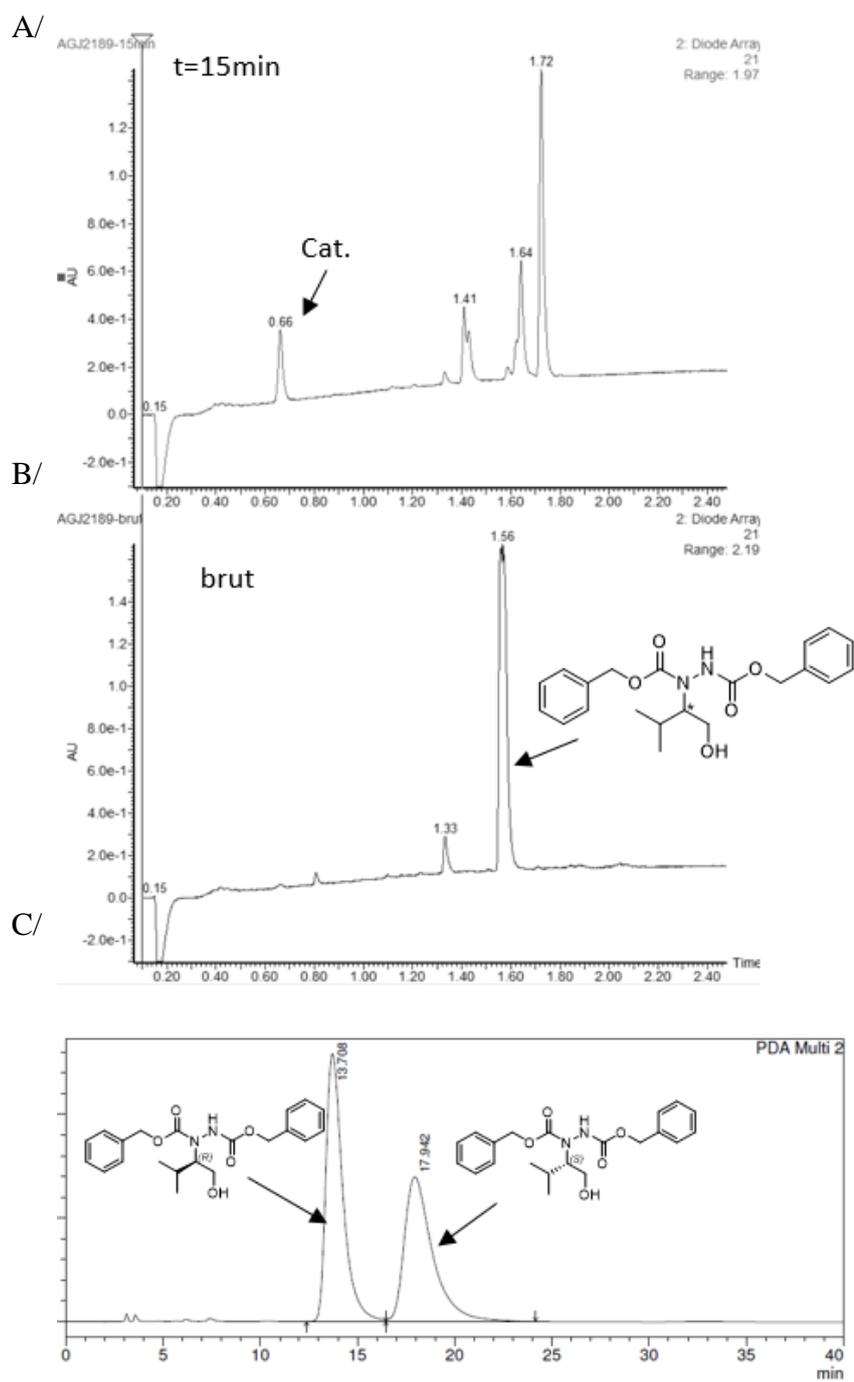
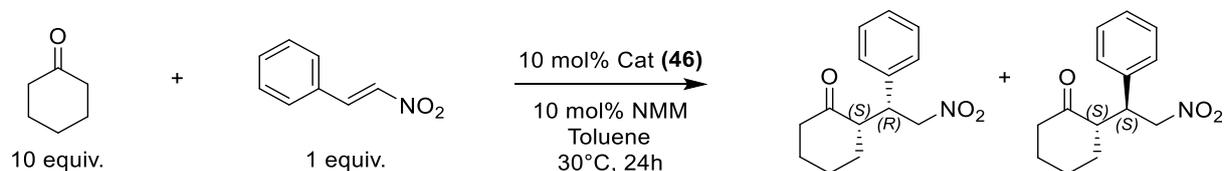


Figure S2: HPLC profiles of the amination of isovaleraldehyde with dibenzylazodicarboxylate. Reaction performed over 3.5 days. A/HPLC after 15 min, B/HPLC of the brut product, C/enantiomeric excess as determined by chiral HPLC

### 3. Nitro-Michael addition reaction of cyclohexanone to $\beta$ -trans-nitrostyrene



Scheme S8: Michael addition reaction of cyclohexanone to  $\beta$  trans-nitrostyrene

The catalyst **(46)** (0.01 mmol, 6.49 mg, 0.1 equiv.) was dissolved in 400  $\mu$ L of a solution of NMM/toluene (11  $\mu$ L in 4mL). The nitrostyrene (14.9 mg, 0.1 mmol, 1.0 equiv.) and the cyclohexanone (103.5  $\mu$ L, 1.0 mmol, 10.0 equiv.) were then added and the reaction mixture was stirred at 30°C for 24 hours. Monitoring was done by diluting 5  $\mu$ L of the reaction mixture in 400  $\mu$ L MeOH. 5  $\mu$ L of this solution were injected for the HPLC analysis performed with Chromolith Speed Rod RP-C18 185 Pm column (50 x 4.6 mm, 5  $\mu$ m) using a flow rate of 3.0 mL/min, and gradients of 100/0 to 0/100 eluents A/B over 3 min, in which eluents solvent A = H<sub>2</sub>O / TFA 0.1% and solvent B = CH<sub>3</sub>CN / TFA 0.1%. Enantiomeric excesses were determined by Chiral HPLC analysis using Chiralpak® IC column (250 x 4.6 mm) with *n*-hexane/isopropanol (80/20) as eluent at a flow rate of 1 ml/min.

Estimated yield by HPLC at 214nm: 54%, ee=27%

## V. HPLC PROFILES OF THE NITRO-MICHAEL ADDITION REACTIONS CATALYSED BY **ATC<sub>5</sub>-3-NH<sub>2</sub>**

The catalyst **ATC<sub>5</sub>-3-NH<sub>2</sub>** (0.0037 mmol, 4.9 mg, 0.1 equiv.) was dissolved in 150  $\mu$ L of a solution of NMM/iPrOH (11  $\mu$ L in 4 mL). The nitrostyrene (0.037 mmol, 1.0 equiv.) and the cyclohexanone (38.5  $\mu$ L, 0.37 mmol, 10.0 equiv.) were then added and the reaction mixture was stirred at 4 °C for 16 hours.

Monitoring was done by diluting 5  $\mu$ L of the reaction mixture in 400  $\mu$ L MeOH. 5  $\mu$ L of this solution were injected for the HPLC analysis performed with Chromolith Speed Rod RP-C18 185 Pm column (50 x 4.6 mm, 5  $\mu$ m) using a flow rate of 3.0 mL/min, and gradients of 100/0 to 0/100 eluents A/B over 3 min, in which eluents solvent A = H<sub>2</sub>O / TFA 0.1% and solvent B = CH<sub>3</sub>CN / TFA 0.1%.

Enantiomeric excesses were determined by Chiral HPLC analysis using method A: Chiralpak® IC column (250 x 4.6 mm) with *n*-hexane/isopropanol (80/20) as eluent at a flow rate of 1 ml/min or method B: Chiralpak® IA column (250 x 4.6 mm) with *n*-hexane/isopropanol (90/10) as eluent at a flow rate of 1 ml/min.

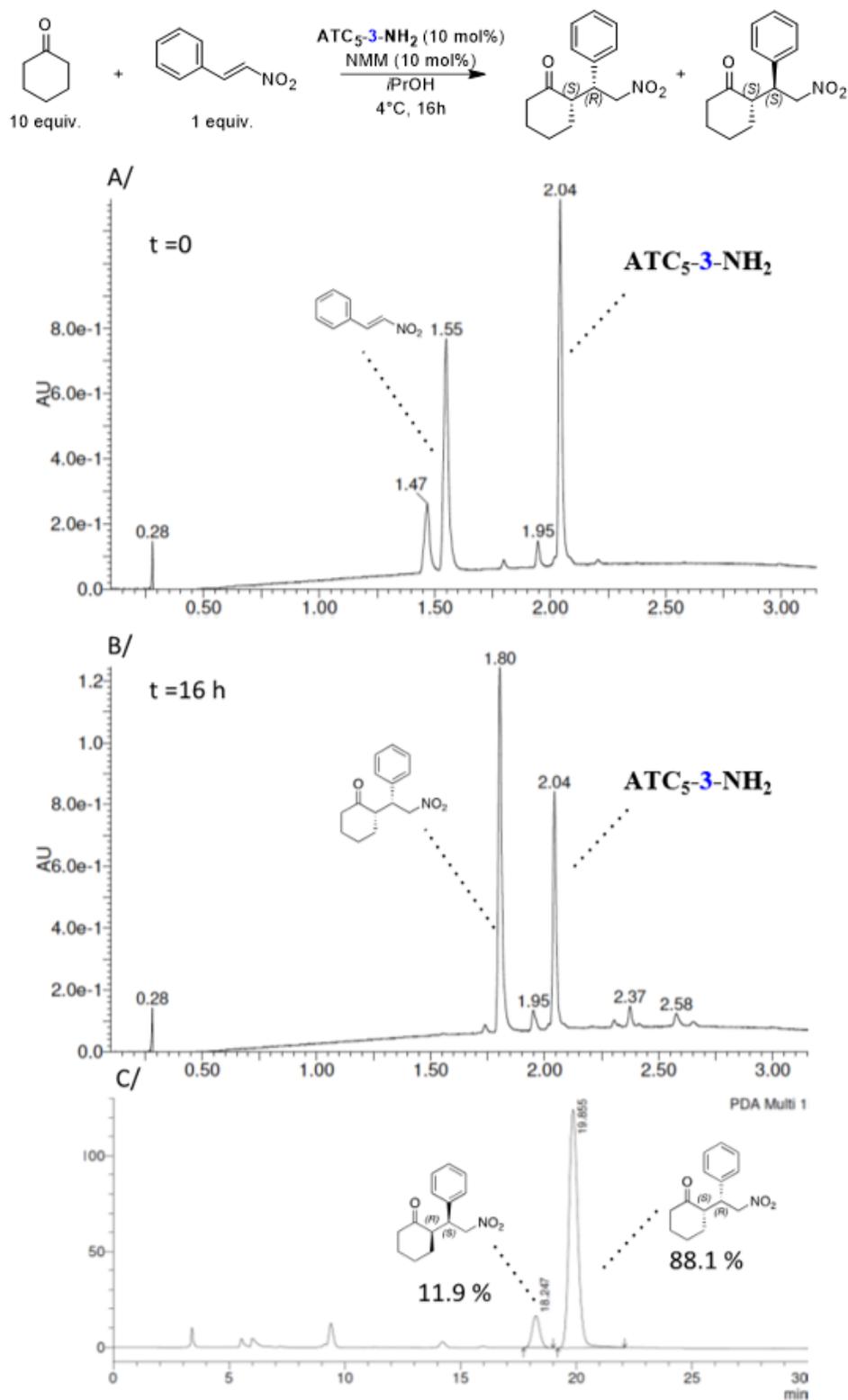


Figure S3: HPLC profiles of the nitro-Michael addition reaction of cyclohexanone to  $\beta$ -trans-nitrostyrene. Reaction performed over 16h with 0.037 mmol nitroolefine, 0.37 mmol cyclohexanone, 0.0037 mmol NMM and 0.0037 mmol catalyst in 150  $\mu$ l solvent. A/, B/monitoring of the nitro-Michael reaction over 16 hours. C/enantiomeric excess as determined by method A.

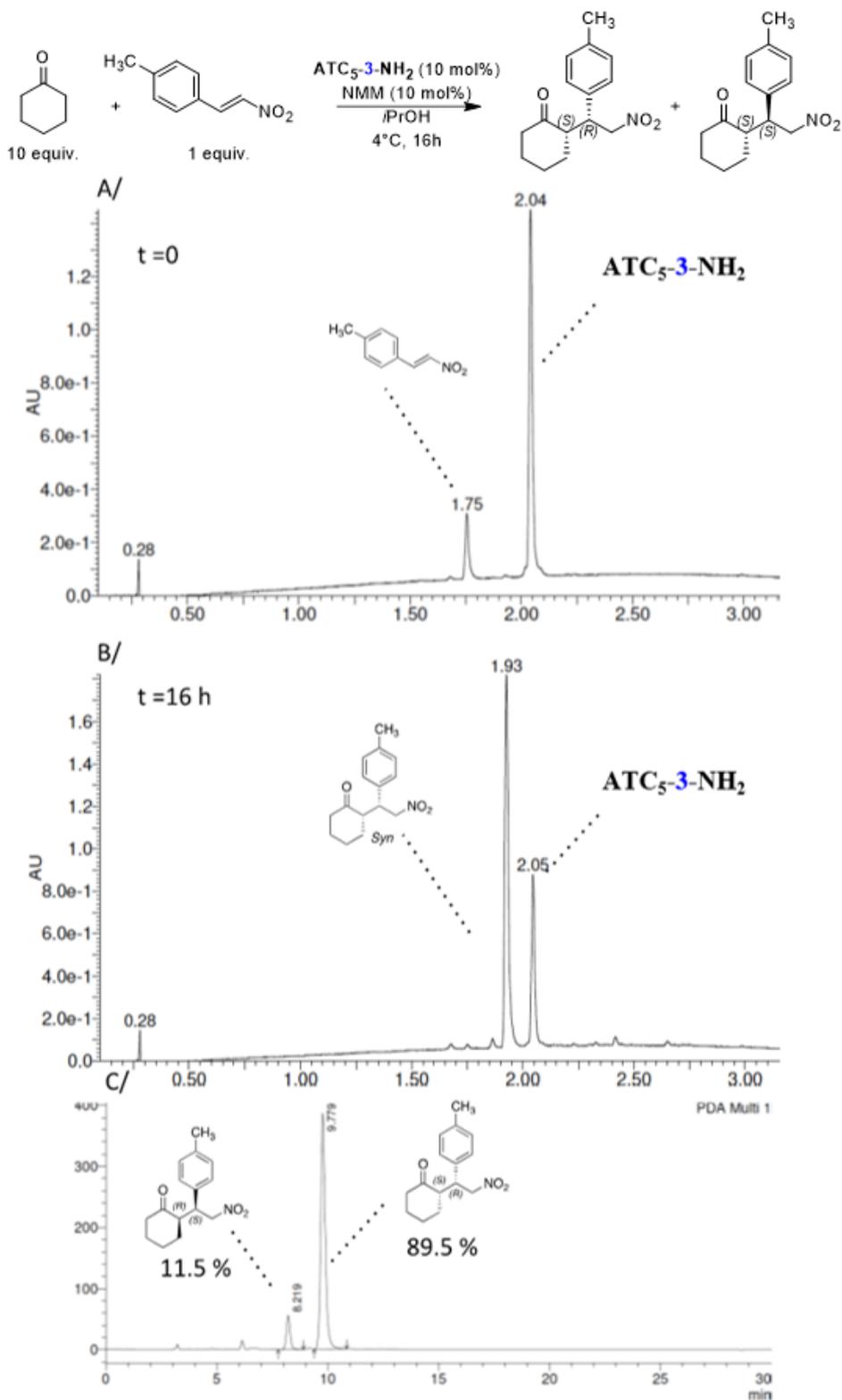


Figure S4: HPLC profiles of the nitro-Michael addition reaction of cyclohexanone to trans-4-methyl-β-nitrostyrene. Reaction performed over 16h with 0.037 mmol nitroolefine, 0.37 mmol cyclohexanone, 0.0037 mmol NMM and 0.0037 mmol catalyst in 150μl solvent. A/, B/monitoring of the nitro-Michael reaction over 16 hours. C/enantiomeric excess as determined by method B.

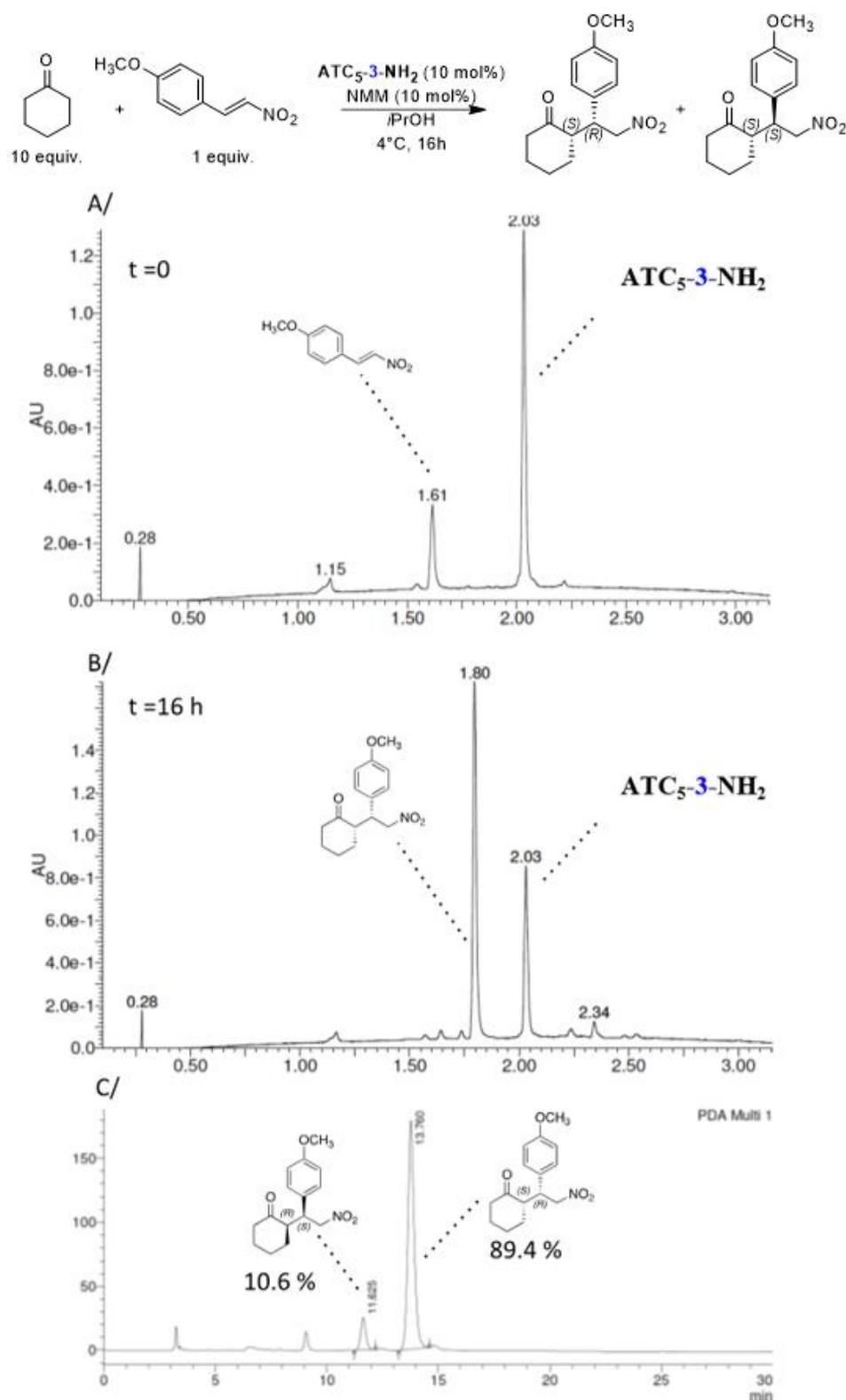


Figure S5: HPLC profiles of the nitro-Michael addition reaction of cyclohexanone to *trans*-4-methoxy- $\beta$ -nitrostyrene. Reaction performed over 16h with 0.037 mmol nitroolefine, 0.37 mmol cyclohexanone, 0.0037 mmol NMM and 0.0037 mmol catalyst in 150 $\mu$ l solvent. A/, B/monitoring of the nitro-Michael reaction over 16 hours. C/enantiomeric excess as determined by method B.

## VI. NMR AND MOLECULAR MODELLING STUDIES

### 1. General procedures

**NMR experiments:** The NMR samples contained 2-10 mM of **ATC<sub>3</sub>-2-NH<sub>i</sub>Pr**, **ATC<sub>3</sub>-2\*-NH<sub>i</sub>Pr** and **ATC<sub>5</sub>-3-NH<sub>i</sub>Pr** dissolved in CD<sub>3</sub>OH and 2 mM of **M(n)-ATC<sub>L</sub>-P-NH<sub>2</sub>** dissolved in iPrOH-d<sub>7</sub>. All spectra were recorded on a Bruker Avance 600 AVANCE III spectrometer equipped with a 5 mm triple-resonance cryoprobe (<sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N). Homonuclear 2-D spectra DQF-COSY, TOCSY (DIPS12) and ROESY were typically recorded in the phase-sensitive mode using the States-TPPI method as data matrices of 300 real (t<sub>1</sub>) × 2048 (t<sub>2</sub>) complex data points; 8-32 scans per t<sub>1</sub> increment with 1.2 s recovery delay and spectral width of 7211 Hz in both dimensions were used. The mixing times were 80 ms for TOCSY and 350 ms for the ROESY experiments. In addition, 2D heteronuclear spectra, <sup>13</sup>C-HSQC and <sup>13</sup>C-HMBC were acquired to fully assign the oligomers (24-32 scans, 256 real (t<sub>1</sub>) × 2048-4096 (t<sub>2</sub>) complex data points). Spectra were processed with Topspin (Bruker Biospin) and visualized with Topspin or NMRview on a Linux station. The matrices were zero-filled to 1024 (t<sub>1</sub>) × 2048 (t<sub>2</sub>) points after apodization by shifted sine-square multiplication and linear prediction in the F1 domain. Chemical shifts were referenced to the TMS.

**Force field libraries for ATC fragments** were generated using the restrained electrostatic potential (RESP) method. Atom charges were derived from the electrostatic potential obtained by DFT calculations with Gaussian 09<sup>[221]</sup> at the M062X/6-31+G(d) level of theory in vacuum. The computed electrostatic potential (ESP) was fit using RESP charge fitting in antechamber.

**Structure calculations.** <sup>1</sup>H chemical shifts were assigned according to classical procedures. NOE cross-peaks were integrated and assigned within the NMRView software<sup>[216]</sup>. The volume of a ROE between methylene pair protons or ortho phenyl protons was used as a reference of 1.8 Å or 2.48 Å respectively. The lower bound for all restraints was fixed at 1.8 Å and upper bounds at 2.7, 3.3 and 5.0 Å, for strong, medium and weak correlations, respectively. Pseudo-atoms corrections of the upper bounds were applied for unresolved aromatic, methylene and methyl protons signals as described previously.<sup>[213]</sup> Structure calculations were performed with AMBER 16<sup>[212]</sup> in three stages: cooking, simulated annealing in vacuum and refinement in a solvent box. The cooking stage was performed at 600 K to generate 100 initial random structures. SA calculations were carried during 20 ps (20 000 steps, 1 fs long) as described elsewhere. First, the temperature was risen quickly and was maintained at 600 K for the first

5000 steps, then the system was cooled gradually from 600 K to 100 K from step 5001 to 18 000 and finally the temperature was brought to 0 K during the 2000 remaining steps. For the 3000 first steps, the force constant of the distance restraints was increased gradually from 2.0 kcal.mol<sup>-1</sup>Å to 20 kcal.mol<sup>-1</sup>Å. For the rest of the simulation (step 3001 to 20 000), the force constant is kept at 20 kcal.mol<sup>-1</sup>Å. The calculations were launched using the generalized born solvation model. The 20 lowest energy structures with no violations > 0.3 Å were considered as representative of the peptide structure. The representation and quantitative analysis were carried out using Ptraj and MOLMOL<sup>[218]</sup>.

### Nomenclature and torsion angles description

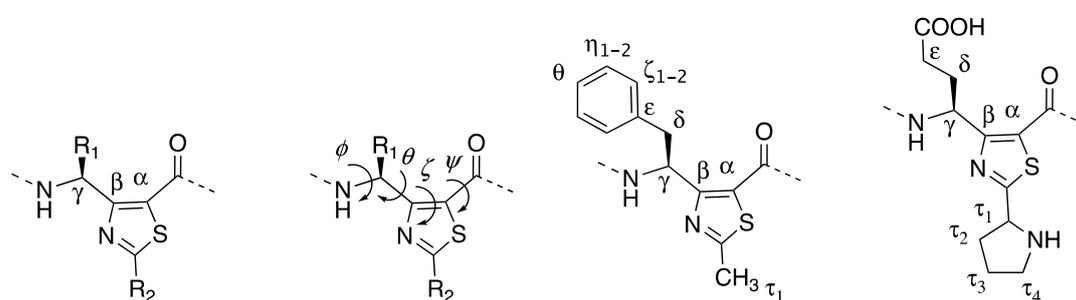


Figure S6: Nomenclature and torsion angles description for ATC units

## 2. NMR analyses of ATC<sub>3</sub>-2-NHiPr

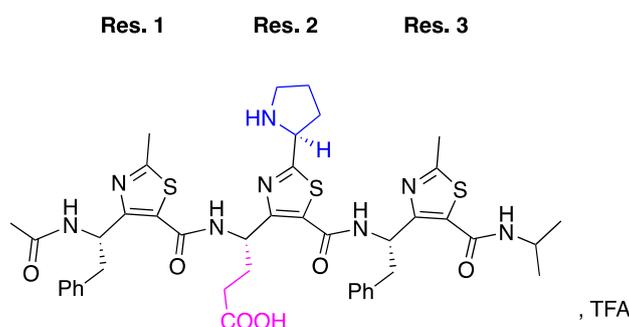


Table S1: <sup>1</sup>H NMR in CD<sub>3</sub>OH at 288K

Residue	HN	<sup>γ</sup> CH	<sup>δ</sup> CH	Others
Ac	-	-	-	CH <sub>3</sub> 1.94
Res. 1	8.94	5.31	3.08	Ph 6.74 <sup>τ1</sup> CH <sub>3</sub> 2.68
Res. 2	9.90	5.63	2.25-2.35	<sup>ε</sup> CH <sub>2</sub> 2.08 <sup>τ1</sup> CH 5.08 <sup>τ2</sup> CH <sub>2</sub> 2.15-2.29 <sup>τ3</sup> CH <sub>2</sub> 2.07-2.21 <sup>τ4</sup> CH <sub>2</sub> 3.42-3.48

<b>Res. 3</b>	10.27	5.61	3.31	Ph 7.13 $^{13}\text{C}$ 2.70
<b>NHiPr</b>	8.85	-	-	CH 4.07 $\text{CH}_3$ 1.24

Table S2 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in  $\text{CD}_3\text{OH}$  at 298K.

Residue	$^3J(\text{NH}, \gamma\text{CH})$
<b>Res. 1</b>	7.5
<b>Res. 2</b>	6.5
<b>Res. 3</b>	6.5
<b>NHiPr</b>	7.5

Table S3- Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in  $\text{CD}_3\text{OH}$  at 288K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	nd
2.NH-3.NH	nd
3.NH-NHiPr	nd
Ac-1.NH	m
1.H $\gamma$ -2.NH	s
2.H $\gamma$ -3.NH	s
3.H $\gamma$ -NHiPr	s
1.H $\gamma$ -2.H $\gamma$	nd
2.H $\gamma$ -3.H $\gamma$	nd
3.H $\gamma$ -iPr	w
1.H $\gamma$ -2.H $\tau$ 1	w
2.H $\gamma$ -3.H $\tau$ 1	s
1.H $\gamma$ -2.H $\delta$	nd
2.H $\gamma$ -3.H $\delta$	nd
1.Ph-2.H $\tau$ 1	w
2.Ph-3.H $\tau$ 1	w

nd: not detected

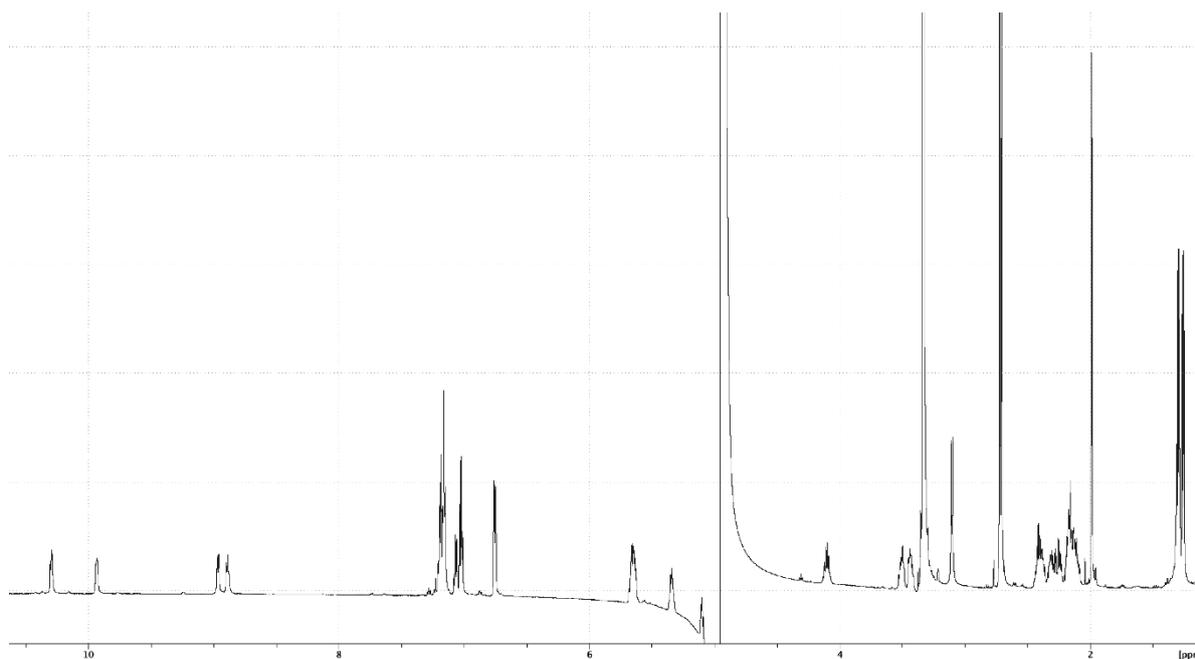


Figure S7 -  $^1\text{H}$  NMR spectrum of **ATC<sub>3</sub>-2-NHPr** in  $\text{CD}_3\text{OH}$  at 288K

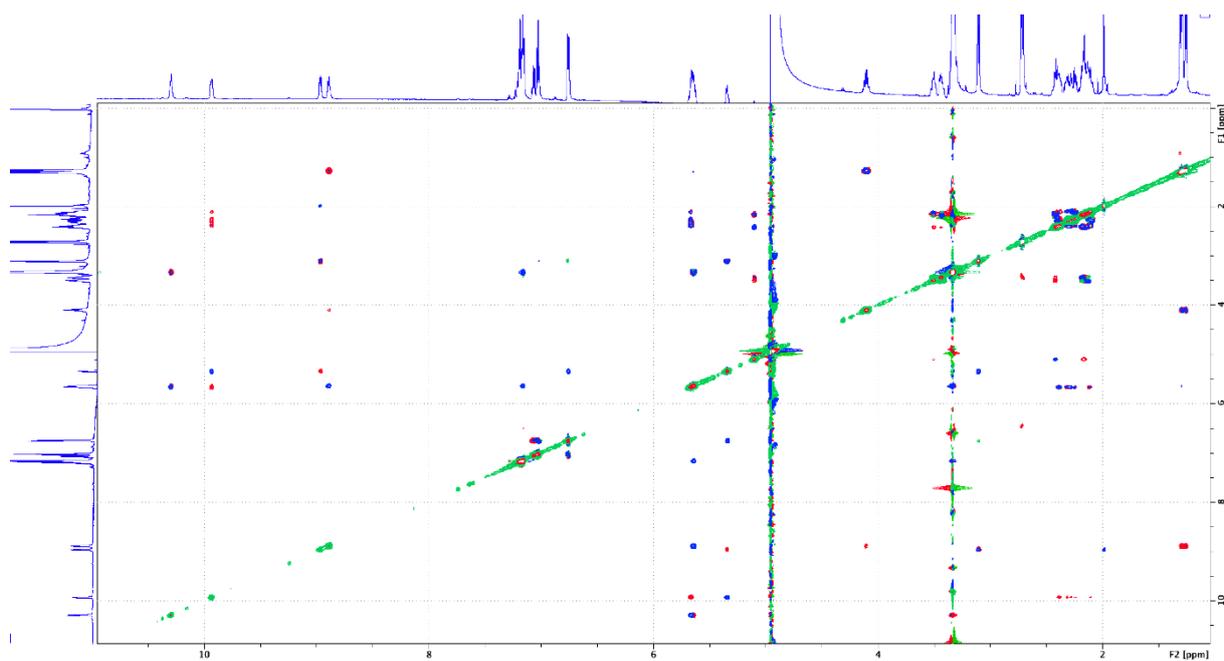
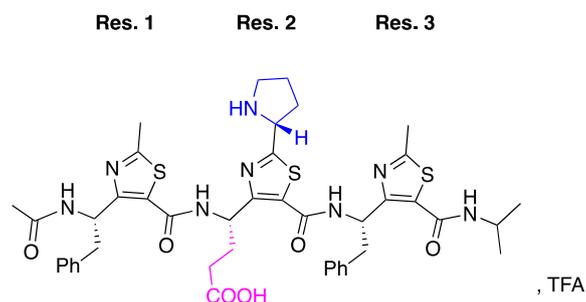


Figure S8 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **ATC<sub>3</sub>-2-NHPr** in  $\text{CD}_3\text{OH}$  at 288K

3. NMR analyses of ATC<sub>3</sub>-2\*-NH<sub>i</sub>PrTable S4 - <sup>1</sup>H NMR in CD<sub>3</sub>OH at 298K

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 1.92
Res. 1	8.76	5.45	3.08	Ph 6.84 <sup>13</sup> C <sub>3</sub> 2.71
Res. 2	9.64	5.52	2.25-2.34	<sup>2</sup> CH <sub>2</sub> 2.07 <sup>1</sup> CH 5.01 <sup>2</sup> CH <sub>2</sub> 2.57-2.29 <sup>3</sup> CH <sub>2</sub> 2.07-2.22 <sup>4</sup> CH <sub>2</sub> 3.42-3.51
Res. 3	10.32	5.64	3.31	Ph 7.13 <sup>13</sup> C <sub>3</sub> 2.70
NH <sub>i</sub> Pr	8.79	-	-	CH 4.07 CH <sub>3</sub> 1.24

Table S5 - Coupling Constants <sup>3</sup>J(NH,  $\gamma$ CH) (in Hz). Values were measured in CD<sub>3</sub>OH at 298K.

Residue	<sup>3</sup> J(NH, $\gamma$ CH)
Res. 1	7.5
Res. 2	6.5
Res. 3	6.5
NH <sub>i</sub> Pr	7.5

Table S6 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in CD<sub>3</sub>OH at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	nd
2.NH-3.NH	nd
3.NH-NH <sub>i</sub> Pr	nd
Ac-1.NH	m
1.H $\gamma$ -2.NH	s

2.H $\gamma$ -3.NH	s
3.H $\gamma$ -NH <i>i</i> Pr	s
1.H $\gamma$ -2.H $\gamma$	nd
2.H $\gamma$ -3.H $\gamma$	nd
3.H $\gamma$ - <i>i</i> Pr	nd
1.H $\gamma$ -2.H $\tau$ 1	nd
2.H $\gamma$ -3.H $\tau$ 1	w
1.H $\gamma$ -2.H $\delta$	nd
2.H $\gamma$ -3.H $\delta$	nd
3.H $\gamma$ - <i>i</i> Pr	nd
1.Ph-2.H $\tau$ 1	m
2.Ph-3.H $\tau$ 1	w

*nd: not detected*

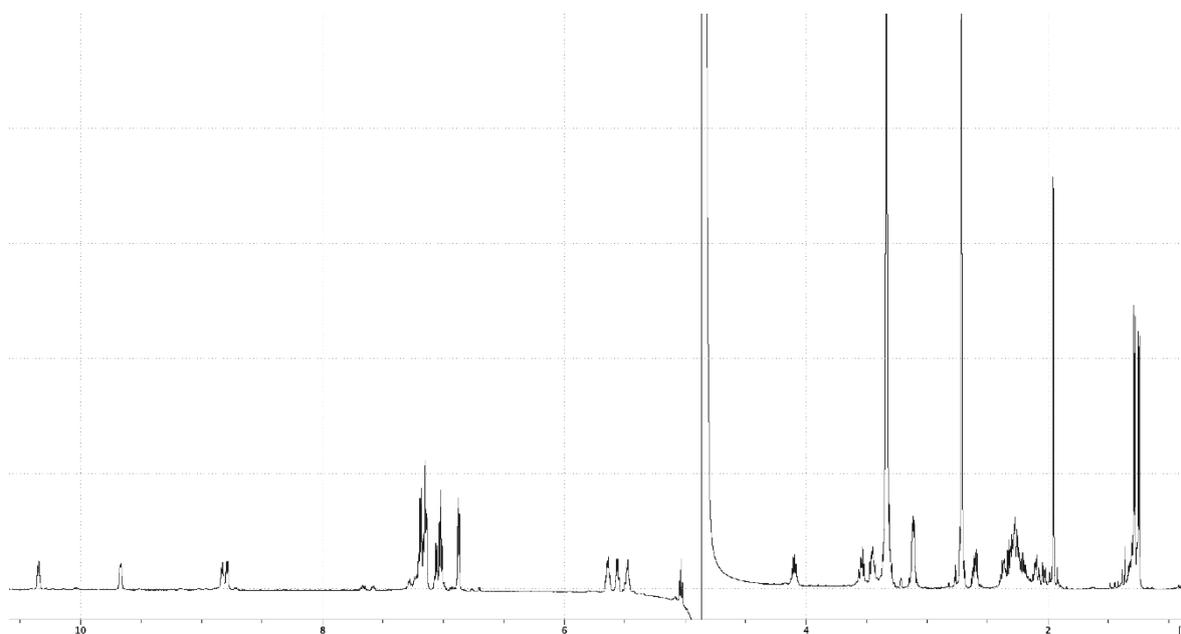


Figure S9 -  $^1\text{H}$  NMR spectrum of  $\text{ATC}_3\text{-2}^*\text{-NHPr}$  in  $\text{CD}_3\text{OH}$  at 298K

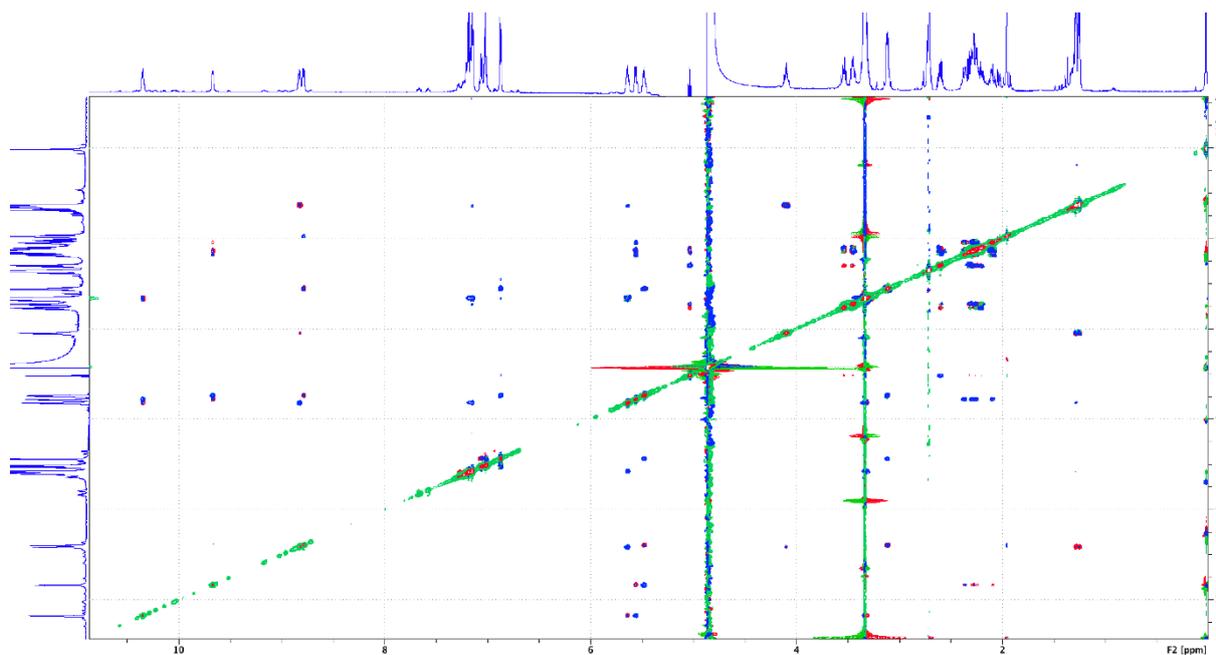


Figure S10 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of ATC<sub>3</sub>-2\*-NHPr in CD<sub>3</sub>OH at 298K

#### 4. NMR analyses of ATC<sub>5</sub>-3-NHPr

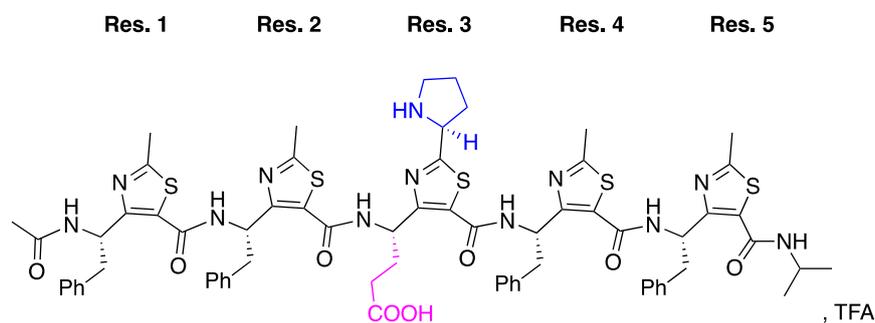


Table S7 - <sup>1</sup>H NMR in CD<sub>3</sub>OH at 293K

Residue	HN	<sup>γ</sup> CH	<sup>δ</sup> CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.06
Res. 1	9.04	5.51	3.18	Ph 7.00-6.78 <sup>τ</sup> 1CH <sub>3</sub> 2.68
Res. 2	10.65	5.41	3.29	Ph 7.03-6.76 <sup>τ</sup> 1CH <sub>3</sub> 2.74
Res. 3	9.89	5.74	2.32	<sup>ε</sup> CH <sub>2</sub> 2.15 <sup>τ</sup> 1CH 5.20 <sup>τ</sup> 2CH <sub>2</sub> 2.38-2.20 <sup>τ</sup> 3CH <sub>2</sub> 2.07-2.14 <sup>τ</sup> 4CH <sub>2</sub> 3.41-3.49
Res. 4	10.34	5.75	3.32	Ph 6.96-6.83 <sup>τ</sup> 1CH <sub>3</sub> 2.72
Res. 5	10.14	5.50	3.36	Ph 7.19-7.15 <sup>τ</sup> 1CH <sub>3</sub> 2.76

<b>NHiPr</b>	9.07	-	-	CH 4.09 CH <sub>3</sub> 1.26-1.24
--------------	------	---	---	-----------------------------------

Table S8 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in  $\text{CD}_3\text{OH}$  at 293K.

<b>Residue</b>	$^3J(\text{NH}, \gamma\text{CH})$
<b>Res. 1</b>	7.1
<b>Res. 2</b>	5.9
<b>Res. 3</b>	6.9
<b>Res. 4</b>	5.6
<b>Res. 5</b>	6.5
<b>NHiPr</b>	7.4

Table S9 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in  $\text{CD}_3\text{OH}$  at 293K. Strong (*s*) < 2.7 Å, 2.7 Å < Medium (*m*) < 3.3 Å, 3.3 Å < Weak (*w*).

<b>NOE correlations</b>	<b>Intensity</b>
1.NH-2.NH	nd
2.NH-3.NH	nd
3.NH-4.NH	nd
4.NH-5.NH	nd
5.NH-NHiPr	nd
Ac-1.NH	m
1.H $\gamma$ -2.NH	s
2.H $\gamma$ -3.NH	s
3.H $\gamma$ -4.NH	s
4.H $\gamma$ -5.NH	s
5.H $\gamma$ -NHiPr	s
1.H $\gamma$ -2.H $\gamma$	nd
2.H $\gamma$ -3.H $\gamma$	nd
3.H $\gamma$ -4.H $\gamma$	nd
4.H $\gamma$ -5.H $\gamma$	nd
1.H $\gamma$ -2.H $\tau$ 1	w
2.H $\gamma$ -3.H $\tau$ 1	w
3.H $\gamma$ -4.H $\tau$ 1	w
4.H $\gamma$ -5.H $\tau$ 1	w
1.H $\gamma$ -2.H $\delta$	nd

2.H $\gamma$ -3.H $\delta$	nd
3.H $\gamma$ -4.H $\delta$	nd
4.H $\gamma$ -5.H $\delta$	nd
1.Ph-2.H $\tau$ 1	m
2.Ph-3.H $\tau$ 1	w
3.Ph-4.H $\tau$ 1	nd
4.Ph-5.H $\tau$ 1	m

*nd: not detected*

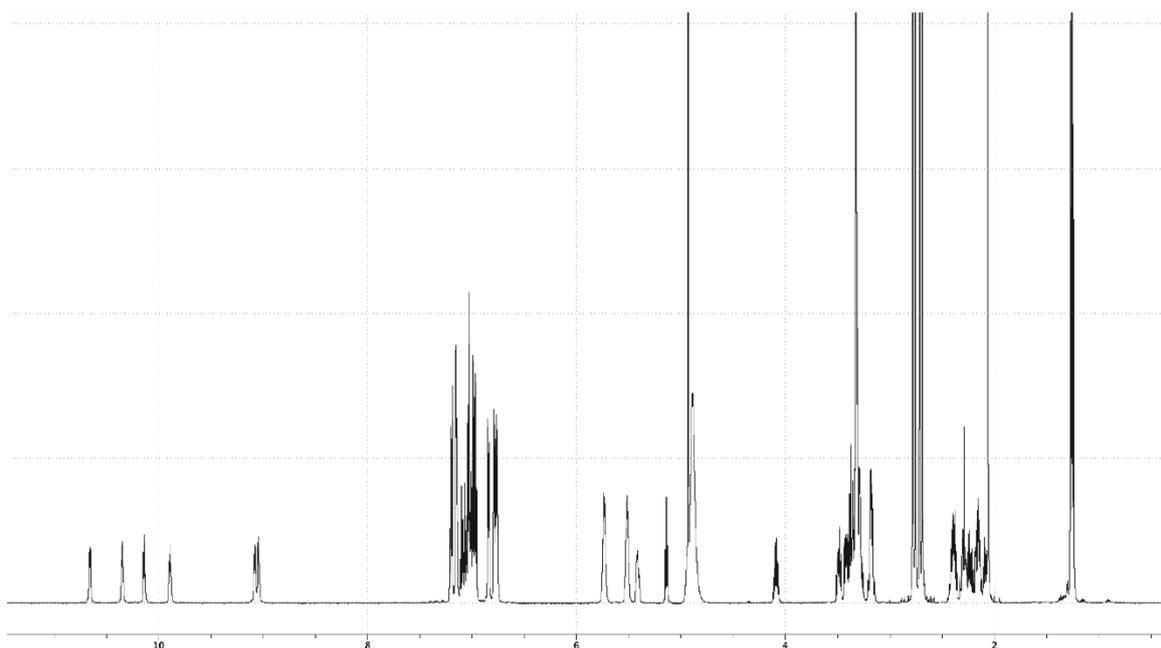


Figure S11 -  $^1\text{H}$  NMR spectrum of *ATC5-3-NHPr* in  $\text{CD}_3\text{OH}$  at 293K

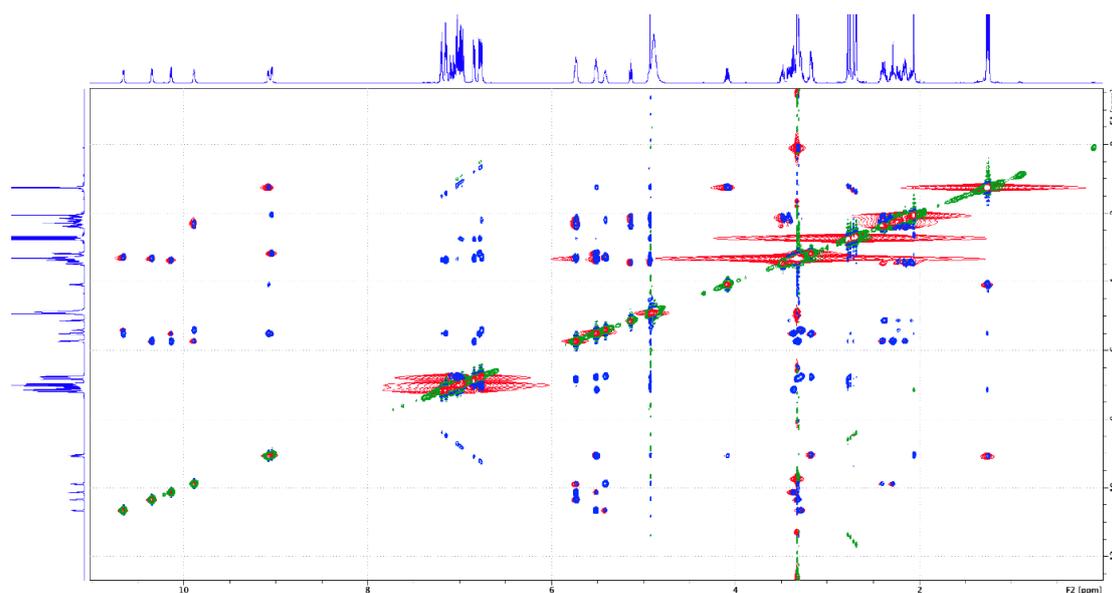


Figure S12 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of *ATC5-3-NHPr* in  $\text{CD}_3\text{OH}$  at 293K

### 5. NMR analyses of ATC<sub>5</sub>-1-NH<sub>2</sub>

ATC<sub>5</sub>-1-NH<sub>2</sub> precipitated in few minutes in deuterated isopropanol making NMR characterization impossible.

### 6. NMR analyses of ATC<sub>5</sub>-2-NH<sub>2</sub>

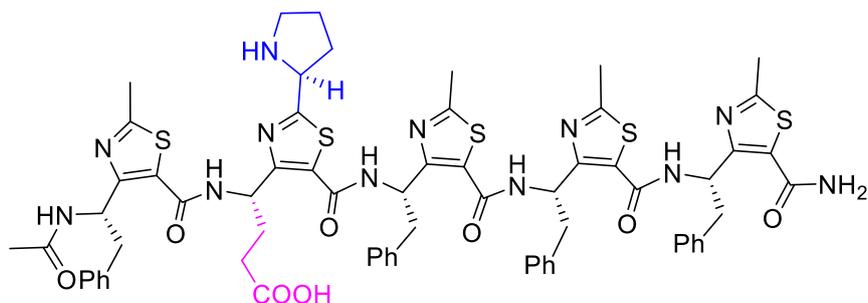


Table S10 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (72%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.01
Res. 1	9.09	5.21	3.11	Ph 76.68 <sup>1</sup> CH <sub>3</sub> 2.66
Res. 2	10.08	5.76	2.29-2.40	<sup>ε</sup> CH <sub>2</sub> 2.09-2.11 <sup>1</sup> CH 4.07 <sup>2</sup> CH <sub>2</sub> na <sup>3</sup> CH <sub>2</sub> na <sup>4</sup> CH <sub>2</sub> na
Res. 3	10.25	5.79	3.35	Ph 7.09 – 6.94 – 6.87 <sup>1</sup> CH <sub>3</sub> 2.68
Res. 4	10.24	5.73	3.35	Ph 6.89 <sup>1</sup> CH <sub>3</sub> 2.66
Res. 5	10.42	5.57	3.41-3.24	Ph 7.24 – 7.16 <sup>1</sup> CH <sub>3</sub> 2.68
NH <sub>2</sub>	8.88-7.34	-	-	-

na: not assigned

Table S11 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (72%)

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	172.8	-	-	-	-	CH <sub>3</sub> na
Res. 1	172.2	50.1	39.4	137.3	154.2	Ph 127.9 <sup>1</sup> C 18.1
Res. 2	161.7	48.8	29.1	nd	nd	<sup>ε</sup> C 31.3 <sup>1</sup> C na <sup>2</sup> C na <sup>3</sup> C na <sup>4</sup> C na
Res. 3	160.8	51.0	39.4	137.0	154.0	Ph 128.6 <sup>1</sup> C 18.1

<b>Res. 4</b>	162.3	51.0	39.4	137.0	154.0	Ph 128.6 <sup>13</sup> C 18.1
<b>Res. 5</b>	161.5	51.8	39.9	137.4	154.3	Ph 128.6 <sup>13</sup> C 18.1

na: not assigned

Table S12 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K

Residue	Major (298 K)
<b>Res. 1</b>	7.6
<b>Res. 2</b>	7.8
<b>Res. 3</b>	o
<b>Res. 4</b>	o
<b>Res. 5</b>	6.3

o: overlaps

Table S13 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	nd
2.NH-3.NH	nd
3.NH-4.NH	o
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.H $\gamma$ -2.NH	s
2.H $\gamma$ -3.NH	s
3.H $\gamma$ -4.NH	s
4.H $\gamma$ -5.NH	s
5.H $\gamma$ -NH <sub>2</sub>	m
5.H $\gamma$ -NH <sub>2</sub>	s
1.H $\gamma$ -2.H $\gamma$	nd
2.H $\gamma$ -3.H $\gamma$	nd
3.H $\gamma$ -4.H $\gamma$	nd
4.H $\gamma$ -5.H $\gamma$	nd
1.H $\gamma$ -2.H $\tau$ 1	nd
2.H $\gamma$ -3.H $\tau$ 1	m
3.H $\gamma$ -4.H $\tau$ 1	m

4.H $\gamma$ -5.H $\tau$ 1	m
1.H $\gamma$ -2.H $\delta$	nd
2.H $\gamma$ -3.H $\delta$	m
3.H $\gamma$ -4.H $\delta$	m
4.H $\gamma$ -5.H $\delta$	nd
1.H $\delta$ -2.H $\tau$ 1	nd
2.H $\delta$ -3.H $\tau$ 1	nd
3.H $\delta$ -4.H $\tau$ 1	o
4.H $\delta$ -5.H $\tau$ 1	o
1.Ph-2.H $\tau$ 1	nd
3.Ph-4.H $\tau$ 1	m
4.Ph-5.H $\tau$ 1	m

*o*: overlap with the solvent signals, *nd*: not detected.

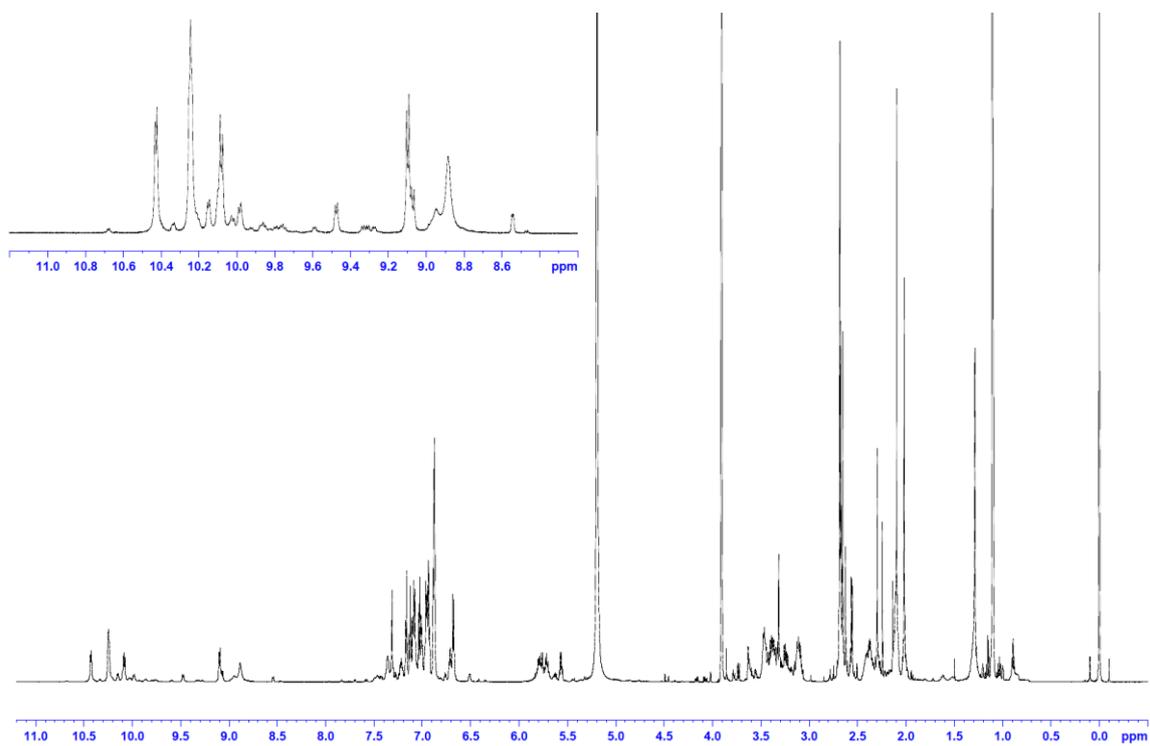


Figure S13 -  $^1\text{H}$  NMR spectrum of **ATC<sub>5</sub>-2-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

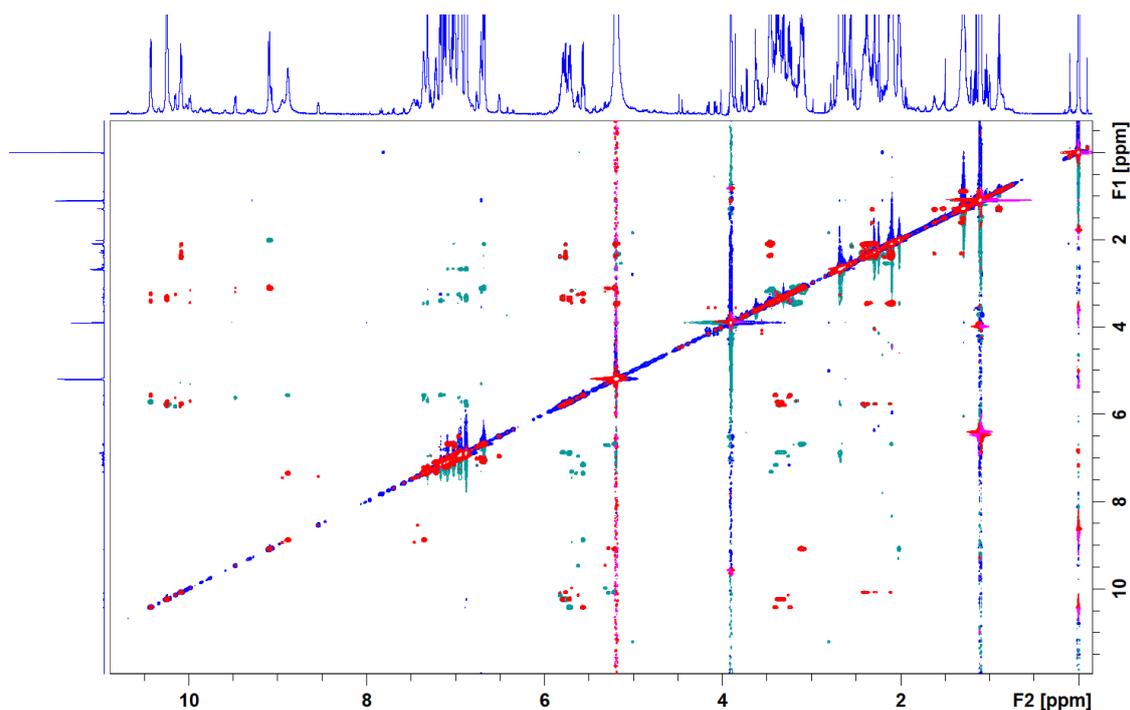


Figure S14 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of ATC<sub>5</sub>-2-NH<sub>2</sub> in iPrOH-d<sub>7</sub> at 298K

### 7. NMR analyses of ATC<sub>5</sub>-3-NH<sub>2</sub>

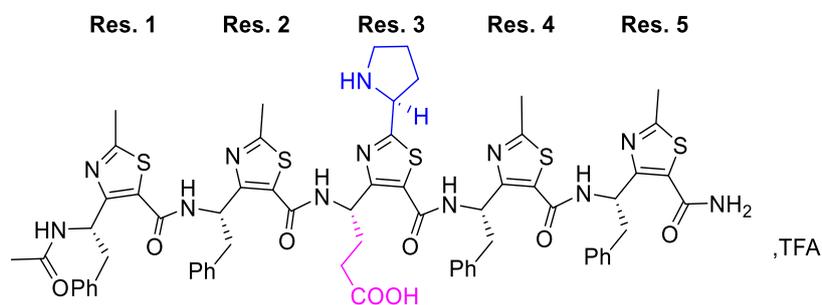


Table S14 - <sup>1</sup>H NMR in iPrOH-d<sub>7</sub> at 298K of the major conformer (82%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.06
Res. 1	9.10	5.54	3.19	Ph 7.04-6.78 <sup>1</sup> CH <sub>3</sub> 2.64
Res. 2	10.81	5.24	3.27	Ph 7.06-6.72 <sup>1</sup> CH <sub>3</sub> 2.71
Res. 3	9.94	5.75	2.39-2.30	<sup>6</sup> CH <sub>2</sub> 2.15 <sup>1</sup> CH 5.20 <sup>2</sup> CH <sub>2</sub> 2.30-2.14 <sup>3</sup> CH <sub>2</sub> 2.05 <sup>4</sup> CH <sub>2</sub> 3.43
Res. 4	10.23	5.69	3.33-3.28	Ph 6.95-6.83 <sup>1</sup> CH <sub>3</sub> 2.66
Res. 5	10.17	5.62	3.41-3.31	Ph 7.19-7.14 <sup>1</sup> CH <sub>3</sub> 2.70

<b>NH<sub>2</sub></b>	8.79-7.39	-	-	-
-----------------------	-----------	---	---	---

Table S15 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 298K of the major conformer (82%)

<b>Residue</b>	<b>C(O)</b>	<b>γC</b>	<b>δC</b>	<b>βC</b>	<b>αC</b>	<b>Others</b>
<b>Ac</b>	172.1	-	-	-	-	CH <sub>3</sub> 22.0
<b>Res. 1</b>	162.1	50.3	39.4	136.8	154.2	Ph 128.1-126.4-126.2 <sup>13</sup> C 18.7
<b>Res. 2</b>	161.7	51.2	39.3	137.0	153.6	Ph 128.2-126.8-126.2 <sup>13</sup> C 18.8
<b>Res. 3</b>	161.1	48.8	29.0	na	na	<sup>6</sup> C 30.6 <sup>13</sup> C 58.0 <sup>12</sup> C na <sup>13</sup> C na <sup>14</sup> C 45.2
<b>Res. 4</b>	161.6	51.9	39.8	137.1	153.7	Ph 128.3-126.1-126.2 <sup>13</sup> C 18.6
<b>Res. 5</b>	na	61.4	40.2	137.3	154.2	Ph 128.7-126.4-126.1 <sup>13</sup> C 19.4

na: not assigned

Table S16 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K of the Minor conformer a (16%)

<b>Residue</b>	<b>HN</b>	<b>γCH</b>	<b>δCH</b>	<b>Others</b>
<b>Ac</b>	-	-	-	CH <sub>3</sub> 2.04
<b>Res. 1</b>	9.03	5.60	3.22-3.13	na
<b>Res. 2</b>	10.90	5.53	3.50	na
<b>Res. 3</b>	9.69	4.98	2.16	<sup>6</sup> CH <sub>2</sub> 1.97-1.86
<b>Res. 4</b>	10.08	5.75	3.42-3.10	na
<b>Res. 5</b>	10.38	5.66	3.41-3.26	na
<b>NH<sub>2</sub></b>	8.90-7.42	-	-	-

na: not assigned

Table S17 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K of the Minor conformer b (2%)

<b>Residue</b>	<b>HN</b>	<b>γCH</b>	<b>δCH</b>	<b>Others</b>
<b>Ac</b>	-	-	-	CH <sub>3</sub> 2.12
<b>Res. 1</b>	9.11	4.94	3.09-2.88	na
<b>Res. 2</b>	10.30	5.61	3.57-3.16	na
<b>Res. 3</b>	10.29	5.78	2.40	<sup>6</sup> CH <sub>2</sub> 2.26-2.23

<b>Res. 4</b>	10.56	5.75	3.41-3.27	na
<b>Res. 5</b>	10.15	5.63	3.40-3.29	na
<b>NH<sub>2</sub></b>	na	-	-	-

na: not assigned

Table S18 - Coupling Constants  $^3J(\text{NH}, ^\gamma\text{CH})$  (in Hz). Values were measured in *iPrOH-d<sub>7</sub>* at 298K.

<b>Residue</b>	<b>Major (298 K)</b>	<b>Minor a (298 K)</b>
<b>Res. 1</b>	7.0	7.9
<b>Res. 2</b>	6.5	5.4
<b>Res. 3</b>	7.5	8.0
<b>Res. 4</b>	5.2	7.2
<b>Res. 5</b>	6.5	6.6

Table S19 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *iPrOH-d<sub>7</sub>* at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

<b>NOE correlations</b>	<b>Intensity</b>
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	o
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.H $\gamma$ -2.NH	s
2.H $\gamma$ -3.NH	m/w
3.H $\gamma$ -4.NH	s
4.H $\gamma$ -5.NH	s
5.H $\gamma$ -NH <sub>2</sub>	m
5.H $\gamma$ -NH <sub>2</sub>	m
1.H $\gamma$ -2.H $\gamma$	nd
2.H $\gamma$ -3.H $\gamma$	nd
3.H $\gamma$ -4.H $\gamma$	nd
4.H $\gamma$ -5.H $\gamma$	nd
1.H $\gamma$ -2.H $\tau$ 1	w
2.H $\gamma$ -3.H $\tau$ 1	o

3.H $\gamma$ -4.H $\tau$ 1	w
4.H $\gamma$ -5.H $\tau$ 1	w
1.H $\gamma$ -2.H $\delta$	nd
2.H $\gamma$ -3.H $\delta$	nd
3.H $\gamma$ -4.H $\delta$	nd
4.H $\gamma$ -5.H $\delta$	nd
1.Ph-2.H $\tau$ 1	m
4.Ph-5.H $\tau$ 1	m

*o*: overlap with the solvent signals; *nd*: not detected

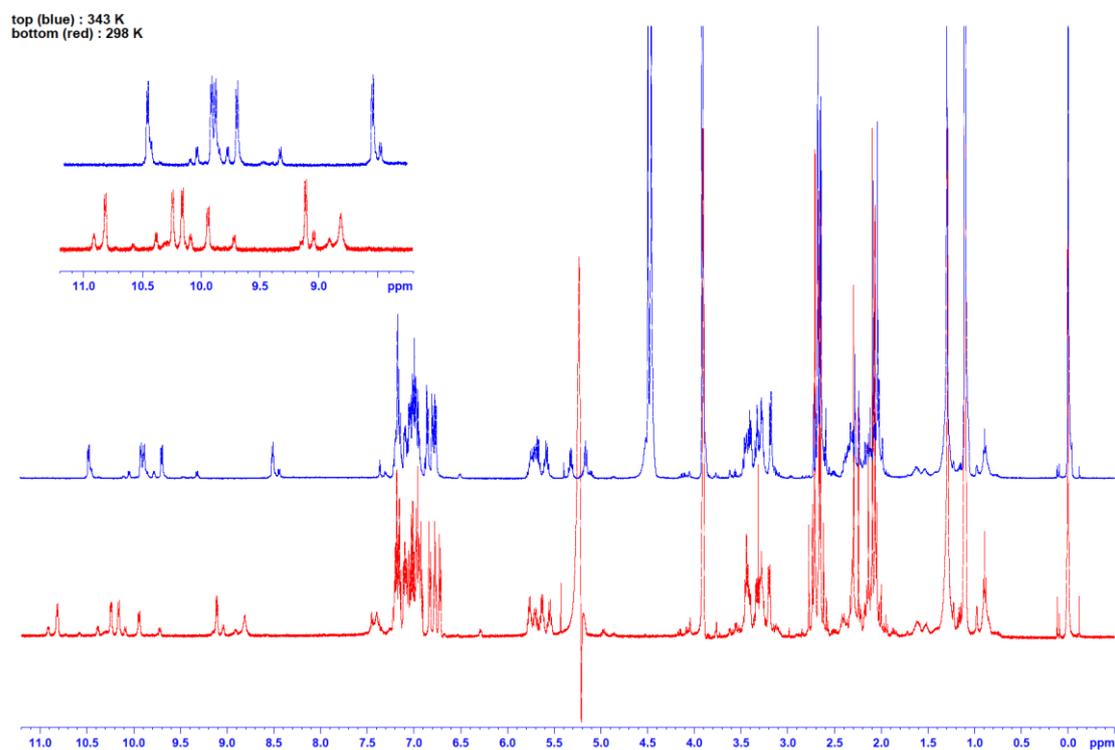


Figure S15 -  $^1\text{H}$  NMR spectrum of  $\text{ATC}_5\text{-3-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K (red) and 343K (blue)

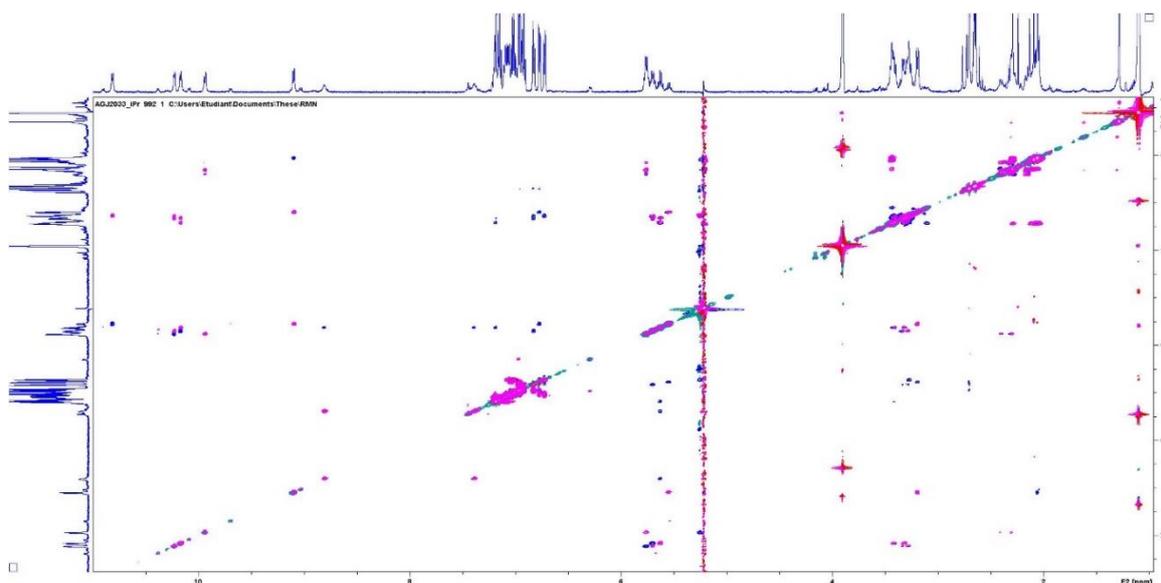


Figure S16 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of ATC<sub>5</sub>-3-NH<sub>2</sub> in iPrOH-d<sub>7</sub> at 298K

### 8. NMR analyses of ATC<sub>5</sub>-4-NH<sub>2</sub>

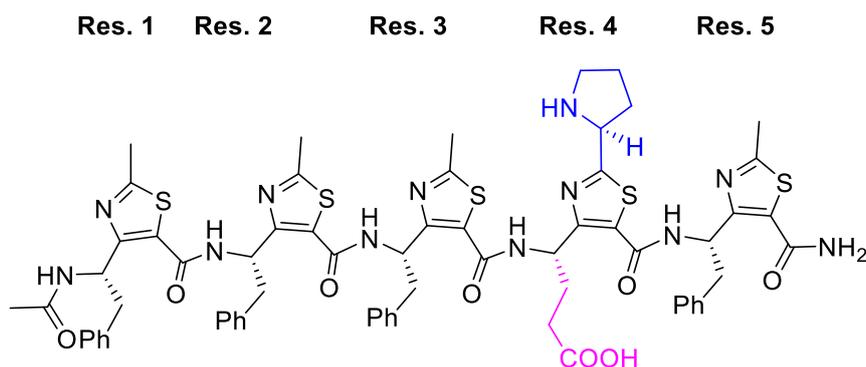


Table S20 - <sup>1</sup>H NMR in iPrOH-d<sub>7</sub> at 298K major conformer (73%)

Residue	HN	<sup>γ</sup> CH	<sup>δ</sup> CH	Others
Ac	-	-	-	CH <sub>3</sub> 1.96
Res. 1	8.95	5.59	3.13	Ph 6.81-6.91-6.26 <sup>τ</sup> 1CH <sub>3</sub> na
Res. 2	10.65	5.55	3.31	Ph 6.70 <sup>τ</sup> 1CH <sub>3</sub> na
Res. 3	10.56	-	3.20-3.01	Ph 6.64-6.70 <sup>τ</sup> 1CH <sub>3</sub> 1.84
Res. 4	9.92	5.65	2.19-2.24	<sup>ε</sup> CH <sub>2</sub> na <sup>τ</sup> 1CH 5.01 <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na
Res. 5	10.01	5.68	3.28-3.21	Ph 7.28-7.15 <sup>τ</sup> 1CH <sub>3</sub> na
NH <sub>2</sub>	8.45-7.30	-	-	-

na: not assigned

Table S21 -  $^{13}\text{C}$  NMR in *i*PrOH- $d_7$  at 298K major conformer (73%)

Residue	C(O)	$\gamma\text{C}$	$\delta\text{C}$	$\beta\text{C}$	$\alpha\text{C}$	Others
Ac	172.0	-	-	-	-	CH <sub>3</sub> 21.2
Res. 1	171.5	49.5	38.8	136.6	153.5	Ph 127.3-125.6 $^{13}\text{C}$ na
Res. 2	162.0	50.6	38.6	136.5	153.7	Ph 127.5 $^{13}\text{C}$ na
Res. 3	162.2	50.7	39.0	136.8	153.8	Ph 127.7 $^{13}\text{C}$ 17.0
Res. 4	161.3	48.5	30.3	135.6	na	$^{\epsilon}\text{C}$ na $^{13}\text{C}$ 57.5 $^{12}\text{C}$ na $^{13}\text{C}$ 22.0 $^{14}\text{C}$ na
Res. 5	160.3	51.8	44.6	137.7	155.5	Ph 128.5-128.4 $^{13}\text{C}$ na

na: not assigned

Table S22 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH- $d_6$  at 298K.

Residue	Major (298 K)
Res. 1	7.0
Res. 2	5.6
Res. 3	6.3
Res. 4	7.3
Res. 5	6.5

o: overlap with the solvent signals; nd: not detected

Table S23 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH- $d_7$  at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	m/w
3.NH-4.NH	w
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.H $\gamma$ -2.NH	s
2.H $\gamma$ -3.NH	s
3.H $\gamma$ -4.NH	nd
4.H $\gamma$ -5.NH	s
5.H $\gamma$ -NH <sub>2</sub>	s

5.H $\gamma$ -NH <sub>2</sub>	s
1.H $\gamma$ -2.H $\gamma$	o
2.H $\gamma$ -3.H $\gamma$	o
3.H $\gamma$ -4.H $\gamma$	o
4.H $\gamma$ -5.H $\gamma$	o
1.H $\gamma$ -2.H $\tau$ 1	m
2.H $\gamma$ -3.H $\tau$ 1	s
3.H $\gamma$ -4.H $\tau$ 1	-
4.H $\gamma$ -5.H $\tau$ 1	-
1.H $\gamma$ -2.H $\delta$	m
2.H $\gamma$ -3.H $\delta$	m
3.H $\gamma$ -4.H $\delta$	m
4.H $\gamma$ -5.H $\delta$	s
1..H $\delta$ -2.H $\tau$ 1	m
2..H $\delta$ -3.H $\tau$ 1	s
1.Ph-2.H $\tau$ 1	m

*o*: overlap with the solvent signals; *nd*: not detected

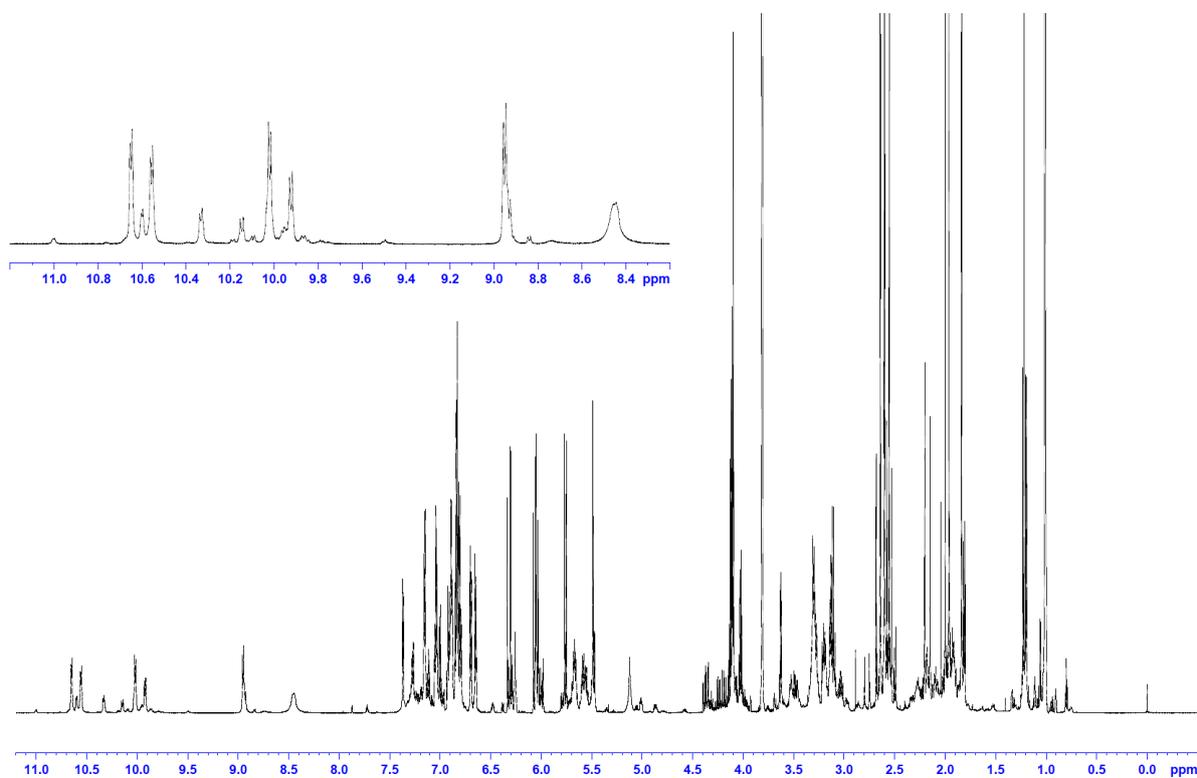


Figure S17 - <sup>1</sup>H NMR spectrum of ATC<sub>5</sub>-4-NH<sub>2</sub> in iPrOH-*d*<sub>7</sub> at 298K

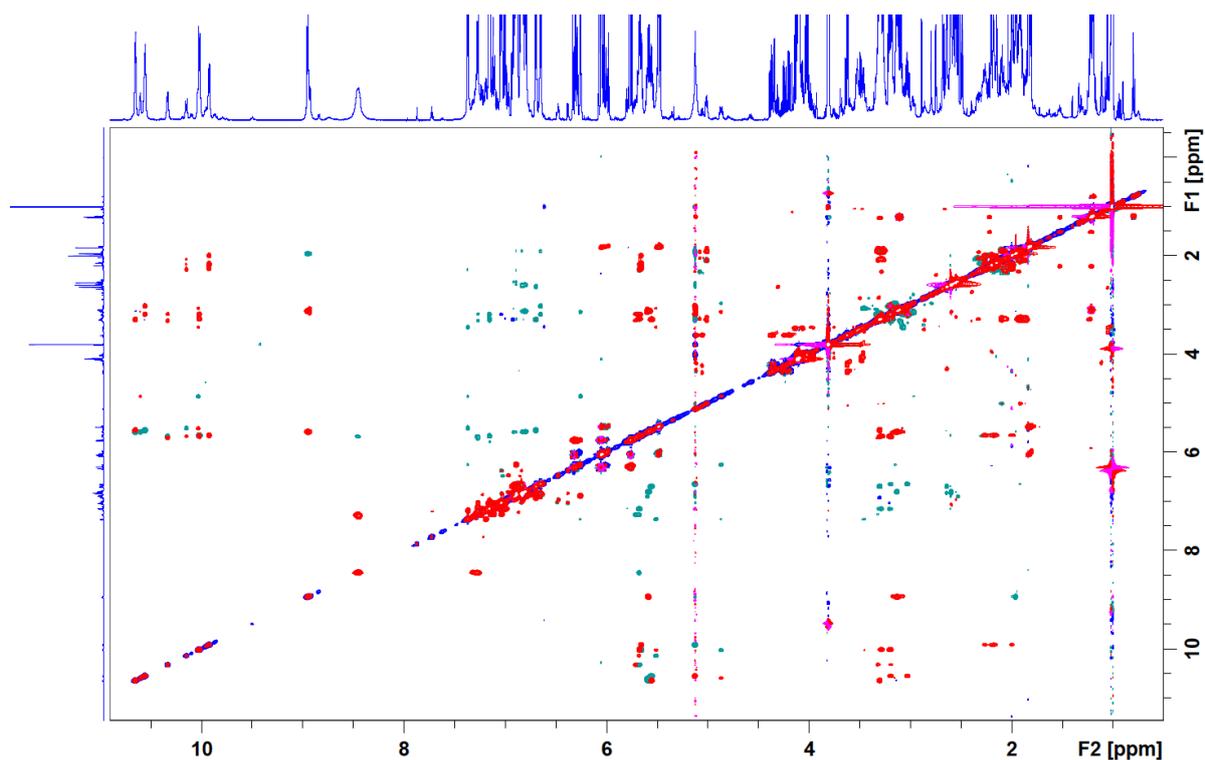


Figure S 18 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of ATC<sub>5</sub>-4-NH<sub>2</sub> in iPrOH-d<sub>7</sub> at 298K

### 9. NMR analyses of ATC<sub>5</sub>-5-NH<sub>2</sub>

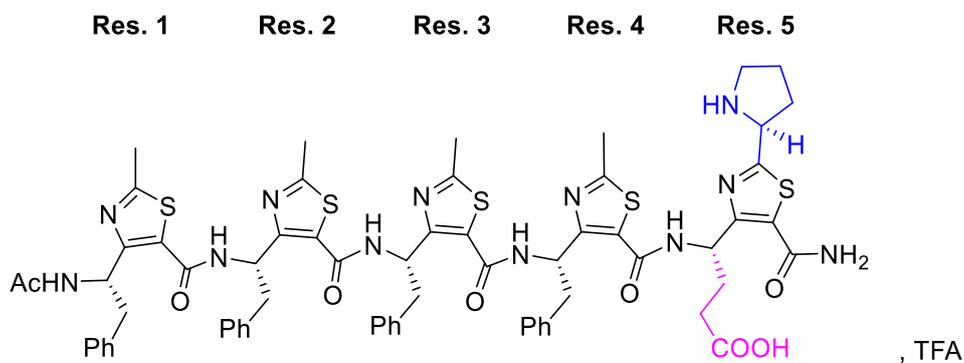


Table S24 - <sup>1</sup>H NMR in iPrOH-d<sub>7</sub> at 298K major conformer (83%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.06
Res. 1	9.03	5.66	3.22	Ph 6.86 <sup>1</sup> CH <sub>3</sub> 2.65
Res. 2	10.68	5.76	3.41	Ph 6.87 <sup>1</sup> CH <sub>3</sub> 2.99
Res. 3	10.55	5.63	3.30-3.38	Ph 6.89 <sup>1</sup> CH <sub>3</sub> 2.72

<b>Res. 4</b>	10.75	5.15	3.09-3.29	Ph 6.74 <sup>1</sup> C <sub>3</sub> 2.72
<b>Res. 5</b>	9.94	5.70	2.37-2.28	<sup>ε</sup> CH <sub>2</sub> na <sup>1</sup> CH 5.14 <sup>2</sup> CH <sub>2</sub> 2.29-2.50 <sup>3</sup> CH <sub>2</sub> 2.03-2.16 <sup>4</sup> CH <sub>2</sub> 3.41-3.42
<b>NH<sub>2</sub></b>	8.71-7.69	-	-	-

Table S25 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (83%)

<b>Residue</b>	<b>C(O)</b>	<b>γC</b>	<b>δC</b>	<b>βC</b>	<b>αC</b>	<b>Others</b>
<b>Ac</b>	172.0	-	-	-	-	CH <sub>3</sub> 21.0
<b>Res. 1</b>	na	50.1	39.2	136.9	154.1	Ph 128.3 <sup>1</sup> C 18.9
<b>Res. 2</b>	na	51.1	38.8	137.0	153.9	Ph 126.0 <sup>1</sup> C 18.9
<b>Res. 3</b>	na	51.3	39.5	137.4	153.6	Ph 127.7 <sup>1</sup> C 18.9
<b>Res. 4</b>	na	58.0	30.8	136.7	154.0	Ph 128.5 <sup>1</sup> C 18.9
<b>Res. 5</b>	161.5	48.4	30.4	na	na	<sup>ε</sup> C na <sup>1</sup> C na <sup>2</sup> C na <sup>3</sup> C na <sup>4</sup> C na

na: not assigned

Table S26 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K minor conformer (14%)

<b>Residue</b>	<b>HN</b>	<b>γCH</b>	<b>δCH</b>	<b>Others</b>
<b>Ac</b>	-	-	-	CH <sub>3</sub> 2.06
<b>Res. 1</b>	9.03	4.98	2.92-3.07	Ph na <sup>1</sup> CH <sub>3</sub> na
<b>Res. 2</b>	9.98	5.79	3.12-3.50	Ph na <sup>1</sup> CH <sub>3</sub> na
<b>Res. 3</b>	10.85	5.66	3.29	Ph na <sup>1</sup> CH <sub>3</sub> na
<b>Res. 4</b>	10.70	5.51	3.44	Ph na <sup>1</sup> CH <sub>3</sub> na
<b>Res. 5</b>	9.55	5.56	2.27-2.35	<sup>ε</sup> CH <sub>2</sub> na <sup>1</sup> CH na <sup>2</sup> CH <sub>2</sub> na <sup>3</sup> CH <sub>2</sub> na <sup>4</sup> CH <sub>2</sub> na
<b>NH<sub>2</sub></b>	8.84-7.69	-	-	-

na: not assigned

Table S27 - Coupling Constants <sup>3</sup>J(NH, γCH) (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

<b>Residue</b>	<b>Major (298 K)</b>	<b>Minor (298 K)</b>
<b>Res. 1</b>	7.9	7.9

<b>Res. 2</b>	5.6	6.6
<b>Res. 3</b>	5.4	4.7
<b>Res. 4</b>	5.6	o
<b>Res. 5</b>	6.9	8.3

*o: overlaps.*

Table S28 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (*s*) < 2.7 Å, 2.7 Å < Medium (*m*) < 3.3 Å, 3.3 Å < Weak (*w*).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	w
1.Hγ-2.NH	s
2.Hγ-3.NH	s
3.Hγ-4.NH	s
4.Hγ-5.NH	s
5.Hγ-NH <sub>2</sub>	s/m
1.Hγ-2.Hγ	o
2.Hγ-3.Hγ	o
3.Hγ-4.Hγ	o
4.Hγ-5.Hγ	o
1.Hγ-2.Hτ1	m
2.Hγ-3.Hτ1	m
3.Hγ-4.Hτ1	m
4.Hγ-5.Hτ1	nd
1.Hγ-2.Hδ	nd
2.Hγ-3.Hδ	w
3.Hγ-4.Hδ	w
4.Hγ-5.Hδ	w

*o: overlap with the solvent signals, nd: not detected.*

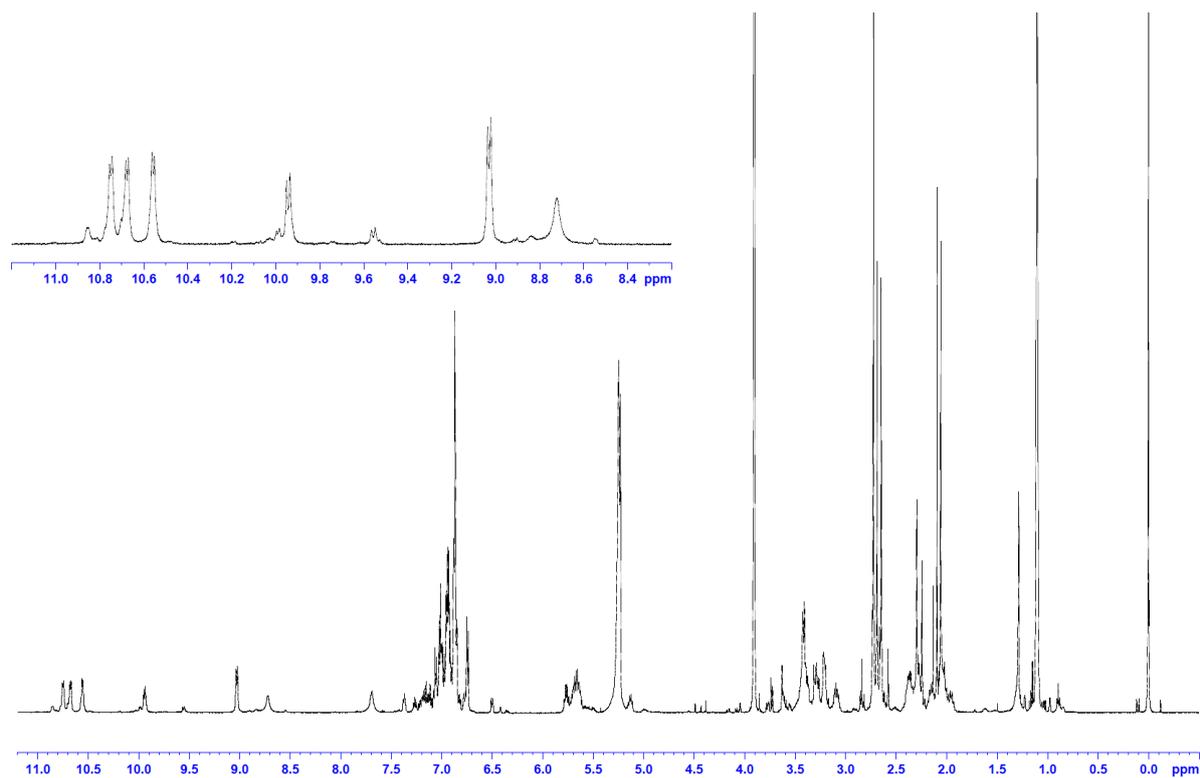


Figure S19 -  $^1\text{H}$  NMR spectrum of  $\text{ATC}_5\text{-5-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

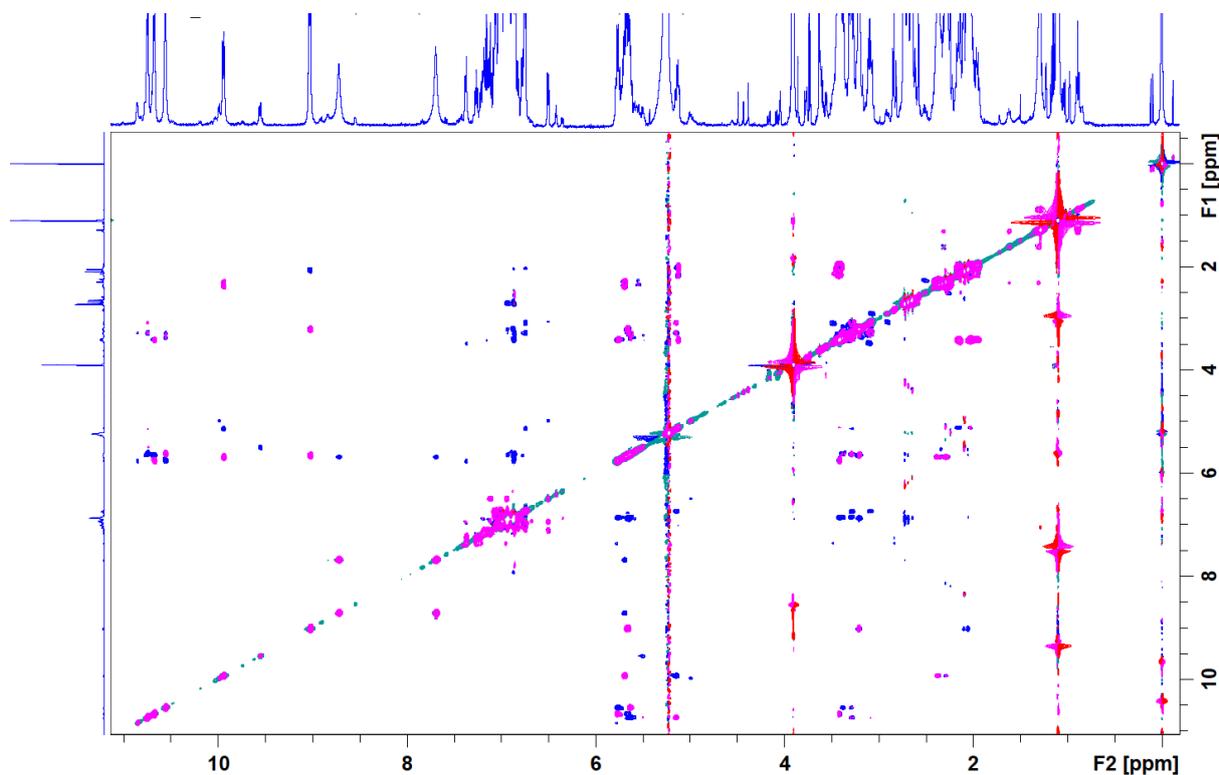
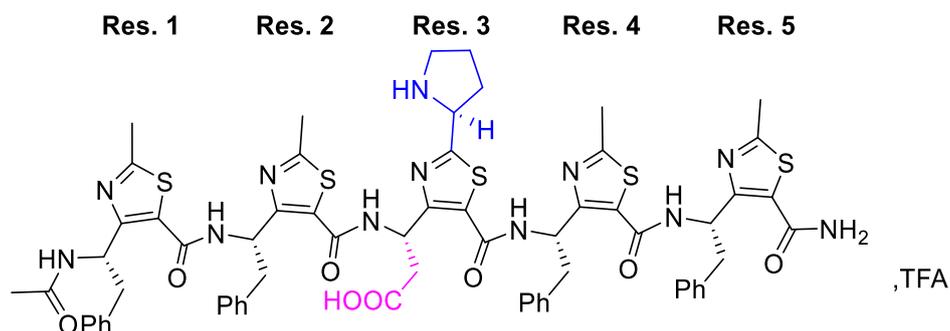


Figure S20 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of  $\text{ATC}_5\text{-5-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

10. NMR analyses of **M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>**Table S29 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (77%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 1.97
Res. 1	9.02	5.48	3.11	Ph 6.86-6.71 <sup>1</sup> CH <sub>3</sub> 2.56
Res. 2	10.82	5.08	3.16	Ph 6.98-6.55 <sup>1</sup> CH <sub>3</sub> 2.60
Res. 3	9.73	6.02	2.94-3.24	<sup>1</sup> CH na <sup>2</sup> CH <sub>2</sub> na <sup>3</sup> CH <sub>2</sub> na <sup>4</sup> CH <sub>2</sub> na
Res. 4	10.26	5.65	3.27	Ph 6.83-6.69 <sup>1</sup> CH <sub>3</sub> 2.51
Res. 5	10.28	5.54	3.21-3.37	Ph 7.13-7.05 <sup>1</sup> CH <sub>3</sub> 2.58
NH <sub>2</sub>	8.73-7.27	-	-	-

na: not assigned

Table S30 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (77%)

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	172.4					CH <sub>3</sub> 23.0
Res. 1	na	50.2	39.7	137.1	154.5	Ph 128.1-126.4-126.2 <sup>1</sup> C 18.7
Res. 2	na	51.3	39.3	137.2	153.5	Ph 128.2-126.8-126.2 <sup>1</sup> C 18.8
Res. 3	na	45.2	39.2	137.5	nd	<sup>ε</sup> C 30.6 <sup>1</sup> C 58.0 <sup>2</sup> C na <sup>3</sup> C na <sup>4</sup> C 45.2
Res. 4	na	51.6	40.0	136.7	153.8	Ph 128.3-126.1-126.2 <sup>1</sup> C 18.6
Res. 5	na	51.7	39.9	137.4	154.5	Ph 128.7-126.4-126.1 <sup>1</sup> C 19.4

na: not assigned

Table S31 -  $^1\text{H}$  NMR in  $i\text{PrOH-}d_7$  at 293K minor conformer a (12%)

<i>Residue</i>	<i>HN</i>	$\gamma\text{CH}$	$\delta\text{CH}$	<i>Others</i>
<b>Ac</b>	-	-	-	CH <sub>3</sub> 1.95
<b>Res. 1</b>	8.97	5.46	3.10	Ph na $^1\text{CH}_3$ na
<b>Res. 2</b>	10.73	5.05	3.16-3.07	Ph na $^1\text{CH}_3$ na
<b>Res. 3</b>	9.42	5.80	3.24-2.99	$^1\text{CH}$ na $^2\text{CH}_2$ na $^3\text{CH}_2$ na $^4\text{CH}_2$ na
<b>Res. 4</b>	9.99	5.23	3.14-3.07	Ph na $^1\text{CH}_3$ na
<b>Res. 5</b>	9.75	5.42	3.40-3.05	Ph na $^1\text{CH}_3$ na
<b>NH<sub>2</sub></b>	8.88-7.37	-	-	-

*na: not assigned*

Table S32 -  $^1\text{H}$  NMR in  $i\text{PrOH-}d_7$  at 293K minor conformer b (11%)

<i>Residue</i>	<i>HN</i>	$\gamma\text{CH}$	$\delta\text{CH}$	<i>Others</i>
<b>Ac</b>	-	-	-	CH <sub>3</sub> 2.02
<b>Res. 1</b>	8.95	4.83	2.97-2.77	Ph na $^1\text{CH}_3$ na
<b>Res. 2</b>	10.06	5.53	3.44-3.10	Ph na $^1\text{CH}_3$ na
<b>Res. 3</b>	10.33	5.93	3.06-2.82	$^1\text{CH}$ na $^2\text{CH}_2$ na $^3\text{CH}_2$ na $^4\text{CH}_2$ na
<b>Res. 4</b>	10.71	5.93	3.06-2.82	Ph na $^1\text{CH}_3$ na
<b>Res. 5</b>	10.17	5.53	3.22-3.32	Ph na $^1\text{CH}_3$ na
<b>NH<sub>2</sub></b>	8.88-7.37	-	-	-

*na: not assigned*

Table S33 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in  $i\text{-PrOH-}d_6$  at 298K.

<b>Residue</b>	<b>Major (298 K)</b>	<b>Minor a (298 K)</b>	<b>Minor b (298 K)</b>
<b>Res. 1</b>	6.8	6.7	6.7
<b>Res. 2</b>	6.2	6.6	7.4
<b>Res. 3</b>	7.9	7.9	7.1
<b>Res. 4</b>	6.1	6.8	6.3
<b>Res. 5</b>	6.8	6.6	6.8

Table S34 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	o
5.NH-NH <sub>2</sub>	nd
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	m
1.Hγ-2.NH	s
2.Hγ-3.NH	m/s
3.Hγ-4.NH	m/s
4.Hγ-5.NH	s
5.Hγ-NH <sub>2</sub>	m/s
5.Hγ-NH <sub>2</sub>	w/m
1.Hγ-2.Hγ	o
2.Hγ-3.Hγ	w
3.Hγ-4.Hγ	nd
4.Hγ-5.Hγ	nd
1.Hγ-2.Hτ1	w
2.Hγ-3.Hτ1	nd
3.Hγ-4.Hτ1	w
4.Hγ-5.Hτ1	w
1.Hγ-2.Hδ	nd
2.Hγ-3.Hδ	nd
3.Hγ-4.Hδ	nd
4.Hγ-5.Hδ	nd
1.Ph-2.Hτ1	w
4.Ph-5.Hτ1	w/m

*o*: overlap with the solvent signals; *nd*: not detected

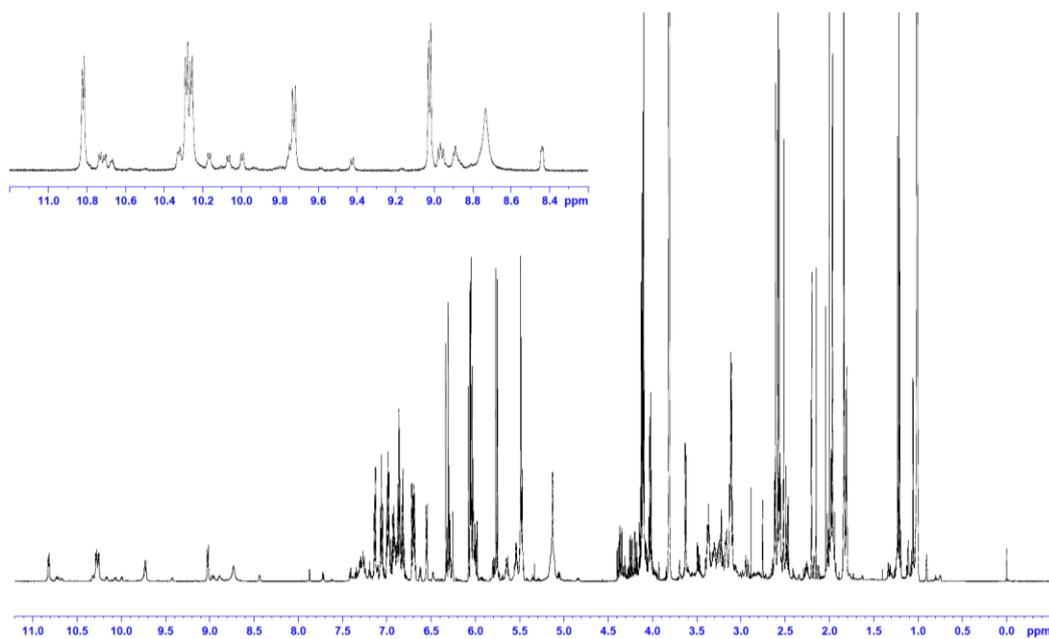


Figure S21 -  $^1\text{H}$  NMR spectrum of  $M(1)\text{-ATC}_5\text{-3-NH}_2$  in  $i\text{PrOH-}d_7$  at 298K

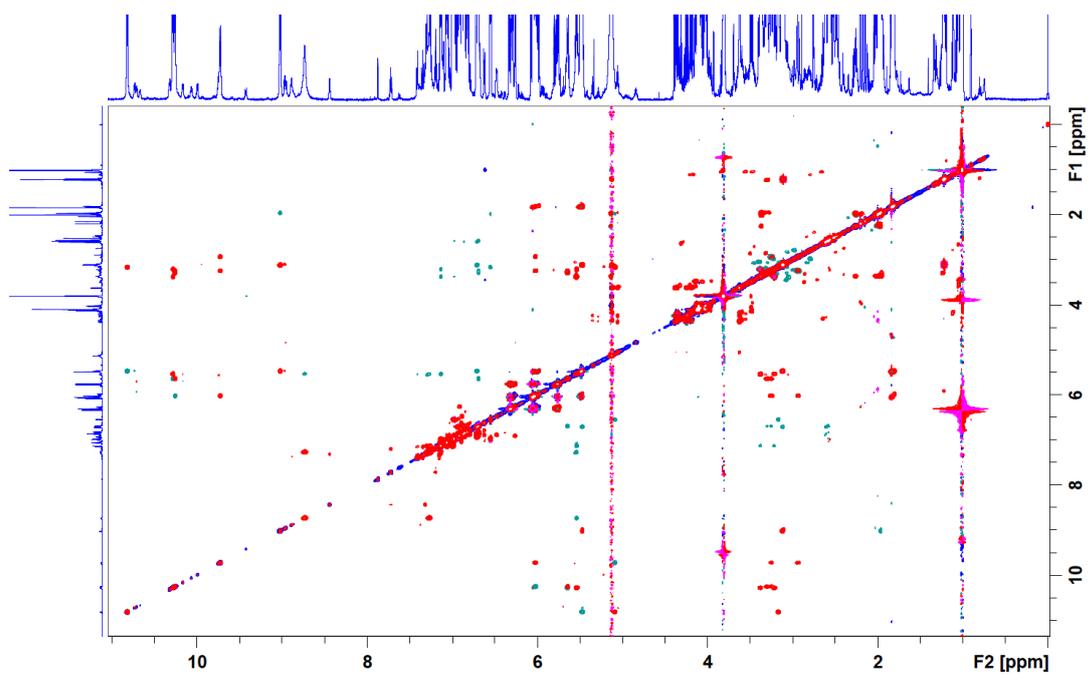
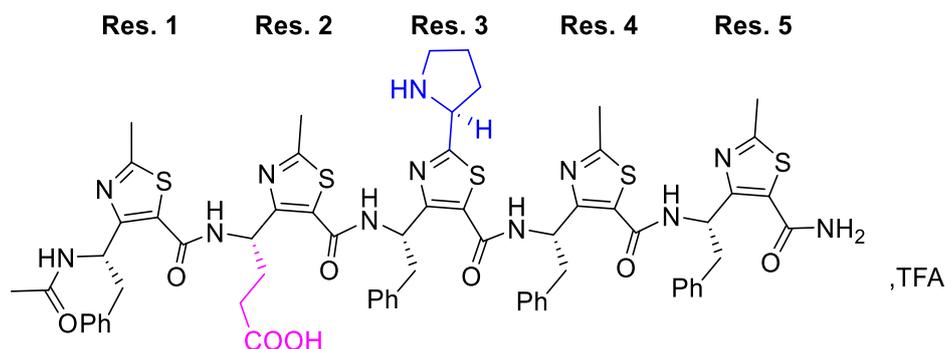


Figure S22 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of  $M(1)\text{-ATC}_5\text{-3-NH}_2$  in  $i\text{PrOH-}d_7$  at 298K

11. NMR analyses of **M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>**Table S35 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (84%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 1.93
Res. 1	8.95	5.39	3.10	Ph 6.71 <sup>1</sup> CH <sub>3</sub> na
Res. 2	10.50	5.42	2.36-2.19	<sup>1</sup> CH <sub>3</sub> 2.60
Res. 3	10.46	5.66	3.30-3.20	Ph 6.75 <sup>1</sup> CH 5.12 <sup>2</sup> CH <sub>2</sub> na <sup>3</sup> CH <sub>2</sub> na <sup>4</sup> CH <sub>2</sub> na
Res. 4	10.58	5.62	3.33-3.19	Ph 6.90-6.86 <sup>1</sup> CH <sub>3</sub> 2.62
Res. 5	10.32	5.47	3.28-3.13	Ph 7.08-7.01 <sup>1</sup> CH <sub>3</sub> 2.59
NH <sub>2</sub>	8.77-7.26	-	-	-

na: not assigned

Table S36 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (84%)

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	172.4	-	-	-	-	CH <sub>3</sub> 21.6
Res. 1	na	50.2	39.2	na	na	Ph 128.6-128.4-126.3 <sup>1</sup> C 18.7
Res. 2	na	48.5	28.4	na	na	Ph na <sup>1</sup> C 18.8
Res. 3	na	51.5	39.5	136.8	na	<sup>2</sup> C 29.8 <sup>1</sup> C na <sup>2</sup> C na <sup>3</sup> C na <sup>4</sup> C na
Res. 4	na	51.2	39.4	136.9	na	Ph 130.4-128.1-126.1 <sup>1</sup> C 18.8
Res. 5	na	52.0	39.9	137.6	na	Ph 128.8-128.1-126.3 <sup>1</sup> C 17.4

na: not assigned

Table S37 - Coupling Constants  ${}^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH- $d_6$  at 298K.

Residue	Major (298 K)
Res. 1	7.0
Res. 2	6.1
Res. 3	4.7
Res. 4	5.7
Res. 5	6.4

Table S38 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH- $d_7$  at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	o
3.NH-4.NH	w
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.H $\gamma$ -2.NH	s
2.H $\gamma$ -3.NH	s
3.H $\gamma$ -4.NH	s
4.H $\gamma$ -5.NH	s
5.H $\gamma$ -NH <sub>2</sub>	m
5.H $\gamma$ -NH <sub>2</sub>	m
1.H $\gamma$ -2.H $\gamma$	o
2.H $\gamma$ -3.H $\gamma$	o
3.H $\gamma$ -4.H $\gamma$	o
4.H $\gamma$ -5.H $\gamma$	o
1.H $\gamma$ -2.H $\tau$ 1	w
2.H $\gamma$ -3.H $\tau$ 1	nd
3.H $\gamma$ -4.H $\tau$ 1	w
4.H $\gamma$ -5.H $\tau$ 1	w
1.H $\gamma$ -2.H $\delta$	m
2.H $\gamma$ -3.H $\delta$	w
3.H $\gamma$ -4.H $\delta$	m

4.H $\gamma$ -5.H $\delta$	m
1.Ph-2.H $\tau$ 1	w
4.Ph-5.H $\tau$ 1	w

*o*: overlap with the solvent signals; *nd*: not detected

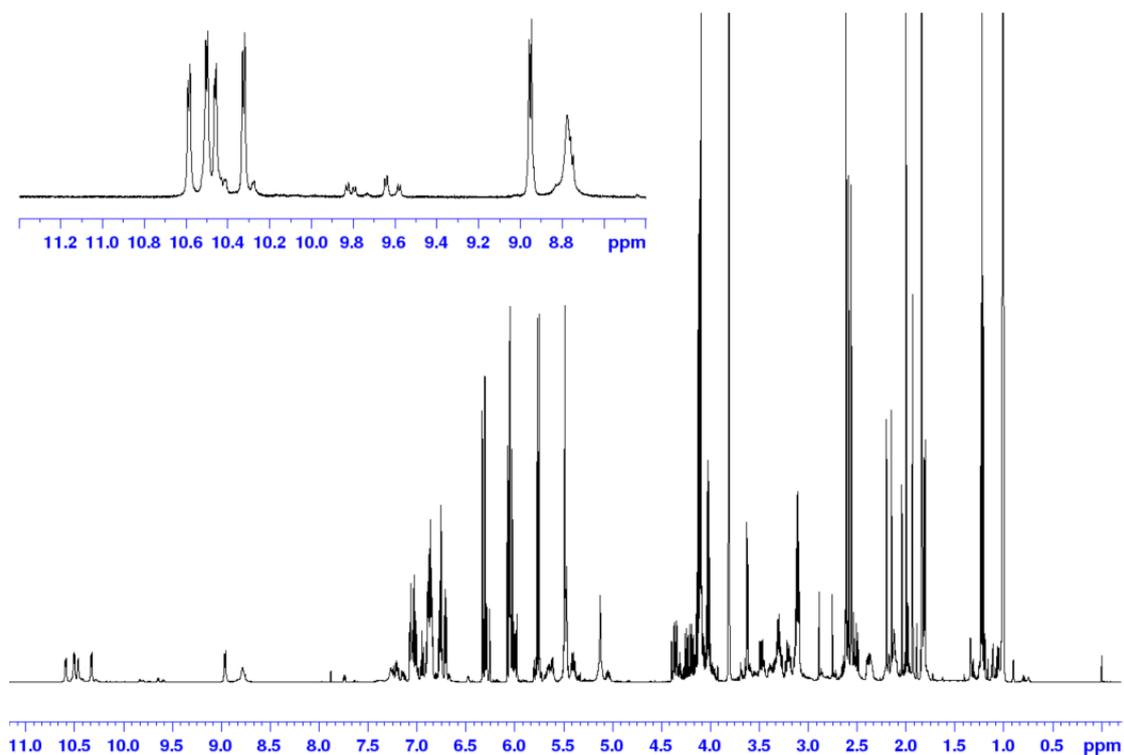


Figure S23 -  $^1\text{H}$  NMR spectrum of **M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

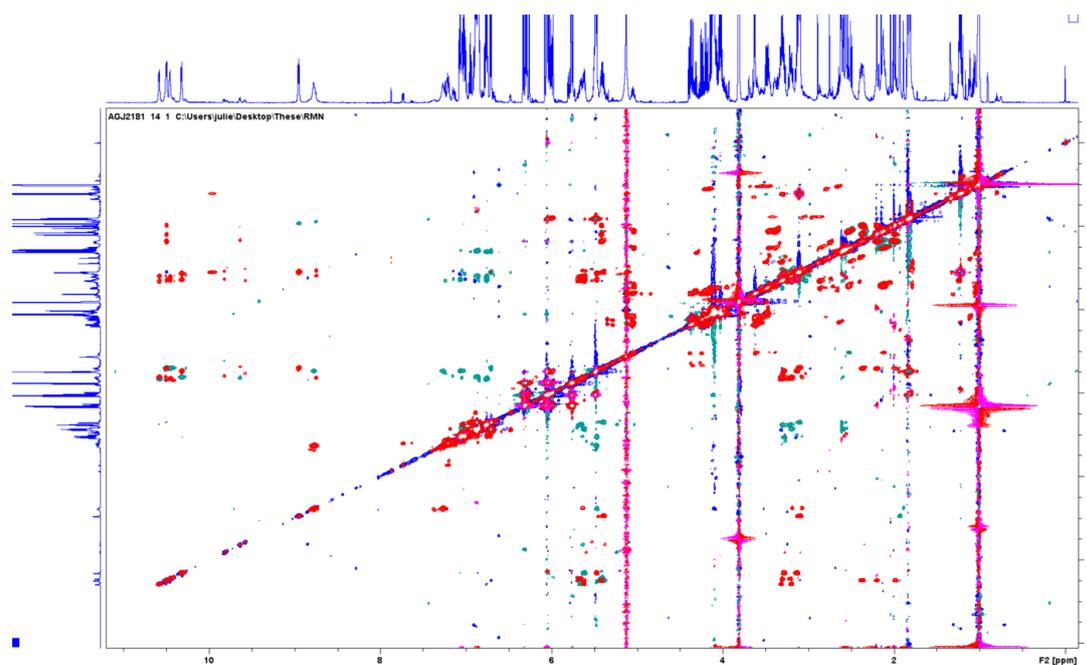
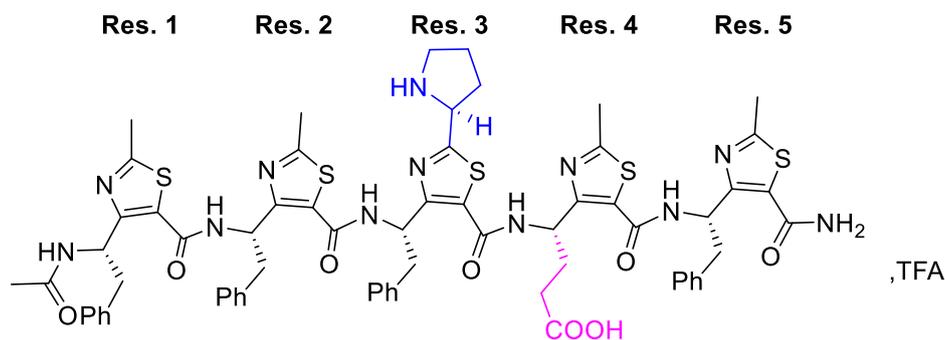


Figure S24 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

12. NMR analyses of **M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>**Table S39 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (89%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 1.98
Res. 1	9.02	5.49	3.12	Ph 6.69 <sup>1</sup> CH <sub>3</sub> na
Res. 2	10.73	5.17	3.18	Ph 6.58 <sup>1</sup> CH <sub>3</sub> 2.57
Res. 3	9.88	5.75	3.32-3.14	Ph 6.92 <sup>1</sup> CH 5.17 <sup>2</sup> CH <sub>2</sub> 2.03 <sup>3</sup> CH <sub>2</sub> na <sup>4</sup> CH <sub>2</sub> na
Res. 4	10.08	5.50	2.33-2.11	<sup>1</sup> CH <sub>3</sub> 2.56
Res. 5	10.24	5.58	3.31-3.21	Ph 7.15-7.06-7.02 <sup>1</sup> CH <sub>3</sub> 2.62
NH <sub>2</sub>	8.54-7.26	-	-	-

na: not assigned

Table S40 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (89%)

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	171.9	-	-	-	-	CH <sub>3</sub> 29.1
Res. 1	161.3	50.2	39.5	137.1	154.5	Ph 128.2 <sup>1</sup> C 18.1
Res. 2	161.5	51.6	39.5	142.8		Ph 128.2 <sup>1</sup> C 19.6
Res. 3	165.4	50.5	39.3	137.0	152.4	Ph 128.3 <sup>1</sup> C na <sup>2</sup> C 31.5 <sup>3</sup> C na <sup>4</sup> C na
Res. 4	166.8	50.2	28.6	137.2	154.4	<sup>1</sup> C 20.2
Res. 5	na	52.3	39.8	137.6	154.8	Ph 129.9-127.3-127.1 <sup>1</sup> C 19.6

na: not assigned

Table S41 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	Major (298 K)
Res. 1	6.7
Res. 2	6.3
Res. 3	7.6
Res. 4	5.6
Res. 5	6.1

Table S42 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.H $\gamma$ -2.NH	s
2.H $\gamma$ -3.NH	m
3.H $\gamma$ -4.NH	s
4.H $\gamma$ -5.NH	s
5.H $\gamma$ -NH <sub>2</sub>	s
5.H $\gamma$ -NH <sub>2</sub>	m
1.H $\gamma$ -2.H $\gamma$	nd
2.H $\gamma$ -3.H $\gamma$	nd
3.H $\gamma$ -4.H $\gamma$	nd
4.H $\gamma$ -5.H $\gamma$	nd
1.H $\gamma$ -2.H $\tau$ 1	nd
2.H $\gamma$ -3.H $\tau$ 1	nd
3.H $\gamma$ -4.H $\tau$ 1	w
4.H $\gamma$ -5.H $\tau$ 1	w
1.H $\gamma$ -2.H $\delta$	w
2.H $\gamma$ -3.H $\delta$	nd
3.H $\gamma$ -4.H $\delta$	m

4.H $\gamma$ -5.H $\delta$	w
1.Ph-2.H $\tau$ 1	m
3.Ph-4.H $\tau$ 1	m

*o*: overlap with the solvent signals; *nd*: not detected

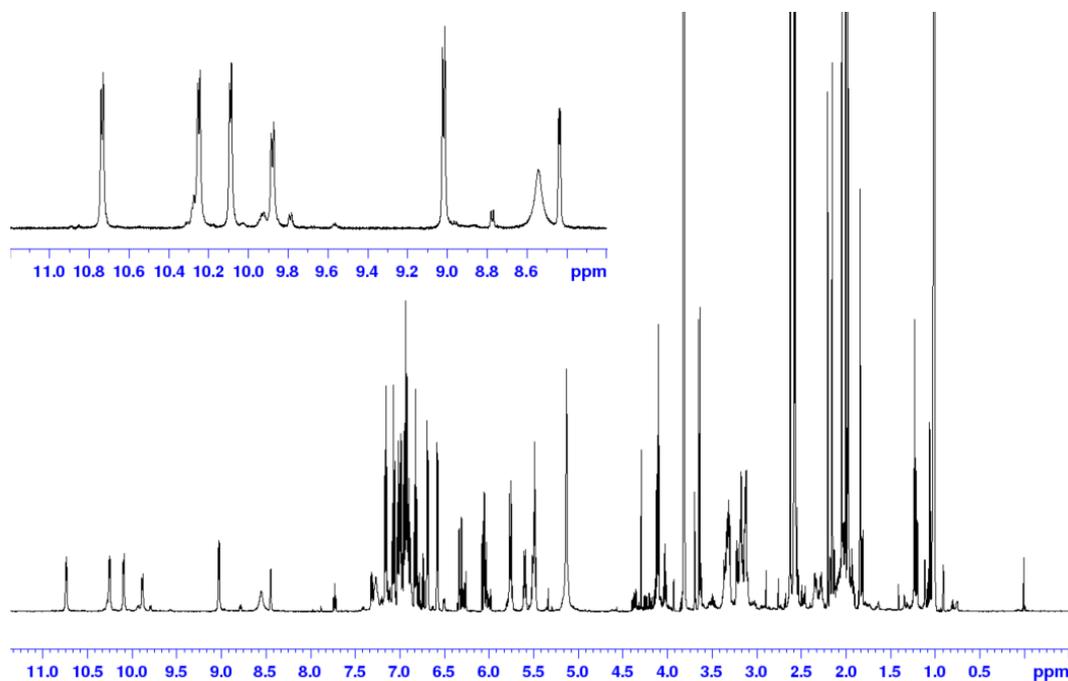


Figure S25 -  $^1\text{H}$  NMR spectrum of **M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

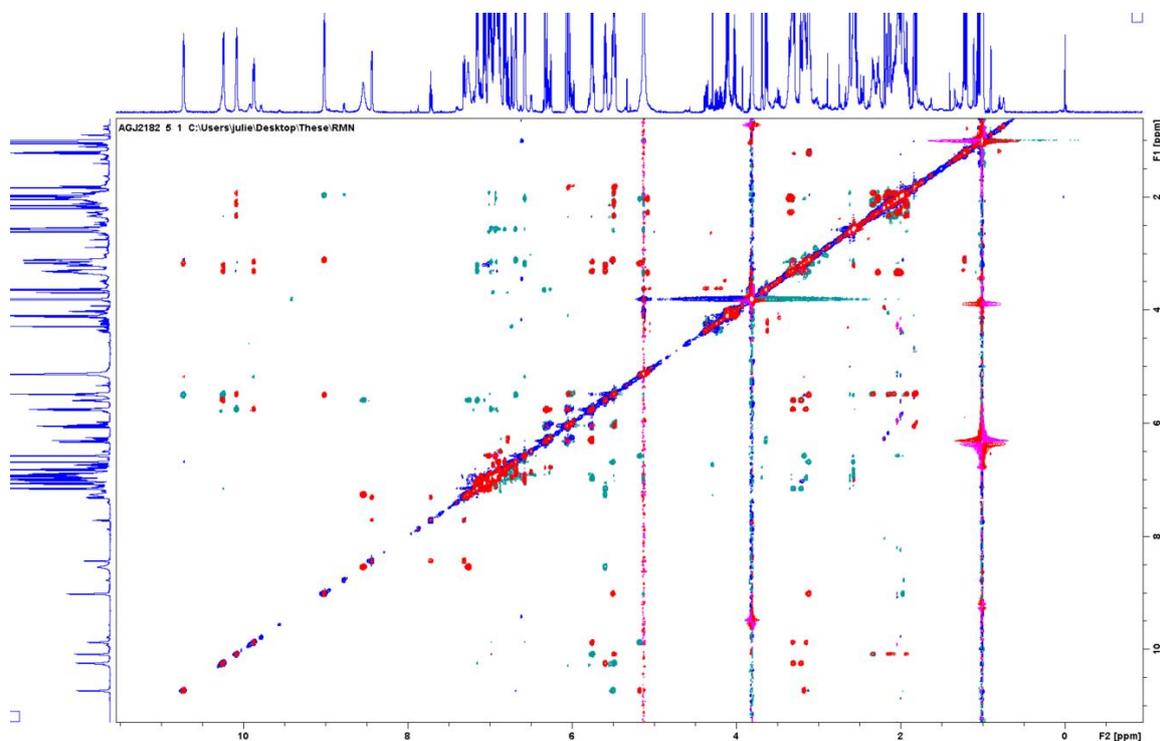
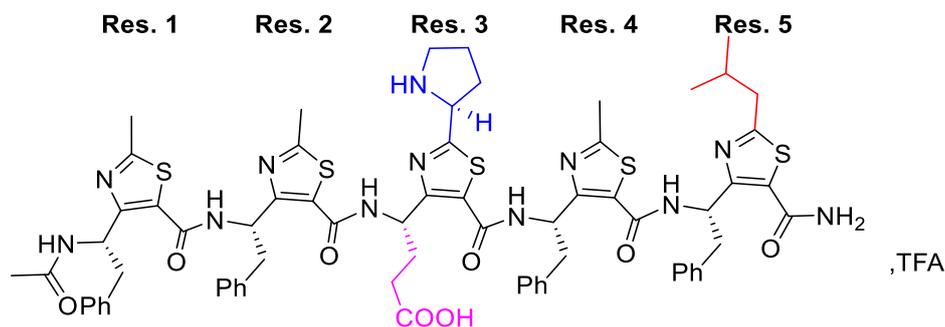


Figure S26 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

13. NMR analyses of **M(4)-ATC<sub>5</sub>-3-NH<sub>2</sub>**Table S43 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (75%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.06
Res. 1	9.07	5.57	3.19	Ph 6.79 <sup>1</sup> CH <sub>3</sub> 2.62
Res. 2	10.79	5.31	3.28	Ph na <sup>1</sup> CH <sub>3</sub> 2.60
Res. 3	9.99	5.75	2.38-2.33	<sup>1</sup> CH na <sup>2</sup> CH <sub>2</sub> na <sup>3</sup> CH <sub>2</sub> na <sup>4</sup> CH <sub>2</sub> na
Res. 4	10.27	5.65	3.26	Ph na <sup>1</sup> CH <sub>3</sub> 2.51
Res. 5	10.18	5.71	3.40	Ph 7.16 <sup>1</sup> CH <sub>3</sub> 2.56 <sup>2</sup> CH <sub>2</sub> na <sup>3</sup> CH <sub>2</sub> na <sup>4</sup> CH <sub>2</sub> na
NH <sub>2</sub>	8.72-7.35	-	-	-

na: not assigned

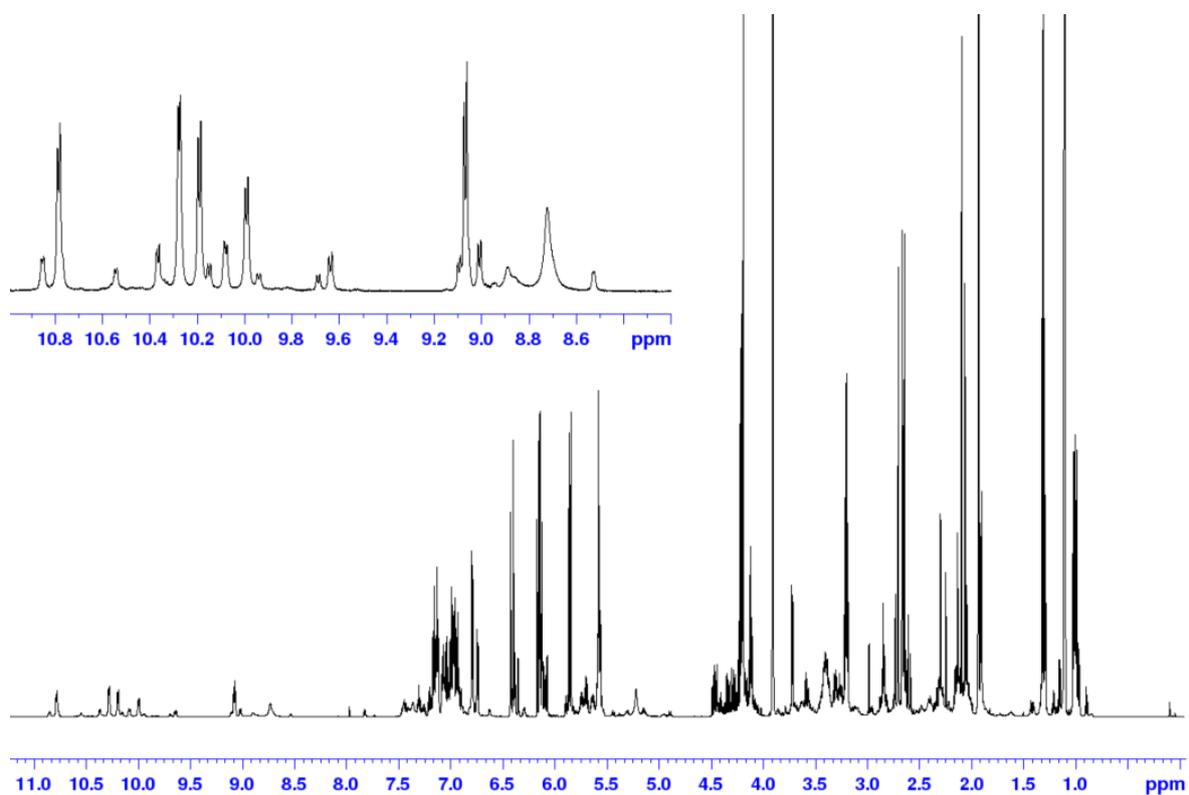
Table S44 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (75%)

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	172.4					CH <sub>3</sub> 21.5
Res. 1	na	52.1	39.3	137.4	152.9	Ph 128.5 <sup>1</sup> C 18.6
Res. 2	na	53.1	40.2	na	154.2	Ph na <sup>1</sup> C 19.0
Res. 3	na	48.9	29.1	na	na	<sup>1</sup> C 58.0 <sup>2</sup> C na <sup>3</sup> C na <sup>4</sup> C 45.2
Res. 4	na	49.2	na	na	153.8	Ph na <sup>1</sup> C 18.6
Res. 5	na	51.4	39.9	140.2	154.5	Ph 128.31 <sup>1</sup> C 18.5 <sup>2</sup> C na <sup>3</sup> C na <sup>4</sup> C na

na: not assigned

Table S45 - Coupling Constants  ${}^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH- $d_6$  at 298K.

Residue	Major (298 K)
Res. 1	6.8
Res. 2	6.3
Res. 3	7.4
Res. 4	5.7
Res. 5	6.8

Figure S27 -  ${}^1\text{H}$  NMR spectrum of **M(4)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *i*PrOH- $d_7$  at 298K

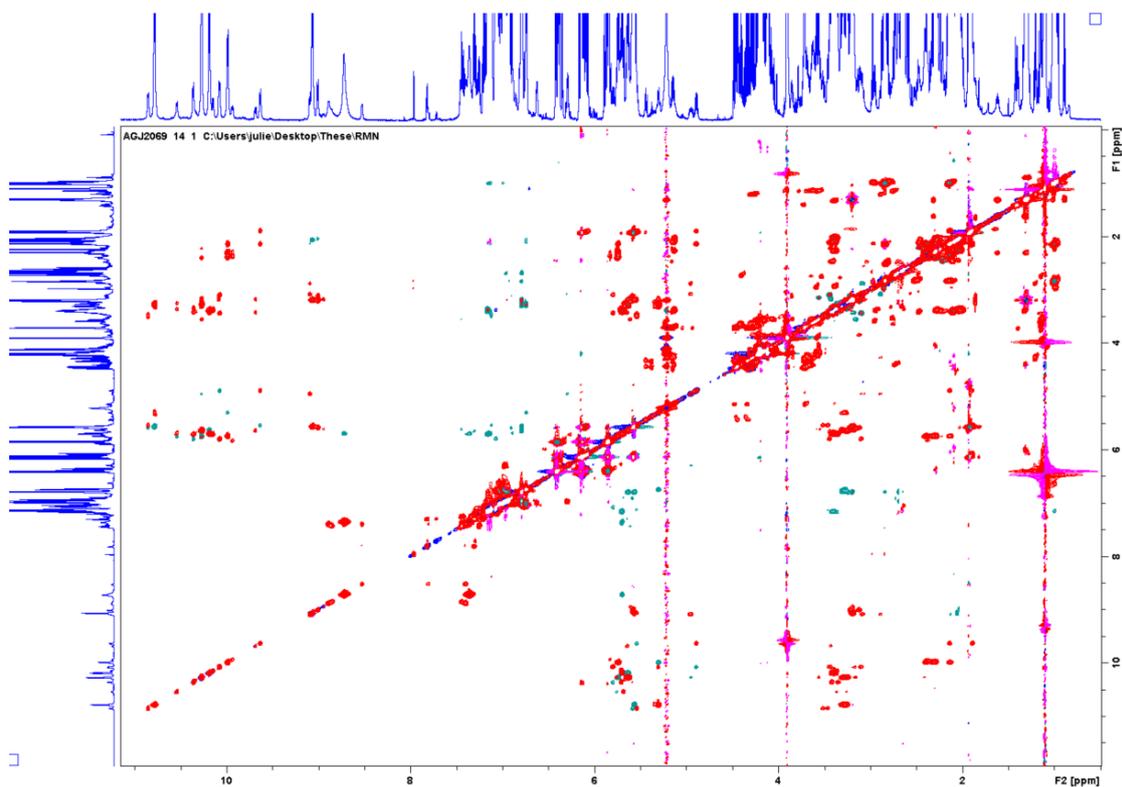


Figure S28 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(4)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *i*PrOH-*d*<sub>7</sub> at 298K

#### 14. NMR analyses of **M(5)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

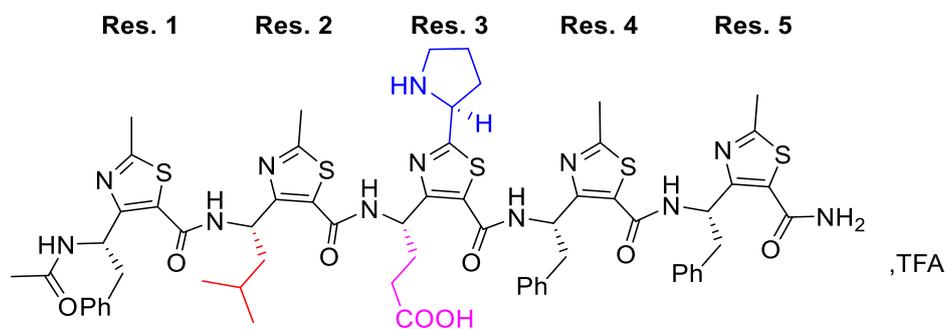


Table S46 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (85%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.01
Res. 1	9.03	5.46	3.17	Ph 6.75 <sup>1</sup> CH <sub>3</sub> 2.64
Res. 2	10.49	5.08	1.68	<sup>6</sup> CH 1.09 <sup>5</sup> CH <sub>3</sub> 0.78-0.44 <sup>1</sup> CH <sub>3</sub> 2.72
Res. 3	10.30	5.79	2.32-2.43	<sup>6</sup> CH <sub>2</sub> 2.16 <sup>1</sup> CH 5.28 <sup>2</sup> CH <sub>2</sub> 2.24-2.31 <sup>3</sup> CH <sub>2</sub> 2.16-2.14 <sup>4</sup> CH <sub>2</sub> 3.51

<b>Res. 4</b>	10.36	5.72	3.56-3.18	Ph na $^1\text{CH}_3$ 2.68
<b>Res. 5</b>	10.11	5.62	3.30-3.38	Ph na $^1\text{CH}_3$ 2.69
<b>NH<sub>2</sub></b>	8.80-7.40	-	-	-

na: not assigned

Table S47 -  $^{13}\text{C}$  NMR in *iPrOH-d<sub>7</sub>* at 293K major conformer (85%)

<b>Residue</b>	<b>C(O)</b>	<b><math>\gamma\text{C}</math></b>	<b><math>\delta\text{C}</math></b>	<b><math>\beta\text{C}</math></b>	<b><math>\alpha\text{C}</math></b>	<b>Others</b>
<b>Ac</b>	172.1	-	-	-	-	CH <sub>3</sub> 18.7
<b>Res. 1</b>	na	51.8	39.2	137.0	154.0	Ph 128.7 $^1\text{C}$ 18.8
<b>Res. 2</b>	na	48.0	41.6	na	154.1	$^6\text{C}$ na $^7\text{C}$ 22.6 $^{13}\text{C}$ na $^{14}\text{C}$ na
<b>Res. 3</b>	174.5	49.2	29.0	137.0	na	$^6\text{C}$ 23.2 $^1\text{C}$ 58.1 $^2\text{C}$ 29.0 $^{13}\text{C}$ 23.2 $^{14}\text{C}$ 45.1
<b>Res. 4</b>	na	52.7	41.1	136.6	153.8	Ph na $^1\text{C}$ 18.6
<b>Res. 5</b>	na	51.4	40.3	137.0	154.5	Ph na $^1\text{C}$ 18.9

na: not assigned

Table S48 -  $^1\text{H}$  NMR in *iPrOH-d<sub>7</sub>* at 298K minor conformer (15%)

<b>Residue</b>	<b>HN</b>	<b><math>\gamma\text{CH}</math></b>	<b><math>\delta\text{CH}</math></b>	<b>Others</b>
<b>Ac</b>	-	-	-	CH <sub>3</sub> 2.01
<b>Res. 1</b>	8.97	5.54	3.11-3.21	Ph na $^1\text{CH}_3$ na
<b>Res. 2</b>	10.68	5.41	2.12-2.05	$^6\text{CH}$ 1.47 $^7\text{CH}_3$ 1.01 $^1\text{CH}_3$ na
<b>Res. 3</b>	10.21	5.15	2.16-2.25	$^6\text{CH}_2$ 1.88 $^1\text{CH}_3$ na $^2\text{CH}_2$ na $^{13}\text{CH}_2$ na $^{14}\text{CH}_2$ na
<b>Res. 4</b>	10.15	5.85	2.18-3.56	Ph na $^1\text{CH}_3$ na
<b>Res. 5</b>	10.41	5.69	3.29-3.44	Ph na $^1\text{CH}_3$ na
<b>NH<sub>2</sub></b>	8.90-7.43	-	-	-

na: not assigned

Table S49 - Coupling Constants  $^3\text{J}(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i-PrOH-d<sub>6</sub>* at 298K.

<b>Residue</b>	<b>Major (298 K)</b>	<b>Minor (298 K)</b>
<b>Res. 1</b>	6.6	6.7

<b>Res. 2</b>	6.0	5.7
<b>Res. 3</b>	7.1	7.7
<b>Res. 4</b>	5.7	6.7
<b>Res. 5</b>	6.7	6.7

Table S50 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	nd
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
5.NH-NH <sub>2</sub>	s/m
Ac-1.NH	m
1.Hγ-2.NH	s
2.Hγ-3.NH	s
3.Hγ-4.NH	s
4.Hγ-5.NH	m
5.Hγ-NH <sub>2</sub>	o
5.Hγ-NH <sub>2</sub>	o
1.Hγ-2.Hγ	o
2.Hγ-3.Hγ	o
3.Hγ-4.Hγ	w
4.Hγ-5.Hγ	w
1.Hγ-2.Hτ1	w
2.Hγ-3.Hτ1	w
3.Hγ-4.Hτ1	nd
4.Hγ-5.Hτ1	nd
1.Hγ-2.Hδ	w
2.Hγ-3.Hδ	nd
3.Hγ-4.Hδ	w
4.Hγ-5.Hδ	w
1.Hδ-2.Hτ1	w
2.Hδ-3.Hτ1	nd
3.Hδ-4.Hτ1	w

4.H $\delta$ -5.H $\tau$ 1	w
1.H $\delta$ -2.NH	w
2.H $\kappa$ -3.H $\gamma$	w
2.H $\kappa$ -4.NH	w
2.H $\kappa$ -3.H $\tau$ 2	w

*o*: overlap with the solvent signals; *nd*: not detected

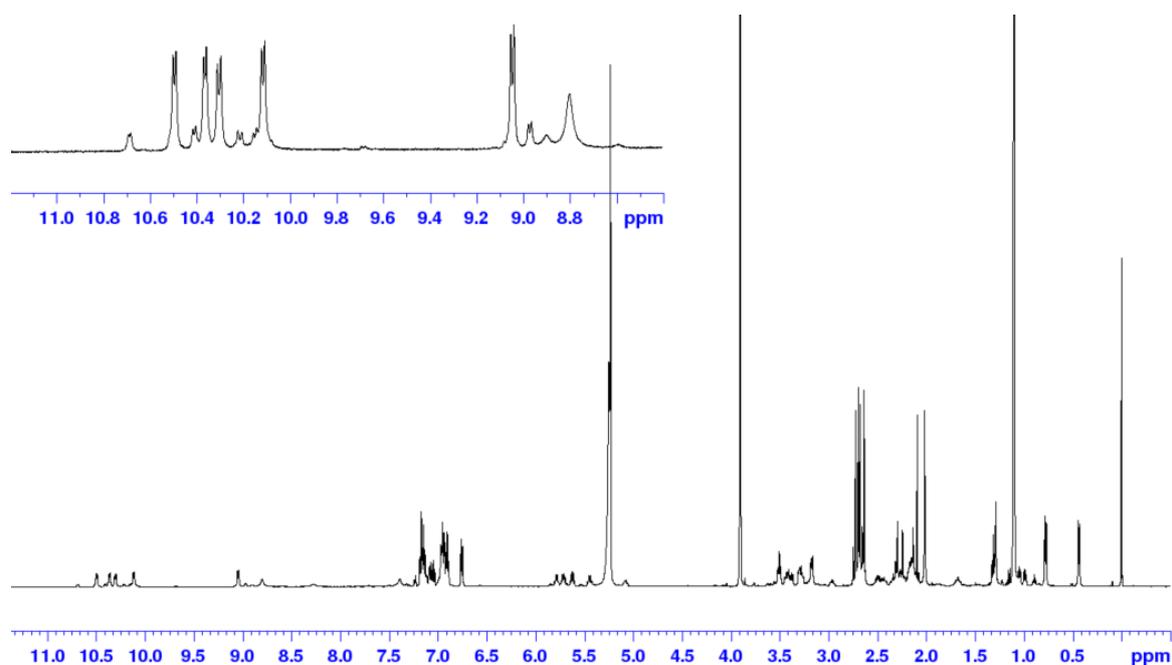


Figure S29 -  $^1\text{H}$  NMR spectrum of  $M(5)\text{-ATC}_5\text{-3-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

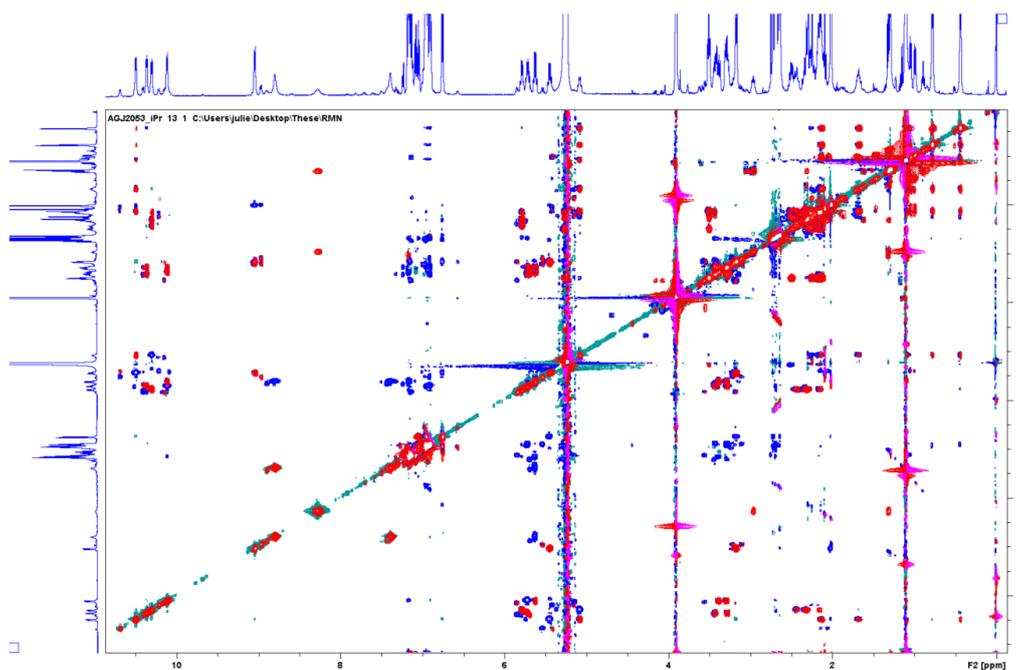


Figure S30 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of  $M(5)\text{-ATC}_5\text{-3-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

### 15. NMR analyses of **M(6)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

The yield obtained for **M(6)-ATC<sub>5</sub>-3-NH<sub>2</sub>** was too small to perform NMR analysis. We used all the compound for the nitro-Michael addition reaction tests.

### 16. NMR analyses of **M(7)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

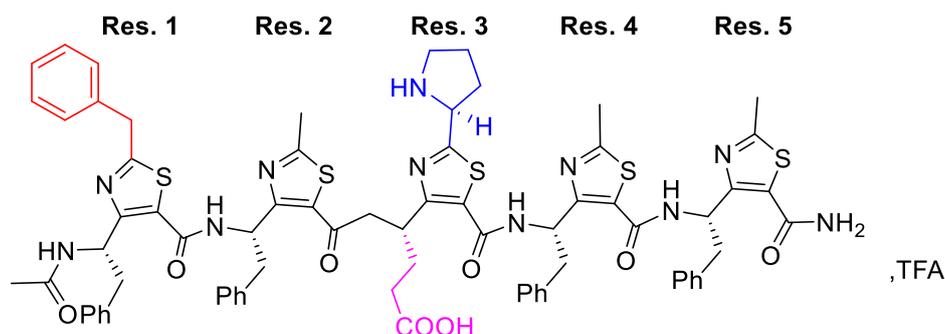


Table S51 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (84%)

Residue	HN	<sup>γ</sup> CH	<sup>δ</sup> CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.00
Res. 1	9.10	5.46	3.11	Ph 6.57 <sup>τ</sup> 1CH <sub>2</sub> na <sup>τ</sup> 3CH na <sup>τ</sup> 4CH na <sup>τ</sup> 5CH na
Res. 2	10.72	5.10	3.18	Ph 6.64 <sup>τ</sup> 1CH <sub>3</sub> na
Res. 3	9.72	5.69	2.23	<sup>ε</sup> CH <sub>2</sub> na <sup>τ</sup> 1CH na <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na
Res. 4	10.02	5.60	3.20	Ph 6.72 <sup>τ</sup> 1CH <sub>3</sub> na
Res. 5	10.09	5.51	3.25	Ph 7.08 <sup>τ</sup> 1CH <sub>3</sub> 2.13
NH <sub>2</sub>	8.71-7.28	-	-	-

na: not assigned

Table S52 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K minor conformer a (11%)

Residue	HN	<sup>γ</sup> CH	<sup>δ</sup> CH	Others
Ac	-	-	-	CH <sub>3</sub> 1.99
Res. 1	9.05	5.51	3.13	Ph na <sup>τ</sup> 1CH <sub>3</sub> na
Res. 2	10.63		3.16	<sup>ε</sup> CH <sub>2</sub> na <sup>ζ</sup> CH <sub>3</sub> na <sup>τ</sup> 1CH <sub>3</sub> na
Res. 3	9.54	5.29	3.11	<sup>ε</sup> CH <sub>2</sub> na <sup>τ</sup> 1CH <sub>3</sub> na <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na
Res. 4	9.76	5.65	3.45-3.07	Ph na <sup>τ</sup> 1CH <sub>3</sub> na

<b>Res. 5</b>	10.48	5.66	3.33-3.19	Ph na $^{13}\text{C}$ na
<b>NH<sub>2</sub></b>	8.83-7.38	-	-	-

na: not assigned

Table S53 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	Major (298 K)	Minor a (298 K)
<b>Res. 1</b>	6.2	6.2
<b>Res. 2</b>	6.5	6.3
<b>Res. 3</b>	7.9	7.9
<b>Res. 4</b>	5.8	7.7
<b>Res. 5</b>	6.6	6.3

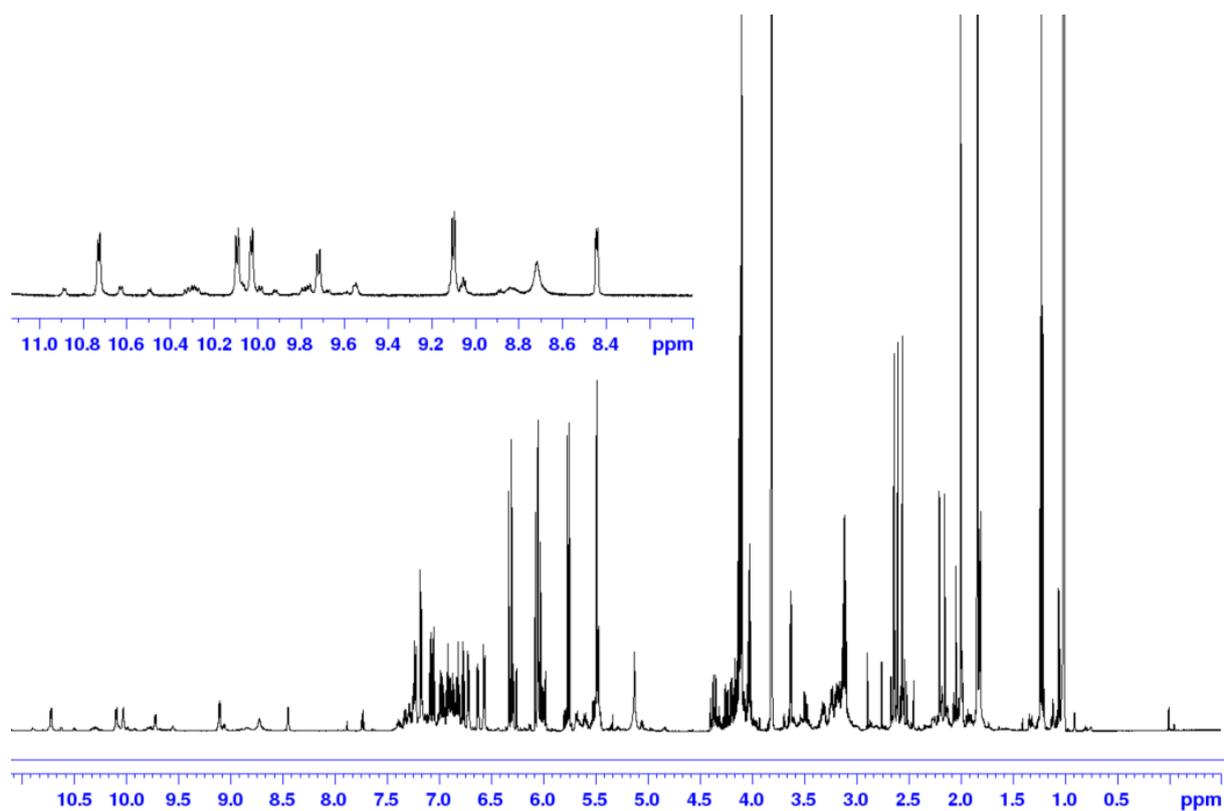


Figure S31 -  $^1\text{H}$  NMR spectrum of **M(7)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *i*PrOH-*d*<sub>7</sub> at 298K

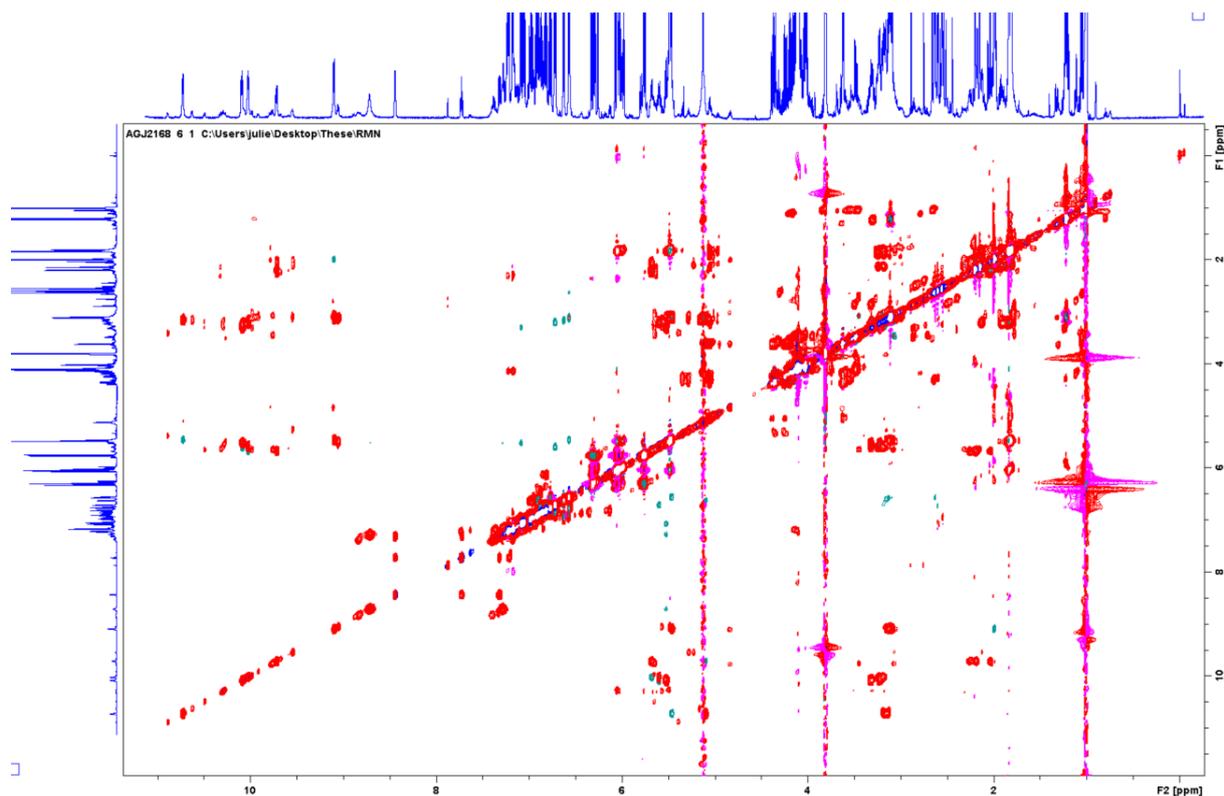


Figure S32 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(7)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

### 17. NMR analyses of **M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

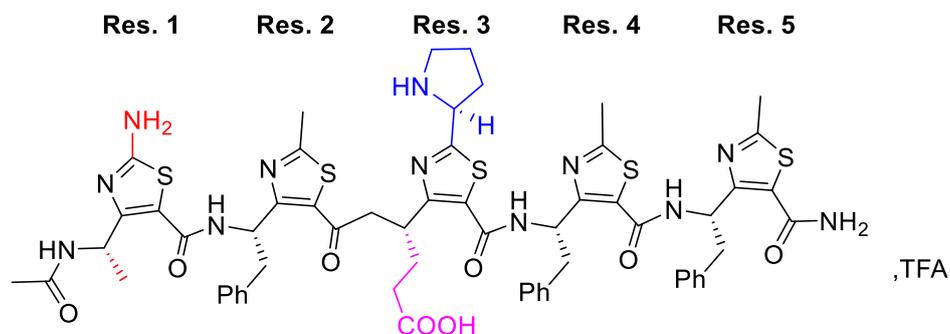


Table S54 - <sup>1</sup>H NMR in *iPrOH-d<sub>7</sub>* at 293K major conformer (84%)

Residue	HN	<sup>γ</sup> CH	<sup>δ</sup> CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.03
Res. 1	8.43	5.43	1.32	NH <sub>2</sub>
Res. 2	10.33	5.23	3.11	Ph 6.93-6.57 <sup>τ</sup> 1CH <sub>3</sub> 1.84
Res. 3	9.79	5.67	2.31-2.23	<sup>ε</sup> CH <sub>2</sub> na <sup>τ</sup> 1CH 5.12 <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na

<b>Res. 4</b>	10.09	5.60	3.19-3.16	Ph 6.82-6.72 <sup>13</sup> C 2.57
<b>Res. 5</b>	10.07	5.53	3.32-3.23	Ph 7.10-6.99 <sup>13</sup> C 2.62
<b>NH<sub>2</sub></b>	8.71-7.29	-	-	-

na: not assigned

Table S55 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 293K major conformer (84%)

<b>Residue</b>	<b>C(O)</b>	<b><math>\gamma</math>C</b>	<b><math>\delta</math>C</b>	<b><math>\beta</math>C</b>	<b><math>\alpha</math>C</b>	<b>Others</b>
<b>Ac</b>	171.2	-	-	-	-	CH <sub>3</sub> 21.2
<b>Res. 1</b>	163.3	43.7	18.9	na	na	-
<b>Res. 2</b>	161.7	na	45.5	na	na	Ph na <sup>13</sup> C 17.1
<b>Res. 3</b>	160.2	48.4	28.6	na	na	<sup>13</sup> C na
<b>Res. 4</b>	161.4	51.4	38.9	na	na	Ph 127.8 <sup>13</sup> C 18.1
<b>Res. 5</b>	na	50.9	39.5	na	na	Ph 128.2-125.7 <sup>13</sup> C 18.3

na: not assigned

Table S56 - Coupling Constants <sup>3</sup>J(NH,  $\gamma$ CH) (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

<b>Residue</b>	<b>Major (298 K)</b>
<b>Res. 1</b>	6.7
<b>Res. 2</b>	5.9
<b>Res. 3</b>	7.8
<b>Res. 4</b>	6.2
<b>Res. 5</b>	7.8

Table S57 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (*s*) < 2.7 Å, 2.7 Å < Medium (*m*) < 3.3 Å, 3.3 Å < Weak (*w*).

<b>NOE correlations</b>	<b>Intensity</b>
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	o
5.NH-NH <sub>2</sub>	w
5.NH-NH <sub>2</sub>	s

Ac-1.NH	s
1.H $\gamma$ -2.NH	nd
2.H $\gamma$ -3.NH	s
3.H $\gamma$ -4.NH	s
4.H $\gamma$ -5.NH	s
5.H $\gamma$ -NH <sub>2</sub>	s
5.H $\gamma$ -NH <sub>2</sub>	nd
1.H $\gamma$ -2.H $\gamma$	nd
2.H $\gamma$ -3.H $\gamma$	o
3.H $\gamma$ -4.H $\gamma$	o
4.H $\gamma$ -5.H $\gamma$	m
1.H $\gamma$ -2.H $\tau$ 1	nd
2.H $\gamma$ -3.H $\tau$ 1	s
3.H $\gamma$ -4.H $\tau$ 1	w
4.H $\gamma$ -5.H $\tau$ 1	w
1.H $\gamma$ -2.H $\delta$	nd
2.H $\gamma$ -3.H $\delta$	nd
3.H $\gamma$ -4.H $\delta$	w
4.H $\gamma$ -5.H $\delta$	w

*o: overlap with the solvent signals; nd: not detected*

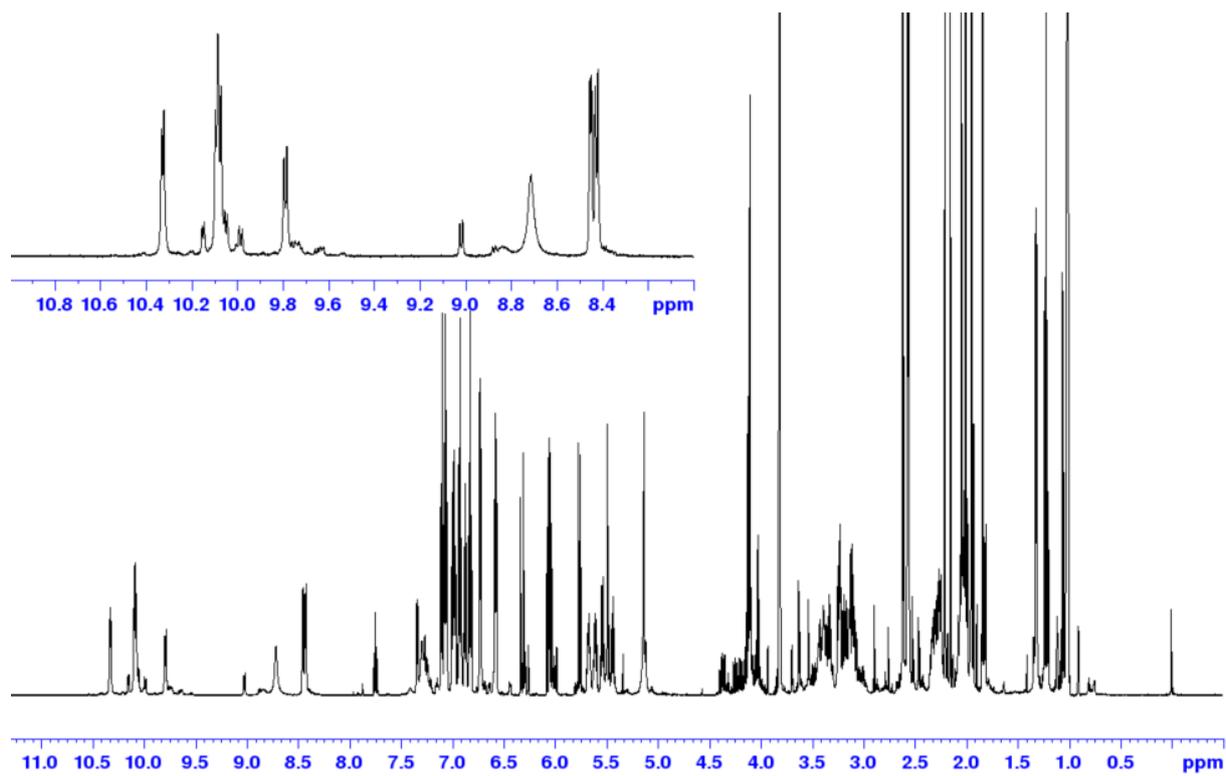


Figure S33 -  $^1\text{H}$  NMR spectrum of **M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

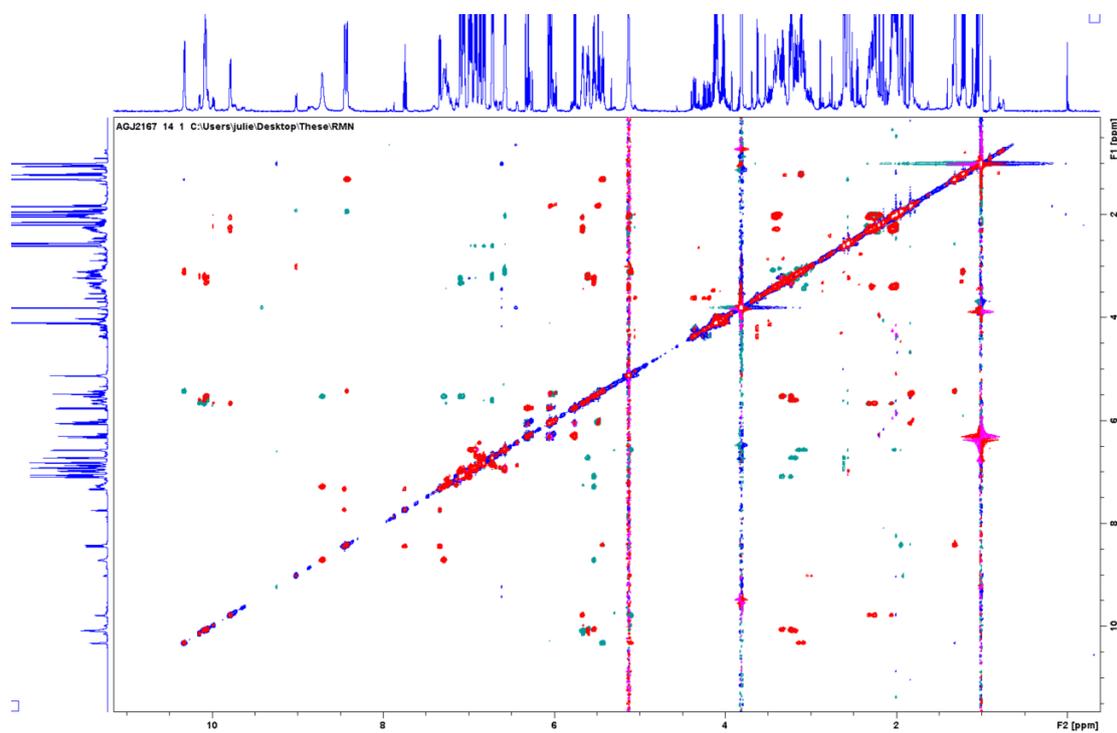
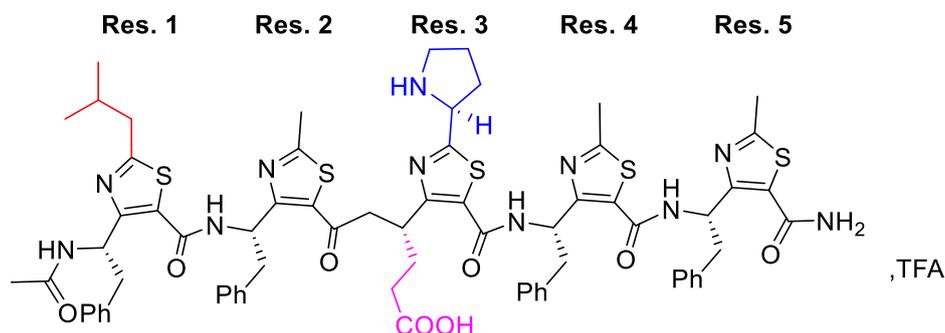


Figure S34 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

18. NMR analyses of **M(9)-ATC<sub>5</sub>-3-NH<sub>2</sub>**Table S58 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (86%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.05
Res. 1	9.17	5.56	3.19	Ph 7.01 <sup>t1</sup> CH <sub>2</sub> 2.78 <sup>t2</sup> CH <sub>2</sub> 2.10 <sup>t3</sup> CH <sub>2</sub> 0.99-0.97
Res. 2	10.85	5.25	3.26	Ph 6.90-6.71 <sup>t1</sup> CH <sub>3</sub> 2.73
Res. 3	9.92	5.76	2.29-2.38	<sup>s</sup> CH <sub>2</sub> 2.14 <sup>t1</sup> CH 5.10 <sup>t2</sup> CH <sub>2</sub> 2.14-2.30 <sup>t3</sup> CH <sub>2</sub> 2.05 <sup>t4</sup> CH <sub>2</sub> 3.41
Res. 4	10.19	5.68	3.30	Ph 6.82 <sup>t1</sup> CH <sub>3</sub> 2.66
Res. 5	10.17	5.61	3.40	Ph na <sup>t1</sup> CH <sub>3</sub> 2.70
NH <sub>2</sub>	8.59-7.48	-	-	-

*na: not assigned*Table S59 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (86%)

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	172.0	-	-	-	-	CH <sub>3</sub> 22.2
Res. 1	162.3	50.5	39.5	na	na	<sup>t1</sup> C 19.9 <sup>t2</sup> C 31.1 <sup>t3</sup> C 20.8
Res. 2	161.9	51.7	39.3	na	na	Ph na <sup>t1</sup> C 18.9
Res. 3	161.1	48.8	30.5	na	na	<sup>s</sup> C 51.6 <sup>t1</sup> C na <sup>t2</sup> C 31.5 <sup>t3</sup> C 30.8 <sup>t4</sup> C 45.4
Res. 4	161.8	52.0	39.5	na	na	Ph na <sup>t1</sup> C 18.9
Res. 5	na	51.5	45.3	na	na	Ph na <sup>t1</sup> C 18.9

*na: not assigned*

Table S60 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	Major (298 K)
Res. 1	6.6
Res. 2	6.2
Res. 3	7.7
Res. 4	5.8
Res. 5	6.8

Table S61 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (*s*) < 2.7 Å, 2.7 Å < Medium (*m*) < 3.3 Å, 3.3 Å < Weak (*w*).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	nd
3.NH-4.NH	w
4.NH-5.NH	o
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.Hγ-2.NH	s
2.Hγ-3.NH	s
3.Hγ-4.NH	s
4.Hγ-5.NH	s
5.Hγ-NH <sub>2</sub>	m
1.Hγ-2.Hγ	nd
2.Hγ-3.Hγ	nd
3.Hγ-4.Hγ	nd
4.Hγ-5.Hγ	nd
1.Hγ-2.Hτ1	m
2.Hγ-3.Hτ1	nd
3.Hγ-4.Hτ1	m
4.Hγ-5.Hτ1	m
1.Hγ-2.Hδ	w
2.Hγ-3.Hδ	nd
3.Hγ-4.Hδ	nd
4.Hγ-5.Hδ	nd

*o*: overlap with the solvent signals, *nd*: not detected

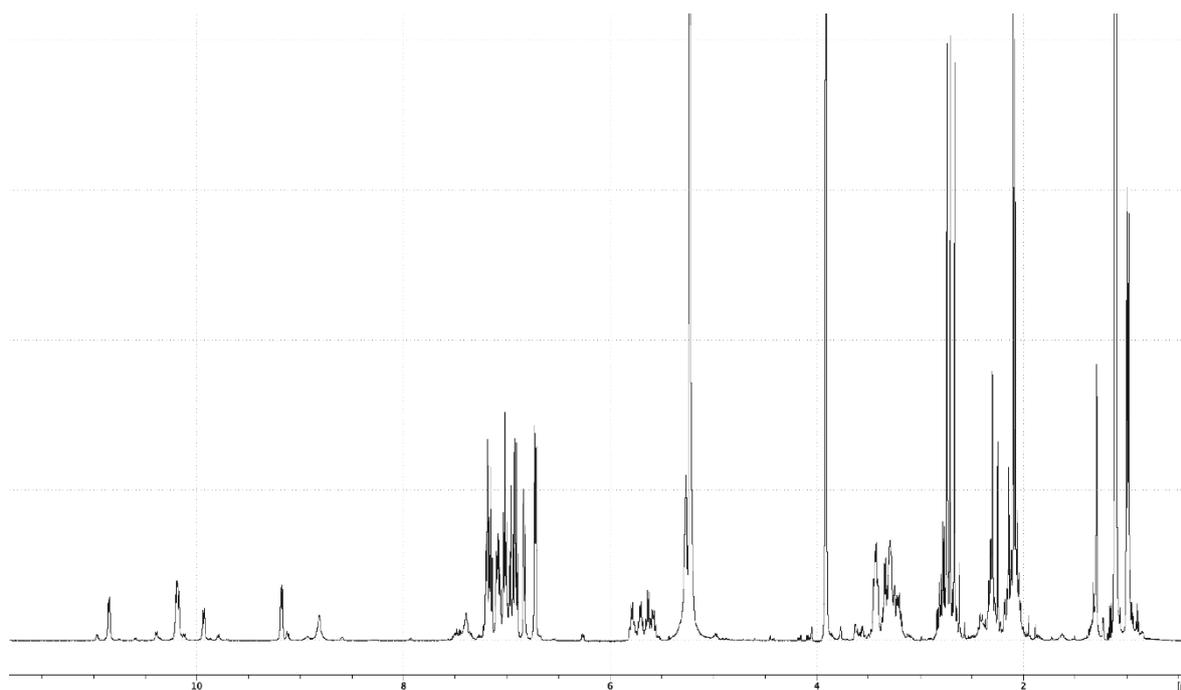


Figure S35 -  $^1\text{H}$  NMR spectrum of **M(9)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

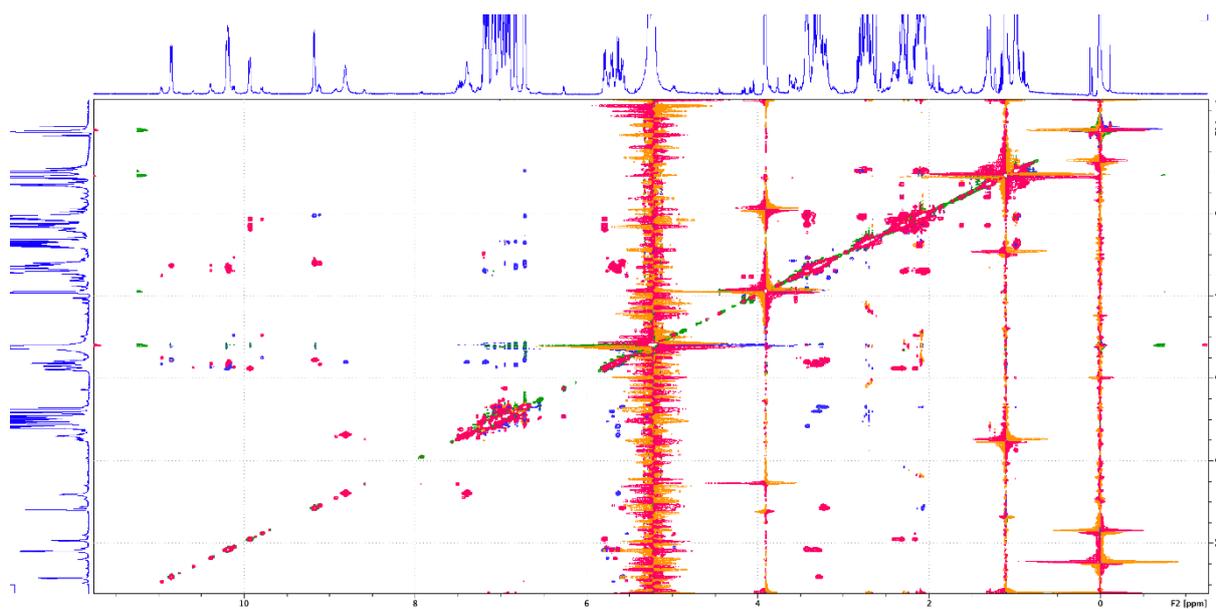
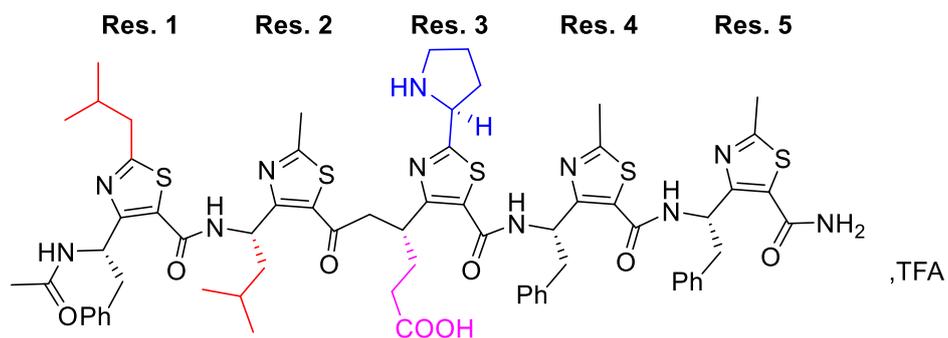


Figure S36 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(9)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

19. NMR analyses of **M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>**Table S62 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (95%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.03
Res. 1	9.12	5.47	3.21-3.18	Ph 6.72 <sup>1</sup> CH <sub>2</sub> 2.78 <sup>2</sup> CH 2.11 <sup>3</sup> CH <sub>3</sub> 0.99-0.97
Res. 2	10.51	5.07	2.14-1.68	$\epsilon$ CH 1.08 $\zeta$ CH <sub>3</sub> 0.78-0.43 <sup>1</sup> CH 2.75
Res. 3	10.27	5.80	2.42-2.32	$\epsilon$ CH <sub>2</sub> 2.42-2.32 <sup>1</sup> CH 5.27 <sup>2</sup> CH <sub>2</sub> 2.50-2.26 <sup>3</sup> CH <sub>2</sub> 2.18 <sup>4</sup> CH <sub>2</sub> 3.50
Res. 4	10.33	5.72	3.42-3.28	Ph 6.92 <sup>1</sup> CH <sub>3</sub> 2.67
Res. 5	10.10	5.63	3.38-3.30	Ph 7.15 <sup>1</sup> CH <sub>3</sub> 2.69
NH <sub>2</sub>	8.80-7.40	-	-	-

na: not assigned

Table S63 - Coupling Constants <sup>3</sup>J(NH,  $\gamma$ CH) (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	Major (298 K)
Res. 1	6.5
Res. 2	6.0
Res. 3	7.2
Res. 4	5.5
Res. 5	6.7

Table S64 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.Hγ-2.NH	s
2.Hγ-3.NH	s
3.Hγ-4.NH	s
4.Hγ-5.NH	s
5.Hγ-NH <sub>2</sub>	s
1.Hγ-2.Hγ	s
2.Hγ-3.Hγ	nd
3.Hγ-4.Hγ	w
4.Hγ-5.Hγ	w
1.Hγ-2.Hτ1	w
2.Hγ-3.Hτ1	w
3.Hγ-4.Hτ1	w
4.Hγ-5.Hτ1	m
1.Hγ-2.Hδ	m
2.Hγ-3.Hδ	w
3.Hγ-4.Hδ	w
4.Hγ-5.Hδ	w
4.Hγ-5.Hδ	m
1.Ph-2.Hτ1	m
4.Ph-5.Hτ1	m
1.Ph-1.Hτ1	s
1.Ph-1.Hτ2	w
1.Ph-1.Hτ3	m
2.Hκ-4.Ph	w
2.Hκ-4.Ph	w
2.Hκ-3.Hγ	nd
2.Hκ-3.Hγ	w

*o*: overlap with the solvent signals, *nd*: not detected

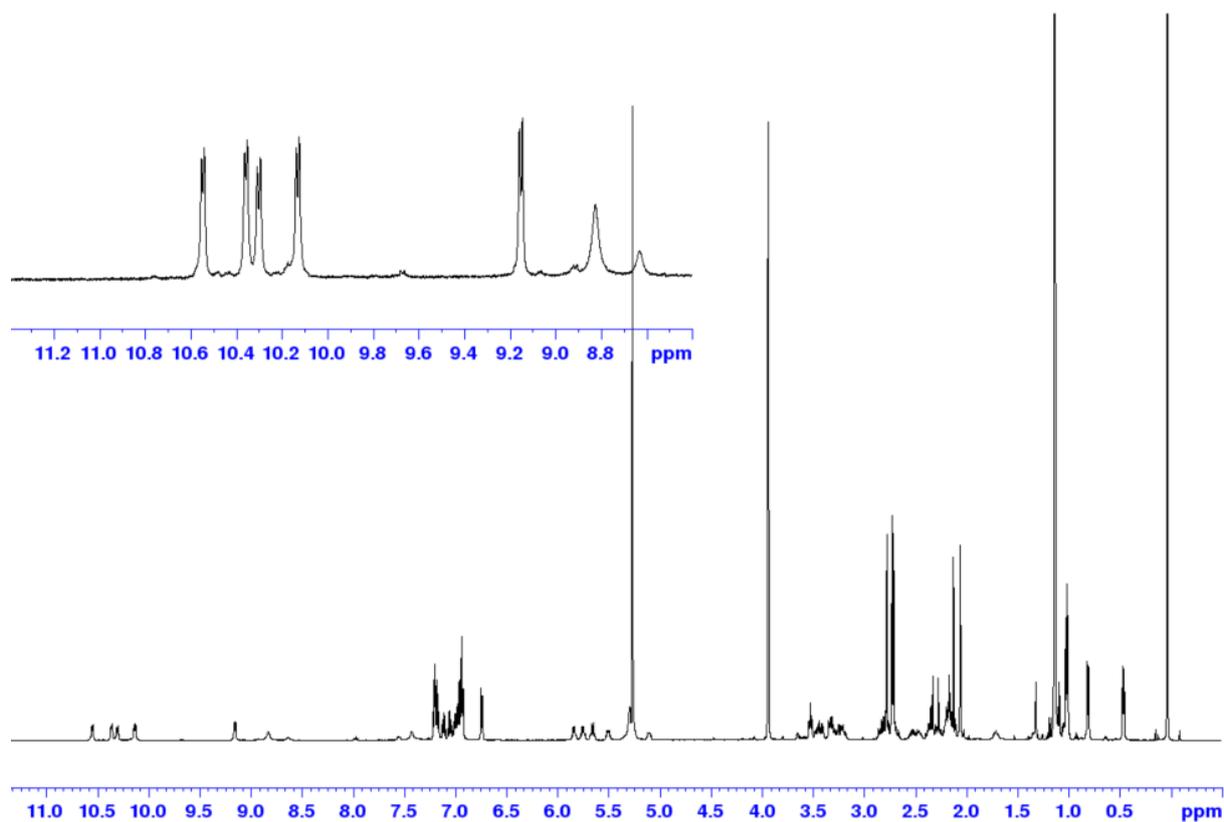


Figure S37 -  $^1\text{H}$  NMR spectrum of  $M(10)\text{-ATC}_5\text{-3-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

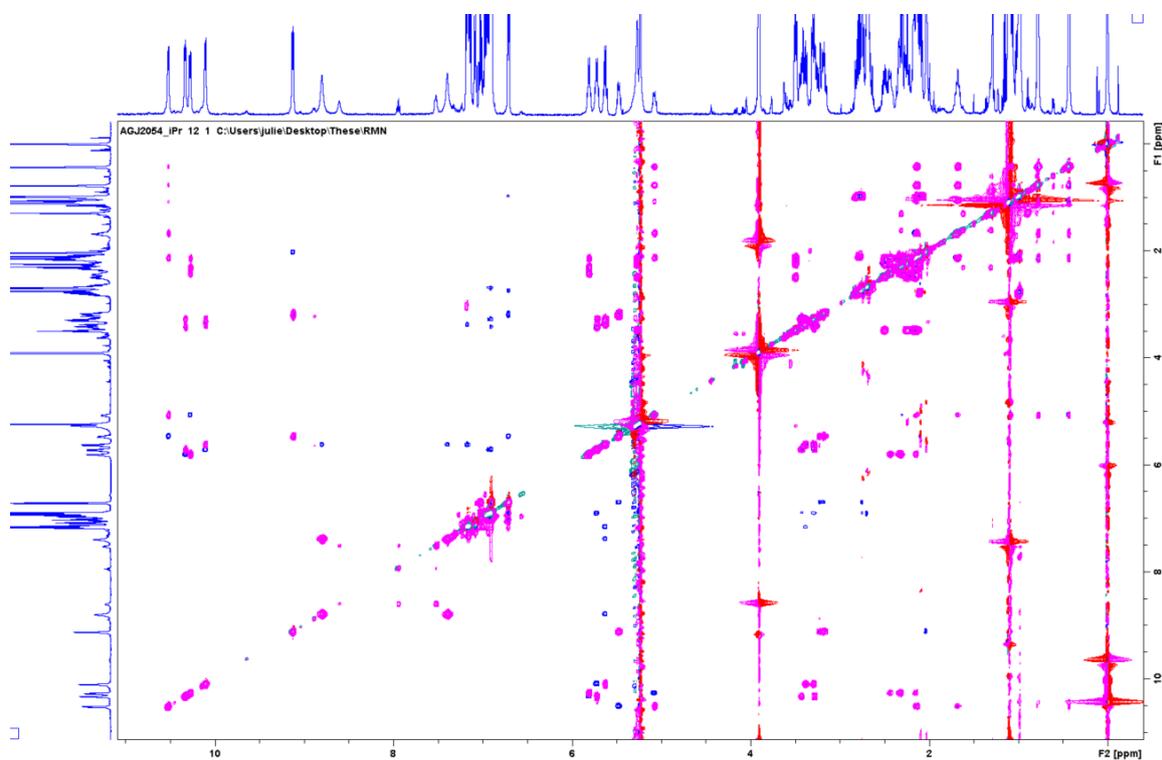
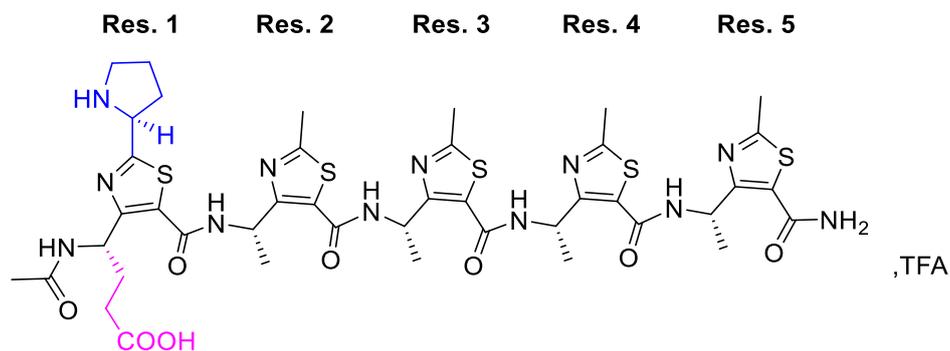


Figure S38 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of  $M(10)\text{-ATC}_5\text{-3-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

20. NMR analyses of **M(11)-ATC<sub>5</sub>-1-NH<sub>2</sub>**Table S65 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (86%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.03
Res. 1	8.96	5.38	2.23-2.15	$\epsilon$ CH <sub>2</sub> 1.90 $\tau^1$ CH <sub>2</sub> na $\tau^2$ CH <sub>2</sub> 2.24 $\tau^3$ CH <sub>2</sub> 2.29
Res. 2	10.60	5.71	1.57	$\tau^1$ CH <sub>3</sub> 2.65
Res. 3	10.10	5.69	1.56	$\tau^1$ CH <sub>3</sub> 2.65
Res. 4	10.18	5.71	1.59	$\tau^1$ CH <sub>3</sub> 2.65
Res. 5	10.20	5.52	1.65	$\tau^1$ CH <sub>3</sub> 2.65
NH <sub>2</sub>	8.81-7.48	-	-	-

na: not assigned

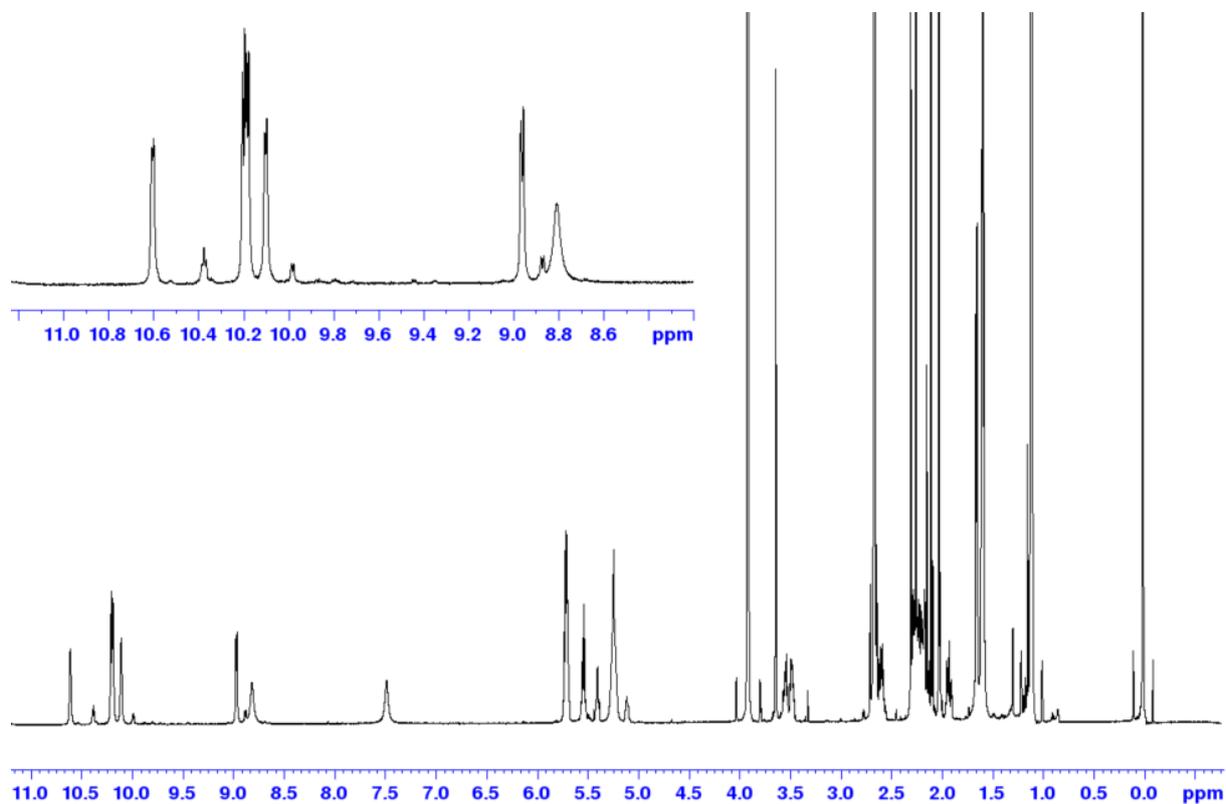
Table S66 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (86%)

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	172.5	-	-	-	-	CH <sub>3</sub> 21.8
Res. 1	161.4	49.1	29.0	na	na	$\epsilon$ C 31.2 $\tau^1$ C 51.8 $\tau^2$ C 52.0 $\tau^3$ C na $\tau^4$ C na
Res. 2	162.4	45.7	19.5	na	na	$\tau^1$ C 18.7
Res. 3	162.4	45.7	19.5	na	na	$\tau^1$ C 18.7
Res. 4	161.6	45.7	19.5	na	na	$\tau^1$ C 18.7
Res. 5	na	45.9	19.6	na	na	$\tau^1$ C 18.7

na: not assigned

Table S67 - Coupling Constants  ${}^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH- $d_6$  at 298K.

Residue	Major (298 K)
Res. 1	6.8
Res. 2	5.1
Res. 3	5.6
Res. 4	6.0
Res. 5	6.2

Figure S39 -  ${}^1\text{H}$  NMR spectrum of **M(11)-ATC<sub>5</sub>-1-NH<sub>2</sub>** in *i*PrOH- $d_7$  at 298K

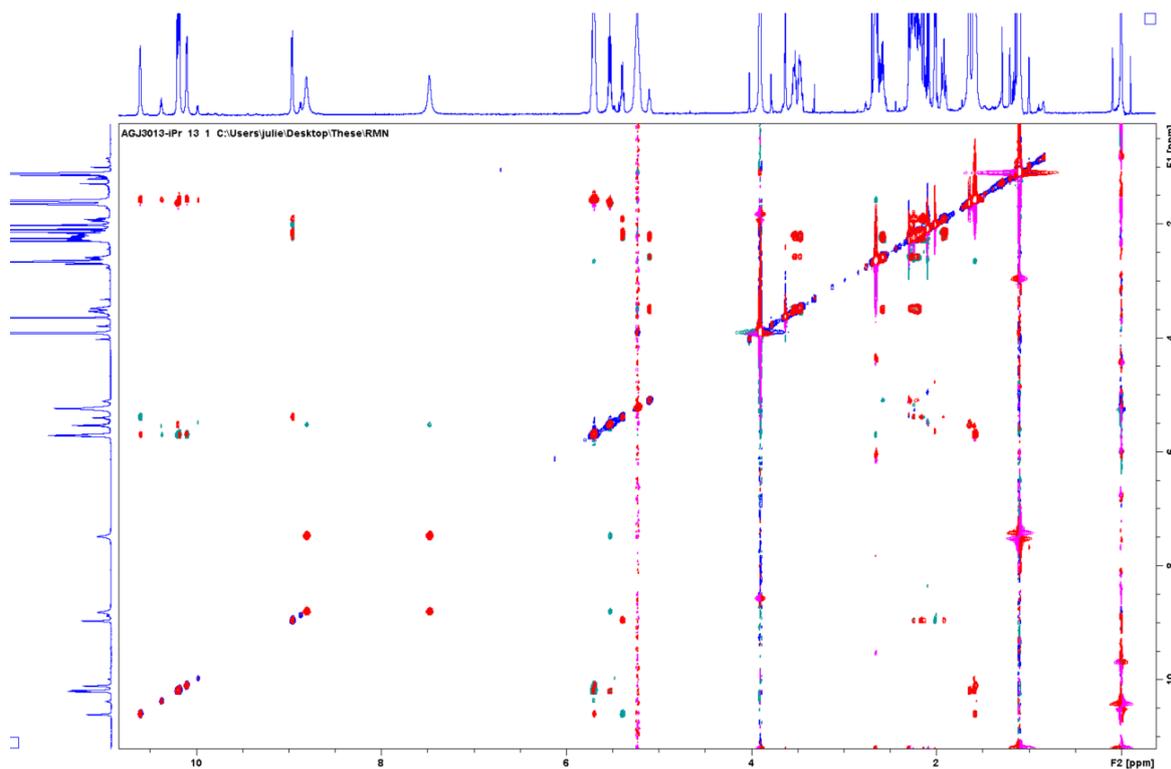


Figure S40 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(11)-ATC<sub>5</sub>-1-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

### 21.NMR analyses of **M(12)-ATC<sub>5</sub>-2-NH<sub>2</sub>**

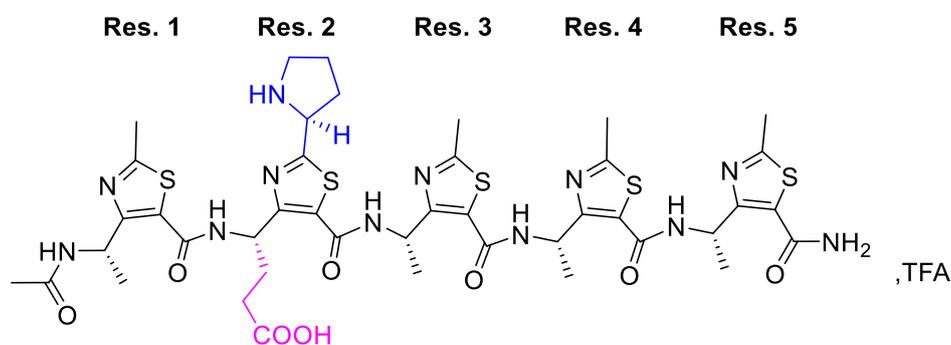


Table S68 - <sup>1</sup>H NMR in *iPrOH-d<sub>7</sub>* at 298K major conformer (88%)

<i>Residue</i>	<i>HN</i>	<i>γCH</i>	<i>δCH</i>	<i>Others</i>
<b>Ac</b>	-	-	-	CH <sub>3</sub> 2.00
<b>Res. 1</b>	8.93	5.21	1.45	<sup>τ</sup> 1CH <sub>3</sub> 2.65
<b>Res. 2</b>	10.48	5.63	2.35-2.26	<sup>ε</sup> CH <sub>2</sub> na <sup>τ</sup> 1CH na <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na
<b>Res. 3</b>	10.26	5.71	1.59	<sup>τ</sup> 1CH <sub>3</sub> 2.65
<b>Res. 4</b>	10.08	5.72	1.60	<sup>τ</sup> 1CH <sub>3</sub> 2.65

<b>Res. 5</b>	10.16	5.52	1.65	$^1\text{CH}_3$ 2.65
<b>NH<sub>2</sub></b>	8.81-7.48	-	-	-

*na: not assigned*

Table S69 -  $^{13}\text{C}$  NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (88%)

<b>Residue</b>	<b>C(O)</b>	<b><math>\gamma\text{C}</math></b>	<b><math>\delta\text{C}</math></b>	<b><math>\beta\text{C}</math></b>	<b><math>\alpha\text{C}</math></b>	<b>Others</b>
<b>Ac</b>	171.8	-	-	-	-	CH <sub>3</sub> 21.8
<b>Res. 1</b>	162.3	44.2	19.7	na	na	$^{\epsilon}\text{C}$ na $^{\tau^1}\text{C}$ na $^{\tau^2}\text{C}$ na $^{\tau^3}\text{C}$ na $^{\tau^4}\text{C}$ na
<b>Res. 2</b>	161.1	49.7	28.8	na	na	$^{\tau^1}\text{C}$ 18.6
<b>Res. 3</b>	162.3	45.8	19.5	na	na	$^{\tau^1}\text{C}$ 18.6
<b>Res. 4</b>	161.6	45.8	19.5	na	na	$^{\tau^1}\text{C}$ 18.6
<b>Res. 5</b>	na	45.8	19.5	na	na	$^{\tau^1}\text{C}$ 18.6

*na: not assigned*

Table S70 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

<b>Residue</b>	<b>Major (298 K)</b>
<b>Res. 1</b>	7.1
<b>Res. 2</b>	6.5
<b>Res. 3</b>	6.1
<b>Res. 4</b>	5.6
<b>Res. 5</b>	5.9

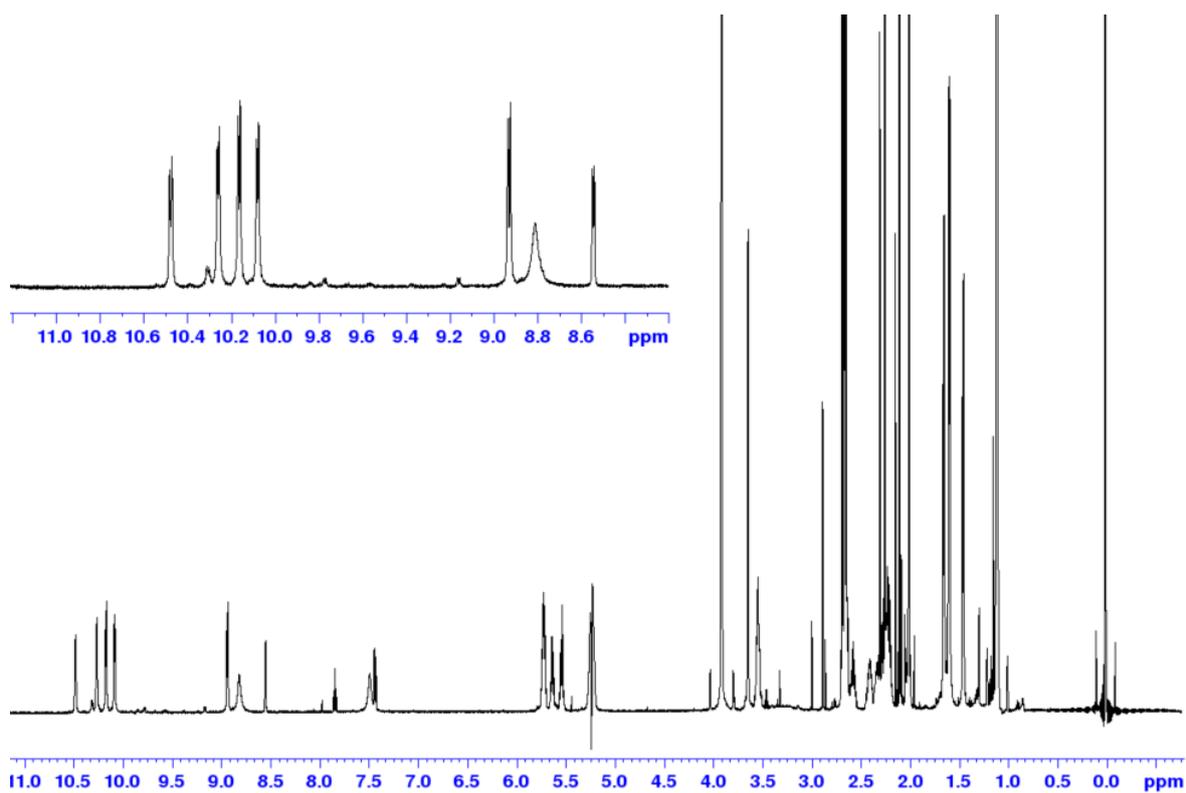


Figure S41 -  $^1\text{H}$  NMR spectrum of  $M(12)\text{-ATC}_5\text{-2-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

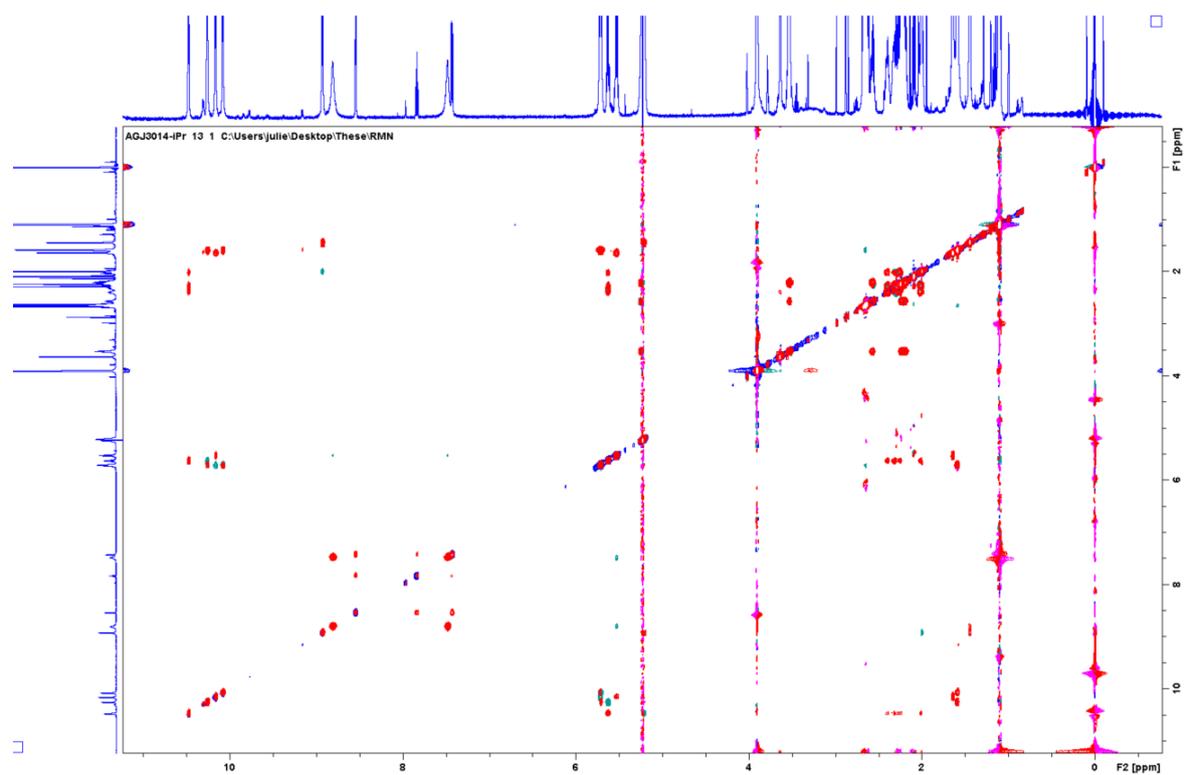
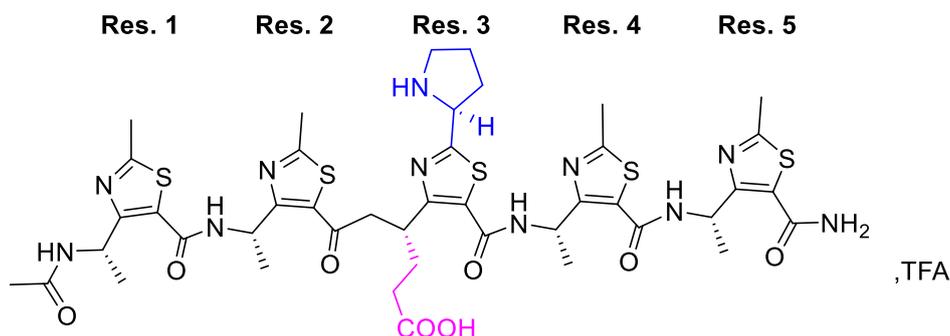


Figure S42 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of  $M(12)\text{-ATC}_5\text{-2-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

22. NMR analyses of **M(13)-ATC<sub>5</sub>-3-NH<sub>2</sub>**Table S71 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.00
Res. 1	8.85	5.49	1.45	$^{\tau 1}$ CH <sub>3</sub> 2.62
Res. 2	10.65	5.37	1.57	$^{\tau 1}$ CH <sub>3</sub> 2.65
Res. 3	10.27	5.60	2.10	$^{\epsilon}$ CH <sub>2</sub> 2.38-2.25 $^{\tau 1}$ CH na $^{\tau 2}$ CH <sub>2</sub> na $^{\tau 3}$ CH <sub>2</sub> na $^{\tau 4}$ CH <sub>2</sub> na
Res. 4	10.29	5.74	1.60	$^{\tau 1}$ CH <sub>3</sub> 2.67
Res. 5	10.04	5.55	1.63	$^{\tau 1}$ CH <sub>3</sub> 2.66
NH <sub>2</sub>	8.76-7.48	-	-	-

na: not assigned

Table S72 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 298K

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	172.0	-	-	-	-	CH <sub>3</sub> 21.7
Res. 1	162.3	44.6	19.7	na	na	$^{\tau 1}$ C 18.7
Res. 2	161.9	45.5	19.6	na	na	$^{\tau 1}$ C 18.9
Res. 3	161.1	49.9	29.7	na	na	$^{\epsilon}$ C 51.6 $^{\tau 1}$ C na $^{\tau 2}$ C 31.5 $^{\tau 3}$ C 30.8 $^{\tau 4}$ C 45.4
Res. 4	161.8	45.6	19.6	na	na	$^{\tau 1}$ C 18.9
Res. 5	na	45.9	19.8	na	na	$^{\tau 1}$ C 18.9

na: not assigned

Table S73 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	(298 K)
Res. 1	6.8
Res. 2	5.8
Res. 3	5.8
Res. 4	6.3
Res. 5	6.3

Table S74 - Inter-residue NOE correlations observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	o
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.Hγ-2.NH	s
2.Hγ-3.NH	s
3.Hγ-4.NH	s
4.Hγ-5.NH	s
5.Hγ-NH <sub>2</sub>	m
1.Hγ-2.Hγ	w
2.Hγ-3.Hγ	nd
3.Hγ-4.Hγ	nd
4.Hγ-5.Hγ	nd
1.Hγ-2.Hτ1	w
2.Hγ-3.Hτ1	nd
3.Hγ-4.Hτ1	w
4.Hγ-5.Hτ1	w
1.Hγ-2.Hδ	nd
2.Hγ-3.Hδ	nd
3.Hγ-4.Hδ	nd
4.Hγ-5.Hδ	nd
4.Hγ-5.Hδ	nd

1.H $\delta$ -2.H $\tau$ 1	m
2.H $\delta$ -3.H $\tau$ 1	nd
3.H $\delta$ -4.H $\tau$ 1	w
4.H $\delta$ -5.H $\tau$ 1	m

*o*: overlap with the solvent signals, *nd*: not detected

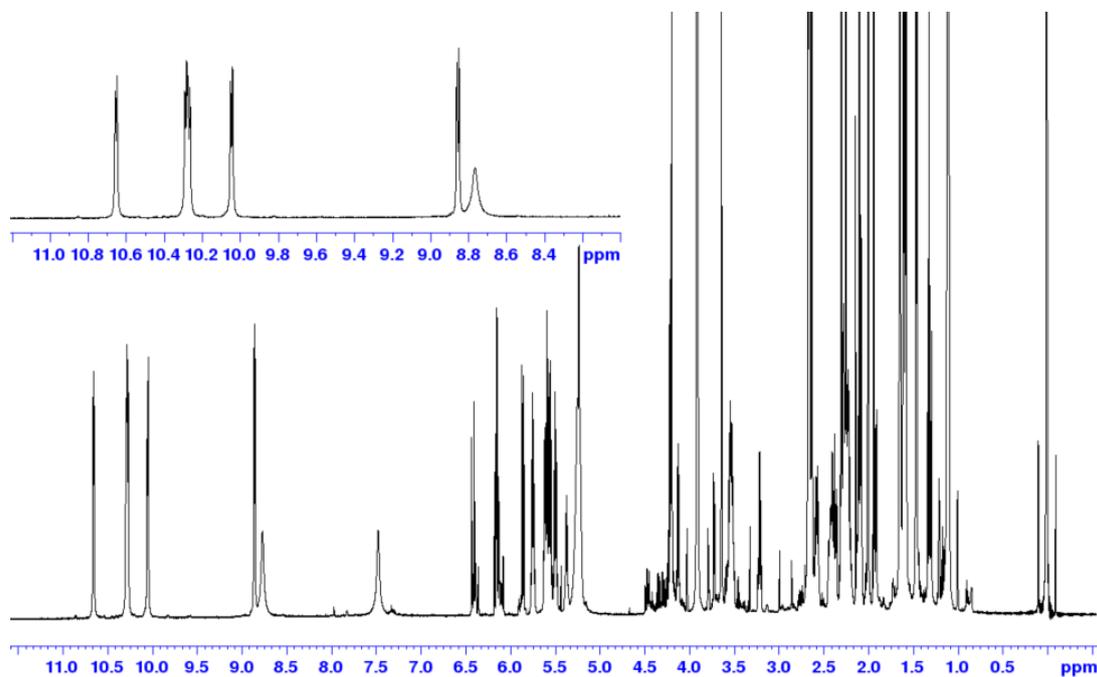


Figure S43 -  $^1\text{H}$  NMR spectrum of **M(13)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

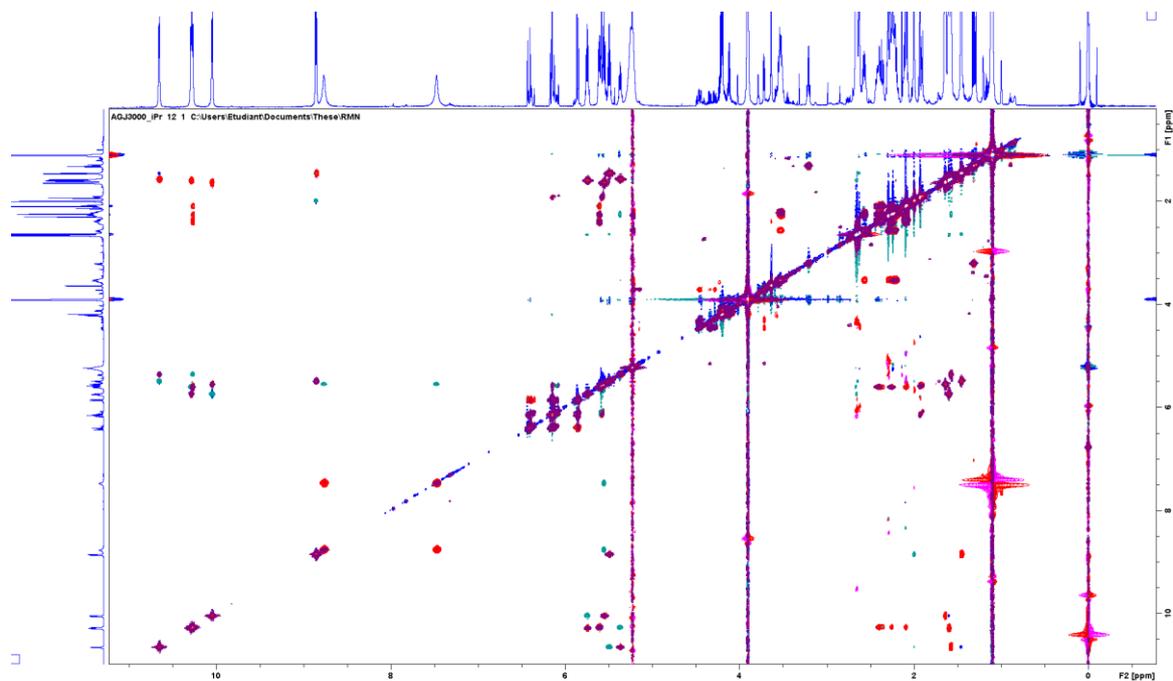
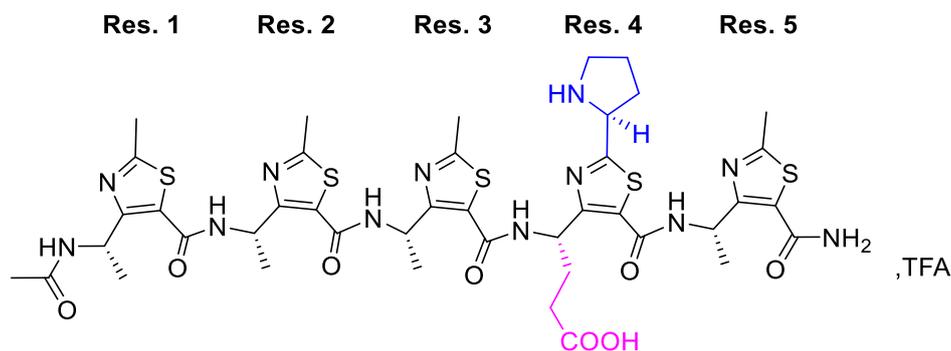


Figure S44 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(13)-ATC<sub>5</sub>-3-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

23. NMR analyses of **M(14)-ATC<sub>5</sub>-4-NH<sub>2</sub>**Table S75 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.00
Res. 1	8.81	5.51	1.46	$\tau^1$ CH <sub>3</sub> 2.65
Res. 2	10.55	5.65	1.58	$\tau^1$ CH <sub>3</sub> 2.65
Res. 3	10.32	5.40	1.57	$\tau^1$ CH <sub>3</sub> 2.65
Res. 4	10.21	5.61	2.40-2.26	$\epsilon$ CH <sub>2</sub> 2.09 $\tau^1$ CH na $\tau^2$ CH <sub>2</sub> na $\tau^3$ CH <sub>2</sub> 2.05 $\tau^4$ CH <sub>2</sub> na
Res. 5	10.26	5.59	1.65	$\tau^1$ CH <sub>3</sub> 2.65
NH <sub>2</sub>	8.66-7.50	-	-	-

*na: not assigned*Table S76 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 298K

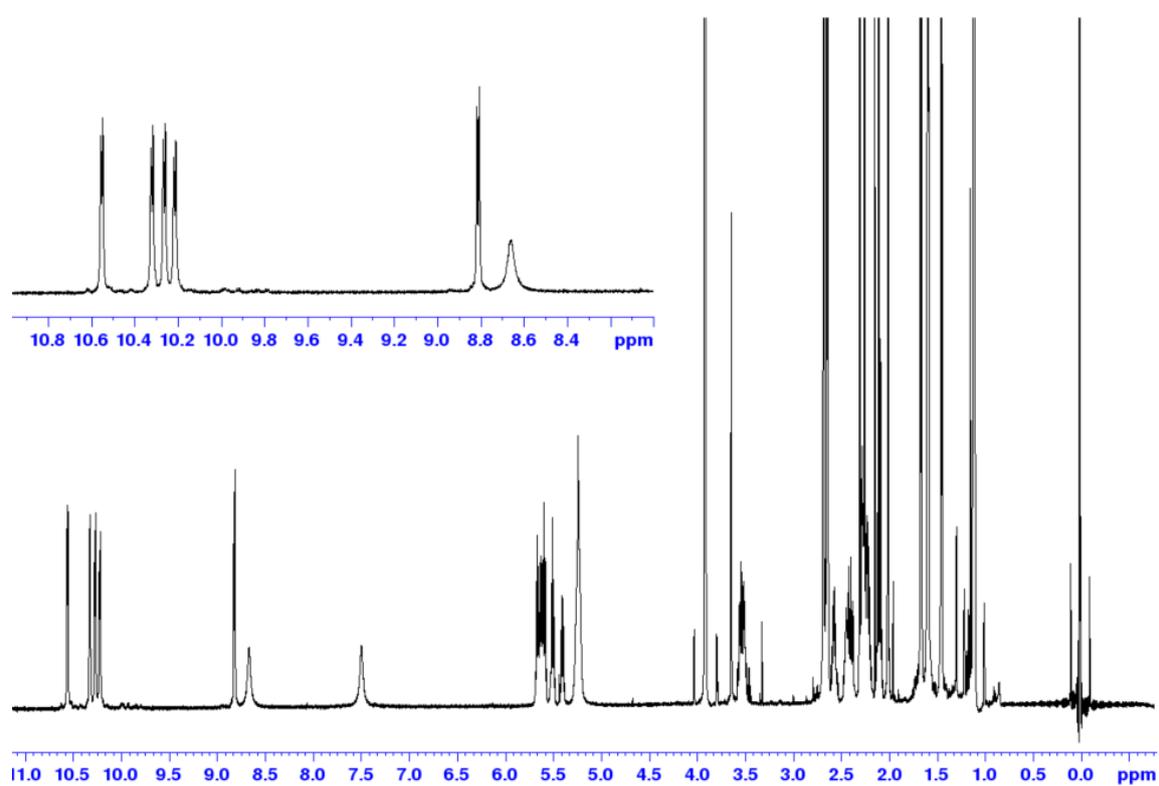
Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	171.7	-	-	-	-	CH <sub>3</sub> 21.7
Res. 1	162.5	44.6	19.5	na	na	$\tau^1$ C 18.8
Res. 2	162.3	45.5	19.5	na	na	$\tau^1$ C 18.8
Res. 3	162.1	45.4	19.5	na	na	$\epsilon$ C na $\tau^1$ C na $\tau^2$ C na $\tau^3$ C na $\tau^4$ C na
Res. 4	160.8	50.0	31.5	na	na	$\tau^1$ C 18.8
Res. 5	na	46.0	19.5	na	na	$\tau^1$ C 18.8

*na: not assigned*

Table S77 - Coupling Constants  ${}^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH- $d_6$  at 298K.

Residue	(298 K)
Res. 1	6.9
Res. 2	5.7
Res. 3	6.0
Res. 4	6.1
Res. 5	6.1

*o*: overlap with the solvent signals, *nd*: not detected

Figure S45 -  ${}^1\text{H}$  NMR spectrum of *M*(14)-ATC<sub>5</sub>-4-NH<sub>2</sub> in *i*PrOH- $d_7$  at 298K

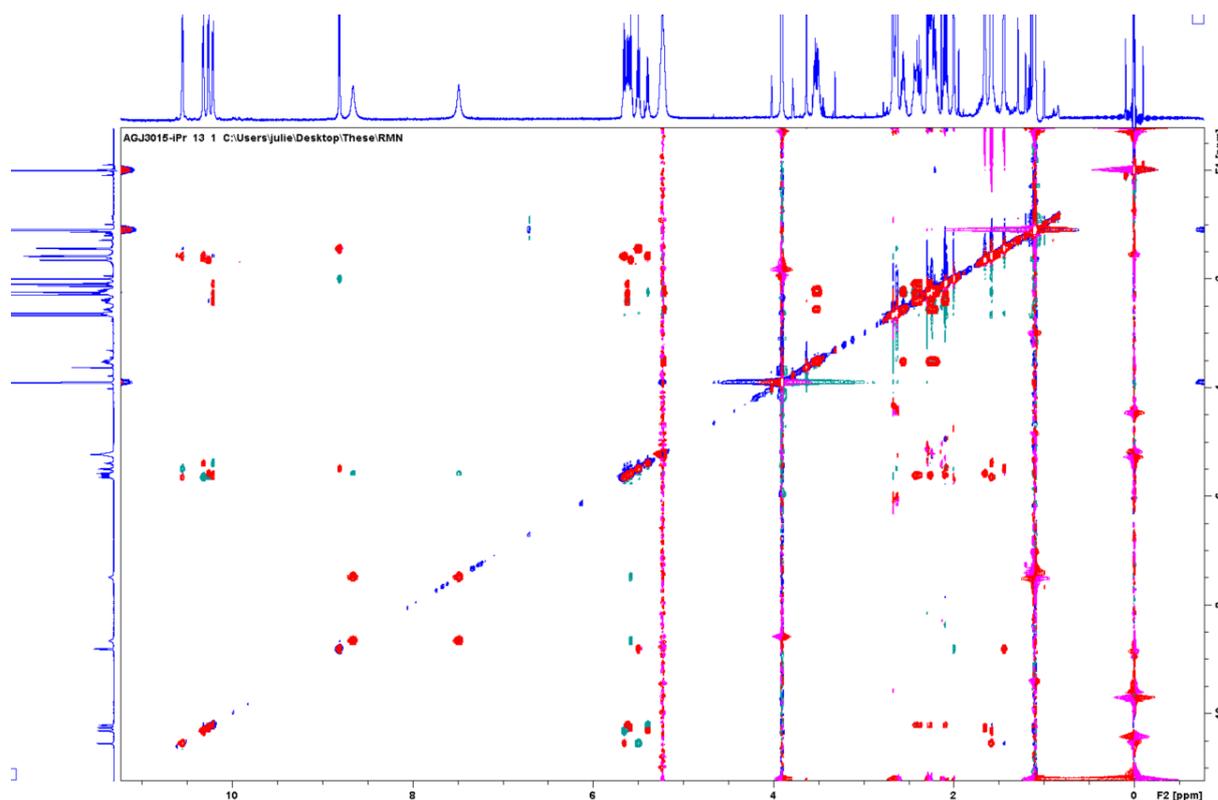


Figure S46 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(14)-ATC<sub>5</sub>-4-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

### 24.NMR analyses of **M(15)-ATC<sub>5</sub>-5-NH<sub>2</sub>**

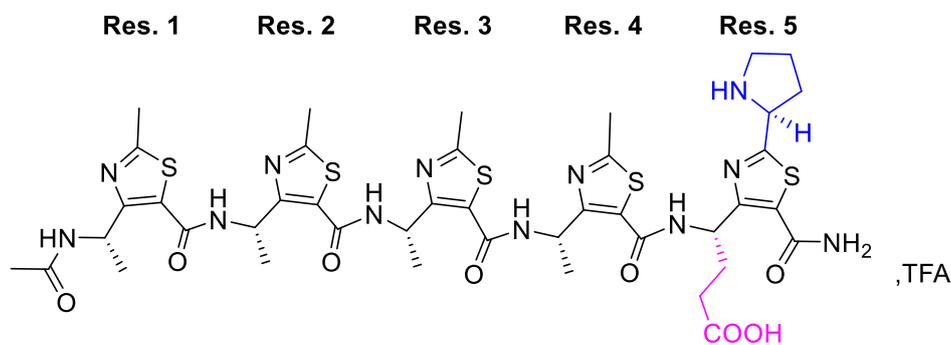


Table S78 - <sup>1</sup>H NMR in *iPrOH-d<sub>7</sub>* at 298K major conformer (91%)

<i>Residue</i>	<i>HN</i>	<i>γCH</i>	<i>δCH</i>	<i>Others</i>
<b>Ac</b>	-	-	-	CH <sub>3</sub> 2.00
<b>Res. 1</b>	8.81	5.51	1.44	<sup>13</sup> CH <sub>3</sub> 2.64
<b>Res. 2</b>	10.52	5.66	1.57	<sup>13</sup> CH <sub>3</sub> 2.64
<b>Res. 3</b>	10.21	5.67	1.61	<sup>13</sup> CH <sub>3</sub> 2.64

<b>Res. 4</b>	10.32	5.39	1.59	$^{\epsilon}\text{CH}_2$ na $^{\tau^1}\text{CH}$ na $^{\tau^2}\text{CH}_2$ na $^{\tau^3}\text{CH}_2$ na $^{\tau^4}\text{CH}_2$ na
<b>Res. 5</b>	10.14	5.56	2.45-2.31	$^{\tau^1}\text{CH}_3$ 2.66
<b>NH<sub>2</sub></b>	8.88-7.75	-	-	-

na: not assigned

Table S79 -  $^{13}\text{C}$  NMR in *iPrOH-d<sub>7</sub>* at 298K major conformer (91%)

<b>Residue</b>	<b>C(O)</b>	<b><math>\gamma\text{C}</math></b>	<b><math>\delta\text{C}</math></b>	<b><math>\beta\text{C}</math></b>	<b><math>^{\alpha}\text{C}</math></b>	<b>Others</b>
<b>Ac</b>	172.0	-	-	-	-	$\text{CH}_3$ 21.7
<b>Res. 1</b>	162.3	44.5	19.5	na	na	$^{\tau^1}\text{C}$ 18.9
<b>Res. 2</b>	162.3	45.7	19.5	na	na	$^{\tau^1}\text{C}$ 18.9
<b>Res. 3</b>	162.4	45.7	19.5	na	na	$^{\tau^1}\text{C}$ 18.9
<b>Res. 4</b>	162.1	45.8	19.5	na	na	$^{\tau^1}\text{C}$ 18.9
<b>Res. 5</b>	na	49.4	23.4	na	na	$^{\epsilon}\text{C}$ 20.6 $^{\tau^1}\text{C}$ na $^{\tau^2}\text{C}$ na $^{\tau^3}\text{C}$ na $^{\tau^4}\text{C}$ na

na: not assigned

Table S80 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i-PrOH-d<sub>6</sub>* at 298K.

<b>Residue</b>	<b>Major (298 K)</b>
<b>Res. 1</b>	6.7
<b>Res. 2</b>	5.5
<b>Res. 3</b>	5.7
<b>Res. 4</b>	4.4
<b>Res. 5</b>	6.8

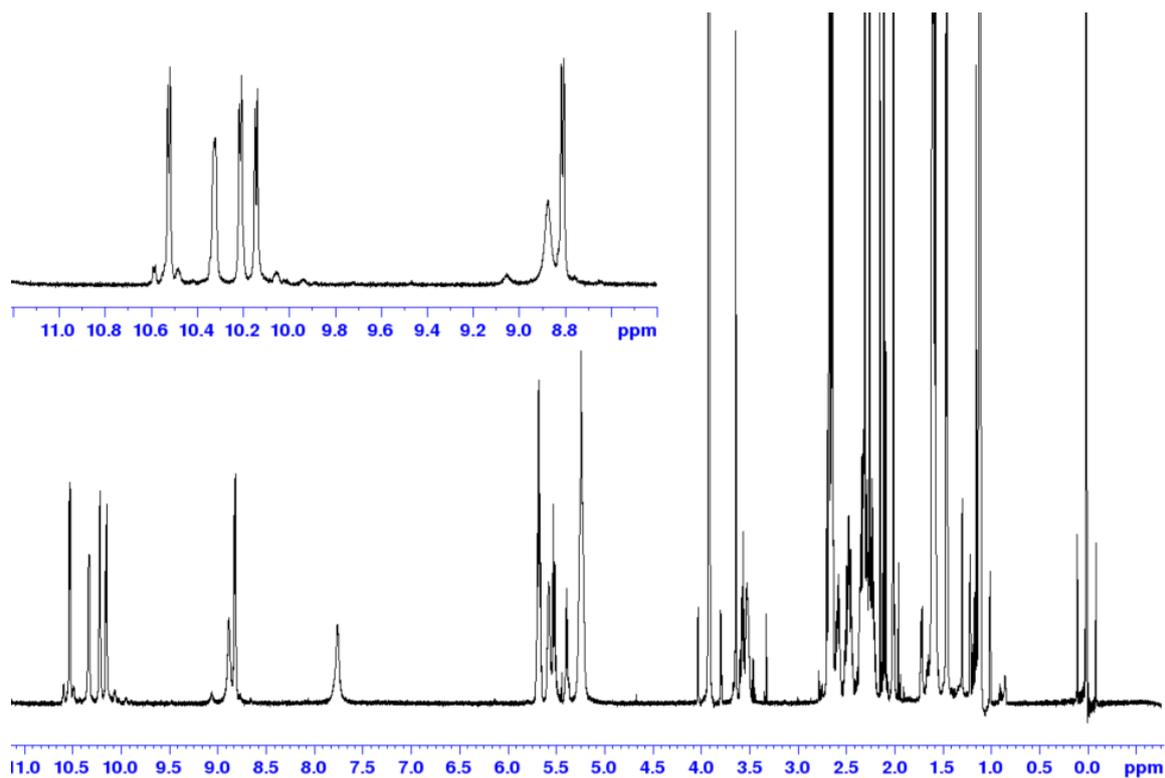


Figure S47 -  $^1\text{H}$  NMR spectrum of  $M(15)\text{-ATC}_5\text{-5-NH}_2$  in  $i\text{PrOH-}d_7$  at 298K

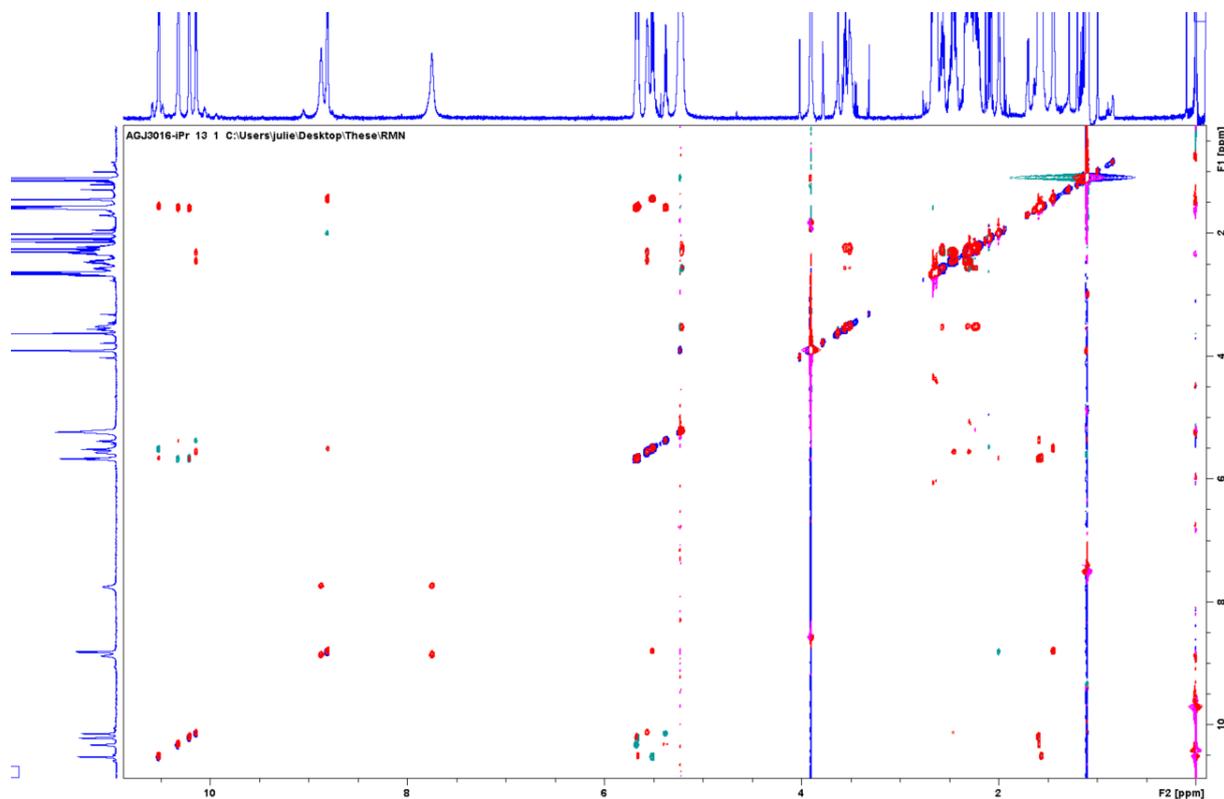
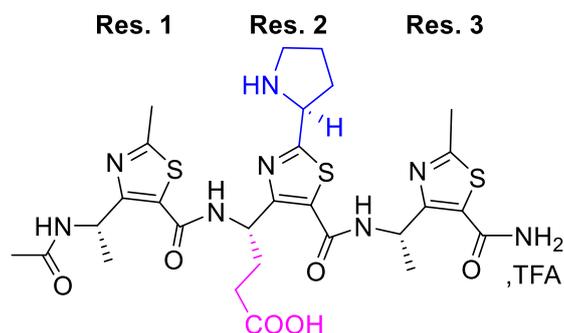


Figure S48 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of  $M(15)\text{-ATC}_5\text{-5-NH}_2$  in  $i\text{PrOH-}d_7$  at 298K

25. NMR analyses of **M(16)-ATC<sub>3</sub>-2-NH<sub>2</sub>**Table S81 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (>95%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 1.99
Res. 1	8.91	5.21	1.43	$^1\text{CH}_2$ 2.63
Res. 2	10.48	5.62	2.39-2.29	$^6\text{CH}_2$ na $^1\text{CH}$ na $^2\text{CH}_2$ 2.24 $^3\text{CH}_2$ 2.09 $^4\text{CH}_2$ 3.51
Res. 3	10.23	5.56	1.64	$^1\text{CH}_3$ 2.66
NH <sub>2</sub>	8.66-7.47	-	-	-

na: not assigned

Table S82 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (>95%)

Residue	C(O)	$\gamma$ C	$\delta$ C	$\beta$ C	$\alpha$ C	Others
Ac	171.7	-	-	-	-	CH <sub>3</sub> 21.7
Res. 1	162.2	44.5	19.8	na	na	$^1\text{C}$ 18.6
Res. 2	160.8	49.8	na	na	na	$^6\text{C}$ 52.1 $^1\text{C}$ na $^2\text{C}$ na $^3\text{C}$ 45.3
Res. 3		46.1	19.7	na	na	$^1\text{C}$ 18.8

na: not assigned

Table S83 - Coupling Constants <sup>3</sup>J(NH,  $\gamma$ CH) (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	Major (298 K)
Res. 1	7.1
Res. 2	6.2

Res. 3

6.1

Table S84 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (*s*) < 2.7 Å, 2.7 Å < Medium (*m*) < 3.3 Å, 3.3 Å < Weak (*w*).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.Hγ-2.NH	m
2.Hγ-3.NH	s
3.Hγ-NH <sub>2</sub>	s
1.Hγ-2.Hγ	nd
2.Hγ-3.Hγ	nd
1.Hγ-2.Hτ1	nd
2.Hγ-3.Hτ1	m
1.Hγ-2.Hδ	nd
2.Hγ-3.Hδ	nd

*o*: overlap with the solvent signals, *nd*: not detected

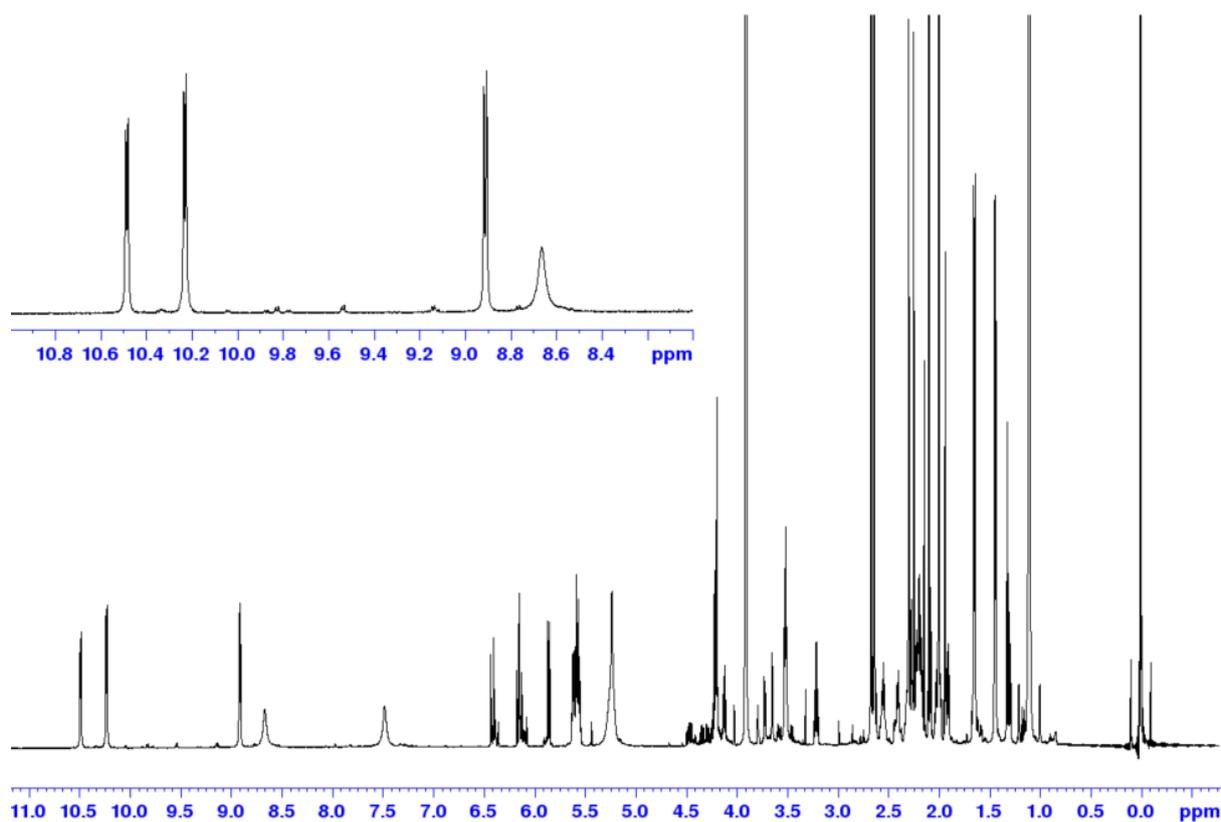


Figure S49 - <sup>1</sup>H NMR spectrum of **M(16)-ATC<sub>3</sub>-2-NH<sub>2</sub>** in *i*PrOH-*d*<sub>7</sub> at 298K

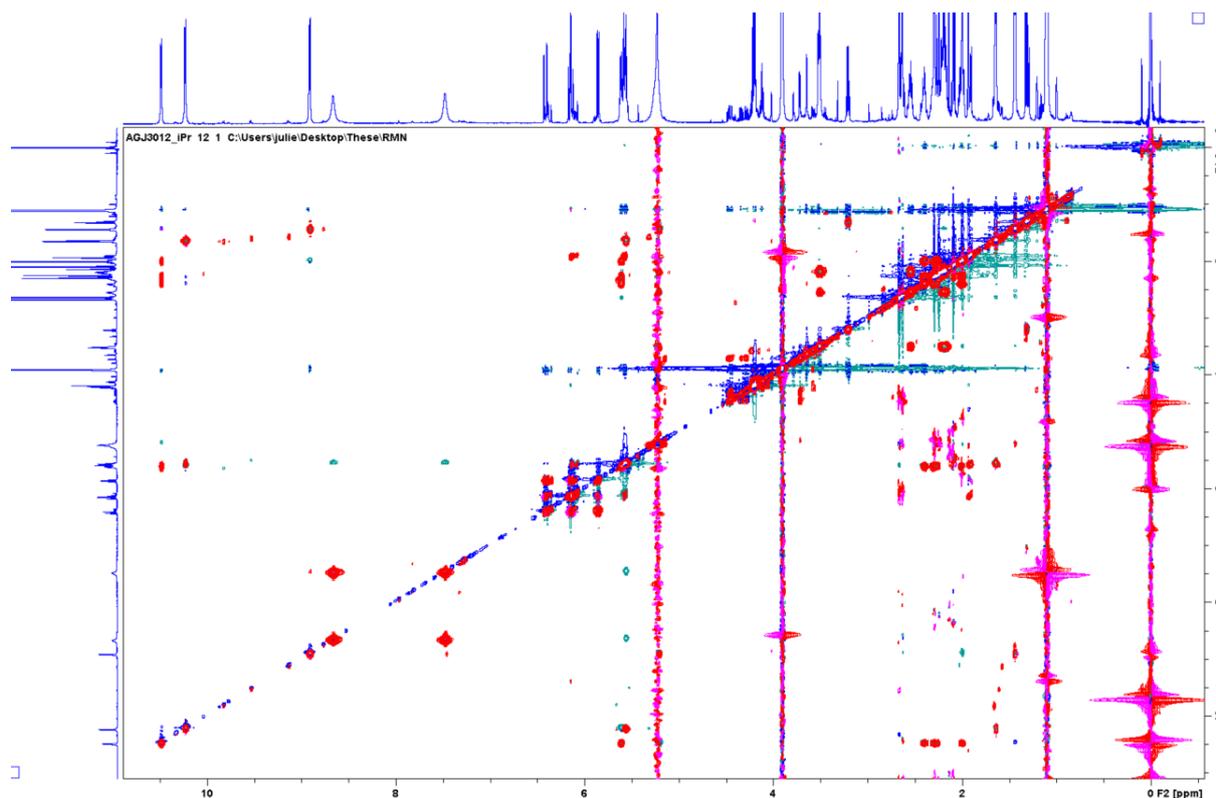


Figure S50 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(16)-ATC<sub>3</sub>-2-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

### 26. NMR analyses of **M(17)-ATC<sub>7</sub>-4-NH<sub>2</sub>**

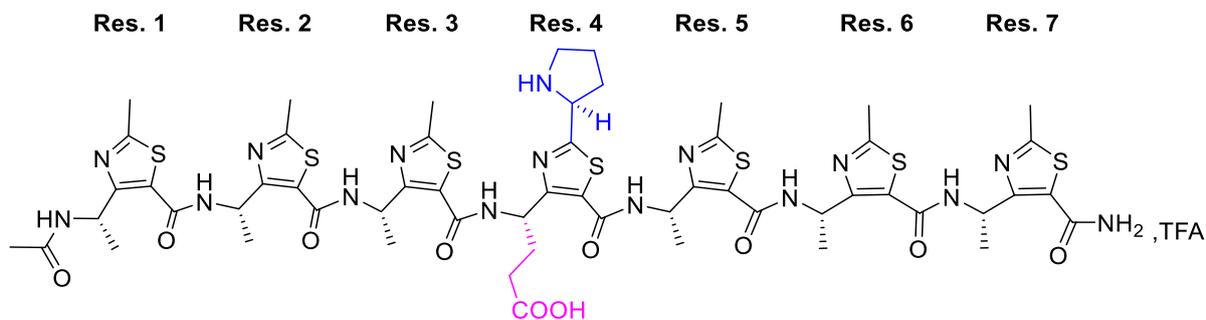


Table S85 - <sup>1</sup>H NMR in *iPrOH-d<sub>7</sub>* at 298K major conformer (95%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.00
Res. 1	8.81	5.49	1.44	$\tau^1$ CH <sub>3</sub> 2.65
Res. 2	10.55	5.66	1.59	$\tau^1$ CH <sub>3</sub> 2.64
Res. 3	10.32	5.40	1.59	$\tau^1$ CH <sub>3</sub> 2.67
Res. 4	10.19	5.63		$\delta$ CH <sub>2</sub> na $\tau^1$ CH na $\tau^2$ CH <sub>2</sub> na $\tau^3$ CH <sub>2</sub> na $\tau^4$ CH <sub>2</sub> na

<b>Res. 5</b>	10.30	5.73	1.61	$^{13}\text{C}$ 2.66
<b>Res. 6</b>	10.09	5.72	1.59	$^{13}\text{C}$ 2.68
<b>Res. 7</b>	10.17	5.53	1.64	$^{13}\text{C}$ 2.66
<b>NH<sub>2</sub></b>	8.81-7.48	-	-	-

na: not assigned

Table S86 -  $^{13}\text{C}$  NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (95%)

<b>Residue</b>	<b>C(O)</b>	<b><math>\gamma\text{C}</math></b>	<b><math>\delta\text{C}</math></b>	<b><math>\beta\text{C}</math></b>	<b><math>\alpha\text{C}</math></b>	<b>Others</b>
<b>Ac</b>	171.9	-	-	-	-	CH <sub>3</sub> 21.7
<b>Res. 1</b>	167.8	44.5	19.6	na	na	$^{13}\text{C}$ 18.9
<b>Res. 2</b>	167.8	45.8	19.5	na	na	$^{13}\text{C}$ 18.9
<b>Res. 3</b>	168.1	45.7	19.5	na	na	$^{13}\text{C}$ 18.9
<b>Res. 4</b>	na	50.1	30.9	na	na	$^{13}\text{C}$ 20.6 $^{13}\text{C}$ na $^{13}\text{C}$ na $^{13}\text{C}$ na $^{13}\text{C}$ na
<b>Res. 5</b>	167.2	45.8	19.5	na	na	$^{13}\text{C}$ 18.9
<b>Res. 6</b>	168.1	45.8	19.5	na	na	$^{13}\text{C}$ 18.9
<b>Res. 7</b>	na	45.8	19.5	na	na	$^{13}\text{C}$ 18.9

na: not assigned

Table S87 - Coupling Constants  $^3J(\text{NH}, \gamma\text{CH})$  (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

<b>Residue</b>	<b>Major (298 K)</b>
<b>Res. 1</b>	6.6
<b>Res. 2</b>	5.3
<b>Res. 3</b>	6.0
<b>Res. 4</b>	6.0
<b>Res. 5</b>	5.5
<b>Res. 6</b>	5.6
<b>Res. 7</b>	6.0

Table S88 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	w
5.NH-6.NH	w
6.NH-7.NH	w
7.NH-NH <sub>2</sub>	nd
7.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.Hγ-2.NH	s
2.Hγ-3.NH	s
3.Hγ-4.NH	s
4.Hγ-5.NH	s
5.Hγ-6.NH	s
6.Hγ-7.NH	s
7.Hγ-NH <sub>2</sub>	s
7.Hγ-NH <sub>2</sub>	s
1.Hγ-2.Hγ	nd
2.Hγ-3.Hγ	nd
3.Hγ-4.Hγ	nd
4.Hγ-5.Hγ	nd
5.Hγ-6.Hγ	nd
6.Hγ-7.Hγ	nd
1.Hγ-2.Hτ1	m
2.Hγ-3.Hτ1	s
3.Hγ-4.Hτ1	nd
4.Hγ-5.Hτ1	w
5.Hγ-6.Hτ1	m
6.Hγ-7.Hτ1	s
1.Hγ-2.Hδ	o
2.Hγ-3.Hδ	o
3.Hγ-4.Hδ	o
4.Hγ-5.Hδ	o
5.Hγ-6.Hδ	o

6.H $\gamma$ -7.H $\delta$	o
1.H $\delta$ -2.H $\tau$ 1	o
2.H $\delta$ -3.H $\tau$ 1	o
3.H $\delta$ -4.H $\tau$ 1	nd
4.H $\delta$ -5.H $\tau$ 1	o
5.H $\delta$ -6.H $\tau$ 1	o
6.H $\delta$ -7.H $\tau$ 1	o

*o*: overlap with the solvent signals, *nd*: not detected

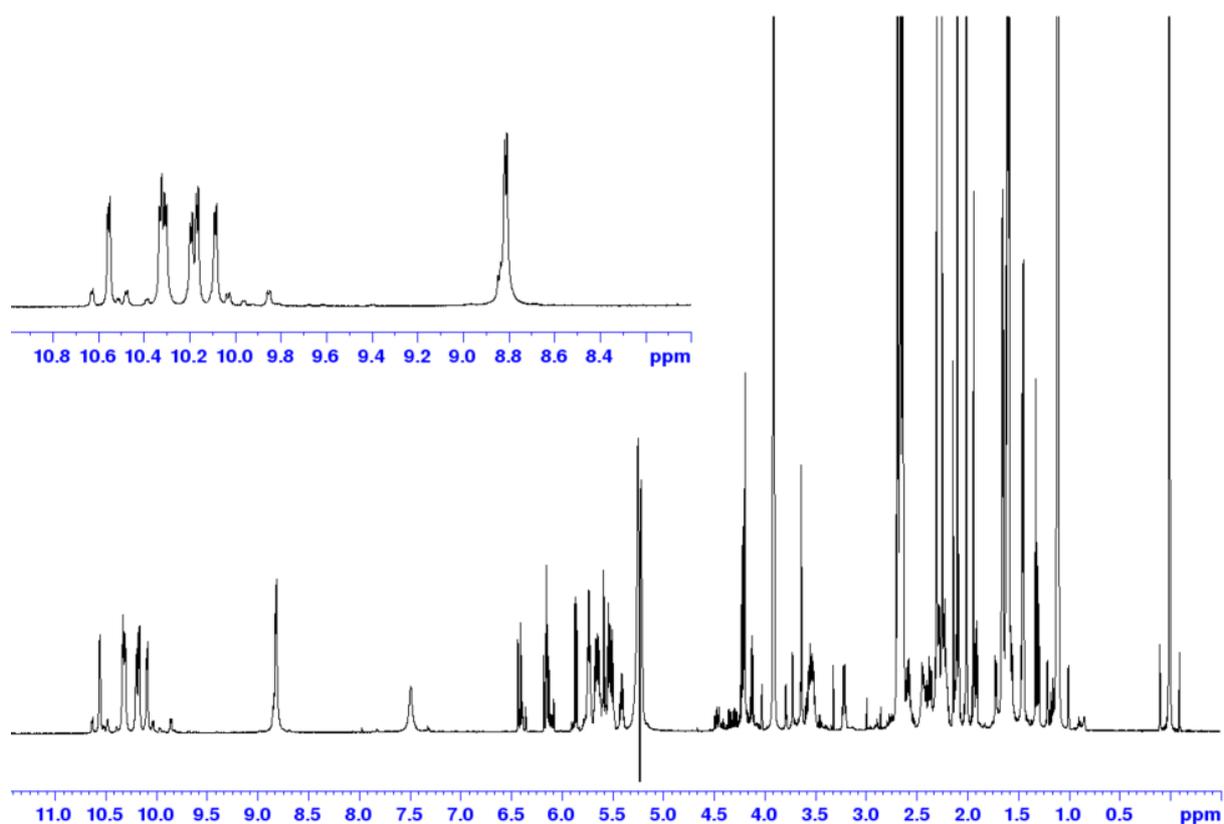


Figure S51 -  $^1\text{H}$  NMR spectrum of **M(17)-ATC7-4-NH<sub>2</sub>** in *i*PrOH-*d*<sub>7</sub> at 298K

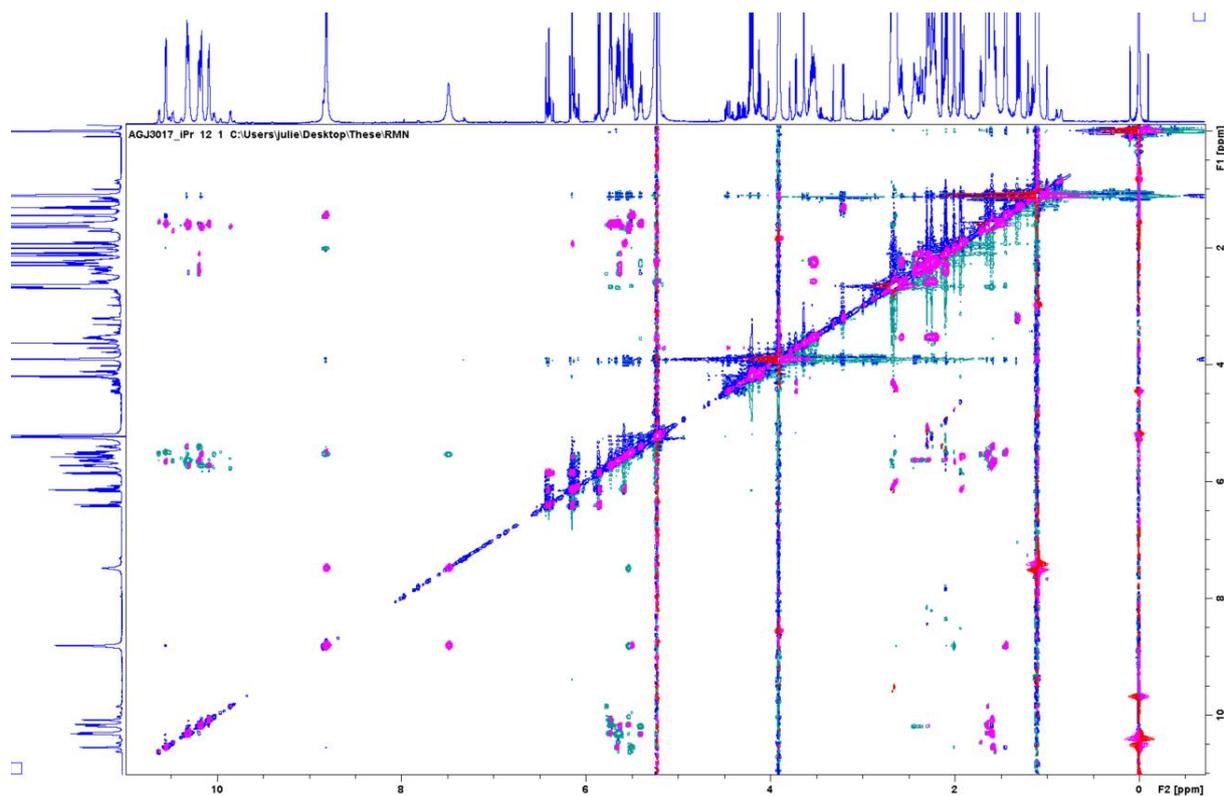


Figure S52 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(17)-ATC<sub>7</sub>-4-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

### 27. NMR analyses of **M(18)-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub>**

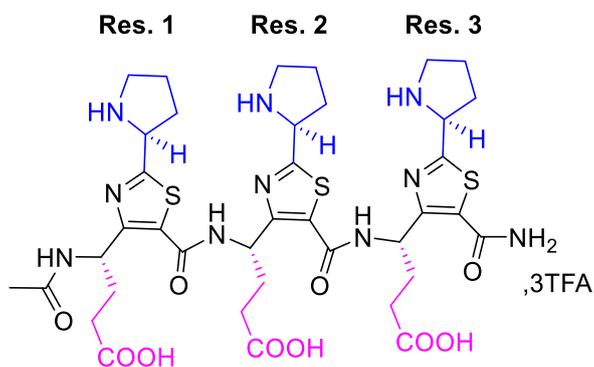


Table S89 - <sup>1</sup>H NMR in *iPrOH-d<sub>7</sub>* at 298K major conformer (>95%)

Residue	HN	<sup>γ</sup> CH	<sup>δ</sup> CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.00
Res. 1	8.95	5.28	2.26-2.16	<sup>ε</sup> CH <sub>2</sub> 1.97 <sup>τ</sup> 1CH 5.09 <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na
Res. 2	10.59	5.47	2.44-2.31	<sup>ε</sup> CH <sub>2</sub> na <sup>τ</sup> 1CH 5.11 <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na
Res. 3	10.32	5.53	2.46-2.29	<sup>ε</sup> CH <sub>2</sub> na <sup>τ</sup> 1CH 5.11 <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na

<b>NH<sub>2</sub></b>	9.00-7.70	-	-	-
-----------------------	-----------	---	---	---

*na: not assigned*

Table S90 - <sup>13</sup>C NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (>95%)

<b>Residue</b>	<b>C(O)</b>	<b>γC</b>	<b>δC</b>	<b>βC</b>	<b>αC</b>	<b>Others</b>
<b>Ac</b>	172.1	-	-	-	-	CH <sub>3</sub> 21.8
<b>Res. 1</b>	na	47.7	28.7	na	na	<sup>ε</sup> C 51.8 <sup>τ</sup> 1C na <sup>τ</sup> 2C na <sup>τ</sup> 3C 45.3
<b>Res. 2</b>	na	48.9	30.4	161.3	na	<sup>ε</sup> C 52.1 <sup>τ</sup> 1C na <sup>τ</sup> 2C na <sup>τ</sup> 3C 45.3
<b>Res. 3</b>	na	49.4	28.6	161.5	na	<sup>ε</sup> C 49.4 <sup>τ</sup> 1C na <sup>τ</sup> 2C na <sup>τ</sup> 3C 45.3

*na: not assigned*Table S91 - Coupling Constants <sup>3</sup>J(NH, γCH) (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

<b>Residue</b>	<b>Major (298 K)</b>
<b>Res. 1</b>	7.7
<b>Res. 2</b>	5.9
<b>Res. 3</b>	5.9

Table S92 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (*s*) < 2.7 Å, 2.7 Å < Medium (*m*) < 3.3 Å, 3.3 Å < Weak (*w*).

<b>NOE correlations</b>	<b>Intensity</b>
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.Hγ-2.NH	s
2.Hγ-3.NH	s
3.Hγ-NH <sub>2</sub>	s
1.Hγ-2.Hγ	nd
2.Hγ-3.Hγ	nd
1.Hγ-2.Hτ1	nd
2.Hγ-3.Hτ1	nd
1.Hγ-2.Hδ	s
2.Hγ-3.Hδ	s

*o: overlap with the solvent signals, nd: not detected*

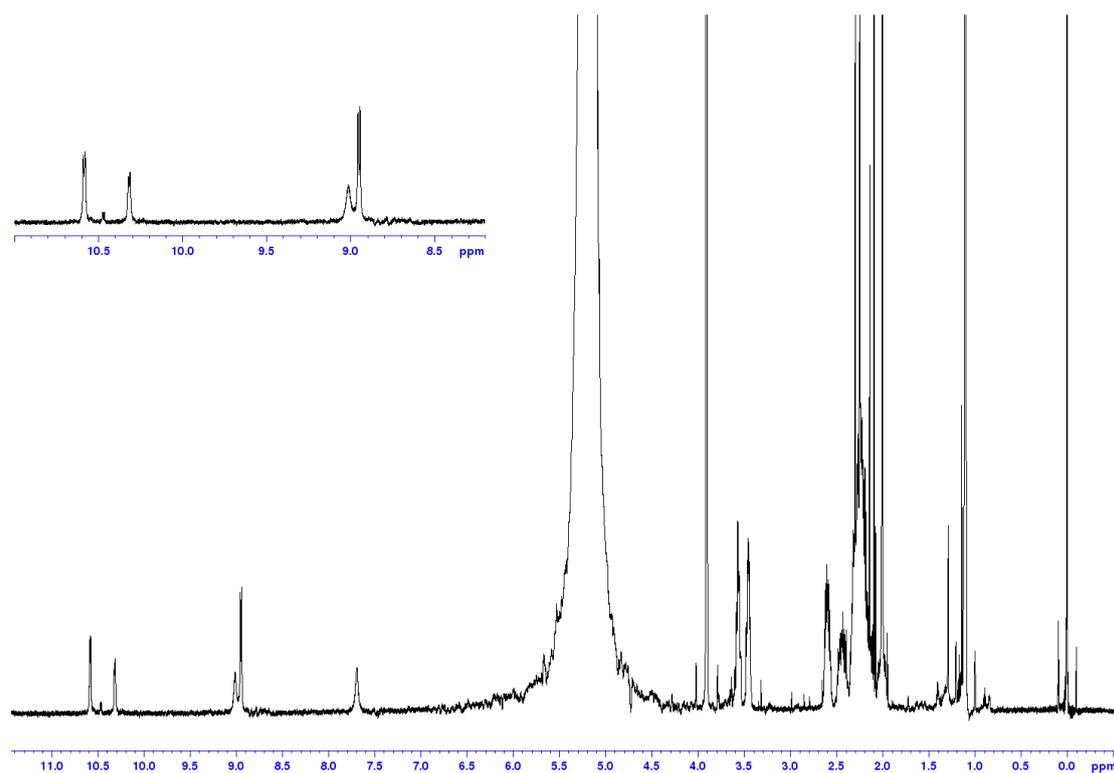


Figure S53 -  $^1\text{H}$  NMR spectrum of **M(18)-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub>** in  $i\text{PrOH-}d_7$  at 298K

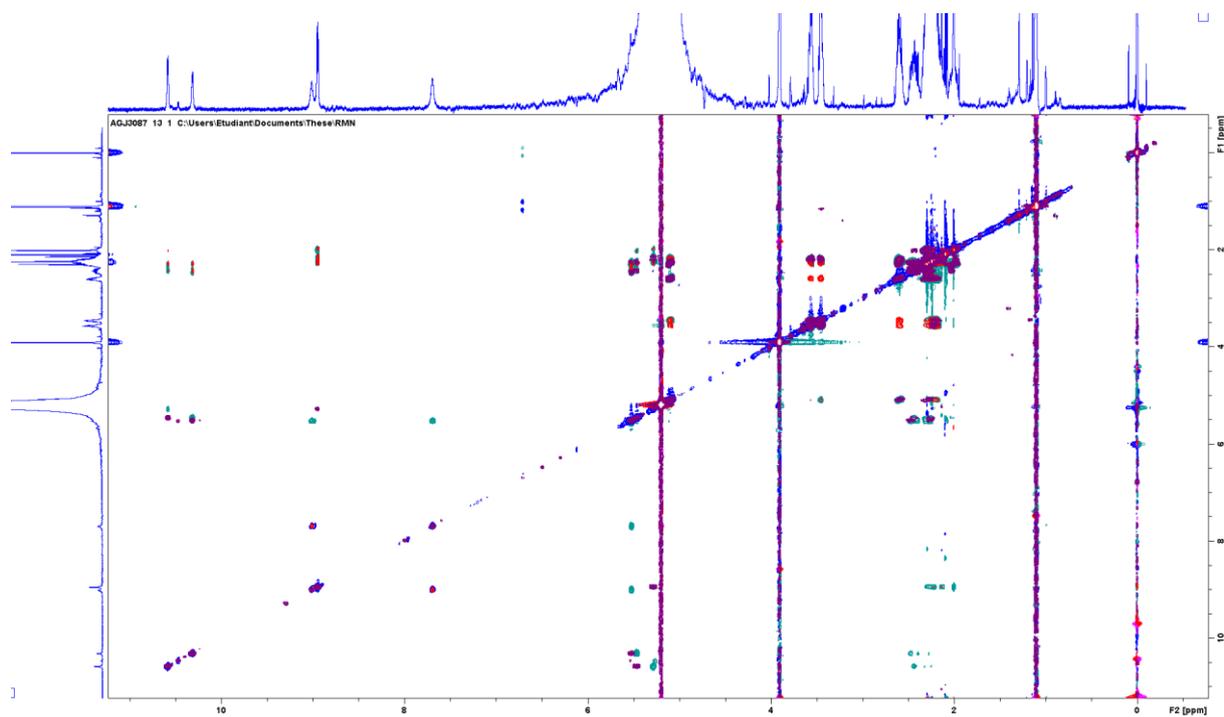
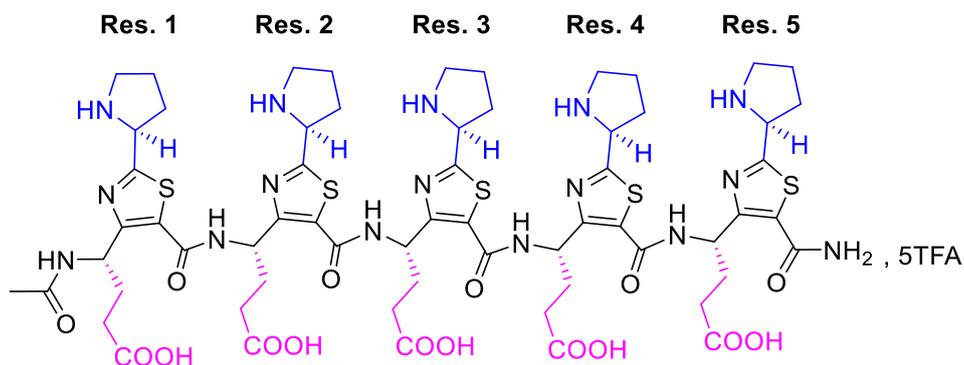


Figure S54 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(18)-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub>** in  $i\text{PrOH-}d_7$  at 298K

28. NMR analyses of **M(19)-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub>**Table S93 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (>95%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.02
Res. 1	8.97	5.27	2.19	$^{\epsilon}$ CH <sub>2</sub> 1.96 $^{\tau^1}$ CH na $^{\tau^2}$ CH <sub>2</sub> na $^{\tau^3}$ CH <sub>2</sub> na $^{\tau^4}$ CH <sub>2</sub> na
Res. 2	10.45	5.54	2.28	$^{\epsilon}$ CH <sub>2</sub> 2.06 $^{\tau^1}$ CH na $^{\tau^2}$ CH <sub>2</sub> na $^{\tau^3}$ CH <sub>2</sub> na $^{\tau^4}$ CH <sub>2</sub> na
Res. 3	10.41	5.46	2.43-2.26	$^{\epsilon}$ CH <sub>2</sub> na $^{\tau^1}$ CH 5.17 $^{\tau^2}$ CH <sub>2</sub> na $^{\tau^3}$ CH <sub>2</sub> na $^{\tau^4}$ CH <sub>2</sub> na
Res. 4	10.36	5.56	2.45-2.31	$^{\epsilon}$ CH <sub>2</sub> na $^{\tau^1}$ CH 5.11 $^{\tau^2}$ CH <sub>2</sub> na $^{\tau^3}$ CH <sub>2</sub> na $^{\tau^4}$ CH <sub>2</sub> na
Res. 5	10.20	5.51	2.43	$^{\epsilon}$ CH <sub>2</sub> na $^{\tau^1}$ CH 5.15 $^{\tau^2}$ CH <sub>2</sub> na $^{\tau^3}$ CH <sub>2</sub> na $^{\tau^4}$ CH <sub>2</sub> na
NH <sub>2</sub>	8.99-7.69	-	-	-

na: not assigned

Table S94 - Coupling Constants <sup>3</sup>J(NH,  $\gamma$ CH) (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	Major (298 K)
Res. 1	7.1
Res. 2	6.1
Res. 3	nd
Res. 4	nd
Res. 5	nd

nd: not detected

Table S95 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	o
3.NH-4.NH	o
4.NH-5.NH	o
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.Hγ-2.NH	m
2.Hγ-3.NH	m
3.Hγ-4.NH	m
4.Hγ-5.NH	m
5.Hγ-NH <sub>2</sub>	m
1.Hγ-2.Hγ	w
2.Hγ-3.Hγ	w
3.Hγ-4.Hγ	nd
4.Hγ-5.Hγ	nd
1.Hγ-2.Hτ1	nd
2.Hγ-3.Hτ1	nd
3.Hγ-4.Hτ1	w
4.Hγ-5.Hτ1	nd
1.Hγ-2.Hδ	nd
2.Hγ-3.Hδ	m
3.Hγ-4.Hδ	m
4.Hγ-5.Hδ	m

*o*: overlap with the solvent signals, *nd*: not detected

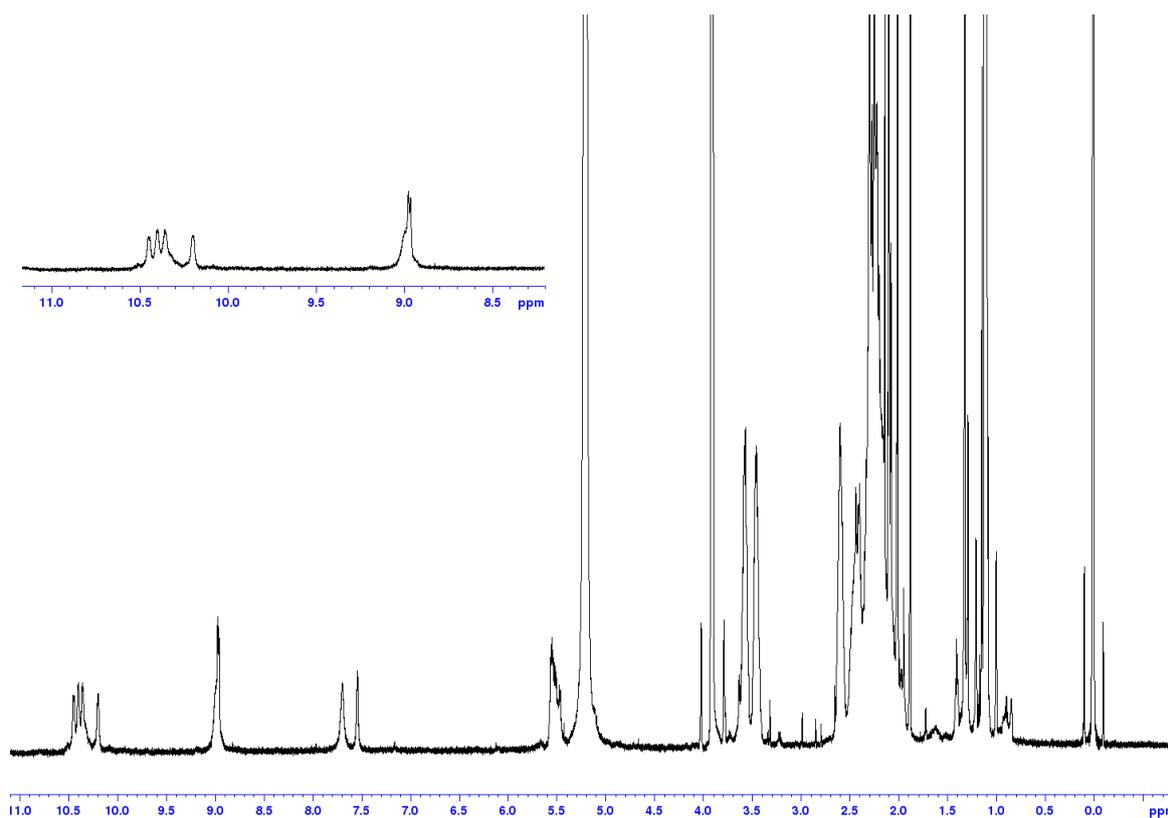


Figure S55 -  $^1\text{H}$  NMR spectrum of **M(19)-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

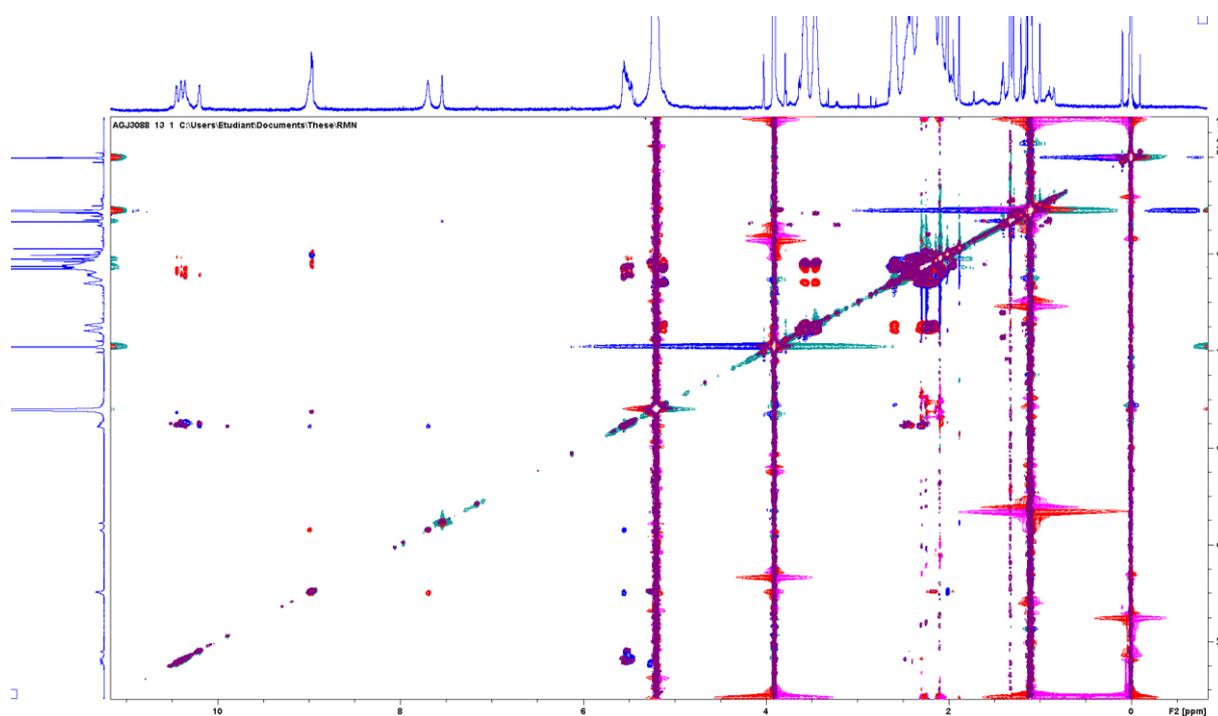
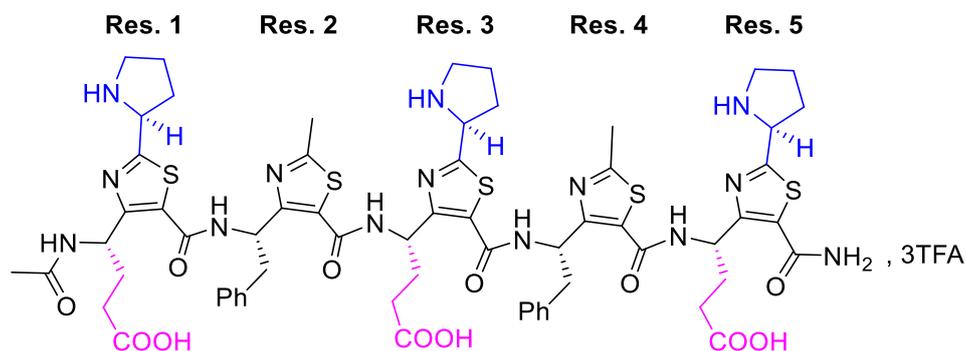


Figure S56 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(19)-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub>** in *iPrOH-d<sub>7</sub>* at 298K

29. NMR analyses of **M(20)-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub>**Table S96 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (89%)

Residue	HN	$\gamma$ CH	$\delta$ CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.05
Res. 1	8.91	5.45	2.25-2.13	$\epsilon$ CH <sub>2</sub> 1.97 $\tau^1$ CH na $\tau^2$ CH <sub>2</sub> na $\tau^3$ CH <sub>2</sub> na $\tau^4$ CH <sub>2</sub> na
Res. 2	10.66	5.57	3.31-3.20	Ph 6.86 $\tau^1$ CH 2.65
Res. 3	9.67	5.73	2.34-2.16	$\epsilon$ CH <sub>2</sub> na $\tau^1$ CH na $\tau^2$ CH <sub>2</sub> na $\tau^3$ CH <sub>2</sub> na $\tau^4$ CH <sub>2</sub> na
Res. 4	10.42	5.34	3.28-3.18	Ph 6.77 $\tau^1$ CH <sub>3</sub> 2.67
Res. 5	9.68	5.67	2.34	$\epsilon$ CH <sub>2</sub> na $\tau^1$ CH na $\tau^2$ CH <sub>2</sub> na $\tau^3$ CH <sub>2</sub> na $\tau^4$ CH <sub>2</sub> na
NH <sub>2</sub>	8.74-7.70	-	-	-

na: not assigned

Table S97 - Coupling Constants <sup>3</sup>J(NH,  $\gamma$ CH) (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	Major (298 K)
Res. 1	7.1
Res. 2	6.2
Res. 3	7.2
Res. 4	5.8
Res. 5	7.2

Table S98 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	o
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.Hγ-2.NH	s
2.Hγ-3.NH	s
3.Hγ-4.NH	s
4.Hγ-5.NH	s
5.Hγ-NH <sub>2</sub>	m
1.Hγ-2.Hγ	nd
2.Hγ-3.Hγ	w
3.Hγ-4.Hγ	nd
4.Hγ-5.Hγ	nd
1.Hγ-2.Hτ1	nd
2.Hγ-3.Hτ1	nd
3.Hγ-4.Hτ1	nd
4.Hγ-5.Hτ1	nd
1.Hγ-2.Hδ	w
2.Hγ-3.Hδ	m
3.Hγ-4.Hδ	m
4.Hγ-5.Hδ	m

*o*: overlap with the solvent signals, *nd*: not detected

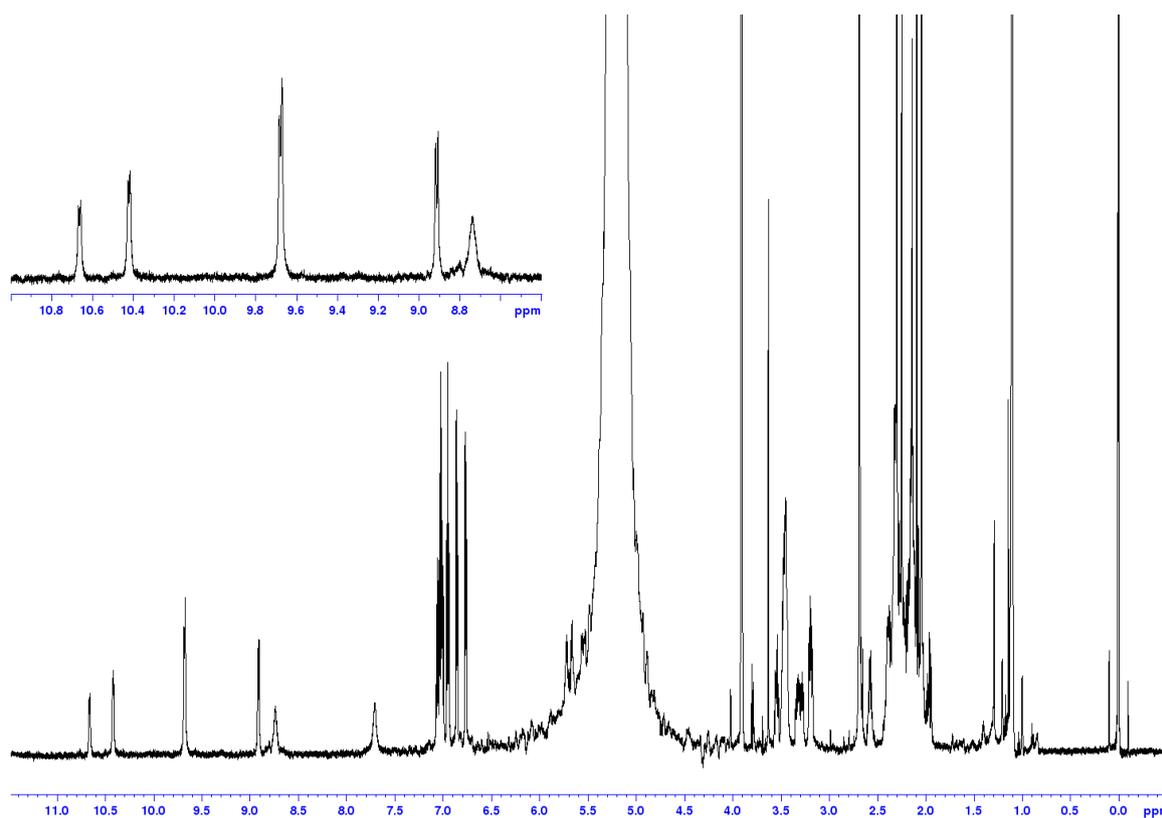


Figure S57 -  $^1\text{H}$  NMR spectrum of **M(20)-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub>** in  $i\text{PrOH-}d_7$  at 298K

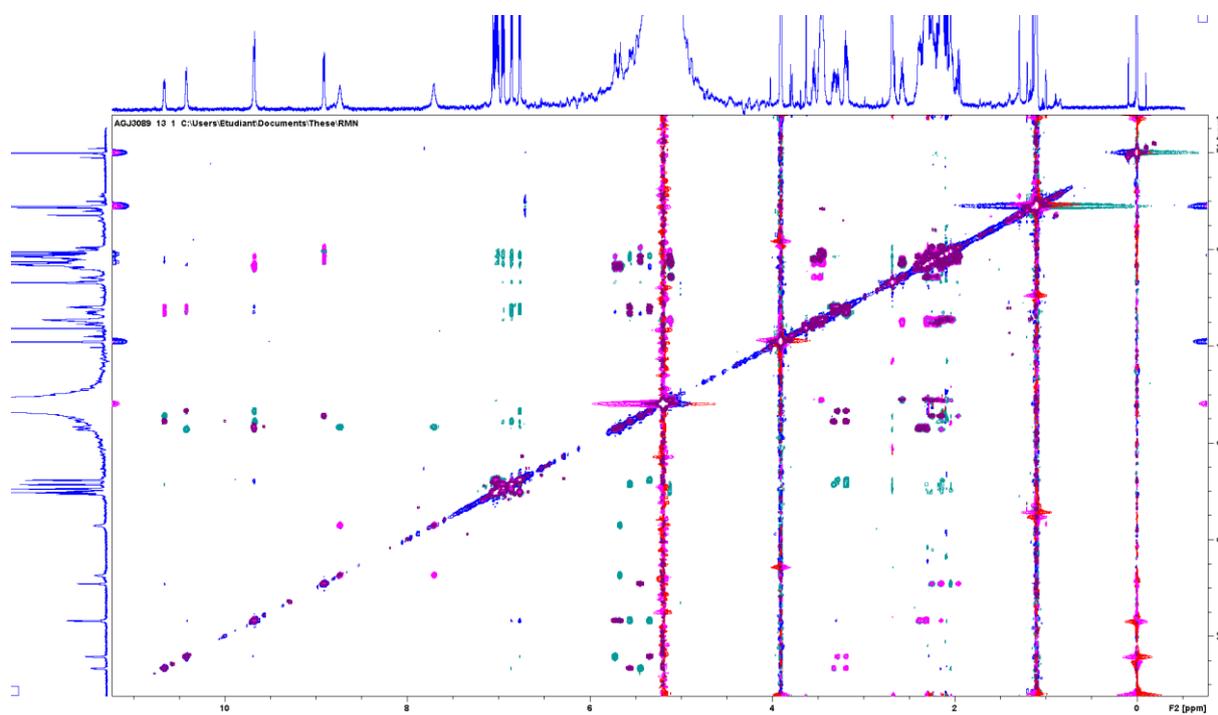
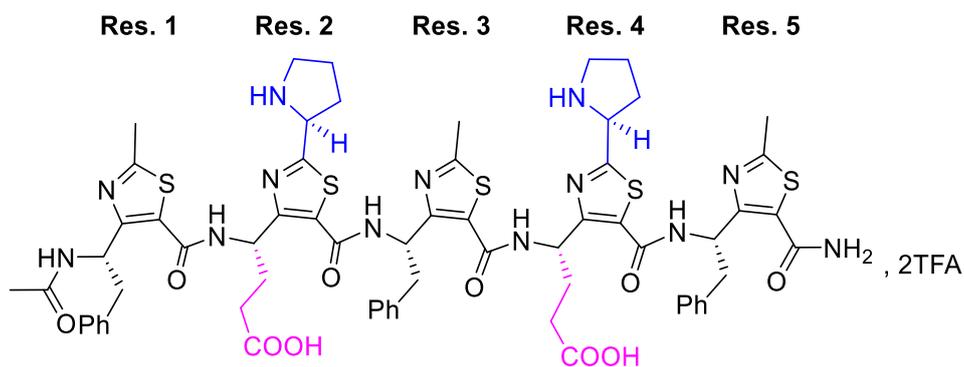


Figure S58 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of **M(20)-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub>** in  $i\text{PrOH-}d_7$  at 298K

30. NMR analyses of **M(21)-ATC<sub>5</sub>-2,4-NH<sub>2</sub>**Table S99 - <sup>1</sup>H NMR in *i*PrOH-*d*<sub>7</sub> at 298K major conformer (89%)

Residue	HN	<sup>γ</sup> CH	<sup>δ</sup> CH	Others
Ac	-	-	-	CH <sub>3</sub> 2.00
Res. 1	9.05	5.35	3.13	Ph 6.72 <sup>τ</sup> 1CH <sub>3</sub> 2.54
Res. 2	10.22	5.73	2.39-2.29	<sup>ε</sup> CH <sub>2</sub> 2.09 <sup>τ</sup> 1CH na <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na
Res. 3	10.38	5.40	3.28-3.16	Ph 6.82 <sup>τ</sup> 1CH <sub>3</sub> 2.58
Res. 4	9.70	5.74	2.36-2.27	<sup>ε</sup> CH <sub>2</sub> 2.11 <sup>τ</sup> 1CH na <sup>τ</sup> 2CH <sub>2</sub> na <sup>τ</sup> 3CH <sub>2</sub> na <sup>τ</sup> 4CH <sub>2</sub> na
Res. 5	10.16	5.80	3.34	Ph 7.24 <sup>τ</sup> 1CH <sub>3</sub> 2.80
NH <sub>2</sub>	8.50-7.36	-	-	-

na: not assigned

Table S100 - Coupling Constants <sup>3</sup>J(NH, <sup>γ</sup>CH) (in Hz). Values were measured in *i*-PrOH-*d*<sub>6</sub> at 298K.

Residue	Major (298 K)
Res. 1	7.2
Res. 2	6.9
Res. 3	6.0
Res. 4	7.2
Res. 5	7.1

Table S101 - Inter-residue NOE correlations in the major conformer observed in the ROESY spectrum in *i*-PrOH-*d*<sub>7</sub> at 298K. Strong (s) < 2.7 Å, 2.7 Å < Medium (m) < 3.3 Å, 3.3 Å < Weak (w).

NOE correlations	Intensity
1.NH-2.NH	w
2.NH-3.NH	w
3.NH-4.NH	w
4.NH-5.NH	w
5.NH-NH <sub>2</sub>	nd
Ac-1.NH	s
1.H $\gamma$ -2.NH	s
2.H $\gamma$ -3.NH	s
3.H $\gamma$ -4.NH	s
4.H $\gamma$ -5.NH	s
5.H $\gamma$ -NH <sub>2</sub>	m
1.H $\gamma$ -2.H $\gamma$	nd
2.H $\gamma$ -3.H $\gamma$	w
3.H $\gamma$ -4.H $\gamma$	nd
4.H $\gamma$ -5.H $\gamma$	nd
1.H $\gamma$ -2.H $\tau$ 1	nd
2.H $\gamma$ -3.H $\tau$ 1	w
3.H $\gamma$ -4.H $\tau$ 1	nd
4.H $\gamma$ -5.H $\tau$ 1	nd
1.H $\gamma$ -2.H $\delta$	w
2.H $\gamma$ -3.H $\delta$	nd
3.H $\gamma$ -4.H $\delta$	w
4.H $\gamma$ -5.H $\delta$	w

*o*: overlap with the solvent signals, *nd*: not detected

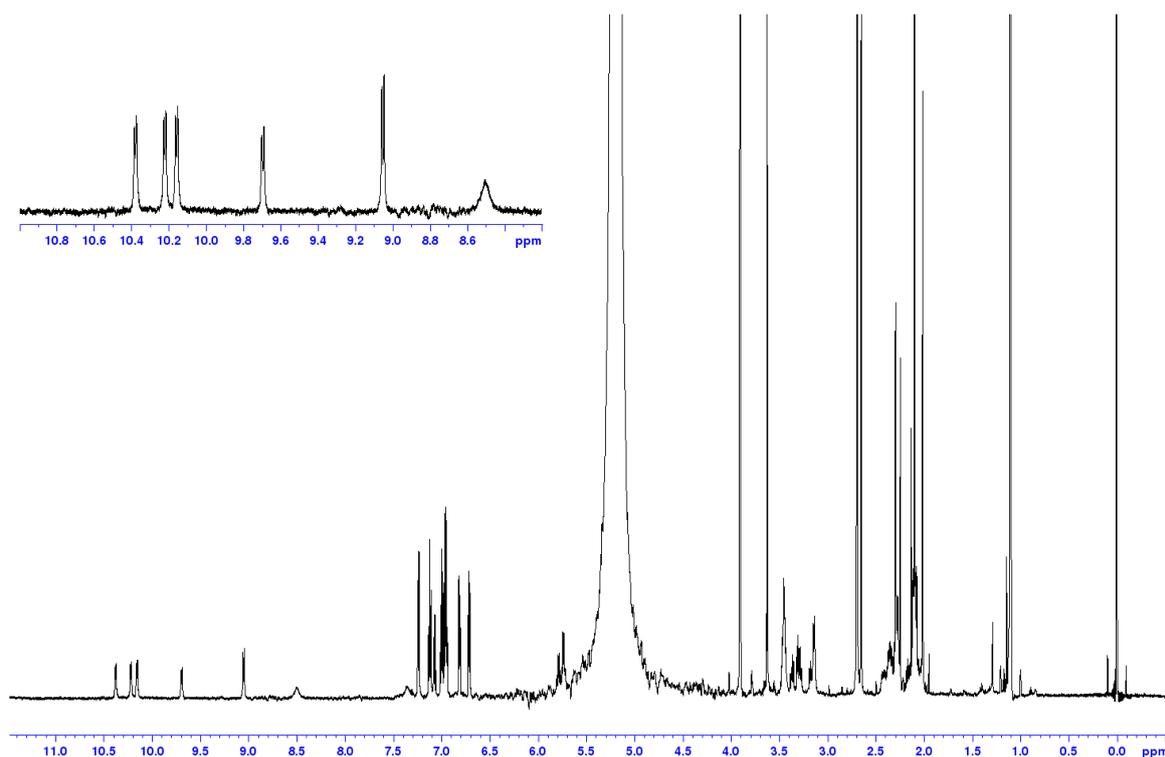


Figure S59 -  $^1\text{H}$  NMR spectrum of  $M(21)\text{-ATC}_5\text{-2,4-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

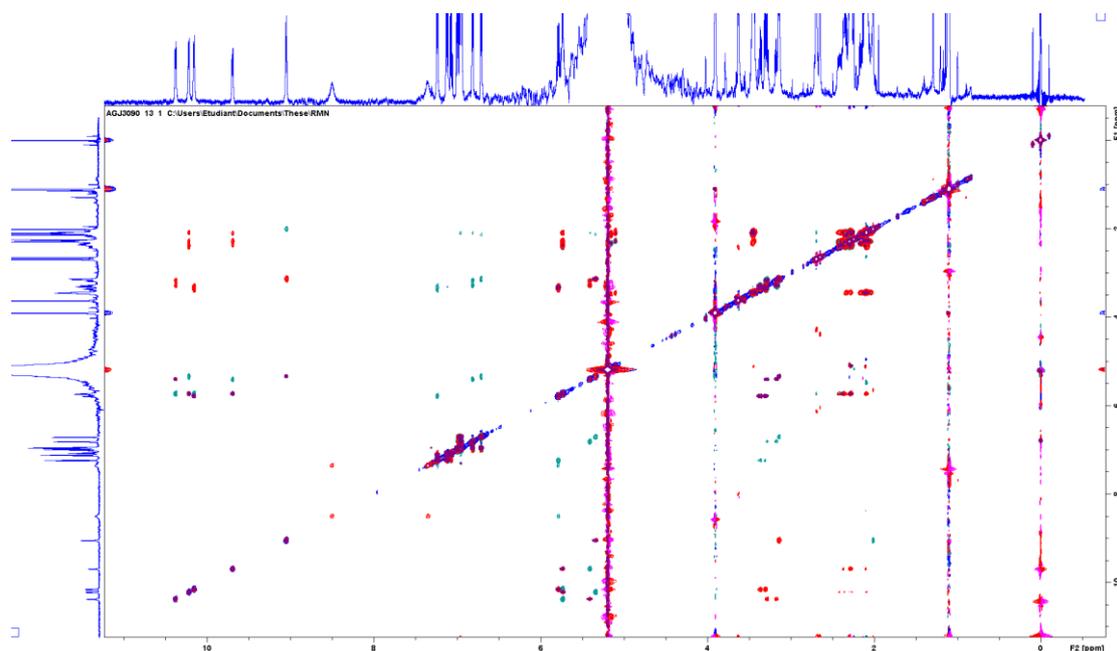
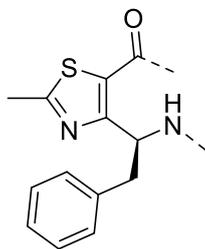


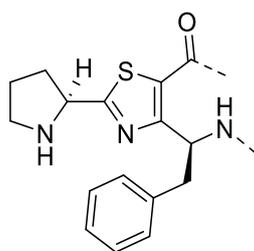
Figure S60 - Superimposition of the ROESY (Blue) and TOCSY (Red) spectra of  $M(21)\text{-ATC}_5\text{-2,4-NH}_2$  in  $i\text{PrOH-d}_7$  at 298K

*31. Restrained electrostatic potential (RESP) atomic partial charges  
calculated for ATC fragments*



Atome type	Atom coordinates			Partial charges
N	5.098439	-3.195375	0.916633	-0.684057
H	5.607378	-4.048365	0.733633	0.377414
CT	4.633614	-2.637496	-0.34998	-0.193615
H1	4.777723	-1.557076	-0.344975	0.158161
CT	5.41675	-3.240645	-1.478213	-0.322849
HC	5.070717	-2.824549	-2.424376	0.144883
HC	5.273753	-4.321213	-1.483213	0.144883
CA	6.911296	-2.922811	-1.294271	0.237704
CA	7.842259	-3.381845	-2.226269	-0.222312
HA	7.512295	-3.964721	-3.081123	0.144091
CA	9.197753	-3.092809	-2.059205	-0.143762
HA	9.92273	-3.450836	-2.784202	0.146288
CA	9.620757	-2.344886	-0.960249	-0.158384
HA	10.676221	-2.119837	-0.830180	0.142681
CA	8.69001	-1.885966	-0.029477	-0.143762
HA	9.020038	-1.302989	0.826554	0.146288
CA	7.336396	-2.174801	-0.195432	-0.222312
HA	6.610648	-1.816878	0.530342	0.144091
CC	3.177946	-2.947346	-0.528937	0.601545

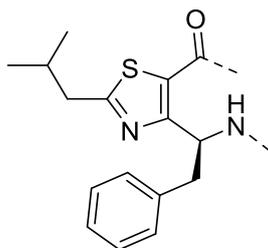
NB	2.761054	-4.197451	-0.425067	-0.578315
CA	1.466495	-4.292372	-0.608991	0.395849
CT	0.70455	-5.583182	-0.552002	-0.397407
HC	1.392482	-6.401147	-0.338035	0.130255
HC	-0.048653	-5.52615	0.235217	0.130255
HC	0.216735	-5.759085	-1.509612	0.130255
S	0.653558	-2.796342	-0.931974	0.070661
C*	2.194916	-2.001231	-0.804951	-0.642903
C	2.336944	-0.518493	-0.981915	0.931836
O	1.758124	0.152366	-1.830705	-0.598206



Atome type	Atom coordinates			Partial charges
N	9.573142	1.298542	2.075876	-0.641985
H	9.788023	0.363977	1.762036	0.366648
CT	9.393195	2.16144	0.911991	-0.292761
H1	9.88044	3.120837	1.092077	0.182286
CT	10.004135	1.509486	-0.292878	-0.471724
HC	9.870225	2.154212	-1.161497	0.174598
HC	9.518206	0.551673	-0.472847	0.174598
CA	11.507786	1.284484	-0.047916	0.349685
CA	12.292702	0.688514	-1.033798	-0.265183
HA	11.843657	0.381489	-1.974895	0.153065
CA	13.654719	0.484462	-0.811773	-0.128028

---

HA	14.266961	0.01924	-1.580082	0.144427
CA	14.230915	0.876585	0.39664	-0.177488
HA	15.292194	0.717505	0.5697	0.146176
CA	13.445678	1.473805	1.382924	-0.128028
HA	13.894596	1.778751	2.324757	0.144427
CA	12.08366	1.677859	1.160903	-0.265183
HA	11.471815	2.142773	1.929714	0.153065
CC	7.92964	2.381414	0.674045	0.960844
NB	7.124661	1.339198	0.572038	-0.677862
CA	5.882224	1.706362	0.364951	-0.024641
CT	4.709389	0.721567	0.20297	0.646208
H1	4.873355	0.146589	-0.533994	-0.075898
CT	3.540003	1.419779	0.000002	0.090184
HC	3.741107	2.49208	0.002995	-0.08611
HC	3.097086	1.129851	-0.952818	-0.08611
CT	2.660854	1.088745	1.058179	0.109299
HC	1.754581	0.63064	0.662062	-0.023257
HC	2.404812	1.98801	1.620359	-0.023257
CT	3.319929	0.198574	1.876313	0.037965
H1	3.411922	0.615522	2.879205	0.019479
H1	2.768667	-0.740811	1.918328	0.019479
NT	4.566338	-0.018542	1.348141	-0.642954
H	4.673349	-0.999925	1.130062	0.284005
S	5.621261	3.41749	0.279937	0.227223
C*	7.326657	3.629488	0.546025	-0.790349
C	7.953735	4.99057	0.611044	0.885041
O	7.676737	5.934703	-0.122042	-0.602566

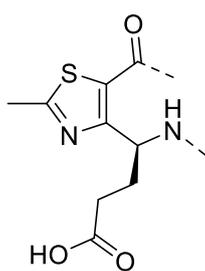


Atome type	Atom coordinates			Partial charges
N	1.371865	-0.593624	1.319873	-0.603664
CT	0.931178	-0.052906	0.037278	-0.344671
CC	-0.561283	0.097428	0.044242	0.783152
CT	1.334438	-0.987564	-1.064472	-0.434138
CA	-1.239878	1.271964	-0.267181	-0.741008
C	-0.681573	2.606143	-0.665164	0.963897
O	-1.140541	3.333804	-1.540665	-0.607009
S	-2.951392	0.998009	-0.132246	0.111959
CC	-2.592878	-0.635644	0.320447	0.331687
NB	-1.316914	-0.937555	0.363821	-0.624959
CT	-3.693848	-1.605224	0.6328	-0.352737
H	1.570492	-1.577911	1.210956	0.369519
H1	1.393687	0.920679	-0.124693	0.202868
HC	1.005491	-0.583856	-2.022057	0.164141
HC	0.871926	-1.961192	-0.90241	0.178847
HC	-3.263544	-2.572078	0.893542	0.124168
HC	-4.280712	-1.230659	1.471753	0.104182
CT	-4.594286	-1.762083	-0.587199	0.522678
CT	-5.196468	-0.410123	-0.952373	-0.613224
CT	-5.713431	-2.747054	-0.269063	-0.568486
HC	-4.007712	-2.137021	-1.425896	-0.036749
HC	-5.28355	-3.714061	-0.007855	0.110748

---

HC	-6.356936	-2.859528	-1.141523	0.129441
HC	-6.224395	-2.420363	0.460978	0.168494
HC	-4.396799	0.294868	-1.179958	0.151202
HC	-5.840151	-0.523012	-1.824701	0.134048
HC	-5.707517	-0.083705	-0.222241	0.172169
CA	3.679327	-0.024648	-1.263318	-0.233557
HA	3.231322	0.954718	-1.407957	0.162077
CA	2.866669	-1.141883	-1.071584	0.313652
CA	3.441766	-2.399113	-0.886176	-0.286891
HA	2.808662	-3.269626	-0.736557	0.162805
CA	4.829527	-2.539024	-0.892683	-0.122726
HA	5.277573	-3.518535	-0.748441	0.144949
CA	5.642274	-1.421691	-1.084369	-0.164289
HA	6.723501	-1.530759	-1.089334	0.144453
CA	5.067071	-0.164553	-1.269603	-0.127605
HA	5.700245	0.706022	-1.41886	0.144488
N	1.533758	-0.947673	0.539746	-0.718778
CT	0.789723	-0.16862	-0.445708	-0.331268
CC	-0.673405	-0.476606	-0.32534	0.766725
CT	1.263494	-0.522783	-1.824174	-0.485129
CA	-1.673041	0.478966	-0.168424	-0.743098
C	-1.545998	1.971288	-0.084965	1.010532
O	-2.298796	2.777791	-0.623007	-0.614702
S	-3.215992	-0.316314	-0.071614	0.07686
CC	-2.381612	-1.824611	-0.24807	0.393032
NB	-1.078313	-1.733035	-0.366978	-0.592437
CT	-3.134346	-3.121936	-0.259149	-0.412166

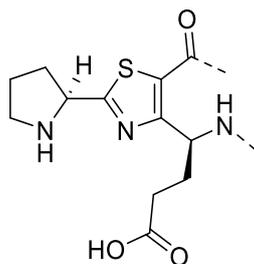
H	1.965822	-1.735258	0.078046	0.404825
H1	0.953224	0.89412	-0.267188	0.193121
HC	0.708062	0.058885	-2.559914	0.145672
HC	1.100015	-1.585592	-2.002625	0.151599
HC	-2.433324	-3.947857	-0.379062	0.1444
HC	-3.673872	-3.237018	0.681181	0.127295
HC	-3.843157	-3.12378	-1.087177	0.128951
CT	2.750382	-0.209766	-1.94642	0.657874
HC	3.050652	-0.433701	-2.818531	-0.083238
CT	3.52689	-1.023596	-0.91754	-0.615309
CT	2.97902	1.276497	-1.696414	-0.572303
HC	3.904733	1.471829	-1.772416	0.131293
HC	2.679399	1.500062	-0.824	0.168827
HC	2.495128	1.783134	-2.336592	0.135658
HC	3.226926	-0.799324	-0.045474	0.194292
HC	3.384249	-1.948927	-1.07301	0.130384
HC	4.452752	-0.828963	-0.993558	0.130437



Atome type	Atom coordinates			Partial charges
N	6.962591	4.954331	1.717126	-0.520275
H	7.66762	4.231249	1.764121	0.348015
CT	6.643564	5.217403	0.316907	-0.26161
H1	6.55162	6.292263	0.16095	0.169001

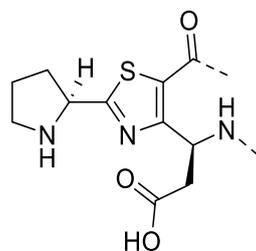
---

CT	7.738311	4.677512	-0.554899	-0.144901
HC	7.500263	4.8746	-1.601212	0.109714
HC	7.830273	3.602653	-0.398947	0.109714
CT	9.054575	5.357585	-0.197792	-0.256671
HC	8.974612	6.29353	-0.332752	0.091258
HC	9.26255	5.186547	0.712204	0.091258
C	10.167219	4.809756	-1.083502	0.750221
O	10.934141	3.939764	-0.647536	-0.546277
OH	10.311261	5.253771	-2.291446	-0.627825
HO	9.717003	5.928142	-2.628614	0.466483
CC	5.348972	4.548678	-0.034999	0.513576
NB	5.194099	3.260553	0.21307	-0.598719
CA	4.007984	2.836567	-0.153961	0.429002
CT	3.540003	1.419779	0.000002	-0.366738
HC	4.331047	0.823712	0.455035	0.122311
HC	2.656209	1.396817	0.635849	0.122311
HC	3.293886	1.009601	-0.979427	0.122311
S	2.975954	4.041695	-0.851004	0.0619
C*	4.260051	5.189692	-0.618966	-0.57005
C	4.117077	6.627992	-1.020027	0.884374
O	3.582239	7.026924	-2.049751	-0.590998



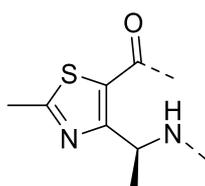
Atome type	Atom coordinates			Partial charges
N	6.962591	4.954331	1.717126	-0.520275
H	7.66762	4.231249	1.764121	0.348015
CT	6.643564	5.217403	0.316907	-0.26161
H1	6.55162	6.292263	0.16095	0.169001
CT	7.738311	4.677512	-0.554899	-0.144901
HC	7.500263	4.8746	-1.601212	0.109714
HC	7.830273	3.602653	-0.398947	0.109714
CT	9.054575	5.357585	-0.197792	-0.256671
HC	8.974612	6.29353	-0.332752	0.091258
HC	9.26255	5.186547	0.712204	0.091258
C	10.167219	4.809756	-1.083502	0.750221
O	10.934141	3.939764	-0.647536	-0.546277
OH	10.311261	5.253771	-2.291446	-0.627825
HO	9.717003	5.928142	-2.628614	0.466483
CC	5.348972	4.548678	-0.034999	0.513576
NB	5.194099	3.260553	0.21307	-0.598719
CA	4.007984	2.836567	-0.153961	0.429002
CT	3.540003	1.419779	0.000002	-0.366738
HC	4.331047	0.823712	0.455035	0.122311
HC	2.656209	1.396817	0.635849	0.122311
HC	3.293886	1.009601	-0.979427	0.122311

S	2.975954	4.041695	-0.851004	0.0619
C*	4.260051	5.189692	-0.618966	-0.57005
C	4.117077	6.627992	-1.020027	0.884374
O	3.582239	7.026924	-2.049751	-0.590998



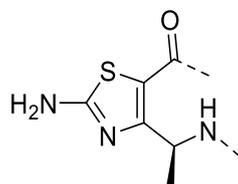
Atome type	Atom coordinates			Partial charges
N	10.068303	4.222476	-1.29704	-0.706307
H	9.180502	4.692401	-1.166083	0.323289
CT	10.104257	2.780113	-1.117012	0.016837
H1	10.853233	2.390108	-1.810972	0.118897
CT	10.54133	2.405965	0.310399	-0.052525
HC	11.458076	2.959767	0.533369	0.083635
HC	9.774047	2.728098	1.023647	0.083635
CT	10.806235	0.913107	0.443375	-0.170879
HC	9.898111	0.314087	0.302328	0.097791
HC	11.519232	0.576085	-0.320599	0.097791
C	11.377359	0.540948	1.781776	0.729613
O	11.59041	1.306037	2.69995	-0.583185
OH	11.640207	-0.773650	1.861734	-0.680775
HO	12.007232	-0.95972	2.747852	0.474049
CC	8.741365	2.215217	-1.434024	0.435808
NB	7.647352	2.967079	-1.063118	-0.513845
CA	6.534704	2.327339	-1.27607	0.313043

CT	5.158954	2.877297	-0.995139	0.019877
H1	5.022019	3.754063	-1.634958	0.092756
CT	4.998923	3.226367	0.494151	0.006117
HC	5.921625	3.601208	0.942012	0.028032
HC	4.222691	3.991656	0.598178	0.028032
CT	4.51201	1.900747	1.074991	-0.121271
HC	5.351372	1.201411	1.17404	0.06154
HC	4.035019	2.002771	2.051994	0.06154
CT	3.540003	1.419779	0.000002	0.188672
H1	3.399007	0.335052	0.017989	0.019034
H1	2.55889	1.886875	0.143018	0.019034
NT	4.144853	1.848642	-1.286644	-0.707477
H	3.423918	2.255639	-1.875593	0.361749
S	6.723609	0.735742	-1.921915	0.100809
C*	8.439428	0.978638	-1.937878	-0.590114
C	9.307308	-0.168291	-2.33084	0.863355
O	9.006137	-1.318736	-1.986724	-0.564065



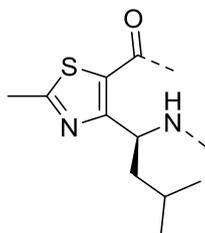
<b>Atome type</b>	<b>Atom coordinates</b>			<b>Partial charges</b>
N	5.102519	2.201464	1.569542	-0.683301
H	4.920461	1.384366	2.134614	0.364863
CT	6.103412	3.028322	2.237464	0.051622
CT	6.189373	2.638346	3.68333	-0.420711
HC	6.936128	3.256097	4.182169	0.133769

HC	6.474409	1.589098	3.760354	0.133769
HC	5.344792	2.768302	4.096121	0.133769
H1	5.817472	4.077252	2.159462	0.109557
CC	7.438214	2.823303	1.584569	0.499553
NB	7.910288	1.59797	1.44054	-0.587123
CA	9.082486	1.597922	0.853451	0.413556
CT	9.869626	0.355596	0.553388	-0.405055
HC	9.315423	-0.517643	0.897498	0.131221
HC	10.034598	0.280596	-0.521441	0.131221
HC	10.830341	0.401542	1.066238	0.131221
S	9.696661	3.161155	0.426388	0.061374
C*	8.247176	3.846169	1.100621	-0.592112
C	8.003237	5.326167	1.110617	0.87973
O	8.847517	6.182399	1.359701	-0.588065



Atome type	Atom coordinates			Partial charges
N	5.166287	2.056028	1.570994	-0.694829
H	4.915367	1.254362	2.131771	0.36509
CT	6.235275	2.789955	2.241975	0.059926
CT	6.289239	2.386998	3.68479	-0.392918
HC	7.087261	2.934955	4.185793	0.125288
HC	6.483241	1.316755	3.755811	0.125288
HC	5.458811	2.586143	4.100999	0.125288
H1	6.040322	3.85996	2.169989	0.113575

CC	7.546387	2.474882	1.584925	0.422438
NB	7.911215	1.215313	1.433975	-0.588345
CR	9.079083	1.118283	0.844047	0.672521
N2	9.680014	-0.038689	0.571059	-1.094466
H	10.580759	-0.043719	0.114186	0.476863
H	9.234783	-0.910074	0.821174	0.476863
S	9.824353	2.625667	0.424935	0.012384
C*	8.440383	3.427809	1.104961	-0.593574
C	8.323427	4.924137	1.122973	0.887974
O	9.238438	5.703067	1.375968	-0.593955



Atome type	Atom coordinates			Partial charges
N	2.605057	2.141749	-2.029169	-0.662169
H	3.037172	1.354483	-2.491325	0.371139
CT	1.861248	2.920612	-3.014928	-0.193046
CT	2.334126	2.566704	-4.392656	-0.424388
HC	1.77922	3.148662	-5.128526	0.13185
HC	2.17112	1.50393	-4.571655	0.13185
CT	3.820673	2.879557	-4.514631	0.546334
HC	4.121581	2.655618	-5.387446	-0.065167
CT	4.59757	2.065598	-3.48678	-0.512069
HC	5.523227	2.2605	-3.562766	0.124626
HC	4.297632	2.290555	-2.613893	0.124626

HC	4.454495	1.140348	-3.641834	0.124626
CT	4.049696	4.365363	-4.264623	-0.512069
HC	4.976129	4.561427	-4.340658	0.124626
HC	3.565616	4.87248	-4.905715	0.124626
HC	3.749704	4.589351	-3.392594	0.124626
H1	2.024247	3.983392	-2.835954	0.153924
CC	0.398198	2.612644	-2.893934	0.62371
NB	-0.006723	1.356988	-2.935961	-0.582099
CA	-1.311077	1.264997	-2.816932	0.395784
CT	-2.062977	-0.03174	-2.827973	-0.411713
HC	-1.361989	-0.857761	-2.947994	0.13352
HC	-2.6029	-0.146737	-1.888101	0.13352
HC	-2.771929	-0.033696	-3.655912	0.13352
S	-2.144998	2.773928	-2.640893	0.064489
C*	-0.601542	3.568438	-2.736958	-0.643104
C	-0.474492	5.060739	-2.653915	0.921537
O	-1.22755	5.867835	-3.191952	-0.597714

### Example of a .gjf file

```
# M062X/6-311++g(d) geom=connectivity
Pop=MK iop(6/33=2) iop(6/42=6) iop(6/50=1)
```

```
Ac-[ATC-pyr(d)-1]-NHMe
```

```
0 1
S      -1.93170000 -1.57420000 -0.80170000
N      -1.57800000  0.95090000 -0.53170000
C      -0.32110000  0.38980000 -0.57730000
C      -2.50890000  0.04870000 -0.65160000
C      -0.30810000 -0.97090000 -0.73570000
C      -3.98880000  0.33480000 -0.65910000
.....
C      3.10350000 -2.51150000 -1.37780000
H      1.93750000 -0.80790000 -1.96300000
H      2.92390000 -3.38920000 -1.96280000
H      3.93770000 -1.97760000 -1.78280000
H      3.31680000 -2.79350000 -0.36790000
```

```
1 4 1.0 5 1.0
2 3 1.5 4 2.0
```

3 5 2.0 19 1.0  
 4 6 1.0  
 5 39 1.0  
 6 7 1.0 8 1.0 18 1.0  
 7 10 1.0 11 1.0 12 1.0  
 8  
 9 10 1.0 13 1.0 14 1.0 18 1.0  
 10 15 1.0 16 1.0  
 11  
 12  
 13  
 14  
 15  
 .....

### *32. Script for NMR solution calculations*

*An example of the script used will be given here, of course, name of residues are arbitrary chosen here and the name of the folder in which files are saved are given for understanding of the script but can be renamed.*

1/ Create .mol2 files of ATC-monomers using Chimera and save them in ./Bureau/Julie/ATC/init

2/ Open terminal : cd Bureau/Julie/ATC

3/ Launch the following script by replacing “name\_of\_residue” with the appropriate three-letter code:

```
#Antechamber for 1 ATC-monomer
#Creation of the parameters and coordinates files from a .mol2 file (Chimera)
antechamber -i ./init/name_of_residue.mol2 -fi mol2 -o ./prep/name_of_residue.prepin -fo prepi -c bcc -s 2 >
./prep/antechamber.out
parmchk -i ./prep/name_of_residue.prepin -f prepi -o ./prep/name_of_residue.frcmod
cat > leap.in <<EOF
source leaprc.gaff
source leaprc.ff99SBnmr
loadamberprep ./prep/name_of_residue.prepin
loadamberparams ./prep/name_of_residue.frcmod
saveamberparm name_of_residue ./prep/name_of_residue.prmtop ./prep/name_of_residue.inpcrd
quit
EOF
tleap -f leap.in > ./prep/leap.out
ambpdb -a atm -p ./prep/name_of_residue.prmtop < ./prep/name_of_residue.inpcrd >
./prep/name_of_residue.pdb
```

```

cat > leap.in <<EOF
source leaprc.gaff
source leaprc.ff99SBnmr
loadamberprep ./prep/name_of_residue.prepin
loadamberparams ./prep/name_of_residue.frcmod
#We identify the atoms to remove and the one that will be used as connectors for the creation of sequences.
remove name_of_residue name_of_residue.1.HH31
remove name_of_residue name_of_residue.1.HH32
remove name_of_residue name_of_residue.1.HH33
remove name_of_residue name_of_residue.1.CX2
remove name_of_residue name_of_residue.1.CX
remove name_of_residue name_of_residue.1.OX
remove name_of_residue name_of_residue.1.NX
remove name_of_residue name_of_residue.1.HNX
remove name_of_residue name_of_residue.1.CX3
remove name_of_residue name_of_residue.1.HH34
remove name_of_residue name_of_residue.1.HH35
remove name_of_residue name_of_residue.1.HH36
set name_of_residue head name_of_residue.1.N
set name_of_residue tail name_of_residue.1.C
set name_of_residue.1 connect0 name_of_residue.1.N
set name_of_residue.1 connect1 name_of_residue.1.C
#We now create a sequence containing only the residue (ATC-monomer) without its extremities to save the
#corresponding .prepin/.prmtop/.inpcrd files
name_of_residue = sequence { name_of_residue }
saveamberprep name_of_residue ./Librairie_ATC_sans_cap/name_of_residue.prepin
saveamberparm name_of_residue ./Librairie_ATC_sans_cap/name_of_residue.prmtop
./Librairie_ATC_sans_cap/name_of_residue.inpcrd
quit
EOF
tleap -f leap.in > ./prep/leap.out
parmchk -i ./prep/name_of_residue.prepin -f prep1 -o ./Librairie_ATC_sans_cap/name_of_residue.frcmod

```

4/ In the terminal: cd ..

```

#We must be in terminal ./Bureau/Julie/
#Replace "name_of_sequence" by the appropriate same
parmchk -i ./ATC/prep/lambda.prepin -f prep1 -o ./Sequences/prep/name_of_sequence1.frcmod

```

```
parmchk -i ./ATC/rep/lambda.prepin -f prepi -o ./Sequences/rep/name_of_sequence2.frmod
parmchk -i ./ATC/rep/lambda.prepin -f prepi -o ./Sequences/rep/name_of_sequence3.frmod
parmchk -i ./ATC/rep/lambda.prepin -f prepi -o ./Sequences/rep/AGJ2180.frmod
parmchk -i ./ATC/rep/lambda.prepin -f prepi -o ./Sequences/rep/AGJ2053.frmod
parmchk -i ./ATC/rep/lambda.prepin -f prepi -o ./Sequences/rep/AGJ2167.frmod
#Association of residues from the ATC-library without capping then minimisation of the obtained structure
cat > leap.in <<EOF
source leaprc.gaff
source leaprc.ff99SBnmr
#load all the prepin and frmod files needed (every useful residue)
#as well as the ACE and NHE extremities modified to match the same nomenclature as ours (ACE_1 and
NHE_1)
loadamberprep ./ATC/Librairie_ATC_sans_cap/LF1.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LF2.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LF3.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LF4.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LF5.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LA1.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LA6.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LL1.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LE1.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LE2.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/LD2.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/ACE_1.prepin
loadamberprep ./ATC/Librairie_ATC_sans_cap/NHE_1.prepin
loadamberparams ./ATC/Librairie_ATC_sans_cap/LF1.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LF2.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LF3.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LF4.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LF5.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LA1.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LA6.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LL1.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LE1.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LE2.frmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/LD2.frmod
```

```

loadamberparams ./ATC/Librairie_ATC_sans_cap/ACE_1.frcmod
loadamberparams ./ATC/Librairie_ATC_sans_cap/NHE_1.frcmod
#change the name of the sequence and the order and name of the specified residues
#As many sequences as we want can be created at the same time, it is just needed to add new lines and
double the following command lines to save the new sequences files
name_of_sequence1 = sequence { ACE LF1 LF1 LE2 LF1 LF1 NHE }
name_of_sequence2 = sequence { ACE LF1 LF1 LF1 LE2 LF1 NHE }
name_of_sequence3 = sequence { ACE LF1 LF1 LF1 LF1 LE2 NHE }
saveamberprep name_of_sequence1 ./Sequences/prep/name_of_sequence1.prepin
saveamberprep name_of_sequence2 ./Sequences/prep/name_of_sequence2.prepin
saveamberprep name_of_sequence3 ./Sequences/prep/name_of_sequence3.prepin
saveamberparm name_of_sequence1 ./Sequences/prep/name_of_sequence1.prmtop
./Sequences/prep/name_of_sequence1.inpcrd
saveamberparm name_of_sequence2 ./Sequences/prep/name_of_sequence2.prmtop
./Sequences/prep/name_of_sequence2.inpcrd
saveamberparm name_of_sequence3 ./Sequences/prep/name_of_sequence3.prmtop
./Sequences/prep/name_of_sequence3.inpcrd
quit
EOF
tleap -f leap.in > ./Sequences/prep/leap.out
#Minimisation of the structure
sander -O -i min.in -o ./Sequences/prep/name_of_sequence1.out -c
./Sequences/prep/name_of_sequence1.inpcrd -p ./Sequences/prep/name_of_sequence1.prmtop -r
./Sequences/prep/name_of_sequence1.rst
ambpdb -aatm -p ./Sequences/prep/name_of_sequence1.prmtop -c ./Sequences/prep/name_of_sequence1.rst
> ./Sequences/prep/name_of_sequence1.pdb
sander -O -i min.in -o ./Sequences/prep/name_of_sequence2.out -c
./Sequences/prep/name_of_sequence2.inpcrd -p ./Sequences/prep/name_of_sequence2.prmtop -r
./Sequences/prep/name_of_sequence2.rst
ambpdb -aatm -p ./Sequences/prep/name_of_sequence2.prmtop -c ./Sequences/prep/name_of_sequence2.rst
> ./Sequences/prep/name_of_sequence2.pdb
sander -O -i min.in -o ./Sequences/prep/name_of_sequence3.out -c
./Sequences/prep/name_of_sequence3.inpcrd -p ./Sequences/prep/name_of_sequence3.prmtop -r
./Sequences/prep/name_of_sequence3.rst
ambpdb -aatm -p ./Sequences/prep/name_of_sequence3.prmtop -c ./Sequences/prep/name_of_sequence3.rst
> ./Sequences/prep/name_of_sequence3.pdb

```

5/ Copy in Bureau/Julie the folder named “minimisation-igb1” containing file 5 empty folders named “init”, “min”, “prep”, “sa1” and “sa2” and parametrizing files (“cook\_heat.in”, “min.in”, “map.C71” and “sa.in”, *scripts will be given at the end of this procedure*)

6/ Create a “NOE restraints” text file named “contraintes.txt” (*example of this file will be given at the end of this procedure*)

7/ Open “prep” folder and copy-paste the 7 previously obtained files (.frcmmod, .inpcrd, .out, .pdb, .prepin, .prmtop, .rst) for the desired sequence (they are located in Bureau/Julie/Sequences/prep)

```
#PART 1
#MD at 1200K to generate initial structures
sander -O -i cook_heat.in -o ./init/cook_heat.out -p ./prep/name_of_sequence.prmtop -c
./prep/name_of_sequence.rst -x ./init/name_of_sequence.mdcrd
cat > trajin <<EOF
trajin ./init/name_of_sequence.mdcrd
trajout ./init/name_of_sequence.rst restart
EOF
cpptraj ./prep/name_of_sequence.prmtop < trajin
#Minimisation of the initial structures
for x in `seq 1 100`
do
sander -O -i min.in -o ./min/out/name_of_sequence_$x.out -c ./init/name_of_sequence.rst.$x -p
./prep/name_of_sequence.prmtop -r ./min/name_of_sequence_$x.rst
done
#Preparation of the distance file
makeDIST_RST -upb contraintes.txt -pdb ./prep/name_of_sequence.pdb -rst name_of_sequence.dist -map
map.C71
#PART 2
#Here we can modify the potential r4 strong between Hgamma(i) and NH(i+1) from 3.20 a 3.50
#edit the sa.in file by changing at the bottom of the page DISTANG=“name.dist” by
DISTANG=“name_of_sequence”.dist
# Simulated annealing under NMR refinement (at least twice, until stabilisation of energies and gradients)
for x in `seq 1 100`
do
sander -O -i sa.in -o ./sa1/out/name_of_sequence_$x.out -c ./min/name_of_sequence_$x.rst -p
./prep/name_of_sequence.prmtop -r ./sa1/name_of_sequence_$x.rst -x ./sa1/traj/name_of_sequence_$x.traj
done
for x in `seq 1 100`
do
sander -O -i sa.in -o ./sa2/out/name_of_sequence_$x.out -c ./sa1/name_of_sequence_$x.rst -p
./prep/name_of_sequence.prmtop -r ./sa2/name_of_sequence_$x.rst -x ./sa2/traj/name_of_sequence_$x.traj
```

```
done
#Conversion from .rst to .pdb
for x in `seq 1 100`
do
ambpdb -aatm -p ./prep/name_of_sequence.prmtop -c ./sa2/name_of_sequence_$.rst >
./sa2/name_of_sequence_$.pdb
done
```

8/ Go to the terminal:

```
cd sa2/out
sviol *out>viol.out
gedit viol.out
```

9/ In this file, we chose the 20 structures of lowest energy without violation greater than 0.3 Å

10/ Creation in folder sa2 of a list with these 20 structures called file\_best.nam, containing name\_of\_sequence\_XX.pdb

11/ In the terminal:

```
cd ..
molmol
```

12/ Save the models as .pdb

“[cook\\_heat.in](#)”

```
Cook from heating up the system equilibration stage 1
&cntrl
nstlim=200000, dt=0.002, ntx=1, irest=0, npr=2000, ntwr=2000, ntwx=2000, nmropt=0,
tempi =1000.0, temp0=1000.0, ntt=1, tautp=2.0, ig=209858,
igb=1, ntb=0, ntp=0, cut=15.0
ntc=2, ntf=2,
&end
nmr restraints
&wt type='REST', istep1=0,istep2=200000,value1=1.0
value2=1.0, &end
&wt type='END' &end
LISTOUP=POUT
DISANG=heat.dist
```

“[min.in](#)”

```
&cntrl
imin=1, maxcyc=1000, ncyc=50, ntp=20,
ntb=0, cut=15.0, /
&ewald
eedmeth=5, /
```

“[map.C71](#)” – *All the atom names must be generated, the script is quite long to cover every possible atom given, a specific selection could actually, a short example will be provided here.*

```
RESIDUE ACE
MAPPING C = C
MAPPING O = O
MAPPING CH3 = CH3
MAPPING HH31 = HH31
MAPPING HH32 = HH32
MAPPING HH33 = HH33
RESIDUE NHE
MAPPING N = N
MAPPING HN1 = HN1
MAPPING HN2 = HN2
RESIDUE LF2
MAPPING C1 = C1
MAPPING C2 = C2
MAPPING C3 = C3
MAPPING C4 = C4
MAPPING C5 = C5
MAPPING C6 = C6
MAPPING C7 = C7
MAPPING C8 = C8
...
```

“[sa.in](#)”

simulated annealing protocol, 20 ps

```
&cntrl
nstlim=20000, pencut=-0.001, nmropt=1,
ntp=200, ntt=1, ntwx=200, igb=1
cut=15.0, ntb=0, vlimit=10, /
&ewald
```

```
eedmeth=5, /
#Simple simulated annealing algorithm:
#from steps 0 to 5000: heat the system to 600K (1000K bl)
#from steps 5001-18000: re-cool to low temperatures with long tautp
#from steps 18001-20000: final cooling with short tautp
&wt type='TEMP0', istep1=0,istep2=5000,value1=1000.,
    value2=1000., /
&wt type='TEMP0', istep1=5001, istep2=18000, value1=1000.0,
    value2=100.0, /
&wt type='TEMP0', istep1=18001, istep2=20000, value1=0.0,
    value2=0.0, /
&wt type='TAUTP', istep1=0,istep2=5000,value1=0.4,
    value2=0.4, /
&wt type='TAUTP', istep1=5001,istep2=18000,value1=4.0,
    value2=4.0, /
&wt type='TAUTP', istep1=18001,istep2=19000,value1=1.0,
    value2=1.0, /
&wt type='TAUTP', istep1=19001,istep2=20000,value1=0.1,
    value2=0.05, /
&wt type='REST', istep1=0,istep2=3000,value1=0.1,
    value2=1.0, /
&wt type='REST', istep1=3001,istep2=20000,value1=1.0,
    value2=1.0, /
&wt type='END' /
LISTOUT=POUT
DISANG=nom.dist
“contraintes.txt”
#limits 2.7, 3.3, 5
#NHi - NHi+1 (penalty +1 for the 2H of NHE)
2 res1 HN 3 res2 HN 5
3 res2 HN 4 res3 HN 5
4 res3 HN 5 res4 HN 5
5 res4 HN 6 res5 HN 5
#6 res5 HN 7 NHE HN1 6
#6 res5 HN 7 NHE HN2 6
#Ac - NH1 (penalite +1)
```

1 ACE HH31 2 res1 HN 3.7

1 ACE HH32 2 res1 HN 3.7

1 ACE HH33 2 res1 HN 3.7

#HG<sub>i</sub> - NH<sub>i+1</sub> (penalty +1 for the 2H of NHE)

2 res1 HG 3 res2 HN 2.7

3 res2 HG 4 res3 HN 3.3

4 res3 HG 5 res4 HN 2.7

5 res4 HG 6 res5 HN 2.7

6 res5 HG 7 NHE HN1 2.7

6 res5 HG 7 NHE HN2 3.3

#HG<sub>i</sub> - HG<sub>i+1</sub>

#2 res1 HG 3 res2 HG 5

#3 res2 HG 4 res3 HG 5

#4 res3 HG 5 res4 HG 5

#5 res4 HG 6 res5 HG 5

#HG<sub>i</sub> - HThiazolei+1 (penalty +1)

2 res1 HG 3 res2 HT1 6

2 res1 HG 3 res2 HT2 6

2 res1 HG 3 res2 HT3 6

#3 res2 HG 4 res3 HT1 6

4 res3 HG 5 res4 HT1 6

4 res3 HG 5 res4 HT2 6

4 res3 HG 5 res4 HT3 6

5 res4 HG 6 res5 HT1 6

5 res4 HG 6 res5 HT2 6

5 res4 HG 6 res5 HT3 6

#HG<sub>i</sub> - Hdelti+1 (penalty +1)

#2 res1 HG 3 res2 HD1 6

#2 res1 HG 3 res2 HD2 6

#3 res2 HG 4 res3 HD1 6

#3 res2 HG 4 res3 HD2 6

#4 res3 HG 5 res4 HD1 6

#4 res3 HG 5 res4 HD2 6

#5 res4 HG 6 res5 HD1 6

#5 res4 HG 6 res5 HD2 6

#HAr(i) - HThiazolei+1 (penalty +1)

2 res1 H1 3 res2 HT1 4.3

2 res1 H1 3 res2 HT2 4.3

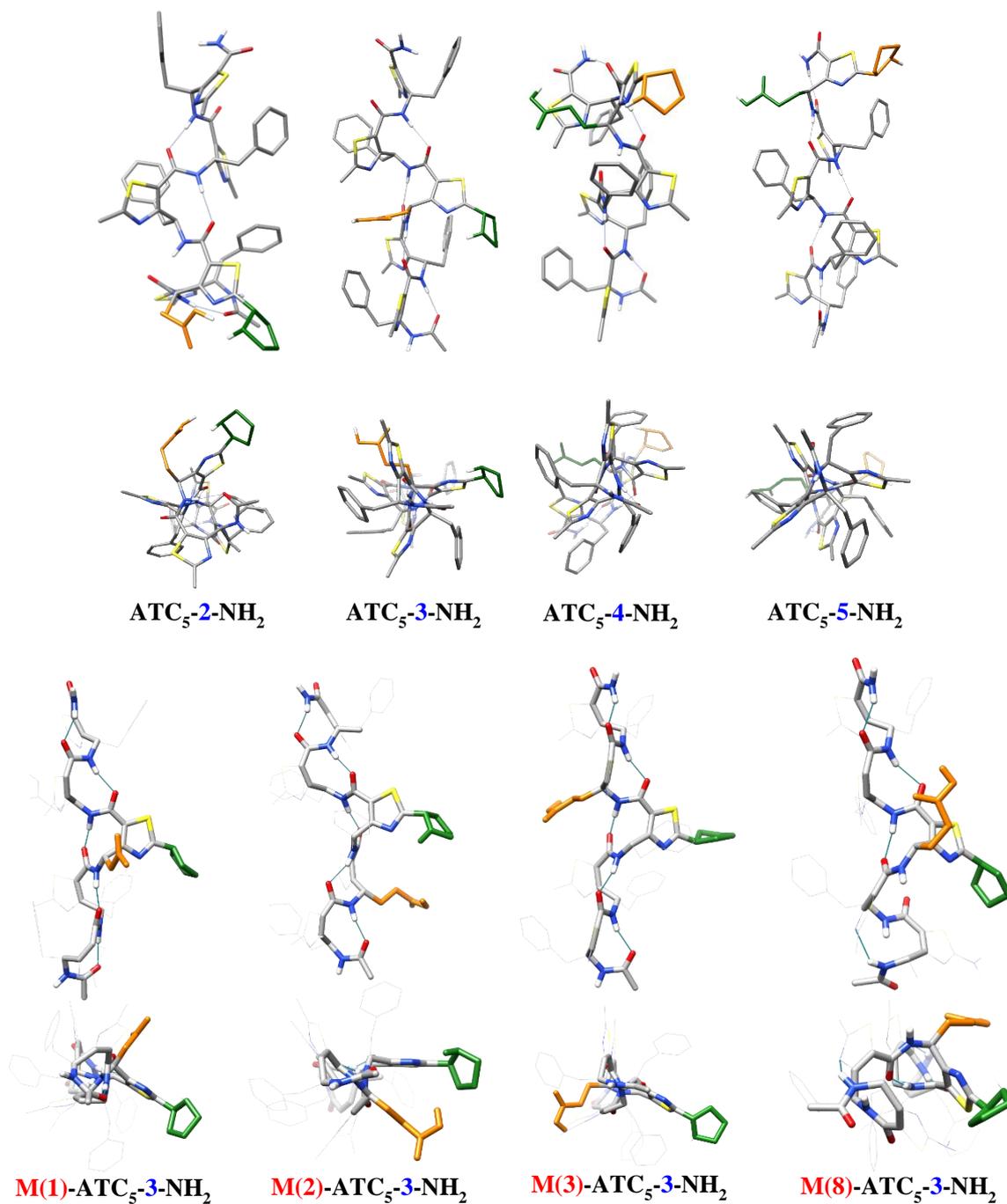
2 res1 H1 3 res2 HT3 4.3

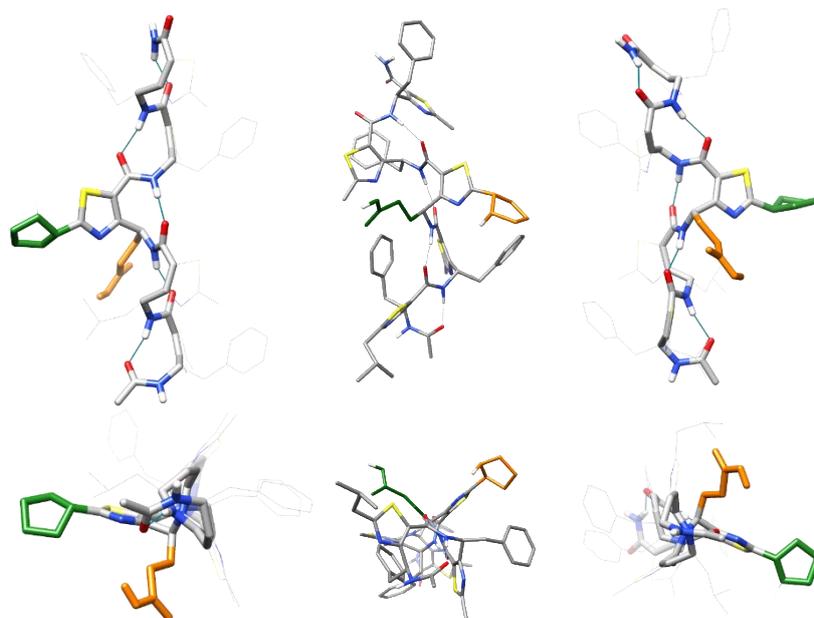
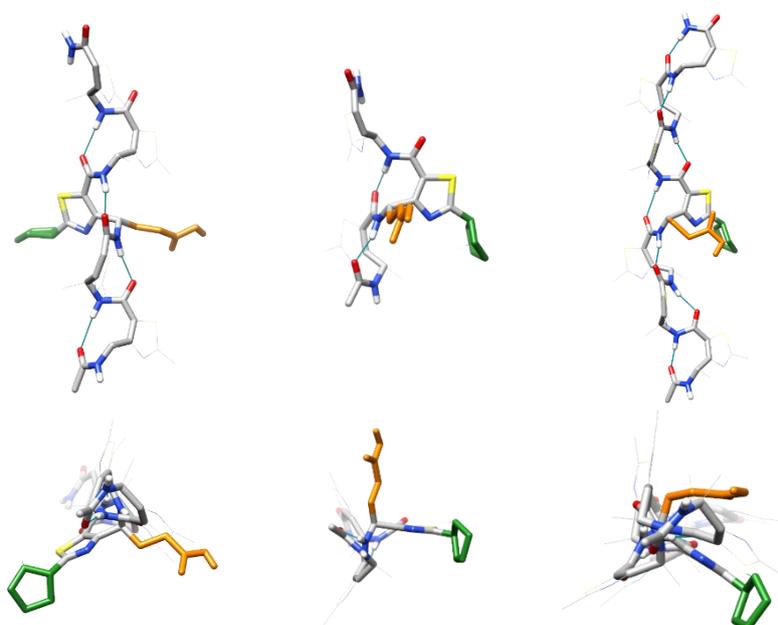
5 res4 H1 6 res5 HT1 4.3

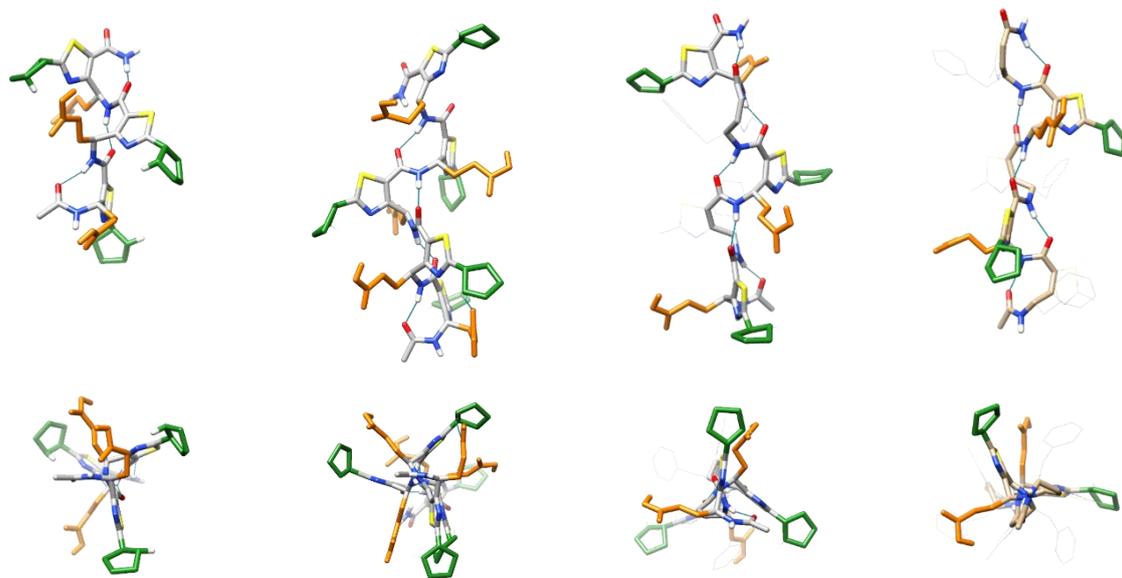
5 res4 H1 6 res5 HT2 4.3

5 res4 H1 6 res5 HT3 4.3

### 33. Calculated NMR solution structures



**M(5)-ATC<sub>5</sub>-3-NH<sub>2</sub>****M(9)-ATC<sub>5</sub>-3-NH<sub>2</sub>****M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>****M(13)-ATC<sub>5</sub>-3-NH<sub>2</sub>****M(16)-ATC<sub>3</sub>-2-NH<sub>2</sub>****M(17)-ATC<sub>7</sub>-4-NH<sub>2</sub>**



**M(18)-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub>** **M(19)-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub>** **M(20)-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub>** **M(21)-ATC<sub>5</sub>-2,4-NH<sub>2</sub>**

*Figure S61 –NMR energy structures optimized by DFT at the M06-2X/6-31G\* level of theory.*

## 34. Structure parameters

Table S102 – Average values (°) of the backbone torsion angles

Compound	Residue	$\varphi$	$\theta$	$\zeta$	$\psi$
<b>ATC<sub>5</sub>-2-NH<sub>2</sub></b>	Res. 1	-67.7 ± 49.5	114.2 ± 22.0	-3.6 ± 3.2	-18.6 ± 24.6
	Res. 2	-92.0 ± 65.8	98.3 ± 17.5	-4.6 ± 1.6	-29.7 ± 7.9
	Res. 3	-67.0 ± 18.7	122.9 ± 7.9	-4.5 ± 1.0	-45.2 ± 4.7
	Res. 4	-64.0 ± 6.2	122.4 ± 2.3	-4.5 ± 0.7	-42.5 ± 4.1
	Res. 5	-71.0 ± 14.6	124.8 ± 10.1	-0.7 ± 4.2	-2.8 ± 26.7
<b>ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	-43.1 ± 42.2	110.6 ± 18.8	-5.0 ± 1.6	-36.6 ± 10.9
	Res. 2	-78.1 ± 27.6	126.9 ± 8.7	-1.9 ± 3.6	-21.6 ± 34.0
	Res. 3	-75.8 ± 59.0	107.4 ± 18.1	-4.5 ± 2.3	-31.0 ± 12.4
	Res. 4	-74.3 ± 26.8	122.4 ± 9.1	-4.0 ± 0.6	-41.7 ± 5.4
	Res. 5	-77.6 ± 14.1	121.2 ± 7.6	-1.1 ± 4.3	-7.1 ± 25.4
<b>ATC<sub>5</sub>-3-NH<i>i</i>Pr</b>	Res. 1	-54.4 ± 33.9	117.3 ± 13.8	-5.1 ± 0.8	-41.6 ± 7.8
	Res. 2	-64.0 ± 7.9	123.3 ± 4.5	-4.1 ± 1.8	-33.7 ± 19.7
	Res. 3	-100.3 ± 34.6	112.4 ± 9.4	-3.3 ± 2.5	-30.1 ± 16.2
	Res. 4	-73.6 ± 27.4	121.8 ± 9.9	-4.1 ± 1.06	-43.3 ± 5.7
	Res. 5	-64.1 ± 3.8	120.0 ± 1.5	-3.4 ± 2.0	-29.2 ± 21.7
<b>ATC<sub>5</sub>-4-NH<sub>2</sub></b>	Res. 1	-59.8 ± 19.1	120.9 ± 2.7	-5.2 ± 0.6	-46.3 ± 2.5
	Res. 2	-64.7 ± 4.0	118.5 ± 3.2	-0.0 ± 0.7	-40.2 ± 4.7
	Res. 3	-86.5 ± 9.7	117.1 ± 3.8	3.9 ± 3.0	11.8 ± 30.2
	Res. 4	-70.3 ± 42.2	112.3 ± 16.1	-5.5 ± 1.2	-35.6 ± 8.8
	Res. 5	-55.0 ± 20.1	121.2 ± 12.7	-2.8 ± 3.8	-17.4 ± 23.1
<b>ATC<sub>5</sub>-5-NH<sub>2</sub></b>	Res. 1	-38.8 ± 36.7	113.6 ± 19.9	-5.3 ± 1.2	-50.9 ± 10.6
	Res. 2	-71.3 ± 7.9	127.1 ± 7.7	-1.2 ± 1.7	-53.9 ± 8.8
	Res. 3	-85.2 ± 19.1	115.1 ± 68.5	-1.5 ± 1.4	-40.4 ± 8.2

	Res. 4	$-111.8 \pm 21.7$	$131.7 \pm 16.2$	$2.4 \pm 2.9$	$2.7 \pm 25.7$
	Res. 5	$-101.1 \pm 58.7$	$89.4 \pm 53.7$	$-0.1 \pm 3.1$	$-6.7 \pm 31.1$
<b>M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	$-53.8 \pm 34.1$	$98.4 \pm 53.5$	$-4.5 \pm 1.9$	$-36.0 \pm 13.7$
	Res. 2	$-77.2 \pm 29.3$	$116.6 \pm 7.1$	$-3.3 \pm 2.1$	$-32.7 \pm 17.6$
	Res. 3	$-59.8 \pm 57.3$	$113.9 \pm 12.5$	$-4.3 \pm 1.6$	$-37.6 \pm 8.2$
	Res. 4	$-71.3 \pm 26.1$	$120.0 \pm 6.2$	$-4.3 \pm 0.9$	$-39.9 \pm 4.3$
	Res. 5	$-66.6 \pm 10.0$	$125.0 \pm 10.8$	$-1.2 \pm 3.8$	$-6.6 \pm 26.4$
<b>M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	$-52.9 \pm 18.7$	$117.3 \pm 10.7$	$-4.9 \pm 1.2$	$-34.4 \pm 12.5$
	Res. 2	$-64.6 \pm 22.7$	$119.3 \pm 5.7$	$-4.0 \pm 1.8$	$-35.3 \pm 13.3$
	Res. 3	$-12.8 \pm 30.0$	$98.3 \pm 17.1$	$-4.9 \pm 1.3$	$-32.4 \pm 12.1$
	Res. 4	$-55.8 \pm 4.2$	$121.8 \pm 2.4$	$-4.7 \pm 0.7$	$-43.0 \pm 2.6$
	Res. 5	$-66.0 \pm 11.5$	$121.1 \pm 8.5$	$-0.9 \pm 4.4$	$-43.0 \pm 99.4$
<b>M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	$-58.1 \pm 28.8$	$115.8 \pm 12.8$	$-4.5 \pm 1.1$	$-38.0 \pm 7.6$
	Res. 2	$-80.8 \pm 38.2$	$94.1 \pm 65.7$	$-3.4 \pm 2.1$	$-29.5 \pm 20.9$
	Res. 3	$-34.5 \pm 27.4$	$109.9 \pm 15.0$	$-5.3 \pm 1.2$	$-39.7 \pm 10.6$
	Res. 4	$-58.6 \pm 6.5$	$119.8 \pm 2.1$	$-4.5 \pm 0.9$	$-40.2 \pm 2.7$
	Res. 5	$-66.8 \pm 15.5$	$127.2 \pm 9.2$	$-1.0 \pm 4.0$	$-8.1 \pm 26.2$
<b>M(5)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	$-39.3 \pm 37.2$	$113.9 \pm 20.8$	$-4.3 \pm 1.8$	$-35.8 \pm 13.6$
	Res. 2	$-81.9 \pm 31.6$	$116.4 \pm 7.8$	$-3.3 \pm 1.6$	$-32.1 \pm 8.8$
	Res. 3	$-99.3 \pm 40.0$	$123.5 \pm 22.6$	$-3.1 \pm 2.0$	$-30.6 \pm 14.3$
	Res. 4	$-73.5 \pm 31.0$	$133.2 \pm 18.5$	$-3.7 \pm 1.9$	$-37.3 \pm 18.0$
	Res. 5	$-88.4 \pm 29.8$	$123.4 \pm 13.4$	$0.8 \pm 3.8$	$3.2 \pm 24.4$
<b>M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	$-26.1 \pm 45.5$	$105.3 \pm 28.9$	$-4.1 \pm 1.9$	$-31.9 \pm 13.7$
	Res. 2	$-111.0 \pm 24.7$	$105.2 \pm 65.3$	$3.5 \pm 3.8$	$-56.1 \pm 77.1$
	Res. 3	$-34.7 \pm 72.6$	$94.4 \pm 30.2$	$-4.1 \pm 2.3$	$-39.0 \pm 17.5$
	Res. 4	$-78.9 \pm 19.1$	$122.2 \pm 9.6$	$1.4 \pm 1.8$	$-29.3 \pm 11.1$
	Res. 5	$-111.0 \pm 28.3$	$109.4 \pm 19.2$	$-1.3 \pm 3.5$	$-7.6 \pm 19.2$

<b>M(9)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	$-57.0 \pm 2.9$	$122.4 \pm 2.2$	$-5.6 \pm 0.7$	$-48.3 \pm 4.6$
	Res. 2	$-83.9 \pm 14.7$	$122.3 \pm 6.7$	$-1.2 \pm 3.1$	$-7.6 \pm 32.6$
	Res. 3	$-83.5 \pm 51.3$	$107.4 \pm 17.8$	$-5.4 \pm 1.6$	$-42.3 \pm 11.0$
	Res. 4	$-69.5 \pm 4.2$	$129.7 \pm 13.8$	$-1.7 \pm 1.3$	$-37.4 \pm 5.7$
	Res. 5	$-119.6 \pm 16.8$	$138.6 \pm 10.5$	$4.0 \pm 2.2$	$11.8 \pm 13.6$
<b>M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	$-52.4 \pm 34.2$	$116.9 \pm 16.1$	$-5.2 \pm 0.9$	$-39.7 \pm 8.0$
	Res. 2	$-67.6 \pm 26.2$	$114.6 \pm 7.7$	$-4.4 \pm 1.6$	$-31.1 \pm 7.8$
	Res. 3	$-90.8 \pm 42.3$	$114.4 \pm 11.5$	$-5.0 \pm 1.8$	$-44.9 \pm 8.3$
	Res. 4	$-71.6 \pm 7.3$	$130.5 \pm 15.1$	$-1.5 \pm 1.5$	$-37.6 \pm 6.0$
	Res. 5	$-120.2 \pm 19.2$	$135.5 \pm 14.0$	$3.5 \pm 3.0$	$6.4 \pm 17.7$
<b>M(13)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	Res. 1	$-67.7 \pm 16.9$	$131.2 \pm 10.8$	$-1.8 \pm 2.0$	$-25.8 \pm 23.1$
	Res. 2	$-82.3 \pm 38.8$	$144.2 \pm 18.3$	$-1.2 \pm 2.7$	$-17.8 \pm 24.2$
	Res. 3	$-92.3 \pm 36.0$	$140.0 \pm 16.7$	$-3.3 \pm 2.6$	$-33.6 \pm 17.7$
	Res. 4	$-71.3 \pm 24.9$	$139.1 \pm 13.1$	$-1.5 \pm 2.0$	$-27.8 \pm 22.9$
	Res. 5	$-83.9 \pm 36.6$	$148.7 \pm 10.3$	$1.7 \pm 1.8$	$7.9 \pm 12.4$
<b>M(16)-ATC<sub>3</sub>-2-NH<sub>2</sub></b>	Res. 1	$-60.5 \pm 36.9$	$27.8 \pm 81.3$	$-2.6 \pm 2.2$	$-18.2 \pm 17.4$
	Res. 2	$-115.0 \pm 39.9$	$113.9 \pm 20.3$	$-3.6 \pm 2.4$	$-29.5 \pm 14.1$
	Res. 3	$-83.0 \pm 32.1$	$151.2 \pm 7.7$	$1.5 \pm 1.6$	$9.1 \pm 8.8$
<b>M(17)-ATC<sub>7</sub>-4-NH<sub>2</sub></b>	Res. 1	$-45.3 \pm 43.6$	$108.8 \pm 18.7$	$-4.8 \pm 1.0$	$-33.7 \pm 7.6$
	Res. 2	$-64.2 \pm 19.2$	$119.4 \pm 7.8$	$-4.1 \pm 1.1$	$-38.5 \pm 5.4$
	Res. 3	$-78.3 \pm 33.2$	$115.9 \pm 19.9$	$-3.3 \pm 2.4$	$-31.3 \pm 9.5$
	Res. 4	$-98.2 \pm 43.4$	$107.6 \pm 19.8$	$-2.9 \pm 2.4$	$-29.9 \pm 11.5$
	Res. 5	$-86.4 \pm 36.5$	$114.9 \pm 13.8$	$-3.6 \pm 1.7$	$-35.4 \pm 8.7$
	Res. 6	$-74.0 \pm 28.7$	$119.3 \pm 12.6$	$-3.4 \pm 1.4$	$-35.7 \pm 11.9$
	Res. 7	$-75.3 \pm 28.2$	$146.0 \pm 12.1$	$1.6 \pm 2.0$	$9.7 \pm 12.8$
<b>M(18)-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub></b>	Res. 1	$-104.5 \pm 51.5$	$112.0 \pm 28.0$	$-1.7 \pm 2.7$	$-14.0 \pm 19.1$
	Res. 2	$-63.2 \pm 35.9$	$111.5 \pm 17.6$	$-2.3 \pm 3.7$	$-15.7 \pm 27.6$

	Res. 3	$-48.6 \pm 59.6$	$100.1 \pm 19.8$	$-3.7 \pm 2.4$	$-19.0 \pm 13.4$
<b>M(19)-ATC<sub>5</sub>- 1,2,3,4,5-NH<sub>2</sub></b>	Res. 1	$-77.6 \pm 70.4$	$91.3 \pm 28.6$	$-1.0 \pm 3.4$	$-47.1 \pm 100.4$
	Res. 2	$-76.5 \pm 48.3$	$109.8 \pm 26.4$	$-0.4 \pm 4.3$	$-5.7 \pm 29.0$
	Res. 3	$-63.8 \pm 44.4$	$114.1 \pm 14.6$	$-4.4 \pm 2.4$	$-29.4 \pm 13.4$
	Res. 4	$-59.5 \pm 25.4$	$117.8 \pm 12.1$	$-3.1 \pm 3.6$	$-22.2 \pm 25.1$
	Res. 5	$-56.7 \pm 68.8$	$95.0 \pm 19.1$	$-1.3 \pm 3.3$	$-25.1 \pm 77.8$
<b>M(20)-ATC<sub>5</sub>-1,3,5- NH<sub>2</sub></b>	Res. 1	$-48.9 \pm 65.3$	$102.0 \pm 30.2$	$-1.7 \pm 4.0$	$-13.7 \pm 25.6$
	Res. 2	$-66.8 \pm 38.4$	$97.7 \pm 57.6$	$-2.0 \pm 3.3$	$-14.1 \pm 32.1$
	Res. 3	$-64.7 \pm 55.3$	$104.3 \pm 18.6$	$-4.6 \pm 2.1$	$-30.3 \pm 11.0$
	Res. 4	$-49.3 \pm 24.1$	$119.2 \pm 12.5$	$-4.8 \pm 2.1$	$-34.0 \pm 16.3$
	Res. 5	$-58.6 \pm 38.5$	$116.2 \pm 18.9$	$0.1 \pm 4.4$	$-58.4 \pm 96.0$
<b>M(21)-ATC<sub>5</sub>-2,4-NH<sub>2</sub></b>	Res. 1	$-37.4 \pm 39.3$	$112.9 \pm 25.6$	$-3.1 \pm 3.3$	$-18.0 \pm 27.6$
	Res. 2	$-92.8 \pm 54.1$	$102.6 \pm 17.5$	$-3.1 \pm 2.9$	$-27.4 \pm 10.7$
	Res. 3	$-80.5 \pm 29.0$	$116.7 \pm 15.0$	$-2.7 \pm 2.5$	$-29.2 \pm 23.8$
	Res. 4	$-62.8 \pm 64.1$	$103.5 \pm 24.7$	$-4.1 \pm 1.7$	$-28.6 \pm 9.4$
	Res. 5	$-70.2 \pm 37.6$	$124.6 \pm 19.6$	$-0.3 \pm 4.6$	$-3.9 \pm 122.2$

## VII. CIRCULAR DICHROISM

Circular dichroism (CD) experiments were carried out using a Jasco J815 spectropolarimeter. The spectra were recorded at 50  $\mu$ M compounds dissolved in *i*PrOH using a 1 mm path length CD cuvette at 4°C, over a wavelength range of 190-300 nm. Continuous scanning mode was used, with a response of 1.0 s with 0.5 nm steps and a bandwidth of 1 nm. The signal to noise ratio was improved by acquiring each spectrum over an average of three scans. Baseline was corrected by subtracting the background from the sample spectrum.

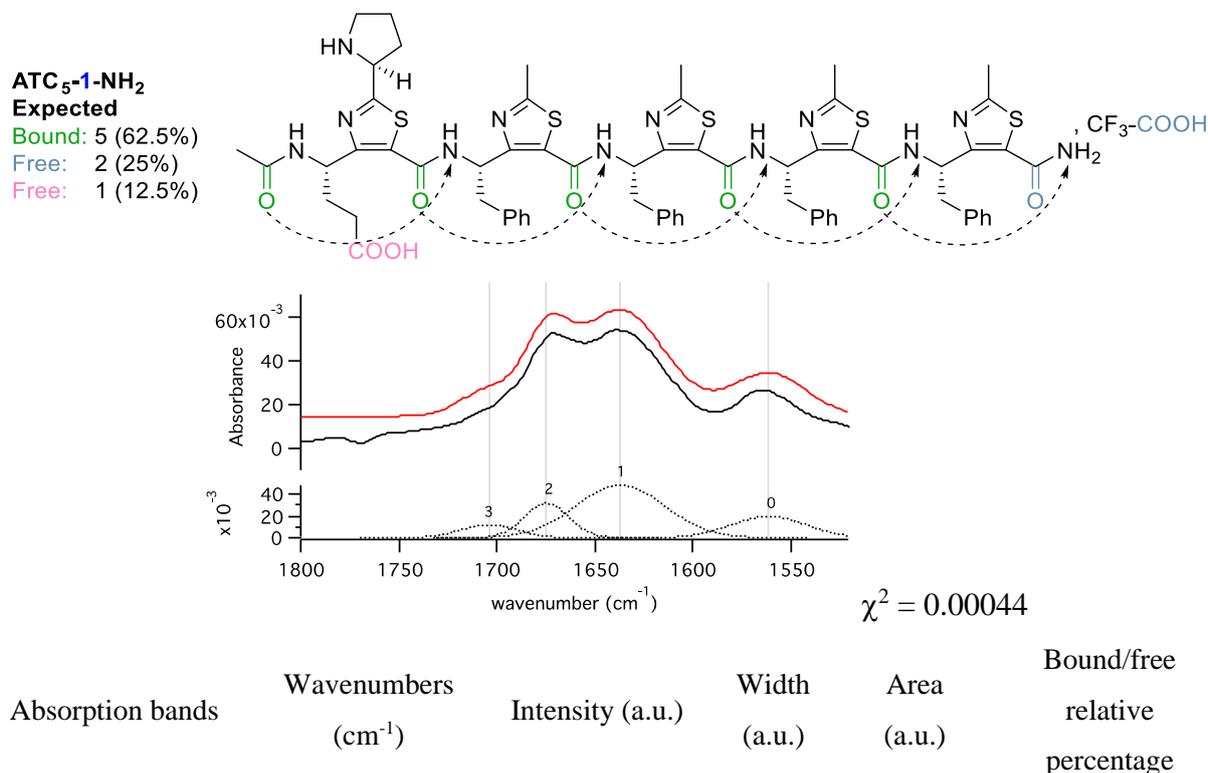
Name of sequence	mre max	$\lambda_{\text{max}}$ (nm)	mre min	$\lambda_{\text{min}}$ (nm)
<b>ATC<sub>5</sub>-1-NH<sub>2</sub></b>	86209.2	265.5	-81998.4	233
<b>ATC<sub>5</sub>-2-NH<sub>2</sub></b>	72830	266	-64264.4	231.5
<b>ATC<sub>5</sub>-3-NH<sub>2</sub></b>	74401.6	266	-63156	233
<b>ATC<sub>5</sub>-4-NH<sub>2</sub></b>	78432.4	266	-73022.4	232.5
<b>ATC<sub>5</sub>-5-NH<sub>2</sub></b>	83087.2	266	-74018	233
<b>M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	74731.2	265.5	-64870.4	231.5
<b>M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	79080.4	265.5	-75256	232
<b>M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	73572.8	266.5	-67835.6	232.5
<b>M(4)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	77548.8	266.5	-66120	232
<b>M(5)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	80620.8	265	-67151.6	232.5
<b>M(7)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	76201.2	267	-62222	233
<b>M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	57302	266.5	-58758.8	231.5
<b>M(9)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	75402.1	265.5	-64980.6	233.5
<b>M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	86895.6	265.5	-72683.2	232
<b>M(11)-ATC<sub>5</sub>-1-NH<sub>2</sub></b>	98783.6	262.5	-73729.6	231
<b>M(12)-ATC<sub>5</sub>-2-NH<sub>2</sub></b>	110201.6	263	-80974.4	230
<b>M(13)-ATC<sub>5</sub>-3-NH<sub>2</sub></b>	80886.4	263	-59871.2	231
<b>M(14)-ATC<sub>5</sub>-4-NH<sub>2</sub></b>	88799.2	264	-64936	230.5
<b>M(15)-ATC<sub>5</sub>-5-NH<sub>2</sub></b>	98618	263	-66733.6	231
<b>M(16)-ATC<sub>3</sub>-2-NH<sub>2</sub></b>	64740.9	262.5	-46391.5	231.5
<b>M(17)-ATC<sub>7</sub>-4-NH<sub>2</sub></b>	102142.6	263	-73600.9	230
<b>M(18)-ATC<sub>3</sub>-1,2,3-NH<sub>2</sub></b>	37286.3	265	-37542	236
<b>M(19)-ATC<sub>5</sub>-1,2,3,4,5-NH<sub>2</sub></b>	69243.8	265.5	-74004.8	234.5
<b>M(20)-ATC<sub>5</sub>-1,3,5-NH<sub>2</sub></b>	25357.6	269.5	-21313.4	234.5



## VIII. FTIR ANALYSIS

Middle-infrared experiments (400-5000  $\text{cm}^{-1}$ ) were recorded in the transmission mode. The measurements were carried out on a Bruker Tensor 27 spectrometer equipped with a deuterated (*L*)-alanine doped triglycine sulphate (DLATGS) pyroelectric detector, a Globar source, and potassium bromide (KBr) beam splitter. The spectral resolution was 4  $\text{cm}^{-1}$ , and 128 scans were co-added for each spectrum. The compounds were dissolved at 2 mM concentration in *i*PrOH transferred in a liquid cell equipped with  $\text{CaF}_2$  windows separated by a teflon spacer (thickness: 50  $\mu\text{M}$ ). A spacer thickness of 50  $\mu\text{M}$  provide exploitable signal in term of signal-to-noise ratio  $S/N$  ( $S/N$  (Peak to peak) =  $4 \times 10^{-5}$ ,  $S/N$  (quadratic mean (RMS)) =  $1 \times 10^{-5}$  between 2000 and 2100  $\text{cm}^{-1}$ ). Below this value, significant compensation problems between solvent spectral bands used as reference and sample spectrum appear. The FT-IR spectra were not smoothed. Linear baseline subtraction and deconvolution were performed using IGOR Pro 7.0 software (WaveMetrics). The deconvolution using a sum of Gaussian functions (in dash lines) and the overall fit (in red) are also shown. Resulted fits are arbitrary vertically shifted for clarity.

Table S103 - Band positions, band intensities, band widths and band areas in the region of 1400-1800  $\text{cm}^{-1}$  (Bound  $\nu(\text{CO})$  are in green while free  $\nu(\text{CO})$  are in blue)



<b>Contribution 1</b> (bound CO)	1632	0.0340	29.9	1.804	64 %
<b>Contribution 2</b> (free CO)	1674	0.0240	16.0	0.677	24 %
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1700	0.0091	20.0	0.324	12 %

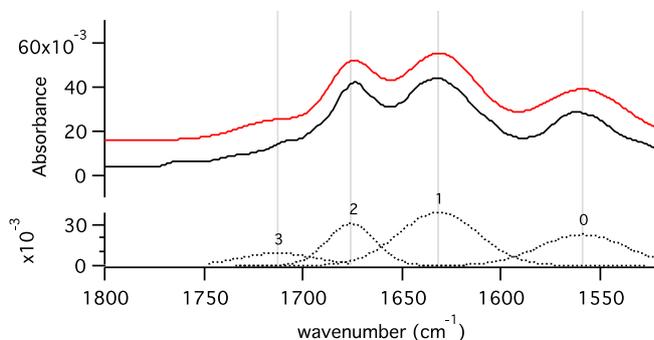
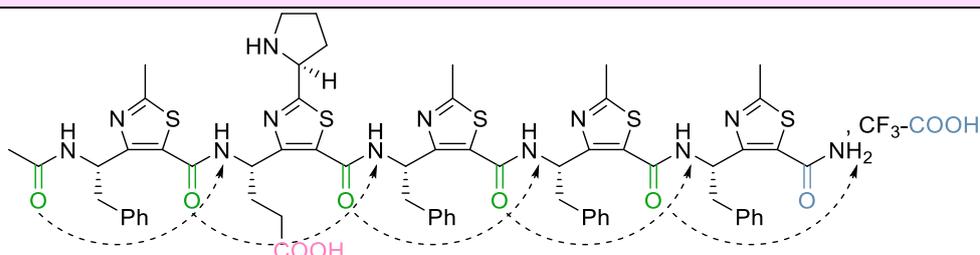
ATC<sub>5</sub>-2-NH<sub>2</sub>

Expected

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



$$\chi^2 = 0.0003$$

Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
------------------	------------------------------------	------------------	-----------------	----------------	--------------------------------------

<b>Contribution 1</b> (bound CO)	1632	0.0379	28.8	1.940	62 %
<b>Contribution 2</b> (free CO)	1676	0.0306	16.0	0.869	27 %
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1708	0.0098	20.0	0.348	11 %

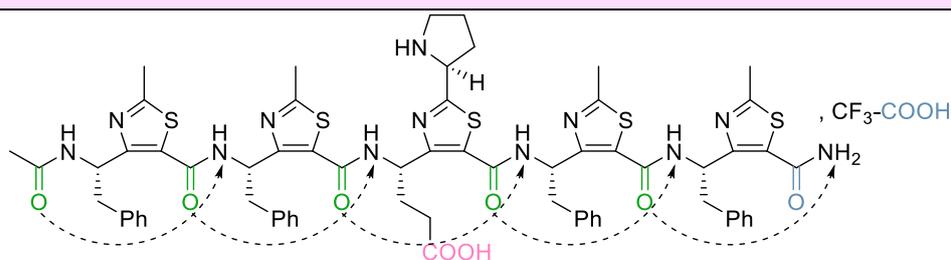
ATC<sub>5</sub>-3-NH<sub>2</sub>

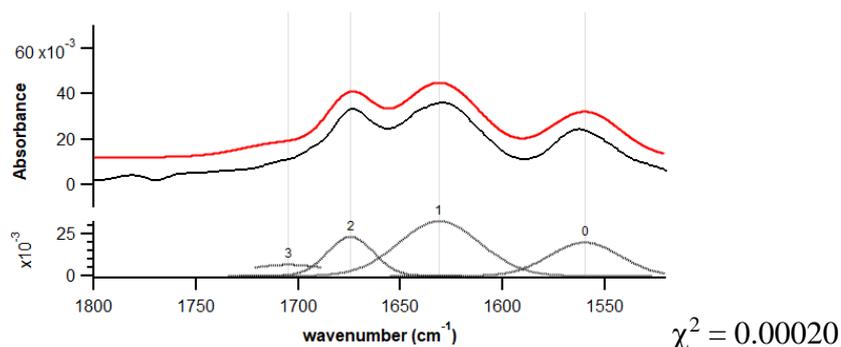
Expected

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)





Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1631	0.0322	27.7	1.580	62 %
<b>Contribution 2</b> (free CO)	1674	0.0251	16.0	0.712	28 %
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1705	0.0068	20.0	0.240	10 %

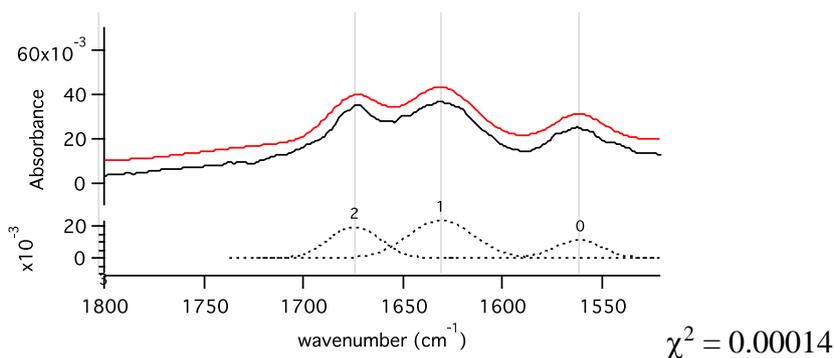
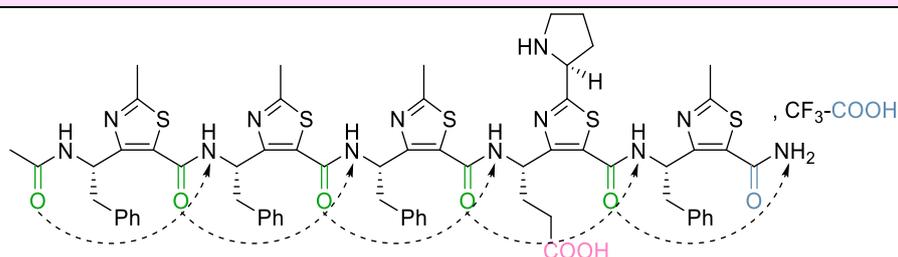
ATC<sub>5</sub>-4-NH<sub>2</sub>

Expected

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1632	0.0302	28.2	1.512	62 %
<b>Contribution 2</b>	1675	0.0238	16.0	0.676	27 %

(free CO)					
ATC5 – TFA salt					
<b>Contribution 3</b>					
(free CO) COOH	1705	0.0076	20.0	0.271	11 %

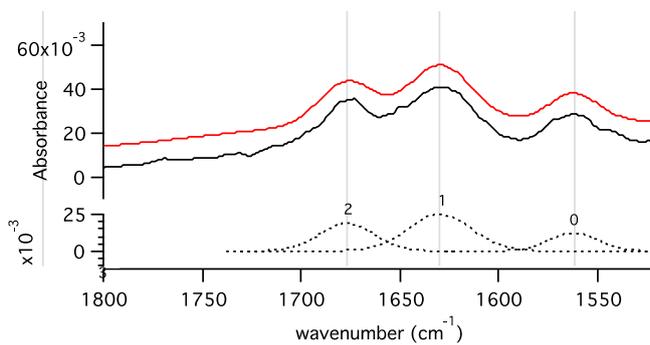
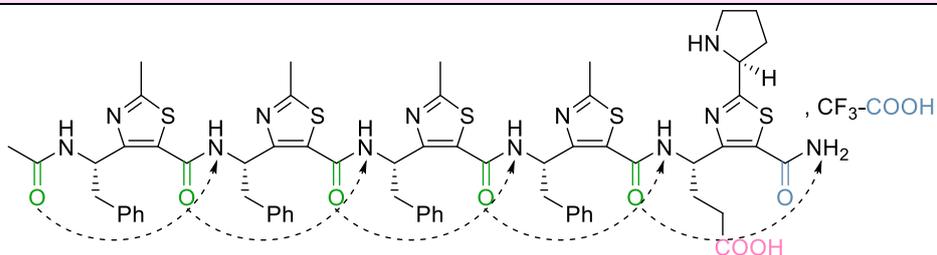
**ATC<sub>5</sub>-5-NH<sub>2</sub>**

**Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



$$\chi^2 = 0.00016$$

Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b>					
(bound CO)	1631	0.0331	27.6	1.620	63 %
<b>Contribution 2</b>					
(free CO)	1677	0.0249	16.0	0.705	27 %
ATC5 – TFA salt					
<b>Contribution 3</b>					
(free CO) COOH	1705	0.0081	18.0	0.258	11 %

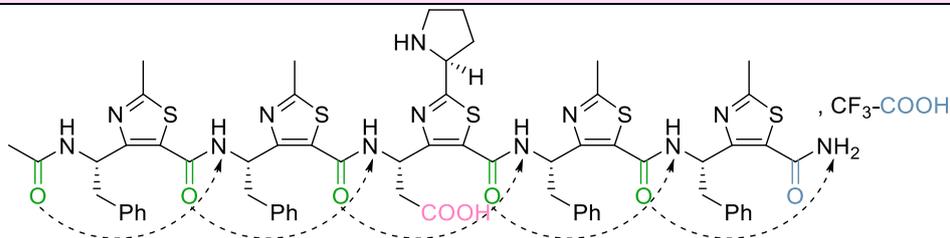
**M(1)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

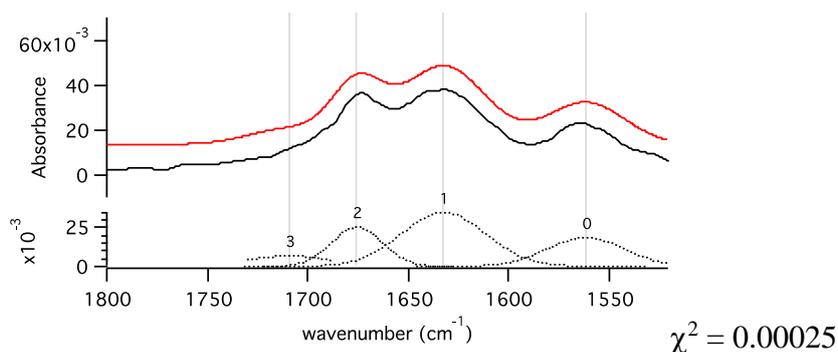
**Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)





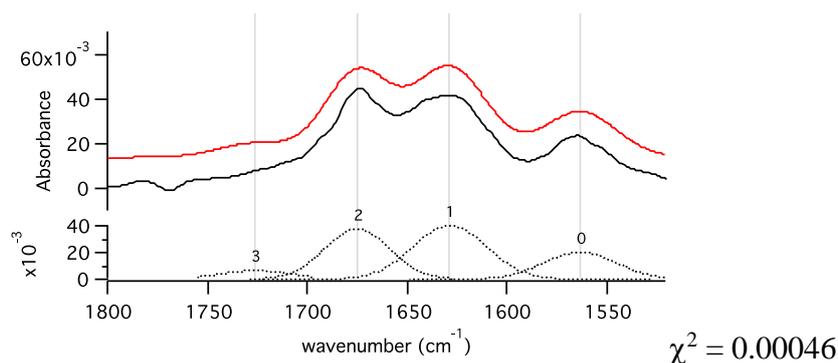
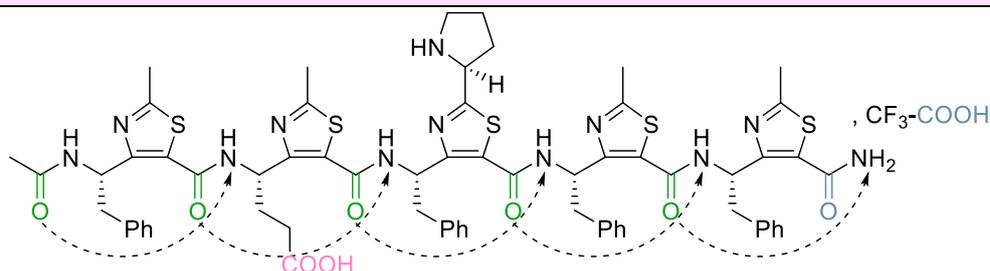
Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1632	0.0340	29.9	1.804	64 %
<b>Contribution 2</b> (free CO)	1674	0.0240	16.0	0.677	24 %
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1700	0.0091	20.0	0.324	12 %

**M(2)-ATC<sub>5</sub>-3-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1629	0.0400	26.0	1.846	50%
<b>Contribution 2</b>	1675	0.0377	23.0	1.535	42%

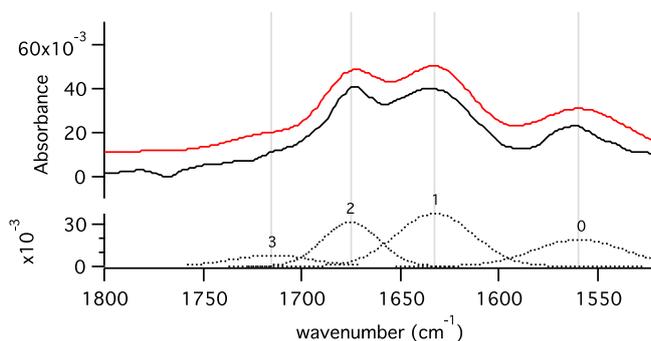
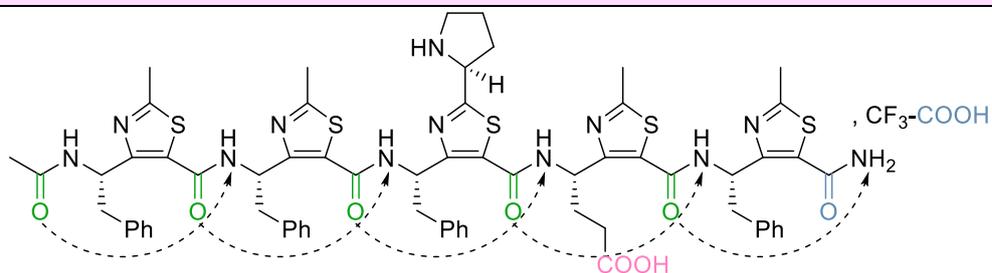
(free CO)					
ATC5 – TFA salt					
<b>Contribution 3</b>					
(free CO) COOH	1726	0.0063	25.4	0.283	8%

**M(3)-ATC<sub>5</sub>-3-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



$\chi^2 = 0.00038$

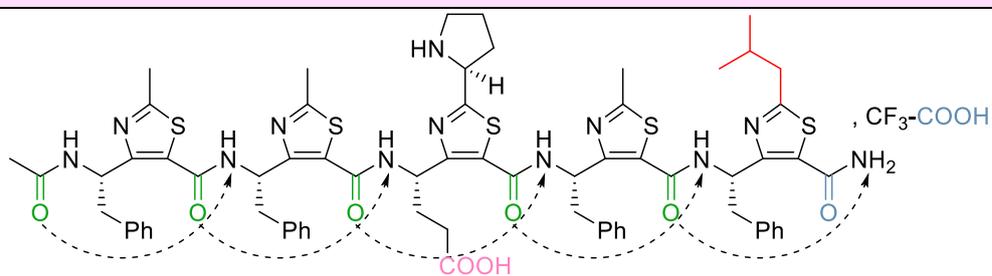
Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b>					
(bound CO)	1632	0.0378	27.8	1.860	55%
<b>Contribution 2</b>					
(free CO)	1675	0.0315	20.2	1.129	33%
ATC5 – TFA salt					
<b>Contribution 3</b>					
(free CO) COOH	1715	0.0079	28.9	0.407	12%

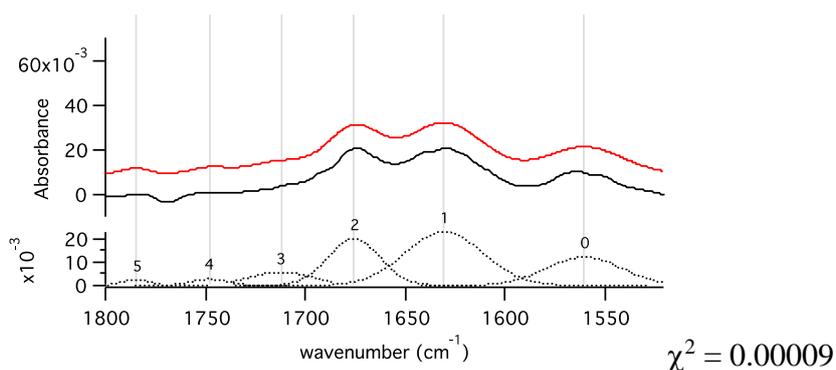
**M(4)-ATC<sub>5</sub>-3-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)

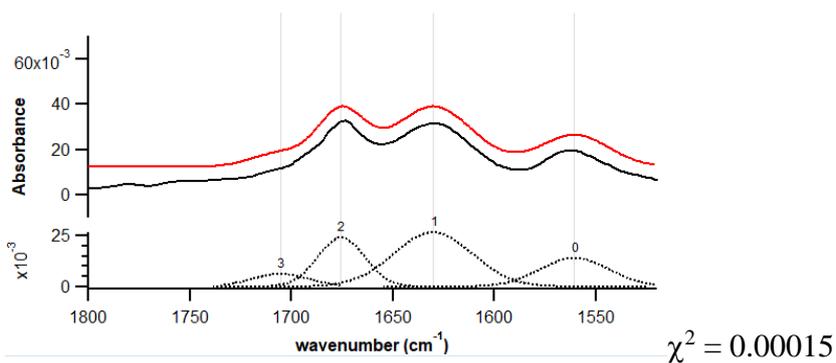
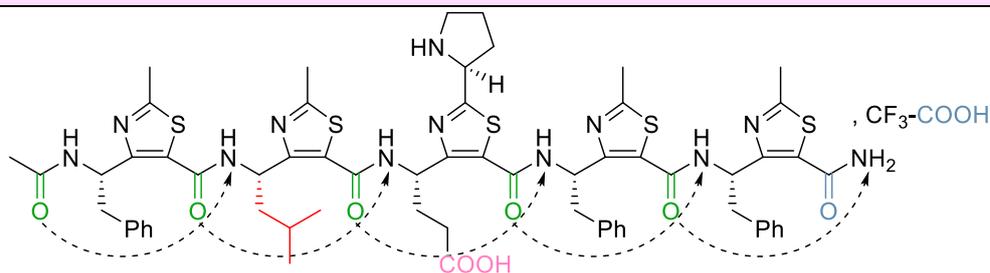




Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1631	0.0227	27.7	1.113	56%
<b>Contribution 2</b> (free CO)	1676	0.0199	18.3	0.647	33%
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1712	0.0056	22.3	0.223	11%

**M(5)-ATC<sub>5</sub>-3-NH<sub>2</sub>**  
**Expected**

Bound: 5 (62.5%)  
Free: 2 (25%)  
Free: 1 (12.5%)



Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1630	0.0266	27.7	1.307	59%
<b>Contribution 2</b>	1676	0.0241	16.5	0.707	32%

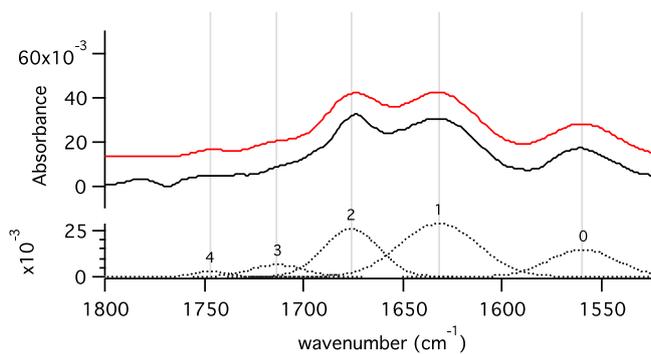
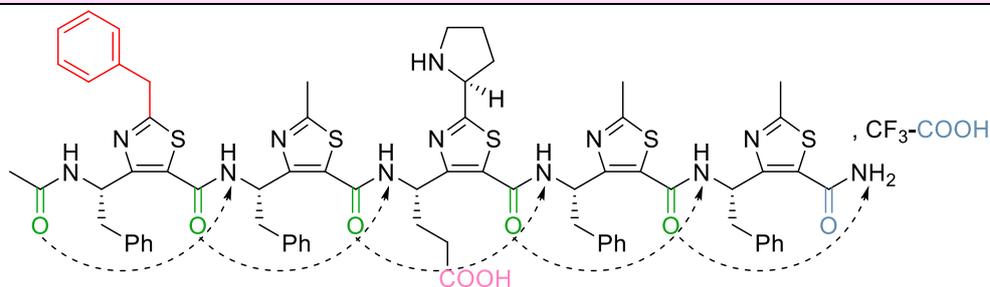
(free CO)					
ATC5 – TFA salt					
<b>Contribution 3</b>					
(free CO) COOH	1705	0.0061	20	0.218	9%

**M(7)-ATC<sub>5</sub>-3-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



$\chi^2 = 0.00017$

Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
------------------	---------------------------------	------------------	--------------	-------------	--------------------------------

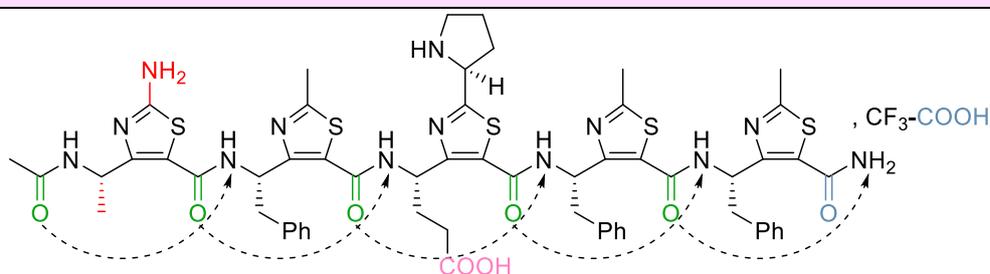
<b>Contribution 1</b>					
(bound CO)	1632	0.0288	27.7	1.409	56%
<b>Contribution 2</b>					
(free CO)	1676	0.0260	19.5	0.898	36%
ATC5 – TFA salt					
<b>Contribution 3</b>					
(free CO) COOH	1713	0.0064	18.7	0.214	8%

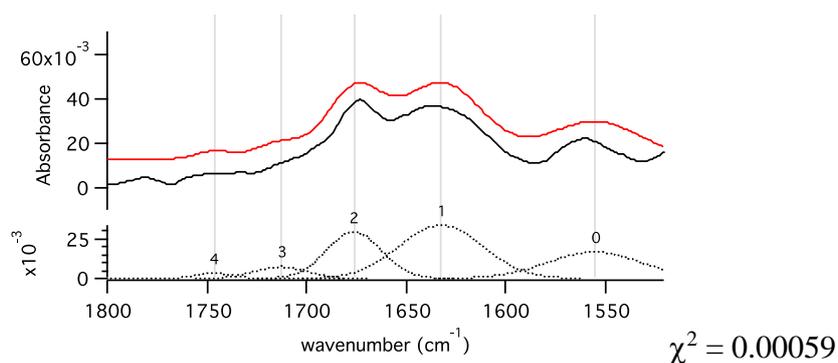
**M(8)-ATC<sub>5</sub>-3-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

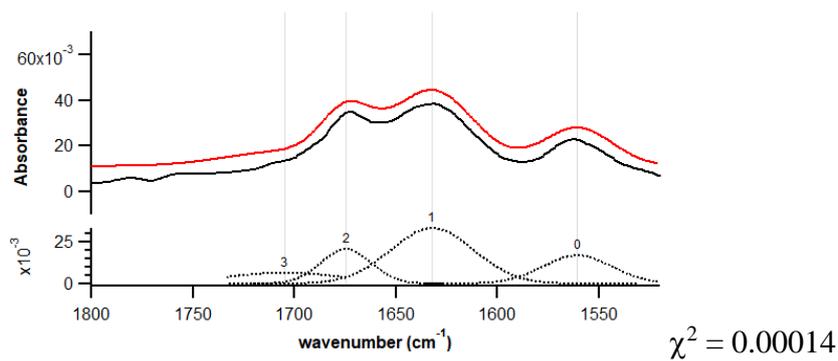
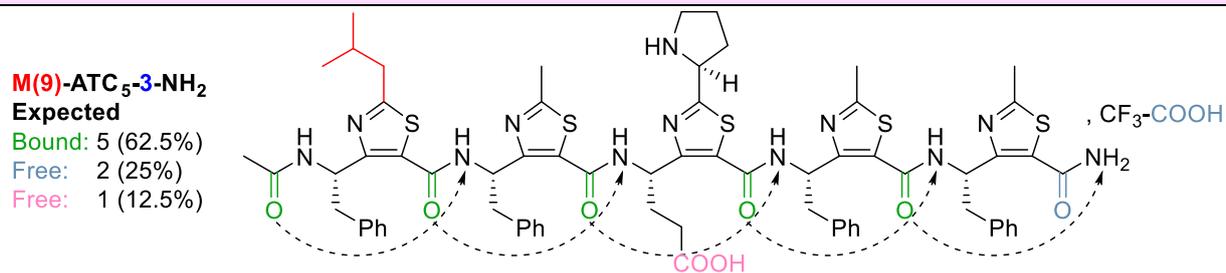
Free: 2 (25%)

Free: 1 (12.5%)





Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1633	0.0339	29.6	1.773	59%
<b>Contribution 2</b> (free CO)	1676	0.0295	19.4	1.015	34%
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1713	0.0073	18.1	0.236	7%



Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1632	0.0330	29.1	1.700	61%
<b>Contribution 2</b>	1675	0.0205	17.7	0.644	23%

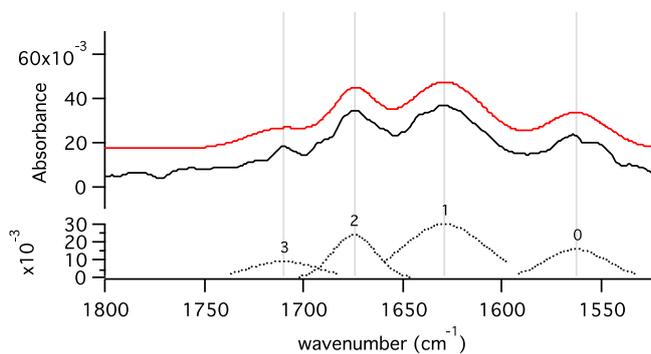
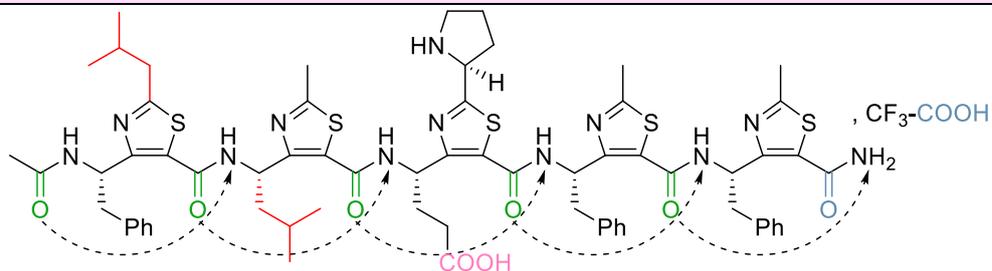
(free CO)					
ATC5 – TFA salt					
<b>Contribution 3</b>					
(free CO) COOH	1705	0.0063	40.7	0.458	16%

**M(10)-ATC<sub>5</sub>-3-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



$\chi^2 = 0.00031$

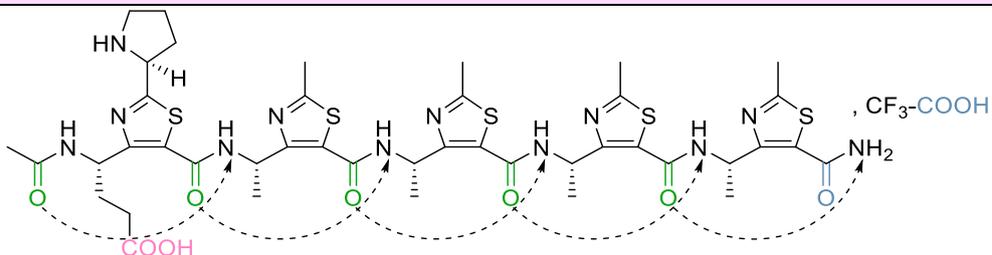
Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b>					
(bound CO)	1629	0.0299	28.6	1.519	59%
<b>Contribution 2</b>					
(free CO)	1674	0.0241	15.3	0.653	25%
ATC5 – TFA salt					
<b>Contribution 3</b>					
(free CO) COOH	1710	0.0093	24.5	0.406	16%

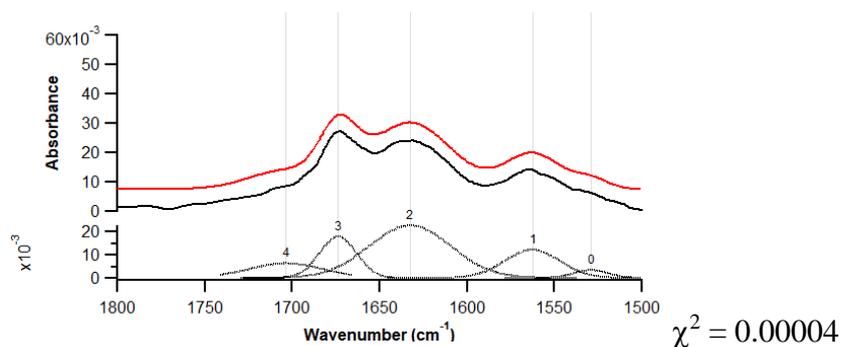
**M(11)-ATC<sub>5</sub>-1-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

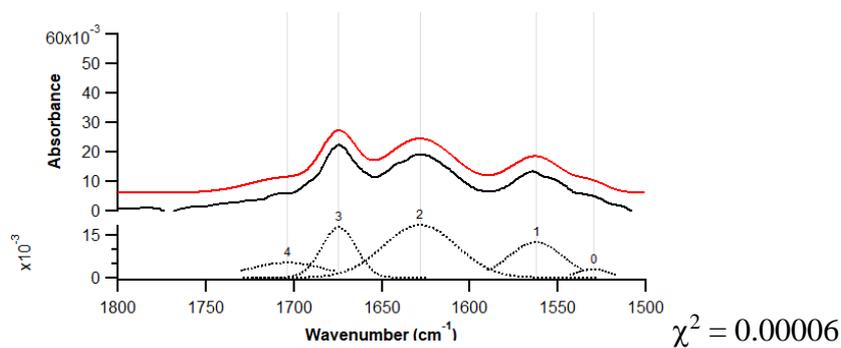
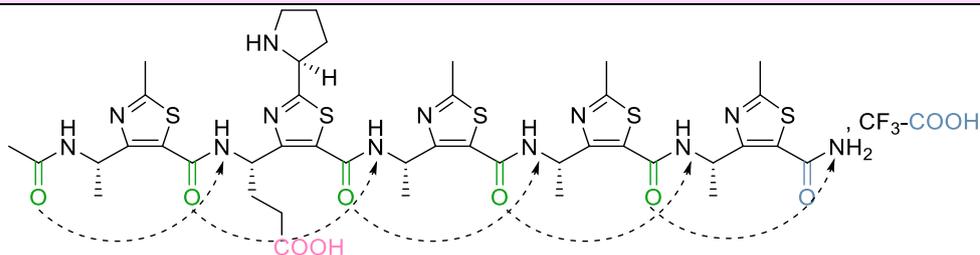
Free: 1 (12.5%)





Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1633	0.0227	33.4	1.341	62%
<b>Contribution 2</b> (free CO)	1674	0.0180	15.0	0.478	22%
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1704	0.0063	30.0	0.337	16%

**M(12)-ATC<sub>5</sub>-2-NH<sub>2</sub>**  
**Expected**  
 Bound: 5 (62.5%)  
 Free: 2 (25%)  
 Free: 1 (12.5%)



Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1628	0.0184	29.6	0.965	58%
<b>Contribution 2</b> (free CO)	1675	0.0176	14.0	0.436	26%

ATC5 – TFA salt

**Contribution 3**

(free CO) COOH

1704

0.0052

30.0

0.277

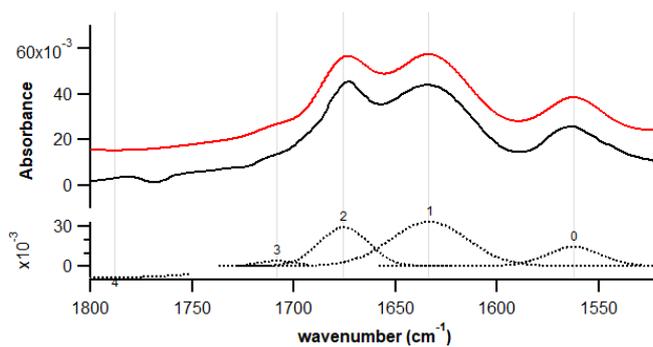
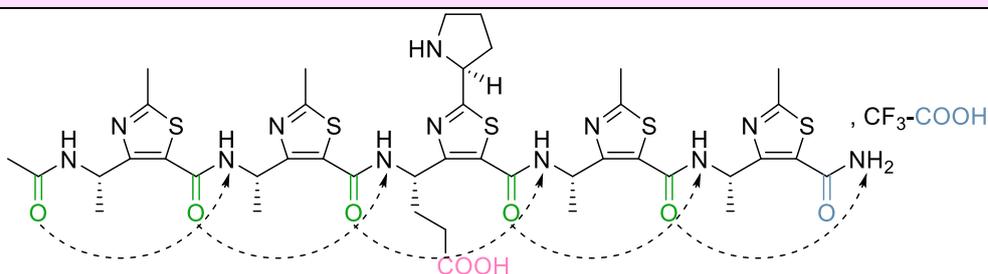
16%

**M(13)-ATC<sub>5</sub>-3-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



$\chi^2 = 0.00016$

Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
------------------	------------------------------------	------------------	-----------------	----------------	--------------------------------------

**Contribution 1**

(bound CO)

1634

0.0333

27.1

1.600

62%

**Contribution 2**

(free CO)

1676

0.0295

17.2

0.900

35%

ATC5 – TFA salt

**Contribution 3**

(free CO) COOH

1708

0.0040

13.4

0.094

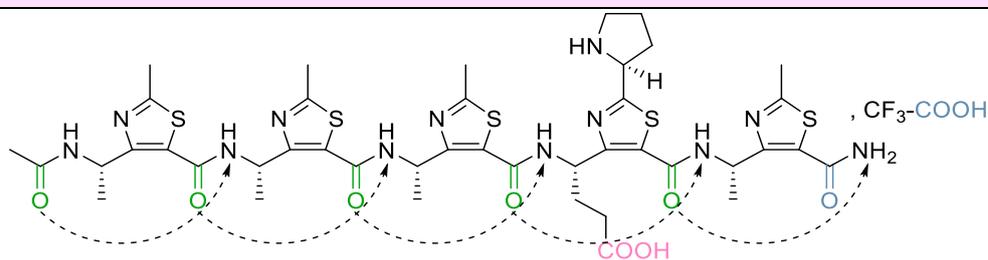
3%

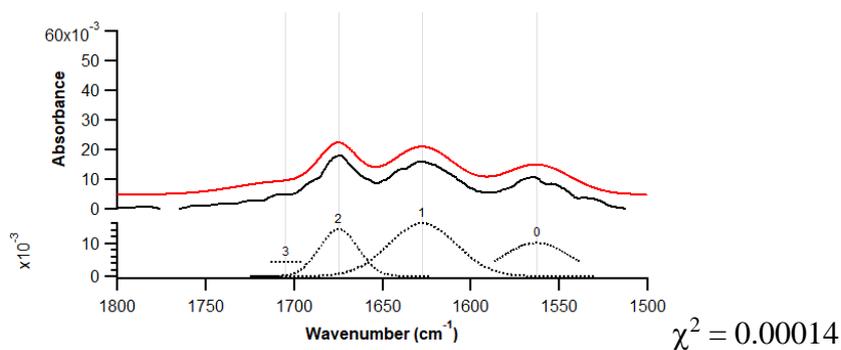
**M(14)-ATC<sub>5</sub>-4-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)





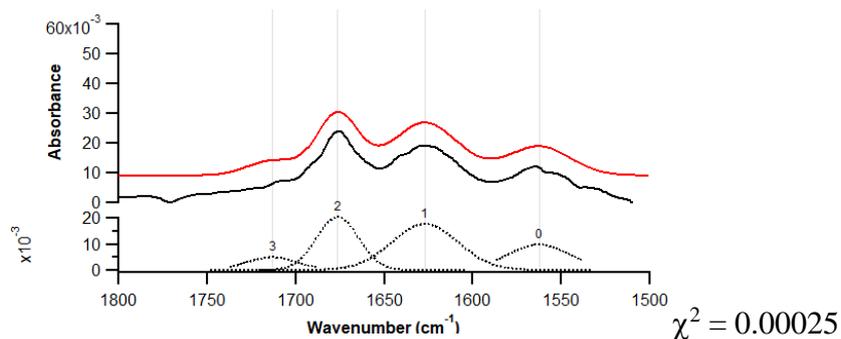
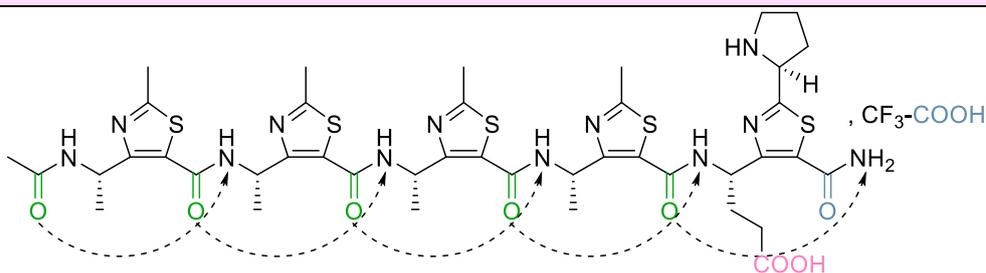
Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1628	0.0161	27.2	0.775	53%
<b>Contribution 2</b> (free CO)	1675	0.0144	25.7	0.393	27%
ATC5 – TFA salt					
<b>Contribution 3</b> (free CO) COOH	1705	0.0044	38.9	0.306	20%

**M(15)-ATC<sub>5</sub>-5-NH<sub>2</sub>****Expected**

Bound: 5 (62.5%)

Free: 2 (25%)

Free: 1 (12.5%)



Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1627	0.0178	25.3	0.847	52%
<b>Contribution 2</b> (free CO)	1676	0.0206	26.9	0.617	38%

ATC5 – TFA salt

**Contribution 3**

(free CO) COOH

1713

0.0050

20.0

0.177

10%

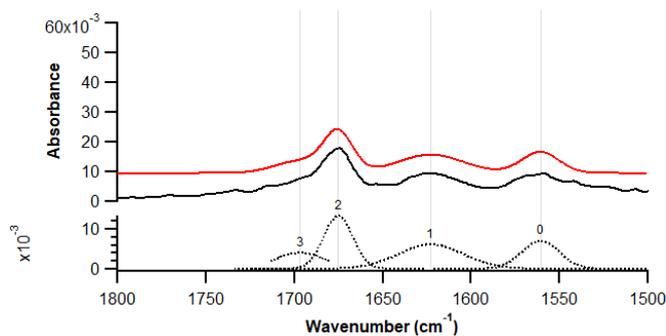
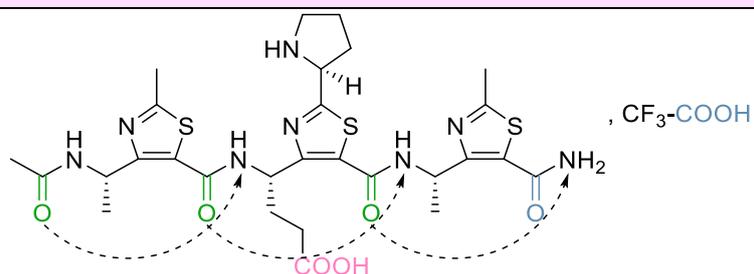
**M(16)-ATC<sub>3</sub>-2-NH<sub>2</sub>**

Expected

Bound: 3 (50%)

Free: 2 (33%)

Free: 1 (17%)



$\chi^2 = 0.00027$

Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
------------------	------------------------------------	------------------	-----------------	----------------	--------------------------------------

**Contribution 1**

(bound CO)

1623

0.0061

26.5

0.288

41%

**Contribution 2**

(free CO)

1675

0.0134

11.8

0.279

39%

ATC5 – TFA salt

**Contribution 3**

(free CO) COOH

1697

0.0040

20.0

0.144

20%

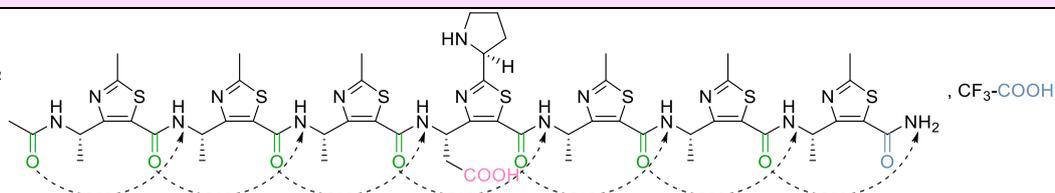
**M(17)-ATC<sub>5</sub>-3-NH<sub>2</sub>**

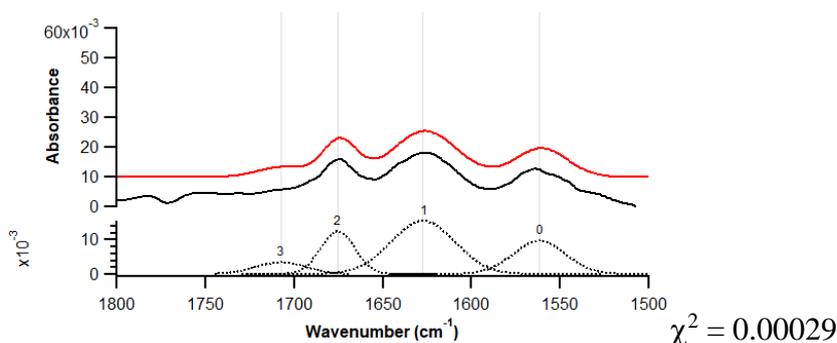
Expected

Bound: 7 (70%)

Free: 2 (20%)

Free: 1 (10%)





Absorption bands	Wavenumbers (cm <sup>-1</sup> )	Intensity (a.u.)	Width (a.u.)	Area (a.u.)	Bound/free relative percentage
<b>Contribution 1</b> (bound CO)	1627	0.0155	21.1	0.709	62%
<b>Contribution 2</b> (free CO) ATC5 – TFA salt	1675	0.0124	25.8	0.305	27%
<b>Contribution 3</b> (free CO) COOH	1708	0.0034	20.0	0.122	11%

## IX. XRD STRUCTURES

X-ray data for compound **NFmoc-{CH<sub>3</sub>;Bn}-ATC-OBn** and (**21b**) were collected at 100K with an Oxford Diffraction Xcalibur 2 diffractometer equipped with copper high-intensity/low-power micro-source X-ray tube. Crystals for compound (**21b**) exhibited weak diffracting power. Data were processed using CrysAlis RED (Oxford Diffraction, 2003). Xray structures were solved with SIR2004 (Burla M.C., Caliendo R., Camalli M., Carrozzini B., Cascarano G.L., De Caro L., Giacovazzo C., Polidori G. and Spagna R. (2005) J. Appl. Cryst. 2005, 38:381-388). The crystallographic refinements were conducted using SHELXL-97 (Sheldrick G.M. Acta Cryst. 2008, A64: 112- 122). In compound (**NFmoc-{CH<sub>3</sub>;Bn}-ATC-OBn**), carbon atoms of the terminal fluorenyl group showed high atomic displacement parameters and the analysis of their displacement ellipsoids revealed some rotational structural disorder in this group. CCDC 915084 (**NFmoc-{CH<sub>3</sub>;Bn}-ATC-OBn**) CCDC 922304 ((**21b**)) contain the supplementary crystallographic data. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre.

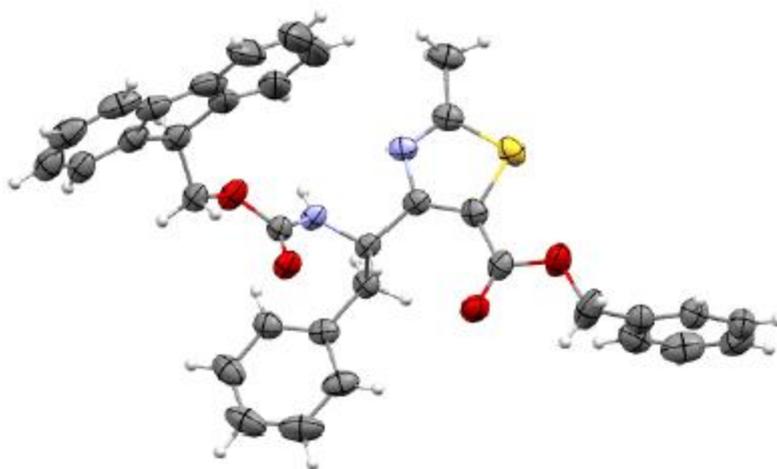


Figure S62: X-ray crystal structure of *NFmoc*-{*CH*<sub>3</sub>;*Bn*}-*ATC-OBn*

The asymmetric unit of **(21b)** crystals consisted of one ATC terramer and one diisopropyl ether molecule. **(21b)** molecules self-assembled along the crystallographic axis to form infinite chains of helices. Infinite chains of diisopropyl ether molecules mediated helix chains. Two helix chains and one diisopropyl ether chain are shown. Hydrogen atoms attached to carbon atoms are not indicated for clarity. Hydrogen, oxygen, nitrogen and sulphur atoms are colored following atom type. Different colors have been used for carbon atoms to distinguish independent molecules. Hydrogen bonds are represented as black dashed bonds.

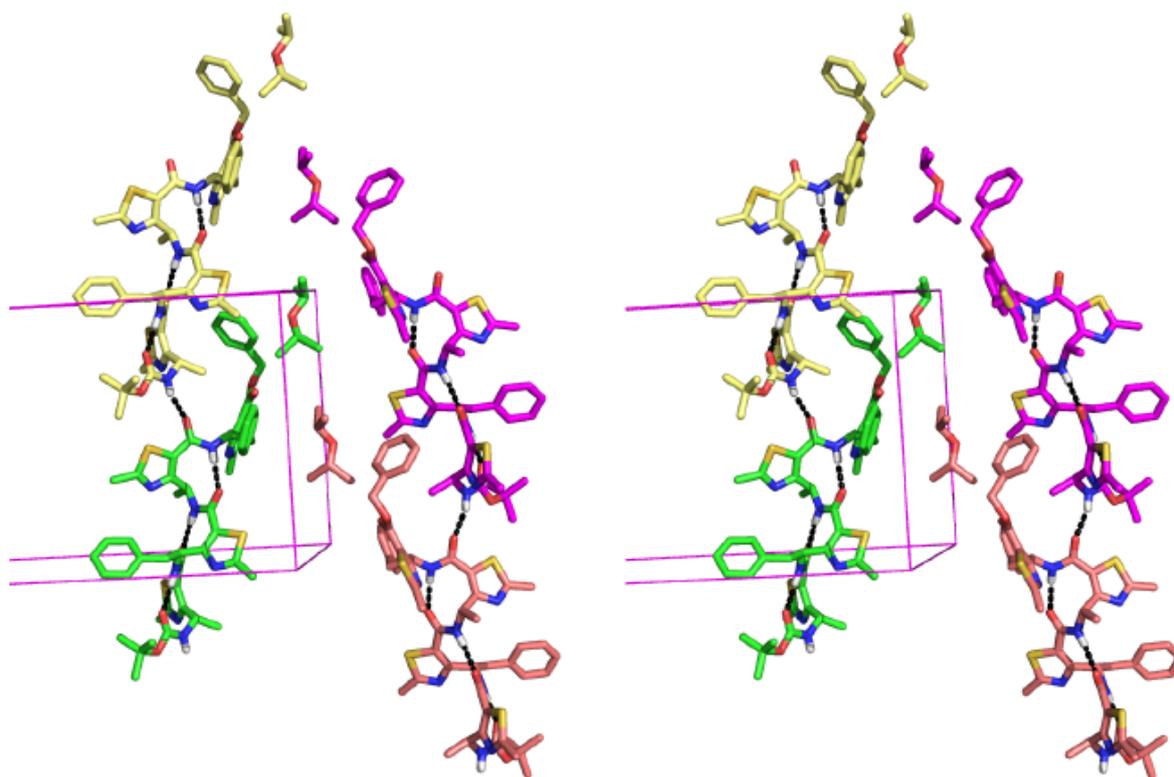


Figure S63: Stereoview of a part of the crystal packing of **(21b)**

Table S104: Backbone dihedral angles ( $^{\circ}$ ) of **(21b)** XRD structure. Values are ordered from C to N and are average on the 20 lowest energy NMR structures

Residue	$\varphi$	$\theta$	$\zeta$	$\psi$
1	-81 $^{\circ}$	+121 $^{\circ}$	-4 $^{\circ}$	-40 $^{\circ}$
2	-78 $^{\circ}$	+119 $^{\circ}$	0 $^{\circ}$	-38 $^{\circ}$
3	-75 $^{\circ}$	+121 $^{\circ}$	-1 $^{\circ}$	-45 $^{\circ}$
4	-80 $^{\circ}$	+147 $^{\circ}$	+3 $^{\circ}$	-165 $^{\circ}$

Crystal data for **M(1)-ATC5-3-NH<sub>2</sub>** - formula: C<sub>139</sub>H<sub>154</sub>N<sub>24</sub>O<sub>19</sub>S<sub>10</sub>, moiety: 2(C<sub>65</sub>H<sub>65</sub>N<sub>12</sub>O<sub>8</sub>S<sub>5</sub>),3(C<sub>3</sub>H<sub>8</sub>O),  $T = 100$  K,  $M_r = 2785.47$  gmol<sup>-1</sup>, crystal size = 0.02x0.03x0.06 mm<sup>3</sup>, monoclinic, spacegroup  $P1_21_1$ ,  $a = 19.014(4)$  Å,  $b = 9.7020(19)$  Å,  $c = 22.324(5)$  Å,  $\alpha = 90^{\circ}$ ,  $\beta = 113.75(3)^{\circ}$ ,  $\gamma = 90^{\circ}$ ,  $V = 3769.5(8)$  Å<sup>3</sup>,  $Z = 1$ ,  $\rho_{\text{calcd}} = 1.227$  gcm<sup>-3</sup>,  $\mu = 0.204$  mm<sup>-1</sup>,  $\Theta_{\text{max}} = 27.791^{\circ}$ , 57444 reflections measured, 18400 unique, 13068 least-squares, 13068 with  $I > 2\sigma(I)$ ,  $R_{\text{int}} = 0.037$ ,  $\langle \sigma(I)/I \rangle = 0.0367$ , 902 refined parameters,  $R_1(I > 2\sigma(I)) = 0.0945$ ,  $wR_2(I > 2\sigma(I)) = 0.1156$ ,  $R_1(\text{all data}) = 0.1165$ ,  $wR_2(\text{all data}) = 0.1531$ , GOF = 0.9490,  $\Delta\rho(\text{min/max}) = -0.76/1.69$  eÅ<sup>-3</sup>.

Crystal evaluation and data collection were performed on the XRD1 beamline of the Elettra synchrotron (Trieste, Italy) on a Huber 4-circles kappa goniometer and a Pilatus 2M detector. The incident wavelength was set at 0.7000 Å. The XDS-program<sup>3</sup> was used for frame integration using default parameters, for the empirical absorption correction using spherical harmonics employing symmetry-equivalent and redundant data, and for the correction for Lorentz and polarization effects. The crystal structure was solved using the *ab-initio* iterative charge flipping method with parameters described by Van der Lee<sup>4</sup> using the *Superflip* program<sup>5</sup> and it was refined using full-matrix least-squares procedures as implemented in *CRYSTALS*<sup>6</sup> on all reflections with  $I > 2\sigma(I)$ .

The H atoms were partially located in a difference map and partially positioned geometrically, and all were repositioned geometrically. They were initially refined with soft restraints on the bond lengths and angles to regularize their geometry (C---H in the range 0.93-0.98 Å) and  $U_{\text{iso}}(\text{H})$  (in the range 1.2-1.5 times  $U_{\text{eq}}$  of the parent atom), after which the positions were refined with riding constraints<sup>7</sup>. The three isopropanol moieties were approximately localized in difference Fourier maps, then regularized using the isopropanol moieties in the structure of CSD refcode DEBROU and finally refined as rigid bodies with the oxygen atoms as pivotal

atoms. The hydrogen atom of the carboxyl could not be located and is most probably exchanging between the two oxygens. The absolute configuration was assigned arbitrarily because of the absence of strong anomalous scatterers for this energy.

Complete crystallographic data – in cif/fcf format -for the structure reported in this manuscript have been deposited with Cambridge Crystallographic Data Centre as supplementary publication N° CCDC 1894172.

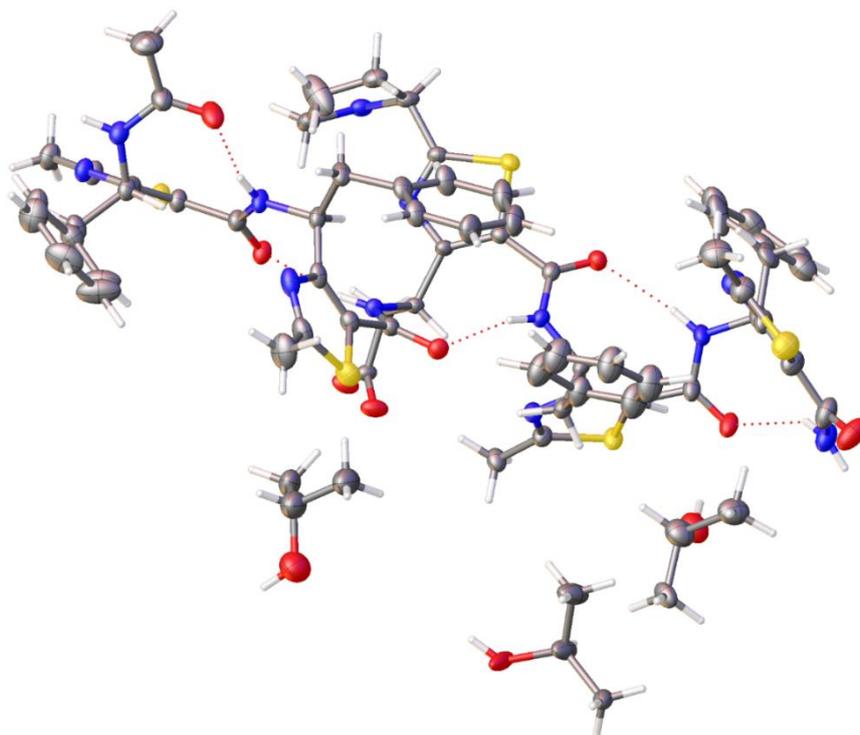


Figure S64 - Ortep style plot of structure **M(1)-ATC5-3-NH<sub>2</sub>** with atomic displacement ellipsoids at the 60% probability level.

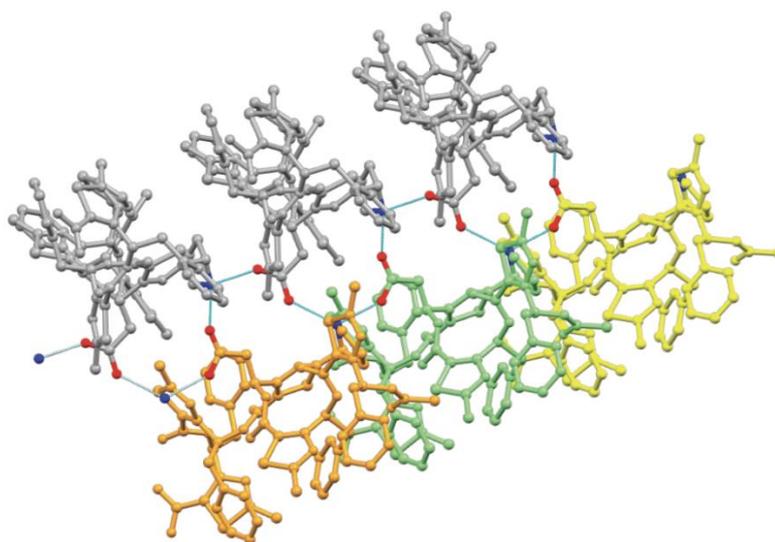
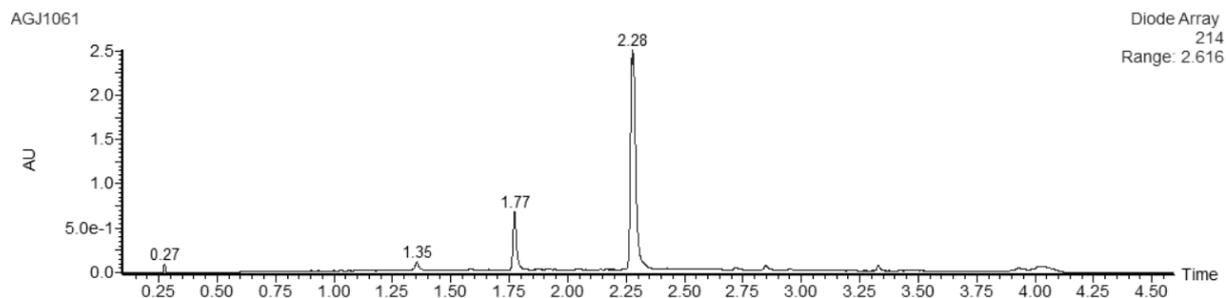


Figure S65 - Part of the crystal packing of **M(1)-ATC5-3-NH<sub>2</sub>**. Infinite columns define independent plans. Each pyrrolidine and each acid side chain form an intra- and an inter-molecular hydrogen-bond.

## X. SPECTRA AND HPLC ANALYSIS

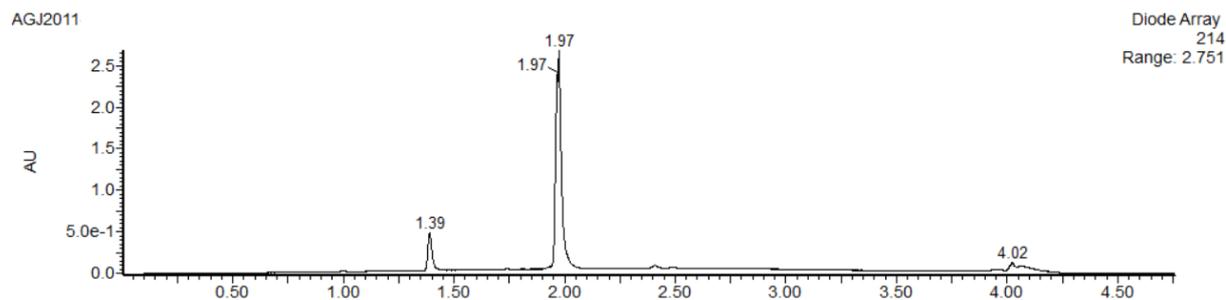
### 1. Compound (45a)

#### C18-RP HPLC analysis



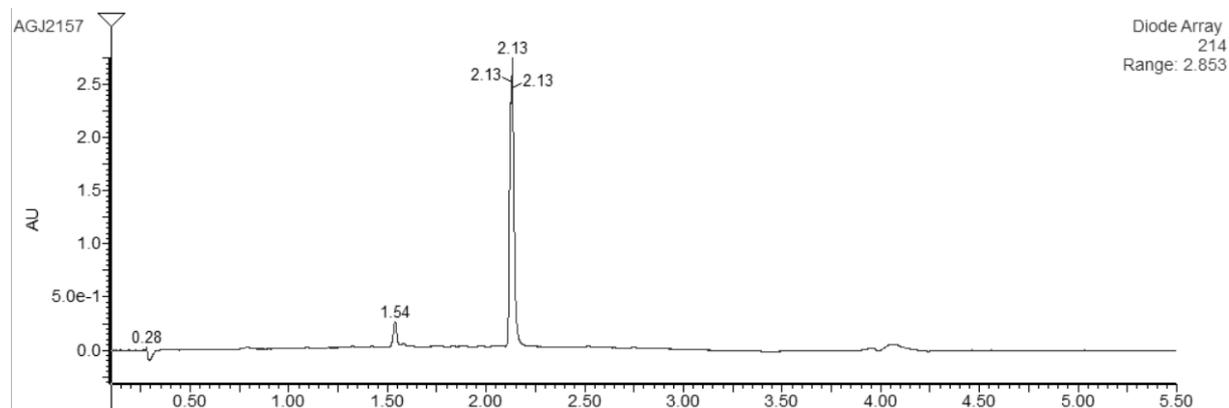
### 2. Compound (45b)

#### C18-RP HPLC analysis



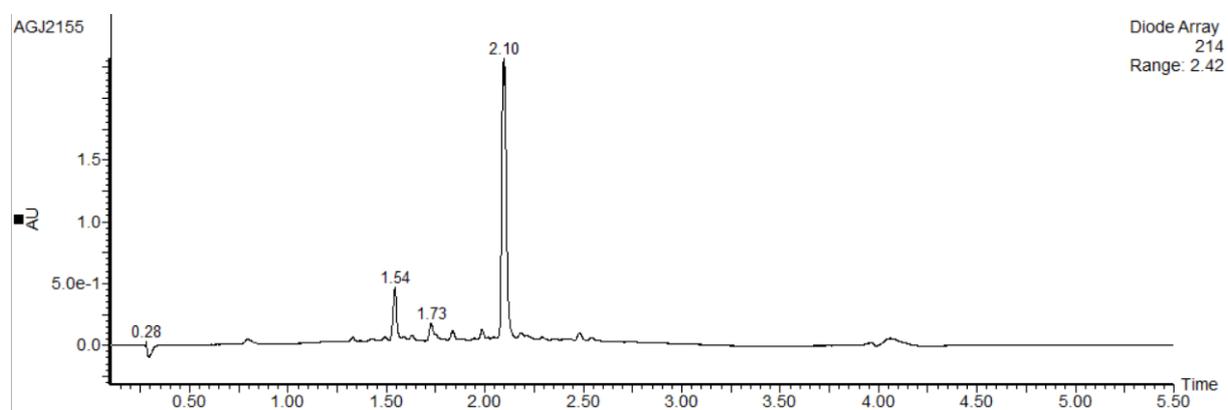
### 3. Compound (45c)

#### C18-RP HPLC analysis



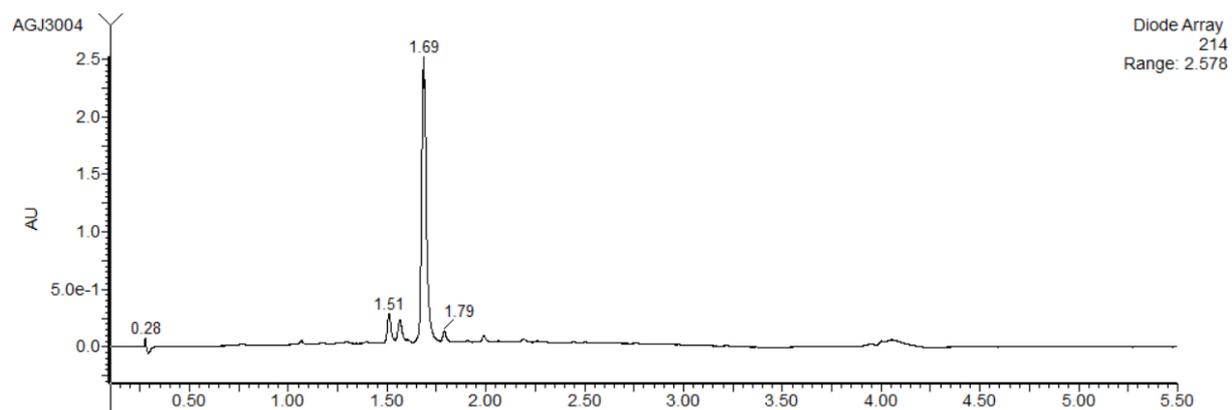
#### 4. Compound (45d)

##### C18-RP HPLC analysis



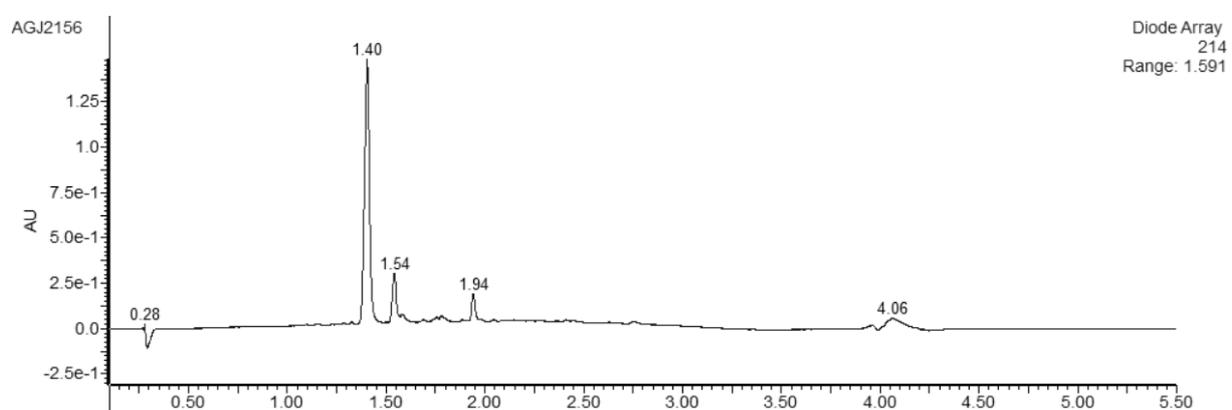
#### 5. Compound (45e)

##### C18-RP HPLC analysis



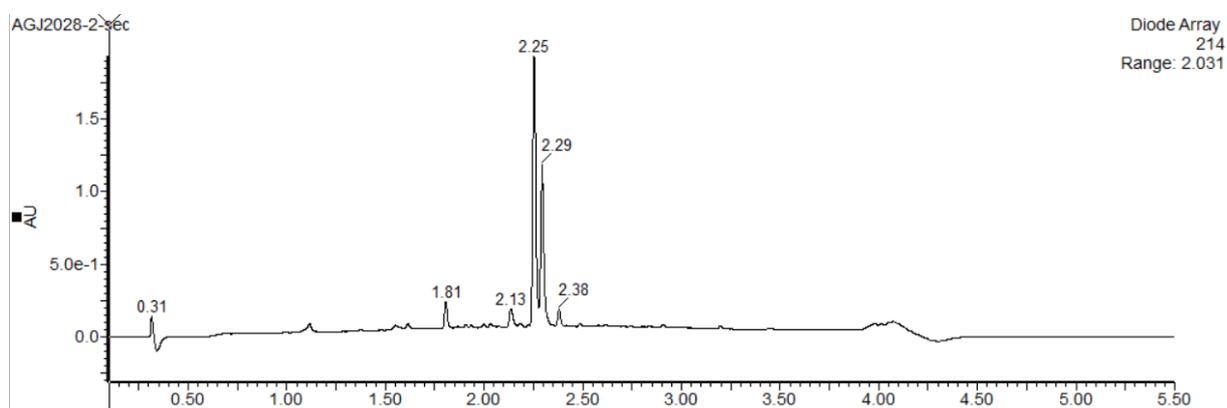
#### 6. Compound (45f)

##### C18-RP HPLC analysis



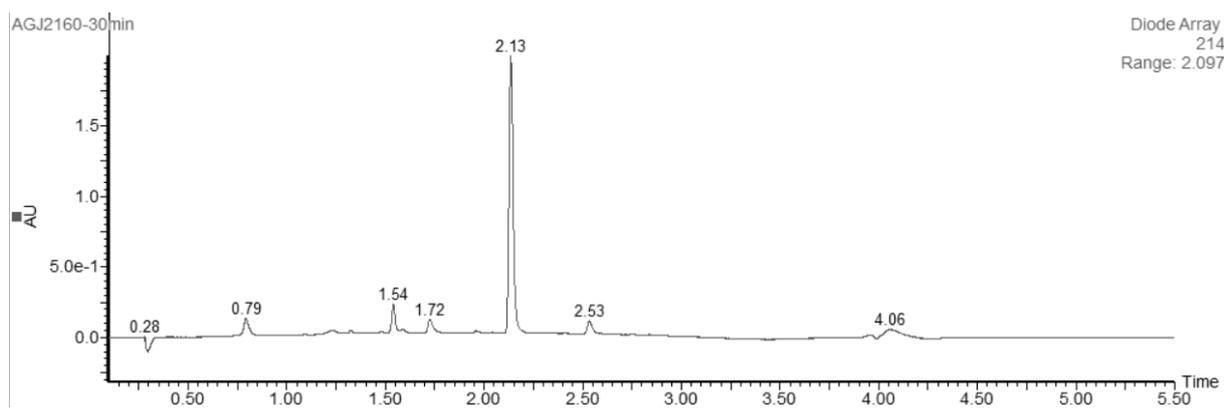
### 7. Compound (45g)

#### C18-RP HPLC analysis



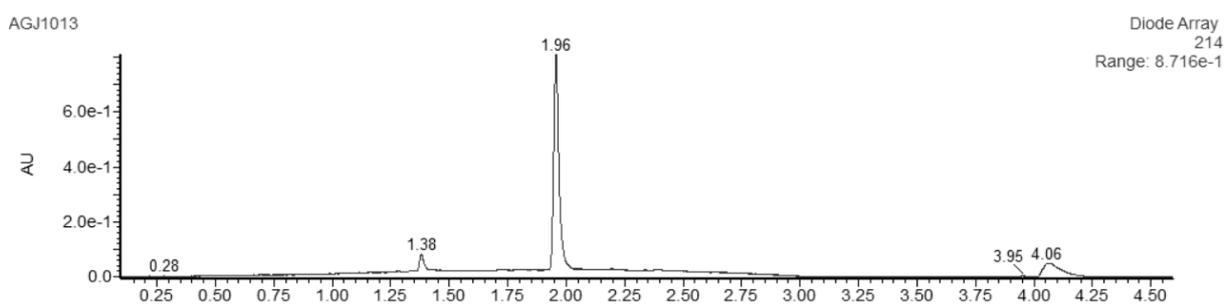
### 8. Compound (45h)

#### C18-RP HPLC analysis



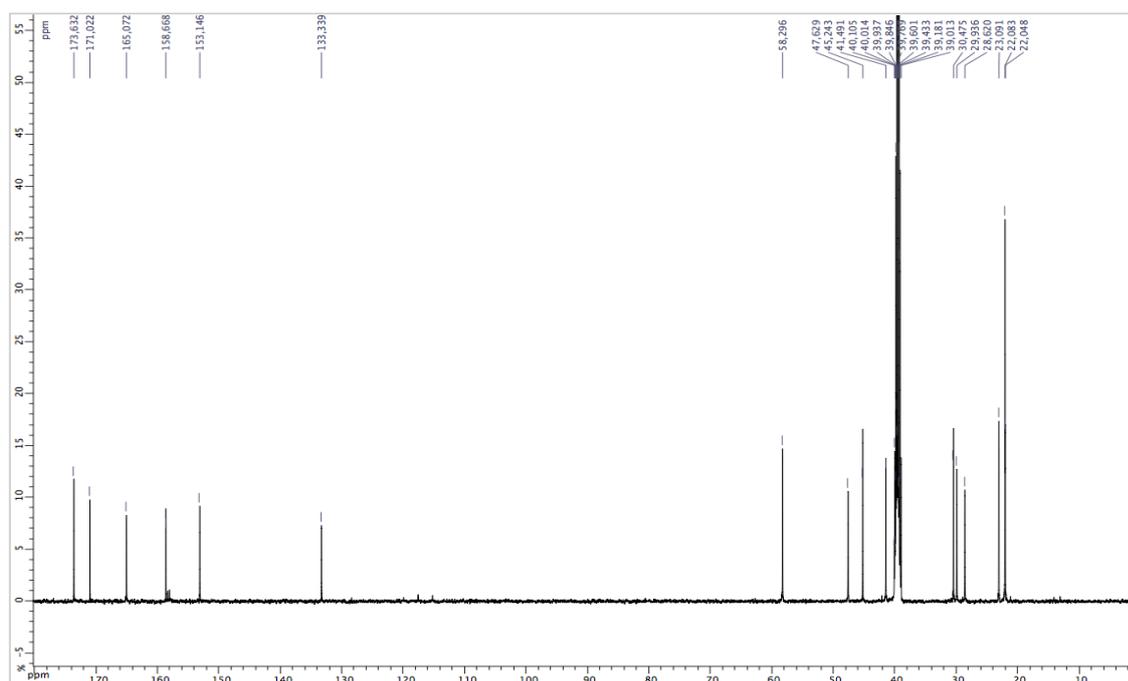
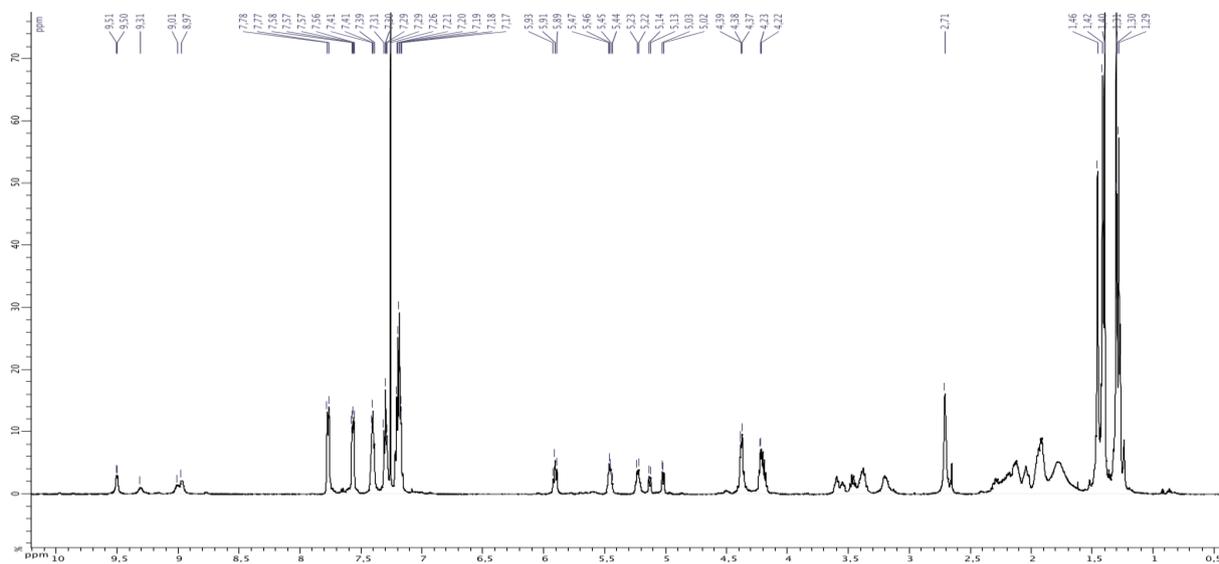
### 9. Compound (45i)

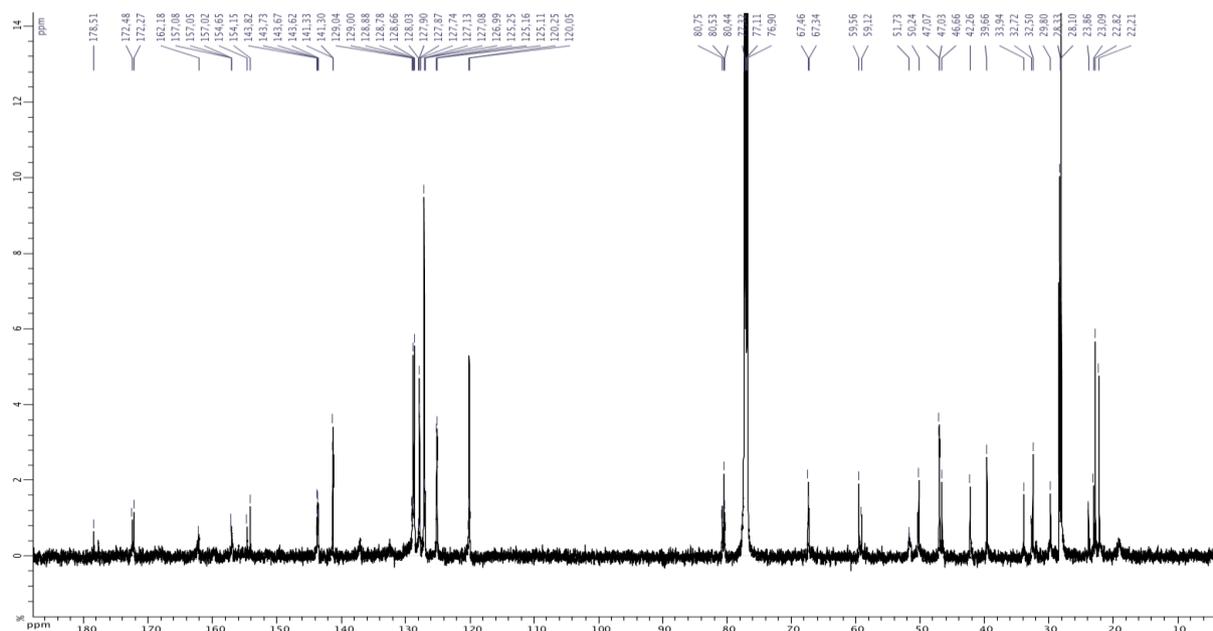
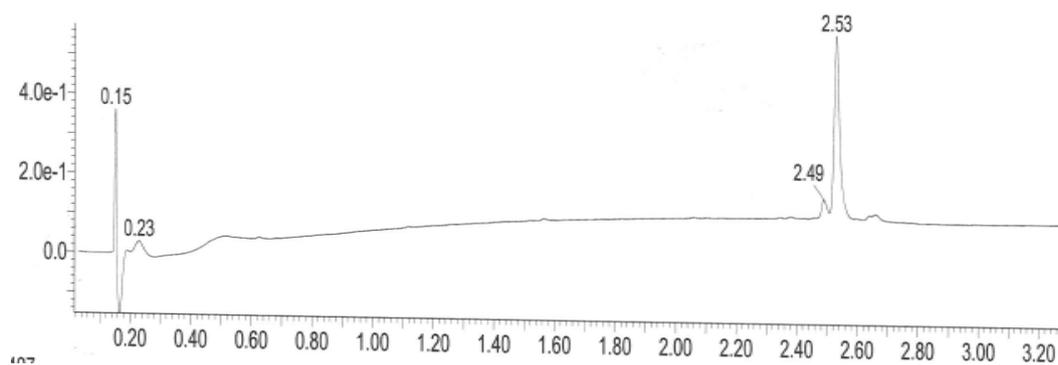
#### C18-RP HPLC analysis



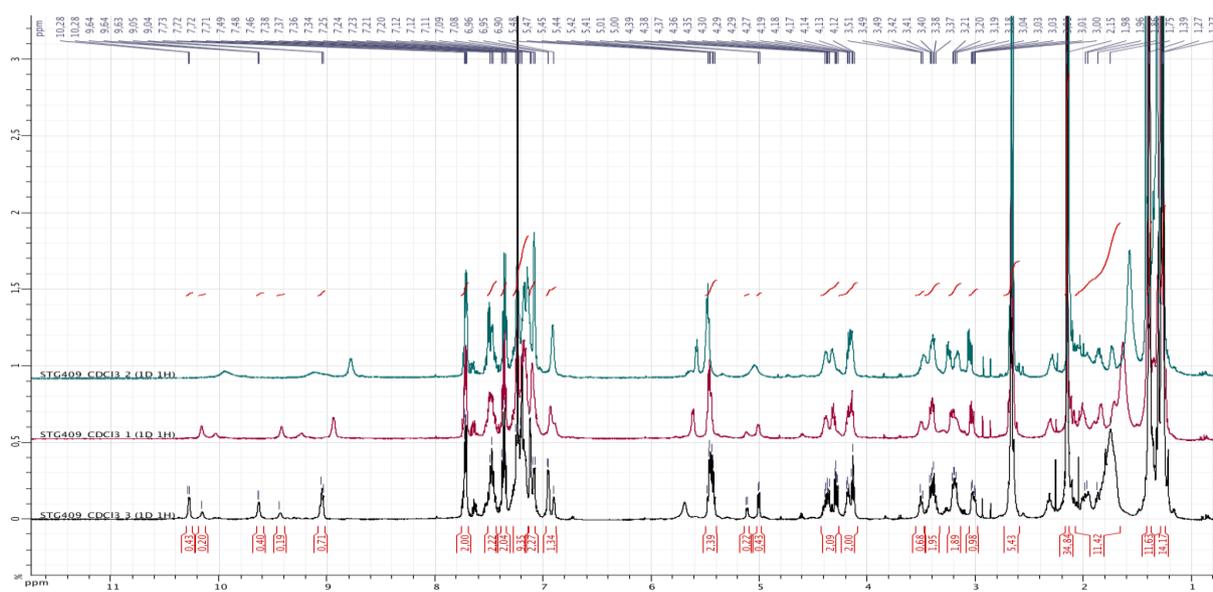




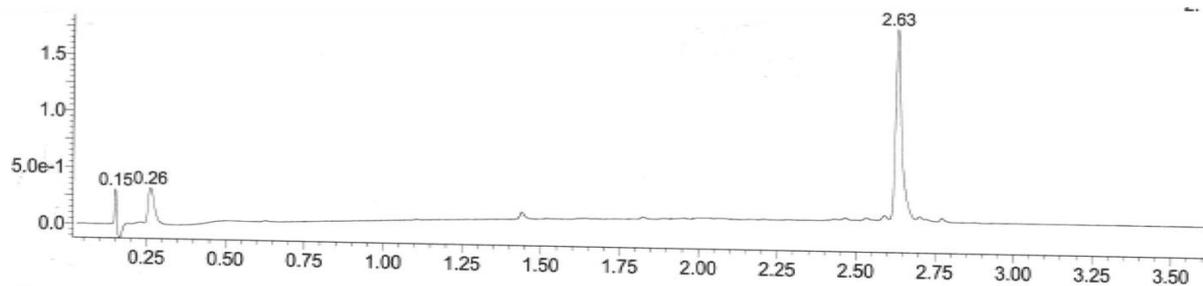
**$^{13}\text{C}$  NMR (d<sub>6</sub>-DMSO, 100 MHz)****14. Compound D****NMR  $^1\text{H}$  (CDCl<sub>3</sub>, 600 MHz) – 278K**

**$^{13}\text{C}$  NMR  $^1\text{H}$  ( $\text{CDCl}_3$ , 150 MHz) – 278K****C18-RP HPLC analysis**

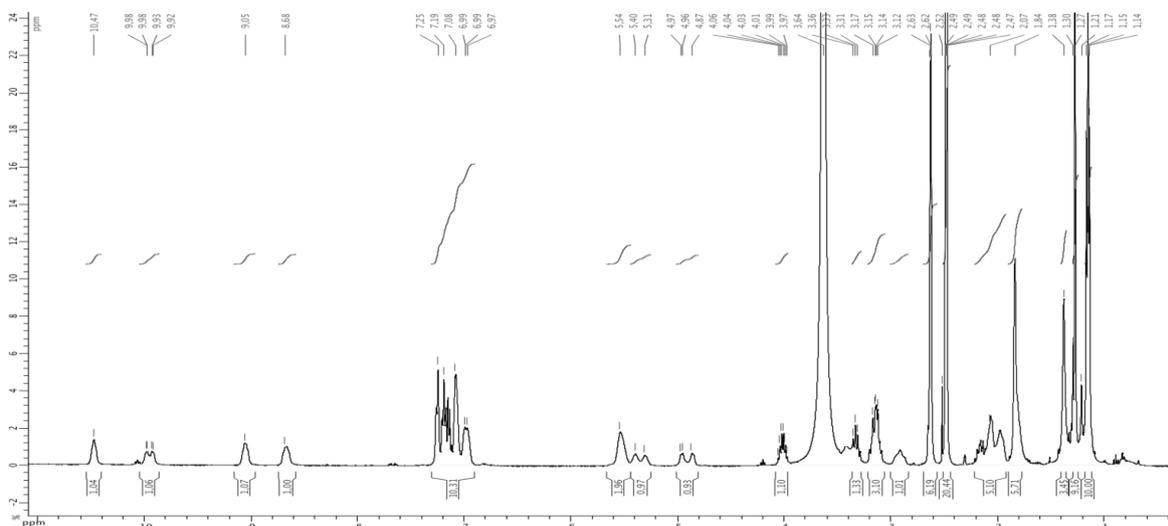


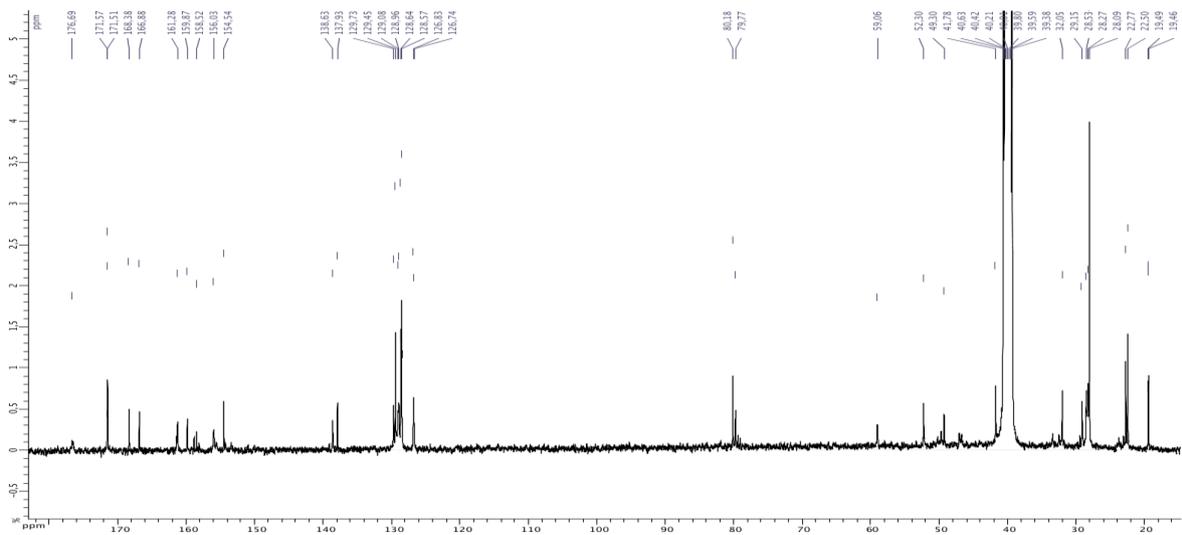
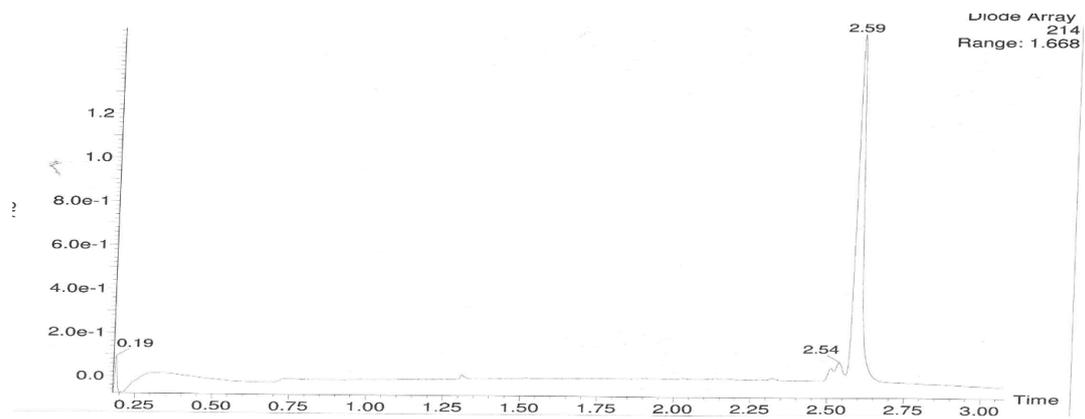
Superimposition of NMR  $^1\text{H}$  ( $\text{CDCl}_3$ , 600 MHz) at 278, 298 and 328 K

## C18-RP HPLC analysis



## 16. Compound F

 $^1\text{H}$  NMR ( $\text{CD}_3\text{OH}$ , 400 MHz)

**$^{13}\text{C}$  NMR (CD<sub>3</sub>OH, 100 MHz)****C18-RP HPLC analysis**

## REFERENCES



1. Nanda, V. & Koder, R. L. Designing Artificial Enzymes by Intuition and Computation. *Nat. Chem.*, **2**, 15–24 (2010).
2. Gellman, S. H. Foldamers: A Manifesto. *Acc. Chem. Res.*, **31**, 173–80 (1998).
3. Berg, J., Stryer, L. & Tymoczko, J. Chapter 3, Protein Structure and Function. in *Biochem. 5th Ed.*, (W H Freeman, 2002).
4. Grauer, A. & König, B. Peptidomimetics – A Versatile Route to Biologically Active Compounds. *Eur. J. Org. Chem.*, **2009**, 5099–5111 (2009).
5. Qvit, N., Rubin, S. J. S., Urban, T. J., Mochly-Rosen, D. & Gross, E. R. Peptidomimetic therapeutics: scientific approaches and opportunities. *Drug Discov. Today*, **22**, 454–62 (2017).
6. Gante, J. Peptidomimetics—Tailored Enzyme Inhibitors. *Angew. Chem. Int. Ed. Engl.*, **33**, 1699–1720 (1994).
7. Liskamp, R. M. J., Rijkers, D. T. S., Kruijtzter, J. A. W. & Kemmink, J. Peptides and Proteins as a Continuing Exciting Source of Inspiration for Peptidomimetics. *ChemBioChem*, **12**, 1626–53 (2011).
8. Allen, F. H. *et al.* Tables of Bond Lengths determined by X-Ray and Neutron Diffraction. Part I. Bond Lengths in Organic Compounds. *J Chem Soc Perkin Trans 2*, 19 (1987).
9. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–12 (2004).
10. Hollingsworth, S. A. & Karplus, P. A. A Fresh Look at the Ramachandran Plot and the Occurrence of Standard Structures in Proteins. *Biomol. Concepts*, **1**, 271–83 (2010).
11. Horton, H. R. Principles of Biochemistry. *Pearson Prentice Hall*, (2003).
12. Barlow, D. J. & Thornton, J. M. Helix Geometry in Proteins. *J. Mol. Biol.*, **201**, 601–19 (1988).
13. Fodje, M. N. & Al-Karadaghi, S. Occurrence, conformational features and amino acid propensities for the  $\pi$ -helix. *Protein Eng. Des. Sel.*, **15**, 353–58 (2002).
14. Adzhubei, A. A., Sternberg, M. J. E. & Makarov, A. A. Polyproline-II Helix in Proteins: Structure and Function. *J. Mol. Biol.*, **425**, 2100–2132 (2013).
15. Astbury, W. T. & Street, A. X-Ray Studies of the Structure of Hair, Wool, and Related Fibres. I. General. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, **230**, 75–101 (1932).
16. Astbury, W. T. & Bell, F. O. Nature of the Intramolecular Fold in Alpha-Keratin and Alpha-Myosin. *Nature*, **147**, 696–99 (1941).
17. Neurath, H. Intramolecular Folding of Polypeptide Chains in Relation to Protein Structure. *J. Phys. Chem.*, **44**, 296–305 (1940).
18. Pauling, L., Corey, R. B. & Branson, H. R. The Structure of Proteins: Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *Proc. Natl. Acad. Sci.*, **37**, 205–11 (1951).
19. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637 (1983).
20. Haimov, B. & Srebnik, S. A Closer Look into the  $\alpha$ -Helix Basin. *Sci. Rep.*, **6**, 38341 (2016).
21. Bragg, L., Kendrew, J. C. & Perutz, M. F. Polypeptide Chain Configurations in Crystalline Proteins. *Proc. R. Soc. Lond. Ser. Math. Phys. Sci.*, **203**, 321–57 (1950).

22. Arnott, S. & Wonacott, A. J. Atomic Co-ordinates for an  $\alpha$ -Helix: Refinement of the Crystal Structure of  $\alpha$ -Poly-L-alanine. *J. Mol. Biol.*, **21**, 371–83 (1966).
23. Kumar, S. & Bansal, M. Geometrical and Sequence Characteristics of alpha-Helices in Globular Proteins. *Biophys. J.*, **75**, 1935–44 (1998).
24. Maestro, Schrödinger, LLC. *Schrödinger Release 2019-1 New-York NY*, (2019).
25. Toniolo, C. & Benedetti, E. The Polypeptide 3.10-Helix. *Trends Biochem. Sci.*, **16**, 350–53 (1991).
26. Cooley, R. B., Arp, D. J. & Karplus, P. A. Evolutionary Origin of a Secondary Structure:  $\pi$ -Helices as Cryptic but Widespread Insertional Variations of  $\alpha$ -Helices That Enhance Protein Functionality. *J. Mol. Biol.*, **404**, 232–46 (2010).
27. Kakinoki, S., Hirano, Y. & Oka, M. On the Stability of Polyproline-I and II Structures of Proline Oligopeptides. *Polym. Bull.*, **53**, 109–15 (2005).
28. Nelson, D. L. & Cox, M. M. *Lehninger Principles of Biochemistry*. (W.H.Freeman & Co Ltd, 2008).
29. Lubell, W. D. *Peptidomimetics II*. (Springer, 2016).
30. Venkatachalam, C. M. Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. *Biopolymers*, **6**, 1425–36 (1968).
31. Chou, K.-C. Prediction of Tight Turns and Their Types in Proteins. *Anal. Biochem.*, **286**, 1–16 (2000).
32. Koch, O. & Klebe, G. Turns revisited: A uniform and comprehensive classification of normal, open, and reverse turn families minimizing unassigned random chain portions. *Proteins Struct. Funct. Bioinforma.*, **74**, 353–67 (2009).
33. Hutchinson, E. G. & Thornton, J. M. A Revised Set of Potentials for Beta-Turn Formation in Proteins. *Protein Sci.*, 2207–16 (1994).
34. Matthews, B. W. The  $\gamma$  Turn. Evidence for a New Folded Conformation in Proteins. *Macromolecules*, **5**, 818–19 (1972).
35. Némethy, G. & Printz, M. P. The  $\gamma$  Turn, a Possible Folded Conformation of the Polypeptide Chain. Comparison with the  $\beta$  Turn. *Macromolecules*, **5**, 755–58 (1972).
36. Smith, J. A., Pease, L. G. & Kopple, K. D. Reverse Turns in Peptides and Protein. *Crit. Rev. Biochem.*, **8**, 315–99 (1980).
37. Valle, G. *et al.* Intramolecularly Hydrogen-bonded Peptide Conformations. Preferred Crystal- state and Solution Conformations of N-Monochloroacetylated Glycines Dialkylated at the  $\alpha$ -Carbon Atom. *J Chem Soc Perkin Trans 2*, 885–89 (1986).
38. Rose, G. D. & Gierasch, L. M. Turns in Peptides and Proteins. *Adv. Protein Chem.*, **37**, 1–109 (1985).
39. Milner-White, E. J., Ross, B. M., Ismail, R., Belhadj-Mostefa, K. & Poet, R. One type of gamma-turn, rather than the other gives rise to chain-reversal in proteins. *J. Mol. Biol.*, **204**, 777–82 (1988).
40. Chou, P. Y. & Fasman, G. D. Conformational Parameters for Amino Acids in Helical,  $\beta$ -Sheet, and Random Coil Regions Calculated from Proteins! *Biochemistry*, **13**, 211–22 (1974).
41. Chou, P. Y. & Fasman, G. D. Prediction of Protein Conformation. *Biochemistry*, **13**, 222–45 (1974).

42. Williams, R. W., Chang, A., Juretić, D. & Loughran, S. Secondary structure predictions and medium range interactions. *Biochim. Biophys. Acta*, **916**, 200–204 (1987).
43. Imai, K. & Mitaku, S. Mechanisms of secondary structure breakers in soluble proteins. *Biophysics*, **1**, 55–65 (2005).
44. Hill, D. J., Mio, M. J., Prince, R. B., Hughes, T. S. & Moore, J. S. A Field Guide to Foldamers. *Chem. Rev.*, **101**, 3893–4012 (2001).
45. Hecht, S. & Huc, I. *Foldamers: Structure, Properties and Applications*. (John Wiley & Sons, 2007).
46. Bartus, É. *et al.* De Novo Modular Development of a Foldameric Protein-Protein Interaction Inhibitor for Separate Hot Spots: A Dynamic Covalent Assembly Approach. *ChemistryOpen*, **6**, 236–41 (2017).
47. Mándity, I. M. & Fülöp, F. An overview of peptide and peptoid foldamers in medicinal chemistry. *Expert Opin. Drug Discov.*, **10**, 1163–77 (2015).
48. Schmitt, M. A., Weisblum, B. & Gellman, S. H. Unexpected Relationships between Structure and Function in  $\alpha,\beta$ -Peptides: Antimicrobial Foldamers with Heterogeneous Backbones. *J. Am. Chem. Soc.*, **126**, 6848–49 (2004).
49. Schmitt, M. A., Weisblum, B. & Gellman, S. H. Interplay among Folding, Sequence, and Lipophilicity in the Antibacterial and Hemolytic Activities of  $\alpha/\beta$ -Peptides. *J. Am. Chem. Soc.*, **129**, 417–28 (2007).
50. Simon, M. *et al.* Can Heterocyclic  $\gamma$ -Peptides Provide Polyfunctional Platforms for Synthetic Glycocluster Construction? *Chem. - Eur. J.*, **24**, 11426–32 (2018).
51. Li, C. *et al.* Strong Stacking between F···H-N Hydrogen-Bonded Foldamers and Fullerenes: Formation of Supramolecular Nano Networks. *Chem. - Eur. J.*, **13**, 9990–98 (2007).
52. Maayan, G., Ward, M. D. & Kirshenbaum, K. Folded Biomimetic Oligomers for Enantioselective Catalysis. *Proc. Natl. Acad. Sci.*, **106**, 13679–84 (2009).
53. Bécart, D. *et al.* Helical Oligourea Foldamers as Powerful Hydrogen Bonding Catalysts for Enantioselective C–C Bond-Forming Reactions. *J. Am. Chem. Soc.*, **139**, 12524–32 (2017).
54. Guichard, G. & Huc, I. Synthetic foldamers. *Chem. Commun.*, **47**, 5933 (2011).
55. Seebach, D. *et al.*  $\beta$ -Peptides: Synthesis by Arndt-Eistert Homologation with Concomitant Peptide Coupling. Structure Determination by NMR and CD Spectroscopy and by X-ray Crystallography. Helical Secondary Structure of a  $\beta$ -Hexapeptide in Solution and Its Stability towards Pepsin. *Helv. Chim. Acta*, **79**, 913–41 (1996).
56. De Pol, S., Zorn, C., Klein, C. D., Zerbe, O. & Reiser, O. Surprisingly Stable Helical Conformations in  $\alpha/\beta$ -Peptides by Incorporation of cis- $\beta$ -Aminocyclopropane Carboxylic Acids. *Angew. Chem. Int. Ed.*, **43**, 511–14 (2004).
57. Hayen, A., Schmitt, M. A., Ngassa, F. N., Thomasson, K. A. & Gellman, S. H. Two Helical Conformations from a Single Foldamer Backbone: “Split Personality” in Short  $\alpha/\beta$ -Peptides. *Angew. Chem. Int. Ed.*, **43**, 505–10 (2004).
58. Bolin, K. A. & Millhauser, G. L.  $\alpha$  and  $\beta$ : The Split Personality of Polypeptide Helices. *Acc. Chem. Res.*, **32**, 1027–33 (1999).
59. Karle, I. L. & Balaram, P. Perspectives in Biochemistry. *Biochemistry*, **29**, 6747–56 (1990).

60. Chiang, Y.-C., Lin, Y.-J. & Horng, J.-C. Stereoelectronic Effects on the Transition Barrier of Polyproline Conformational Interconversion. *Protein Sci. Publ. Protein Soc.*, **18**, 1967–77 (2009).
61. Sreerama, N. & Woody, R. W. Molecular Dynamics Simulations of Polypeptide Conformations in Water: A Comparison of  $\alpha$ ,  $\beta$  and poly(Pro)II Conformations. *Proteins Struct. Funct. Genet.*, **36**, 400–406 (1999).
62. Sun, J. & Zuckermann, R. N. Peptoid Polymers: A Highly Designable Bioinspired Material. *ACS Nano*, **7**, 4715–32 (2013).
63. Banerjee, A. & Balaram, P. Stereochemistry of Peptides and Polypeptides Containing Omega Amino Acids. *Curr. Sci.*, **73**, 1067–77 (1997).
64. Bode, K. A. & Applequist, J. Poly( $\beta$ -amino acid) Helices. Theoretical  $\pi$ - $\pi^*$  Absorption and Circular Dichroic Spectra. *Macromolecules*, **30**, 2144–50 (1997).
65. Bella, J., Alemán, C., Fernández-Santin, J. M., Alegre, C. & Subirana, J. A. Conformation of the Helical Polyamide Poly(cu-isobutyl L-aspartate). *Macromolecules*, **25**, 5225–30 (1992).
66. Appella, D. H., Christianson, L. A., Karle, I. L., Powell, D. R. & Gellman, S. H.  $\beta$ -Peptide Foldamers: Robust Helix Formation in a New Family of  $\beta$ -Amino Acid Oligomers. *J. Am. Chem. Soc.*, **118**, 13071–72 (1996).
67. Cheng, R. P., Gellman, S. H. & DeGrado, W. F.  $\beta$ -Peptides: From Structure to Function. *Chem. Rev.*, **101**, 3219–32 (2001).
68. Seebach, D. *et al.* Probing the Helical Secondary Structure of Short-Chain  $\beta$ -Peptides. *Helv. Chim. Acta*, **79**, 2043–66 (1996).
69. Appella, D. H. *et al.* Residue-Based Control of Helix Shape in  $\beta$ -Peptide Oligomers. *Nature*, **387**, 381–84 (1997).
70. Seebach, D. *et al.* NMR-Solution Structures in Methanol of an  $\alpha$ -Heptapeptide, of a  $\beta$ 3/ $\beta$ 2-Nonapeptide, and of an all- $\beta$ 3-Icosapeptide Carrying the 20 Proteinogenic Side Chains. *Helv. Chim. Acta*, **88**, 1969–82 (2005).
71. DeGrado, W. F., Schneider, J. P. & Hamuro, Y. The Twists and Turns of  $\beta$ -peptides. *J. Pept. Res.*, **54**, 206–17 (2008).
72. Appella, D. H. *et al.* Synthesis and Structural Characterization of Helix-Forming  $\beta$ -Peptides: trans-2-Aminocyclopentanecarboxylic Acid Oligomers. *J. Am. Chem. Soc.*, **121**, 7574–81 (1999).
73. Fernandes, C. *et al.* 12-Helix Folding of Cyclobutane  $\beta$ -Amino Acid Oligomers. *Org. Lett.*, **12**, 3606–9 (2010).
74. Hetényi, A. *et al.* Sculpting the  $\beta$ -peptide foldamer H12 helix via a designed side-chain shape. *Chem Commun*, 177–79 (2009).
75. Seebach, D. *et al.* ‘Mixed’  $\beta$ -peptides: A unique helical secondary structure in solution. Preliminary communication. *Helv. Chim. Acta*, **80**, 2033–38 (1997).
76. Wu, Y.-D., Han, W., Wang, D.-P., Gao, Y. & Zhao, Y.-L. Theoretical Analysis of Secondary Structures of  $\beta$ -Peptides. *Acc. Chem. Res.*, **41**, 1418–27 (2008).
77. Rueping, M., Schreiber, J. V., Jaun, B. & Seebach, D. Mixed  $\beta$ 2/ $\beta$ 3-Hexapeptides and  $\beta$ 2/ $\beta$ 3-Nonapeptides Folding to (P)-Helices with Alternating Twelve- and Ten-Membered Hydrogen-Bonded Rings. *Helv. Chim. Acta*, **85**, 2577–93 (2002).

78. Claridge, T. D. W. *et al.* 10-Helical conformations in oxetane  $\beta$ -amino acid hexamers. *Tetrahedron Lett.*, **5** (2001).
79. Ragab, S. S. *et al.* Cooperative 5- and 10-membered ring interactions in the 10-helix folding of oxetin homo-oligomers. *Chem. Commun.*, **54**, 1968–71 (2018).
80. Günther, R., Hofmann, H.-J. & Kuczera, K. Searching for Periodic Structures in  $\beta$ -Peptides. *J. Phys. Chem. B*, **105**, 5559–67 (2001).
81. Gademann, K., Häne, A., Rueping, M., Jaun, B. & Seebach, D. The Fourth Helical Secondary Structure of  $\beta$ -Peptides: The (P)-2.8-Helix of a  $\beta$ -Hexapeptide Consisting of (2R,3S)-3-Amino-2-hydroxy Acid Residues. *Angew. Chem. Int. Ed.*, **42**, 1534–37 (2003).
82. Altmayer-Henzien, A. *et al.* Solution State Conformational Preferences of Dipeptides Derived from N-Aminoazetidincarboxylic Acid: An Assessment of the Hydrazino Turn. *J. Org. Chem.*, **78**, 6031–39 (2013).
83. Altmayer-Henzien, A. *et al.* Fine Tuning of  $\beta$ -Peptide Foldamers: a Single Atom Replacement Holds Back the Switch from an 8-Helix to a 12-Helix. *Angew. Chem. Int. Ed.*, **54**, 10807–10 (2015).
84. Declerck, V. & Aitken, D. J. Strategic C to N Replacement in  $\beta$ -Peptides: Atomic Level Control of Helical Folding. *J. Org. Chem.*, **83**, 8793–8800 (2018).
85. Baldauf, C., Günther, R. & Hofmann, H.-J. Helix Formation and Folding in  $\gamma$ -Peptides and Their Vinylogues. *Helv. Chim. Acta*, **86**, 2573–88 (2003).
86. Hanessian, S., Luo, X., Schaum, R. & Michnick, S. Design of Secondary Structures in Unnatural Peptides: Stable Helical  $\gamma$ -Tetra-, Hexa-, and Octapeptides and Consequences of  $\alpha$ -Substitution. *J. Am. Chem. Soc.*, **120**, 8569–70 (1998).
87. Hintermann, T., Gademann, K., Jaun, B. & Seebach, D.  $\gamma$ -Peptides Forming More Stable Secondary Structures than  $\alpha$ -Peptides: Synthesis and helical NMR-solution structure of the  $\gamma$ -hexapeptide analog of H-(Val-Ala-Leu)2-OH. *Helv. Chim. Acta*, **81**, 983–1002 (1998).
88. Hanessian, S., Luo, X. & Schaum, R. Synthesis and Folding Preferences of  $\gamma$ -Amino Acid Oligopeptides: Stereochemical Control in the Formation of a Reverse Turn and a Helix. *Tetrahedron Lett.*, **40**, 4925–29 (1999).
89. Seebach, D., Brenner, M., Rueping, M., Schweizer, B. & Jaun, B. Preparation and Determination of X-ray-Crystal and NMR-Solution Structures of  $\gamma$ 2,3,4-Peptides. *Chem. Commun.*, 207–8 (2001).
90. Seebach, D., Brenner, M., Rueping, M. & Jaun, B.  $\gamma$ 2-,  $\gamma$ 3-, and  $\gamma$ 2,3,4-Amino Acids, Coupling to  $\gamma$ -Hexapeptides: CD Spectra, NMR Solution and X-ray Crystal Structures of  $\gamma$ -Peptides. *Chem. – Eur. J.*, **8**, 573–84 (2002).
91. Seebach, D., Beck, A. K. & Bierbaum, D. J. The World of  $\beta$ - and  $\gamma$ -Peptides Comprised of Homologated Proteinogenic Amino Acids and Other Components. *Chem. Biodivers.*, **1**, 1111–1239 (2004).
92. Seebach, D., Hook, D. F. & Glättli, A. Helices and other secondary structures of  $\beta$ - and  $\gamma$ -peptides. *Pept. Sci.*, **84**, 23–37 (2005).
93. Vasudev, P. G., Shamala, N., Ananda, K. & Balaram, P. C9 Helices and Ribbons in  $\gamma$ -Peptides: Crystal Structures of Gabapentin Oligomers. *Angew. Chem. Int. Ed.*, **44**, 4972–75 (2005).

94. Sharma, G. V. M. *et al.* A Left-Handed 9-Helix in  $\gamma$ -Peptides: Synthesis and Conformational Studies of Oligomers with Dipeptide Repeats of C-Linked Carbo- $\gamma$ 4-amino Acids and  $\gamma$ -Aminobutyric Acid. *Angew. Chem. Int. Ed.*, **45**, 2944–47 (2006).
95. Guo, L. *et al.* Characteristic Structural Parameters for the  $\gamma$ -Peptide 14-Helix: Importance of Subunit Preorganization. *Angew. Chem. Int. Ed.*, **50**, 5843–46 (2011).
96. Farrera-Sinfreu, J. *et al.* A New Class of Foldamers Based on cis- $\gamma$ -Amino-1-proline1,2. *J. Am. Chem. Soc.*, **126**, 6048–57 (2004).
97. Kothari, A., Qureshi, M. K. N., Beck, E. M. & Smith, M. D. Bend-ribbon forming  $\gamma$ -peptides. *Chem. Commun.*, **0**, 2814–16 (2007).
98. Smith, E. M., Holmes, D. L., Shaka, A. J. & Nowick, J. S. An Artificial Antiparallel  $\beta$ -Sheet Containing a New Peptidomimetic Template. *J. Org. Chem.*, **62**, 7906–7 (1997).
99. Jones, C. R. *et al.* A Nonpeptidic Reverse Turn that Promotes Parallel Sheet Structure Stabilized by C-H...O Hydrogen Bonds in a Cyclopropane Gamma-Peptide. *Angew. Chem. Int. Ed Engl.*, **47**, 7099–7102 (2008).
100. Baldauf, C., Günther, R. & Hofmann, H.-J. Control of Helix Formation in Vinylogous  $\gamma$ -Peptides by (E)- and (Z)-Double Bonds: A Way to Ion Channels and Monomolecular Nanotubes. *J. Org. Chem.*, **70**, 5351–61 (2005).
101. Grison, C., Coutrot, P., Genève, S., Didierjean, C. & Marraud, M. Structural Investigation of “cis” and “trans” Vinylogous Peptides: cis-Vinylog Turn in Folded cis-Vinylogous Peptides, an Excellent Mimic of the Natural  $\beta$ -Turn. *J. Org. Chem.*, **70**, 10753–64 (2005).
102. Mann, E. & Kessler, H. New Oxazole-Based Peptidomimetics: Useful Building Blocks for the Synthesis of Orthogonally Protected Macrocyclic Scaffolds. *Org. Lett.*, **5**, 4567–70 (2003).
103. Mathieu, L. *et al.* Helical Oligomers of Thiazole-Based  $\gamma$ -Amino Acids: Synthesis and Structural Studies. *Angew. Chem. Int. Ed.*, **52**, 6006–10 (2013).
104. Bonnel, C. *et al.* FT-IR and NMR Structural Markers for Thiazole-Based  $\gamma$ -Peptide Foldamers. *Org. Biomol. Chem.*, **14**, 8664–69 (2016).
105. Case, D. A. *et al.* AMBER 10. (2008).
106. Ganesh Kumar, M. *et al.* Non-classical Helices with cis Carbon-Carbon Double Bonds in the Backbone: Structural Features of  $\alpha,\gamma$ -Hybrid Peptide Foldamers. *Angew. Chem. Int. Ed.*, **55**, 7847–51 (2016).
107. Bonnel, C. *et al.* C9/12 Ribbon-Like Structures in Hybrid Peptides Alternating  $\alpha$ - and Thiazole-Based  $\gamma$ -Amino Acids. *Chem. - Eur. J.*, **23**, 17584–91 (2017).
108. Nowick, J. S., Powell, N. A., Martinez, E. J., Smith, E. M. & Noronha, G. Molecular scaffolds. I. Intramolecular Hydrogen Bonding in a Family of Di- and Triureas. *J. Org. Chem.*, **57**, 3763–65 (1992).
109. Nowick, J. S. *et al.* Molecular Scaffolds. 2. Intramolecular Hydrogen Bonding in 1,2-Diaminoethane Diureas. *J. Am. Chem. Soc.*, **117**, 89–99 (1995).
110. Burgess, K., Shin, H. & Linthicum, D. S. Solid-Phase Syntheses of Unnatural Biopolymers Containing Repeating Urea Units. *Angew. Chem. Int. Ed. Engl.*, **34**, 907–9 (1995).
111. Burgess, K. *et al.* Solid Phase Syntheses of Oligoureas. *J. Am. Chem. Soc.*, **119**, 1556–64 (1997).

112. Kim, J.-M., Bi, Y., Paikoff, S. J. & Schultz, P. G. The Solid Phase Synthesis of Oligoureas. *Tetrahedron Lett.*, **37**, 5305–8 (1996).
113. Semetey, V. *et al.* Stable Helical Secondary Structure in Short-Chain N,N'-Linked Oligoureas Bearing Proteinogenic Side Chains. *Angew. Chem. Int. Ed.*, **41**, 1893–95 (2002).
114. Hemmerlin, C. *et al.* Helix-Forming Oligoureas: Temperature-Dependent NMR, Structure Determination, and Circular Dichroism of a Nonamer with Functionalized Side Chains. *Helv. Chim. Acta*, **85**, 3692–3711 (2002).
115. Violette, A. *et al.* N,N'-Linked Oligoureas as Foldamers: Chain Length Requirements for Helix Formation in Protic Solvent Investigated by Circular Dichroism, NMR Spectroscopy, and Molecular Dynamics. *J. Am. Chem. Soc.*, **127**, 2156–64 (2005).
116. Violette, A. *et al.* Exploring Helical Folding of Oligoureas During Chain Elongation by High-Resolution Magic-Angle-Spinning (HRMAS) NMR Spectroscopy. *Chem. – Eur. J.*, **14**, 3874–82 (2008).
117. Tew, G. N., Scott, R. W., Klein, M. L. & DeGrado, W. F. De Novo Design of Antimicrobial Polymers, Foldamers, and Small Molecules: From Discovery to Practical Applications. *Acc. Chem. Res.*, **43**, 30–39 (2010).
118. Hamuro, Y., Schneider, J. P. & DeGrado, W. F. De Novo Design of Antibacterial  $\beta$ -Peptides. *J. Am. Chem. Soc.*, **121**, 12200–201 (1999).
119. Liu, D. & DeGrado, W. F. De Novo Design, Synthesis, and Characterization of Antimicrobial  $\beta$ -Peptides. *J. Am. Chem. Soc.*, **123**, 7553–59 (2001).
120. Porter, E. A., Wang, X., Lee, H. S., Weisblum, B. & Gellman, S. H. Non-haemolytic beta-amino-acid oligomers. *Nature*, **404**, 565 (2000).
121. Porter, E. A., Weisblum, B. & Gellman, S. H. Mimicry of Host-Defense Peptides by Unnatural Oligomers: Antimicrobial  $\beta$ -Peptides. *J. Am. Chem. Soc.*, **124**, 7324–30 (2002).
122. Arvidsson, P. I. *et al.* Exploring the Antibacterial and Hemolytic Activity of Shorter- and Longer-Chain  $\beta$ -,  $\alpha,\beta$ -, and  $\gamma$ -Peptides, and of  $\beta$ -Peptides from  $\beta$ 2-3-Aza- and  $\beta$ 3-2-Methylidene-amino Acids Bearing Proteinogenic Side Chains – A Survey. *Chem. Biodivers.*, **2**, 401–20 (2005).
123. Patch, J. A. & Barron, A. E. Helical Peptoid Mimics of Magainin-2 Amide. *J. Am. Chem. Soc.*, **125**, 12092–93 (2003).
124. Violette, A. *et al.* Mimicking helical antibacterial peptides with nonpeptidic folding oligomers. *Chem. Biol.*, **13**, 531–38 (2006).
125. Claudon, P. *et al.* Consequences of isostructural main-chain modifications for the design of antimicrobial foldamers: helical mimics of host-defense peptides based on a heterogeneous amide/urea backbone. *Angew. Chem. Int. Ed Engl.*, **49**, 333–36 (2010).
126. Kritzer, J. A. *et al.* Relationship between Side Chain Structure and 14-Helix Stability of  $\beta$ 3-Peptides in Water. *J. Am. Chem. Soc.*, **127**, 167–78 (2005).
127. Bautista, A. D., Appelbaum, J. S., Craig, C. J., Michel, J. & Schepartz, A. Bridged  $\beta$ 3-Peptide Inhibitors of p53-hDM2 Complexation: Correlation between Affinity and Cell Permeability. *J. Am. Chem. Soc.*, **132**, 2904–6 (2010).
128. Sadowsky, J. D. *et al.* Chimeric ( $\alpha/\beta + \alpha$ )-Peptide Ligands for the BH3-Recognition Cleft of Bcl-xL: Critical Role of the Molecular Scaffold in Protein Surface Recognition. *J. Am. Chem. Soc.*, **127**, 11966–68 (2005).

129. Davie, E. A. C., Mennen, S. M., Xu, Y. & Miller, S. J. Asymmetric Catalysis Mediated by Synthetic Peptides. *Chem. Rev.*, **107**, 5759–5812 (2007).
130. Miller, S. J. In Search of Peptide-Based Catalysts for Asymmetric Organic Synthesis. *Acc. Chem. Res.*, **37**, 601–10 (2004).
131. Mogharabi, M., Rezaei, S. & Faramarzi, M. A. Peptide-Catalysis in Asymmetric Organic Synthesis. *Trends Pept. Protein Sci.*, **1**, 89–98 (2017).
132. Zozulia, O., Dolan, M. A. & Korendovych, I. V. Catalytic Peptide Assemblies. *Chem. Soc. Rev.*, (2018).
133. Coffey, P. E., Drauz, K. H., Roberts, S. M., Skidmore, J. & Smith, J. A. Beta-Peptides as Catalysts: Poly-Beta-Leucine as a Catalyst for the Juliá-Colonna Asymmetric Epoxidation of Enones. *Chem. Commun. Camb. Engl.*, 2330–31 (2001).
134. Müller, M. M., Windsor, M. A., Pomerantz, W. C., Gellman, S. H. & Hilvert, D. A Rationally Designed Aldolase Foldamer. *Angew. Chem.*, **121**, 940–43 (2009).
135. Wang, P. S. P., Nguyen, J. B. & Schepartz, A. Design and High-Resolution Structure of a  $\beta$ 3-Peptide Bundle Catalyst. *J. Am. Chem. Soc.*, **136**, 6810–13 (2014).
136. Wang, P. S. P. & Schepartz, A.  $\beta$ -Peptide bundles: Design. Build. Analyze. Biosynthesize. *Chem. Commun.*, **52**, 7420–32 (2016).
137. Girvin, Z. C. & Gellman, S. H. Exploration of Diverse Reactive Diad Geometries for Bifunctional Catalysis via Foldamer Backbone Variation. *J. Am. Chem. Soc.*, **140**, 12476–83 (2018).
138. Le Bailly, B. A. F., Byrne, L. & Clayden, J. Refoldable Foldamers: Global Conformational Switching by Deletion or Insertion of a Single Hydrogen Bond. *Angew. Chem. Int. Ed Engl.*, **55**, 2132–36 (2016).
139. Smaldone, R. A. & Moore, J. S. Reactive Sieving with Foldamers: Inspiration from Nature and Directions for the Future. *Chem. – Eur. J.*, **14**, 2650–57 (2008).
140. Mukherjee, S., Yang, J. W., Hoffmann, S. & List, B. Asymmetric Enamine Catalysis. *Chem. Rev.*, **107**, 5471–5569 (2007).
141. List, B. The Ying and Yang of Asymmetric Aminocatalysis. *Chem. Commun.*, 819–24 (2006).
142. Hajos, Z. G. & Parrish, D. R. *Ger. Pat. DE2102623*, (1971).
143. Eder, U., Sauer, G. & Wiechert, R. *Ger. Pat. DE2014757*, (1971).
144. Eder, U., Sauer, G. & Wiechert, R. New Type of Asymmetric Cyclization to Optically Active Steroid CD Partial Structures. *Angew. Chem. Int. Ed. Engl.*, **10**, 496–97 (1971).
145. Clemente, F. R. & Houk, K. N. Computational Evidence for the Enamine Mechanism of Intramolecular Aldol Reactions Catalyzed by Proline. *Angew. Chem. Int. Ed.*, **43**, 5766–68 (2004).
146. Hajos, Z. G. & Parrish, D. R. Asymmetric Synthesis of Bicyclic Intermediates of Natural Product Chemistry. *J. Org. Chem.*, **39**, 1615–21 (1974).
147. List, B., Hoang, L. & Martin, H. J. New mechanistic studies on the proline-catalyzed aldol reaction. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 5839–42 (2004).
148. Jung, M. E. A Review of Annulation. *Tetrahedron*, **32**, 3–31 (1976).
149. Agami, C. Mechanism of the Proline-Catalyzed Enantioselective Aldol Reaction. Recent Advances. *Bull. Société Chim. Fr.*, 499–507 (1988).

150. Hoang, L., Bahmanyar, S., Houk, K. N. & List, B. Kinetic and Stereochemical Evidence for the Involvement of Only One Proline Molecule in the Transition States of Proline-Catalyzed Intra- and Intermolecular Aldol Reactions. *J. Am. Chem. Soc.*, **125**, 16–17 (2003).
151. List, B., Lerner, R. A. & Barbas, C. F. Proline-Catalyzed Direct Asymmetric Aldol Reactions. *J. Am. Chem. Soc.*, **122**, 2395–96 (2000).
152. Dalako, P. I. & Moisan, L. In the Golden Age of Organocatalysis. *Angew. Chem. Int. Ed.*, **43**, 5138–75 (2004).
153. List, B., Peter Pojarliev, and & Castello, C. Proline-Catalyzed Asymmetric Aldol Reactions between Ketones and  $\alpha$ -Unsubstituted Aldehydes. *Org. Lett.*, **3**, 573–75 (2000).
154. Marquez, C. & Metzger, J. O. ESI-MS Study on the Aldol Reaction Catalyzed by L-Proline. *Chem. Commun.*, 1539 (2006).
155. Tanaka, F. & Barbas, C. F. Enamine Catalysis: Aldol and Mannich-Type Reactions. in *Enantioselective Organocatalysis*, (ed. Dalako, P. I.), 19–55 (Wiley-VCH Verlag GmbH & Co. KGaA, 2007).
156. Kleinman, E. F. The Bimolecular Aliphatic Mannich and Related Reactions. in *Compr. Org. Synth.*, **2**, 893–951 (Elsevier Ltd, 1991).
157. List, B. Asymmetric Aminocatalysis. *Synlett*, **2001**, 1675–86 (2001).
158. List, B. The Direct Catalytic Asymmetric Three-Component Mannich Reaction. *J. Am. Chem. Soc.*, **122**, 9336–37 (2000).
159. Córdoba, A. The Direct Catalytic Asymmetric Mannich Reaction. *Acc. Chem. Res.*, **37**, 102–12 (2004).
160. List, B., Pojarliev, P., Biller, W. T. & Martin, H. J. The Proline-Catalyzed Direct Asymmetric Three-Component Mannich Reaction: Scope, Optimization, and Application to the Highly Enantioselective Synthesis of 1,2-Amino Alcohols. *J. Am. Chem. Soc.*, **124**, 827–33 (2002).
161. List, B., Peter Pojarliev, and & Martin, H. J. Efficient Proline-Catalyzed Michael Additions of Unmodified Ketones to Nitro Olefins. *Org. Lett.*, **3**, 2423–25 (2001).
162. Seebach, D. & Golinski, J. Synthesis of Open-Chain 2,3-Disubstituted 4-nitroketones by Diastereoselective Michael-addition of (E)-Enamines to (E)-Nitroolefins. A topological rule for C, C-bond forming processes between prochiral centres. Preliminary communication. *Helv. Chim. Acta*, **64**, 1413–23 (1981).
163. Enders, D. & Seki, A. Proline-Catalyzed Enantioselective Michael Additions of Ketones to Nitrostyrene. *Synlett*, **2002**, 0026–0028 (2002).
164. Sharma, A. K. & Sunoj, R. B. Refined Transition-State Models for Proline-Catalyzed Asymmetric Michael Reactions under Basic and Base-Free Conditions. *J. Org. Chem.*, **77**, 10516–24 (2012).
165. Kofoed, J., Nielsen, J. & Reymond, J.-L. Discovery of new peptide-Based catalysts for the direct asymmetric aldol reaction. *Bioorg. Med. Chem. Lett.*, **13**, 2445–47 (2003).
166. Martin, T. J. & List, B. Mining Sequence Space for Asymmetric Aminocatalysis: N-Terminal Prolyl- Peptides Efficiently Catalyze Enantioselective Aldol and Michael Reactions. *Synlett*, **12**, 1091–1902 (2003).
167. Jarvo, E. R. & Miller, S. J. Amino Acids and Peptides as Asymmetric Organocatalysts. *Tetrahedron*, **58**, 2481–95 (2002).

168. Andrae, M. R. M. & Davis, A. P. Heterogeneous Catalysis of the Asymmetric Aldol reaction by Solid-Supported Proline-Terminated Peptides. *Tetrahedron Asymmetry*, **16**, 2487–92 (2005).
169. Luppi, G. *et al.* Dipeptide-Catalyzed Asymmetric Aldol Condensation of Acetone with (N-Alkylated) Isatins. *J. Org. Chem.*, **70**, 7418–21 (2005).
170. Tang, Z. *et al.* A Highly Efficient Organocatalyst for Direct Aldol Reactions of Ketones with Aldehydes. *J. Am. Chem. Soc.*, **127**, 9285–89 (2005).
171. Krattiger, P., Kovasy, R., Revell, J. D., Ivan, S. & Wennemers, H. Increased Structural Complexity Leads to Higher Activity: Peptides as Efficient and Versatile Catalysts for Asymmetric Aldol Reactions. *Org. Lett.*, **7**, 1101–3 (2005).
172. Revell, J. D. & Wennemers, H. Functional Group Requirements Within the Peptide H-Pro-Pro-Asp-NH<sub>2</sub> as a Catalyst for Aldol Reactions. *Tetrahedron*, **63**, 8420–24 (2007).
173. Wiesner, M., Revell, J. D. & Wennemers, H. Tripeptides as Efficient Asymmetric Catalysts for 1,4-Addition Reactions of Aldehydes to Nitroolefins—A Rational Approach. *Angew. Chem. Int. Ed.*, **47**, 1871–74 (2008).
174. Bächle, F., Duschmalé, J., Ebner, C., Pfaltz, A. & Wennemers, H. Organocatalytic Asymmetric Conjugate Addition of Aldehydes to Nitroolefins: Identification of Catalytic Intermediates and the Stereoselectivity-Determining Step by ESI-MS. *Angew. Chem. Int. Ed.*, **52**, 12619–23 (2013).
175. Duschmalé, J., Wiest, J., Wiesner, M. & Wennemers, H. Effects of internal and external carboxylic acids on the reaction pathway of organocatalytic 1,4-addition reactions between aldehydes and nitroolefins. *Chem. Sci.*, **4**, 1312–18 (2013).
176. Duschmalé, J. & Wennemers, H. Adapting to Substrate Challenges: Peptides as Catalysts for Conjugate Addition Reactions of Aldehydes to  $\alpha,\beta$ -Disubstituted Nitroolefins. *Chem. – Eur. J.*, **18**, 1111–20 (2012).
177. Kastl, R. & Wennemers, H. Peptide-Catalyzed Stereoselective Conjugate Addition Reactions Generating All-Carbon Quaternary Stereogenic Centers. *Angew. Chem. Int. Ed.*, **52**, 7228–32 (2013).
178. Patora-Komisarska, K., Benohoud, M., Ishikawa, H., Seebach, D. & Hayashi, Y. Organocatalyzed Michael Addition of Aldehydes to Nitro Alkenes - Generally Accepted Mechanism Revisited and Revised. *Helv. Chim. Acta*, **94**, 719–45 (2011).
179. Seebach, D. *et al.* 1,2-Oxazine N-Oxides as Catalyst Resting States in Michael Additions of Aldehydes to Nitro Olefins Organocatalyzed by  $\alpha,\alpha$ -Diphenylprolinol Trimethylsilyl Ether. *Helv. Chim. Acta*, **95**, 1064–78 (2012).
180. Burés, J., Armstrong, A. & Blackmond, D. G. Mechanistic Rationalization of Organocatalyzed Conjugate Addition of Linear Aldehydes to Nitro-olefins. *J. Am. Chem. Soc.*, **133**, 8822–25 (2011).
181. Burés, J., Armstrong, A. & Blackmond, D. G. Curtin–Hammett Paradigm for Stereocontrol in Organocatalysis by Diarylprolinol Ether Catalysts. *J. Am. Chem. Soc.*, **134**, 6741–50 (2012).
182. Burés, J., Armstrong, A. & Blackmond, D. G. Correction to Curtin–Hammett Paradigm for Stereocontrol in Organocatalysis by Diarylprolinol Ether Catalysts. *J. Am. Chem. Soc.*, **134**, 14264–14264 (2012).

183. Sahoo, G. *et al.* Dihydrooxazine Oxides as Key Intermediates in Organocatalytic Michael Additions of Aldehydes to Nitroalkenes. *Angew. Chem. Int. Ed.*, **51**, 13144–48 (2012).
184. Maillard, L. T., Park, H. S. & Kang, Y. K. Organocatalytic Asymmetric Addition of Aldehyde to Nitroolefin by H-d-Pro-Pro-Glu-NH<sub>2</sub>: A Mechanistic Study. *ACS Omega*, **4**, 8862–73 (2019).
185. Wiesner, M., Revell, J. D., Tonazzi, S. & Wennemers, H. Peptide Catalyzed Asymmetric Conjugate Addition Reactions of Aldehydes to Nitroethylene—A Convenient Entry into  $\gamma$ -Amino Acids. *J. Am. Chem. Soc.*, **130**, 5610–11 (2008).
186. Rigling, C. *et al.* Conformational Properties of a Peptidic Catalyst: Insights from NMR Spectroscopic Studies. *J. Am. Chem. Soc.*, **140**, 10829–38 (2018).
187. Schnitzer, T. & Wennemers, H. Influence of the Trans/Cis Conformer Ratio on the Stereoselectivity of Peptidic Catalysts. *J. Am. Chem. Soc.*, **139**, 15356–62 (2017).
188. Schnitzer, T. & Wennemers, H. Effect of  $\beta$ -3-Amino Acids on the Performance of the Peptidic Catalyst H-DPro-Pro-Glu-NH<sub>2</sub>. *Helv. Chim. Acta*, hlca.201900070 (2019).
189. Mathieu, L. *et al.* Cross-Claisen Condensation of N-Fmoc-Amino Acids – A Short Route to Heterocyclic  $\gamma$ -Amino Acids. *Eur. J. Org. Chem.*, 2262–70 (2015).
190. Yang, D., Yan, Y.-L. & Lui, B. Mild  $\alpha$ -Halogenation Reactions of 1,3-Dicarbonyl Compounds Catalyzed by Lewis Acids. *J. Org. Chem.*, **67**, 7429–31 (2002).
191. Kumaragurubaran, N., Juhl, K., Zhuang, W., Bøgevig, A. & Jørgensen, K. A. Direct 1-Proline-Catalyzed Asymmetric  $\alpha$ -Amination of Ketones. *J. Am. Chem. Soc.*, **124**, 6254–55 (2002).
192. Belot, S., Quintard, A., Krause, N. & Alexakis, A. Organocatalyzed Conjugate Addition of Carbonyl Compounds to Nitrodienes/Nitroenynes and Synthetic Applications. *Adv. Synth. Catal.*, **352**, 667–95 (2010).
193. Roca-Lopez, D. *et al.* Asymmetric Organocatalytic Synthesis of  $\gamma$ -Nitrocarbonyl Compounds through Michael and Domino Reactions. *Tetrahedron Asymmetry*, **21**, 2561–2601 (2010).
194. Marcia de Figueiredo, R. & Christmann, M. Organocatalytic Synthesis of Drugs and Bioactive Natural Products. *Eur. J. Org. Chem.*, **2007**, 2575–2600 (2007).
195. Xu, Y., Zou, W., Sundén, H., Ibrahim, I. & Córdova, A. Small Peptide-Catalyzed Enantioselective Addition of Ketones to Nitroolefins. *Adv. Synth. Catal.*, **348**, 418–24 (2006).
196. Bonnel, C. Oligopeptides Construits autour du  $\gamma$ -Aminoacide ATC: Syntheses, Analyses Structurales et Evaluation Biologique. *Thesis*, (2016).
197. Diehl, B. Principles in NMR Spectroscopy. in *NMR Spectrosc. Pharm. Anal.*, 1–41 (Elsevier, 2008).
198. Wagner, G., Kumar, A. & Wuthrich, K. Systematic Application of Two-Dimensional <sup>1</sup>H Nuclear-Magnetic-Resonance Techniques for Studies of Proteins. 2. Combined Use of Correlated Spectroscopy and Nuclear Overhauser Spectroscopy for Sequential Assignments of Backbone Resonances and Elucidation of Polypeptide Secondary Structures. *Eur. J. Biochem.*, **114**, 375–84 (1981).

199. Marion, D. & Wiithrich, K. Application of Phase Sensitive Two-Dimensional Correlated Spectroscopy (Cosy) for Measurements Of Lh-1h Spin-Spin Coupling Constants in Proteins. *Biochem. Biophys. Res. Commun.*, **113**, 967–74 (1983).
200. Cung, M. T., Marraud, M. & Neel, J. Experimental Calibration of a Karplus Relationship in Order to Study the Conformations of Peptides by Nuclear Magnetic Resonance. *Macromolecules*, **7**, 606–13 (1974).
201. Ramachandran, G. N., Chandrasekaran, R. & Kopple, K. D. Variation of the NH-CαH Coupling Constant with Dihedral Angle in the NMR Spectra of Peptides. *Biopolymers*, **10**, 2113–31 (1971).
202. Bystrov, V. F. Spin—Spin Coupling and the Conformational States of Peptide Systems. *Prog. Nucl. Magn. Reson. Spectrosc.*, **10**, 41–82 (1976).
203. Pardi, A., Billeter, M. & Wthrich, K. Calibration of the Angular Dependence of the Amide Proton-Cα Proton Coupling Constants, 3JHNα, in a Globular Protein. *J. Mol. Biol.*, **180**, 741–51 (1984).
204. Byler, D. M. & Susi, H. Examination of the Secondary Structure of Proteins by Deconvolved FTIR Spectra. *Biopolymers*, **25**, 469–87 (1986).
205. Haris, P. I. & Chapman, D. The Conformational Analysis of Peptides Using Fourier Transform IR Spectroscopy. *Biopolymers*, **37**, 251–63 (1995).
206. Surewicz, W. K. & Mantsch, H. H. New Insight into Protein Secondary Structure from Resolution-Enhanced Infrared Spectra. *Biochim. Biophys. Acta*, **952**, 115–30 (1988).
207. Vass, E., Hollósi, M., Besson, F. & Buchet, R. Vibrational Spectroscopic Detection of Beta- and Gamma-Turns in Synthetic and Natural Peptides and Proteins. *Chem. Rev.*, **103**, 1917–54 (2003).
208. Braiman, M. S. & Rothschild, K. J. Fourier Transform Infrared Techniques for Probing Membrane Protein Structure. *Annu. Rev. Biophys. Biophys. Chem.*, **17**, 541–70 (1988).
209. Aguesseau-Kondrotas, J. *et al.* Prospect of Thiazole-Based γ-Peptide Foldamers in Enamine Catalysis – Exploration of the Nitro-Michael Addition. *Chem. – Eur. J.*, **25**, 7396–7401 (2019).
210. Max, J.-J. & Chapados, C. Infrared Spectroscopy of Aqueous Carboxylic Acids: Comparison between Different Acids and Their Salts. *J. Phys. Chem. A*, **108**, 3324–37 (2004).
211. Roux, S. *et al.* Elimination and exchange of trifluoroacetate counter-ion from cationic peptides: a critical evaluation of different approaches. *J. Pept. Sci.*, **14**, 354–59 (2008).
212. Case, D. A. *et al.* AMBER 16. (2016).
213. Wüthrich, K. *NMR of Proteins and Nucleic acids*. (1986).
214. Clore, G. & Gronenborn, A. Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. *Science*, **252**, 1390–99 (1991).
215. Gronenborn, A. M. & Clore, G. M. Where is NMR taking us? *Proteins Struct. Funct. Genet.*, **19**, 273–76 (1994).
216. Johnson, B. A. & Blevins, R. A. NMR View: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–14 (1994).
217. Tsui, V. & Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers*, **56**, 275–91 (2000).

- 
218. Koradi, R., Billeter, M. & Wüthrich, K. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55 (1996).
219. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.*, **25**, 1157–74 (2004).
220. Wang, J., Cieplak, P. & Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.*, **21**, 26
221. Frisch, M. J. *et al.* Gaussian 09. *Gaussian Inc Wallingford CT*, (2016).
222. Zhao, Y. & Truhlar, D. G. Applications and Validations of the Minnesota Density Functionals. *Chem. Phys. Lett.*, **502**, 1–13 (2011).
223. Singh, U. C. & Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.*, **5**, 129–45 (1984).

## ELABORATION DE CATALYSEURS BIO-INSPIRES CONÇUS AUTOUR D'UNE ARCHITECTURE

### $\gamma$ -PEPTIDIQUE AUTO STRUCTUREE

**Résumé :** Les travaux décrits dans ce manuscrit concernent la synthèse d'oligomères de  $\gamma$ -amino acides hétérocycliques contraints, appelés ATCs (acides 4-Amino-(méthyl)-1,3-Thiazole-5-Carboxyliques), leur application en catalyse énamine et leur étude structurale. Les monomères d'ATC sont construits autour d'un noyau thiazole inséré entre les carbones  $C_{\alpha}$ - $C_{\beta}$ , permettant de limiter la valeur de l'angle dièdre  $\zeta$  à  $0^{\circ}$ . La présence de deux points de substitution, sur le carbone  $\gamma$  asymétrique et en position 2 du noyau aromatique, permet une large diversification structurale des ATCs. Ainsi, plusieurs séries d'oligomères ont été synthétisées par couplages peptidiques sur support solide. Une étude structurale de ces oligomères par RMN, IR-TF, cristallographie RX et dichroïsme circulaire a démontré qu'ils adoptaient une structure en hélice  $C_9$ , résultant d'un réseau de liaisons hydrogène de type  $CO_{i-1}---NH_{i+2}$  s'établissant tout au long de la séquence. L'objectif du projet présenté ici vise à étudier l'impact de la conformation des architectures développées, à la fois sur la sélectivité et sur l'induction asymétrique dans la réaction de nitro-Michael pour trois réactifs différents. Le dernier axe de ce travail a été de développer une méthode de modélisation sous contraintes RMN spécifique à la génération de modèles tridimensionnels d'oligomères d'ATCs.

---

### DESIGN OF BIO-INSPIRED CATALYSTS BASED ON A $\gamma$ -PEPTIDE FOLDAMER ARCHITECTURE

**Summary:** The work described in this manuscript is devoted to the synthesis of heterocyclic constrained  $\gamma$ -amino acids, named ATCs (4-Amino-(methyl)-1,3-Thiazole-5-Carboxylic acids), their application in enamine catalysis and their structural study. ATC monomers are built around a thiazole ring providing a conformational limitation around the  $C_{\alpha}$  and  $C_{\beta}$  at  $0^{\circ}$ . The presence of two diversification points both on the  $\gamma$  asymmetric carbon and on the position 2 of the aromatic ring, allows a large structural diversification of the ATCs. Therefore, several oligomers were synthesized using solid phase peptide synthesis. A structural study of these oligomers, employing NMR, FTIR, circular dichroism and crystallography RX, demonstrated that they adopt a  $C_9$ -right-handed helix stabilized by a hydrogen bond pattern between  $CO_{i-1}---NH_{i+2}$  along the helix. The objective of the project presented in this manuscript was the design and the structural characterization of molecular edifices with predictable folding properties and the systematic study of structure-function relationships in the nitro-Michael addition reaction, for three different substrates. Eventually, the last part of this work focused on the development of a new methodology, specific to ATC-oligomers, to perform 3D-modelling studies using NMR refinement.

---

**Keywords:** Foldamers, constrained  $\gamma$ -amino acids, organocatalysis, peptides, structural studies