Synthetic Analogues of Intrinsically Disordered Polypeptides : Copolymerization, Sequence-Control and Stimuli-Responsiveness

Mostafa Badreldin

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

SPÉCIALITÉ POLYMÈRES

Par Mostafa BADRELDIN

ANALOGUES SYNTHÉTIQUES DE POLYPEPTIDES INTRINSÈQUEMENT DÉSORDONNÉS

Copolymérisation, Contrôle de Séquence et Réponse aux Stimuli

Sous la direction du Dr. Colin BONDUELLE, Université de Bordeaux

Soutenue le 25 octobre 2023

Membres du jury :

Mme Blanca MARTIN-VACA	Professeure, Université Paul Sabatier	Rapporteur
M. Matthew GIBSON	Professeur, University of Manchester	Rapporteur
Mme Mona SEMSARILAR	Directrice de recherche, Université de Montpellier	Examinateur
Mme Brigitte BIBAL	Professeure des universités, Université de Bordeaux	Examinateur
Mme Sophie GUILLAUME	Directrice de recherche, Université de Rennes	Examinateur
M. Simon HARRISSON	Directeur de recherche, Université de Bordeaux	Invité
M. Colin BONDUELLE	Chargé de recherche, Université de Bordeaux	Directeur

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

SPÉCIALITÉ POLYMÈRES

Par Mostafa BADRELDIN

ANALOGUES SYNTHÉTIQUES DE POLYPEPTIDES INTRINSÈQUEMENT DÉSORDONNÉS

Copolymérisation, Contrôle de Séquence et Réponse aux Stimuli

Sous la direction du Dr. Colin BONDUELLE, Université de Bordeaux

Soutenue le 25 octobre 2023

Membres du jury :

Mme Blanca MARTIN-VACA	Professeure, Université Paul Sabatier	Rapporteur
M. Matthew GIBSON	Professeur, University of Manchester	Rapporteur
Mme Mona SEMSARILAR	Directrice de recherche, Université de Montpellier	Examinateur
Mme Brigitte BIBAL	Professeure des universités, Université de Bordeaux	Examinateur
Mme Sophie GUILLAUME	Directrice de recherche, Université de Rennes	Examinateur
M. Simon HARRISSON	Directeur de recherche, Université de Bordeaux	Invité
M. Colin BONDUELLE	Chargé de recherche, Université de Bordeaux	Directeur

Titre : Analogues Synthétiques de Polypeptides Intrinsèquement Désordonnés : Copolymérisation, Contrôle de Séquence et Réponse aux Stimuli.

Résumé :

Les protéines sont une classe très importante de biomacromoléules naturelles. Elles remplissent de nombreuses fonctions dans les systèmes vivants mais sont composés de seulement 20 types d'acides aminés. Mimer la structure chimique des protéines en produisant des polypeptides synthétiques est d'un grand intérêt pour produire des matériaux fonctionnels, biodégradables et biocompatibles. Le moyen le plus économique et le plus efficace pour produire des polymères à base d'acides aminés est la polymérisation par ouverture de cycle des N-carboxyanhydrides (NCA) dérivés de ces acides aminés. L'objectif de cette thèse de doctorat a été d'imiter la structure et les propriétés de protéines intrinsèquement désordonnées en utilisant cette polymérisation. La séquence des acides aminés dans ces protéines est relativement simple si on la compare à d'autres régions du protéome et se caractérise par des séquences répétitives enrichies en glycine et en proline. En utilisant les monomères NCA dérivés de ces deux acides aminés, ce travail de doctorat développe la préparation de copolymères statistiques par copolymérisation pour produire des polypeptides avec une séquence monomère contrôlée stochastiquement. Un tel contrôle permet d'induire de manière statistique certaines séquences (AA, AB, BB) dans la distribution des monomères le long de la macromolécule. Les études cinétiques d'un modèle de copolymérisation comprenant de la glycine et un glutamate protégé ont d'abord permis de calculer leurs rapports de réactivité respectifs. Le contrôle des différentes conditions de réaction (solvant, température, catalyse) a permis d'orienter stochastiquement le système de copolymérisation vers différentes séquences en monomères. L'influence de cette séquence sur les propriétés macromoléculaires, notamment la solubilité et la structure secondaire, a ensuite été élucidé. Dans un deuxième temps, l'approche biomimétique basée sur la chimie des polymères a été utilisé pour reproduire la séquence de protéines thermosensibles riches en glycine et en proline. En biologie, les polypeptides de type élastine (ELP), sont des protéines thermosensibles qui présentent un caractère thermosensible proche de la température corporelle (37 °C) et sont donc utilisés dans de nombreuses applications biomédicales. Cette classe de protéine a donc été choisie comme modèle de séquence riche en glycine et en proline à mimer. Des homopolymères de polyproline (PLP) de masses moléculaires élevés ont d'abord été préparés en milieu aqueux et étudiés pour comprendre leur comportement thermosensible peu exploré et révélant une hystérèse exceptionnellement large. Avec ce polymère, la structure, la chiralité et la structuration secondaire sont des paramètres importants qui expliquent un effet mémoire qui a permis de produire des formulations originales telles que des sondes de température ou des hydrogels programmables. Dans des réactions de copolymérisation en milieu aqueux, introduire de la proline avec de la glycine permet d'affiner son comportement thermosensible. En réalisant la polymérisation dans l'eau, cette copolymérisation produit des macromolécules à gradient élevé dont on peut ajuster le comportement thermosensible. Ces copolymères s'auto-assemblent spontanément en nanoparticules thermosensibles. Des réactions de polymérisation

plus complexes, des terpolymérisations entre la glycine, la proline et la valine ont ensuite été explorées pour mieux imiter le comportement des ELP, mais cette réaction a montré certaines limites et des problèmes de solubilités. Finalement, d'autres approches de polymérisation ont été explorées, y compris l'organocatalyse et la catalyse organométallique. Ces systèmes permettront d'obtenir des copolymères contenant de la proline avec une thermosensibilité plus modulable.

Mots clés :

Polymérisation à ouverture de cycle, Monomères N-carboxyanhydride, Contrôle de séquence en copolymérisation, Polypeptides thermosensibles, Copolymères biomimétiques

---

Title: Synthetic Analogues of Intrinsically Disordered Polypeptides: Copolymerization, Sequence-Control and Stimuli-Responsiveness.

Abstract:

Proteins are the most diverse class of natural biomacromolecules, and they carry out many different roles in living biology. Their diverse functional and structural portfolio is made up of only 20 different types of monomers, naturally occurring amino acids. Mimicking the chemical structure of proteins by producing synthetic poly(amino-acids) is thus of great interest to produce highly functional, biodegradable and biocompatible materials. The most economical and efficient way to produce high molecular weight amino-acid based polymers is through the ring opening polymerization (ROP) of amino acid N-carboxyanhydrides (NCAs). This polymerization is very similar to classical ROP and can be easily scaled up in synthetic chemical facilities. The goal of this PhD is to be able to mimic the structure and properties of intrinsically disordered proteins (IDPs). The sequence of amino acids in IDPs is relatively simple compared to other regions of the proteome and is characterized by highly repeating patterns. These simple protein structures are considered an easy first step to mimic proteins through polymerization procedures. The primary sequence of these proteins must first be replicated in a polymerization process. This PhD work presents the preparation of statistical copolymers via the ring opening copolymerization (ROCOP) of NCAs to produce polypeptide polymers with a stochastically controlled monomer sequence. This technique aims to enrich the frequency with which desirable monomer sequences appear in the overall sequence distribution. Kinetic studies of a model copolymerization system comprising glycine (a disorder-inducing monomer) and a protected glutamate were used to calculate their respective reactivity ratios. Reaction conditions (solvent, temperature, catalysis) were varied to enrich the resulting copolymers with different monomer sequences and composition profiles. Their consequent effect on the macromolecular properties, notably solubility and secondary structure, was further analysed. Stimuli-responsive polypeptides were then studied, focusing on thermoresponsive polymers. Like glycine, proline units are common in IDPs. The synthesis of high molecular weight

and controlled PLPs via the recently developed water-assisted mechanism led to the study of the little-explored thermoresponsive behaviour of these polymers, which exhibit an unusually broad hysteresis. The structure, chirality and secondary structuring were found to be the underlying reason for this property. This behaviour was exploited to produce simple memory driven formulations such as temperature probes or programable hydrogels. Elastin-like polypeptide (ELP), a thermoresponsive IDP that contains both glycine and proline units, has a lower critical solution temperature (LCST) close to body temperature (37°C) and is used in many biomedical applications. This class of IDP is extensively studied in the laboratory and was chosen as a model IDP to be mimicked. Using the knowledge accumulated from previous studies, the copolymerization of proline with glycine to fine tune the thermoresponsive behaviour was first explored. It was found that the water-assisted ROP mechanism produced copolymers with a strong gradient that showed similar LCST behaviour to polyproline. Lower molecular weight copolymers showed block-like structures and formed thermoresponsive nanoparticles. Terpolymerizations of glycine, proline and valine were further explored to mimic the behaviour of ELP but the use of water-assisted ROP limited the range of accessible compositions. Finally, other ROP mechanisms were explored including organocatalysis and organometallic catalysis. These systems will allow to obtain proline containing copolymers with highly tuneable thermoresponsivness.

Keywords:

Ring-opening polymerization, N-carboxyanhydride monomers, Biomimetic copolymers, Sequence control in copolymerization, Thermoresponsive polypeptides

---

Unité de recherche

Laboratoire de Chimie des Polymères Organiques, UMR 5629

Ecole Nationale Supérieure de Matériaux, d'Agroalimentaire et de Chimie, ENSMAC (ex ENSCBP)

16 Avenue Pey-Berland

33607 Pessac Cedex

France

Acknowledgements

I would like to begin this manuscript by first expressing my gratitude to the members of my defence jury for accepting to take part of the final steps of my academic journey.

I would also like to acknowledge the financial support given to my research starting with the MESRI scholarship who is funding this doctoral program through the University of Bordeaux. Acknowledgments are also due to EDSC, TGIR-RMN platform and Prof. Deming. I would also like to extend my acknowledgments to Prof. Lecommandoux (Seb) the lab director. I appreciated our work together on different projects and I certainly enjoyed our discussion on various topics, and I appreciated your commitment to making this lab a decent working space for everyone.

I am profoundly thankful to my PhD advisors Colin and Simon (I can write it here if I want).

Colin, je veux te remercier pour m'avoir recruté sur ce projet, pour ton soutien, tes conseils et ta bienveillance. Travailler avec toi était un plaisir, tant pour ta passion et ta persévérance que pour la confiance que tu me donnes et que tu mets en moi. Ce projet, nous l'avons bâti ensemble, et je n'aurais jamais imaginé que nous arriverions à cela au début. C'est en grande partie grâce à ta croyance dans la recherche. Je sais que tu es déçu que je ne parle pas espagnol, mais Gracias quand même.

Simon, it has been an honor working alongside you during these three years. From the very beginning, even before embarking officially on the project, you have always been eager to teach and help me and I really appreciate our discussions. You SHOWCASE a huge amount scepticism that is much needed in the scientific word while always having new ideas and solutions. I am extremely grateful we could share this journey together and to the results our collaborations have given. *No proofreading was needed for this section of this manuscript.*

I would like to extend my appreciation to all my collaborators: Mme. Grélard, Mme. Morvan and Mme. Minder at the IECB; Dr. Perrot, Mme. Daou and Prof. Giuseppone at the ICS; Mme. Absalon at the CESAMO platform. I would also like to thank Prof. Guillaume and Ali at the ISCR for the fruitful collaboration and the nice discussions we had, both scientifical and personal. I really believe in this project and hope it will bear fruit one day. I am also grateful for the collaboration with Prof. Ballet and Dr. Lamouroux. Thanks are also due to my collaborators in Los Angeles, starting with Prof. Deming who supported my short and interesting stay in his lab. I wish all the best to all the people who helped me there: Vivian, Casey, Lily, Aryan, Matt, Joseph, Roberto and Katie. I wish you all the best in the future and hope to meet you again.

To Dr. Salas (Pedrito mi amigo), I can’t thank you enough for all you have done to me during these last 3 years. We have shared so many moments together starting from breaking glassware in the lab to surfing on the waves of Santa Monica. I really appreciated learning from you and working with you and I am grateful to the friendship that ensued. I am sure you’ll do great things in the future and can’t wait to read about them. All my best to Angie and to el gatito.

Merci Rosanna pour ton travail. Les 6 mois de stage pendant lesquels je t'ai encadrée ont été d'une immense valeur pour ma recherche et ma thèse. C'était aussi un plaisir (à part quand tu as failli mourir à cause d'un moustique), vu comment tu es efficace et professionnelle. Je te souhaite une bonne continuation et je suis sûr que ton futur sera très brillant. Je regrette encore que tu n'aies pas suivi mon conseil de fuir, donc je te souhaite surtout bon courage pour ta thèse.

Sifan (Sisi) thank you (Xièxiè) for everything it has been a pleasure and I am sure we'll meet again in the future. I am still trying to learn Chinese maybe I will be able to say polyamino acid one day thanks to you. Wish you all the best.

Thank you (soon to be) Dr. Kazarian (Dianita) for everything we shared these past three years. It was a pleasure to have you as a colleague (although you are annoying sometimes) and hopefully our friendship will continue for many years to come.

Je voudrais aussi remercier les chercheurs de notre équipe (l'équipe 3) pour leur aide et leurs conseils. Merci Elisabeth pour tes conseils sur les ELP et pour les moments sympas passés avec toi. Merci Angela pour nos discussions, tes conseils sur le CD et les souvenirs d'organisation du BPC. Merci à Olivier, Bertrand, Jeff et Christophe pour tous vos conseils et votre aide dans et en dehors du laboratoire. C'était un plaisir de travailler avec vous tous et d'apprendre de vos connaissances.

Thanks to all my friends/labmates from team 3 past and present for all the great moments, for your kindness and for the atmosphere in the lab: Antoine, Ségolène, Hannah, Anouk, Leslie, Clémence, Matthieu, Fanny, Megi, Elise, Eloïse, Manon, Yue, Marie, Tingting, Dongxhu, Florian, Tim, Vusalita, Romane, Vaïana, Pierre, Léa, Hang, Julien, Alban, Justine, Barbara, Yupei, Valentin, Sarah, Julien, Nadia, Flavia, Johanna and all the interns that passed through here.

Thanks also to students from other teams, the list is long, and I will not be able to thank each and every one you here. But be assured that I will not forget the nice moments that I shared with you all of you. Thanks particularly to Prof. Samir, Chloé, Peter, Corentin, Dijwar and Virgil(e) for all your help and your friendship.

Je remercie aussi tout le personnel technique du laboratoire. Merci à Manu (pas très fort au tennis), Cédrick et Frédérique. Merci particulièrement à Anne-Laure, Loïc, Oriane, Léna et Nina pour leur aide et leur soutien sur les RMN, autant de temps perdu mais autant d'expérience gagnée. Merci à Amélie pour son aide sur la SEC, j'espère que tout se passe bien pour toi. Merci aussi à Sylvain (à qui je souhaite surtout du courage) pour son aide, sa bienveillance, son soutien et son dévouement. Finalement, je remercie l'équipe administrative du laboratoire ainsi que le personnel de l'école. Sans leurs efforts, nous n'aurions pas pu travailler dans de bonnes conditions.

Chimbonios we Sandra men gheirko makontesh habtedy el meshwar el tawil da. Thank you, Christopher, Rimbo and Shiva, for also being here from the beginning. A7la messa 3ala a7la shabab Hisho wabou el Joe we Omar beh (we Béa) mestaneyenko dayman fi Bordeaux we fi ay 7etta. Ayman basha shed 7elak we rabena yewafa2ak. Izabelly basha we Titi basha wogodkom fi bordeaux menawar el denya. Francis et Hugo merci

pour vous et j’espère que vous allez comprendre quelque chose à cette thèse vous resterez à toujours des ingénieurs chimistes. Kalvin, bon courage mon pote, hate d’assister à ta thèse. To my californian friends Memma el 3azim, Fady and Monica gayelko tany akid, and Panos hope to see you soon my friend. Merci à mes amis lyonnais (69) Elliott, Laura, Laetitia, Charly, Valentin et Elise. Merci aussi à mes amis bordelais Arthur, Simon, Eloïse, Claudia, Ewen, Zuzanna et Hugo. El shabab kolo yasta7ek ta7eya sari3a gamila talik bih.

Merci à ma famille et à ma belle-famille pour leur soutien. Merci à Ali et Julie et bon courage à vous futurs docteurs. Merci à Timo, Judo, Aya et Nouna. Une pensée aux personnes qui ont toujours cru en moi et qui ne sont plus parmi nous aujourd’hui mes grand-parents, Helmy, الله يرحمهم.

To my parents who have supported me during my academic career culminating in this moment, I would like to say thank you and I love you for all your support and for being my inspiration.

Akhiran wa lays akheran (bal 7atta awalan) a7eb ashkor Elo meraty we sharikat 7ayaty 3ala wo2ofha ma3aya we mossandetha leya fi 7ayaty we shoghly. Esmek yasta7ek wogoodoh gamb esmy fi awel saf7a. Merci, thank you, shokran, ba7ebek.

“Preserving knowledge is easy. Transferring knowledge is also easy. But making new knowledge is neither easy nor profitable in the short term. Fundamental research proves profitable in the long run, and, as importantly, it is a force that enriches the culture of any society with reason and basic truth.”[1]

(1) Zewail, A. *Nature* **2010**, *468* (7322), 347–347.

و الحمدُ للَّهِ الَّذي بنعمتِه تتمُّ الصَّالحاتُ

Table of Contents

Objectif de la thèse

Les protéines sont des macromolécules biologiques (biomacromolecules) très diversifiées aussi bien au niveau de leurs structures que de leurs fonctions, bien qu'elles soient composées de seulement 20 types différents de monomères, les acides aminés. C'est en fait la séquence et la distribution spatiale de ces mêmes acides aminés qui permet de générer la grande diversité des structures protéiques dans les systèmes vivants et de promouvoir au mieux leurs propriétés biologiques. Dans le cadre d'une approche de chimie bio-inspirée, mimer la structure chimique des protéines à travers la synthèse de poly(acides aminés) est devenu d'un grand intérêt pour produire des matériaux hautement fonctionnels, biodégradables et biocompatibles. Dans ce travail, nous nous sommes intéressés aux protéines intrinsèquement désordonnées (IDP).[1,2] Les IDP sont caractérisées par des séquences simples,[3,4] une forte solubilité et se comportent comme des chaines polymères solvatées.[5] De plus, elles jouent un rôle biologique important dans la reconnaissance moléculaire et l'auto-assemblage grâce à leur propriétés physico-chimiques uniques.[5–7] Les IDP sont d'excellents modèles pour développer des analogues polymères synthétiques, l'objectif sous-jacent à l'ensemble de ces travaux de thèse.

Cette thèse de doctorat se concentre sur la synthèse de polymères appelés polypeptides (dont les unités répétitives sont des acides α-aminés) en utilisant la polymérisation par ouverture de cycle (ROP) des *N*-carboxyanhydrides (NCA) (**Figure 1**).

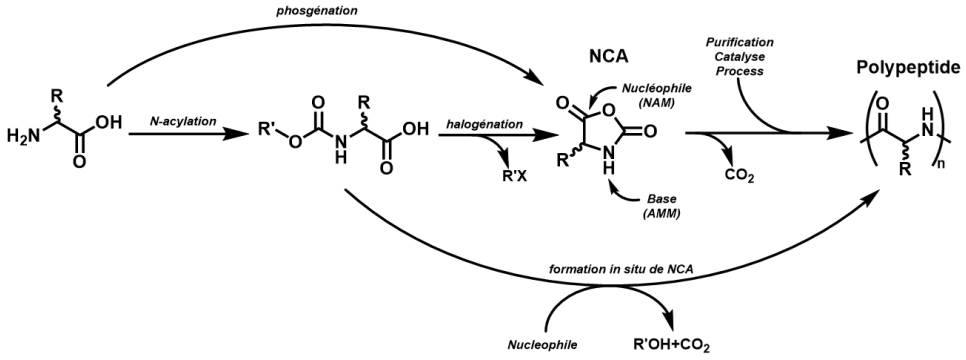


***Figure 1**: Formation de NCA à partir d'acide aminés et leur dérivés. La ROP des NCA pour produire des polypeptides est au centre du développement récent en se focalisant sur les différentes méthodes pour optimiser cette polymérisation.*

Une étude bibliographique sur les progrès récents dans les méthodologies de synthèse des NCA et de la ROP (**Figure 1**) constitue l'état de l'art dans le chapitre **1** de cette thèse. Cette méthode de polymérisation est très similaire à d'autres ROP (lactide, oxazolines, etc.) et peut être facilement mise en œuvre à grande échelle, ouvrant ainsi des perspectives d'application dans le domaine des matériaux, des perspectives d'applications que n'offrent pas souvent les protéines.[8]

Les polymères peptidiques (polypeptides dans **Figure 1**) ont longtemps été étudiés afin de reproduire certaines caractéristiques structurales modèles des protéines naturelles.[9] Par exemple, ces polymères sont considérés comme des analogues synthétiques simplifiés de la structuration des protéines puisqu'ils peuvent adopter des structures secondaires typiques (α-hélice, β-feuille…).[10]

Le chapitre **2** se focalise sur une copolymérisation particulière, la ROP simultanée de deux types de monomères NCA. En s'inspirant du fait que la structure des protéines dépend de l'arrangement des unités monomères le long de la chaîne peptidique (appelé séquence primaire), l'objectif de ce chapitre était de contrôler la distribution des deux types de monomères (en bloc, en gradient, alternées…) au sein des chaînes de copolymères. En utilisant un système binaire pour démontrer comment l'insertion d'un type d'unités peut perturber l'hélicité d'un bloc peptidique en tant que modèle simplifié d'une IDP. Plus précisément, un polypeptide modèle qui se structure en hélice a été étudié, le poly(γ-benzyl-L-glutamate) (PBLG) issu du NCA de BLG (**Figure 2**).[11] Afin de désordonnér ce polymère, nous l'avons copolymérisé avec le NCA de la glycine (**Figure 2**), un résidu qui crée des barrières stériques à l'apparition de la structuration secondaire dans les séquences d'IDP.[12] Pour cela, il a été essentiel de comprendre comment les deux monomères réagissent et s'insèrent dans une chaîne polymère grâce à une étude cinétique du système de copolymérisation.

Le contrôle statistique de la distribution des deux monomères le long de la chaine a été étudié dans différentes conditions de réaction (solvant, température, ratio de mélange et catalyse). Notre travail a permis d'établir une relation entre la distribution de la glycine et la conformation de la chaîne polypeptidique (**Figure 2**). Cette étude a donc permis de démontrer qu'il était possible d'obtenir des "copolymères à base de PBLG intrinsèquement désordonnés" grâce à la copolymérisation contrôlée de glycine pour former des copolymères à gradient.

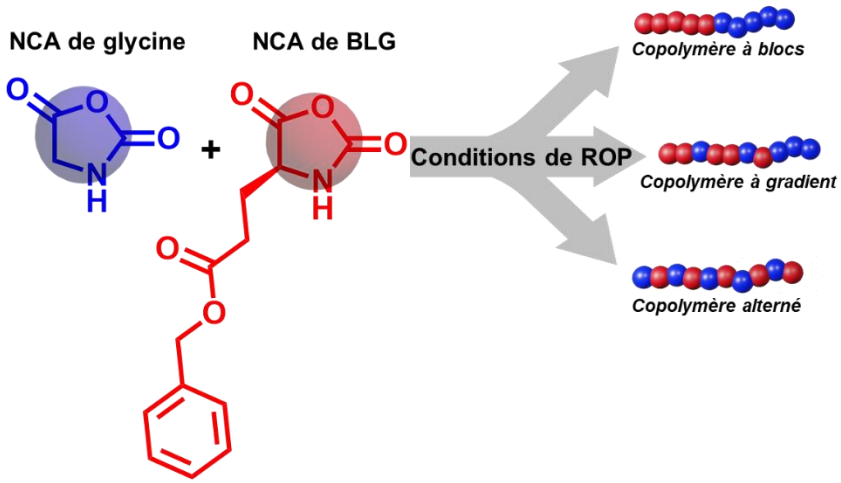


***Figure 2**: Copolymérisation des NCA de glycine et BLG pour obtenir des copolymères à microstructures différentes en contrôlant les conditions de la réaction. Notre étude a démontré que le solvant, la température et le mécanisme de réaction peuvent orienter cette copolymérisation vers une ou autre microstructure.*

Dans le chapitre **3**, nous nous sommes intéressés aux IDPs dérivés de l'élastine, qui sont sensibles à la température et qui contiennent un nombre significatif de résidus proline dans leur séquence. Cet acide aminé que l'on retrouve dans beaucoup d'IDPs n'a pas été vraiment étudié en tant qu'homopolymère notamment vis à vis de ses propriétés thermosensibles. Dans le cas de la ROP des NCAs, des travaux novateurs concernant la polymérisation[13,14] nous ont permis d'étudier les propriétés thermosensibles de polyproline (PLP) à haute masse moléculaire.[15] Ces polymères présentent une température de solution critique (LCST) associée à une hystérèse exceptionnellement large. Les chaines solubles à faible température s'agrègent au-delà de leur point trouble (T_{CP}) et ne se resolubilisent qu'à une température de resolubilization nettement plus faible (T_{CL}) (**Figure 3**).

Une étude approfondie des conditions influençant cette propriété telle que la concentration ou la force ionique du milieu a été réalisée. La compréhension des effets moléculaires sous-jacents a démontré que le comportement LCST des PLP est fortement lié à leur conformation (structure secondaire) crée par la chiralité et les liaisons peptidiques proline-proline.[16] L'exploitation de l'hystérèse a aussi été explorée pour produire des matériaux reprogrammables avec effet mémoire (**Figure 3**).

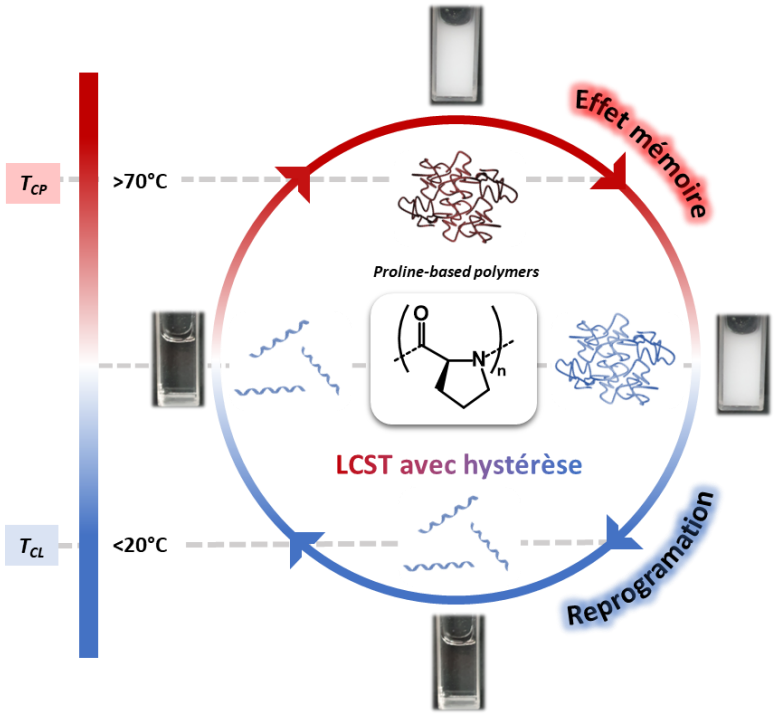


***Figure 3**: Résumé graphique présentant l'effet de mémoire dans les formulations de polyproline à comportement LCST hystérétique et leur reprogrammation*

En allant plus loin dans notre approche bio-inspirée, nous avons considéré que les polypeptides de type élastine (ELP) sont une classe d'IDP thermosensible qui contiennent à la fois de la glycine et des unités de proline.[17] Leur comportement LCST peut être très finement modulé par génie génétique grâce au choix et la séquence de ces deux types d'acides aminés.[18] Le chapitre **4** se focalise donc sur la copolymérisation de la proline afin de moduler le comportement LCST étudié dans le chapitre **3** et pour se rapprocher de la thermosensibilité des ELP (**Figure 4**).

Dans un premier temps, des essais de copolymérisation sur des systèmes binaires ont été abordés. Les copolymères que nous avons pu produire ont révélé un caractère à fort gradient similaire à ceux discutés dans le chapitre **2**. Ceci a été attribué à la forte réactivité de la proline par rapport aux autre comonomères dans diverses conditions de ROP. Ces copolymères (Pro-Gly et Pro-Val) montrent un comportement LCST avec une hystérèse similaire à la PLP. En modulant les masses molaires, il a même été possible de produire des nanoparticules thermosensibles sans hystérèse. Dans un deuxième temps, des terpolymérisations de la glycine, de la proline et de la valine ont aussi été explorées pour mieux reproduire la séquence des ELPs (**Figure 4**). L'étude de ces systèmes dans les conditions assistées par l'eau et avec des catalyseurs organométalliques ont permis d'obtenir des structures polymériques contrôlées. Ces

systèmes permettront d'obtenir des copolymères contenant de la proline et possédant une thermosensibilité facilement modulable.

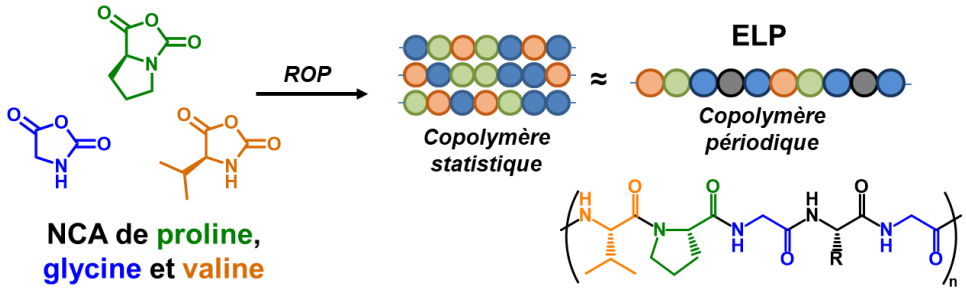


***Figure 4**: La copolymérisation des NCA de proline, glycine et valine pour produire des copolymères statistiques. Ces travaux ont pour objectifs de reproduire statistiquement la séquence périodique des ELP et donc leurs propriétés.*

Les connaissances acquises sur la catalyse et la copolymérisation de la ROP des NCAs ont aussi permis d'ouvrir d'autres perspectives intéressantes. Par exemple, l'utilisation de catalyseurs organométalliques à base d'yttrium pour produire des polymères avec une tacticité contrôlée a été exploré.[19] Ces complexes ont permis la ROP des NCAs pour produire des polypeptides avec les structures souhaitées même si l'efficacité de la stéréosélectivité lors de la polymérisation reste limitée. Finalement, la copolymérisation des NCA avec d'autre monomère cyclique (ROCOP), tel des esters cycliques, a été brièvement étudié afin de produire de nouvelles structures polymériques appelées poly(ester-amide)s.

Aim of the PhD

Proteins are a highly diverse class of biological macromolecules, both in terms of their structures and functions, even though composed of only 20 different types of monomers, known as amino acids. The sequence and spatial distribution of these amino acids generate the vast diversity of protein structures in living systems and determine their biological properties. Mimicking the chemical structure of proteins through the synthesis of poly(amino-acids) is of great interest for producing highly functional, biodegradable, and biocompatible materials. In this work, we focused on intrinsically disordered proteins (IDPs).[1,2] IDPs are characterized by simple sequences[3,4] and high solubility, and behave like solvated polymer chains.[5] They play an important biological role in molecular recognition and self-assembly due to their unique physicochemical properties.[5–7] The objective of this thesis is to develop synthetic polymeric analogues of IDPs.

This doctoral work focuses on the synthesis of polymers known as polypeptides (with repeating units being α-amino acids) using the ring-opening polymerization (ROP) of N-carboxyanhydrides (NCA) (**Figure 1**).

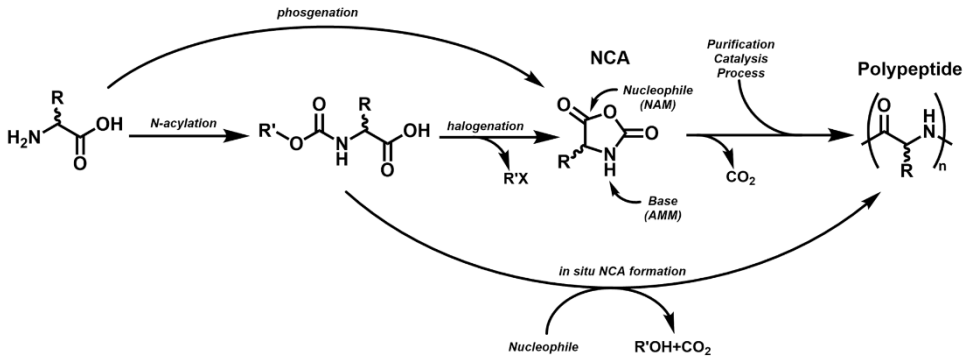


Figure 1: *Synthesis of NCA from amino acids and their derivatives. The ROP of NCAs to produce polypeptides is at the centre of recent developments, focusing on different methods to optimize this polymerization.*

A review of recent advancements in the synthesis of *N*-carboxyanhydrides (NCAs) and their ring-opening polymerization (ROP) (**Figure 1**) is presented in Chapter **1** of this manuscript. This polymerization method is very similar to other ROP techniques (such as lactide, oxazoline, etc...) and can easily be implemented on a large scale, thus opening prospects for applications in the field of materials—a potential not often offered by natural proteins.[8]

Peptidic polymers (polypeptides in **Figure 1**) have long been known to replicate certain structural features of natural proteins.[9] For example, these polymers are considered to be simplified synthetic analogues of protein structures since they can adopt typical secondary structures (alpha-helix, beta-sheet, etc.)[10].

Chapter **2** of this thesis focuses on a specific copolymerization process, the simultaneous ring-opening polymerization (ROP) of two types of NCA monomers. Drawing inspiration from the fact that protein structure depends on the arrangement of monomer units along the peptide chain (referred to as primary sequence), the objective of this chapter was to control the distribution of both monomers (block, gradient, alternating, etc...) in copolymer chains. A binary system was used to demonstrate how the insertion of one type of units can disrupt the helicity of a peptide block as a simplified model of IDP. More specifically, a model polypeptide that forms a helical structure was studied, poly(γ-benzyl-*L*-glutamate) (PBLG) derived from BLG NCA (**Figure 2**).[11] To disorder this polymer, it was copolymerized with glycine NCA (**Figure 2**), a residue that creates steric barriers to the formation of secondary structure in IDP sequences.[12] To achieve this, it was essential to understand how the two monomers react and integrate into a polymer chain, which was accomplished through a kinetic study of the copolymerization system.

We conducted a systematic investigation into the statistical control of the distribution of the two monomers along the polymer chain under various reaction conditions (solvent, temperature, feed ratio, and catalysis). Our work allowed us to establish a relationship between the distribution of glycine and the conformation of the polypeptide chain (**Figure 2**). Consequently, this study demonstrated the feasibility of obtaining "intrinsically disordered PBLG-based copolymers" through controlled glycine copolymerization to form gradient copolymers.

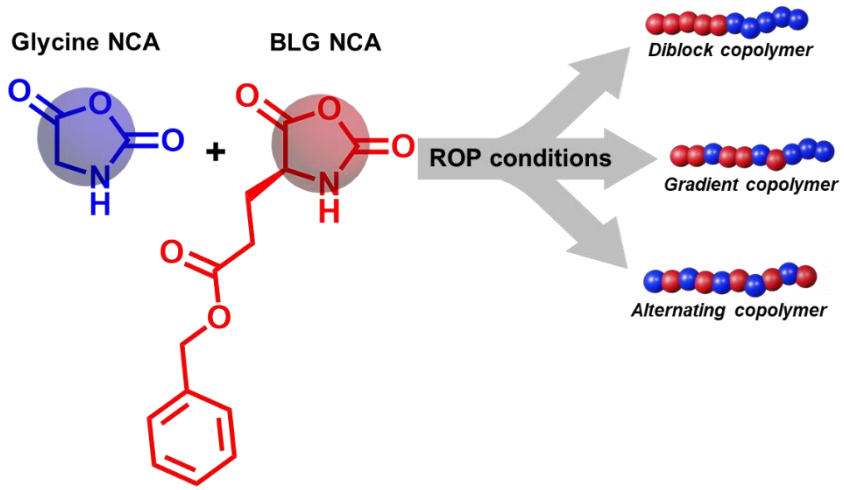


Figure 2: *Copolymerization of glycine and BLG NCA to obtain copolymers with different microstructures by controlling reaction conditions. Our study has shown that the choice of solvent, temperature, and reaction mechanism can influence the copolymerization towards one microstructure or another.*

In Chapter **3**, we shifted our focus to IDPs derived from elastin, which are temperature-sensitive and contain a significant number of proline residues in their sequence. This amino acid, commonly found in many IDPs, has not been extensively studied as a homopolymer, especially concerning its thermosensitive properties. In the case of NCA ROP, innovative work in polymerization[13,14] has enabled us to investigate the thermoresponsive properties of high molecular weight polyproline (PLP).[15] These polymers exhibit a lower critical solution temperature (LCST) associated with an exceptionally wide hysteresis. Chains are soluble at low temperatures but aggregate when heated above their cloud point temperature (T_{CP}) and only resolubilize at a significantly lower clearing point temperature (T_{CL}) (**Figure 3**).

An in-depth study of the conditions influencing this property, such as concentration or ionic strength of the medium, was carried out. These experiments demonstrated that the LCST behaviour of PLPs is strongly linked to their conformation (secondary structure) created by the chirality and the proline-proline peptide bonds.[16] We also exploited the hysteresis to produce reprogrammable materials with memory effects (**Figure 3**).

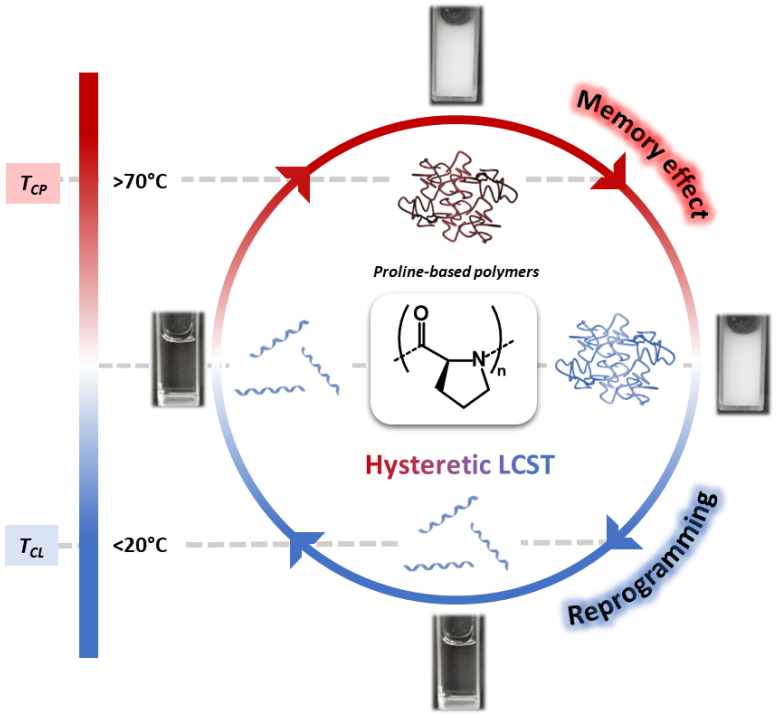


Figure 3: *Graphic summary of the memory effect in hysteric LCST polyproline formulations and their reprogramming.*

Taking a step further, we considered elastin-like polypeptides (ELP), a class of thermosensitive IDPs that contain both glycine and proline units.[17] Their LCST behaviour can be finely tuned through genetic engineering by choosing and sequencing these two types of amino acids.[18] Therefore, Chapter **4** focuses on the copolymerization of proline to modulate the LCST behaviour studied in Chapter **3** and to better mimic the thermoresponsive abilities of ELPs (**Figure 4**).

Copolymerization trials on binary systems were initially addressed. The produced copolymers exhibited a strong gradient character like those discussed in Chapter **2**. This was observed due to the high reactivity of proline compared to other comonomers under various ROP conditions. These copolymers (Pro-Gly and Pro-Val) presented LCST behaviour with hysteresis similar to PLP. By modulating their molecular weights, it was even possible to produce thermoresponsive nanoparticles. Terpolymerizations of glycine, proline, and valine were also explored in more depth to replicate the behaviour of ELPs (**Figure 4**). The study of these systems under water-assisted conditions and with organometallic catalysts allowed for controlled polymeric structures. These systems will enable the synthesis of copolymers containing proline with easily modulated thermoresponsiveness.

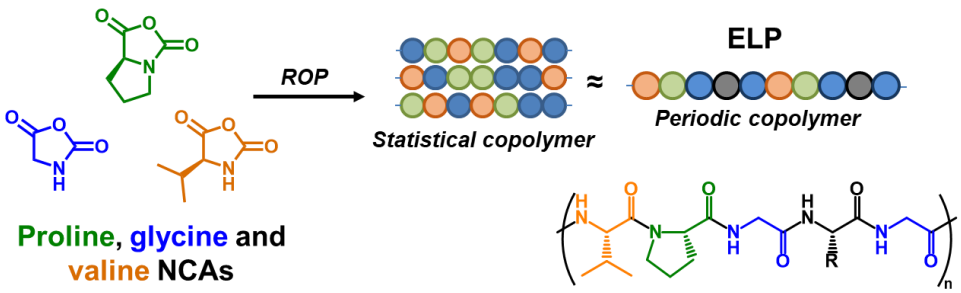


Figure 4: *Copolymerization of proline, glycine, and valine NCA to produce statistical copolymers. This work aims to statistically replicate the periodic sequence of ELPs and their properties.*

The knowledge gained in catalysis and copolymerization of NCA ROP has also opened up other interesting avenues. For example, the use of organometallic yttrium-based catalysts to produce polymers with controlled tacticity has been explored.[19] These complexes have demonstrated the ability to produce polypeptides with the desired structure, although the efficiency of stereoselectivity during polymerization remains limited. Finally, the copolymerization of NCAs with other cyclic monomers (ROCOP), such as cyclic esters, has been briefly studied to produce new polymeric structures called poly(ester-amide)s.

References

(1) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106* (5), 1877–1897.
(2) Alston, J. J.; Ginell, G. M.; Soranno, A.; Holehouse, A. S. *J. Phys. Chem. B* **2023**, *127* (21), 4746–4760.
(3) Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. *Proc. Natl. Acad. Sci.* **2010**, *107* (18), 8183–8188.
(4) Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. *Proc. Natl. Acad. Sci.* **2010**, *107* (33), 14609–14614.
(5) Theillet, F.-X.; Kalmar, L.; Tompa, P.; Han, K.-H.; Selenko, P.; Dunker, A. K.; Daughdrill, G. W.; Uversky, V. N. *Intrinsically Disord. Proteins* **2013**, *1* (1), e24360.
(6) Rath, A.; Davidson, A. R.; Deber, C. M. *Biopolym. - Pept. Sci. Sect.* **2005**, *80* (2–3), 179–185.
(7) Martin, E. W.; Holehouse, A. S. *Emerg. Top. Life Sci.* **2020**, *4* (3), 307–329.
(8) Song, Z.; Tan, Z.; Cheng, J. *Macromolecules* **2019**, *52* (22), 8521–8539.
(9) Rasines Mazo, A.; Allison-Logan, S.; Karimi, F.; Chan, N. J. A.; Qiu, W.; Duan, W.; O'Brien-Simpson, N. M.; Qiao, G. G. *Chem. Soc. Rev.* **2020**, *49* (14), 4737–4834.
(10) Bonduelle, C. *Polym. Chem.* **2018**, *9* (13), 1517–1529.
(11) Alekperov, D.; Shirosaki, T.; Sakurai, T.; Popova, G.; Kireev, V.; Ihara, H. *Polym.*

J. **2003**, *35* (5), 417–421.
(12) Hull, W. E.; Muller, D. **1985**, 2135–2140.
(13) Hu, Y.; Tian, Z.-Y.; Xiong, W.; Wang, D.; Zhao, R.; Xie, Y.; Song, Y.-Q.; Zhu, J.; Lu, H. *Natl. Sci. Rev.* **2022**, *9* (8), nwac033.
(14) Detwiler, R. E.; Schlirf, A. E.; Kramer, J. R. *J. Am. Chem. Soc.* **2021**, *143* (30), 11482–11489.
(15) Badreldin, M.; Le Scouarnec, R.; Lecommandoux, S.; Harrisson, S.; Bonduelle, C. *Angew. Chemie Int. Ed.* **2022**, *61* (46), e202209530.
(16) Kurtz, J.; Berger, A.; Katchalski, E. *Nature* **1956**, *178* (4541), 1066–1067.
(17) Meyer, D. E.; Chilkoti, A. *Biomacromolecules* **2004**, *5* (3), 846–851.
(18) Garanger, E.; Lecommandoux, S. *Adv. Drug Deliv. Rev.* **2022**, *191*, 114589.
(19) Ligny, R.; Guillaume, S. M.; Carpentier, J. F. *Chem. - A Eur. J.* **2019**, *25* (25), 6412–6424.

Abbreviations

AA	allylamine
ACN	acetonitrile
AMM	activated monomer mechanism
ATR-FTIR	attenuated total reflectance-Fourier transform infrared
BA	benzylamine
BLG	γ-benzyl-*L*-glutamate
Boc	tert-butyloxycarbonyl
CD	circular dichroism spectroscopy
DCM	dichloromethane
DCTB	trans-2-[3-(4-tert-Butylphenyl)-2-methyl-2-propenylidene]malononitrile
DMEA	dimethylethanolamine
DMF	dimethylformamide
DMSO	dimethysulfoxide
DNA	deoxyribonucleic acid
DP	degree of polymerization
DPPA	diphenyl phosphoryl azide
ELP	elastin like polypeptides
ESI-MS	electrospray ionization-mass spectrometry
EtOAc	ethyl acetate
EtOHA	2-(2-aminoethoxy) ethanol
FDA	Food and Drug Administration
$Fmes_2BF$	tris[2,4,6-tris(trifluoromethyl)phenyl]borane
GHS	globally harmonized system of classification and labelling of chemicals
HA	hexylamine
HFAB	1,3-Bis(2-hydroxyhexafluoroisopropyl)benzene
HFIP	hexafluoro-2-propanol
HMBC	heteronuclear multiple bond correlation
HMDS	hexamethyldisilazane
HMDS	bis(trimethylsilyl)amine
HPLC	high-performance liquid chromatography
HSQC	seteronuclear single quantum coherence
IDP	intrinsically disordered proteins/polypeptides
IDPR	intrinsically disordered protein regions
IDR	intrinsically disordered regions
*i*PrOH	isopropanol
LCST	lower critical solution temperature
LiHDMS	mithium hexamethyldisilazide
MALDI-TOF	matrix-assisted laser desorption/ionization coupled to time of flight analysis
MALS	multi-angle light scattering
Mn	number average molecular weight
mRNA	messenger ribonucleic acid
MSA	methanesulfonic acid
MW	molecular weight
NAM	normal amine mechanism

NCA	*N*-carboxyanhydride
NHC	*N*-heterocyclic carbenes
NMM	*N*-methylmorpholine
NMR	nuclear magnetic resonance
NNCA	*N*-alkylated-*N*-carboxyanhydride
NOE	nuclear Overhauser effect
NPCA	*N*-phenoxycarbonyl-functionalized α-amino acid
NTA	N-thiocarboxyanhydride
PBLG	poly(γ-benzyl-*L*-glutamate)
PDP	poly-*D*-proline
PEG	polyethyleneglycol
PGA	poly(glutamic acid)
PISA	polymerization induced self-assembly
PLA	poly(lactic acid)
PLP	poly-*L*-proline
rac-LA	racemic lactide
RE	rare earth metals
RI	refractive index
RNA	ribonucleic acid
ROCOP	ring-opening copolymerization
ROMP	ring-opening metathesis polymerization
ROP	ring-opening polymerization
SEC	size exclusion chromatography
SPPS	solid phase peptide synthesis
srCD	synchrotron radiation
T3P	n-propanephosphonic acid anhydride
TBBA	tetrabutylammonium benzoate
TBD	triazabicyclodecene
T_{CL}	clearing point temperature
T_{CP}	cloud point temperature
TEA	triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TMG	tetramethylguanidine
TMS	trimethylsilyl
TU-S	bis-(trifluoromethyl)phenyl]thiourea
UCST	upper critical solution temperature
Ugi 4CC	Ugi multi-component reaction
UV-Vis	ultraviolet-visible light spectroscopy
X_H	extent of hysteresis
α-CHCA	α-cyano-4-hydroxycinnamic acid

Chapter I: Synthetic amino acid-based polymers as simplified protein mimics

1. From proteins to synthetic polymers

Natural biomacromolecules are complex heteropolymers composed of several types of repeating units, arranged along the polymer chain in a sequence-controlled fashion. For example, deoxyribonucleic acid (DNA) is composed of 4 different monomer units called nucleotides and displaying cytosine, adenine, guanine or thymine nucleobases. The information contained in these sequences guides all biological functions. For example, nucleic acid strands carry the code needed to produce proteins. The proteins are the other face of life, as these macromolecules support the metabolism of living organisms. First, proteins' structural role in many living organisms makes them the largest class of biomacromolecules to be found in living systems such as bacteria. Second, proteins carry out many different and diverse functions: they have a catalytic role in biochemical reactions, they can generate energy, they store many compounds, and they form the dynamic skeleton of the cytoplasm. In metabolism, enzymatic reactions are achieved thanks to the unique structures of proteins: this results in specific interactions that bring reagents together or stabilize transition states. Their flexible conformations allow them to switch their three-dimensional structure in response to an external stimulus such as pH or temperature making them an inspiration to design smart materials.[1]

1.1. The structures of proteins

This diverse portfolio of natural proteins is made up of only 20 different types of monomers, the naturally occurring amino acids. These building blocks are composed of an amine and a carboxylic acid, connected by a single carbon called the α-carbon. Except for glycine, which is achiral, the natural amino acids are in the *L*-configuration, with the same relative orientation of acid, amine, and side chain. This relative configuration corresponds to the S-absolute configuration except for cysteine. They only differ from one another by the nature of their side chains, which affect the bending of the polymer backbone and stabilize its folding via different interactions. In cells, amino acids are

‘polymerized’ within the ribosomal complex from mRNA sequences. By translating the genetic information, the ribosome connects the amino acids one by one in a specific order through the formation of amide bonds. These are commonly called peptidic bonds which give the name polypeptide to linear chains of connected amino acids. Due to this amide bond, polypeptides have unique properties in aqueous media. It gives stability to the polymer chain in physiological conditions and allows cooperative hydrogen bonding with other polypeptides and the surrounding water molecules.

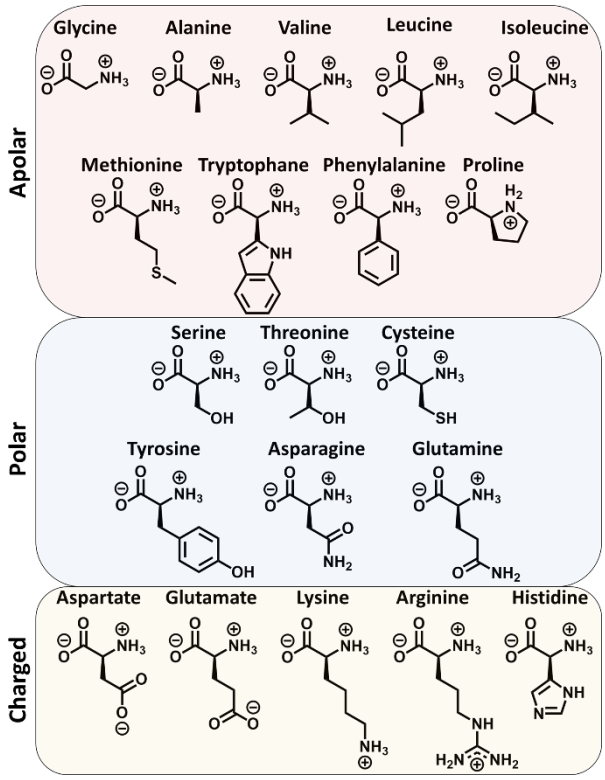


***Figure 1**: The structure of the 20 canonical amino acids found in the human proteome. The amino acids are classified according to the nature of their side chains. The simplest classification is shown in the figure: apolar or hydrophobic side chains, polar or hydrophilic side chains and charge bearing side chains. In the latter, amino acids can present acidic or basic functions. The protonation of all functional groups is shown at physiological conditions (pH=7.4).*

Different levels of protein structuring have been identified. First, the information carried by the nucleic acids is translated in the ribosome and determines the sequence of amino acids inside a protein. This **primary sequence** is key to most proteins’ functions in catalysis, binding, and structure as it is the base for all subsequent supramolecular interactions. The nature of the side chains and their alignment are of major importance

to the protein properties. For instance, hydrophilic side chains may stabilize polypeptides through hydrogen bonding with the aqueous environment. Some side chains present charges which depending on the pH and the environment, may change the solubility of the polypeptide chain. Other amino acids have hydrophobic side chains that undergo weak interactions such as the Van der Waals force and avoid interactions with water, thus bundling together in certain part of the protein (**Figure 1**). Finally, side chains can present amphipathic behaviour, which allows the formation of interfaces.[2]

The second level of protein structuring is the **secondary structure** which takes it form through macromolecular interactions. The chains organize themselves via hydrogen bonding interactions and the bending of the peptide backbone.[3] The nature of the peptide bond restricts the bending possibilities. Delocalisation of the π electrons in the amide gives the C-N bond a partial double bond nature and prevents the rotation of this bond. The other bonds in the backbone have more freedom to rotate. Two torsion angles can be identified in the repeating unit: the φ angle around the Cα–N bond between the C_{i-1}-N_i-$C\alpha_i$ half plane and the N_i-$C\alpha_i$-C_i half plane. The second angle, ψ, is around the C-Cα bond between the N_i-$C\alpha_i$- C_i and $C\alpha_i$- C_i-N_{i+1} half planes (**Figure 2 b.**). The conformation of the polypeptide chains can be roughly determined by the average values of the torsion angles. Since little rotational freedom can be allowed in the polypeptide chain due to its electronic configuration[4] and steric hindrance, two main known regions of angle values were determined.[5] These regions can be attributed to either alpha helices or beta sheets. The Ramachandran plot gives a probability of both angles in a protein sequence and its configurations. These two regions are thus the most prevalent in structured proteins (**Figure 2 c.**).

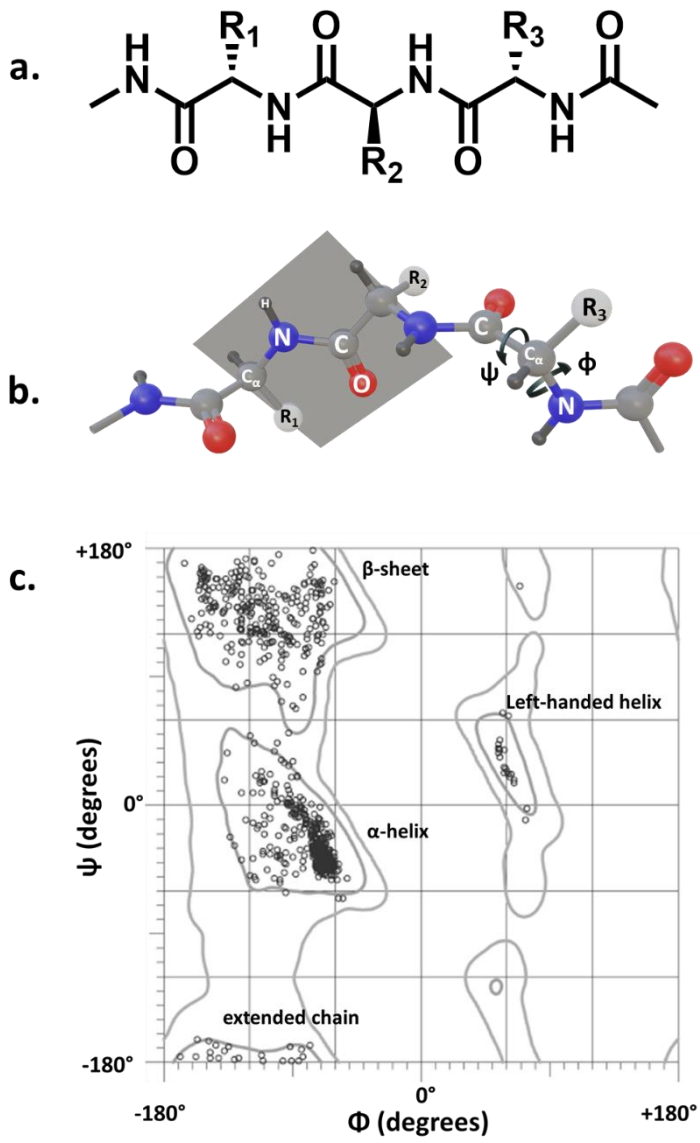


Figure 2: *a. Illustration of part of a peptide chain comprising of three amino acids and the primary structure from left to right; b. the same scheme represented in 3D as to show how the polypeptide backbone bends around the different peptide bond. The plane represented at the second amide linkage shows its coplanar alignment with the neighbouring α-C. The torsion angles φ and ψ of the backbone are represented around the Cα of the third amino acid; c. a general case Rama-chandran plot generated by MolProbity[6] of the human olfactory receptor OR51E2 bound to propionate in complex with miniGs399 protein complex. The lines delimit the "allowed" angle regions and the corresponding secondary structure. Each dot represents a repeat unit from the protein backbone.*

These secondary structures are stabilized by the hydrogen bonding of the peptide backbone. In the α-helix the main chain finds itself at the centre of the helix and the side chains are pointing outwards. Hydrogen bonding happens between the oxygen

of the carbonyl of each peptide bond and the amine group four monomer units further. This interaction forces the torsion angles to be around -60° for the ϕ angle and −45° for the ψ angle, forming the helical compact rod-like structure (**Figure 3**) which makes a complete turn every 3.6 amino acids. These helices are mostly right-handed, as this orientation is favoured by the chirality of the amino acids. The β-sheet is a more extended conformation in which the side chains project in alternating directions from the main chain, which is called the β-strand (**Figure 3**). The torsion angles of the β-sheets are ϕ = −130° and ψ = 120° and adjacent strands are linked by hydrogen bonds. The β-turn falls into the angle values of the β region and, as its name suggests, forces a turn in the direction of the polypeptide chain. It is the simplest conformation as it consists only of three or four residues. Due to their constrained nature, β-turns contain small amino acids such as glycine.[7] When a peptide chain's conformation does not fall into a certain category, it is usually considered to be in an extended or coil conformation. This can be compared to a random coil in a polymer solution.

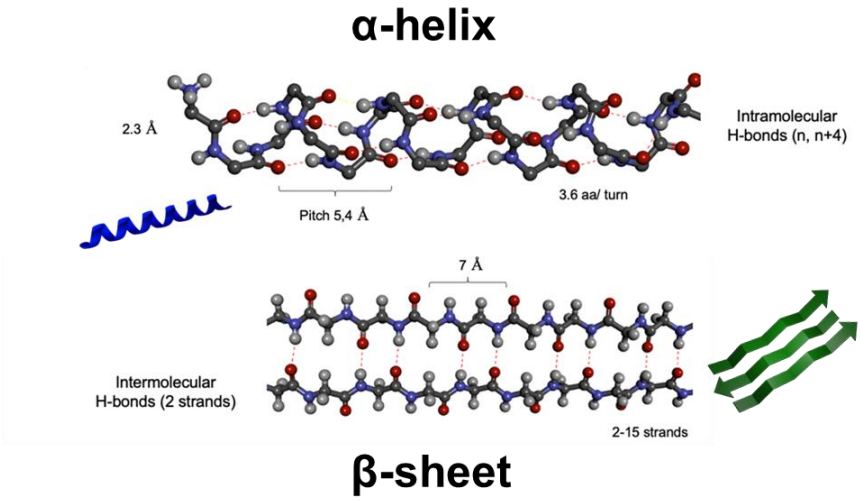


***Figure 3**: Balls and sticks molecular structure of the α-helix (on top) and its schematic representation (in blue); balls and sticks molecular structure of the β-sheet (on the bottom) and its schematic representation (in green). The distinct values and properties are highlighted.*

Other conformations are possible but are less commonly found in natural samples. Since the amide groups are usually involved in hydrogen bonding to form the secondary structures, the nature and position of the side chains typically govern interactions with the environment, which consists mostly of water molecules.[8] When combined,

secondary structures and other compartments of the protein backbone, the macromolecule takes a more globular form that corresponds to another level of structuring known as **tertiary structure**. Several tertiary structured polypeptide chains can eventually interact or combine to give the **quaternary structure** of a functional protein. These two higher order structures are more complex and will not be treated further in this manuscript focused on polymers that are able to mimic the structural and functional abilities of proteins.

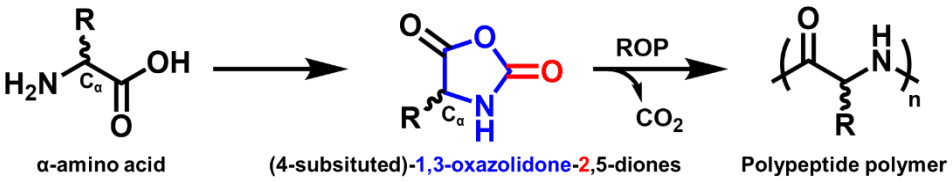
2. Amino-acid based synthetic polymers

Polymer chemists have long dreamed of reproducing the structural and functional diversity of complex biomacromolecules such as proteins. On another hand, polymer science could impart protein properties: one could envision polypeptides backbones comprising different block segments with different properties, similarly to synthetic block copolymers.[9] Recent works suggest that complex biological fluids can be mimicked using synthetic copolymers that can replicate the macroscopic properties of proteins (aggregation, etc.).[10] This has some limits, and the chemical tools available to produce amino-acid based macromolecules are still very basic in comparison to the biological tools available from living organisms. Indeed, an efficient and precise technique that is widely used in the industrial and pharmaceutical sectors[11] is to harvest the productivity of ribosomes inside bacterial hosts. The recombinant expression of proteins in *Escherichia coli* is widely used for this purpose. This can be done by cloning a gene expressing a protein or a polypeptide of interest and then injecting it into the bacteria. The microorganisms produce the polypeptides by translating the genetic information via the ribosomal complex. Finally, the targeted polypeptides can be purified from other components present in the cells.[12] This shows great possibilities and versatility in the production of high molecular weight, uniform and well-defined polypeptide chains with desired sequences.[11] This method is widely employed to produce IDPs which will be briefly presented in Chapter **4**. Nevertheless, it is time and cost-intensive and requires specialized biotechnological equipment.[13,14] Chemical tools that are more suitable for synthetic laboratories are also available for the production of peptidic backbones. One such tool is the solid phase peptide synthesis (SPPS).[15] The principle of this technique is to add monomers iteratively to a substrate that is covalently attached to a solid support via a cycle of reaction, deprotection and washing steps. The first

monomer reacts, in excess, with functional groups on the matrix then the excess is washed away. Then comes the deprotection technique and the monomer is then available to react with the next. This procedure could be repeated *n* times to produce peptidic chains with *n* amino acids, all with the same sequence, as long as quantitative yields are obtained and there are no side reactions. This aspect will be further explored in Chapter **2**. SPPS still suffers from drawbacks as high chain lengths are difficult to achieve. Furthermore, both the recombinant synthesis and SPPS have very low yield, low scalability and are highly costly in both materials and time.[16]

2.1. N-carboxyanhydrides and polypeptide polymers

The most economical and efficient way to produce amino-acid based polymers is through the ring-opening polymerization (ROP) (4-substituted)-1,3-oxazolidine-2,5-diones (Scheme 1), commonly known as *N*-carboxyanhydrides (NCAs). This polymerization is very similar to classical ROP procedures and can be employed in any synthetic chemical facility which facilitates scalability.



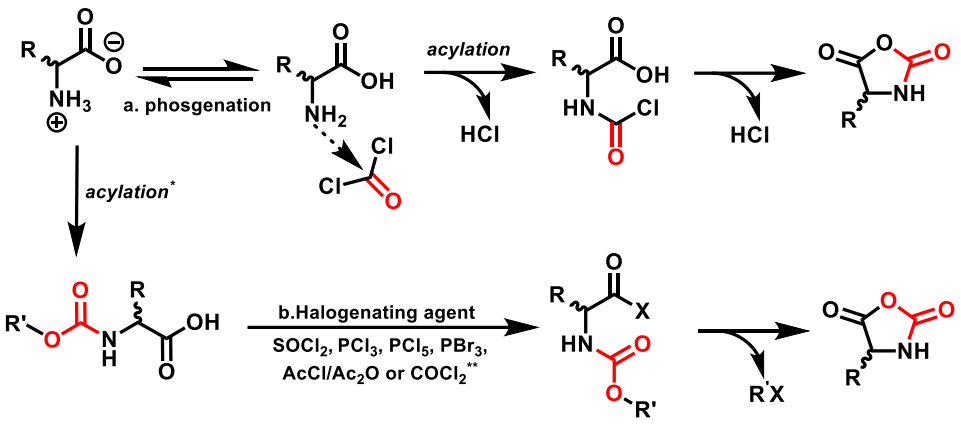
Scheme 1: *Schematic representation of the transformation of an α-amino acid into its corresponding N-carboxyanhydride, highlighting the oxazolidine ring (in blue) and the extra carbonyl group (in red) while keeping the chiral α-carbon. These monomers can be ring opened to produce polypeptide polymers.*

NCAs are highly active in comparison to amino acids due to the highly electrophilic 5-position carbonyl and the acidity of the nitrogen proton. Moreover, liberation of CO_2 during ring opening results in a large increase in entropy, making these monomers unstable under ambient conditions. Thus, they must be handled and used under inert conditions, either in a Schlenk line under nitrogen or argon, or in a glovebox. Synthetic routes to these monomers will be discussed in the following sections. Amino acids used to produce NCAs typically have their functional side groups protected as not to react with the newly formed NCAs. Furthermore, their synthesis typically requires extensive purification and solvents, if used, must be dry.

2.1.1. On NCA synthesis

As shown in **Scheme 2** the preparation of *N*-carboxyanhydrides from amino acids requires the incorporation of a carbonyl moiety that can later be liberated in the form of a CO_2 molecule. The most direct way of doing so was introduced by F. Fuchs in the 1920s by reacting *N*-benzylglycine with phosgene ($COCl_2$)[17] providing the *N*-alkylated NCA (NNCA). This method was later adapted to other amino acids by A. C. *Farthing*[18] and is to this day the most commonly used technique,[13] known as the Fuchs-Farthing method. Earlier in the same century, H. *Leuchs* first stumbled upon what was later called the "Leuchs' anhydrides" or NCAs. Exposing different *N*-alkyloxycarbonyl-amino acids to thionyl chloride provided the corresponding acyl chloride, which upon heating cyclizes into the NCA, releasing an alkyl chloride.[19,20] Although this older method produces NCAs from most amino acids[21], it was widely replaced by the Fuchs-Farthing method that does not employ amino acid derivatives and thus reduces the overall steps. Nevertheless, there are some recent examples that employ the Leuchs method to prepare NCAs of glycine[22,23], histidine[24] and arginine[25,26]. Other halogenating compounds have found success in synthesizing some recalcitrant amino acid NCAs. For example, proline and protected hydroxyproline NCAs were readily synthesized from their *N*-Boc protected amino acid precursors using a mixture of triphosgene and triethylamine, that provide the corresponding acyl chloride (**Scheme 2**).[27,28] Similarly, N_α-benzyloxycarbonyl, N_ε-modified-lysine reacts quantitatively with dichloromethyl methyl ether to give its corresponding NCA.[29] *Johnston* and *Mobashery* have also reported the use of the Vilsmeier reagent (combining oxalyl chloride with DMF) to cyclize *N*-Boc-silylated alanine to its NCA counterpart.[30] Lewis acids such as PCl_3 are commonly used to produce NCAs, while the less reactive PBr_3 has been used to prepare glutamine and asparagine NCAs.[31,32] *Akssira* et al. reported the synthesis of the NCAs of alanine, glycine, phenylalanine, phenylglycine, proline, tryptophane and valine by reacting their *N*-Boc forms with PCl_3 in DCM achieving high yields of NCAs.[33] Similarly, PCl_5 has been widely used to produce NCAs of amino acids including cysteine[34], tyrosine[34], glycine[35] and alanine[36]. *Zhang* et al. reacted amino acids with dimethylcarbonate producing the *N*-methoxycarbonate derivative that underwent the cyclization in the presence of glacial acetic acid and heat.[37] The more stable, thiolated counterparts of NCAs, *N*-thiocarboxyanhydrides (NTA), are produced via a similar mechanism to produce similar structures with a sulphur atom between the two carbonyls. Their use in polypeptide formation has been reported but has not been widely adopted, since they are not easily

crystallized.[38] Furthermore, the release of toxic and flammable carbonyl sulfide in the polymerization process could be a reason to favour NCAs.



Scheme 2: *Synthesis pathways of NCAs and reagents used. a. Fuchs-farthing pathway using phosgene and derivatives for the N-acylation process, b. Leuchs pathway using N-protected amino acids with halogenating agents to produce acyl chlorides. *acylation of amino acids could be done via several pathways[39–42]but they are usually commercially available. **phosgene and derivatives are widely used for the formation of acyl chlorides[43–45]*

The Fuchs-Farthing method, though widely used nowadays, has its disadvantages. Gaseous phosgene is mortal by inhalation (H330: Fatal if inhaled, GHS hazard statements[46]) and a Schedule 3 substance under the Chemical Weapons Convention[47], it is thus mostly used in the form of standardized solutions.[48] Easier to handle alternatives to phosgene include trichloromethyl chloroformate or diphosgene, a liquid and (trichloromethyl)carbonate or triphosgene, a solid. The latter seems to be the most popular choice among NCA synthesizers given its stability and reduced toxicity.[49] Both di- and triphosgene are phosgene precursors and must be activated by heating or the addition of catalyst to liberate $COCl_2$.[49,50] *Endo* et al. proposed an alternative to phosgene by using bisarylcarbonates as cyclizing agents. This follows a similar mechanism but releases phenols instead of HCl limiting possible acidic hydrolysis.[51,52] A novel method has also been proposed by *Sugimoto* et al. using photophosgenation.[53] The authors showed the *in-situ* formation of transient phosgene by exposing chloroform to O_2 and irradiation allowing for the formation of NCAs without the use of actual phosgene. Nonetheless, the use of O_2 seems to add an underlying danger due to the risk of explosion as well as the need for specialized apparatus. Similarly, the nitrosation of

N-carbamoyl-amino acids using different NO_x gas mixtures has proven efficient in producing NCAs.[54,55] This speculatively prebiotic route[56] produces water as byproduct of cyclization and thus needs extensive drying for NCA storage and controlled polymerization.

Other mechanisms, such as the internal cyclization of acyl azides from amino acids after a Curtius rearrangement,[57] have also been explored. Recently *Laconde* et al. have shown that exposing the *N*-alkyloxycarbonyls of amino acids to n-propanephosphonic acid anhydride (T3P), in the presence of either a base or heat, successfully yielded the corresponding NCA in high yields and high purity.[58] This synthetic route requires no dangerous reagents and does not produce polymerization inhibitors as side products, such as the alkyl chlorides of the Leuchs method. Nonetheless efficient polymerization of these monomers is still to be proven. This technique has recently been optimized further by using amino acids and CO_2 with T3P to yield NCAs.[59] This method seems very promising because while *N*-alkyloxycarbonyls of amino acids are cheap and readily available, their synthesis from amino acids usually employs phosgene derivatives. This new method based on CO_2 might thus prove to be a green process, producing no undesired byproducts and allowing for a neutral carbon footprint of the direct synthesis of polypeptides. Nevertheless, much like the previous method, no polymers were investigated. Furthermore, the use of gaseous CO_2 requires high pressures and thus special equipment.

2.1.2. On NCA purification

The development of new techniques for NCA synthesis has been driven by the flaws that plague the Leuchs and Fuchs-Farthing methods, still the most widely used methods to produce NCAs. Their main problem is the elimination of byproducts that can be toxic and can ring open the newly synthesized NCAs. For example, both phosgene and $SOCl_2$ produce highly corrosive HCl which can transform the NCAs to their amino acid hydrochloride salts. Moreover, any amino acids or their salts, whether unreacted reagents or degradation products of NCAs, must be removed.[60] If not, these molecules can initiate the ring opening of the NCAs, reducing their shelf life and limiting the controllability of the polymerization. The Leuchs method generates halogenated byproducts which must also be removed from the crude product. Furthermore, any traces of acidity or chlorine can result in retardation of polymerization and loss of control over the molecular weight distribution as they can hinder the propagating species or even

deactivate it. Extensive washing/filtering and repeated recrystallizations are typically used to purify NCAs but are sometimes ineffective. Cold aqueous liquid/liquid extraction is also widely employed to remove nucleophilic compounds.[59,61] Although effective, the use of water in contact with water sensitive NCAs seems risky. Some highly hydrophilic NCAs are themselves water soluble and can be extracted into the aqueous phase. A microfluidic system with an acidic/basic flash switching has recently been developed to produce pure NCAs easily without further need for purification.[62] Through a microflow reactor, the biphasic mixture with on one side the amino acid sodium salt in aqueous solution and on the other triphosgene in an organic solvent, produces once mixed the NCA in a very rapid manner (**Figure 4**). In these basic conditions, the carbonate is overly reactive towards phosgene and its subsequent cyclization to NCA. The water present in the aqueous media thus decomposes any extra phosgene and, by adding a mild base such as *N*-methylmorpholine (NMM), is neutralized by the resulting HCl. This brings the overall aqueous pH to a more acidic value. Further along the microfluidic tract, the addition of further organic phase separates the NCA from its more water-soluble undesired by products. Although this method requires the use of specialized equipment and knowledge of microfluidics, it has shown promising results in producing several NCAs and a potential for scale up.[62]

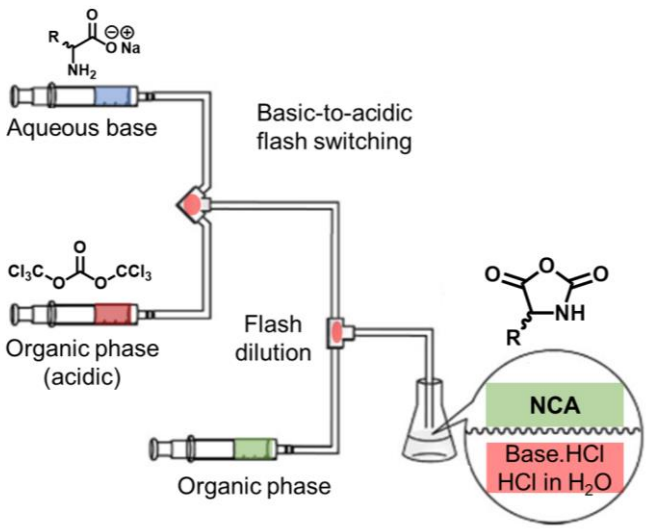
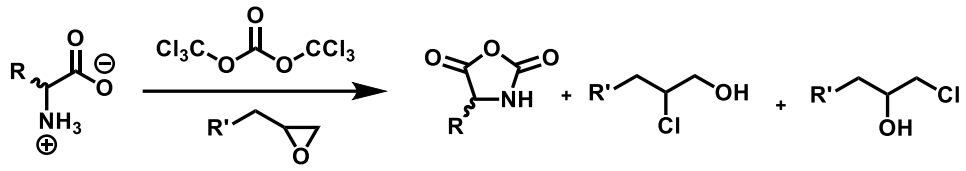


***Figure 4**: Schematic representation of the microfluidic system developed for the synthesis of NCA through basic to acidic switching. Colors reflect the intended pH as would be reflected by pH paper (red for acid, blue for basic and green for neutral). Figure was adapted from reference[62].*

In a simpler approach, the addition of acid scavengers such as triethylamine, (+)-limonene or α-pinene was effective in reducing the contaminants and the risk of degradation of NCAs, specifically by precipitating during the reaction with HCl.[61] More recently, *Tian* et al. reported that the use of highly reactive epoxy compounds (**Scheme 3**) provides a clean and moisture-tolerant synthesis in an open flask. This method could ring close amino acids bearing unprotected reactive side chains, allowing the synthesis of previously unreported NCAs.[61] The resulting chlorinated alcohols can be evaporated easily with the solvent to provide the NCAs with minimal workup.



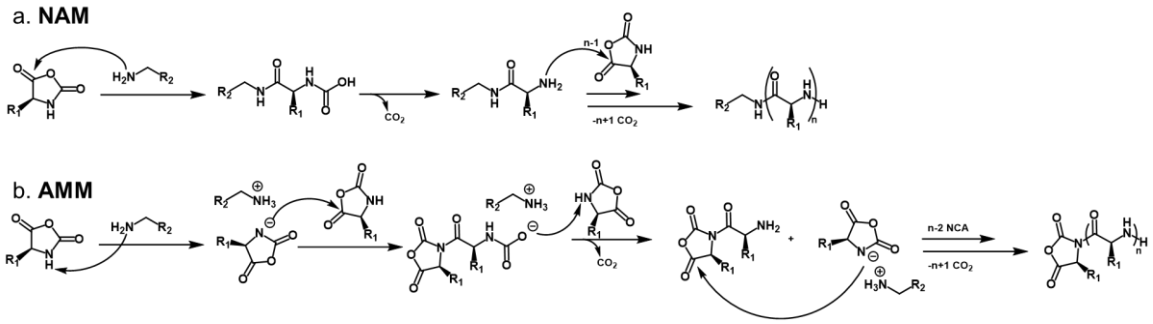
Scheme 3: *Reaction scheme of NCA synthesis using triphosgene using epoxy compounds (R'=H or Cl) as acid scavengers.*

Another easy yet very effective way was developed by *Kramer* and *Deming* using anhydrous flash chromatography on dried silica gel inside a glovebox.[60] Similarly, passing the crude reaction medium over dried celite will capture any charged and hydrophilic moieties while keeping the NCAs in solution.[63]

2.2. ROP of NCA

Nucleophilic attack on the ester-like carbonyl of NCAs proceeds via a simple ring opening, CO_2 releasing mechanism. The nucleophilic attack of a primary amine results in ring opening and the newly formed free primary amine can thus undergo a nucleophilic attack on another NCA affording the formation of the amide or peptide bond with the living chain end. This is the normal amine mechanism (NAM) (**Scheme 4**).[64] A competing mechanism is known as the activated monomer mechanism or AMM. In this second pathway the amine is deprotonated by a base affording an "active monomer" NCA anion. This undergoes a nucleophilic attack from this NCA anion, producing a new active moiety with the liberation of CO_2 (**Scheme 4**).[22] Although this second mechanism leads to faster polymerizations, the ROP is usually uncontrolled[21] and has wider dispersities.[65] When employing amines as initiator for the ROP, their basic nature must be considered and the AMM is arguably better eliminated. Thus primary amines,

which are more nucleophilic and less basic than their secondary or tertiary counterparts, are usually preferred.[16]



Scheme 4: *a. Normal amine initiation mechanism (NAM) of NCAs using a primary amine, b. activated monomer mechanism (AMM) that can occur during the ROP of NCA using primary amines.*

Alcohols can also be used to ring open NCAs with simple acid catalysis. The corresponding molecule having a dormant NH_3^+ species can be activated by adding a base allowing for the propagation that is much faster thus ensuring the control over the MW through the initiation step.[66] Among other nucleophiles, thiols have also been shown to produce the NAM with NCAs.[67] In the following different mechanisms of the ROP of NCA will be discussed as well as other catalytic routes.

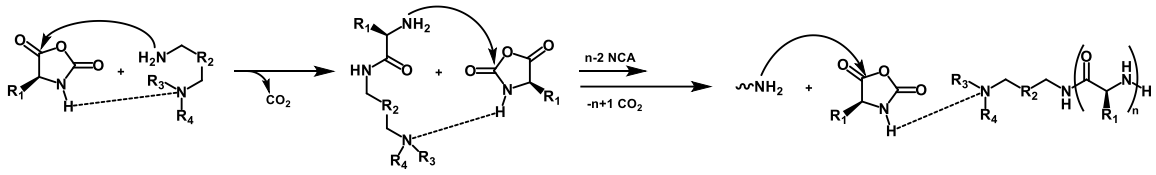
2.2.1. NAM

Much work has been done to optimize this mechanism to ensure living conditions and to achieve predictable and high molecular weights while minimizing dispersity. Other nucleophiles, like water, can also initiate the polymerizations so it is usually advisable to N_2 or Ar atmosphere…). Improving CO_2 elimination, either by conducting the polymerization under high vacuum[68,69] or increasing the flow of inert gas[16,70], has been shown to improve control and speed-up kinetics of the polymerization. This way minimizing any side reactions can be minimized and long reaction times are avoided as well as eventual NCA degradation.

Reducing or eliminating the anionic mechanism (**Scheme 4**) that can terminate the chain propagation processes and provide uncontrolled polymerizations is also another concern.[71]The choice of initiator also influences the polymerization process. For example, using a primary amine in its hydrochloride salt form can improve the control of the polymerization.[72] This is attributed to a shift in the acid-base equilibrium towards the

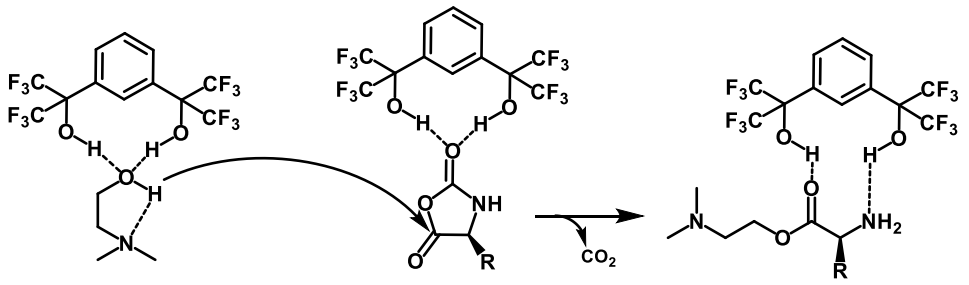
amine through the reaction of the hydrogen chlorine with any NCA anion that can be formed which allows to avoid the AMM mechanism.

Other analogues or derivatives of amines have also been shown to enhance the ROP mechanism. Using a synergetic effect between a primary and a secondary[73]/ tertiary[74] amine on the same molecule can highly enhance the rate and the control of the ROP (**Scheme 5**). The monomer is activated by the interaction with the secondary amine, allowing for a more effective attack by the primary amine. The same effect is continued in the propagation of the polymerization either by the same molecule or by a neighbouring polymer chain (**Scheme 5**).



Scheme 5: *Allied amine mechanism for the ROP of NCAs. R_3 and $R_4 \neq H$.*

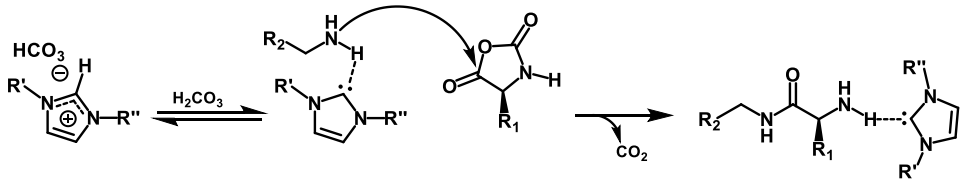
Still some evidence shows the possibility of the coexistence of both these mechanisms while affording well controlled polypeptides. Using a combined initiator containing a primary ammonium and a tertiary amine for the ROP of NCA resulted in polymerizations deviating from first order kinetics.[65] This could be explained by fast AMM at the beginning of the polymerization that is followed by predominance of the NAM in later stages.[65,75] The use of an alcohol and a tertiary amine on the same initiating molecule was also possible using dimethylethanolamine (DMEA). The nucleophilicity of the DMEA hydroxyl group is enhanced by intramolecular hydrogen bonding to the tertiary amine, enabling efficient initiation of the NCA.[76] In this study, the authors coupled this effect to an organocatalyst. The use of 1,3-bis-(2-hydroxyhexafluoroisopropyl)benzene (HFAB) coupled with DMEA enhanced the ROP of NCAs avoiding any side reactions. Thanks to hydrogen bonding between HFAB and the initiator, monomers, and propagating chain ends, (**Scheme 6**) the polymerization proceeded faster and with higher control. Similar conclusions were made by using thioureas such as N,N'-bis[3,5- bis(trifluoromethyl)phenyl]thiourea with DMEA.[77] These results show an effective initiation through the alcohol of DMEA activated by the presence of both the tertiary amine and the catalyst. This much faster polymerization led to well controlled, high molecular weight polypeptides.[76]



Scheme 6: *Proposed mechanism of the catalysis of HFAB during the ROP of NCAs initiated by DMEA.*

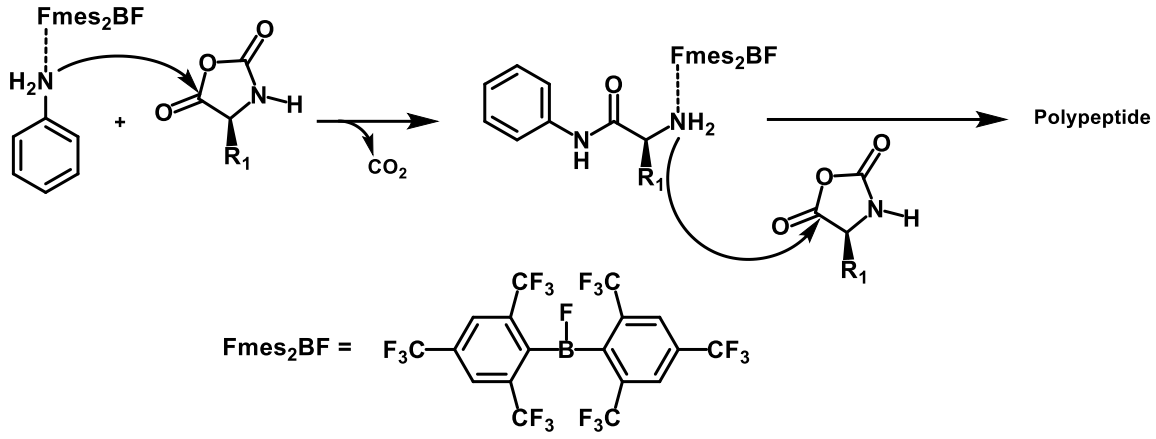
(S)-1,1′-binaphthyl-based phosphoric acid were also good catalysts for the ROP of NCAs initiated by primary amines. The activation of the initiation step and the establishment of an equilibrium between actively propagating amine species and dormant ammonium phosphate species allowed for higher control over the polymerizations.[78]

N-heterocyclic carbenes (NHC) have also been explored as organocatalysts for the NAM. Combining imidazolium hydrogen carbonates (easy to handle NHC precursors) with an amine resulted in enhanced polymerization rates and control (**Scheme 7**). Free NHC alone also resulted in NCA polymerization, producing cyclic polypeptides through a zwitterionic mechanism.[79]



Scheme 7: *NHC mediated NAM.*

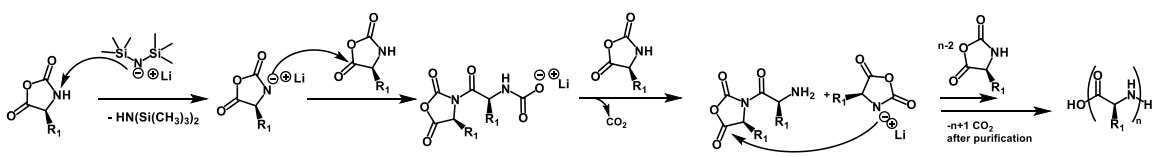
Using tris[2,4,6-tris(trifluoromethyl)phenyl]borane ($Fmes_2BF$) as Lewis acid combined with aniline as Lewis base (**Scheme 8**), achieved polymers with predictable molar masses and dispersities lower than 1.3.[80] Zinc acetate is also an effective Lewis acid in combination with aniline. [81] Their combination with NCAs was also successful in producing polypeptides with similar structures. The mechanism of this ROP has yet to be proven with NCAs but has been studied on the ROP of other cyclic monomers. In the case of lactide, the interaction between the zinc and the carbonyl catalyses the nucleophilic attack of the alcohol and the ring opening.[82] Similar mechanism can thus be envisioned for NCAs.



Scheme 8: *Proposed mechanism for the use of a frustrated Lewis pair (aniline and $Fmes_2BF$) in the ROP of NCAs.*

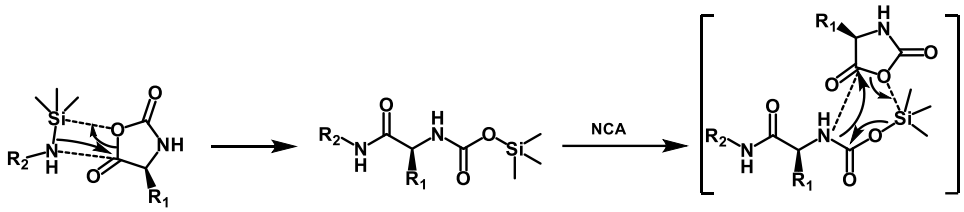
2.2.2. Other mechanisms

Other mechanisms than the NAM have also been shown in the literature. Propagation in the AMM mechanism proceeds via a carbamate chain end.[22] Tertiary amines such as triethylamine give this mechanism but in a highly uncontrolled manner, producing polypeptides with very high dispersities.[21,65,75] This is because deprotonation of the NCA is slow and reversible. New chains are generated throughout the reaction, leading to broad molecular weight distributions. Other moieties have also been investigated, including LiCl and sodium methoxide,[38] as well as complexes of palladium and platinum[83] without significant success in obtaining controlled polymerization. Recently, a controlled AMM has been reported using lithium hexamethyldisilazide (LiHMDS). The ROP using LiHMDS afforded very high MW polypeptides with very low dispersity in just a few minutes. First the organobase deprotonates the amine producing the anionic NCA with a lithium counter ion and hexamethyldisilazane (HMDS). The anionic NCA can undergo a nucleophilic attack on a second monomer producing a terminal lithium carbamate which propagates in a similar manner (**Scheme 9**). While this technique is very rapid due to the strong basicity of LiHMDS, the obtained MW deviated from theoretical M/I values.[84] This could be due to the presence of trimeric form of LiHMDS instead of its active monomeric version.



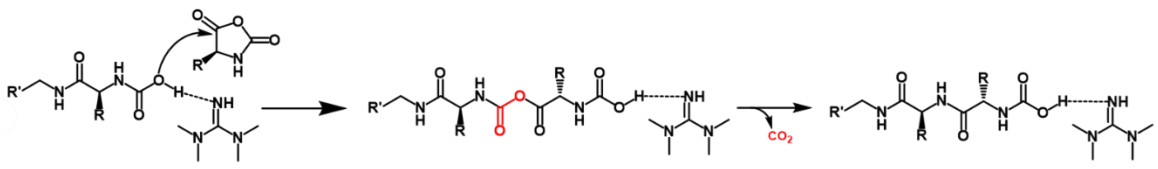
Scheme 9: *Mechanism of the ROP of NCA using LiHMDS.*

These results are related to the discovery of the ROP of NCAs using *N*-trimethylsilyl (TMS) amines such as bis(trimethylsilyl)amine (HMDS) [85], trimethylsilyl dimethylcarbamate[85] and various other *N*-TMS amines[86]. HMDS induces a polymerization mechanism that is neither NAM nor AMM. *Cheng*'s group have shown that HMDS reacts with the NCA to provide a TMS carbamate (**Scheme 10**). This carbamate is the propagating species, and transfers in a similar manner to newly added monomers.[85] This mechanism was extended to trimethylsilyl sulfides, which form a thioester group on one end of the polypeptide and a reactive TMS carbamate on the other.[87]



Scheme 10: *TMS mediated initiation of NCA and propagating intermediate.*

Similarly, by using trimethylstannyl phenyl sulfide faster controlled polymerizations were possible.[88] Tetramethylguanidine (TMG) can catalyse both the propagation step of NCA polymerization through a carbamate chain end and its initiation by carboxylic acids, alcohols and primary amines. Given its basicity and its interaction with the protons present on the initiator, the initiation step is highly enhanced and proceeds similarly to the NAM via a nucleophilic attack. As demonstrated by MALDI-TOF analysis, the stabilization of the carbamate at the end chain by TMG does not directly eliminate the CO_2 and allows higher control over the polymerization. After the addition of a second monomer, an intramolecular CO_2 elimination can then occur (**Scheme 11**).[89]



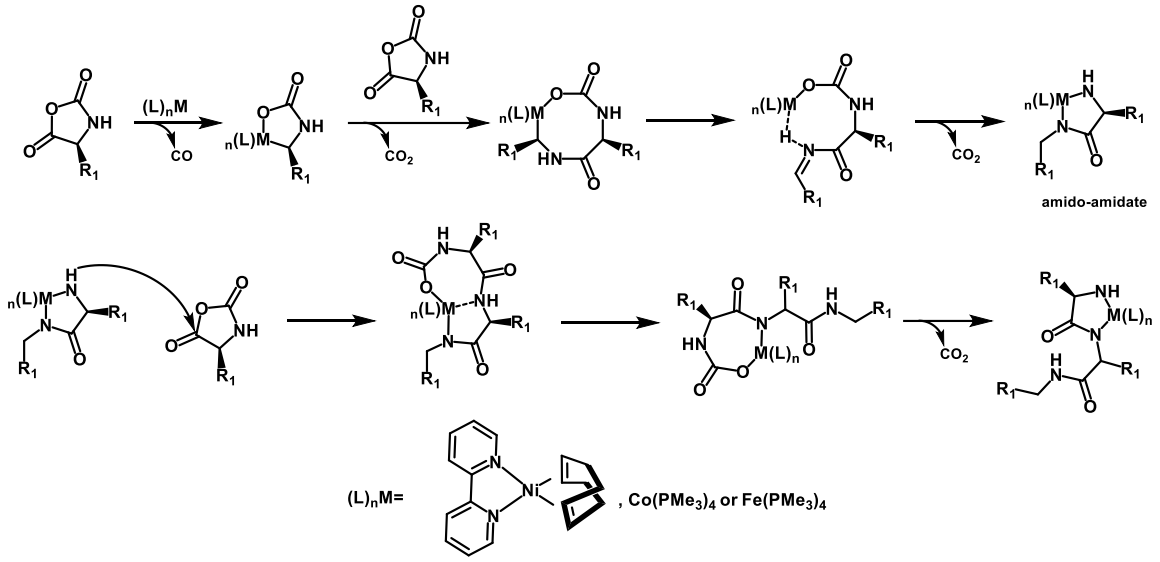
Scheme 11: *TMG catalysis propagation mechanism during the ROP of NCAs.*

The anionic ring opening polymerization of NCAs via a similar propagating moiety was reported using a catalyst containing three thiourea functions.[90] When a highly nucleophilic anionic initiator such as sodium phenolate was used, the ROP proceeded very rapidly and with good control over molecular weight. Through anion binding and hydrogen bonding, the trithiourea catalyst with electron withdrawing substituents can activate

the initiator, allow proximity with the monomer, and help with ring opening much like previously described organocatalysts. The carbamate moiety thus formed is stabilized by interaction with the catalyst that then helps with the decarboxylation and nucleophilic attack for the propagation. The polymerization was even successfully conducted in an open flask and in the presence of up to 6% w/w water in DCM.

2.2.3. Organometallic catalysis

The field of ROP of NCAs has also seen the development of organometallic catalysis. Metal acetate-tri-n-butylphosphine complexes showed promising polymerization activity with BLG NCA.[91] The mechanism involved in this pathway using organonickel compounds to prepare well defined and controlled polypeptides with a highly living character was first described by *Deming*. NCA monomers were activated through oxidative ring opening by a zero valent nickel complex, (2,2'-bipyridyl)(1,5-cyclooctadiene)nickel(0) (bipyNi(COD)), allowing for the polymerization. The polymers produced were of high molecular weight and very low dispersity.[92] Alternative first row transition metals, cobalt and iron, were also investigated. Polymerizations using (trimethylphosphine)$_4$cobalt(0) or iron(0) showed more efficient and rapid ROP of NCAs and followed a similar mechanism.[93] The connection between the complex and the NCA at the 5-carbonyl and the 4-ether yields a divalent nickel 6-membered amido-alkyl metallacycle, which upon reacting with a second NCA and losing CO_2 contracts to a 5-membered amido-amidate metallacycle.[13,94] This new active species can then attack the electrophile carbonyl of a new NCA monomer via its nucleophilic amido group and result in CO_2 loss. A proton migration from the free amide to the amidate group would bring back the amido-amidate propagating species and the metal would thus move along to the active chain end (**Scheme 12**).[95]



Scheme 12: *Initiation (top) and propagation (bottom) mechanisms when using zero valent metal initiators for the ROP of NCAs.*

Furthermore, pre-formed divalent nickel amido-amidate complexes can be used as initiators to produce end-group functionalized polypeptides.[96] Recently, more studies elucidating the mechanism and the use of such complexes with *N*-alkylated NCAs such as proline and protected hydroxyproline have emerged.[27] Ruthenium and iridium complexes were also investigated and showed similar successful results.[97] Furthermore, by using ruthenium or iridium based amido-sulfinamidate complexes bearing chiral chelates, it was possible to selectively polymerize different NCA enantiomers. Similarly to the ruthenium complexes, platinum amido-amidate mechanism was explored in the preparation of high molecular weight polypeptides with low dispersity.[98] This also englobed rare earth metal complexes that were similarly explored. The activity of several yttrium, scandium, lanthanum, and dysprosium complexes were shown to produce polypeptides with narrow dispersities and good control over the molar masses. Using these $M(BH_4)_3(THF)_3$ was shown to follow a more classical nucleophilic ring opening polymerization. On the other hand, using $M(NTMS)_3$ the mechanism was more similar to that of HMDS explained earlier.[99] *Bhaw-Luximon* et al. reported the use of Schiff's base/aluminum alkoxide complexes to successfully initiate the polymerization of NCAs. A possible mechanism involving an interaction between the Al atom and the -NH of NCA while the nucleophilic alkoxide attacked on the C-5 carbonyl is proposed. Nonetheless, the propagation step seems to be similar to the NAM.[100]

2.2.4. Mild conditions

The high activity of NCAs entails significant moisture sensitivity, which is why they should be handled under dry conditions. Although aminolysis is typically faster than hydrolysis, most of the previously mentioned mechanisms and systems require dry inert conditions to eliminate the possible initiation by water in the case of NAM, termination in the case of the AMM and to avoid degradation of moisture sensitive catalysts.[101,102] Nonetheless, it has been shown that controlled polymerizations can be obtained in less stringent conditions. Most of these techniques rely on the rapidity of the polymerization that outpaces the monomer degradation. This is the case for LiHMDS, described earlier.[84] Tetraalkylammonium carboxylate initiators behave in a similar manner.[103] The resulting polymerizations afford high molecular weight and low dispersity polypeptides in very short times (less than one minute). The carboxylate group of the initiator was observed at the chain end, excluding a possible AMM involving deprotonation. The authors present a plausible pathway in which the sterically hindered tetraalkylammonium counter cation stabilizes the chain end and the entering monomer. This concerted addition and decarboxylation means that a carbamate moiety is the propagating species.

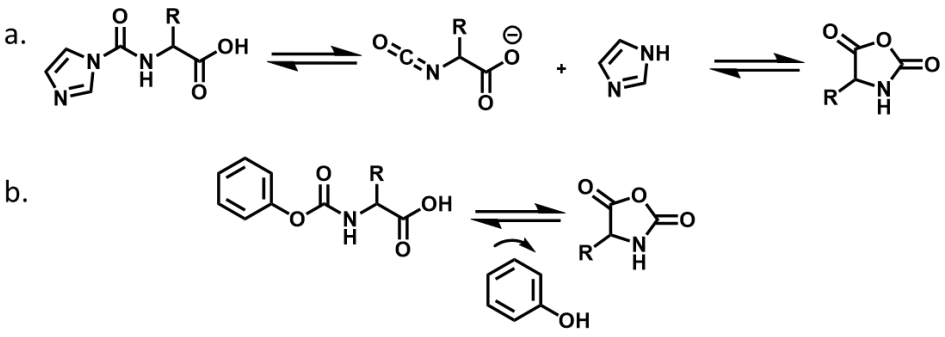
Using crown ethers showed the acceleration of the ROP of NCAs initiated by amines from hours to mere minutes. Using 18-C-6 crown ether, this method produced the desired polypeptides in a highly controlled manner in lenient conditions, such as biphasic mixture of DCM/aqueous buffer and using unpurified NCAs.[104]

Amphiphilic diblock macroinitiators have also been used in biphasic conditions. *Cheng* et al. have shown that amine-terminated PEG-*b*-PBLG that are localized at the interface of the biphasic DCM/water mixture enable a rapid polymerization that is unaffected by the presence of water. The authors used a similar biphasic mixture to purify NCAs from the reaction between phosgene and amino acid, reducing the overall procedure to just two steps from NCA synthesis to polypeptide production.[105] Nonetheless, these methods have only been reported to produce helical polypeptides, like polyglutamates, in a helix inducing solvent like DCM[106]. Our laboratory developed the polymerization induced self-assembly (PISA) of NCAs, called ROPISA in strictly aqueous media.[107] Although the polymerization of NCAs in buffered aqueous solution has been reported[108], the formation of high molecular weight polypeptides was still not reported. This procedure eliminates the use of organic solvents while directly producing highly

concentrated polymeric nanomaterials. When hydrophilic macroinitiators are used, rapid polymerization occurs as a result of self-assembly of the polypeptide blocks and their secondary structuring. This self-assembly allows for concentrated clusters of monomers in the dispersed phase that accelerates the ROP and protects them from hydrolysis.

2.2.5. Bypassing NCAs

Other strategies to overcome the sensitivity of NCAs to moisture and impurities that might induce a premature ROP have been developed. In this case the preparation of stable precursors that, under certain conditions, spontaneously produce NCAs that can further polymerize have gained a lot of interest. *Buess* and coworkers first showed that α-carboxy benzoylhydroxamic acids can undergo a Lossen rearrangement to give the corresponding isocyanate (**Scheme 13**).[109] The proximity of the isocyanate and carboxylic acid allows for the ring closing into NCA that will hydrolyse into polypeptides spontaneously in aqueous conditions. Later *Orgel* et al. showed that imidazole groups on such compounds allowed for a facile way to produce polypeptides in aqueous buffers.[110] *N,N'*-carbonyldiimidazole can react with amino acids in aqueous conditions to give *N*-[imidazolyl-(1)-carbonyl]-amino acids. They showed that the formation of oligomers is attributed to the intermittent formation of NCAs that ring open in the presence of water. The equilibrium between the different charged states of the amine, imidazolium and carboxylate allows for a similar rearrangement to that discussed earlier.



Scheme 13: *a. Lossen rearrangement of an α-carboxy benzoylhydroxamic acids. The isocyanate form can be in equilibrium with an NCA. b. N-phenoxycarbonyl-functionalized α-amino acid and the formation of corresponding NCA*

Endo's group developed a similar synthesis by preparing activated urethanes of α-amino acids, *N*-carbamoyl-amino acids, which were easily worked up and purified (**Scheme 13**).[111] These monomers can be stored without special precautions. The

group further demonstrated this techniques' ability to yield many types of homopolypeptides and copolypeptides with amine initiations and defined structures.[112–115] This *in-situ* formation of NCA makes the purification and storage of the monomers much more convenient and also bypasses the use of dangerous reagent like phosgene derivatives.[113] This polymerization pathway is robust with respect to water and other hydrophilic impurities. Notably, the polymerization of tyrosine could be conducted without protecting its hydroxyl group.[116] Building on this field, another group have recently shed more light on the mechanism of the method by using *N*-phenoxycarbonyl-functionalized α-amino acids (NPCA) as precursors for NCAs.[117] The authors show that using these precursors with primary amine hydrochloride as initiators, the polymerization proceeds under starved conditions. The low transient NCA concentration allows for their rapid consumption avoiding any undesired side reactions. Since the NPCA needed to be heated to convert to NCAs, the authors employed ammonium chlorides as initiators. In these conditions the dormant ammonium form is more stable than the amine one leading to higher polymerization control. Furthermore, the released phenols from the NPCA cyclization help in quenching any anionic species that can plague the reaction. In their study the authors also showcased the moisture insensitive nature of this pathway by conducting the polymerization under open flask conditions and wet solvent (DMA).

In the following chapters the synthesis and ROP of NCAs will be used to yield polymers with amino acid-based backbones as a first step to mimic the structural and the functional features of a specific class of proteins called intrinsically disordered proteins (IDPs). The first challenge of this work was to study the chemical leverage available to control the monomer distribution during the ROP process to mimic the primary sequence of these proteins (chapter **2**). Details in the literature about the microstructure of random copolymers prepared via the ROP of NCA are scarce. Furthermore, very little information was found regarding the effect of the different ROP mechanisms and catalysts on the reactivity of comonomers. That is why the copolymerization of NCAs, the effect of different reaction conditions on their statistical distribution will be considered in more detail in the next chapter.

2.3. IDPs: a simple model for chemists

IDPs, as their name suggests, these proteins don't express a folded macroscopic structure and behave as extended polymer chains in solution.[118,119] Nonetheless, they play major structural and functional roles in living organisms which are reflected in their primary structure.[120,121] They also play an important part in some notable human diseases.[121,122] Unfolded IDPs mainly consist of random coil segments whose torsion angles don't have fixed values like common secondary structures.[123] Instead their amino acids can access any of the allowed regions of the dihedral angles in the Ramachandran space (**Figure 2c.**).[124] IDPs are also characterized by low complexity sequences,[120,125] containing repeating patterns of amino acids that promote extended coil conformations and disfavour tertiary folding. These features that prevent folding lead to highly soluble and non-aggregated proteins. The prevalence of polar amino acids (**Figure 1**) promote the solvation of the extended chains while charged residues generate repulsion between chain segments.[120,126] Proline and glycine, which disrupt the secondary structure, are also commonly found in IDPs.[127] Recent studies have also shown that large portions of the biologically active proteome are unstructured. These are the intrinsically disordered protein regions (IDPR).[121,128]

The simple sequence of IDPs and the lack of higher order structuring makes them prime candidates to be mimicked by synthetic polymers. In the following chapters, the possibility of producing polymers that can mimic the structural and functional capabilities of certain IDPs will be explored and discussed.

3. Conclusions

Mimicking the structural and functional abilities of nature's most versatile biomacromolecules, proteins, is long thought after by polymer chemists. It is nonetheless not an easy feat considering the structural complexity of proteins. One interesting class of proteins, IDPs, are characterized by their lack of higher order structure and behave as free polymer chains in solution. Their sequence of amino acids, or the primary structure, is also relatively simple compared to other regions of the proteome and are characterized by highly repeating patterns. These simple protein structures could be an easy step to mimic the biomacromolecules through polymerization. To this end, the ring-opening polymerization of NCAs is employed to produce high molecular weight and controlled amino-acid based polymers. Recent developments in the synthesis of

NCA and subsequent purification allow to produce highly defined polymeric structures. Many polymerization mechanisms and catalytic systems proved efficient in the ROP of NCA. Interestingly, in most of the above-mentioned studies no statistical copolypeptides were investigated and the effect of initiation type and/or catalysis has yet to be properly investigated. In chapter **2**, the stochastic control of the monomer distribution in copolymerization systems will be discussed and focused on the ROP of NCAs. A model system was developed and the effect of several reaction conditions including catalysis on the primary structure were explored. The difference in monomer distribution was also reflected in the difference of macroscopic properties of copolymers including their secondary structures. Chapter **3** introduces thermoresponsive polypeptides as simplified analogues of IDPs. Indeed, polyproline was found to produce a thermoresponse with unusually broad hysteresis that was comprehensively studied. These findings led us to copolymerize proline with other monomers to fine tune the thermoresponsive behaviour. This is discussed in chapter **4**, applying the combined knowledge from the two previous studies towards synthetic analogues of IDP proteins derived from elastin. Indeed, the final goal of these three chapters was to prepare synthetic analogues of the elastin-like polypeptides (ELP), a thermoresponsive class of IDPs that can be obtained by genetic engineering and that holds many promises for biomedical applications.

4. References

(1) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Peter, W. ; "*Molecular Biology of the Cell*"; 4th ed.; Garland Science, **2002**.

(2) Creighton, T. ; "*Proteins: Structures and Molecular Properties*"; 2nd ed.; W. H. Freeman, **1993**.

(3) Martin, R. B. *Met. Ions Biol. Syst.* **2001**, *38*, 1–23.

(4) Hol, W. G. J. *Prog. Biophys. Mol. Biol.* **1985**, *45* (3), 149–195.

(5) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7* (1), 95–99.

(6) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B.; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *Protein Sci.* **2018**, *27* (1), 293–315.

(7) Mattos, C.; Petsko, G. A.; Karplus, M. *J. Mol. Biol.* **1994**, *238* (5), 733–747.

(8) Jaenicke, R. *J. Biotechnol.* **2000**, *79* (3), 193–203.

(9) Deming, T. J. *Nat. Mater.* **2010**, *9* (7), 535–536.

(10) Ruan, Z.; Li, S.; Grigoropoulos, A.; Amiri, H.; Hilburg, S. L.; Chen, H.; Jayapurna, I.; Jiang, T.; Gu, Z.; Alexander-Katz, A.; Bustamante, C.; Huang, H.; Xu, T. *Nature* **2023**, *615* (7951), 251–258.

(11) van Hest, J. C. M.; Tirrell, D. A. *Chem. Commun.* **2001**, *19*, 1897–1904.

(12) Rosano, G. L.; Ceccarelli, E. A. *Front. Microbiol.* **2014**, *5* (APR), 1–17.

(13) Rasines Mazo, A.; Allison-Logan, S.; Karimi, F.; Chan, N. J. A.; Qiu, W.; Duan, W.; O'Brien-Simpson, N. M.; Qiao, G. G. *Chem. Soc. Rev.* **2020**, *49* (14), 4737–4834.

(14) Deng, C.; Wu, J.; Cheng, R.; Meng, F.; Klok, H. A.; Zhong, Z. *Prog. Polym. Sci.* **2014**, *39* (2), 330–364.

(15) Mitchell, A. R. *Biopolym. - Pept. Sci. Sect.* **2008**, *90* (3), 175–184.

(16) Liu, Y.; Li, D.; Ding, J.; Chen, X. *Chinese Chem. Lett.* **2020**, *31* (12), 3001–3014.

(17) Fuchs, F. *Berichte der Dtsch. Chem. Gesellschaft (A B Ser.* **1922**, *55* (9), 2943–2943.

(18) Farthing, A. C. *J. Chem. Soc.* **1950**, 3213–3217.

(19) Leuchs, H. *Berichte der Dtsch. Chem. Gesellschaft* **1906**, *39* (1), 857–861.

(20) Leuchs, H.; Geiger, W. *Berichte der Dtsch. Chem. Gesellschaft* **1908**, *41* (2), 1721–1726.

(21) Katchalski, E.; Sela, M. *Adv. Protein Chem.* **1958**, *13* (C), 243–492.

(22) Kricheldorf, H. R.; Hull, W. E. *Macromolecules* **1980**, *13* (1), 87–95.

(23) Salas-Ambrosio, P.; Tronnet, A.; Badreldin, M.; Ji, S.; Lecommandoux, S.; Harrisson, S.; Verhaeghe, P.; Bonduelle, C. *Polym. Chem.* **2022**, *13* (43), 6149–6161.

(24) Skoulas, D.; Stavroulaki, D.; Santorinaios, K.; Iatrou, H. *Polymers (Basel).* **2017**, *9* (11), 564.

(25) Hayakawa, T.; Kondo, Y.; Yamamoto, H.; Murakami, Y. *Bull. Chem. Soc. Jpn.* **1969**, *42* (2), 479–482.

(26) Hofmann, K.; Rheiner, A.; Peckham, W. D. *J. Am. Chem. Soc.* **1953**, *75* (23), 6083–6084.

(27) Detwiler, R. E.; Schlirf, A. E.; Kramer, J. R. *J. Am. Chem. Soc.* **2021**, *143* (30), 11482–11489.

(28) Gkikas, M.; Iatrou, H.; Thomaidis, N. S.; Alexandridis, P.; Hadjichristidis, N. *Biomacromolecules* **2011**, *12* (6), 2396–2406.

(29) Yu, M.; Nowak, A. P.; Deming, T. J.; Pochan, D. J. *J. Am. Chem. Soc.* **1999**, *121* (51), 12210–12211.

(30) Mobashery, S.; Johnston, M. *J. Org. Chem.* **1985**, *50* (12), 2200–2202.

(31) Hirschmann, R.; Schwam, H.; Strachan, R. G.; Schoenewaldt, E. F.; Barkemeyer, H.; Miller, S. M.; Conn, J. B.; Garsky, V.; Veber, D. F.; Denkewalter, R. G. *J. Am. Chem. Soc.* **1971**, *93* (11), 2746–2754.

(32) Bossion, A.; Nicolas, J. *Eur. Polym. J.* **2020**, *140* (July), 110033.

(33) Akssira, M.; Boumzebra , M.; Kasmi, H; Dahdouh, A; Roumestant, M L; Viallefont, P. *Tetrahedron* **1994**, *50* (30), 9051–9060.

(34) Ohkawa, K.; Nagai, T.; Nishida, A.; Yamomoto, H. *J. Adhes.* **2009**, *85* (11), 770–791.

(35) ZALIPSKY, S.; ALBERICIO, F.; SLOMCZYNSKA, U.; BARANY, G. *Int. J. Pept. Protein Res.* **1987**, *30* (6), 740–783.

(36) Devarayan, K.; Nakagami, S.; Suzuki, S.; Yuki, I.; Ohkawa, K. *Polymers (Basel).* **2020**, *12* (2), 1–15.

(37) Zhang, Z.; Su, K.; Li, Z. *Org. Lett.* **2019**, *21* (3), 749–752.

(38) Kricheldorf, H. R. ; "*α-Aminoacid-N-Carboxy-Anhydrides and Related Heterocycles*"; 1st ed.; Springer Berlin Heidelberg, **1987**.

(39) Marcantoni, E.; Massaccesi, M.; Torregiani, E.; Bartoli, G.; Bosco, M.; Sambri, L. *J. Org. Chem.* **2001**, *66* (12), 4430–4432.

(40) Sarkar, A.; Roy, S. R.; Parikh, N.; Chakraborti, A. K. *J. Org. Chem.* **2011**, *76* (17), 7132–7140.

(41) Hioki, K.; Kinugasa, M.; Kishimoto, M.; Fujiwara, M.; Tani, S.; Kunishima, M. *Synthesis (Stuttg).* **2006**, No. 12, 1931–1933.

(42) Ibrahim, T. S.; Tala, S. R.; El-Feky, S. A.; Abdel-Samii, Z. K.; Katritzky, A. R. *Synlett* **2011**, No. 14, 2013–2016.

(43) Eckert, H.; Forster, B. *Angew. Chemie Int. Ed. English* **1987**, *26* (9), 894–895.

(44) Yang, J.; Han, X.; Zhou, L.; Xiong, C. *Asian J. Chem.* **2011**, *23* (4), 1615–1617.

(45) Ganiu, M. O.; Nepal, B.; Van Houten, J. P.; Kartika, R. *Tetrahedron* **2020**, *76* (47), 131553.

(46) INRS. *Phosgène Fiche toxicologique n° 72*.

(47) OPCW. *Annex on Chemicals Schedule 3*.

(48) Wilder, R.; Mobashery, S. *J. Org. Chem.* **1992**, *57* (9), 2755–2756.

(49) Daly, W. H.; Poché, D. *Tetrahedron Lett.* **1988**, *29* (46), 5859–5862.

(50) Oya, M.; Katakai, R.; Nakai, H.; Iwakura, Y. *Chem. Lett.* **1973**, *2* (11), 1143–1144.

(51) Fujita, Y.; Koga, K.; Kim, H.-K.; Wang, X.-S.; Sudo, A.; Nishida, H.; Endo, T. *J. Polym. Sci. Part A Polym. Chem.* **2007**, *45* (22), 5365–5370.

(52) Koga, K.; Sudo, A.; Endo, T. *J. Polym. Sci. Part A Polym. Chem.* **2010**, *48* (19), 4351–4355.

(53) Sugimoto, T.; Kuwahara, T.; Liang, F.; Wang, H.; Tsuda, A. *ACS Omega* **2022**,

7 (43), 39250–39257.

(54) Collet, H.; Bied, C.; Mion, L.; Taillades, J.; Commeyras, A. *Tetrahedron Lett.* **1996**, *37* (50), 9043–9046.

(55) Vayaboury, W.; Giani, O.; Collet, H.; Commeyras, A.; Schué, F. *Amino Acids* **2004**, *27* (2), 161–167.

(56) Boiteau, L.; Collet, H.; Lagrille, O.; Taillades, J.; Vayaboury, W.; Giani, O.; Schué, F.; Commeyras, A. *Polym. Int.* **2002**, *51* (10), 1037–1040.

(57) Curtius, T.; Hochschwender, K.; Meier, H.; Lehmann, W.; Benckiser, A.; Schenck, M.; Wirbatz, W.; Gaier, J.; Mühlhäusser, W. *J. für Prakt. Chemie* **1930**, *125* (1), 211–302.

(58) Laconde, G.; Amblard, M.; Martinez, J. *Org. Lett.* **2021**, *23* (16), 6412–6416.

(59) Tran, T. V.; Shen, Y.; Nguyen, H. D.; Deng, S.; Roshandel, H.; Cooper, M. M.; Watson, J. R.; Byers, J. A.; Diaconescu, P. L.; Do, L. H. *Green Chem.* **2022**, 9245–9252.

(60) Kramer, J. R.; Deming, T. J. *Biomacromolecules* **2010**, *11* (12), 3668–3672.

(61) Tian, Z. Y.; Zhang, Z.; Wang, S.; Lu, H. *Nat. Commun.* **2021**, *12* (1), 1–11.

(62) Otake, Y.; Nakamura, H.; Fuse, S. *Angew. Chemie - Int. Ed.* **2018**, *57* (35), 11389–11393.

(63) Semple, J. E.; Sullivan, B.; Sill, K. N. *Synth. Commun.* **2017**, *47* (1), 53–61.

(64) Song, Z.; Tan, Z.; Cheng, J. *Macromolecules* **2019**, *52* (22), 8521–8539.

(65) Vacogne, C. D.; Schlaad, H. *Polymer (Guildf).* **2017**, *124*, 203–209.

(66) Gradišar, Š.; Žagar, E.; Pahovnik, D. *ACS Macro Lett.* **2017**, *6* (6), 637–640.

(67) Zhang, X.; Oddon, M.; Giani, O.; Monge, S.; Robin, J. J. *Macromolecules* **2010**, *43* (6), 2654–2656.

(68) Aliferis, T.; Iatrou, H.; Hadjichristidis, N. *Biomacromolecules* **2004**, *5* (5), 1653–1656.

(69) Habraken, G. J. M.; Wilsens, K. H. R. M.; Koning, C. E.; Heise, A. *Polym. Chem.* **2011**, *2* (6), 1322.

(70) Zou, J.; Fan, J.; He, X.; Zhang, S.; Wang, H.; Wooley, K. L. *Macromolecules* **2013**, *46* (10), 4223–4226.

(71) Deming, T. J. *Adv. Mater.* **1997**, *9* (4), 299–311.

(72) Dimitrov, I.; Schlaad, H. *Chem. Commun.* **2003**, *3* (23), 2944–2945.

(73) Zhao, W.; Gnanou, Y.; Hadjichristidis, N. *Biomacromolecules* **2015**, *16* (4), 1352–1357.

(74) Zhao, W.; Gnanou, Y.; Hadjichristidis, N. *Chem. Commun.* **2015**, *51* (17), 3663–3666.

(75) Vacogne, C. D.; Schlaad, H. *Chem. Commun.* **2015**, *51* (86), 15645–15648.

(76) Zhao, W.; Lv, Y.; Li, J.; Feng, Z.; Ni, Y.; Hadjichristidis, N. *Nat. Commun.* **2019**, *10* (1), 3590.

(77) Zhao, W.; Gnanou, Y.; Hadjichristidis, N. *Polym. Chem.* **2015**, *6* (34), 6193–6201.

(78) Liang, J.; Zhi, X.; Zhou, Q.; Yang, J. *Polymer (Guildf).* **2019**, *165* (November 2018), 83–90.

(79) Zhang, Y.; Liu, R.; Jin, H.; Song, W.; Augustine, R.; Kim, I. *Commun. Chem.* **2018**, *1* (1), 1–7.

(80) Zhang, H.; Nie, Y.; Zhi, X.; Du, H.; Yang, J. *Chem. Commun.* **2017**, *53* (37), 5155–5158.

(81) Nie, Y.; Zhi, X.; Du, H.; Yang, J. *Molecules* **2018**, *23* (4).

(82) Gowda, R. R.; Chakraborty, D. *J. Mol. Catal. A Chem.* **2010**, *333* (1–2), 167–172.

(83) Goodwin, A. A.; Bu, X.; Deming, T. J. *J. Organomet. Chem.* **1999**, *589* (1), 111–114.

(84) Wu, Y.; Zhang, D.; Ma, P.; Zhou, R.; Hua, L.; Liu, R. *Nat. Commun.* **2018**, *9* (1), 1–10.

(85) Lu, H.; Cheng, J. *J. Am. Chem. Soc.* **2007**, *129*, 14114–14115.

(86) Lu, H.; Cheng, J. *J. Am. Chem. Soc.* **2008**, *130* (38), 12562–12563.

(87) Yuan, J.; Sun, Y.; Wang, J.; Lu, H. *Biomacromolecules* **2016**, *17* (3), 891–896.

(88) Yuan, J.; Zhang, Y.; Li, Z.; Wang, Y.; Lu, H. *ACS Macro Lett.* **2018**, *7* (8), 892–897.

(89) Chan, B. A.; Xuan, S.; Horton, M.; Zhang, D. *Macromolecules* **2016**, *49* (6), 2002–2012.

(90) Lv, W.; Wang, Y.; Li, M.; Wang, X.; Tao, Y. *J. Am. Chem. Soc.* **2022**.

(91) Yamashita, S.; Tani, H. *Macromolecules* **1974**, *7* (4), 406–409.

(92) Deming, T. J. *Nature* **1997**, *390* (6658), 386–389.

(93) Deming, T. J. *Macromolecules* **1999**, *32* (13), 4500–4502.

(94) Deming, T. J. *J. Am. Chem. Soc.* **1998**, *120* (17), 4240–4241.

(95) Deming, T. J.; Curtin, S. A.; Barbara, S. **2000**, No. eq 2, 5710–5717.

(96) Curtin, S. A.; Deming, T. J. **1999**, No. 8, 7427–7428.

(97) Seidel, S. W.; Deming, T. J. **2003**, *1* (Table 1), 969–972.

(98) Peng, Y. L.; Lai, S. L.; Lin, C. C. *Macromolecules* **2008**, *41* (10), 3455–3459.

(99) Peng, H.; Ling, J.; Shen, Z. *J. Polym. Sci. Part A Polym. Chem.* **2012**, *50* (6), 1076–1085.

(100) Bhaw-Luximon, A.; Jhurry, D.; Belleney, J.; Goury, V. *Macromolecules* **2003**,

36 (4), 977–982.

(101) Hadjichristidis, N.; Iatrou, H.; Pitsikalis, M.; Sakellariou, G. *Chem. Rev.* **2009**, *109* (11), 5528–5578.

(102) Coleman, D.; Farthing, A. C. *J. Chem. Soc.* **1950**, No. 3218, 3218.

(103) Wu, Y.; Chen, K.; Wu, X.; Liu, L.; Zhang, W.; Ding, Y.; Liu, S.; Zhou, M.; Shao, N.; Ji, Z.; Chen, J.; Zhu, M.; Liu, R. *Angew. Chemie Int. Ed.* **2021**, *60* (50), 26063–26071.

(104) Xia, Y.; Song, Z.; Tan, Z.; Xue, T.; Wei, S.; Zhu, L.; Yang, Y.; Fu, H.; Jiang, Y.; Lin, Y.; Lu, Y.; Ferguson, A. L.; Cheng, J. *Nat. Commun.* **2021**, *12* (1), 732.

(105) Song, Z.; Fu, H.; Wang, J.; Hui, J.; Xue, T.; Pacheco, L. A.; Yan, H.; Baumgartner, R.; Wang, Z.; Xia, Y.; Wang, X.; Yin, L.; Chen, C.; Rodríguez-López, J.; Ferguson, A. L.; Lin, Y.; Cheng, J. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *166* (22), 10658–10663.

(106) Song, Z.; Fu, H.; Baumgartner, R.; Zhu, L.; Shih, K. C.; Xia, Y.; Zheng, X.; Yin, L.; Chipot, C.; Lin, Y.; Cheng, J. *Nat. Commun.* **2019**, *10* (1), 1–7.

(107) Grazon, C.; Salas-Ambrosio, P.; Ibarboure, E.; Buol, A.; Garanger, E.; Grinstaff, M. W.; Lecommandoux, S.; Bonduelle, C. *Angew. Chemie - Int. Ed.* **2020**, *59* (2), 622–626.

(108) Becker, R. R.; Stahmann, M. A. *J. Biol. Chem.* **1953**, *204* (2), 737–744.

(109) Hurd, C. D.; Buess, C. M. *J. Am. Chem. Soc.* **1951**, *73* (6), 2409–2412.

(110) Ehler, K. W.; Orgel, L. E. *BBA - Protein Struct.* **1976**, *434* (1), 233–243.

(111) Kamei, Y.; Nagai, A.; Sudo, A.; Nishida, H.; Kikukawa, K.; Endo, T. *J. Polym. Sci. Part A Polym. Chem.* **2008**, *46* (8), 2649–2657.

(112) Kamei, Y.; Sudo, A.; Nishida, H.; Kikukawa, K.; Endo, T. *J. Polym. Sci. Part A Polym. Chem.* **2008**, *46* (7), 2525–2535.

(113) Yamada, S.; Koga, K.; Sudo, A.; Goto, M.; Endo, T. *J. Polym. Sci. Part A Polym. Chem.* **2013**, *51* (17), 3726–3731.

(114) Yamada, S.; Sudo, A.; Goto, M.; Endo, T. *J. Polym. Sci. Part A Polym. Chem.* **2015**, *53* (6), 829–829.

(115) Akbulut, H.; Ando, S.; Yamada, S.; Endo, T. *J. Polym. Sci. Part A Polym. Chem.* **2018**, *56* (22), 2522–2530.

(116) Yang, Z.; Bai, T.; Ling, J.; Shen, Y. *J. Polym. Sci. Part A Polym. Chem.* **2019**, *57* (8), 907–916.

(117) Li, L.; Cen, J.; Pan, W.; Zhang, Y.; Leng, X.; Tan, Z.; Yin, H.; Liu, S. *Research* **2021**, *2021*, 1–16.

(118) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106* (5), 1877–1897.

(119) Alston, J. J.; Ginell, G. M.; Soranno, A.; Holehouse, A. S. *J. Phys. Chem. B* **2023**, *127* (21), 4746–4760.

(120) Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. *Proc. Natl. Acad. Sci.* **2010**, *107* (33), 14609–14614.

(121) Theillet, F.-X.; Kalmar, L.; Tompa, P.; Han, K.-H.; Selenko, P.; Dunker, A. K.; Daughdrill, G. W.; Uversky, V. N. *Intrinsically Disord. Proteins* **2013**, *1* (1), e24360.

(122) Schweitzer-Stenner, R. *Mol. Biosyst.* **2012**, *8* (1), 122–133.

(123) Radivojac, P.; Iakoucheva, L. M.; Oldfield, C. J.; Obradovic, Z.; Uversky, V. N.; Dunker, A. K. *Biophys. J.* **2007**, *92* (5), 1439–1456.

(124) Serrano, L. *J. Mol. Biol.* **1995**, *254* (2), 322–333.

(125) Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. *Proc. Natl. Acad. Sci.* **2010**, *107* (18), 8183–8188.

(126) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. *Proteins Struct. Funct. Genet.* **2000**, *41* (3), 415–427.

(127) Imai, K.; Mitaku, S. *Biophysics (Oxf).* **2005**, *1*, 55–65.

(128) Rath, A.; Davidson, A. R.; Deber, C. M. *Biopolym. - Pept. Sci. Sect.* **2005**, *80* (2–3), 179–185.

Chapter II: Disordered polypeptide polymers

1. Introduction

As depicted in chapter **1**, intrinsically disordered proteins (IDPs) are a special class of proteins without an organized secondary or tertiary structure. The goal of this chapter is to show how such intrinsically disordered polypeptides could be produced using simple synthetic chemistry and copolymerization processes. As in IDPs, the amino acid sequence explains the lack of structuring, our objective was to show that controlling the sequence distribution of monomers inside copolymers could induce the same structural feature. In many IDPs, glycine residues create steric barriers to the appearance of secondary structure and glycine-rich sequences are associated with "intrinsically disordered regions" of the proteome. Inspired by this feature, we have studied how the introduction of glycine residues influences the secondary structure of a model polypeptide, poly(γ-benzyl-*L*-glutamate) or PBLG, a helical polymer. For this purpose, it is important to understand how the comonomers react and insert themselves in a polymer chain using kinetic studies of the copolymerization system. That is why we chose to use two N-carboxyanhydride (NCA) monomers, γ-benzyl-*L*-glutamate and glycine NCA to undergo ring opening copolymerization in which the monomer distribution could be controlled by changing the reaction conditions (temperature, solvent, etc.). The relationship between the glycine distribution and secondary structure evidenced the importance of these reaction conditions and ultimately afforded “intrinsically disordered PBLG-based copolymers” (**Figure 1**).

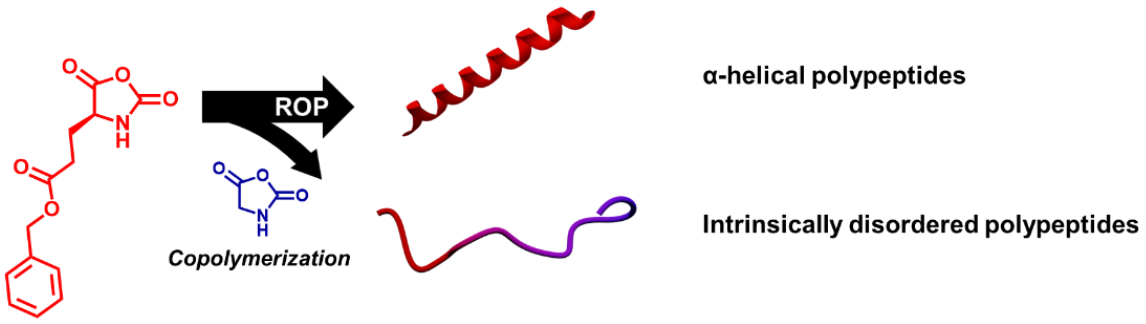


***Figure 1**: Schematic representation of the goal of this chapter. ROP of BLG NCAs (in red) give rise to structured polymers in the α-helical form. By introducing glycine units through copolymerization with glycine NCAs (in blue), intrinsically disordered polyglutamates are targeted.*

1.1. Sequence control in polymer science

In proteins, amino acids are linearly arranged along the macromolecular chain in what is known as the **primary sequence.** This structural arrangement is the fundamental driver of biological activity, pivotal for catalytic processes, binding affinities, and structural formations of most proteins. The organization of distinct amino acids within a polypeptide chain underpins the emergence of various natural proteins and their functionalities. Polymer chemists have long sought to reproduce this control over the primary structure of synthetic macromolecules. Control over monomer insertion would be a powerful tool for tailoring the properties and functions of the ensuing copolymers. Many examples of how to tailor the distribution of different repeat units along the copolymer chains have been developed. These advances include sequence regulation in classical polymerization processes (step- and chain-growth), or methods inspired by other fields such as iterative synthesis (biochemistry) and template polymerization (biology).[1,2] While these synthetic macromolecules do not yet match their natural counterparts in terms of functional capabilities, they introduce novel capabilities for their self-assembly processes, and the macroscopic traits exhibited by polymer materials.[3] In the following, only the most useful and pertinent works will be cited, with a focus on examples of polymers with amino acids in the backbone.

1.1.1. Iterative method

In the iterative method much like in natural processes, one monomer unit is added at a time in the intended order. To achieve acceptable yields, a solid phase is generally

employed and decorated with reactive sites. The monomers used must have two reactive sites, one of which is protected. The first step is the reaction between the first type of monomer, usually reacted in molar excess, and the reactive sites of the solid phase. This reaction must be fast and able to reach 100% conversion. The second step is to wash the unreacted monomers leaving all reactive sites with one connected monomer having a protected group. The third step is the deprotection of the attached monomer. The cycle can now be repeated by the addition of a second type of monomer. Repeating this cycle n times will give polymers with n repeat units, all with in theory the same sequence and length (**Figure 2**). This protocol was introduced by Merrifield and is known as solid phase peptide synthesis (SPPS).[4]

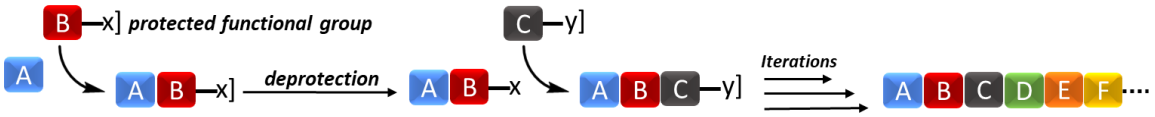


Figure 2: *Iterative addition of monomers on a growing polymer chain. This method could be done on solid support (here would be A unit) and the growing polymer would be cleaved at the end steps.*

This method can be completely automated through robotization providing a high throughput process,[5] but has some drawbacks such as the high consumption of toxic solvent.[6] Furthermore, it is ideally used to produce low molecular weight (MW) polymers that are defined in sequence and monodisperse but is less effective for higher molecular masses. This is because some chains can stop being reactive especially as a result of their collapse or folding.[7,8] This leads to an increase in the MW dispersity and deviations from the desired sequence. Bioactive peptides are often prepared via this pathway. SPPS has also been applied to many types of polymers including synthetic DNA[9], polyesters[10], polyurethanes[11], polyamides[11,12] and polypeptoids[6]. Polymers prepared in this way may find applications in data storage.[2] Yet much work is still needed for this field to be mature enough for applicative use.[13] Automation has helped reduce wasted solvent, while alternative “greener” solvents have proven useful.[4] Although the new developments in this field allow the preparation of higher MW polymers,[3] the current work will focus on more classical polymerization techniques that easily accesses these longer chains.

1.1.2. Batch methods

Although classical polymerization techniques produce long polymeric chains, they still lack the ability to produce sequence-defined polymers. Nonetheless, some tools are available to polymer chemists to be able to induce a sequence in a copolymer.

Monomer engineering method

One way to produce sequence-controlled polymers is by designing monomers that will bring a periodic sequence once polymerized.

Figure 3: *Examples of how a tailored synthesis of a monomer and the choice of the polymerization pathway can induce a sequence in the resulting polymer. On the left: cyclooctene bearing different functional groups undergoes ring-opening metathesis polymerization (ROMP) to produce a periodic copolymer. On the right: designed monomers bearing polymerizable functional groups (*) can give similar periodic copolymers.*

Regioselective ring-opening metathesis polymerization (ROMP) (**Figure 3**) is one way to produce such copolymers. Large monomers, such as cyclooctenes bearing three functional groups, have been polymerized using Grubbs' catalyst to ensure regioselectivity.[14] The product is equivalent to a periodic copolymer of the three functional monomers. Similarly, monomers bearing an unconjugated carbon double bond on one side and a reactive C-Cl on the other can produce periodic copolymers. *Kamigaito*'s group have shown this possibility using tailored monomers that combine vinyl chloride with different styrenes and acrylates.[15] Different copper catalysts were used to produce periodic copolymers of vinyl chloride-*co*-styrene-*co*-acrylate (**Figure 3**). This is not possible with classical radical copolymerization of the three comonomers. Another example is the ROP of morpholine-2,5-diones to produce perfectly alternating poly(ester-amide)s (**Figure 4**). Although this class of polymer can be obtained by polycondensation, ROP offers the possibility of producing high MW polymers with controlled morphology and low dispersity.[16]

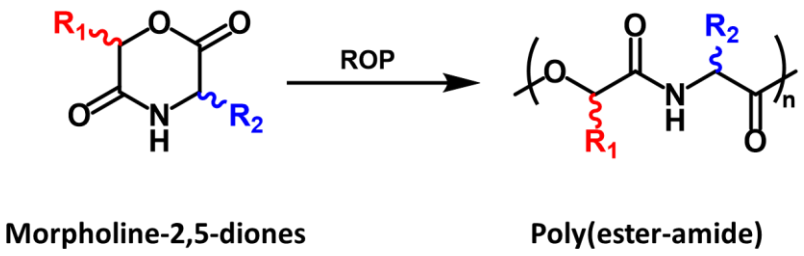


Figure 4: *ROP of morpholine-2,5-diones and the resulting poly(ester-amide).*

The polymers produced with this pathway are both enzymatically degradable and hydrolysable through the amide and ester linkages respectively. Nonetheless, the synthesis of the monomers is cumbersome. Furthermore, the requirement for specific catalysts and techniques as shown above limits the possible monomers.

Living polymerizations: reactivity and kinetics

Using controlled and living polymerization, polymers with well-defined structures can be produced. In contrast to the iterative method, both step-growth and chain-growth polymerizations produce polymers with a distribution of MW/structure/composition. In general, these distributions are governed by a statistical nature related to the reaction conditions and reactivities of the monomers involved.[17] Different monomers typically react at different rates and the final product contains a statistical distribution of the monomers along the different polymer chains. The first step towards understanding this distribution is to calculate the reactivity ratios of the comonomers. Taking the example of a copolymerization between monomer A and monomer B, the reactivity ratio of monomer A noted r_A is equal to the ratio between the reaction constants k_{AA} and k_{AB}. The first corresponds to the reaction of a newly incorporated monomer A in the growing polymer chain with similar monomer (A) from the reaction media. The second corresponds to its reaction with the other comonomer (B). Thus, at equimolar ratio, an A monomer at the end of the polymer chain will more likely react with a similar monomer if $r_A > 1$. On the other hand, if $r_A < 1$ it will more likely react with the comonomer B inducing the alternating sequence AB (**Figure 5**).

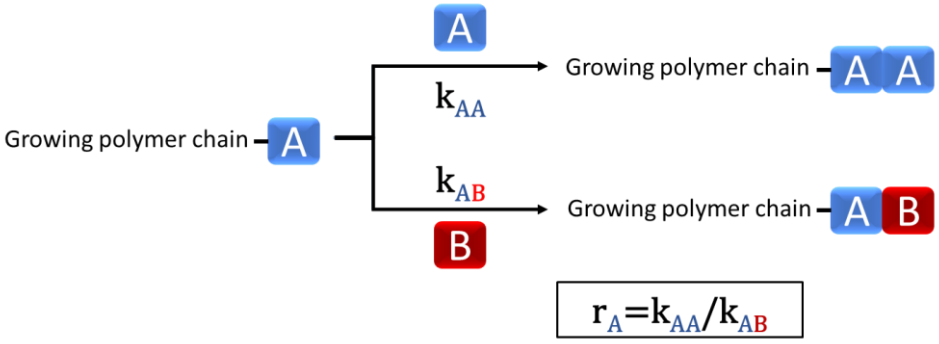


Figure 5: *Illustration of the concept of reactivity ratio of a pair of comonomers A and B. The reactivity ratio of comonomer A will be equal to the ratio between the rate coefficients of a monomer A at the growing chain-end to react with either a similar monomer or with the opposing comonomer respectively.*

The concept of reactivity ratios can also be extrapolated to terpolymerizations and more complex systems.[18] The Mayo-Lewis equation[19] (**Equation 1**) describes the relative rate at which the two monomers A and B incorporate into the copolymer as a function of their reactivity ratios.

$$\frac{d[M_A]}{d[M_B]} = \frac{[M_A]}{[M_B]} \times \frac{r_A.[M_A]+[M_B]}{[M_A]+r_B.[M_B]} \qquad (1)$$

$$f_A = \frac{[M_A]}{[M_A]+[M_B]} = 1 - f_B \qquad (2)$$

$$\frac{df_A}{df_B} = \frac{r_A.f_A^2+f_A.f_B}{r_B.f_B^2+f_B.f_A} = \frac{F_A}{F_B} \qquad (3)$$

Equation 1: *The Mayo-Lewis equation expressed in function of monomer concentration [M_X] and instantaneous monomer feeds (f_X). F_X is the copolymer composition in monomer X.*

The composition of the produced copolymers can be estimated by considering the reactivity ratios and the initial monomer feed for a given set of reaction conditions. The probability of finding certain sequences inside the polymer chain can also be calculated. The final copolymer microstructure will fall into one of four categories depending the reactivity ratios of comonomers A and B: multiblock, alternating, statistical or gradient copolymers (**Figure 6**).[20]

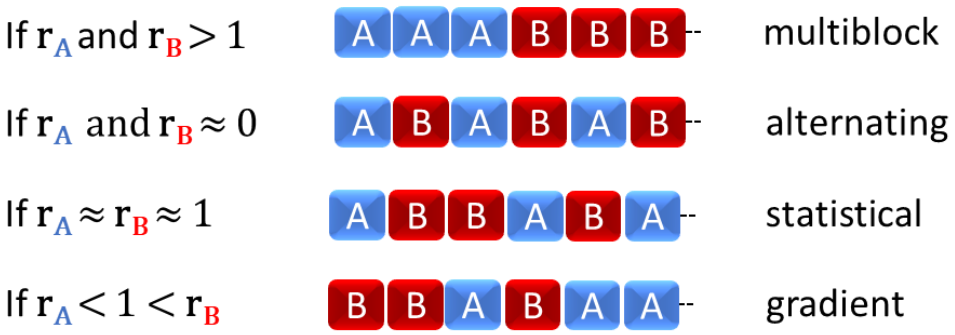


Figure 6: *Schematic representation of the microstructure of copolymers comprising of two monomers A and B depending on the relative values of their reactivity ratios.*

It is important to keep in mind that the polymer chains formed will be disperse (different chain lengths) and non-uniform in their primary sequence (the order of the monomers in the chain is different from one chain to the other).[3,21] Thus the overall composition, much like the chain length, is reflected in a statistical manner in all the chains. This makes it impossible to determine a single primary sequence.

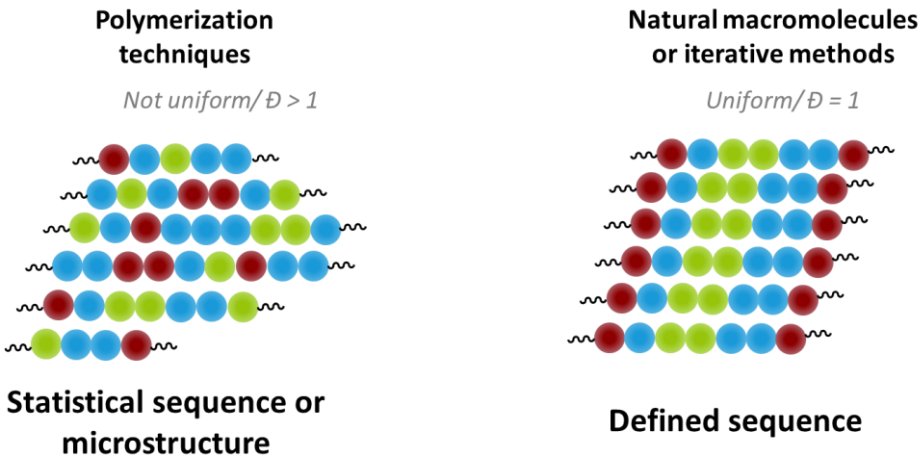


Figure 7: *Differences between a copolymer produced via classic polymerization techniques and macromolecules produced by biotechnology, iterative methods or found in nature.*

By varying the reactivity ratios, a certain stochastic control over the distribution of monomers along the chain can be exerted. In most cases reactivity ratios are relatively insensitive to reaction conditions, as rate constants will vary similarly to changes in solvent polarity or temperature. Nonetheless, it is possible to change the reactivity

ratios' values and resulting copolymer microstructures by varying reaction conditions like solvent polarity[22,23] and temperature[24,25].[26]

Semi-batch methods can be used to exert a higher degree of control over the sequence of the copolymers. This can be done by taking two comonomers with reactivity ratios close to zero. Undertaking a homopolymerization of monomer A while following its kinetics, one can add a single unit of monomer B at a certain point during the polymerization. Since an A unit at the growing chain end will most definitely react with the monomer B that was just added, all the chains will have one monomer B inserted at that position of the polymer inducing a discrete sequence along the chain. Nonetheless, due to the statistical nature of such copolymerizations there is bound to be a structural variation between the chains (**Figure 7**).[27] This method has been for example used for copolymerization of styrene and maleimides.[1,28] Continuous homogenous feed of the comonomers in reaction or by varying the feed ratios along the polymerization can also be used to better control the statistical distribution of monomers along the chain.

1.1.3. Sequence control in NCA copolymerization

As reviewed in chapter **1**, the most economical and efficient route to prepare synthetic polypeptides is the ring-opening polymerization (ROP) of amino acids using N-carboxyanhydride (NCA) monomers. Compared to proteins, peptidic polymers are much simpler macromolecules in which amino acids are statistically repeated. However, they combine the advantageous features of synthetic polymers (solubility, process, elasticity, etc.) with those of natural proteins (secondary structure, functionality, biocompatibility, etc.). So far, batch copolymerization is the most direct method to exert a stochastic control over the monomer distribution along the polymer backbone through the control of reactivity ratios. As for proteins, one would expect the primary structure of polypeptide polymers to have a strong influence on their structural and functional properties. This is relevant in mimicking tandem-repeating sequences such as the ones found in IDPs, yet to our knowledge, literature on this subject is still limited.

Reactivity ratios of NCA copolymers

Recent advances in the polymerization of NCAs have greatly improved the control over the livingness and the end-chain fidelity of the polypeptides produced via this pathway.[29,30] Of the many examples of NCA copolymerization, most have explored

block or graft copolymers[31–38] Relatively few have studied the copolymerization kinetics of NCAs. These will be discussed in the following section and will also serve as a basis for development of our experimental procedures.

Low conversion calculations

Reactivity ratios are typically calculated using the Mayo-Lewis equation (**Equation 1**). For this purpose, copolymerization must be stopped at very low conversions. This can be done by undertaking multiple copolymerizations with different monomer feeds under the same reaction conditions. The reaction must be quenched at low conversion and the formed copolymers must be separated from the unreacted monomers and analysed. The fraction of each monomer incorporated in the copolymer chains gives the information needed to calculate the copolymerization parameters. The copolymer composition can be obtained in multiple ways. One possibility is to precipitate the copolymer from the excess non-reacted monomers via precipitation and filtration/washing. The copolymers can then be analysed for their content in each comonomer,[39] for example using nuclear magnetic resonance (NMR).[40] Another possibility is to hydrolyse the separated copolymers into the amino acids that are then analysed with automatic amino acid analysers to determine the copolymer composition.[41–43] Nonetheless, this technique only gives information about low molecular weight compositions as polymerizations have to be stopped at low conversions. Since polypeptides' secondary structuring is highly dependent on chain length,[44] it is important to monitor any deviation from models based on low molecular weight oligomers. Such calculations can be extended to terpolymerization systems., using the Alfrey-Goldfinger equations.[18] *Wamsley* et al. applied this technique to terpolymers of valine, leucine and β-benzyl aspartate.[45]

Copolymerization monitoring

Although the low conversion approach is widely used in radical copolymerizations, it is not necessarily applicable to copolypeptides. This is because low and high molecular weight chains may behave differently, as their macromolecular conformation may differ, which could thus affect the copolymerization kinetics. It is thus necessary to conduct at least one copolymerization at higher conversion to validate the calculated reactivity ratios.[43,45] *Heise* et al. devised a technique based on high-performance liquid chromatography (HPLC) monitoring to achieve this.[46,47] Using acid quenching, they converted the remaining unreacted NCAs to amino acids. By taking aliquots at different

conversions from copolymerization reactions, they were able to quantify the unreacted NCAs as amino acids using HPLC calibrations. This has the advantage of only analysing the small amino acids without interference from soluble oligomers. The group further extended this methodology to terpolymerization[47] and it has also been applied for hybrid systems containing NCAs and NTAs[48]. The HPLC monitoring technique was also used to ensure the quality of commercial Glatiramer copolymers.[49] Glatiramer acetate is a Food and Drug Administration (FDA) approved medication for multiple sclerosis.[50] Its active component is a copolymer of alanine, lysine, glutamate and tyrosine that is prepared via ROP of the corresponding NCAs. Unfortunately, this method is not universally applicable. Parameters for each system must be determined to detect and correctly separate the amino-acid peaks. Furthermore, tedious calibrations are necessary to obtain precise and accurate results. *In-situ* NMR monitoring is the most convenient method for such calculations, as polymerizations can be conducted directly in the NMR spectrometer using only a small amount of material. The disappearance of monomer peaks and/or the appearance of polymer peaks can then be used to monitor the monomer conversion and copolymer composition. Nonetheless, well defined peaks must also be determined for correct integration. *Mbizana* et al. used this technique to study the terpolymerization of β-benzyl-*L*-aspartate, glycine and N^{δ}-carbobenzyloxy-*L*-ornithine NCAs.[51] For the calculations they used the integrated form of the terminal copolymerization model which is adapted to high conversion. This was done using the Contour program that employs the terminal copolymerization model and a non-linear least squares curve fitting method.[52] In this work, the calculations are based on these two methods and will be detailed in the experimental section.

Beyond kinetic studies: post-polymerization characterizations

The pioneering works of *Kricheldorf* et al. using NMR spectroscopy opened the field of sequence analysis for polypeptide copolymers. They first applied the analysis of the chemical shifts of carbonyls using ^{13}C NMR previously developed for polyamide copolymers.[53] Standards were used to identify the chemical shifts corresponding to different carbons. These standards can be the homopolymers; for example, homo-polyglycine can be used to determine the chemical shift of the central glycine carbonyl in a Gly-**Gly**-Gly sequence. This can also be done for the other comonomer, thus discriminating between the two homo-sequences of the copolymer. The hetero-

sequences can then be deduced from these where Xaa-**Gly**-Xaa glycine's carbonyl will either be shielded or deshielded depending on the amino acid (Xaa). The group further used small peptides with a defined sequence to determine the hetero-sequences when using comonomers with very close carbonyl shifts. They used these model peptides to show that longer range influence was not present and thus only neighbouring residues affect the carbonyls' chemical shift.[54] By proceeding to random copolymerization of different ratios of a pair of NCA monomers, the NMR analysis of copolymers can give an estimate of the reactivity ratios and the microstructural nature of the formed copolymers. The authors identified the different peaks of the carbonyls for ^{13}C NMR. With the simplification that the carbonyls' ^{13}C chemical shift is only affected by the preceding monomer they were able to identify the following peaks: Gly-**Gly**, Ala-**Ala**, Ala-**Gly** and Gly-**Ala** using model peptides and copolymers with excess of each comonomer. Thus, in a comonomer with $f_{Gly}=2f_{Ala}$ more Gly-Gly sequence will be observed in the carbonyl region of glycine and more Gly-Ala will be observed in that of the alanine thus easily identifying each peak. The analysis of copolymers using glycine NCA and alanine NCA with equal initial feed ratio $f_{Gly}=f_{Ala}=1$ in different conditions, gave equal amounts of all four peaks indicating essentially random copolymers ($r_{Gly}\approx r_{Ala}\approx 1$).[54] This also shows the major shortcoming of using this technique as only an estimate of the reactivity ratios can be calculated in most cases. This technique also requires concentrated samples (around 200mg/ml in deuterated trifluoroacetic acid (TFA)) and long analysis times due to the low natural abundance of ^{13}C and its longer relaxation times. This is especially true for carbonyls with no Nuclear Overhauser effect (NOE) effect from adjacent protons. Nevertheless, it can be optimized to some extent with more powerful modern equipment. More accurate information and direct ratio calculations can also be obtained by studying the ^{15}N NMR of certain copolymers. This is due to better separation in the different amide protons' signals.[55,56] Nevertheless, such analysis must be done on molecules that are isotopically enriched in ^{15}N. At the time of these studies, modern analysis techniques such as size exclusion chromatography (SEC) to verify the controllability of the polymerization and their dispersities were not commonly employed. Since these works, these NMR techniques have mainly been used to validate the presence of copolymers in the final sample.[51,57,58] *Wamsley* and coauthors showed the possibility of such a technique to obtain more information about terpolymers using a statistical method based on the Poisson distribution. In their study, they investigated the terpolymerization system

consisting of valine, leucine and β-benzyl aspartate to determine the probability of having an LDV sequence.[59] Another notable technique is the use of a blob-based model and fluorescent decay analysis on polypeptides having fluorescent probes along the chains.[60] *Duhamel*'s group studied the folding of several polypeptide polymers using this technique which gave an estimate of the neighbouring amino acids to the fluorescent unit on the polymer chain and was applied to several copolypeptide systems.[61–66] It can thus give information on both the primary sequence (neighbouring amino acids on the chain) and secondary structure (neighbouring amino acids in the space).[64]

Controlling the primary sequence of polypeptide polymers

Although the primary sequence is directly related to the reactivity ratios of the comonomers, these ratios are in turn sensitive to the reaction conditions. Most of the above-mentioned studies do not show the effect of the reaction conditions on the copolymerization kinetics and only data on homopolymerizations can be found in the literature. *Kricheldorf*'s work on alanine and glycine copolymerizations using different reaction conditions (catalyst and solvent) was inconclusive[54,56] as the post-polymerization NMR monitoring used was only able to give a rough estimate of the reactivity ratios. Another group compared copolymerization results using as initiators Al-Schiff's base complexes or classical amines. The study was conducted using both low conversion calculations and post-polymerization NMR analysis.[58] Their results showed little difference between the studied systems on the copolymerization of *γ*-methylglutamate and leucine NCA even with the difference in ROP mechanism[67]. *Hayashi*'s group investigated the effect of solvent on copolymerizations of BLG and valine NCA.[43] The solvent affects the secondary structuring of the polymers which in turn affects the reactivity of the different comonomers.[39,43] Furthermore, to avoid the insolubility due to such secondary structuring and the change in kinetics, racemic mixtures of NCAs, which can disrupt secondary structure, were employed.[53] Nonetheless, copolypeptides based on *DL*-alanine and *DL*-tryptophane were found to form α-helical structures and precipitated during polymerization making for highly disperse and non-uniform copolymers.[68] On the other hand, secondary structuring was avoided when copolymerizing racemic mixtures of NCAs based on lysine and aspartic acid allowing for homogenous processes. This behaviour depended on the monomers'

initial feed ratio and only influenced the rate of copolymerization without affecting the copolymer composition.[40]

Going one step further, studies on the relationship between the primary structure, the secondary structure and the macroscopic properties are scarce. *Kricheldorf*'s group showed that the initiator type, the solvent and temperature affected the copolymerization kinetics of a binary system comprising of valine NCA and leucine NCA. By means of ^{15}N NMR analysis, the synthesized copolymers showed different average block lengths of each monomer depending on the conditions. Valine and leucine induce β-sheets and α-helice respectively in a polypeptide chain. The analysis of the secondary structure using solid state ^{13}C NMR showed that copolymers having different monomer reactivities also possess different α-helix/β-sheet ratios.[69] The group applied similar principles to study the effect of incorporating glycine into a terpolymer containing valine and alanine as an α-helix inducing monomer. The slower polymerization rate of valine led to its homopolymerization at the chain ends without full conversion. This technique does not allow differentiation between the different primary structures, but the difference in secondary structures using different reaction conditions was established. Furthermore, the secondary structure analysis led to the conclusion that while glycine is an α-helix breaking amino acid, this secondary structure can still be detected during polymerization or in the solid state. However, when the copolymers were dissolved in TFA, the conformation was more of a random coil.[70] Another group conducted a similar study on the dependence of secondary structure on the sequence of BLG and valine in their copolymers.[41] Due to the higher reactivity of BLG, the highly gradient copolymers showed β-sheet like structures consistent with short BLG sequences and valine sequences of variable length.[41]

The primary sequence can also induce differences in macromolecular properties. For example, the biodegradability of poly-α/β-peptides formed from NCA/NTA systems depended on their microstructure.[71] Furthermore, statistical copolymers of NCAs can roughly reproduce desired properties of natural proteins having periodic primary sequences, also known as tandem repeat proteins.[72] These proteins,[73] a form of IDPs, have inspired the studies of the terpolymerization systems discussed above. *Wamsley*'s group showed that the presence of the LDV sequence affects the cellular adhesion properties in the Leu, Asp and Val terpolymer system.[59] Similarly, the work of *Mbizana* et al. obtained RGD sequences from the terpolymerization of Arg, Gly and Asp NCAs giving similar cell adhesion properties.[51] In both examples, the study of the

kinetics allowed for the possibility to exert a stochastic control over the sequence of the copolymers through reaction conditions. The probability of obtaining the desired sequence was increased by choosing the appropriate feed ratios. This led to an increase in the desired cell adhesion property in both cases.

2. Results and discussion

In the following parts of this chapter, the model system developed to achieve an understanding of the copolymerization process of NCA monomers will be presented. This study will focus on two axes:

1. The effect of copolymerization conditions on the monomer insertion
2. The effect of the monomer distribution (primary sequence) on the secondary structure of the copolymers

The copolymerization of BLG NCA, an α-helix inducing monomer, and glycine NCA, a disorder inducing monomer, will comprise this model system. **As sequences of glutamate usually structure themselves in a helical conformation, the addition of glycine is expected to disrupt this conformation, producing intrinsically disordered polyglutamates.** The less sterically hindered monomer, glycine, has a higher reactivity in ROP than BLG. By means of reaction conditions, including feed ratio, temperature, solvent, catalysis, and ROP mechanism, a stochastic control over the monomer distribution can be exerted. This can be done by improving the distribution of the monomers along the chain and/or through changes in reactivity ratios. In the following parts, the ROP of BLG NCA, will be studied and the characteristics of the resulting polymer poly(γ-benzyl-*L*-glutamate) (PBLG) presented (section 2.1.1). This polymer can also be made water soluble by removing benzyl side chains to form a polypeptide bearing water soluble carboxylic side chains known as polyglutamate (PGA). In this form, the polymer may adopt either α-helix or random coil secondary structure depending on the protonation of the side chain, making it pH responsive. The effect of introducing glycine units in the PBLG/PGA chains on this secondary structuring will then be studied. For this purpose, glycine NCA is first synthesized, and its ROP briefly studied and discussed (section **2.1.2**). This was followed by the copolymerization study in which the effect of the reaction conditions on the statistical primary sequence was estimated using the terminal copolymerization model (section **2.2.1**). A correlation between this sequence and the secondary structure

on water soluble copolymers will also be detailed (section 2.2.3). Thus, presenting a model where the stochastic control of the primary sequence of copolypeptides can help steer their macromolecular conformation and other properties like solubility.

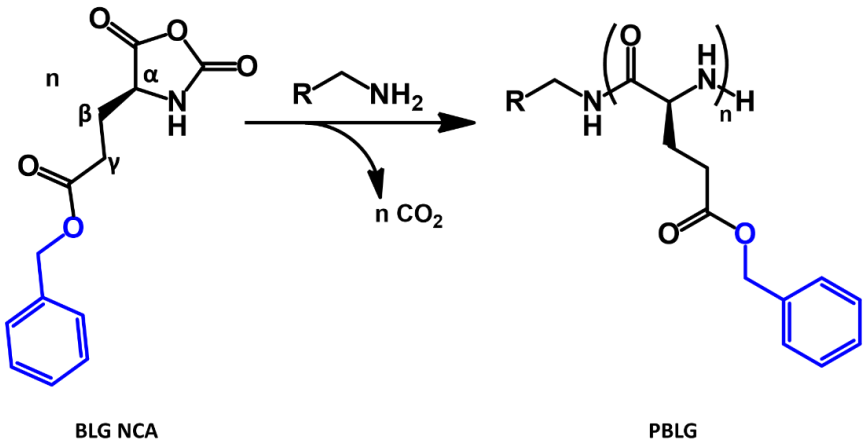
2.1. Preparation of the model copolymerization system

2.1.1. BLG NCA

Glutamate is the highest produced amino acid industrially[74] due to its use in the food sector[75]. This makes its derivatives, like BLG NCA, readily available commercially. Furthermore, the study of its polymers, PBLG, and its deprotected form polyglutamic acid (PGA), a water soluble and pH-responsive polymer,[76,77] have been at the heart of polypeptide polymers' advances. Thus BLG/PGA based polymers[31,75,78–81] (PBLG/PGA) and copolymers[82,83] are widely studied and have potential applications in bioactive substance delivery,[84–90] as antibacterial agents[91] and as nanomaterials[92,93]. The functionality of the carboxylic acid/ester side-chain has made PBLG based polymers even more versatile[90,94–96] and employed for other applications such as temperature responsive or energy storage[97] materials. Almost all published examples of ROP and catalysis use BLG NCA as a benchmark. In the following sections are presented, the implementation of the ROP of pure BLG-NCA, the deprotection of the corresponding PBLG and the study the secondary structure of the PGA at different pH values. These results will serve as a reference for understanding the influence of introducing glycine later in the research work.

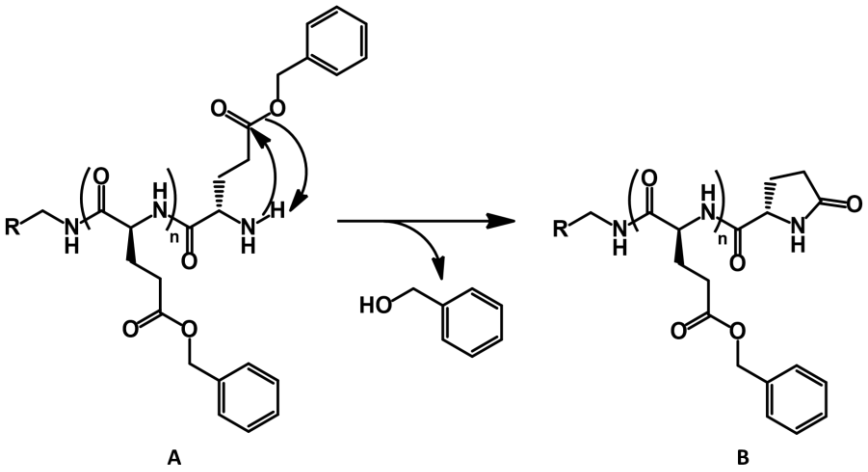
2.1.1.1. ROP of BLG NCA and PBLG characterization

In most cases, the ROP of NCA must be done with protected side groups in anhydrous solvents due to sensitivity to water and protic solvents.[98] In the case of glutamate, γ-benzyl-*L*-glutamate (BLG NCA in **Scheme 1**) is arguably the most studied NCA form.[99] This is due to the easy preparation of the protected amino acid, the stability of the protecting group and the versatility of its removal as will be shown in subsequent parts.[100]



Scheme 1: *General scheme of the ROP of BLG by a primary amine initiator. In blue, the benzyloxy protecting group of the carboxylic acid side chain functionality. This scheme could be extended to other NCA monomers.*

Successful polymerizations of BLG-NCA require carefully controlled reaction conditions.[31,76] Highly polar aprotic solvents, such as DMF or DMSO, are usually used as they weaken intramolecular hydrogen bonding in the polymer chains: indeed such interactions may induce the formation of less soluble secondary structures. Additionally, the polymerization must be done at low temperatures to avoid the backbiting reaction (**Scheme 2**) that produces dead chains and reduces the control over the polymerization.[31]



Scheme 2: *Termination/backbiting of a PBLG chain during or post polymerization in solution. The terminal primary amine can attack the ester bond forming an end-capping pyroglutamate group whilst releasing a Benzylalcohol molecule. PBLG polymers can thus be found in two forms (A or B) depending on the terminal group's structure.*

A first PBLG was synthesized (polymer **1**) using propargylamine as initiator at 0°C in dry DMF. The targeted DP, determined by the ratio of BLG NCA to propargylamine (M/I), was 120. The initial monomer concentration was 0.6 M in DMF, and the reactions were left to stir under argon protection on a Schlenk line to allow for the CO_2 release. The reaction completion was verified by Fourier transform infrared using attenuated total reflectance sampling (ATR-FTIR) analysis. This was done by monitoring the disappearance of the carbonyl bands of NCAs as they are reacting and being incorporated in the polypeptides. These peaks can be observed at 1850cm^{-1} and 1790cm^{-1}, peaks a and b respectively in **Figure 8**.

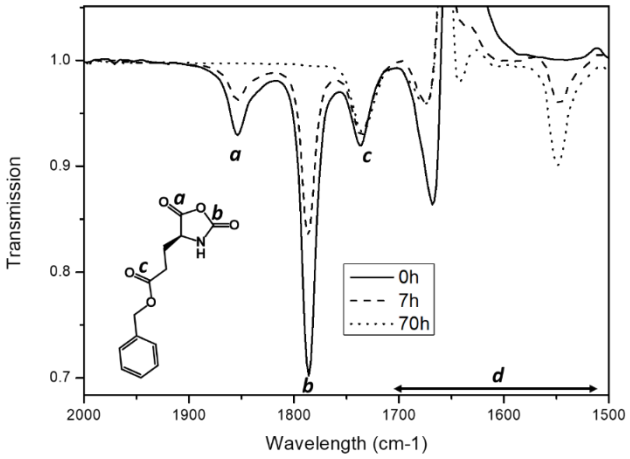


Figure 8: *FTIR spectra of the polymerization mixture (polymer **1**) of BLG NCA in DMF initiated by propargylamine. Three different aliquots from the sample are represented at different times with t=0h before the addition of the initiator. The window represented from 2000cm^{-1} to 1500cm^{-1} shows the stretching bands of different carbonyls from BLG NCA and PBLG. a. is the stretching band of the anhydride carbonyl from the NCA, b. that of the anhydride/amide one, c. the ester one from the side chain of both monomer and polymer (stable concentration) and d. wavelengths where the amide bonds from the polypeptide can be found but where the solvent (DMF) also absorbs.*

The polymer was then purified by precipitation in cold ether and centrifugation and finally dried under vacuum (see **Supporting information**). The resulting PBLG was dissolved in $CDCl_3$ for NMR analysis. 15% TFA is usually added to help solubilize the polymer as it breaks hydrogen bonds that may form insoluble secondary structuring. The ^{1}H NMR analysis of polymer **1** (PA-PBLG$_{120}$) showed the polymer backbone peaks

(a, f), side chain peaks (b,c) and benzyl protection peaks (d,e) as shown in **Figure 9** thus verifying its chemical composition.

DP calculation using ^{1}H NMR was not accurate since the initiator peaks could not be easily identified as the targeted DP value is high. Chromatograms from SEC in DMF showed one peak in all three detectors using multi-angle light scattering (MALS), refractive index (RI) and ultraviolet-visible light spectroscopy (UV-Vis) respectively verifying the formation of high molecular weight polymers with narrow dispersities (**Figure 10**). It was then possible to determine the molar mass of the polymer using its dn/dC (see **Supporting information**). The number average molecular weight (Mn) was thus calculated to be 27 kg/mol (DP=123) and this molar mass is associated with a narrow dispersity value of 1.05.

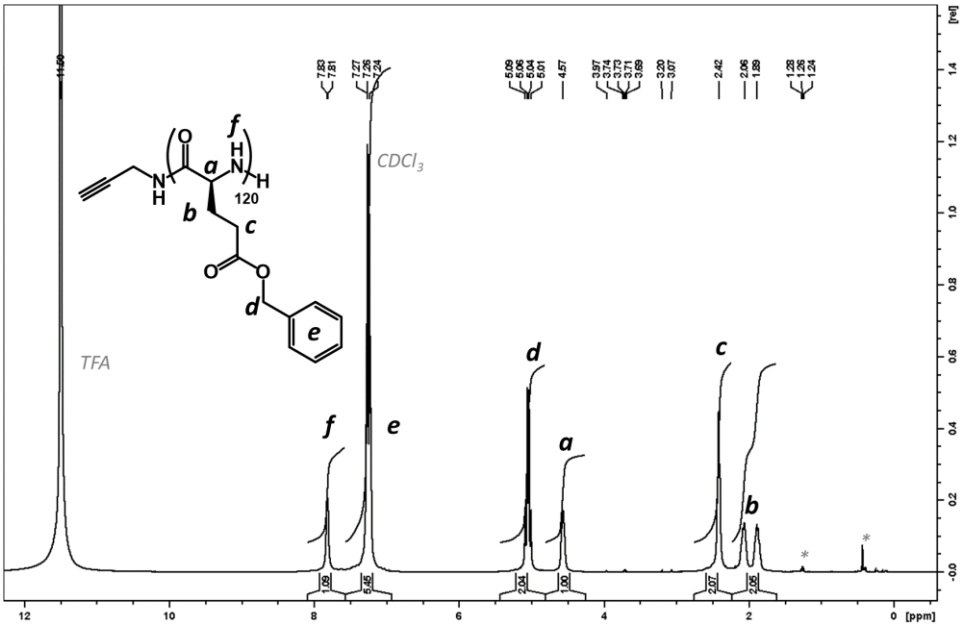


Figure 9: *^{1}H NMR spectrum of polymer **1** in $CDCl_3$+15%TFA. Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be identified and integrated for DP calculation. Peak f is visible due to the use of non-deuterated TFA. *Impurities due to the use of use of non-deuterated TFA.*

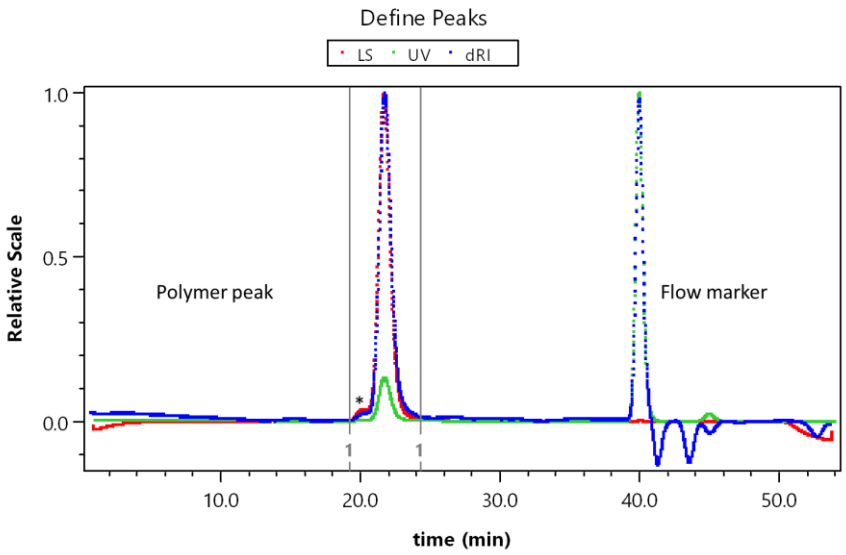


Figure 10: *SEC chromatograms of polymer **1** using triple detection MALS, RI and UV. MALS detection verifies the formation of high molecular weight polymers using dn/dC, RI the relative concentration of species in solution as well as the elution time compared to the flow marker and UV to confirm the presence of certain functions, in this case aromatic rings on the side chains.*

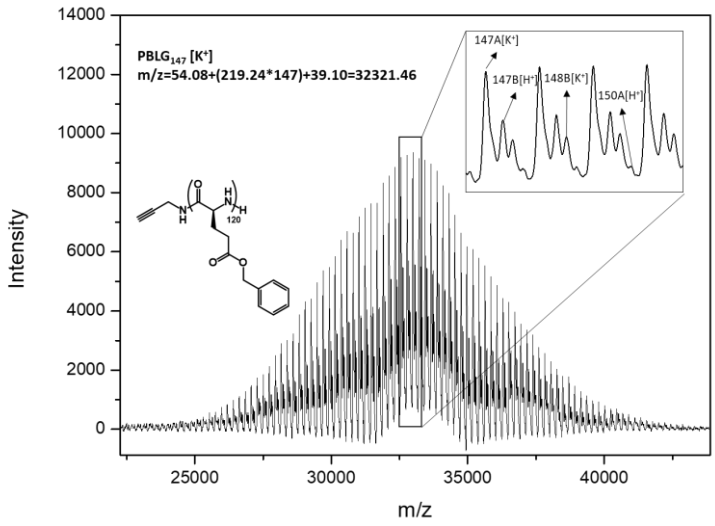
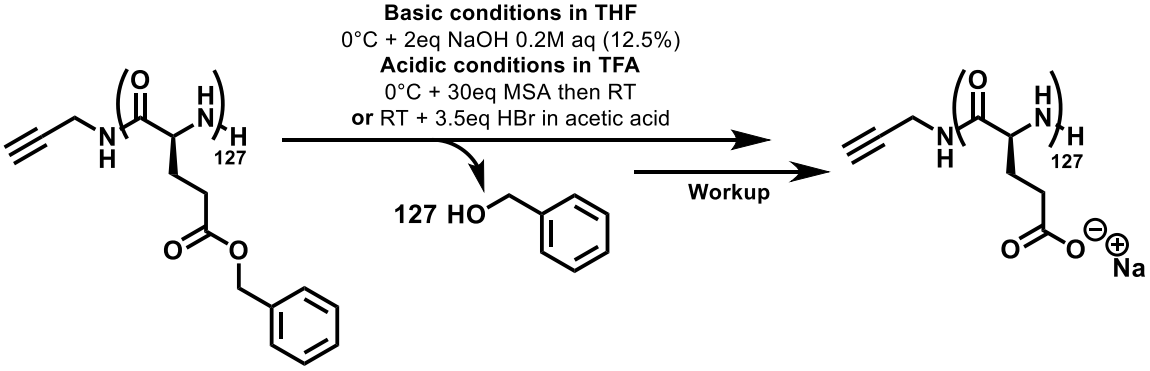


Figure 11: *MALDI-TOF spectra of polymer 1 using DCTB as matrix. The different peaks were assigned to a certain monomer unit in either form A or form B polymers (as shown in **Scheme 2**) as the analysis was done at room temperature. The initiator propargylamine was taken into consideration during the calculation confirming both chain ends. The average molar mass is an approximation as no calibration at such high values of molar mass was possible.*

Both NMR and SEC results were also confirmed using MALDI-TOF analysis (**Figure 11**). The mass spectrum evidenced a peak molecular weight near 32.5kg/mol and it revealed the presence of a set of ions with a mass difference of 219 g/mol verifying the repetition units of BLG with an initiation of the ROP attributable to propargylamine (54 g/mol).

2.1.1.2. Deprotection

In the previous part, the synthesis and characterization of PBLG in a controlled manner was shown. This hydrophobic side chain bearing polymer can be used in many applications including nanomaterials[101], nanofibers[79], organogels[37], and drug delivery application[84]. PBLG polymers are also more frequently synthesized as precursors for their deprotected water soluble version poly(*L*-glutamic acid), PGA. This anionic polypeptide holds great promise for biomedical applications[102] and catalysis[93]. Once polymerized, the hydrophobic side chains of PBLG can be deprotected to give its water-soluble counterpart bearing carboxylic acid moieties. This reaction can be implemented in acidic or in basic conditions (**Scheme 3**). In this part, some of the different ways of deprotecting PBLG into PGA will be compared aiming to preserve the chirality of the backbone, prerequisite for PGA secondary structure.



Scheme 3: *Deprotection pathways of PBLG (Polymer **1**) to obtain PGA (Polymers **1D**$_X$) and optimized conditions. In all three cases, the workup involves neutralization of the pH of the medium with $NaHCO_3$. This gives carboxylate side chains with Na^+ counterions.*

Basic conditions

PGA can be produced from PBLG (**Scheme 3**) using NaOH in aqueous solvents, but epimerization[76] and even backbone hydrolysis can occur in such conditions[103]. For these reasons, mild basic conditions must be used:[76] PBLG is first solubilized in an

organic solvent, THF, at low concentration then cooled to 0°C. An aqueous NaOH solution is then added dropwise for the saponification reaction (see **Supporting information**). The THF/water ratio was kept at 8:1 and the reaction time >16h. Furthermore, 2eq of NaOH to benzyl gave optimal results, meaning that the initial synthesis and characterization must be precise as to succeed in the subsequent deprotection. After reaction completion, the solvent was evaporated, and the NaOH neutralized with saturated $NaHCO_3$ aqueous solution. The crude product was subsequently dialyzed against Milli-Q® water and finally lyophilized to obtain pure PGA, polymer **1D₂** (the **D** corresponds to the deprotected form of the polymer) which was completely soluble in water. A first deprotection of polymer **1** was conducted as to optimize the conditions, polymer **1D₁**, and is shown in the **Supporting information**.

The deprotection was initially verified by NMR in D_2O containing a small amount of NaOD to ensure solubility. Polymer **1D₂** showed a peak consistent with PGA backbone (peak a) and side chain peaks (b and c) as shown in **Figure 12**. Furthermore, the region around 8 ppm showed no residual benzyl groups, indicating a full deprotection of the polymer.

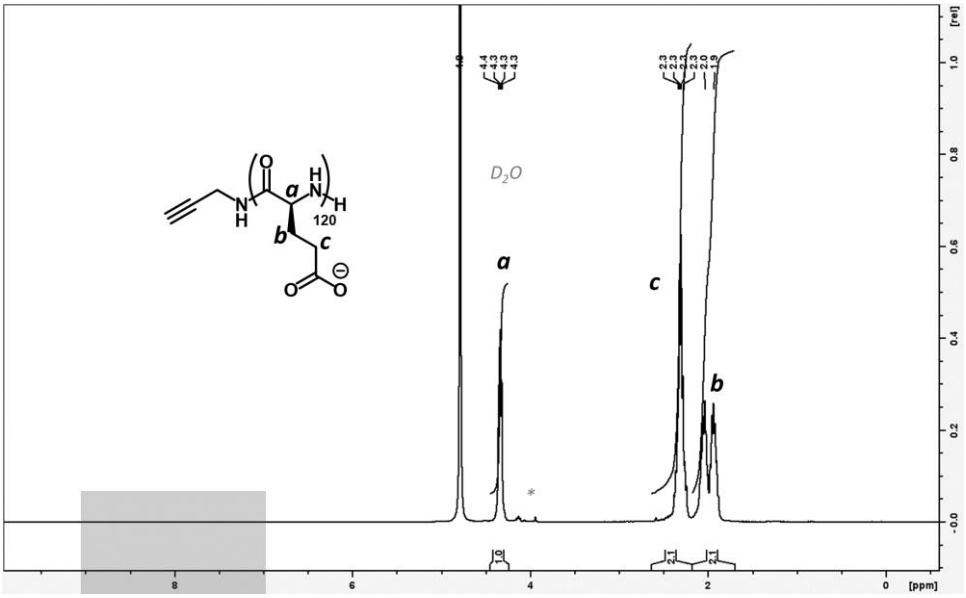


Figure 12: *^{1}H NMR spectrum of polymer 1D$_2$ in D$_2$O+NaOD. Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be identified and integrated for DP calculation. No benzyl groups were detected in the highlighted area around 8ppm. *end chain groups*

Although this deprotection technique proved to be useful, a simpler workup was found with other techniques as will be discussed in the following part.

Acidic conditions

Two acidic pathways were implemented for this study, and in both cases the polymers were solubilized in TFA. Although TFA is an acid, it is not strong enough to deprotect the stable benzyloxy pendant chains at room temperature.[100] Nonetheless, it has the advantage of solubilizing most, if not all, polypeptides. The acidic hydrolysis can then be done using either methanesulfonic acid (MSA)[77] or HBr[76] (**Scheme 3**).

MSA

The polymer was dissolved in TFA at 100mg/ml and then ≈ 30 eq MSA to BLG units was added. A small amount of anisole is usually added in these conditions to avoid side reactions such as alkylation or oxidation.[100] The reaction mixture was stirred for 20 minutes at 0°C then 40 minutes at room temperature (see **Supporting information**). The polymer obtained, polymer **1D$_3$**, was not completely soluble in water meaning that the deprotection was certainly not fully achieved. NMR analysis in different solvents failed to determine the insoluble compounds but the D_2O-soluble fraction was confirmed to be PGA (see **Supporting information**). This pathway was not taken further.

HBr

A first HBr driven deprotection was conducted as to optimize the conditions, polymer **1D$_4$**, and is shown in the **Supporting information**. Polymer **1** was dissolved in TFA at 30 mg/ml, 3.5 eq of aq. HBr was added and the reaction was left to stir overnight. The polymer was precipitated from cold ether, centrifuged, neutralized using saturated $NaHCO_3$ aqueous solution then dialyzed against Milli-Q® water and finally lyophilized to obtain pure PGA, polymer **1D$_5$** as a white fluffy powder (see **Supporting information**). The obtained polymer was completely soluble in water and ^{1}H NMR spectra backbone (peak a) and side chain peaks (b and c) are shown in **Figure 13**.

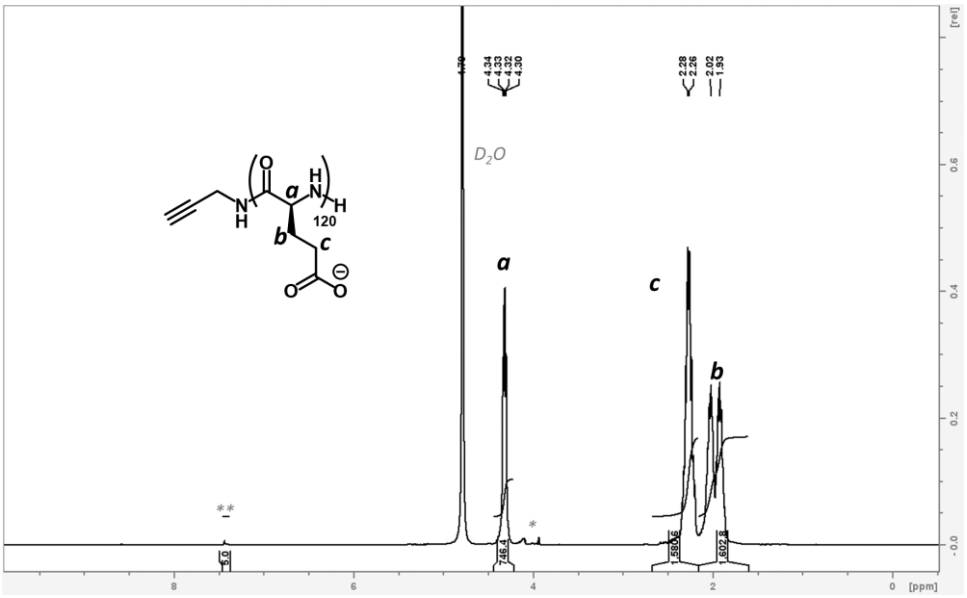


Figure 13: *^{1}H NMR spectrum of polymer **1D$_5$** in D_2O+NaOD. Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be identified and integrated for DP calculation. *end chain groups **residual benzyl groups, integration shows a ratio of 750 glutamate unit to 1 benzyl making it negligeable*

For SEC analysis, the water-soluble PGAs were dissolved in a basic phosphate buffer and passed through an aqueous chromatography setup. The basic conditions results in the deprotonation of the carboxylic acid side chains,[104] to ensure solubility. Furthermore, in these conditions secondary structuring of the polypeptide chains can be avoided as protonated PGA forms insoluble helical structures. SEC results for polymer **1D$_5$** (**Figure 14**) showed a number average molecular weight (Mn) of 18 kg/mol (DP=129) and this molar mass associated to a narrow dispersity of 1.1. Overall, SEC results for the deprotected polymers (polymers **1D$_2$** and **1D$_5$**) showed signs of aggregation by LS detection that we attributed (**Figure S7** and **Table S1**) to the low concentration of the buffer related to the pendant carboxylate groups. This led to aggregation and inaccurate Mn calculations due to the peak shapes in MALS detection.

The HBr method was chosen for further study. Its use of TFA as solvent will ensure its applicability on a wide range of copolymers. Furthermore, the simpler work up will allow the screening of several copolymers.

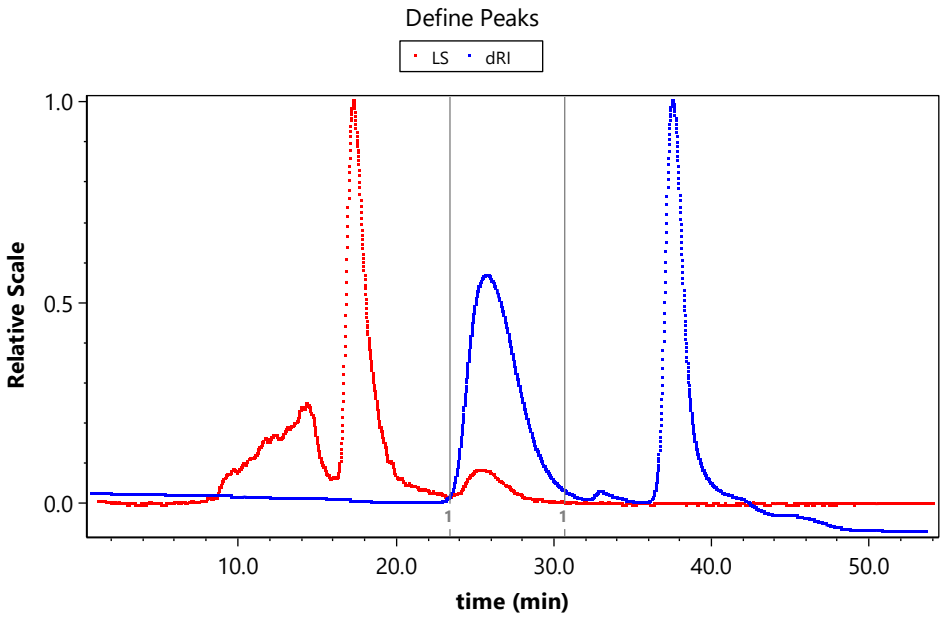


Figure 14: *SEC chromatograms of polymer* ***1D₄*** *using MALS and RI detection. MALS detection shows aggregated state at lower elution time. Even though the aggregates are detected by MALS, they form a negligible fraction of the mass of the polymer, as can be seen from the dRI trace.*

For the rest of this study, copolymers of a targeted DP=50 were examined. Polymer **2,** a homoPBLG$_{50}$ initiated by hexylamine and its deprotected form polymer **2D**, will serve as standard for all copolymers. Its synthesis was done in the same way as described before and its structure verified by NMR and SEC. The initiator peaks (a, b, c, and d) were readily identified by ^{1}H NMR due to the smaller molecular weight and the easier to identify alkyl chain protons (**Figure 15**). By integrating the polymer backbone peaks and dividing them by the initiator ones, a DP of 53-54 was calculated for polymer **2**.

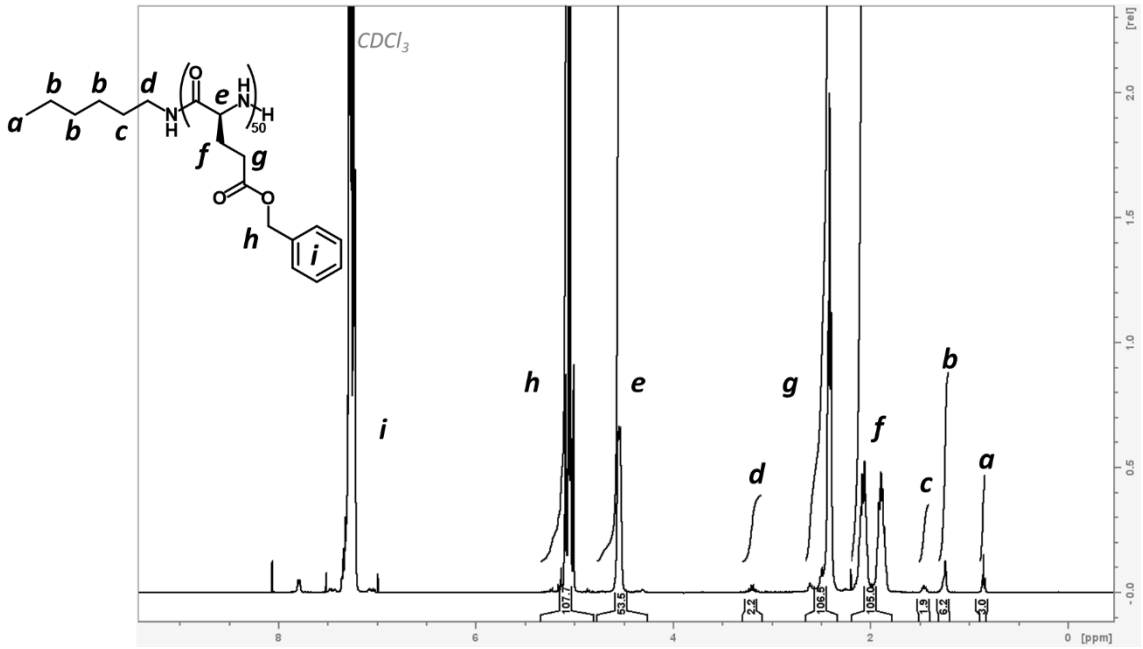


Figure 15 : *^1H NMR* spectrum of polymer **2** in $CDCl_3$+15%TFA-d. Peak attributions are indicated by the letters. Integration of peak *a* from hexylamine allows to calculate a DP of 53-54.

Both SEC of the protected form in DMF (polymer **2**) and the deprotected form (polymer **2D**) in aqueous SEC showed high molecular weight peaks that are well defined (**Figure 16**).

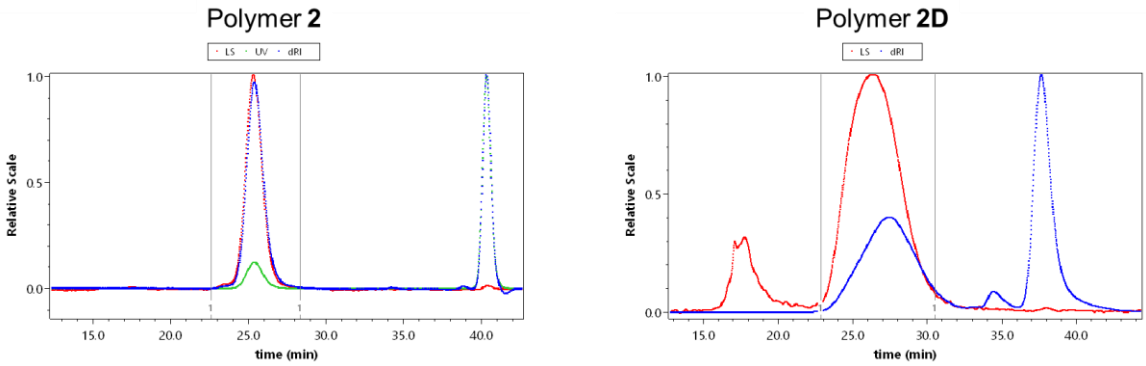


*Figure 16: On the left: SEC of polymer **2** in DMF. MALS (red), RI(blue) and UV (green) detections are presented. On the right: Aqueous SEC of polymer 2D. MALS and RI detections are presented. Some aggregation can be observed in MALS detection though its concentration is very low as suggested by the RI detection.*

Using dn/dC values in both cases, it was possible to calculate the DP. For polymer **2** a DP of 51 with a dispersity of 1.03 were calculated and a DP of 56 with a dispersity of 1.29 for polymer **2D**. These results show the possibility to obtain PBLG and PGA

with well-defined and controlled structure that can be used as a benchmark for the rest of the study.

2.1.1.3. Secondary structuring

Peptide-based polymers can adopt similar secondary structures to those of proteins.[105] As mentioned in the previous chapter, both classes of macromolecules are constituted by a repetition of two neighbouring planes that are connected by the α-carbon. Only some torsion angles (ψ, ϕ) are allowed for each amino acid residue, corresponding to favourable low-energy regions, and limited by the chirality of the backbone and its pendant chains. The structure of a given polypeptide chain, will be the one which is most thermodynamically favourable and will depend on its environment.[106] Ramachandran highlighted the torsion angles of natural proteins in what is known as the Ramachandran plot (chapter **1**).[107] Two main zones have been highlighted: the α-helix and β-sheet, as the most naturally abundant structures. Focusing on the α-helix which is a rod-like structure and has the pendant side chains extending outward in a helical array. This structure is stabilized by a hydrogen bonding network created between the oxygen of the carbonyl group of each peptide bond in the strand and the proton present on the amide of the peptide bond four units distant. The most common angles found in the α-helix geometry are $\phi = -60°$ and $\psi = -45°$ with 3.6 amino acids (1.5 Å) per helix step and a radius of 2.3 Å, excluding side chains.[105] Other interactions may play a role in this secondary structuring such as weaker electrostatic interactions, hydrophobic affinity and even intermolecular bonding.

PBLG exhibits secondary structuring that depends on its molecular weight and on the solvent. Oligopeptides can adopt beta-sheet like structures, but above a DP of 10-15, they adopt an α-helical conformation.[108] This conformation reveals itself in helicogenic solvents, such as tetrahydrofuran (THF) or chloroform, and is usually insoluble.[109,110] Circular dichroism spectroscopy (CD) is a powerful tool to analyse this secondary structure. CD is widely used in structural biology to evaluate the conformation of proteins.[111] Solutions of the analytes are irradiated with UV-Vis light. Optically active chiral molecules interact differently with left- and right-handed circularly polarized light and the differential absorption between the right- and left-handed light is thus measured. The chiral conformation of peptidic samples can then be monitored via this technique. This is usually done in the far-UV region where the electronic transitions of the amide group of the peptidic backbone occur. The CD spectra of most secondary

structures, such as α-helices, β-sheets and random coil, have well documented shapes and magnitudes.[105] In the case of PBLG, the α-helical form is soluble in hexafluoro-2-propanol (HFIP) (**Figure 17**), a solvent that does not absorb in the analysed wavelength window.

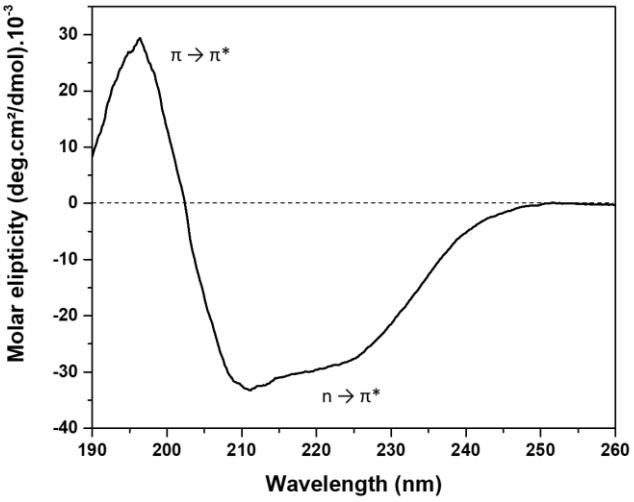


Figure 17: *CD spectra of a homoPBLG$_{50}$ sample (polymer **2**) in HFIP. The α-helical signal is consistent with literature data.[108] The positive absorption band around 190 nm corresponds to a π→π* transition while the broader one around 210-220 corresponds to an n→π* electronic transition.*

The double minimum in the area between 208 nm and 224 nm is attributable to a α-helical structure.[108] In the case of the deprotected, water-soluble form, PGA, different conformations can be observed by CD analysis in aqueous buffers. Indeed, the secondary structure of this polymer depends on the pH of the solution (**Figure 18**). This is due to the charged state of the side chain. When they are deprotonated, the negative charges repel each other and extend the polyanion in a random coil form. On the other hand, the carboxylic acid moieties form a network of intramolecular hydrogen bonding, stabilizing the α-helical form.[77]

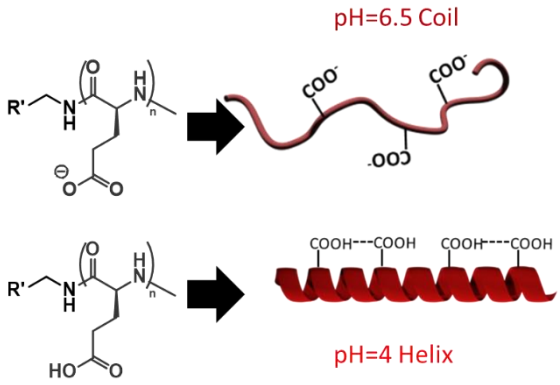


***Figure 18**: Schematic representation of the two forms of PGA: On top, at pH>pKa the side chains are in the carboxylate forms and the polymer chain is in its coil form. On the bottom, the side chains are protonated at pH<pKa and stabilize the helical form.*

This behaviour has prompted researchers to consider the pH responsiveness of PGA in the formulation of smart materials.[112] The analysis of the two forms of polymer **2D** were analysed using aqueous CD. The obtained spectra of the sample at pH=6.5 had a small maximum at 218 nm and a sharp minimum at 198 nm (**Figure 19**) attributable to a random coil conformation. The sample at pH=4 had two minima at 209 nm and 224 nm and a maximum at 195 nm (**Figure 19**) corresponding to the α-helical form. These results were also in accordance with data found in the literature.[77]

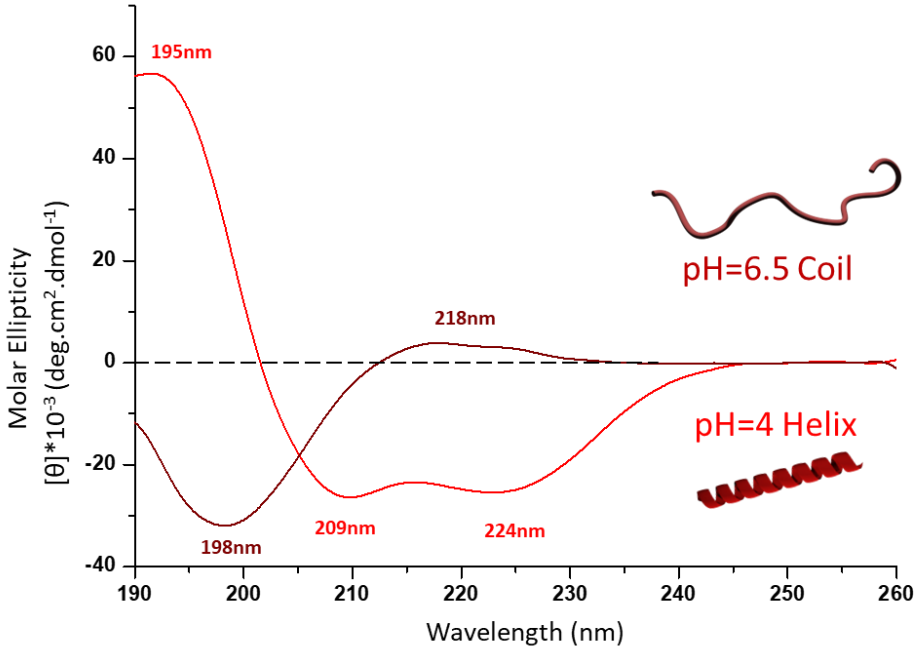


***Figure 19**: CD spectra of the two forms of PGA (polymer 2D). The distinctive positive/negative peaks of each form are represented.*

The successful synthesis of PBLG and its deprotection to PGA were explored in this part. The analysis of both the polymers' structures and conformations were also developed. In the next part, the second monomer, glycine, will be discussed. Focus on the synthesis of the NCA form, the ROP of the homopolymer and its analysis will be presented.

2.1.2. Glycine NCA

Glycine is the simplest amino-acid and is highly abundant in natural proteins. It is not chiral. Glycine units are known to avoid folding in peptide sequences and are usually found in disordered domains in proteins.[113] Thus, copolymerizing BLG and Gly NCA, a structure inducing monomer and a disorder inducing monomer respectively, in a controlled manner will have a direct influence on the secondary structuring of the produced copolymers.[66] Furthermore, the absence of a side chain on glycine makes it easier to distinguish from other amino acids through characterization, notably by NMR.[53] This will simplify the development of the calculations and experiments needed to determine the copolymerization kinetics.

2.1.2.1. Synthesis and ROP of Glycine NCA

Unlike BLG NCA, Gly NCA is not commercially available and had to be synthesized. While the Fuchs-Farthing pathway is most commonly used to produce glycine NCA (**Table 1**), we chose not to use phosgene or any of its derivatives due to their acute toxicity.

Table 1: *Synthesis examples of Gly NCA found in the literature and their main characteristics. In bold, pathways that were successfully used in-house for other glycine-like monomers[114] and used in the synthesis at hand.*

Amino acid precursor	Synthesis pathway	Main reagents	Reaction conditions	Ref.
Glycine	Fuchs-Farthing	Phosgene	THF, 45°C	115
			Dioxane, 50°C	116–118
		Phosgene + SO_2	Dioxane/acetic anhydride 45°C	119
		Diphosgene	THF, 45°C	37,120–122
		Triphosgene	THF, 70°C	123
			Ethyl acetate, 70°C	35

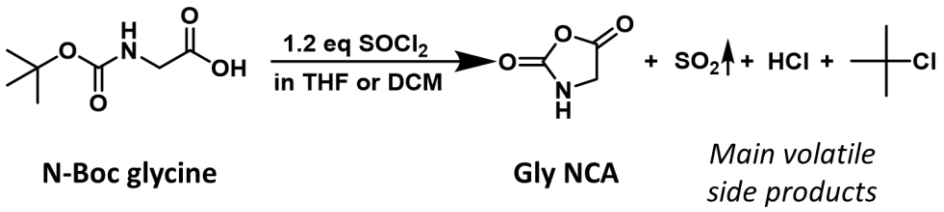
Amino acid precursor	Synthesis pathway	Main reagents	Reaction conditions	Ref.
Glycine	Fuchs-Farthing	Triphosgene + triethylamine	Ethyl acetate	124
		Triphosgene + α-pinene	THF, 65°C	51
		Triphosgene + limonene	Acetonitrile, 70°C	125
		Triphosgene + epichlorohydrin	THF, RT	98
		Triphosgene	Acetonitrile/aqueous sodium salts (microfluidics)	126
		Chloroform	Chloroform/ACN, high energy UV irradiation, O_2	127
		1,1'-carbonyldiimidazole	D_2O	117
2-isocyanatoacetyl chloride		Water/dioxane	Diethyl ether	113
N-ethoxycarbonyl glycine	Leuchs	**$SOCl_2$**	$SOCl_2$, 40°C	128
N-benzyloxycarbonyl glycine			Acetic anhydride, reflux	116
		PCl_5	Ethyl acetate, ice bath	129,130
			Diethyl ether	131,132
***N*-Boc glycine**		**PCl_3**	DCM, 0°C	133

Other pathways can be found in the literature describing the synthesis of Gly NCA from different precursors and using different reagents (**Table 1**).

Synthesis of Gly NCA

The protocol using N-Boc glycine and PCl_3[133] (**Table 1**) was not reproducible (see **Supporting information**). Only trace amounts of Gly NCA were detected even after a prolonged reaction time and at higher temperatures. Similarly, the pathway using T3P, though useful for the synthesis of other NCAs, showed little conversion over prolonged reaction times for glycine. Furthermore, the purification of the reagents, notably pyridine, proved complicated (see **Supporting information**). Finally, N-Boc glycine was used as the amino acid precursor for Gly NCA and was dissolved in THF or DCM. $SOCl_2$ was chosen as the cyclizing agent (**Scheme 4**) and novel synthesis and

purification protocols were developed. When THF was used as solvent, the obtained powder was brownish if THF was used for filtration and yellowish if ethyl-acetate was used instead. When DCM was used as the reaction solvent, the powder was perfectly white, but the yield was always below 60% (see **Supporting information**). Otherwise, all products had very similar reactivities and were chemically comparable.



Scheme 4: *Developed synthesis pathway for glycine NCA using $SOCl_2$ and starting from N-Boc glycine. The main side products resulting from the reaction are all gaseous or volatile.*

Building on previous knowledge of NCA synthesis from glycine derivatives[114], model conditions were selected; 0.36 M and 1.2 eq of $SOCl_2$. The optimal reaction time was determined to be 4 hours using ^{1}H NMR in DMSO-d_6 (**Figure 20**).

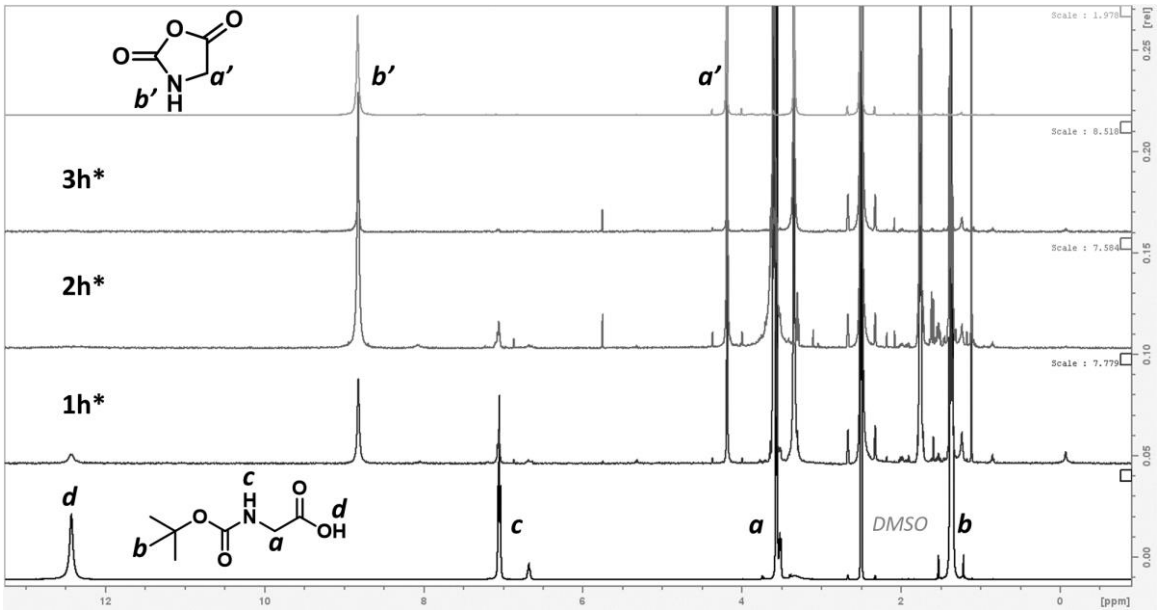


Figure 20: *^{1}H NMR spectrum of Gly NCA synthesis at different times. ^{1}H NMR spectra (ordered from bottom to top): N-Boc glycine, reaction mixture after 1 hour, 2 hours, 3 hours, and glycine NCA after purification. All spectra were recorded in DMSO-d_6. *reaction solvent was evaporated, and crude mixture was analysed as is. Untagged signals are unidentified side products.*

After 4 hours of reaction, the reaction medium was a dark brown colour. The first step of purification was to evaporate most of the reaction solvent and almost all volatile side

product (**Scheme 4**). Solubility of reactants and products were determined as to correctly select the washing steps (Table S2). When the reaction mixture was concentrated enough to see a precipitate, it was washed with excess hexane and DCM. ^{1}H NMR showed removal of many unidentified byproducts and only Gly-NCA as precipitate (**Figure 21**).

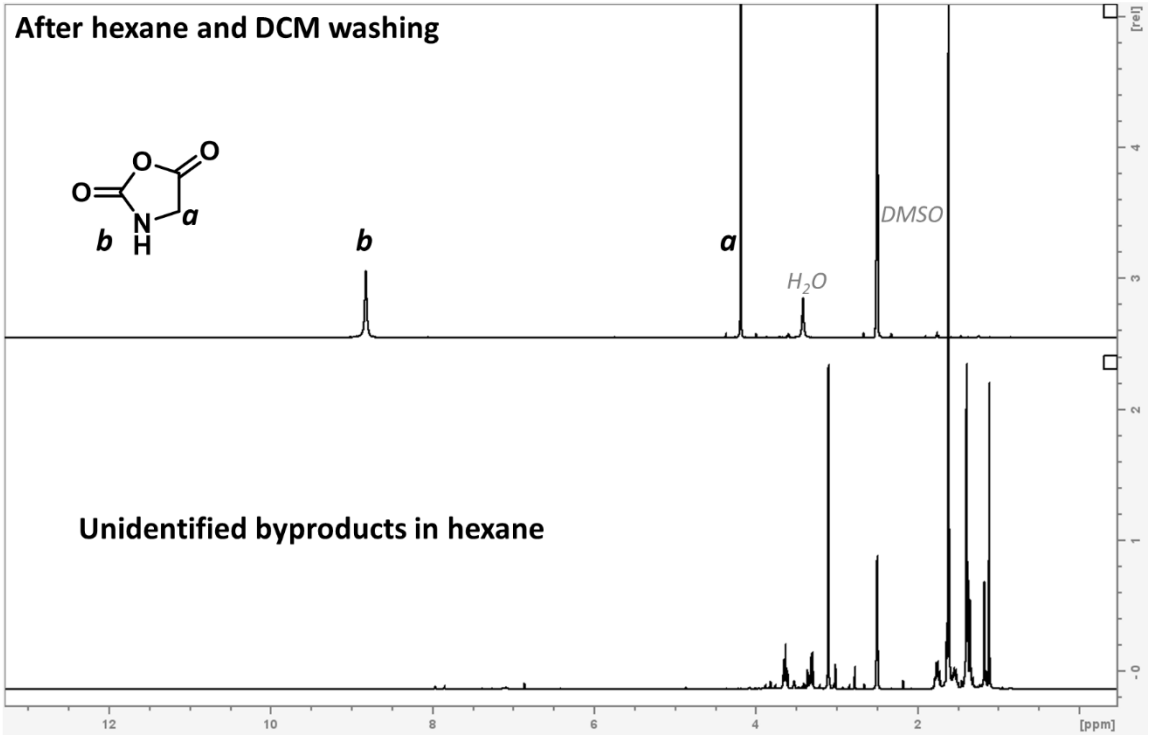


Figure 21: *^{1}H NMR of hexane soluble byproducts (bottom) and precipitate left after hexane/DCM washing. Both spectra were recorded in DMSO-d_6.*

As shown in **Figure 21**, water can be observed in the NMR solvent as DMSO is highly hygroscopic. Yet, the NCA molecules remain intact as the medium is very acidic. This was verified by pH paper that quickly gave acidic colours. This presence of excess chlorine was confirmed by the addition of silver nitrate in a THF solution and formation of a silver chloride precipitate. Celite filtration[134] was selected as the most convenient and high yielding method to remove both contaminants.[114] In order to do so, the brownish precipitates were suspended in ethyl acetate (EtOAc) and passed over a dried celite column. The undissolved matter was shown to be glycine (**Figure S9**) and was readily filtered out on top of the celite. After two celite filtrations, the final yellowish powder gave no visible precipitate using the silver nitrate technique described above and the pH paper was neutral. Furthermore, subsequent ^{1}H NMR in DMSO-d_6 containing traces of water, NCAs can sometime be seen to degrade to their amino acid form or to polymers that precipitated over time (**Figure S10**), confirming the usability

of this product for ROP. Analysis of pure Gly NCA using ^{1}H NMR (**Figure 22**) , ^{13}C NMR (**Figure 23**), ATR-FTIR (**Figure S11**) and ESI-MS (**Figure S12**) all showed the desired product with high purity (see **Supporting information**). Other purification methods, like recrystallization, gave low yields.

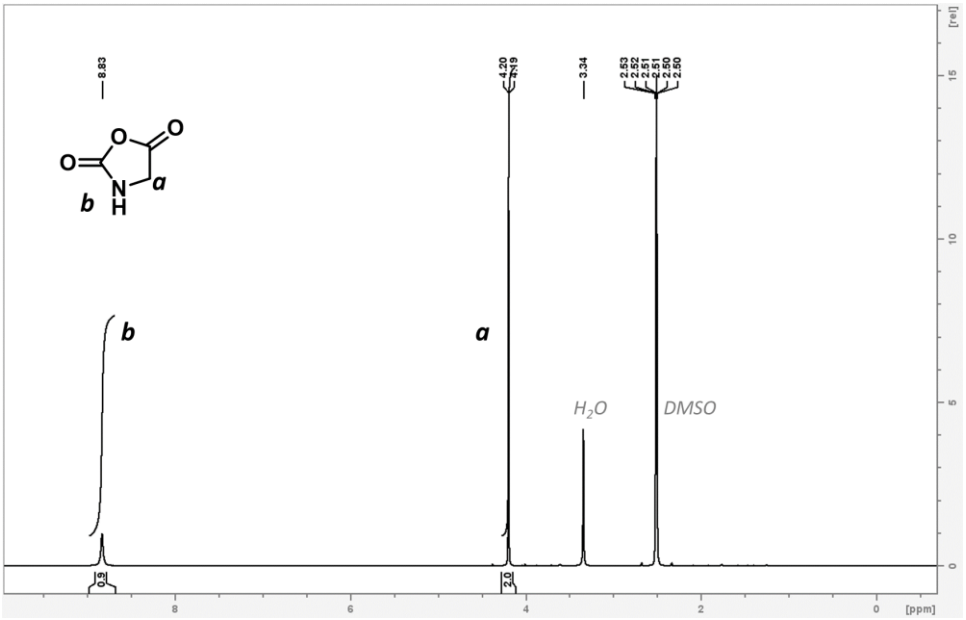


***Figure 22**: ^{1}H NMR in DMSO-d_6 of Gly NCA after purification through celite.*

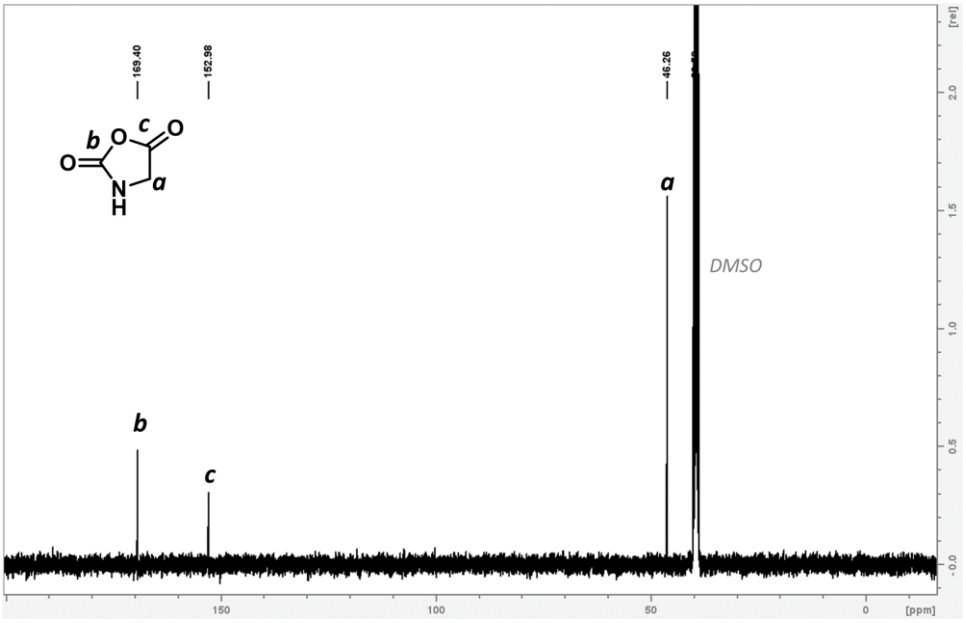


***Figure 23**: ^{13}C NMR in DMSO-d_6 of Gly NCA after purification through celite.*

The presence of two carbonyls in the ^{13}C NMR (**Figure 23**), is characteristic of the formation of an NCA. To further consolidate the successful synthesis, 2D-$^{1}H/^{13}C$ HMBC was carried out in dry THF-d_8 (**Figure 24**).

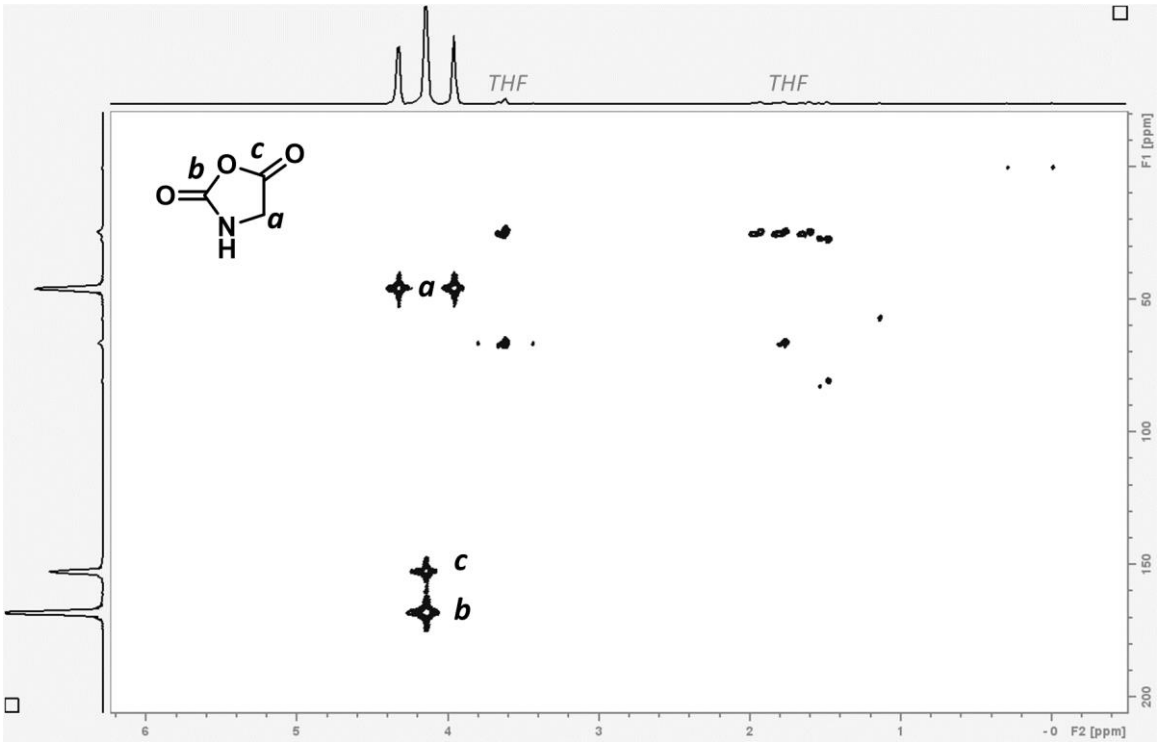


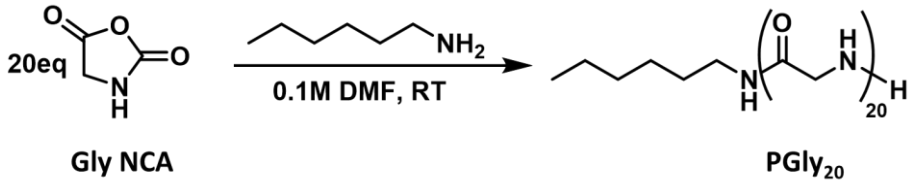
***Figure 24**: 2D H-C HMBC NMR of Gly NCA in THF-d.*

In this chapter, all Gly-NCA employed was synthesized via this optimized procedure. Ultimately the best way to probe the purity of NCAs is to ultimately evaluate their ability to form polypeptides via ROP (see the following section). For this purpose, all batches of Gly-NCA were polymerized and the kinetics of the polymerization followed and compared. If a batch had slower kinetics than previous batches, it was either purified again over celite or discarded.

ROP of Gly NCA

The ROP of glycine *N*-carboxyanhydride (Gly-NCA), even though it's a very simple monomer, is known to be much more complicated than that of BLG. The polymerization of Gly-NCA proceeds in a heterogenous reaction medium due to the precipitating PGly.[135,136] This is mainly due to the formation of insoluble secondary structures that arises from even short repeating glycine sequences.[137] As glycine is achiral, such secondary structures are studied using X-ray diffraction and solid state NMR.[138,139] The peculiar secondary structures of polyglycine (PGly) have been extensively characterized: PGly can adopt either an anti-parallel β-sheet form (PGI) or a 3_1-helical form (PGII).[139] In dry DMF, low concentration of initial monomers were used and an M/I ratio of 20 was targeted to synthesize polyglycines from synthesized NCAs (**Scheme 5**). Using hexylamine as initiator allowed for easy DP calculation by ^{1}H NMR,

if no other initiation occurred. The reaction mixture quickly became turbid on addition of hexylamine as the reaction proceeded rapidly (see **Supporting information**).



***Scheme 5**: Standardized polymerization of Gly NCA, initiated by hexylamine in dry DMF to produce polyglycine of DP 20.*

As discussed in the section presenting the synthesis of Gly-NCA, three polyglycines (PGly) were synthesized from different batches of Gly-NCA at different purification step. PGly **1** from Gly NCA that was not purified by celite, PGly **2** from Gly NCA that was passed through one dry celite column and PGly **3** from Gly NCA that was passed through two dry celite columns. The ROP was slightly faster for PGly **1** while little to no difference in kinetics was observed between PGly **2** and PGly **3**. The final polymers of PGly **1** and PGly **2** were of lower molar mass than PGly **3** as shown by ^{1}H NMR (**Figure 25**). ^{1}H NMR showed that polymers PGly **1** and PGly **2** had lower DP than the targeted one (11 and 15 respectively) while PGly **3** had a DP of 22 (**Figure 25**). Furthermore, both monomers and polymers from PGly **1** and PGly **2** experiments were of brownish color, while PGly **3** was white. It was thus established that two filtrations over dry celite columns were necessary to produce Gly-NCA that will be suitable for the ROP.

Although the characterization of PGly using SEC has not yet been reported in the literature, the analysis in HFIP was developed here. However, polymers with a targeted DP higher than 20-25 units were completely insoluble due to secondary structuring. PGly **3** showed higher elution volume by SEC than PGly **1-2** (**Figure 26**). The molecular weights of the polymers and their dispersities were calculated using PMMA calibration. PGly **1-3** had respectively Mn=2.3, 2.4 and 3.7 kg/mol and their dispersities were of 1.50, 1.31 and 1.32 respectively.

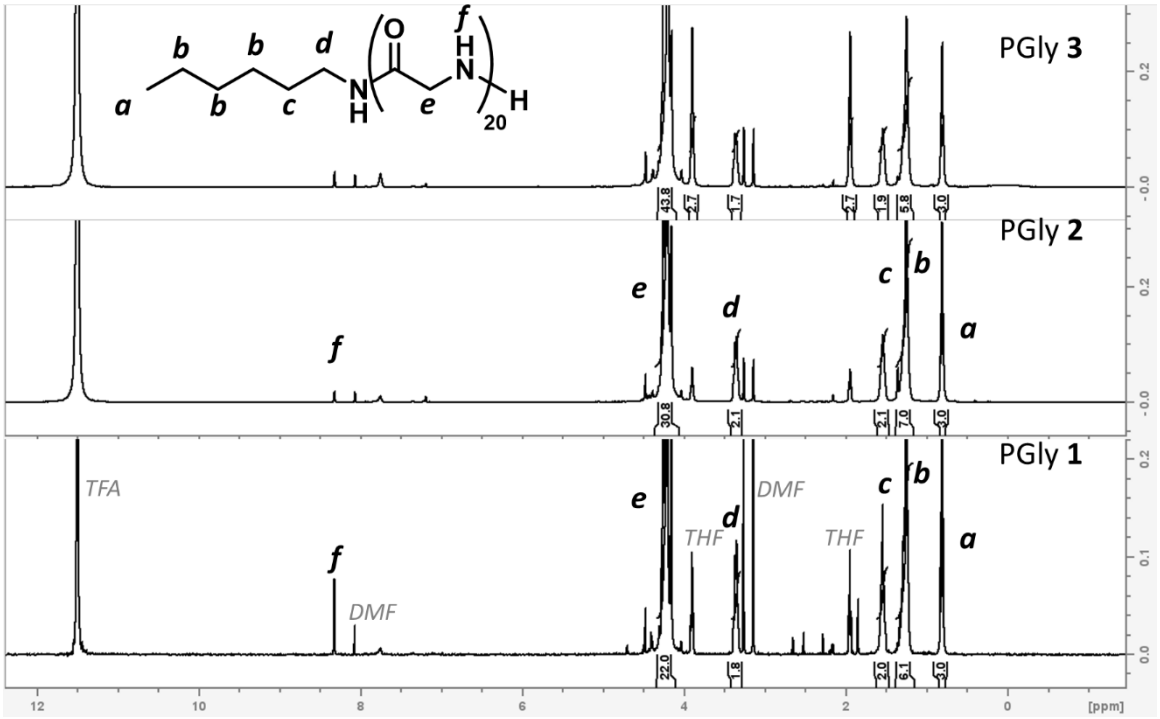


Figure 25: *^{1}H NMR spectra of polymers (from bottom to top) PGly **1-3** in TFA-d. Peak attributions are shown with the letters corresponding to the chemical structure represented.* Peaks f could not be correctly integrated as pure TFA-d was used for clean analysis.

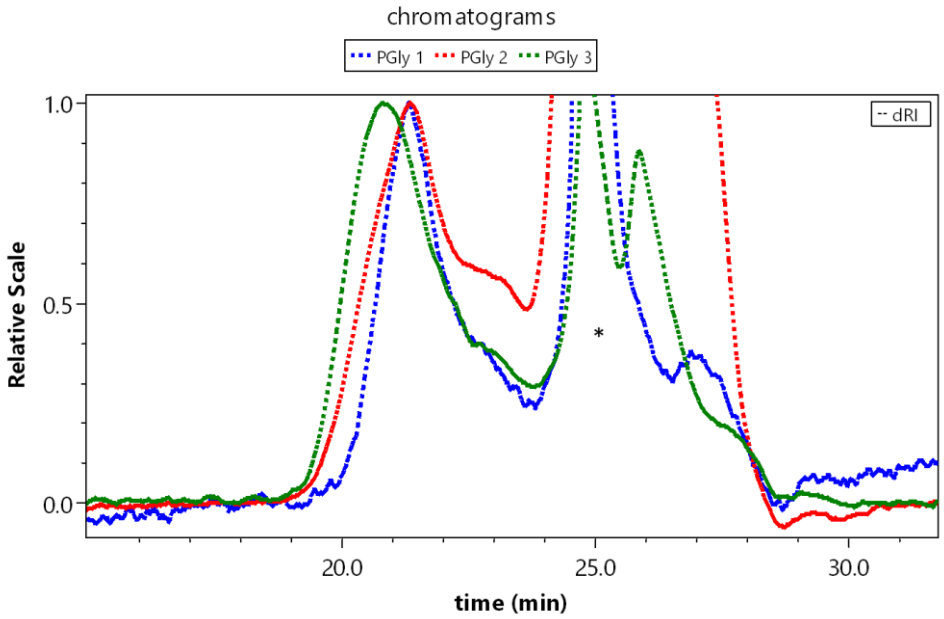


Figure 26: *SEC chromatograms of polymers PGly **1-3** in HFIP. RI detection is represented, MALS detection was not possible and thus no dn/dC could be used for Mn calculation. Chromatograms were normalized at the polymer peaks for easier visualization. *Peaks from flow marker and solvent.*

MALDI-TOF analysis also proved complicated for PGly, as for other aliphatic polyamides,[140] but it revealed the presence of a set of a series of ions matching with a weight difference of 57 g/mol verifying the repeat units of glycine with an initiation of the ROP attributable to hexylamine (101.19 g/mol) (**Figure 27**).

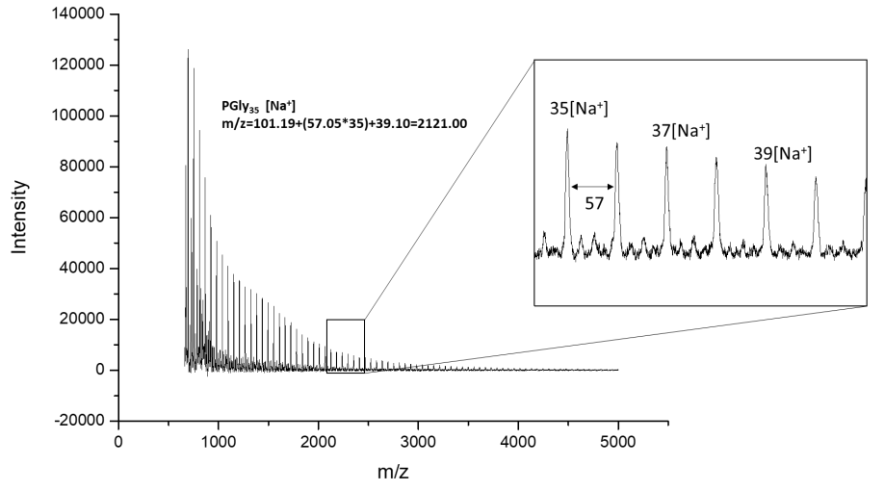


Figure 27: *MALDI-TOF spectra of PGly **3** using α-CHCA as matrix. The different peaks were assigned to a certain monomer unit and its counter ion considering chain end hexylamine.*

The synthesis of Gly NCA was thus fully developed in this research project and its subsequent polymerization explored. Outside this work, this knowledge can prove useful in future projects and perspectives; for example glycine and polyglycine are known to have electrical properties such as piezoelectricity[141] or electrocatalytic effects.[142] With both monomers and their ROP studied and analysed, copolymerizations based on both glycine and BLG NCA will be discussed in the next part. The results presented will thus serve as a basis for the synthesis and the analysis of the copolymers in the rest of this manuscript.

2.2. Copolymerization, kinetics, and secondary structures

The copolymerization of BLG and Gly NCA will be discussed in this section. The goal of this study is to produce glutamate-based copolymers that are intrinsically disordered by breaking their secondary structures using the incorporation of glycine. As a first step, the statistical control over the distribution of the two monomers in the copolymer chains will be explored. The analysis and the calculation of the reactivity ratios as well as the microstructures of the synthesized copolymers will be presented. These results are

used to establish the influence of different experimental conditions on the distribution of the comonomers. Finally, a correlation between the primary sequences and secondary structures of the synthesized copolymers will be established.

2.2.1. Copolymer kinetics

An *in-situ* NMR technique was used to calculate the reactivity ratios of the comonomers, BLG and Gly NCA (see section 1.1.3 for bibliographic context). This technique allows to monitor, concomitantly and independently, the disappearance of the peaks of each monomer and thus quantify their incorporation in the polymer chains. Polymerizations were monitored to high conversions to consider the effect that secondary structure can have on the copolymerization kinetics. The Meyer-Lowry integrated copolymerization equation (**Equation 2**) was used to calculate the copolymerization parameters.[143] It was developed for radical copolymerization and assumes that only the terminal unit affects the reactivity, and that the kinetics are independent of the chain length.

$$(1-x) = \left(\frac{f_A}{f_{A,0}}\right)^{\alpha}\left(\frac{f_B}{f_{B,0}}\right)^{\beta}\left(\frac{f_{A,0}-\delta}{f_A-\delta}\right)^{\gamma}$$

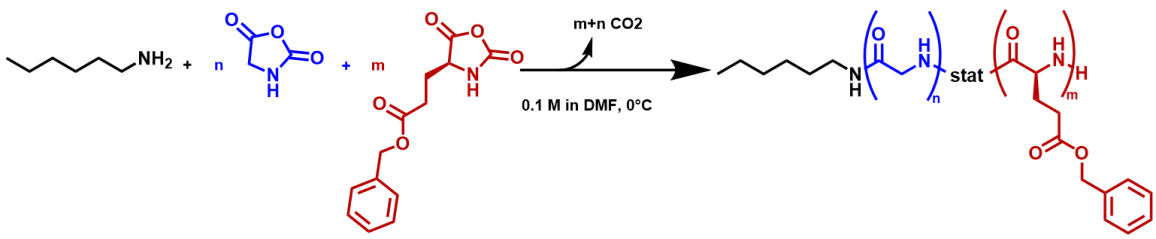
$$\alpha = \frac{r_B}{1-r_B} \qquad \beta = \frac{r_A}{1-r_A}$$

$$\gamma = \frac{1-r_A r_B}{(1-r_A)(1-r_B)} \qquad \delta = \frac{1-r_B}{2-r_A r_B}$$

f is instantaneous monomer feed
f_0 is initial monomer feed
r is reactivity ratio
x is reaction conversion
$f_A + f_B = 1$

Equation 2: *Integrated copolymerization equation and the values of the constants.*

Reactivity ratios are estimated by fitting **Equation 2** to experimental data of monomer conversion in function of monomer fraction. To show the technique used, the study of copolymer **1** (**Scheme 6**) will be detailed. All other copolymers were explored in a similar manner.



Scheme 6: *Copolymer **1**'s general scheme. Gly NCA and BLG NCA were copolymerized in DMF at 0°C. n=m=25 for a copolymer of total theoretical DP 50 (M/I=50).*

This polymer was studied at equal initial monomer feeds ($f_{G,0}$ = $f_{B,0}$ = 0.5) for Gly NCA and BLG NCA. An advantage of the in-situ NMR technique is that only small amounts of chemicals need be used. This allows to run several experiments at different $f_{x,0}$ to be able to determine the copolymer composition (**Fx**) in all conditions. In the first experiment where $f_{G,0}$ = $f_{B,0}$ = 0.5, the disappearance of the monomer peaks can be observed and used to calculate the conversion of each monomer as well as the overall conversion (**Figure 28**). Gly-NCA reacts faster than BLG-NCA and thus is preferentially incorporated in the copolymer at lower conversion.

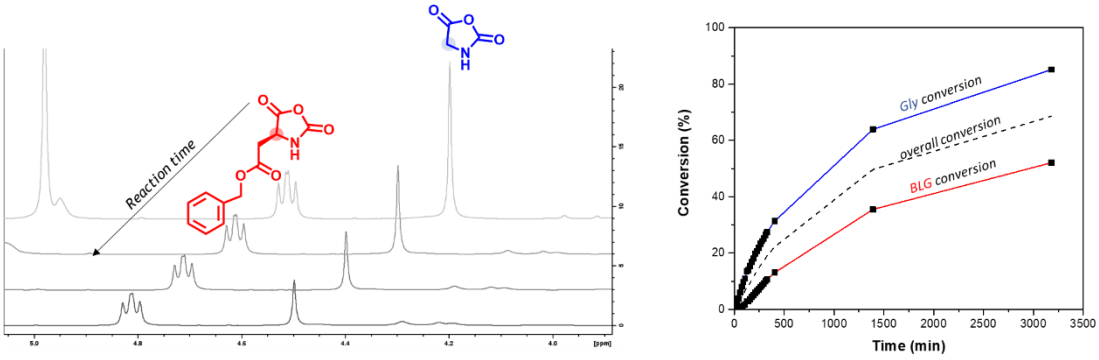


Figure 28: *In-situ NMR monitoring of a copolymerization of initial composition* $f_{G,0}$ = $f_{B,0}$ *= 0.5. On the left:* ^{1}*H NMR spectra at different points of the copolymerization reaction highlighting the peaks of the* C_α *of each monomer. On the right is plotted the conversion of each monomer in function of the reaction time. The overall conversion was consequently determined and plotted.*

A copolymerization system was then studied (copolymerization system **1**) to calculate the reactivity ratios. The conditions of copolymerization system **1** were DMF (0.1 M concentration of monomers), 0°C and hexylamine initiation and M/I=50 for all experiments. In total, five experiments were done similarly with varying initial monomer feeds: $f_{B,0}$ *= 0.90, 0.80, 0.50, 0.36 and 0.20* The experimental data was then fitted to the copolymerization equation (**Equation 2**) to obtain the reactivity ratios. Since both conversion and fraction are experimentally determined it is necessary to consider the induced error in each. For this purpose, a previously developed[144] nonlinear least squares method, was used assuming non-negligible error in all variables.[145] Using the visualization of the sum of squares space method, the point estimates of r_{Gly} and r_{BLG} were found to be 1.75 and 0.47 respectively (**Figure 29**). This result corroborates the experimental observations mentioned above and the copolymer will have a gradient

like microstructure (**Figure 6**) with a higher concentration of glycine units at one end of the chain.

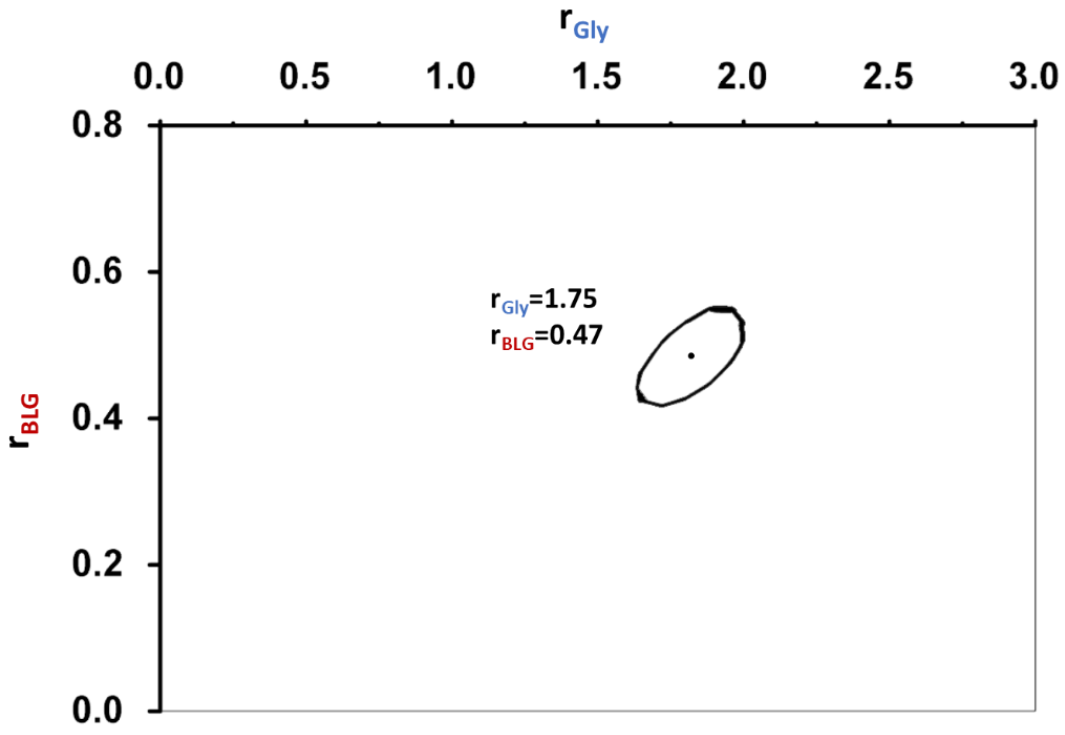


Figure 29: *Point estimate and 95% joint confidence interval (contour) of reactivity ratios of copolymerization system 1.* r_{Gly} *= 1.75 and* r_{BLG} *= 0.47.*

In order to confirm the accuracy of the results obtained it is important to compare the experimental data to the modelled results (obtained using **Equation 2**). This is plotted in **Figure 30** and shows reasonable agreement between the two, although some deviations can be observed at higher conversion rate. Therefore, the calculated reactivity ratios reflect the composition of the copolymers obtained through this system. While the instantaneous feed ratio of BLG varies little during the beginning of the reaction, the reaction mixture is getting slowly enriched with BLG NCA. As Gly-NCA is consumed more rapidly, the molar fraction of BLG NCA will increase towards the end of the copolymerization.

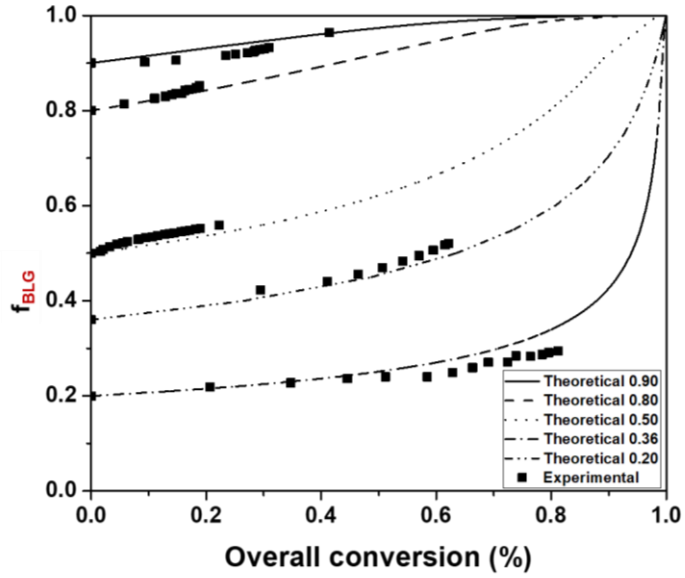


Figure 30: *Modelled monomer (BLG NCA) compositions in function of the overall copolymerization conversion in the case of the five experiments listed above with varying initial monomer feeds (copolymerization system **1**). These plots are compared with the experimental data from all the experiments.*

The gradient copolymer **1** obtained from this system could also be schematically represented on the scale of an individual chain using a Monte Carlo simulation to calculate the probability of insertion of one monomer unit at a time. In **Figure 31**, 20 chains are presented, with an average number of units equal to the number-average degree of polymerization of the polymer and a length distribution with dispersity of 1.1 (**Figure 7**). The probability that a given unit is assigned as Gly or BLG is determined by the identity of the previous unit and the monomer ratio at the conversion corresponding to that chain length calculated from the reactivity ratios.[144]

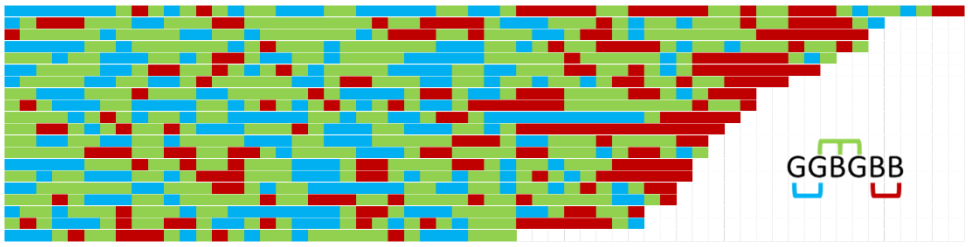


Figure 31: *Simulated 20 polymer sequences of copolymer 1. The direction of polymerization is from left to right. Gly-Gly dyads are shown in blue (28%), Gly-BLG and BLG-Gly dyads in green (51%) and BLG-BLG dyads in red (28%). Since the dispersity is taken into account, the polymer chains are represented in order of their length from top to bottom.*

Copolymer **1** has thus a small gradient nature with a higher concentration of Gly-Gly dyads at the beginning of the chain and more BLG-BLG dyads at the other end with an overall high concentration of alternating dyads (**Figure 31**). The copolymer was revealed to have highly different macroscopic properties from polymer **2** ($PBLG_{50}$). The copolymer was insoluble in DMF and copolymerization system **1** allowed to map the solubility of the system depending on the initial feed ratio and the conversion (**Figure 32**).

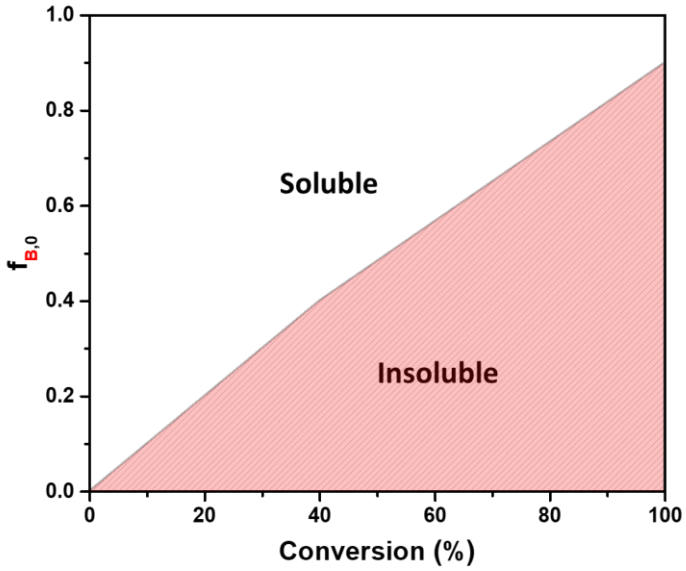


Figure 32: *Phase diagram of copolymerization system 1. The initial feed ratio was plotted in function of the conversion and the cloud points were determined by looking at the state of the solubility of the polymer in NMR tube. At low $f_{B,0}$ the copolymer became insoluble after a certain degree of conversion in contrast to high $f_{B,0}$ systems where it was soluble until high conversion. $f_{B,0}$ = 1 is the homopolymerization where the polymer is soluble at all stages.*

2.2.2. Effect of reaction conditions on the reactivity ratios

Effect of feed ratio

As a first step to assess the effect that the reaction conditions could have on the copolymerization of BLG and Gly-NCA, a series of copolymers was synthesized based on the copolymerization system **1** described above. For this purpose, $f_{B,0}$ value was varied from 0 to 0.5 (**Table S3**). From the reactivity ratios calculated the copolymer chains were represented to visualize differences in composition profile. As the initial

feed ratio of BLG NCA decreases, alternating dyads start to appear and are dispersed along the chain. For copolymers **4-6**, more Gly-Gly dyads start to appear and are more concentrated at the beginning of the chains.

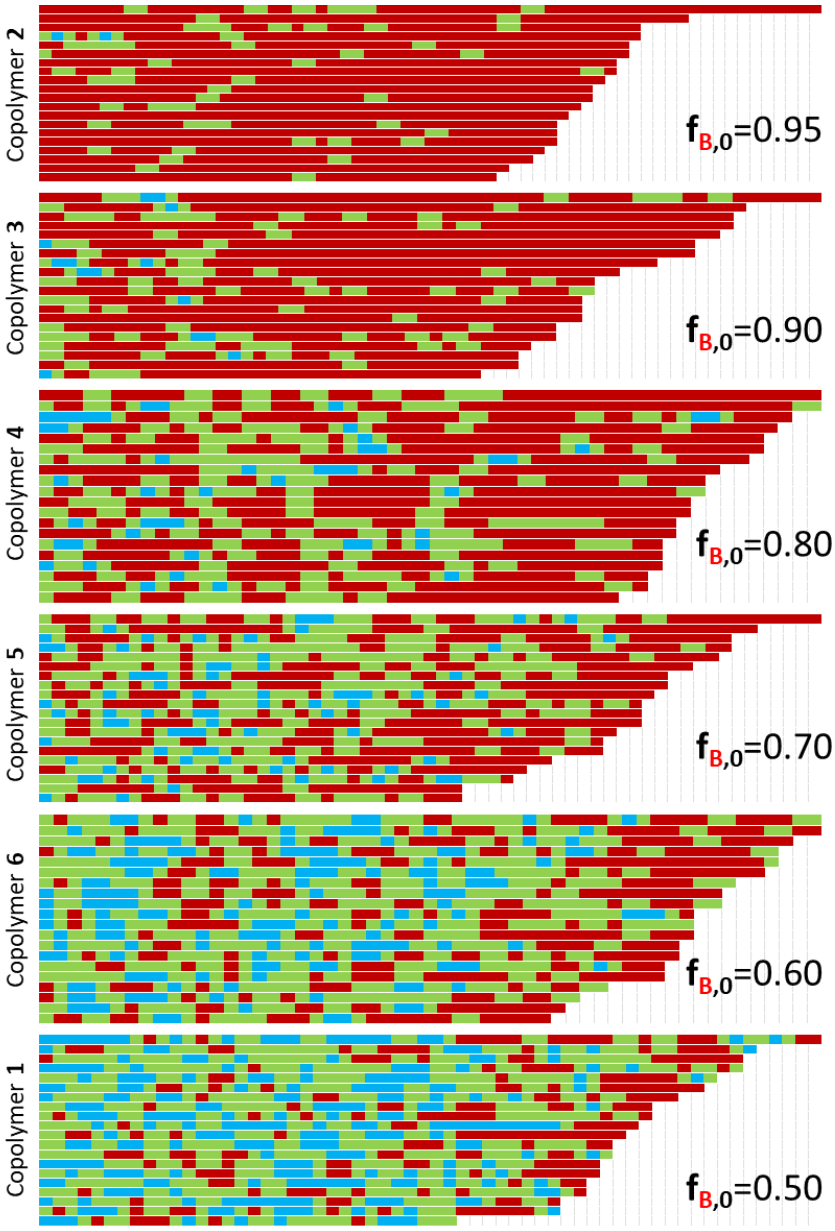


Figure 33: *Simulated 20 polymer sequences of copolymers 1-6. The direction of polymerization is from left to right. Gly-Gly dyads are shown in blue, Gly-BLG and BLG-Gly dyads in green and BLG-BLG dyads in red. Since the dispersity is considered, the polymer chains are represented in order of their length from top to bottom.*

Effect of solvents

The change of reaction solvent might influence the kinetics of the copolymerization. In radical polymerization, the polarity of the solvent and its dielectric properties can

stabilize different propagating species thus changing the reactivity ratios.[23] As mentioned in the introduction of this chapter, the solvent nature dictates the secondary structuring and the solubility of the growing copolymers' chain and this can influence the reactivity ratios.[39,43] To study this effect in the case of BLG and Gly NCA copolymerization, three new copolymerization systems (**2** – **4**) were studied: DMF (copolymerization system **2**), THF (copolymerization system **3**) and DMSO (copolymerization system **4**). The use of DMSO makes it difficult to conduct the polymerization at low temperature as it solidifies. For this reason, the copolymerizations were studied at 25°C and all other criteria kept the same as copolymerization system **1** (0.1M concentration of monomers and hexylamine initiation and M/I=50). Using the technique described above the reactivity ratios were calculated for the three systems (see **Supporting information**).

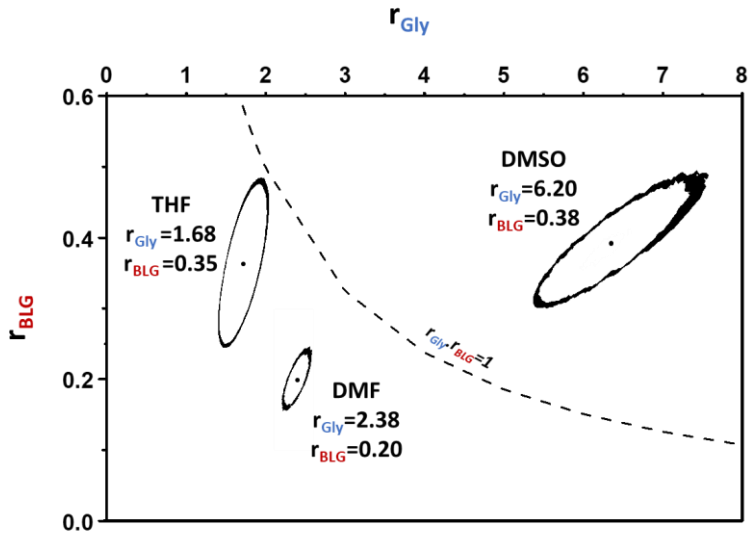


Figure 34: *Point estimate and 95% joint confidence interval (contour) of reactivity ratios of copolymerization system* ***2-4****. The values of system* r_{Gly} *and* r_{BLG} *from each solvent system are shown. The dotted line corresponds to the case of an ideal copolymerization where the difference in reactivity between the comonomers is independent of the propagating chain end.*[146]

The solvent had a high influence on the reactivity ratios of Gly and BLG NCA (**Figure 34**). The product of the resulting reactivity ratios ($r_{Gly} \times r_{BLG}$) gives an indication of the behaviour of the copolymerization. If this value is equal to unity, the copolymerization is considered to be ideal meaning that difference in reactivity is independent of the type of propagating species but only on the inherent reactivities of the monomers (**Figure**

34).[146] Values of $r_{Gly} \times r_{BLG}$ lower than one indicates a tendency of the copolymerization towards alternating repeat units. The lowest $r_{Gly} \times r_{BLG}$ was obtained in THF (0.59) while in DMF $r_{Gly} \times r_{BLG}$ was found to be 0.82. In the case of DMSO $r_{Gly} \times r_{BLG}$ was equal to 2.36 which indicates a tendency to form blocks of homopolymer. In all three cases the copolymers have the same gradient nature, this is more pronounced for DMF and DMSO which have a higher difference in reactivity ratios (**Figure 35**). Glycine is highly reactive and the difference in reactivity is expected. The difference in results between the solvents is explained by the solubility of the polymer chains, the polarity of the solvents, their protic nature, and their capacity to form hydrogen bonds. DMF and DMSO are more polar and can form hydrogen bonds with the growing polymer chains thus preventing the α-helical secondary structuring of PBLG homosequences during the reaction. This structuring is known to accelerate the ROP of BLG NCA due to macrodipolar affinity.[78] In contrast, formation of secondary structure could be observed in THF, a helicogenic solvent,

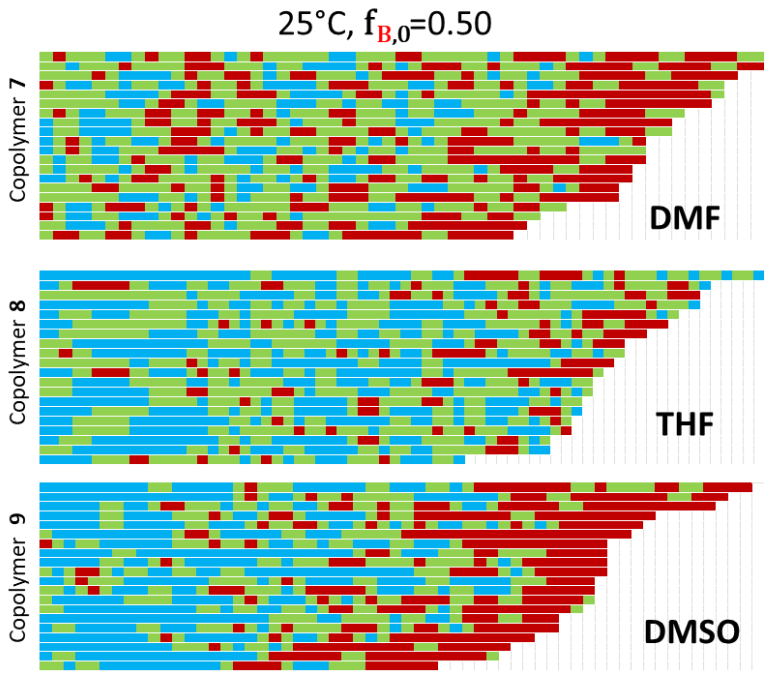


Figure 35: *Simulated 20 polymer sequences of copolymer systems **2-4**. The direction of polymerization is from left to right. Gly-Gly dyads are shown in blue, Gly-BLG and BLG-Gly dyads in green and BLG-BLG dyads in red. Since the dispersity is considered, the polymer chains are represented in order of their length from top to bottom.*

Three copolymers (**7-9**) were synthesized in these three solvents to elucidate the difference of behaviour (**Table 2**).

Table 2: *Synthesized copolymers to show the effect of the reaction solvent on the properties of the copolymers. All synthesis were conducted in the following reaction: $f_{B,0}$=0.5, 0.1M of initial monomer concentration, 25°C and initiated with hexylamine(M/I=50).*

	Reaction solvent	Yield (%)	BLG% by ^{1}H NMR[a]	Overall Dp by ^{1}H NMR[b]
Copolymer **7**	DMF	85	50	49
Copolymer **8**	THF	86	48	50
Copolymer **9**	DMSO	69	50	52

[a]as determined by the integration of the glycine and the PBLG peaks in ^{1}H NMR in TFA-d; [b]as determined by the ratio between the integration of the HA peaks and the backbone peaks in ^{1}H NMR in TFA-d.

The three copolymers had different behaviours in each solvent. Only copolymer **9** was soluble during polymerization in DMSO, while copolymer **7** and **8** precipitated at 40% and 50% conversion respectively. While copolymers **7-8** were only soluble in HFIP for SEC analysis, copolymer **9**'s SEC analysis was done in DMF (**Table S4**) since it was not soluble in HFIP. These copolymers were further characterized by ^{13}C NMR in TFA-d to elucidate carbonyl peaks of homo-sequences and hetero-sequences (**Figure 36**). The homo-sequences peaks were determined by analysing a pure PBLG (Polymer **2**) and PGly **3**.

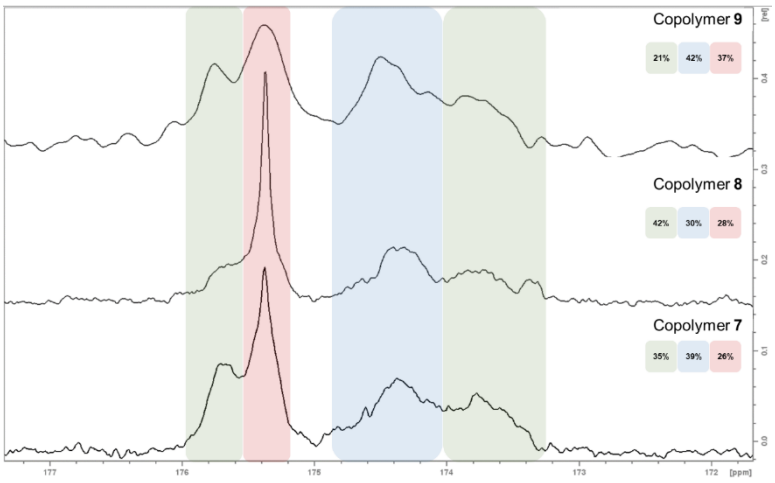


Figure 36: *Quantitative ^{13}C NMR of copolymers **7-9**. The spectra are zoomed on the carbonyl region between 170-180 ppm. The different peak zones are highlighted in colours; Gly-Gly carbonyl peaks are shown in blue, Gly-BLG and BLG-Gly carbonyl peaks in green and BLG-BLG carbonyl peaks in red. The percentages of the different peak regions are shown and were determined by peak fitting and deconvolution.*

BLG repeat units' carbonyl peaks are deshielded when the previous unit is glycine (blue peaks at 175.4ppm). On the other hand, glycine repeat units' carbonyl peaks are shielded when the previous unit is BLG (red peaks at 174.4 ppm). By fitting and deconvoluting the different peaks (**Figure S20**), it was possible to calculate the percentage of homo-dyads and hetero-dyads (**Figure 36** and Table 3).

Table 3 : *Dyads percentages as determined by both the copolymer chains' simulation and* ^{13}C *NMR for copolymers* ***7-9***

	Monte Carlo simulation			^{13}C NMR		
	Gly-BLG% BLG-Gly%	BLG-BLG%	Gly-Gly%	Gly-BLG% BLG-Gly%	BLG-BLG%	Gly-Gly%
Copolymer **7**	51	19	29	35	26	39
Copolymer **8**	54	20	26	42	28	30
Copolymer **9**	31	28	41	21	37	42

Although the values from both techniques were not the same, the tendencies were similar. Both techniques confirmed that copolymer **9**, with highest gradient effect, had the lowest alternating dyads content. This means that copolymer **9** had the highest content of homosequences of both monomers. For all copolymers and in both techniques, more Gly-Gly dyads were observed than the BLG-BLG dyads. This is consistent with the reactivity ratios obtained where $r_{Gly} > r_{BLG}$. The results obtained by ^{13}C NMR confirm the difference in microstructure of copolymers **7-9** confirming that the solvent influences the reactivity ratios of Gly NCA and BLG NCA. Thus, by changing the reaction solvent the primary structure of Gly-BLG copolymers can be stochastically controlled to obtain higher alternation.

Reaction temperature

In copolymerization system **1** (0°C in DMF, section 2.2.1), the reactivity ratios were calculated to be r_{Gly}=1.75 **and** r_{BLG}=0.47 while in the same solvent at 25°C (copolymerization system **2**,0) the values were 2.38 and 0.2 respectively. The difference in reactivity ratios obtained for these two copolymerization systems show that the temperature can also influence the kinetics. Similar studies were done at 50°C in DMF (copolymerization system **5**) and at 0°C in THF (copolymerization system **6**) (**Table S5**) to further highlight the effect of reaction temperature on copolymerization kinetics in both these solvents. As shown in **Figure 37**¸ the increase in temperature

significantly increases the value of r_{Gly} and has little effect on the value of r_{BLG}. This effect was more significant in DMF than in THF at both these temperatures.

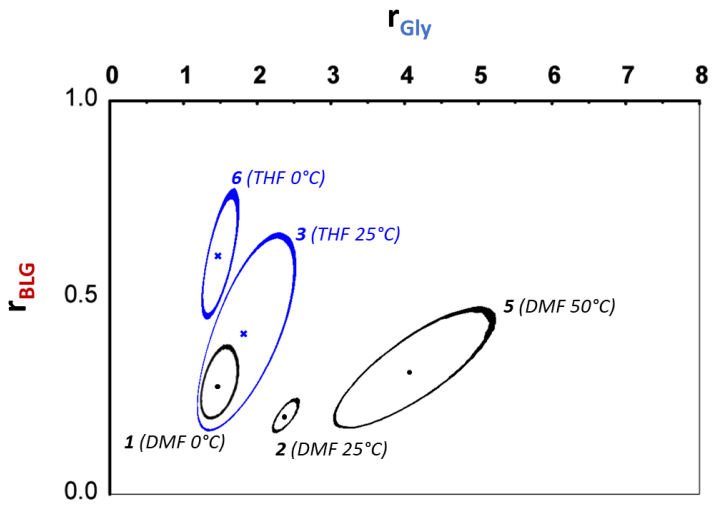


Figure 37: *Point estimate (black circles or blue crosses for DMF and THF systems respectively) and 95% joint confidence interval (contour in black or blue for DMF and THF systems respectively) of reactivity ratios of copolymerization systems* ***1,2,3,5*** *and* ***6****.*

By looking at these results, the increase in temperature increased the difference between the two reactivity ratios leading to an increase in the gradient extent (see **Figure 38**).

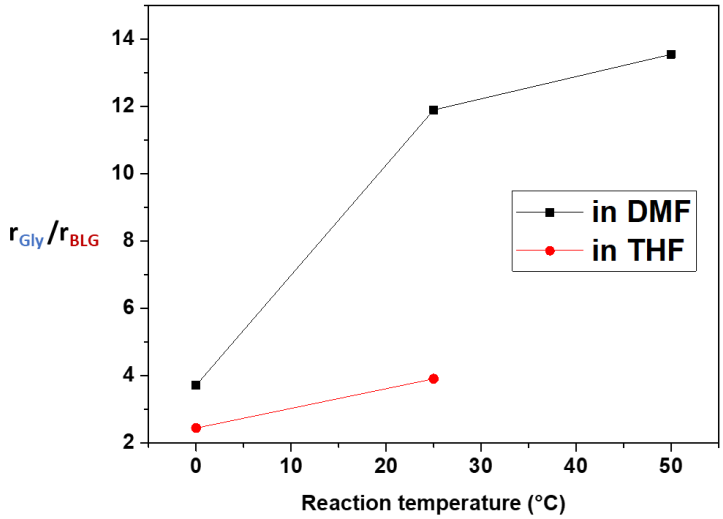
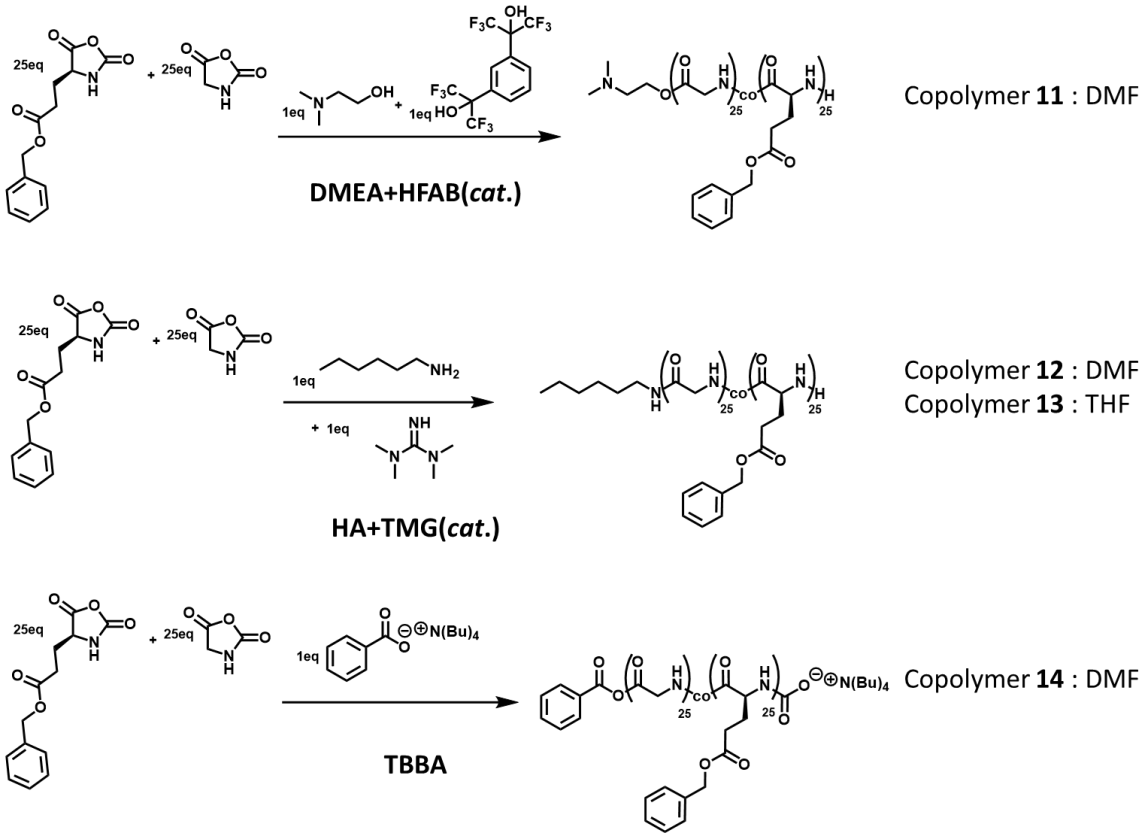


Figure 38: *Effect of the temperature on the relative ratio between* r_{Gly} *and* r_{BLG} *for synthesis done in THF and DMF. These results are for copolymerization systems* ***1,2,3,5*** *and* ***6****.*

Increasing the reaction temperature increases the overall reactivity of the copolymerization. In the studied system, glycine had its reactivity increase more than that of BLG, producing copolymers with higher gradient nature. Although this behaviour could be useful it was not investigated further as of concerns of chain end capping due to the BLG units at higher temperatures discussed earlier.

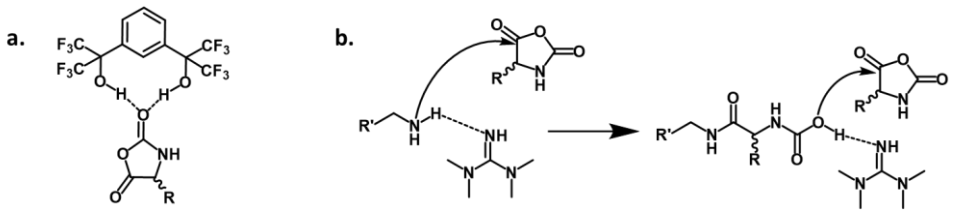
ROP mechanism

The ROP mechanism could also act as a lever to change the copolymerization kinetics. For this purpose, three different ROP systems were explored to be compared to the classical amine-initiated ROP discussed above (**Scheme 7**)



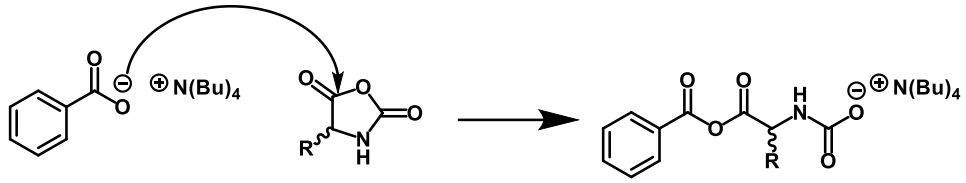
***Scheme 7**: General reaction schemes for copolymers **11-14**. All copolymers were conducted in DMF except for copolymer 12. DMEA=dimethylethanolamine; HFAB=1,3-Bis(2-hydroxyhexafluoroisopropyl)benzene; TMG=tetramethylguanidine; TBBA=tetrabutylammonium benzoate.*

These three systems were chosen for the difference of their reaction mechanisms. In the first catalytic system, the fluorinated alcohol (HFAB) creates hydrogen bonds with the NCAs (**Scheme 8**) allowing for an easier ring opening reaction when attacked by the initiator or the growing chain end.[147]



Scheme 8: *Postulated mechanism of a) HFAB on NCA[147] and b) TMG on the ROP of NCA[148].*

For the second system, tetramethylguanidine (TMG) mediated catalysis, the mechanism involves the interaction between the organobase and the initiator or the propagating chain end increasing their nucleophilicity. This catalyst was shown to work in different solvents[148,149] and was employed in both THF (copolymer **12**) and DMF (copolymer **13**) as the polarity of the solvent can affect the reaction.



Scheme 9*: Initiation of the ROP of NCA using TBBA.*[150]

In the last system, copolymer **14** (TBBA), activating effect comes from the stabilization of a carbamate anion (**Scheme 9**) as propagating species due to the large alkylammonium counter ion.[150] This is also very similar to the effect TMG has on the chain end shown above.

The ROP for all these systems was extremely fast and the use of *in-situ* NMR to calculate the reactivity ratios proved therefore difficult. Indeed, using TBBA and DMEA/HFAB in catalysis, it was not possible to determine the reactivity ratio. On the other hand, when using TMG as a catalyst, reaction monitoring was possible, and the results obtained varied depending on the solvent that was used. In this case the use of TMG in DMF (copolymer **12**) induces a smaller gradient effect than without the use of the catalyst as the reactivity ratio of BLG is increased from 0.2 to 0.96 while for glycine no significant change was observed. For THF the reactivity ratios were both found to be above 1 (r_{Gly}=4.50 and r_{BLG}=1.42). This last result suggests an interesting multiblock microstructure of the resulting copolymer that was not characterized yet in this study.

***Table 4** : Dyads percentages as determined by both the copolymer chains' simulation and ^{13}C NMR for copolymers **11-14** when possible.*

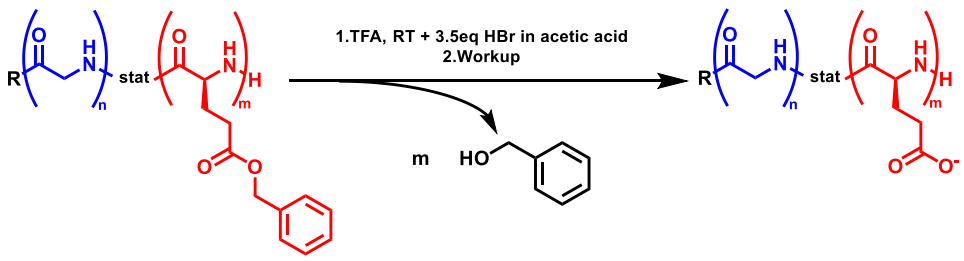
	Monte Carlo simulation			^{13}C NMR		
	Gly-BLG% BLG-Gly%	BLG-BLG%	Gly-Gly%	Gly-BLG% BLG-Gly%	BLG-BLG%	Gly-Gly%
Copolymer **11**	*a.*			42	29	29
Copolymer **12**	39	25	36	*b.*		
Copolymer **13**	28	31	41	40	29	31
Copolymer **14**	*a.*			48	28	24

[a] The speed of the reaction did not allow for reactivity ratios calculations; [b] the copolymer precipitated out of the solution during analysis in TFA-d.

The resulting copolymers were then analysed by ^{13}C NMR to establish the relative content of each dyad as described above. As shown in **Table 4**, copolymer chain simulations were only obtained for copolymers **12** and **13**. For the latter, the ^{13}C NMR analysis did not give values similar to those of the simulation and thus no conclusions can be made. Furthermore, the values obtained were close to those obtained for **8**. It was possible to see from ^{13}C NMR analysis of copolymers **11** and **14** that the use of catalysis allowed to increase the percentage of alternating dyads in comparison to copolymer **7**. The change in ROP mechanism and catalysis thus has an effect on the primary structure of copolymerizations of Gly NCA and BLG NCA. It was also possible to access new microstructures, multiblock copolymers, when TMG catalysis was used in THF.

2.2.3. Effect of reaction conditions on the secondary structure

In this part, the secondary structures of copolymer synthesized in the different conditions discussed above was analysed. The copolymers were deprotected for analysis in aqueous conditions (**Scheme 10**).



***Scheme 10:** Deprotection of copolymers for secondary structure analysis.*

Effect of feed ratio

Copolymers **1-6** were deprotected to yield glutamate-containing PGly-*co*-PGA copolymers **1D-6D**. The synthesized copolymers were in reasonable agreement with their theoretical composition by ^{1}H NMR (**Table S3** and **Table 5**). The difference in solubility of the copolymers in their initial and deprotected states complexified the SEC analysis (**Table S4**). Furthermore, since dn/dC calculations are necessary to determine Mn values, this could not be done for each copolymer independently. Nonetheless, the formation of high molecular weight copolymers could be verified, and the values presented in **Table S4** are indicative.

Table 5: *Synthesized copolymers to show the effect of monomers' initial feed ratio on the properties of the copolymers. All synthesis were conducted in the reaction conditions of copolymerization system 1 (0.1M in DMF, 0°C and initiated with hexylamine with M/I= 50) and subsequently deprotected according to previously discussed methods. Polymer 2D is inserted for comparison.*

	$f_{B,0}$[a]	Structure (theoretical)	PGA% by ^{1}H NMR[b]	Overall DP by ^{1}H NMR[c]
Polymer **2D**	100	HA-PGA$_{50}$	100	55
Copolymer **2D**	95	HA-PGA$_{47}$-co-PGly$_3$	93	57
Copolymer **3D**	90	HA-PGA$_{45}$-co-PGly$_5$	85	60
Copolymer **4D**	80	HA-PGA$_{40}$-co-PGly$_{10}$	80	64
Copolymer **5D**	70	HA-PGA$_{35}$-co-PGly$_{15}$	65	63
Copolymer **6D**	60	HA-PGA$_{30}$-co-PGly$_{20}$	57	50
Copolymer **1D**	50	HA-PGA$_{25}$-co-PGly$_{25}$	47	51

[a]*initial monomer feed and expected PGA% in final copolymer; HA: hexylamine;* [b]*as determined by the integration of the glycine and the PGA peaks in ^{1}H NMR in D_2O;* [c]*as determined by the ratio between the integration of the HA peaks and the backbone peaks in ^{1}H NMR in D_2O;* [d]*as determined by the Mn value from dn/dC through MALS detection, the used value was that of homoPGA (0.1489) thus inducing an error in the calculation of the copolymers' Mn value, which are just indicative;* [e]*as determined by MALS detection.*

The copolymers were then analysed by CD to assess their secondary structures. **Figure 39** shows the different CD spectra of the copolymers at pH=4 where PGA adopts an α-helical structure. In comparison to polymer **2D** spectra (**Figure 19**) the intensity of the peaks is much lower even with low glycine incorporation (copolymer **2D**). Since α-helical structuring requires long sequences of PGA units,[108] small incorporations of glycine units may disrupt the structure, leading to lower helical content.

This can be visualized by the simulation of chains from copolymer **2D**, as shown in **Figure 33**, where the 5% copolymerized glycine units are incorporated at different points along the different chains, possibly disrupting long distance helical form.

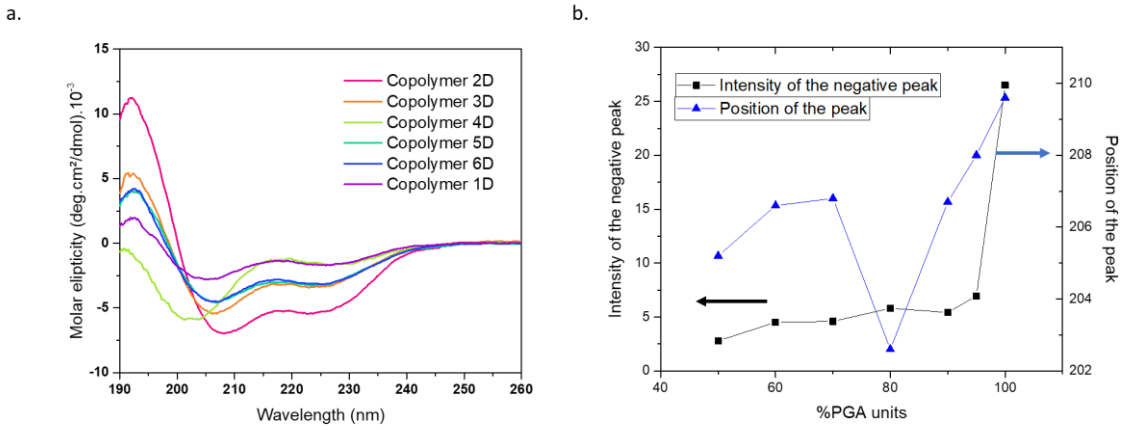


Figure 39: *a) CD spectra of copolymers **1-6(D)** in helix inducing media (pH=4); b) data from a. and Figure 17 was plotted to show the effect of the %PGA units in polymer 2D and copolymers **1-6(D)** on the intensity of the negative peak around 209nm and its position.*

The CD spectra also show the gradual decrease of the α-helical peaks with decreasing PGA% in the copolymers as well as the shifting of the peak maximum towards lower wavelengths (**Figure 39** b.). Since the signals from CD spectra are additive,[111] the coexistence of the two forms, α-helix and random coil, will give an intermediate signal. Thus the negative peak is shifting from an α-helical form (>209nm) towards a more random coil conformation (<200nm) (**Figure 19**) and all the detailed peaks in **Figure 39** b. could be a combination of both signals. Deconvolution could also be done to calculate the percentage of each, but they are model-dependant[111] and were not conducted for this study. Copolymer **4D** has a highly random conformation compared to other copolymers even those that have a higher glycine content. This composition could be consistent with high incorporation of glycine at the beginning of the copolymer destructuring the copolymer and not allowing for formation of long PGA sequences at the other end of the copolymer to allow for helical structuring. This could also be an outlier and the experiment needs to be repeated to confirm any hypothesis.

Effect of reaction solvent

Copolymers were prepared in the different solvents at 25°C and were subsequently deprotected and analysed by CD. The resulting copolymers are presented in **Table 6**.

Table 6: *Synthesized copolymers to show the effect of the reaction solvent on the properties of the copolymers. All synthesis were conducted in the following reaction:* $f_{B,0}$*=0.5, 0.1M of initial monomer concentration, 25°C and initiated with hexylamine(M/I=50).*

	Reaction solvent	Yield (%)	BLG% by ^{1}H NMR[b]
Copolymer **7** Copolymer **7D**	DMF	85 82[a]	50
Copolymer **8** Copolymer **8D**	THF	86 90 [a]	48
Copolymer **9** Copolymer **9D**	DMSO	69 80 [a]	50
Copolymer **10** Copolymer **10D**	DCM	90 97 [a]	49

[a]*after HBr deprotection;* [b]*average value as determined by the integration of the glycine and the BLG peaks in* ^{1}*H NMR in both TFA-d (before deprotection) and* D_2O *(after deprotection).*

A copolymer was synthesized in DCM, a less polar solvent than THF. The calculation of the reactivity ratios using the same method was not possible for copolymerization in DCM, as Gly-NCA is only slightly soluble (Table S2). This could help reduce the relative reactivity of Gly-NCA to BLG-NCA. While copolymer **9** was soluble from start to finish and copolymer **10** was synthesized in a heterogenous manner, copolymers **7** and **8** started out homogeneously but the chains precipitated at different times. The secondary structure analysis of copolymers **7D-10D** showed varying behaviours in helix-inducing conditions. Copolymer **8D**, synthesized in THF showed the highest helical conformation, followed by copolymer **10D** synthesized in DCM. In contrast, copolymers **7D** and **9D** both showed more random coil behaviour with the negative peak shifted towards lower wavelengths. This was more pronounced for copolymer **9D** synthesized in DMSO with the highest gradient nature (**Figure 40** a.). These results could corroborate the hypothesis that the reactivity of the comonomers can be affected by the nature of the solvent and the secondary structuring of the growing chains. Using these results, the relative reactivity ratios of copolymer **10D** could be estimated to be between that of copolymer **7D** and of copolymer **8D** (**Figure 40** b.). More data is of course needed to establish a complete model for such calculations.

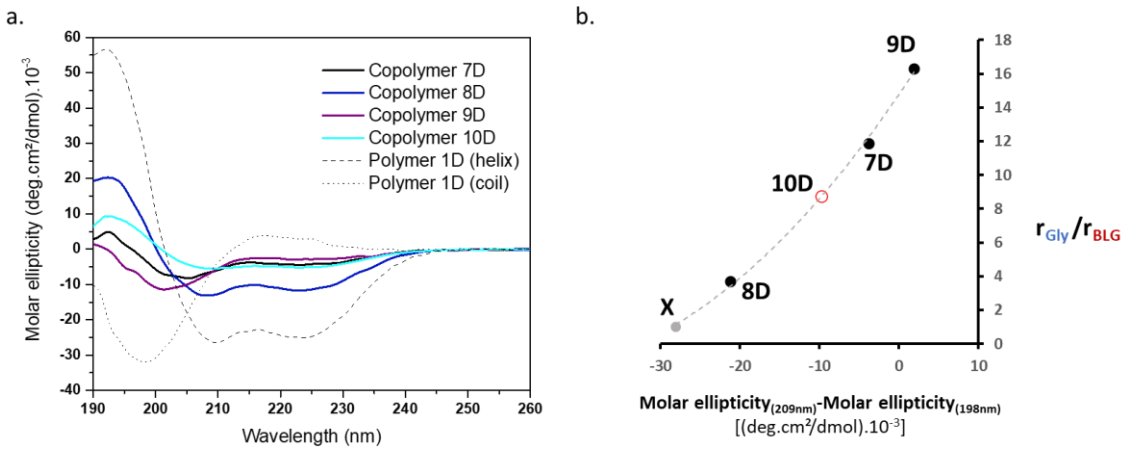


Figure 40: *a.CD spectra of copolymers **7D-10D** in helix-inducing aqueous media (pH=4). The two extreme conformations of Polymer **1D** are shown for easier understanding of the reader (helix form at pH=4 and coil form at pH=6). b. Plot of the ratio between r_{Gly} to r_{BLG} in function of the difference of the molar ellipticity of the copolymers **7D-9D** at 209nm (α-helix peak) and 198nm (random coil peak). The line is drawn to easily visualize the trend. The red circle corresponds to copolymer **10D** and is an estimation based on its CD signal. The grey circle (X) is an estimation of the values for a perfectly alternating copolymer of 25/25 glycine/glutamate. r_{Gly}/r_{BLG}=1 and its molar ellipticity value is equal to half of polymer 1D's value.*

Copolymer **8D** having the most random distribution by the closer value of reactivity ratios has the most helical nature. It is thus hypothesized that an even distribution of glycine along the PGA chain does not disrupt its helical structure. In other words, replacing, in a regular matter along the chain, glutamate units by achiral units without side chains in a PGA polymer still allows the formation of an alpha helical conformation. The helicity signal from copolymer **8D** is around half of that of polymer **1D**. This is because the copolymer contains half as many glutamate units that give out the chiral signal. On the other hand, if the monomer distribution is less even, the conformation is disrupted. A highly gradient copolymer, like copolymer **9D**, has a part with high concentration of glycine that does not form alpha helical structure, and on the other end of the copolymer, a higher concentration of glutamate unit that do not form a long enough chain to form a helix structure. This hypothesis can also explain the behaviour observed for copolymer **4D** highlighted above. Having small amounts of glycine had little effect on the helicity, but at a certain glycine percentage the conformation was disrupted. Adding more glycine allowed for structuring again as they were more

dispersed over a longer range. More experiments are needed to elucidate the conformational behaviour of such copolymers and such knowledge could help in comprehending how natural proteins adopt their different secondary structures.

Effect of catalysis and ROP mechanism

Copolymers were prepared using the different catalytic conditions and were subsequently deprotected to be analysed by CD.

Table 7: *Synthesized copolymers to show the effect of catalysis and ROP mechanism on the final properties of the copolymers. All synthesis were conducted in the following reaction:* $f_{B,0}$*=0.5, 0.1M of initial monomer concentration and 25°C.*

	Reaction solvent	Catalyst	Initiator
Copolymer **11D**	DMF	HFAB	DMEA
Copolymer **12D**	DMF	TMG	HA
Copolymer **13D**	THF	TMG	HA
Copolymer **14D**	DMF	-	TBBA

Only copolymer **13D** showed a high helical form thanks to synthesis in THF catalysed by TMG (**Figure 41**). The other copolymers showed moderate helicity comparable to copolymers synthesized with classical NAM.

The values of reactivity ratios for copolymer **12D** were close to copolymerization systems in DMF without catalysis while those in THF (copolymer **13D**) were both higher than 1. This multiblock copolymer shows increased helicity in comparison to other systems suggesting the possibility of glutamate sequences long enough for this secondary structuring. The difference in the different systems also suggests that the catalyst and/or the ROP mechanism could depend on the nature of the comonomer or the propagating species. Although the copolymerization is much faster in all these conditions glycine still reacts faster than BLG. The steric hindrance around the BLG units thus reduces its rate of reaction.

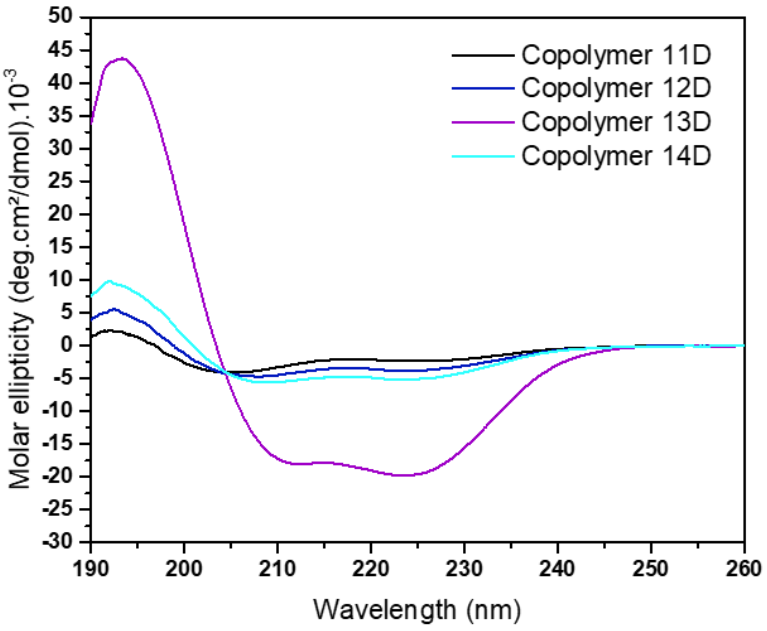


Figure 41: *CD spectra of copolymers* ***11D-14D*** *in helix-inducing aqueous media (pH=4).*

3. Conclusions

In this chapter the copolymerization of BLG-NCA and glycine-NCA was studied as a model system to design synthetic IDP using ROP of NCA. This model system has enabled us to develop a convenient method to calculate the reactivity ratios during a ROP of different NCAs. These ratios determine the microstructure of the obtained copolymers and can be used to steer the copolymerization towards a desired statistical sequence. The levers that can affect the reactivity ratios are the reaction conditions and their effect was explored. Feed ratio, solvent, temperature, catalysis, and ROP mechanism were changed to observe their effect on the copolymerization kinetics. It was found that by changing one condition at a time, different behaviours can be obtained from the ROP. This method can then be used to obtain copolymers with different microstructural properties using the same batch reaction of the two comonomers. This primary sequence was analysed by Monte Carlo simulations and ^{13}C quantitative NMR analysis to highlight the differences between the different synthesized copolymers. This information was then used to compare their secondary structure after deprotection. First the feed ratio of both comonomers was varied with all other conditions kept the same. By increasing the glycine ratio, the helical nature

decreased linearly. At a ratio of 50%, the polymer had low helicity compared to homoPGA but still retained a helical CD signature. Then, the polarity of the solvent was changed at a feed ratio of 50% and it was found that in higher polarity solvents the gradient nature of the produced copolymers was accentuated. This resulted in the formation of copolymers based on glutamate with a random coil conformation as the alternating dyads were decreased as measured by ^{13}C NMR and by the reactivity ratio-based simulations. Thus, homosequences with an asymmetrical distribution along the copolymer chain allowed to synthesize disordered PGA based copolymers. Furthermore, when higher alternating sequences were observed, and their distribution was more spread out along the chains, the copolymer retained high α-helical content. Similar conclusions were made when the temperature was changed. The increase in temperature increased the gradient nature. Using different ROP mechanisms and catalysis it was possible to synthesize copolymers with varying microstructures. A multiblock copolymer using TMG catalysis in THF produced a highly helical copolymer. Although it was hypothesized that distributing glycine units evenly (increasing alternating dyads) will destructure the helical nature of PGA, it was found that removing evenly along the copolymer the side chains of glutamate unit it was possible to retain its α-helical structure. We hypothesize that replacing glutamate with glycine at random points along the polypeptide chain without having high concentration of glycine homo-sequence allows the structuring. Experimental conditions thus allow to exert a degree of control over the distribution of comonomers along the polymer chains. The choice of the two comonomers allow to mimic the properties of IDPs. Production of intrinsically disordered polyglutamates by adding glycine units was shown to be possible using the appropriate reaction conditions.

4. Supporting information

4.1. Materials and methods

All chemicals and solvents in this work were purchased from Sigma Aldrich, Fluorochem, Acros, TCI, Strem and VWR, and were used without any purification unless otherwise described. DMF, THF, DCM, acetonitrile and cyclohexane were obtained from a solvent system purifier (PureSolv, Innovative Technology), kept under argon atmosphere and freshly used. Milli-Q water was obtained from a Purelab Prima purification system (ELGA) with a resistivity of 18.2 MΩ cm-1. The N-carboxyanhydride

monomers were purchased from PMC Isochem, stored at 20 °C under argon atmosphere and weighed in the glove box Jacomex GP13 no. 2675 at the Laboratoire de Chimie des Polymères Organiques (LCPO, Bordeaux, France). Dialysis membranes were purchased from SpectrumLabs. Hexylamine and TMG were freeze-thawed and cryo-distilled on the Schlenk line prior to use. Deuterated solvents used for in-situ NMR monitoring were dried before use. THF-d was dried using benzophenone/Na and cryodistilled before use. DMF-d was stored on activated 3A° molecular sieves for at least one week and cryodistilled before use. DMSO-d was stored on activated 3A° molecular sieves for at least one week before use.

4.1.1. Nuclear Magnetic Resonance (NMR)

^{1}H NMR 400 MHz spectra were obtained using a Bruker Avance I (Liquid-state 400 MHz NMR spectrometer with 5 mm BBFO probe). In-situ ^{1}H NMR and quantitative ^{13}C NMR were both conducted on a spectra were recorded on a 500 MHz Prodigy NMR AVANCE NEO 500. Solvent peaks were used for reference for the locking. The spectra obtained were calibrated using the residual solvent signals. The signals were categorized as follows: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m) and broad (br).

4.1.2. Size Exclusion Chromatography (SEC) Analyses

SEC DMF/DMSO

Polymer molar masses were determined by Size Exclusion Chromatography (SEC) using dimethyformamid (DMF + lithium bromide LiBr 1g/L) or DMSO as the eluent. Measurements in DMF/DMSO were performed on an Ultimate 3000 system from Thermoscientific equipped with diode array detector DAD. The system also include a multi-angles light scattering detector MALS and differential refractive index detector dRI from Wyatt technology. Polymers were separated on two Shodex Asahipack gel columns GF310 and GF510 (300 x 7.5 mm) (exclusion limits from 500 Da to 300 000 Da) at a flowrate of 0.5 mL/min. Columns temperature was held at 50°C. Mn values were calculated using dn/dC value of 0.1197 calculated for polymer **2**. Pullulan from Agilent was used as the standard.

SEC HFIP

Two methods were used:

1. Polymer molar masses were determined by Size Exclusion Chromatography (SEC) using hexafluoroisopropanol (HFIP + 0,05% KTFA) as the eluent. Measurements in HFIP were performed on an Ultimate 3000 system from Thermoscientific equipped with diode array detector DAD. The system also includes a multi-angles light scattering detector MALS and differential refractive index detector dRI from Wyatt technology. Polymers were separated on **two PL HFIP gel columns (300 x 7.5 mm)** (exclusion limits from 200 Da to 2 000 000 Da) at a flowrate of 0.8mL/min. Columns temperature was held at 40°C. Easivial kit of PMMA from Agilent was used as the standard.

2. Polymer molar masses were determined by Size Exclusion Chromatography (SEC) using hexafluoroisopropanol (HFIP + 0,05% KTFA) as the eluent. Measurements in HFIP were performed on an Ultimate 3000 system from Thermoscientific equipped with UV detector. The system also includes differential refractive index detector dRI from Wyatt technology. Polymers were separated on **PSS PFG 100Å and PFG 1000Å columns** (300 x 8 mm, 7µM) (exclusion limits from 100 Da to12 000 000 Da) at a flowrate of 1mL/min. Columns temperature was held at 30°C. Easivial kit of PMMA from Agilent was used as the standard.

SEC Water

Polymer molar masses were determined by Size Exclusion Chromatography (SEC) using aqueous buffer as the eluent. Measurements in water were performed on an Ultimate 3000 system from Thermoscientific equipped with diode array detector DAD. The system also include a multi-angles light scattering detector MALS and differential refractive index detector dRI from Wyatt technology. Polymers were separated on two Shodex OH Pack SB802.5 and SB803 (8*300) (exclusion limit from 500da to 100 000da) columns at a flowrate of 0.6 mL/min with Phosphate buffer 0.01M Na2HNO3, 0.1M HPO42-, NaN3 0.02% pH9. Columns temperature was held at 25°C. Mn values were calculated using dn/dC value of 0.1489 calculated for polymer **2D**.

4.1.3. FTIR spectroscopy

The IR spectra were recorded using an FTIR spectrometer (Vertex 70, Bruker), and the samples were measured with an ATR (GladiATR, Pike Technologies) from Fisher technologies performing 32 scans at the LCPO (Bordeaux, France). The raw data were

obtained with Opus7.5 software. The analyzed compounds were dissolved in anhydrous solvent in the Schlenk vessel under an argon atmosphere at 20 °C at 0.4 M. Samples (50 µL) were taken with an argon purged syringe at predetermined intervals. The samples were dropped onto the ATR unit of the IR spectrometer and measured in the range of 400-4000 cm-1. The monomer conversion was followed using the band height of the C=O stretching at about 1850 cm-1, whose increase is linear in the examined concentration range. The initial value for the band at 100% conversion was corrected using THF or DMF as blank. Samples were taken until no further change in band height was observed in order to determine the monomer conversion and corresponding reaction times (the point were no further change occurred was assumed to correspond to 100% conversion). Data analysis was carried out using Microsoft Excel.

4.1.4. MALDI-MS measurements

MALDI-MS spectra were performed at the CESAMO facility (Bordeaux, France) on an Autoflex maX TOF mass spectrometer equipped with a frequency tripled Nd:YAG laser emitting at 355 nm. Spectra were recorded in the positive-ion mode with an accelerating voltage of 19 kV. For MALDI-MS analyses, polymers' deposits were prepared by dissolving 10 mg of polymer in 1 ml of THF or TFA, and an aqueous solution of 10mg/ml of α-CHCA. Subsequently 2 µL of polymer solution was mixed with 20 µL of matrix solution. 1 µL of this mixture was deposited on the MALDI plate and dried.

4.1.5. CD spectroscopy

CD measurements were performed on a JASCO J-815 spectropolarimeter between 190 nm and 260 nm (far-UV). A quartz cell of 1 mm path length (type: 21/10/Q/1) was purchased from Starna Scientific, Ltd. Spectra were recorded at desired temperatures: 20 °C for standard measurements. The measurement parameters were optimized as follows: sensitivity between 5 and 200 mdeg, 0.01 mdeg resolution, 8 seconds response time (Digital Integration Time), 1 nm bandwidth and 10 $nm.min^{-1}$ scanning rate. Polymer solutions at a monomer unit concentration of 2.5 µM in buffered solutions were used for the measurements. The concentration of buffered solution was of 2.5mM. For pH=4 sodium acetate/acetic acid was used, while for pH=6.5 sodium acetate was used.

4.1.6. ROP of BLG NCA

*Polymer **2***

Typical ROP of BLG NCA was done as follows: BLG NCA (3 g, 11.4 mmol,50eq) was weighed in a previously flame dried Schlenk in a glovebox under pure argon. The powder was then dissolved in 19 mL of anhydrous DMF (0.6 M) at 4°C. Hexylamine (30 µL, 0.23 mmol,1eq) was added and this was considered starting time of the reaction. The solution was stirred for 10 days at 4°C under argon. FTIR allowed to monitor the reaction completion by monitoring the depletion of the bands at 1859 cm^{-1} and 1787 cm^{-1} characteristic of the C=O stretching of NCAs. The polymer was then recovered by precipitation in diethyl ether, twice, and dried under high vacuum. Yield: 1690 mg; 67%, white powder. Sample was then analyzed by ^{1}H NMR (10 mg/ml in CDCl3 containing 15% trifluoroacetic acid). ^{1}H-NMR of the polypeptide backbone (**Figure 15**, 400MHz, δ, ppm): 0.87 (t, 3H, CH3 hexylamine), 1.25 (br, 6H, CH3-(CH2)3-hexylamine), 1.46 (t, 2H, -CH2-CH2-NH- hexylamine), 1.81-2.19 (m, 2H, CH2), 2.44 (t, 2H, CH2, J=6.8 Hz), 3.2 (m,2H, -CH2-NH- hexylamine), 4.59 (m, 1H, CH), 5.03-5.11 (m, 2H, CH2O), 7.25-7.29 (m, 5H, ArH).

4.1.7. Deprotection of PBLG into PGA

Hydrolysis in basic conditions

This protocol was adapted from the literature.[76] 100mg (0.5mmol of BLG units, 1eq) of polymer **1** were solubilized in 41ml THF (2.5mg/ml) and cooled to 0°C. 5ml of an aqueous solution of NaOH (0.95mmol, 2eq) were added dropwise to the mixture, the medium became turbid. In a first attempt the reaction mixture was left to stir for 10 hours. THF was then evaporated, and the remaining aqueous reaction mixture was neutralized with a saturated aqueous solution of $NaHCO_3$. The crude mixture was dialyzed against Milli-Q® water using 3.5kDa dialysis membranes and finally lyophilized to obtain pure PGA as white fluffy powder, polymer **$1D_1$** (60mg, >99% yield). The powder was not readily soluble in water even after considerably basifying the medium. The polymer was analysed in D_2O+NaOD and in $CDCl_3$+15%TFA where it was not completely soluble in both. The NMR spectra (Figure **S1**), verify the presence of deprotected PGA but also show that some unprotected PBLG remained in the sample.

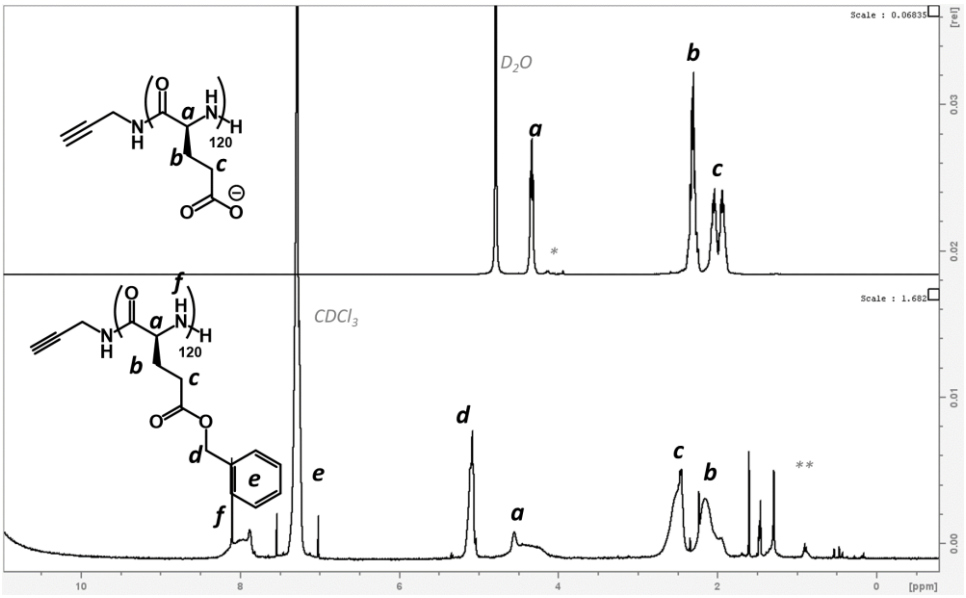


Figure S1: *^{1}H NMR spectra of polymer **1D$_1$** in D_2O+NaOD(top) and $CDCl_3$+15%TFA(bottom). Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be. In both cases the polymer was not completely soluble, and the spectra mostly reflected the chemical structure of the soluble species; PGA in D_2O and remaining unprotected PBLG in the organic solvents. In D_2O, no benzyl groups could be observed meaning that the soluble polymer chains are completely deprotected. *end chain groups **Impurities due to the use of use of non-deuterated TFA.*

As an optimization, a second reaction was done following the same procedure and left to react under vigorous stirring for 16 hours. The obtained polymer (polymer **1D$_2$**, white powder, 45mg,73% yield) was completely soluble (see section 2.1.1.2). The lower yield of this second deprotection consistent with the incomplete deprotection of the first one.

Hydrolysis in acidic conditions

MSA

This protocol was adapted from the literature.[77] 100mg (0.5mmol of BLG units, 1eq) of polymer **1** were solubilized in 1ml TFA (100mg/ml) and cooled to 0°C. 1ml of MSA (15mmol, 30eq) and 0.2ml anisole were added dropwise to the mixture under stirring. After 20 minutes of reaction, the reaction was left to stir for another 40 minutes at room temperature. Solution was precipitated twice in cold Et_2O and centrifuged. The obtained white powder was precipitated in 5ml MilliQ water and neutralized using a saturated aqueous solution of $NaHCO_3$. The crude mixture was dialyzed against Milli-Q® water using 3.5kDa dialysis membranes and finally lyophilized to obtain PGA as white fluffy powder, polymer **1D$_3$** (60mg, >99% yield). The powder was not readily

soluble in water even after considerably basifying the medium. The polymer was analyzed in D_2O+NaOD and in $CDCl_3$+15%TFA where it was not completely soluble in both. The NMR spectra in D_2O (**Figure S2**), verified the presence of deprotected PGA. When the sample was suspended in the deuterated organic solvents, no PBLG chains were found as in the case of polymer **1D$_1$** and only a compound that looks similar in structure to benzyl alcohol using ^{1}H NMR (**Figure S2**).

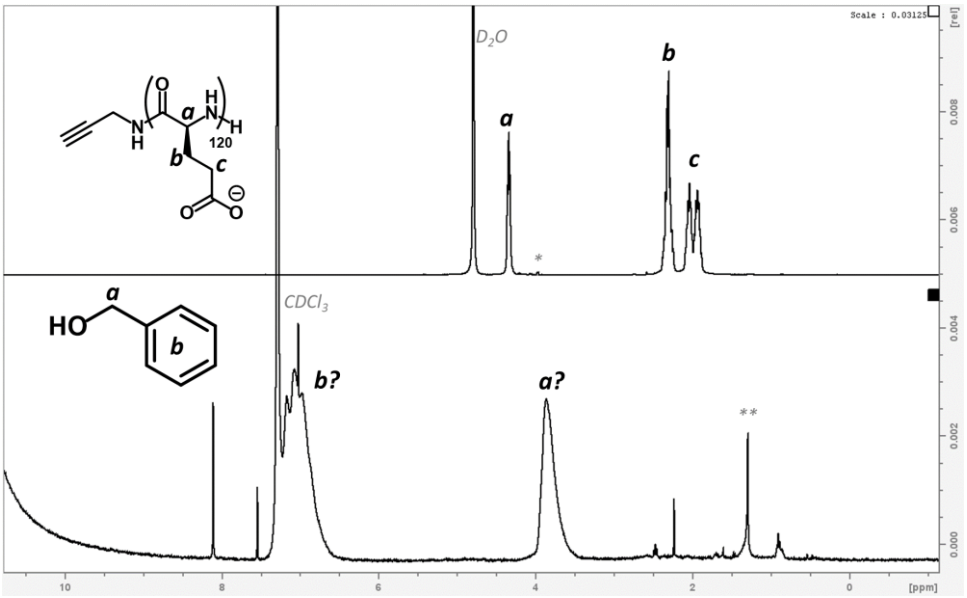


Figure S2: *^{1}H NMR spectra of polymer **1D$_3$** in D_2O+NaOD(top) and $CDCl_3$+15%TFA (bottom). Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be identified. In both cases the sample was not completely soluble, and the spectra mostly reflected the chemical structure of the soluble species. PGA peaks could be identified in D_2O(top). No remaining unprotected PBLG were investigated in the organic solvents(bottom) and peaks could be attributed to benzyl alcohol. In D_2O, no benzyl groups could be observed meaning that the soluble polymer chains are completely deprotected. *end chain groups **Impurities due to the use of use of non-deuterated TFA.*

To further investigate the insoluble part, the sample was suspended in DCM then the precipitate and the supernatant were separated. The first should consist of only PGA and the latter should be the suspected benzyl alcohol. This was confirmed by the ^{1}H NMR spectra in D_2O+NaOD and in $CDCl_3$+15%TFA of both parts respectively (**Figure S3**).

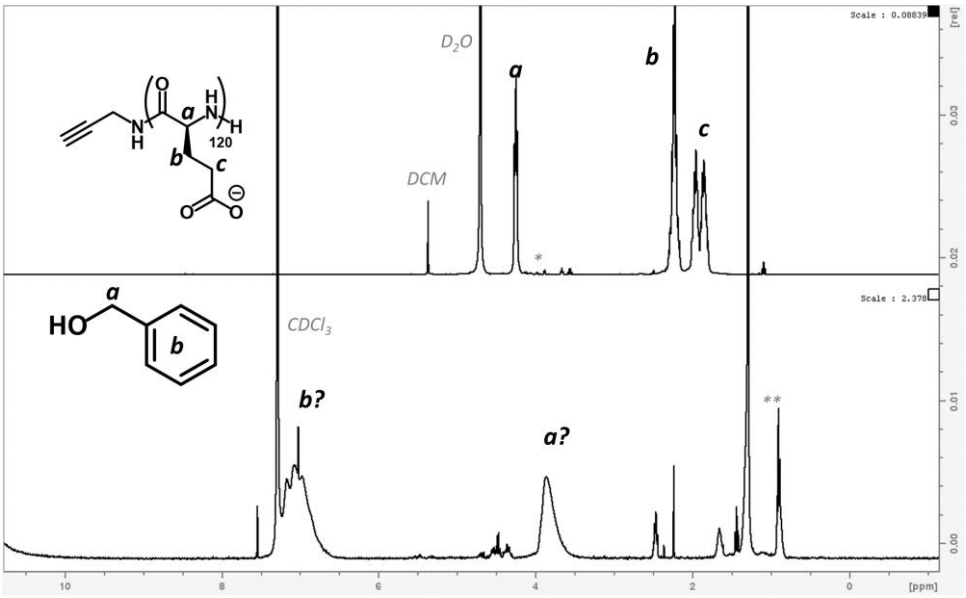


Figure S3: *^{1}H NMR spectra of polymer **1D$_3$** in D_2O+NaOD(top) and $CDCl_3$+15%TFA(bottom) after separation using DCM. Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be identified. PGA peaks could be identified in D_2O(top). In the organic solvents(bottom) peaks could be attributed to benzyl alcohol. In D_2O, no benzyl groups could be observed meaning that the soluble polymer chains are completely deprotected. *end chain groups **Impurities due to the use of use of non-deuterated TFA and the separation using DCM.*

The dialysis step should have eliminated any remaining benzyl alcohol from the reaction mixture and was the polymer was redialyzed using 5kDa dialysis membrane. Nonetheless, the resulting polymer was still not completely soluble in water and the NMR analysis (**Figure S4**) confirmed again the presence of the mystery compound suspected to be benzyl alcohol. It is also unlikely that the compound is pure benzyl alcohol as it should be soluble in water. This matter deserves further investigation, but it does not advance the study at hand.

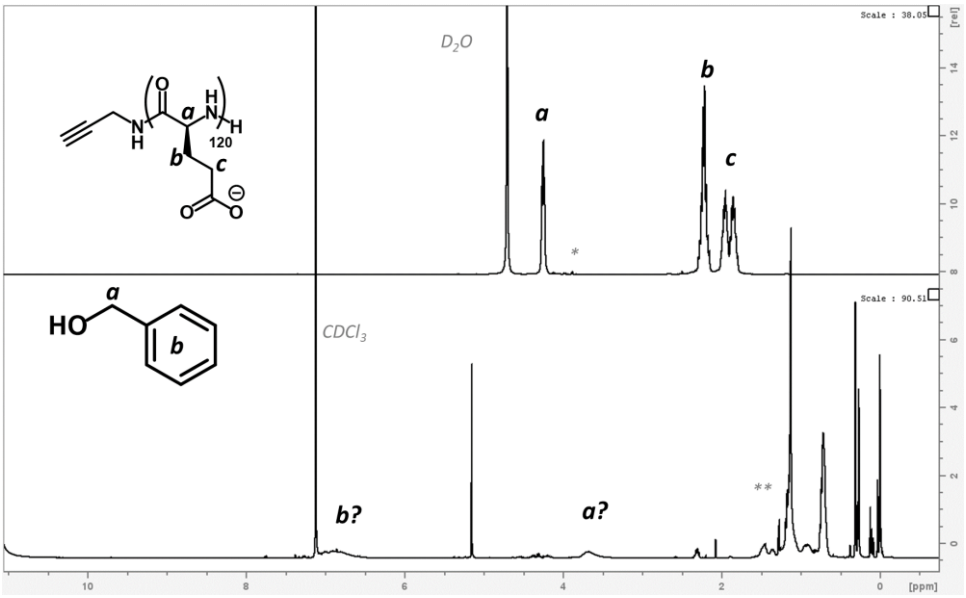


Figure S4: *^{1}H NMR spectra of polymer **1D$_3$** in D_2O+NaOD(top) and $CDCl_3$+15%TFA(bottom). Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be identified. In both cases the sample was not completely soluble, and the spectra mostly reflected the chemical structure of the soluble species. PGA peaks could be identified in D_2O(top). No remaining unprotected PBLG were investigated in the organic solvents(bottom) and peaks could be attributed to benzyl alcohol. In D_2O, no benzyl groups could be observed meaning that the soluble polymer chains are completely deprotected. *end chain groups **Impurities due to the use of use of non-deuterated TFA and successive workups.*

TFA

This protocol was adapted from the literature.[76] 100mg (0.5mmol of BLG units, 1eq) of polymer **1** were solubilized in 3ml TFA (33mg/ml) then HBr 48% aqueous solution (0.18 ml, 1.6 mmol, 3.5 eq) were added under vigorous stirring. The reaction mixture was stirred for 5 hours at room temperature. Solution was precipitated twice in cold Et_2O and centrifuged. The obtained white powder was precipitated in 5ml MilliQ water then neutralized with a saturated aqueous solution of $NaHCO_3$. The crude mixture was dialyzed against Milli-Q® water using 3.5kDa dialysis membranes and finally lyophilized to obtain pure PGA as white fluffy powder, polymer **1D$_4$** (70mg, >100% yield). The polymer was not completely soluble in aqueous medium and the high yield hints on a non-complete reaction. The ^{1}H NMR in D_2O+NaOD clearly shows residual aromatic peaks from the benzyl groups and the ^{1}H NMR in $CDCl_3$+15%TFA showed trace amounts of PBLG (**Figure S5**).

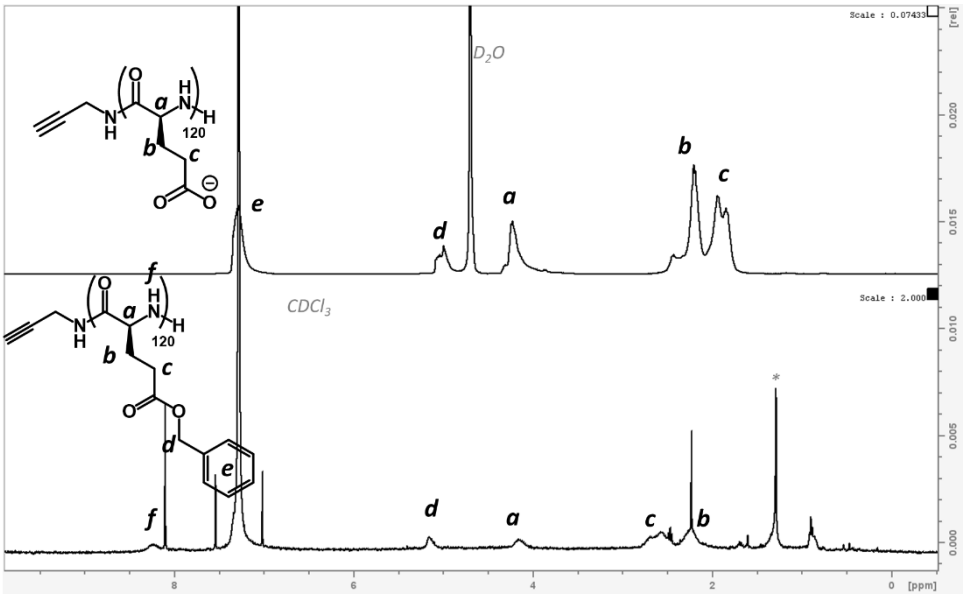


Figure S5: *^{1}H NMR spectra of polymer **1D$_4$** in D_2O+NaOD(top) and $CDCl_3$+15%TFA(bottom). Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be. In both cases the polymer was not completely soluble. PGA In D_2O, PGA peaks but also pendant benzyloxy groups could be identified. The poor quality of the spectrum means the polymers are not completely soluble. In organic solvents, only trace amounts of PBLG could be observed and *Impurities due to the use of use of non-deuterated TFA.*

Polymer **1D$_4$** was redeprotected using the same procedure but left for another 5 hours. The obtained polymer, polymer **1D$_4$'**, was readily soluble in water. The resulting ^{1}H NMR proved the presence of PGA with fewer remaining benzyl groups. For further optimization, a second experiment was undergone and left to react overnight (polymer **1D$_5$**). Negligeable amounts of benzyl groups were detected (**Figure 13**) and this procedure is to be used for subsequent deprotection experiments.

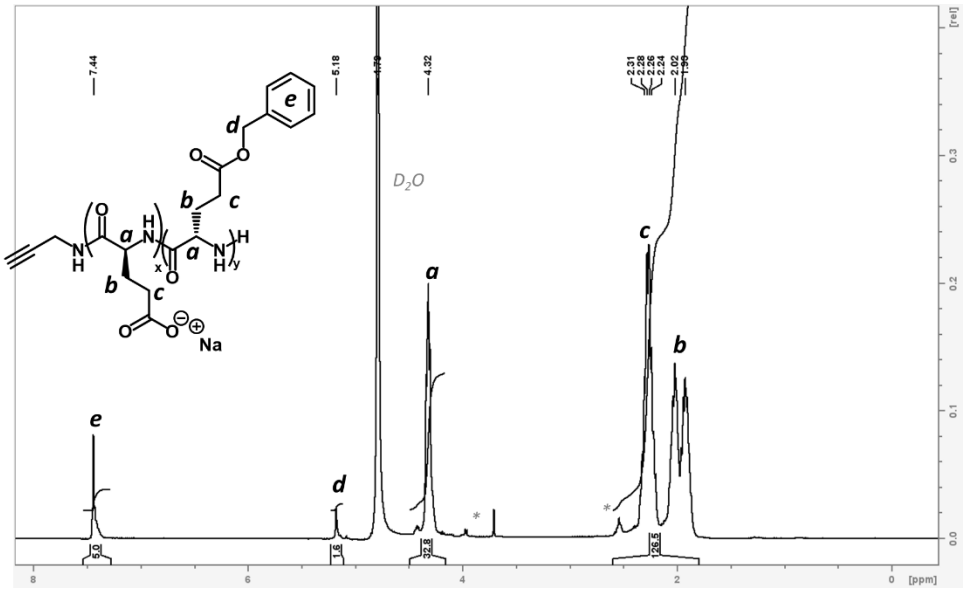


Figure S6: *^{1}H NMR spectra of polymer **1D$_4$'** in D_2O+NaOD(top). Peak attributions are indicated by the letters. As the targeted DP is high, the peaks from the initiator (propargylamine) could not easily be integrated. The spectrum clearly shows an incomplete deprotection. The integration and the ratio of benzyl groups can give an estimate about the deprotection extent. In this case, x=123 and y=4. *end chain groups.*

It is interesting to see the behavioural difference between the three proposed deprotection pathways. In the first two cases, completely deprotected PGA chains were obtained along with nearly intact PBLG. In the latter, deprotection seems to advance similarly in all chains and thus we can obtain partially deprotected PBLG chains. This might be useful in obtaining certain copolymers, but a detailed analysis of their composition will be needed for further use.

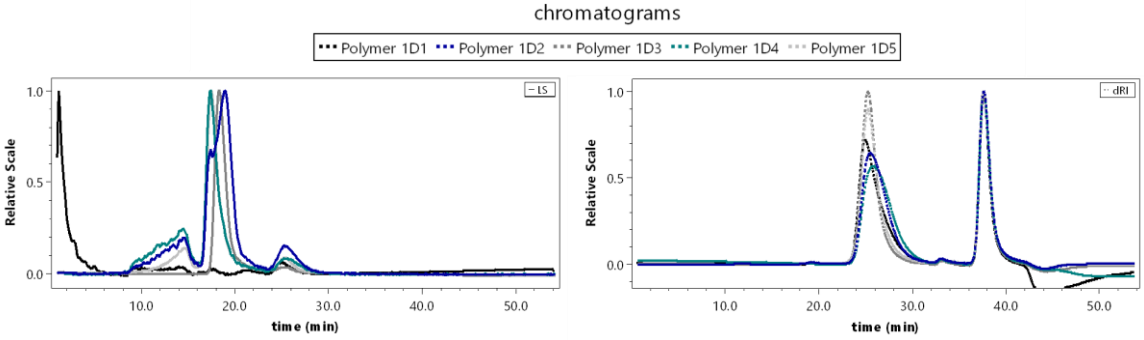


Figure S7: *SEC chromatograms of polymers **1D$_1$** to **1D$_5$** using MALS (on the left) and RI (on the right) detections. MALS detection shows aggregated state at lower elution time.*

Table S1: *Aqueous SEC results of all deprotected polymers from polymer **1**.*

Polymer	Mn (kg/mol)[a]	DP[b]	Đ[c]
Polymer **1D$_1$**	12	86	1.2
Polymer **1D$_2$**	36	260	1.1
Polymer **1D$_3$**	26	186	1.1
Polymer **1D$_{4'}$**	18	129	1.2
Polymer **1D$_5$**	12	86	1.3

a.as calculated using dn/dC=0.1675 from MALS detection, b.as calculated by dividing Mn value obtained from SEC by the molecular weight of the repeat unit (140g/mol), c. from SEC peaks.

4.1.8. Synthesis of Gly NCA

PCl_3

*Failed synthesis of Glycine-NCA, t*his protocol was based on the literature.[133] In a clean and previously dried Schlenk flask *N*-Boc-glycine (2 g, 11.4 mmol, 1 equiv.) was dried during 15 min in the Schlenck line, then 22 mL of anhydrous THF were added, followed of thionyl chloride (1 mL, 13.7 mmol, 1.2 equiv.). The reaction was stirred for 2h at 0°C and a faint white precipitate was observed. ^{1}H NMR showed formation of glycine NCA, but *N*-Boc-glycine peaks still remained meaning the reaction was not complete (**Figure S8**). The reaction took 3 days to complete and celite filtration did not remove all byproducts (data not shown).

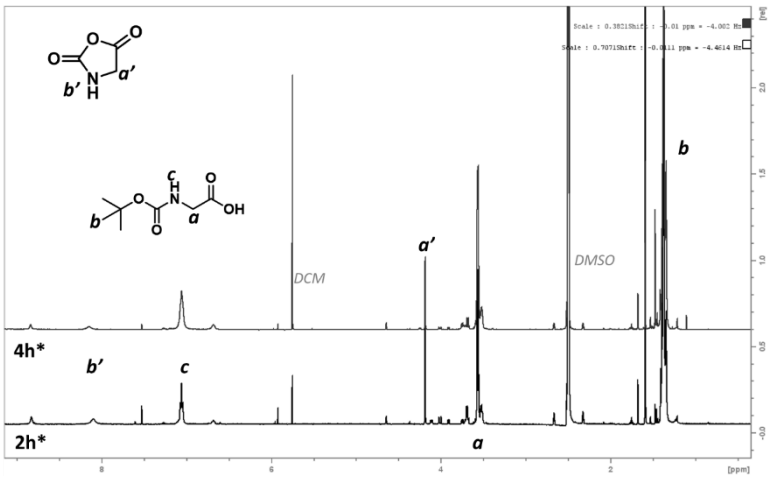


***Figure S8** : ^{1}H NMR of Gly NCA synthesis using PCl_3 at 2 hours and 4 hours. Peaks from the precursor and Gly NCA can be observed glycine NCA after purification. All spectra were recorded in DMSO-d_6. *reaction solvent was evaporated, and crude mixture was analysed as is. Unidentified peaks are unknown side products.*

T3P

Failed synthesis of Glycine-NCA, this protocol was adapted from the literature for glycine NCA.[151] In a clean and previously dried Schlenk flask *N*-Boc-glycine (2 g, 11.4 mmol, 1 equiv.) was dried during 15 min in the Schlenck line, then 29 mL of anhydrous ethyl acetate were added. Finally, T3P (8.15 mL, 13.7 mmol, 1.2 equiv.) and pyridine (1.1ml, 13.7 mmol, 1.2 equiv.) were added. The reaction was stirred for 24h at RT and celite filtration did not remove all byproducts (data not shown).

$SOCl_2$

Synthesis of Glycine-NCA. In a clean and previously dried Schlenk flask *N*-Boc-glycine (2 g, 11.4 mmol, 1 equiv.) was dried during 15 min in the Schlenck line, then 32 mL of anhydrous THF were added, followed of thionyl chloride (1 mL, 13.7 mmol, 1.2 equiv.). The reaction was stirred for 4 h at RT. The solvent was evaporated using vacuum Schlenck line, washed three times with 35 mL of anhydrous hexane, the suspension was filtered using a cannula and the powder was dried. Then 90 mL of anhydrous ethyl acetate were added, and the suspension was filtered over previously dried celite. The system was washed with 60 mL of anhydrous ethyl acetate and the solution concentrated. The solution was filtered again over previously dried celite and the system washed with 60 ml of anhydrous ethyl acetate. The solid was washed two times with 10 mL of dry diethyl ether before being dried. This last step proved efficient in removing any Boc-containing side products that remained when scaling up the process. The product was isolated as a white yellowish crystalline powder in 79% yield (700 mg). ^{1}H-NMR, DMSO-d_6 400 MHz δ: 4.20 (s, 2H, CH_2), 8.83 (s, 1H, NH). ^{13}C-NMR, DMSO-d_6 400 MHz δ: 46.26 (s, CH_2), 153.97 (s, NCOO), 169.40 (s, COO). ESI-MS: m/z[M-H]=100 $C_3H_2NO_3$.

Table S2: *Solubility of N-Boc glycine and glycine NCA in different solvent as perceived by eye. Not soluble: as low as 1mg/ml solutions were tested, slightly soluble: as high as 50mg/ml solutions were tested, soluble/highly soluble: as high as 100mg/ml solutions were tested.*

	cyclo-hexane	**Et_2O**	**DCM**	**THF**	**ACN**	**AcOEt**	**DMF**
N-Boc glycine	Not soluble	Highly soluble	Soluble	Highly soluble	Highly soluble	Highly soluble	Highly soluble
Gly NCA	Not soluble	Slightly soluble	Slightly soluble	Highly soluble	Highly soluble	Highly soluble	Highly soluble

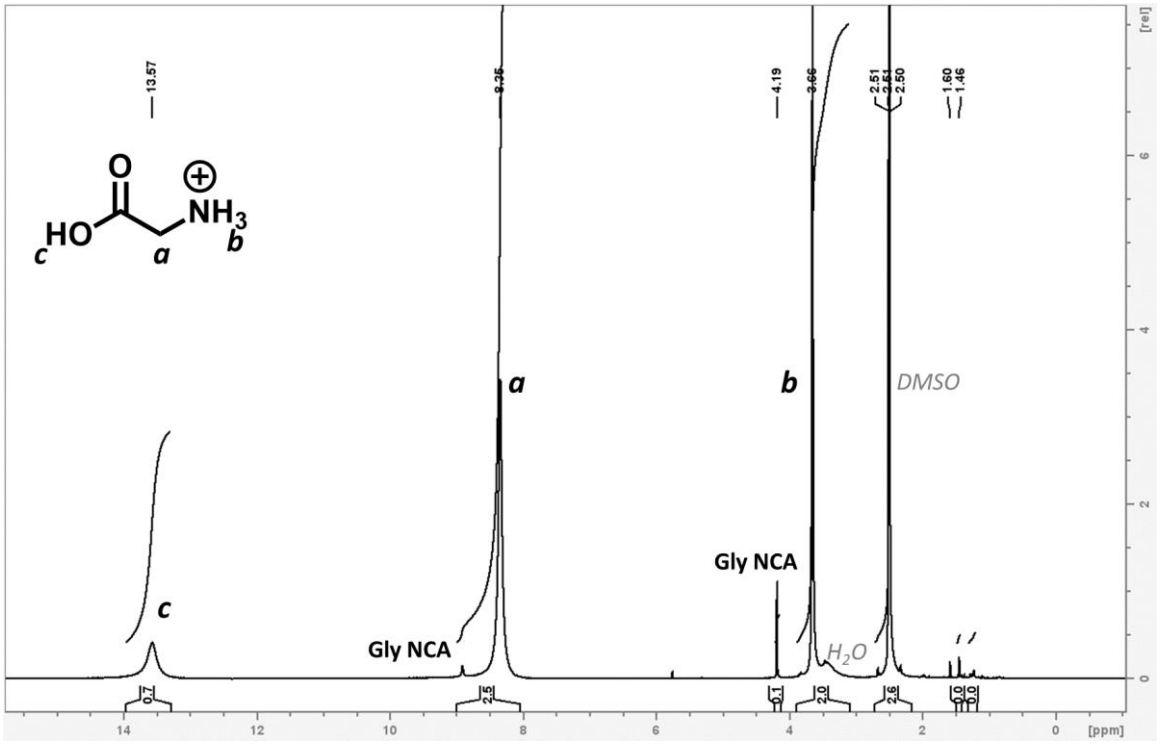


Figure S9: *^1H NMR in in DMSO-d_6 of precipitate in AcOEt after synthesis reaction of NCA Gly. Main product is glycine amino acid. Some traces of Gly NCA could be observed.*

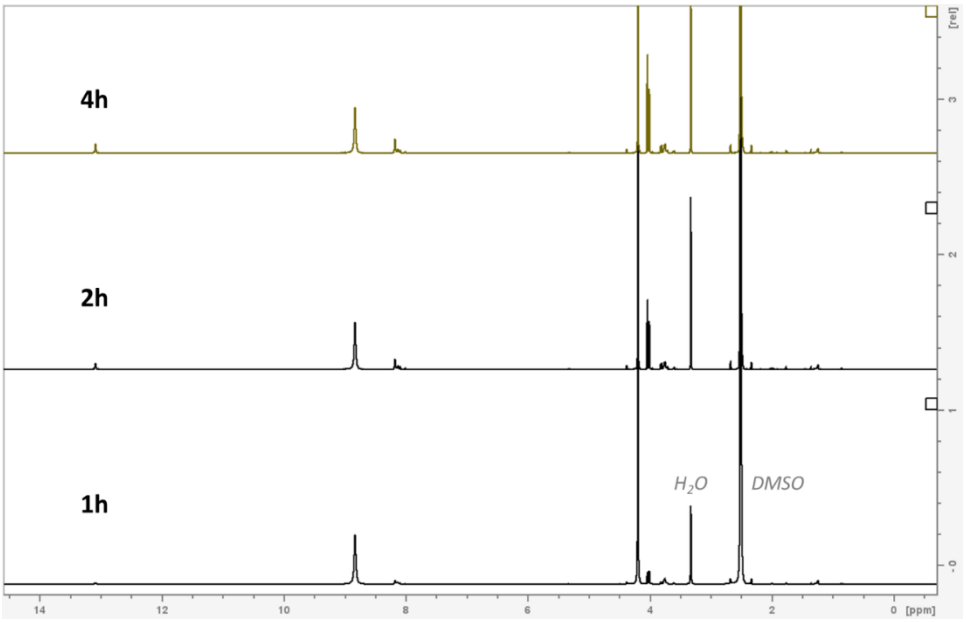


***Figure S10**: ^1H NMR in DMSO-d_6 of Gly NCA. Spectra were obtained at different times and precipitate appeared after 30 minutes in solution, suspected to be oligoglycine.*

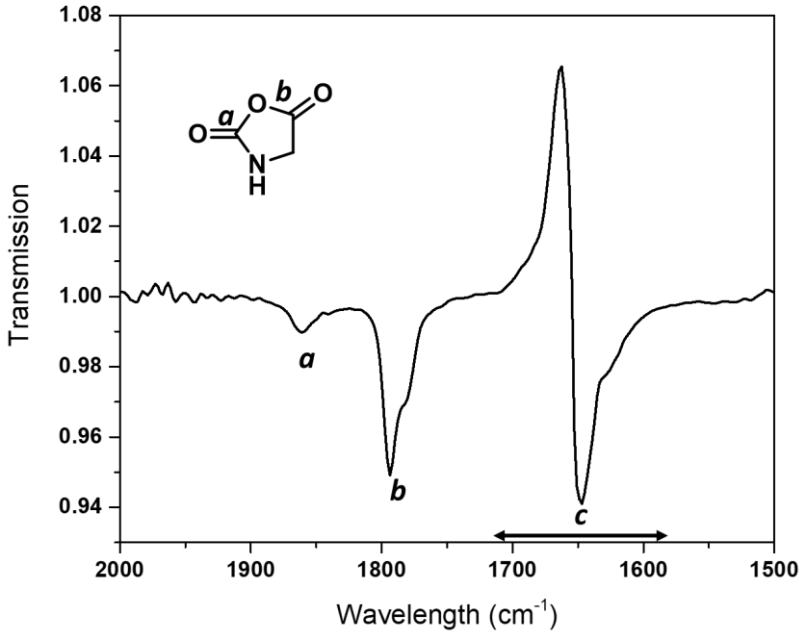


***Figure S11:** FTIR spectra of Gly NCA in dry DMF. The window represented from 2000cm^{-1} and 1500cm^{-1} shows the stretching bands of different carbonyls from Gly NCA: a. is the stretching band of the anhydride carbonyl from the NCA and b. that of the anhydride/amide one. c. wavelengths where the solvent (DMF) absorbs distorting the baseline.*

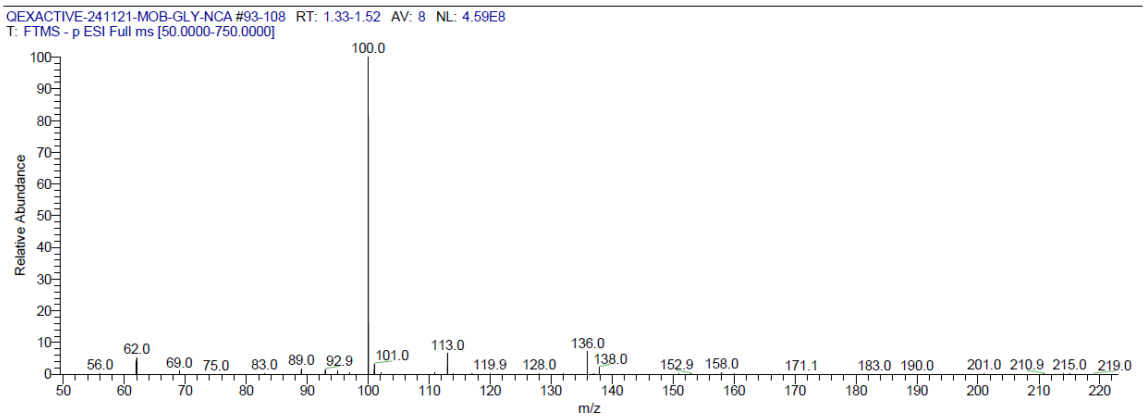


Figure S12: *ESI-MS of Gly NCA. Analysis and spectrum were provided by CESAMO analysis platform.*

4.1.9. ROP of Gly NCA

Glycine NCA monomer (38.2 mg, 0.378x10-3 mol) was weighted in a glovebox under pure argon, introduced in a flame-dried Schlenck vessel, and dissolved with 0.944 mL of anhydrous DMF. A solution of allylamine (1.3 M in DMF) was prepared and 100 µL (0.018x10-3 mol) were added with an argon purged syringe. The solution was stirred at room temperature under argon until completion. The polymer was then recovered as a white solid by precipitation in THF and dried under high vacuum. Yield: 90% (20 mg). ^{1}H-NMR TFA-d 400 MHz δ (ppm): 3.86-3.99 (m, 2H, CH2 allylamine), 3.80-4.42 (m, 32H, CH2), 5.0-5.20 (m, 2H, CH2 allylamine), 5.64-5.82 (m, 1H, CH allylamine). NH terminal missing in these experimental conditions.

4.1.10. Copolymerizations

Table S3: *Synthesized copolymers to show the effect of monomers' initial feed ratio on their properties. All synthesis were conducted in the reaction conditions of copolymerization system* ***1*** *(0.1M in DMF, 0°C and initiated with hexylamine and M/I= 50). Polymer* ***2*** *is included for comparison.*

	$f_{B,0}$[a]	Structure (theoretical)	Yield (%)	BLG% by ^{1}H NMR[b]	Overall DP by ^{1}H NMR[c]
Polymer **2**	100	HA-PBLG$_{50}$	87	0	54
Copolymer **2**	95	HA-PBLG$_{47}$-co-PGly$_{3}$	60	94	35
Copolymer **3**	90	HA-PBLG$_{45}$-co-PGly$_{5}$	70	85	39

	$f_{B,0}$[a]	Structure (theoretical)	Yield (%)	BLG% by ^{1}H NMR[b]	Overall DP by ^{1}H NMR[c]
Copolymer **4**	80	HA-PBLG$_{40}$-co-PGly$_{10}$	50	80	40
Copolymer **5**	70	HA-PBLG$_{35}$-co-PGly$_{15}$	94	65	61
Copolymer **6**	60	HA-PBLG$_{30}$-co-PGly$_{20}$	79	60	50
Copolymer **1**	50	HA-PBLG$_{25}$-co-PGly$_{25}$	78	50	50

[a]initial monomer feed and expected BLG% in final copolymer; HA: hexylamine; [b]as determined by the integration =, the glycine and the BLG peaks in ^{1}H NMR in TFA-d; [c]as determined by the ratio between the HA integration of the HA peaks and the backbone peaks in ^{1}H NMR in TFA-d.

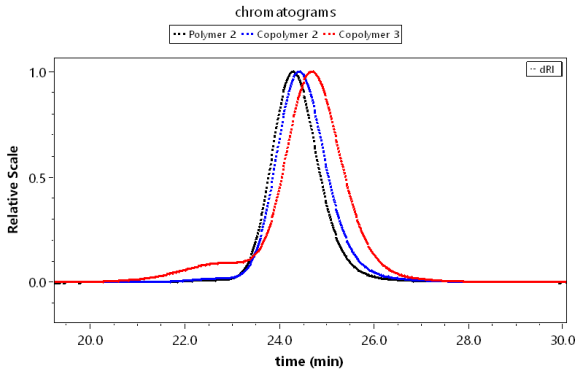


Figure S13 : *SEC chromatograms in DMF of copolymers **2-3** with RI detection. Polymer **2**'s chromatogram is shown for comparison.*

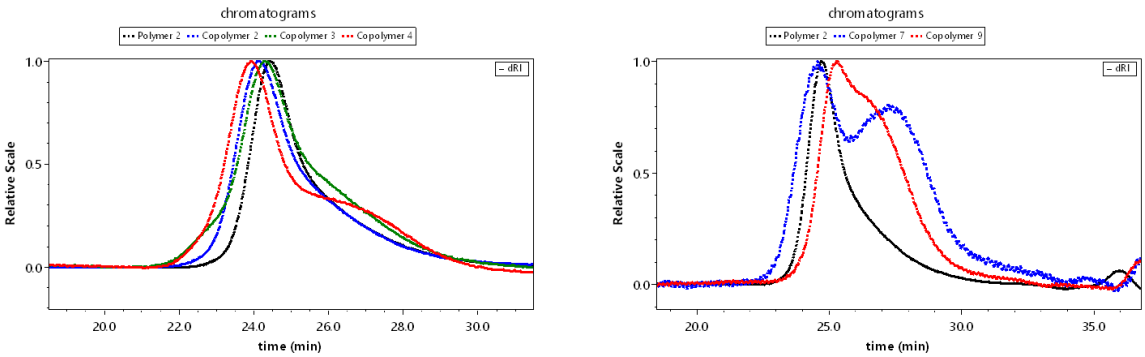


Figure S14 : *SEC chromatograms in DMSO of copolymers **2-4** (on the left) and copolymers **7-8** (on the right) with RI detection. Polymer **2**'s chromatogram is shown for comparison.*

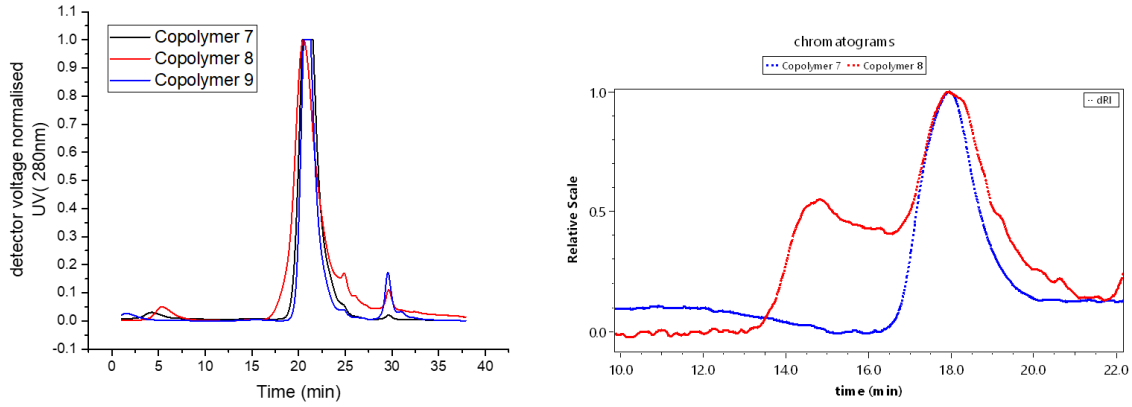


Figure S15: *On the left, SEC chromatograms in HFIP (PL column) of copolymers* ***7-9*** *with UV detection (280nm). On the right, SEC chromatograms in HFIP (PSS column) of copolymers* ***7-8*** *with RI detection.*

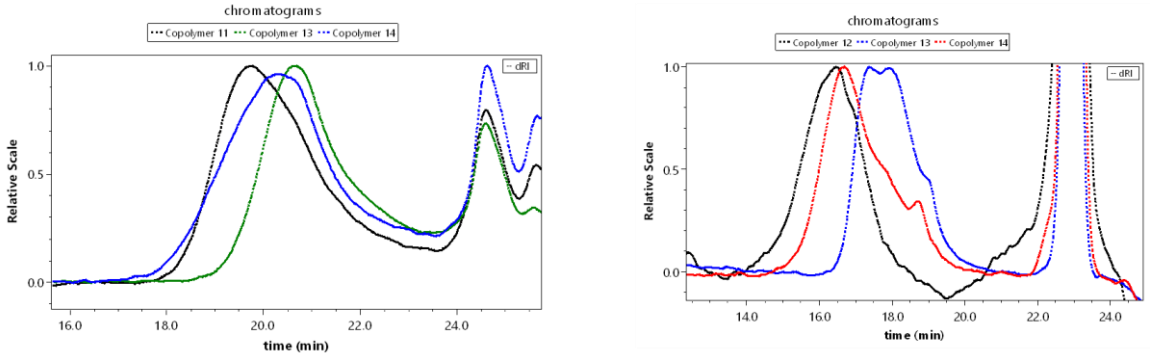


Figure S16 *: On the left, SEC chromatograms in HFIP (PL column) of copolymers* ***11****,* ***13*** *and* ***14*** *with RI detection. On the right, SEC chromatograms in HFIP (PSS column) of copolymers* ***12-14*** *with RI detection.*

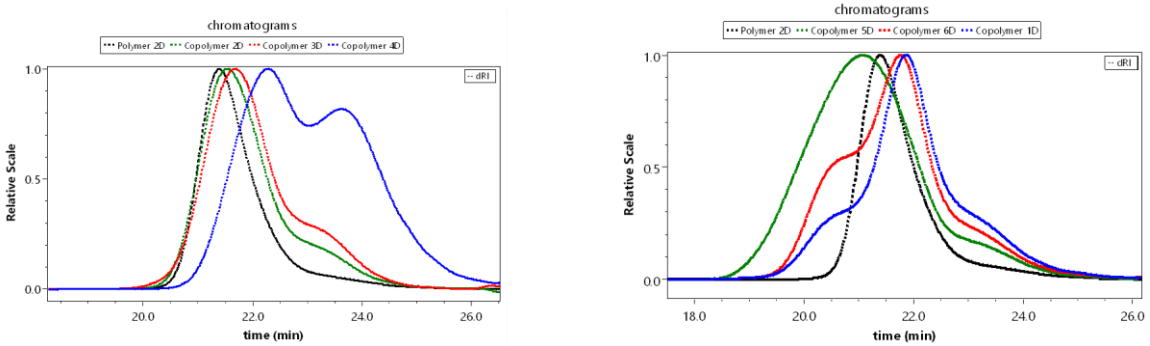


Figure S17 *: SEC chromatograms in aqueous buffer of copolymers* ***1D-6D*** *with RI detection. Polymer* ***2D****'s chromatogram is shown for comparison.*

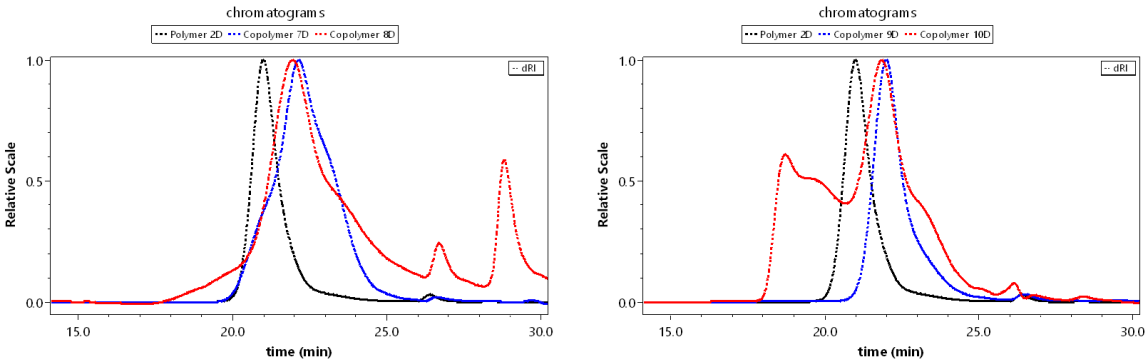


Figure S18 *: SEC chromatograms of copolymers* ***7D-10D*** *with RI detection. Polymer* ***2D*** *is represented for comparison. All copolymers have peaks at high elution time consistent with a Mn lower than the homopolymer.*

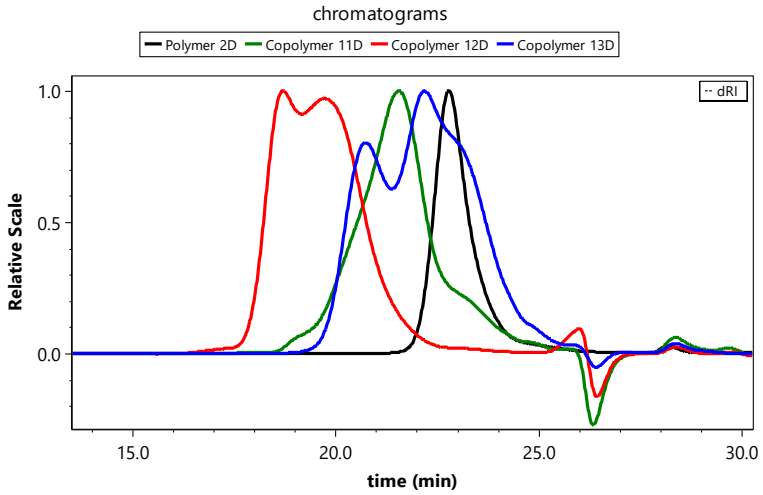


Figure S19 *: SEC chromatograms of copolymers* ***11D-13D*** *with RI detection. Polymer 2D is represented for comparison. All copolymers have peaks consistent with a Mn higher than the homopolymer and some aggregation.*

Table S4 *: SEC analysis of all synthesized copolymers (**1-14**) and their deprotected form (**1D-14D**) compared to polymer **2(D)**.*

	SEC DMF/DMSO Mn (g/mol) / Đ[a]	SEC HFIP Mn (g/mol) / Đ[a]	SEC aqueous conditions Mn (g/mol) / Đ[a]
Polymer **2(D)**	11.1[b] / 1.03 2.3[c] / 1.11	*d*	7.9[j] / 1.03 5.4[k] / 1.08
Copolymer **2(D)**	10.1[b] / 1.03 2.6[c] / 1.15	*d*	5.3[j] / 1.01
Copolymer **3(D)**	9.2[b] / 1.11 2.5[c] / 1.37	*d*	4.6[j] / 1.04
Copolymer **4(D)**	2.5[c] / 1.37	*d*	3.4[k] / 1.20[l]
Copolymer **5(D)**	*d*	-	9.8[j] / 1.16

	SEC DMF/DMSO Mn (g/mol) / Đ[a]	SEC HFIP Mn (g/mol) / Đ[a]	SEC aqueous conditions Mn (g/mol) / Đ[a]
Copolymer **6(D)**	*d*	-	6.8[j] / 1.15
Copolymer **1(D)**	*d*	-	4.8[j] / 1.20
Copolymer **7(D)**	4.7[c] / 1.22[e]	14.6[f] / 1.35 3.7[g] / 1.46	4.5[j] / 1.12
Copolymer **8(D)**	*d*	14.2[f] / 1.45[h] 3.1[g] / 1.91	4.8[j] / 1.28
Copolymer **9(D)**	2.2[c] / 1.21	4.4[g] / 1.20	4.1[k] / 1.02[l]
Copolymer **10(D)**	*d*	17.5[f] / 1.23[h]	5.0[k] / 1.15[l]
Copolymer **11(D)**	*d*	7.2[g] / 1.49	8.6[j] / 1.20
Copolymer **12(D)**	*d*	60.5[f] / 1.34	20.1[j] / 1.24
Copolymer **13(D)**	*d*	17.4[f] / 1.3[i] 4[g] / 1.11	4.0[j] / 1.11
Copolymer **14(D)**	*d*	36.8[f] / 1.41[h] 5.4[g] / 1.76	*d*

[a] as determined by SEC peaks using the adequate detection; [b] as determined by the Mn value from dn/dC of homoPBLG (0.1197) through MALS detection in SEC DMF; [c] as determined by pullulan calibration in SEC DMSO using RI detection; [d]sample was insoluble in these conditions; [e] peak showed shouldering; [f] as determined by PMMA calibration in SEC HFIP and PSS PFG columns using RI detection; [g] as determined by PMMA calibration in SEC HFIP and PL HFIP gel columns using RI detection; [h] sample showed high concentration of aggregates; [i] sample showed a double peak; [j] as determined by the Mn value from dn/dC of homoPGA (0.1489) through MALS detection in aqueous SEC; [k] as determined by PEG calibration in aqueous SEC using RI detection; [l] aggregates were detected by MALS not allowing for the use of the detector.

Table S5: *Explored copolymerization systems to calculate reactivity ratios. Their main features have been listed and other reaction conditions were: 0.1M initial monomer concentration and M/I=50.*

	Solvent	Temp.	Initiator	Catalyst	$r_{Gly}xr_{BLG}$	50/50 copolymer[a]	ΔMolar ellipticity[b]
Copo. system **1**	DMF	0°C	HA[c]	-	0.82	Yes	-1.70
Copo. system **2**	DMF	25°C	HA[c]	-	0.48	Yes	-3.84
Copo. system **3**	THF	25°C	HA[c]	-	0.59	Yes	-21.21
Copo. system **4**	DMSO	25°C	HA[c]	-	2.36	Yes	1.88

	Solvent	Temp.	Initiator	Catalyst	$r_{Gly}xr_{BLG}$	50/50 copolymer[a]	ΔMolar ellipticity[b]
Copo. system **5**	DMF	50°C	HA[c]	-	1.14	No	-
Copo. system **6**	THF	0°C	HA[c]	-	0.88	No	-
Copo. system **7**	DMF	25°C	HA[c]	TMG[d]	2.30	Yes	-5.59
Copo. system **8**	THF	25°C	HA[c]	TMG[d]	6.39	Yes	-46.00

[a]If a synthesis comprising a 50/50 Gly/BLG ratio was done, it was subsequently deprotected and analyzed with CD for secondary structuring; [b]Molar ellipticity$_{(209nm)}$-Molar ellipticity$_{(198nm)}$ [(deg.cm²/dmol).10^{-3}] indicating the helical nature; [c]hexylamine; [d]tetramethylguanidine.

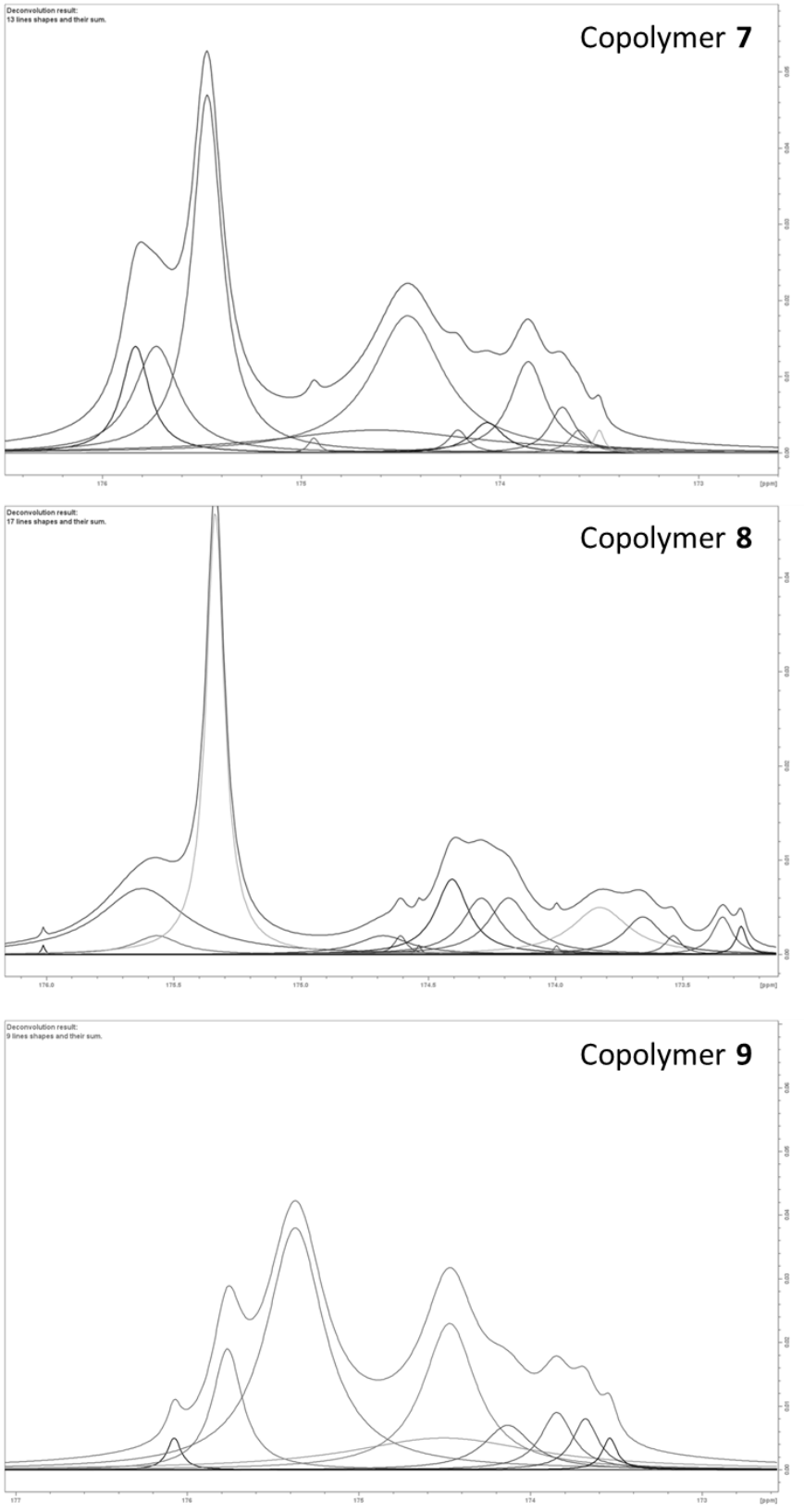


***Figure S20** : deconvolution and fitting of ^{13}C NMR of polymer **7-9.** The fitted peaks are shown from the deconvolution of the peaks' shape shown in **Figure 36**.*

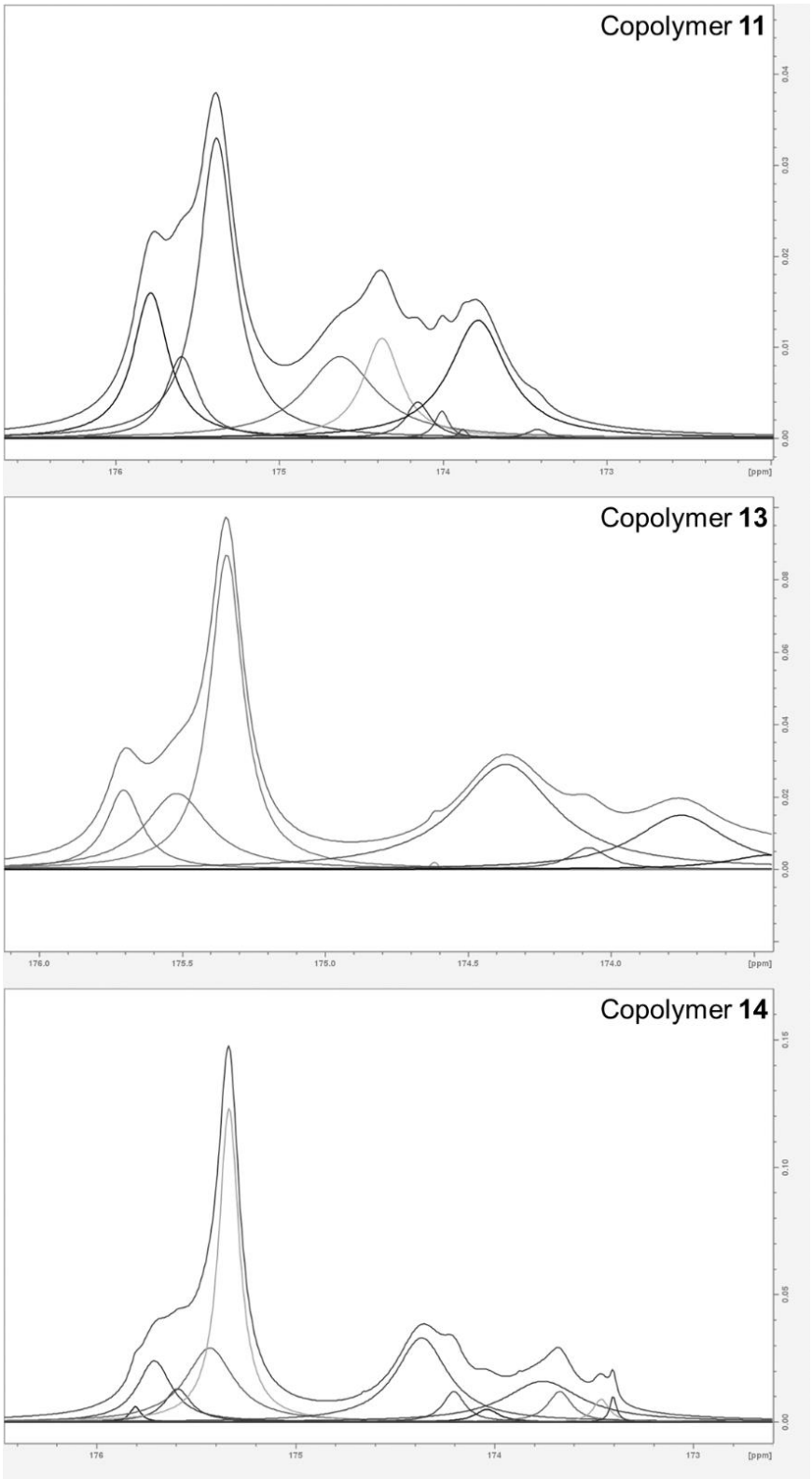


***Figure S21** : deconvolution and fitting of ^{13}C NMR of polymers **11,13 and 14**. The fitted peaks are shown from the deconvolution of the peaks' shape and the results are illustrated in **Table 4**.*

5. References

(1) Lutz, J. F.; Schmidt, B. V. K. J.; Pfeifer, S. *Macromol. Rapid Commun.* **2011**, *32* (2), 127–135.
(2) Lutz, J.-F. *Macromol. Rapid Commun.* **2017**, *38* (24), 1700582.
(3) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. *Science (80-.).* **2013**, *341* (6146), 1238149.
(4) Merrifield, R. B. *Am. Chem. Soc.* **1963**, *85* (14), 2149–2154.
(5) Kallmyer, N. E.; Rider, N. E.; Reuel, N. F. *PLoS One* **2020**, *15* (8), e0237473.
(6) Clapperton, A. M.; Babi, J.; Tran, H. *ACS Polym. Au* **2022**, *2* (6), 417–429.
(7) Gutman, I.; Gutman, R.; Sidney, J.; Chihab, L.; Mishto, M.; Liepe, J.; Chiem, A.; Greenbaum, J.; Yan, Z.; Sette, A.; Koşaloğlu-Yalçın, Z.; Peters, B. *ACS Omega* **2022**, *7* (27), 23771–23781.
(8) Al-Warhi, T. I.; Al-Hazimi, H. M. A.; El-Faham, A. *J. Saudi Chem. Soc.* **2012**, *16* (2), 97–116.
(9) Church, G. M.; Gao, Y.; Kosuri, S. *Science (80-.).* **2012**, *337* (6102), 1628–1628.
(10) Trinh, T. T.; Oswald, L.; Chan-Seng, D.; Lutz, J.-F. *Macromol. Rapid Commun.* **2014**, *35* (2), 141–145.
(11) Martens, S.; Landuyt, A.; Espeel, P.; Devreese, B.; Dawyndt, P.; Du Prez, F. *Nat. Commun.* **2018**, *9* (1), 4451.
(12) Roy, R. K.; Meszynska, A.; Laure, C.; Charles, L.; Verchin, C.; Lutz, J.-F. *Nat. Commun.* **2015**, *6* (1), 7237.
(13) Colquhoun, H.; Lutz, J. F. *Nat. Chem.* **2014**, *6* (6), 455–456.
(14) Zhang, J.; Matta, M. E.; Hillmyer, M. A. *ACS Macro Lett.* **2012**, *1* (12), 1383–1387.
(15) Satoh, K.; Ozawa, S.; Mizutani, M.; Nagai, K.; Kamigaito, M. *Nat. Commun.* **2010**, *1* (1), 0–5.
(16) Han, S.; Wu, J. *Biomacromolecules* **2022**, *23* (5), 1892–1919.
(17) Odian, G. ; "*Principles of Polymerization: Chap. 2 Step Polymerization*"; 4th ed.; Wiley, Ed.; **2004**.
(18) Alfrey, T.; Goldfinger, G. *J. Chem. Phys.* **1946**, *14* (2), 115–116.
(19) Mayo, F. R.; Lewis, F. M. *J. Am. Chem. Soc.* **1944**, *66* (9), 1594–1601.
(20) Lutz, J. F. *Polym. Chem.* **2010**, *1* (1), 55–62.
(21) Lutz, J. F. *ACS Macro Lett.* **2014**, *3* (10), 1020–1023.
(22) Yamada, K.; Nakano, T.; Okamoto, Y. *J. Polym. Sci. Part A Polym. Chem.* **2000**, *38* (1), 220–228.
(23) Fernández-García, M.; Fernández-Sanz, M.; Madruga, E. L.; Cuervo-Rodriguez, R.; Hernández-Gordo, V.; Fernández-Monreal, M. C. *J. Polym. Sci. Part A Polym. Chem.* **2000**, *38* (1), 60–67.
(24) Barson, C. A.; Turner, M. J. *Eur. Polym. J.* **1973**, *9* (8), 789–793.
(25) Lewis, F. M.; Walling, C.; Cummings, W.; Briggs, E. R.; Mayo, F. R. *J. Am. Chem. Soc.* **1948**, *70* (12), 4277.
(26) Hou, C.; Liu, J.; Ji, C.; Ying, L.; Sun, H.; Wang, C. *J. Appl. Polym. Sci.* **2006**, *102* (4), 4045–4048.
(27) Gody, G.; Zetterlund, P. B.; Perrier, S.; Harrisson, S. *Nat. Commun.* **2016**, *7*, 1–8.
(28) Zamfir, M.; Lutz, J.-F. *Nat. Commun.* **2012**, *3* (1), 1138.
(29) Song, Z.; Tan, Z.; Cheng, J. *Macromolecules* **2019**, *52* (22), 8521–8539.
(30) Rasines Mazo, A.; Allison-Logan, S.; Karimi, F.; Chan, N. J. A.; Qiu, W.; Duan, W.; O'Brien-

Simpson, N. M.; Qiao, G. G. *Chem. Soc. Rev.* **2020**, *49* (14), 4737–4834.
(31) Habraken, G. J. M.; Peeters, M.; Dietz, C. H. J. T.; Koning, C. E.; Heise, A. *Polym. Chem.* **2010**, *1* (4), 514–524.
(32) Kataoka, K.; Harada, A.; Nagasaki, Y. *Adv. Drug Deliv. Rev.* **2012**, *64* (SUPPL.), 37–48.
(33) Gaspard, J.; Silas, J. A.; Shantz, D. F.; Jan, J. **2010**, *22* (3), 178–185.
(34) Deming, T. J. *Nature* **1997**, *390* (6658), 386–389.
(35) Papadopoulos, P.; Floudas, G.; Schnell, I.; Aliferis, T.; Iatrou, H.; Hadjichristidis, N. *Biomacromolecules* **2005**, *6* (4), 2352–2361.
(36) Aliferis, T.; Iatrou, H.; Hadjichristidis, N. *Biomacromolecules* **2004**, *5* (5), 1653–1656.
(37) Fan, J.; Zou, J.; He, X.; Zhang, F.; Zhang, S.; Raymond, J. E.; Wooley, K. L. *Chem. Sci.* **2014**, *5* (1), 141–150.
(38) Hadjichristidis, N.; Iatrou, H.; Pitsikalis, M.; Sakellariou, G. *Chem. Rev.* **2009**, *109* (11), 5528–5578.
(39) Atreyi, M.; Rao, M. V. R.; Kumar, S. *Biopolymers* **1983**, *22* (2), 747–753.
(40) Oya, M.; Takahashi, T. *J. Polym. Sci. A1.* **1982**, *20* (2), 529–539.
(41) Saudek, V.; Stejskal, J.; Schmidt, P.; Zimmermann, K.; Škarda, V.; Kratochvíl, P.; Drobník, J. *Biopolymers* **1987**, *26* (5), 705–725.
(42) Hill, D. J. T.; Cardinaux, F.; Scheraga, H. A. *Biopolymers* **1977**, *16* (11), 2469–2480.
(43) Sederel.W, Deshmane.S, Hayashi.T, A. J. M. **1978**, *17*, 2835–2849.
(44) Bonduelle, C. *Polym. Chem.* **2018**, *9* (13), 1517–1529.
(45) Wamsley, A.; Jasti, B.; Phiasivongsa, P.; Li, X. *J. Polym. Sci. Part A Polym. Chem.* **2004**, *42* (2), 317–325.
(46) Zelzer, M.; Heise, A. *Polym. Chem.* **2013**, *4* (13), 3896–3904.
(47) Zelzer, M.; Heise, A. *J. Polym. Sci. Part A Polym. Chem.* **2014**, *52* (9), 1228–1236.
(48) Zou, J.; Zhou, M.; Ji, Z.; Xiao, X.; Wu, Y.; Cui, R.; deng, shuai; Liu, R. *Polym. Chem.* **2021**, 388–394.
(49) Iyer, M. S. US20220025115A1, **2022**.
(50) Campos-García, V. R.; Herrera-Fernández, D.; Espinosa-De La Garza, C. E.; González, G.; Vallejo-Castillo, L.; Avila, S.; Muñoz-García, L.; Medina-Rivero, E.; Pérez, N. O.; Gracia-Mora, I.; Pérez-Tapia, S. M.; Salazar-Ceballos, R.; Pavón, L.; Flores-Ortiz, L. F. *Sci. Rep.* **2017**, *7* (1), 1–12.
(51) Mbizana, S.; Hlalele, L.; Pfukwa, R.; Du Toit, A.; Lumkwana, D.; Loos, B.; Klumperman, B. *Biomacromolecules* **2018**, *19* (7), 3058–3066.
(52) Van Herk, A. M. *J. Chem. Educ.* **1995**, *72* (2), 138–140.
(53) Schilling, G.; Kricheldorf, H. *Die Makromol. Chemie* **1977**, *178* (3), 885–892.
(54) Kricheldorf, H.; Schilling, G. *Die Makromol. Chemie* **1978**, *179* (5), 1175–1191.
(55) Kricheldorf, H. R.; Hull, W. E. *J Polym Sci Polym Chem Ed* **1978**, *16* (3), 583–595.
(56) Kricheldorf, H. R.; Hull, W. E. *Macromolecules* **1980**, *13* (1), 87–95.
(57) Frisch, H. L. **1987**, 1747–1752.
(58) Goury, V.; Jhurry, D.; Bhaw-Luximon, A.; Novak, B. M.; Belleney, J. *Biomacromolecules* **2005**, *6* (4), 1987–1991.
(59) Wamsley, A.; Phiasivongsa, P.; Jasti, B.; Li, X. *J. Polym. Sci. Part A Polym. Chem.* **2006**, *44* (14), 4328–4337.
(60) Farhangi, S.; Duhamel, J. *Macromolecules* **2016**, *49* (17), 6149–6162.
(61) Casier, R.; Duhamel, J. *Macromolecules* **2020**, *53* (22), 9823–9835.
(62) Casier, R.; Duhamel, J. *Macromolecules* **2021**, *54* (2), 919–929.
(63) Casier, R.; Duhamel, J. *Macromolecules* **2020**, *53* (22), 9811–9822.

(64) Casier, R.; Duhamel, J. *Macromolecules* **2018**, *51* (9), 3450–3457.
(65) Casier, R.; Duhamel, J. *J. Phys. Chem. B* **2023**, *127* (6), 1325–1337.
(66) Casier, R.; Duhamel, J. *Macromolecules* **2021**, *54* (18), 8904–8912.
(67) Bhaw-Luximon, A.; Jhurry, D.; Belleney, J.; Goury, V. *Macromolecules* **2003**, *36* (4), 977–982.
(68) Iizuka, Y.; Uchida, C.; Wakamatsu, K.; Oya, M. *Bulletin of the Chemical Society of Japan*. ,; "*Synthesis and Properties of High Molecular Weight Polypeptides Containing tryphtophane*"; , ,**1991**, pp 1336–1341.
(69) Kricheldorf, H. R.; Müller, D.; Hull, W. E. *Int. J. Biol. Macromol.* **1986**, *8* (1), 20–26.
(70) Hull, W. E.; Muller, D. **1985**, 2135–2140.
(71) Zhou, M.; Zou, J.; Liu, L.; Xiao, X.; Deng, S.; Wu, Y.; Xie, J.; Cong, Z.; Ji, Z.; Liu, R. *iScience* **2021**, *24* (10), 103124.
(72) Salas-Ambrosio, P.; Tronnet, A.; Verhaeghe, P.; Bonduelle, C. *Biomacromolecules* **2021**, *22* (1), 57–75.
(73) Javadi, Y.; Itzhaki, L. S. *Curr. Opin. Struct. Biol.* **2013**, *23* (4), 622–631.
(74) Wendisch, V. F. *Metab. Eng.* **2020**, *58* (March 2019), 17–34.
(75) Johnson, L. C.; Akinmola, A. T.; Scholz, C. *Biocatal. Agric. Biotechnol.* **2022**, *40* (September 2021), 102292.
(76) Conejos-Sánchez, I.; Duro-Castano, A.; Birke, A.; Barz, M.; Vicent, M. J. *Polym. Chem.* **2013**, *4* (11), 3182–3186.
(77) Aujard-Catot, J.; Nguyen, M.; Bijani, C.; Pratviel, G.; Bonduelle, C. *Polym. Chem.* **2018**, *9* (30), 4100–4107.
(78) Song, Z.; Fu, H.; Baumgartner, R.; Zhu, L.; Shih, K. C.; Xia, Y.; Zheng, X.; Yin, L.; Chipot, C.; Lin, Y.; Cheng, J. *Nat. Commun.* **2019**, *10* (1), 1–7.
(79) Niehoff, A.; Mantion, A.; McAloney, R.; Huber, A.; Falkenhagen, J.; Goh, C. M.; Thünemann, A. F.; Winnik, M. A.; Menzel, H. *Colloid Polym. Sci.* **2013**, *291* (6), 1353–1363.
(80) Baumgartner, R.; Fu, H.; Song, Z.; Lin, Y.; Cheng, J. *Nat. Chem.* **2017**, *9* (7), 614–622.
(81) Guo, J.; Huang, Y.; Jing, X.; Chen, X. *Polymer (Guildf).* **2009**, *50* (13), 2847–2855.
(82) Baumgartner, R.; Kuai, D.; Cheng, J. *Biomater. Sci.* **2017**, *5* (9), 1836–1844.
(83) Tian, H. Y.; Deng, C.; Lin, H.; Sun, J.; Deng, M.; Chen, X.; Jing, X. *Biomaterials* **2005**, *26* (20), 4209–4217.
(84) Sugimoto, H.; Nakanishi, E.; Hanai, T.; Yasumura, T.; Inomata, K. *Polym. Int.* **2004**, *53* (7), 972–983.
(85) Van Lysebetten, D.; Malfanti, A.; Deswarte, K.; Koynov, K.; Golba, B.; Ye, T.; Zhong, Z.; Kasmi, S.; Lamoot, A.; Chen, Y.; Van Herck, S.; Lambrecht, B. N.; Sanders, N. N.; Lienenklaus, S.; David, S. A.; Vicent, M. J.; De Koker, S.; De Geest, B. G. *ACS Appl. Mater. Interfaces* **2021**, *13* (5), 6011–6022.
(86) Alsehli, M.; Gauthier, M. *Materials (Basel).* **2023**, *16* (6), 2461.
(87) Mahi, B.; Gauthier, M.; Hadjichristidis, N. *Biomacromolecules* **2022**, *23* (6), 2441–2458.
(88) Tinajero-Díaz, E.; Kimmins, S. D.; García-Carvajal, Z. Y.; Martínez de Ilarduya, A. *React. Funct. Polym.* **2021**, *168* (September), 105040.
(89) Zhang, Y.; Song, W.; Lu, Y.; Xu, Y.; Wang, C.; Yu, D. G.; Kim, I. *Biomolecules* **2022**, *12* (5).
(90) Chen, X.; Yang, H.; Song, X.; Liang, H.; Wei, Y.; Lu, J.; Barz, M.; Jin, R.; Nie, Y. *Chinese Chem. Lett.* **2022**, *34* (5), 107753.
(91) Ijadi Bajestani, M.; Mousavi, S. M.; Mousavi, S. B.; Jafari, A.; Shojaosadati, S. A. *Int. J. Biol. Macromol.* **2018**, *113*, 142–149.

(92) Ikeda, M.; Akagi, T.; Yasuoka, T.; Nagao, M.; Akashi, M. *J. Pharm. Biomed. Anal.* **2018**, *150*, 460–468.
(93) Bonduelle, C.; Makni, F.; Severac, L.; Piedra-Arroni, E.; Serpentini, C. L.; Lecommandoux, S.; Pratviel, G. *RSC Adv.* **2016**, *6* (88), 84694–84697.
(94) Brunato, S.; Mastrotto, F.; Bellato, F.; Bastiancich, C.; Travanut, A.; Garofalo, M.; Mantovani, G.; Alexander, C.; Preat, V.; Salmaso, S.; Caliceti, P. *J. Control. Release* **2021**, *335* (May), 21–37.
(95) Su, C. F.; Chen, Y. F.; Tsai, Y. J.; Weng, S. M.; Jan, J. S. *Eur. Polym. J.* **2021**, *152* (March), 110497.
(96) Klemm, P.; Solomun, J. I.; Rodewald, M.; Kuchenbrod, M. T.; Hänsch, V. G.; Richter, F.; Popp, J.; Hertweck, C.; Hoeppener, S.; Bonduelle, C.; Lecommandoux, S.; Traeger, A.; Schubert, S. *Biomacromolecules* **2022**, *23* (11), 4718–4733.
(97) Nguyen, T. P.; Easley, A. D.; Kang, N.; Khan, S.; Lim, S. M.; Rezenom, Y. H.; Wang, S.; Tran, D. K.; Fan, J.; Letteri, R. A.; He, X.; Su, L.; Yu, C. H.; Lutkenhaus, J. L.; Wooley, K. L. *Nature* **2021**, *593* (7857), 61–66.
(98) Tian, Z. Y.; Zhang, Z.; Wang, S.; Lu, H. *Nat. Commun.* **2021**, *12* (1), 1–11.
(99) Hanby, W. E.; Waley, S. G.; Watson, J. *J. Chem. Soc.* **1950**, 3239.
(100) Isidro-Llobet, A.; Álvarez, M.; Albericio, F. *Chem. Rev.* **2009**, *109* (6), 2455–2504.
(101) Grazon, C.; Salas-Ambrosio, P.; Antoine, S.; Ibarboure, E.; Sandre, O.; Clulow, A. J.; Boyd, B. J.; Grinstaff, M. W.; Lecommandoux, S.; Bonduelle, C. *Polym. Chem.* **2021**, *12* (43), 6242–6251.
(102) Zhang, Y.; Song, W.; Lu, Y.; Xu, Y.; Wang, C.; Yu, D.-G.; Kim, I. *Biomolecules* **2022**, *12* (5), 636.
(103) Hou, Y.; Wu, Z.; Dai, Z.; Wang, G.; Wu, G. *J. Anim. Sci. Biotechnol.* **2017**, *8* (1), 24.
(104) Rinaudo, M.; Domard, A. *Biopolymers* **1973**, *12*, 2211–2224.
(105) Bonduelle, C. *Polym. Chem.* **2018**, *9* (13), 1517–1529.
(106) O'Neil, K. T.; DeGrado, W. F. *Science (80-.).* **1990**, *250* (4981), 646–651.
(107) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7* (1), 95–99.
(108) Alekperov, D.; Shirosaki, T.; Sakurai, T.; Popova, G.; Kireev, V.; Ihara, H. *Polym. J.* **2003**, *35* (5), 417–421.
(109) Kratohvil, J. P. *Kolloid-Zeitschrift Zeitschrift für Polym.* **1970**, *238* (1–2), 455–459.
(110) Block, H. ; "*Poly(γ-benzyl-L-glutamate) and other glutamic-acid-containing polymers*"; 1st ed.; Breach, G. and, Ed.; Gordon and Breach, **1983**; Vol. 9.
(111) Miles, A. J.; Janes, R. W.; Wallace, B. A. *Chem. Soc. Rev.* **2021**.
(112) Agut, W.; Brûlet, A.; Schatz, C.; Taton, D.; Lecommandoux, S. *Langmuir* **2010**, *26* (13), 10546–10554.
(113) Lamy, C.; Lemoine, J.; Bouchu, D.; Goekjian, P.; Strazewski, P. *ChemBioChem* **2008**, *9* (5), 710–713.
(114) Salas-Ambrosio, P.; Tronnet, A.; Badreldin, M.; Ji, S.; Lecommandoux, S.; Harrisson, S.; Verhaeghe, P.; Bonduelle, C. *Polym. Chem.* **2022**, *13* (43), 6149–6161.
(115) Oya, M.; Uno, K.; Iwakura, Y. *J. Polym. Sci. A1.* **1970**, *8* (7), 1851–1862.
(116) Farthing, A. C. *J. Chem. Soc.* **1950**, 3213–3217.
(117) Iwakuha, Y.; Uno, K.; Kang, S. *J. Org. Chem.* **1965**, *30* (4), 1158–1161.
(118) Grimshaw, J.; Perera, S. D. *J. Chem. Soc., Perkin Trans. 2* **1989**, No. 11, 1711–1718.
(119) Heyns, K.; Walter, W.; Grützmacher, H. F. *Justus Liebigs Ann. Chem.* **1957**, *609* (1), 209–224.

(120) Kanazawa, H.; Matsuura, Y.; Tanaka, N.; Kakudo, M.; Komoto, T.; Kawai, T. *Bull. Chem. Soc. Jpn.* **1976**, *49* (4), 954–956.
(121) Oya, M.; Katakai, R.; Nakai, H.; Iwakura, Y. *Chem. Lett.* **1973**, *2* (11), 1143–1144.
(122) Fan, J.; Li, R.; Wang, H.; He, X.; Nguyen, T. P.; Letteri, R. A.; Zou, J.; Wooley, K. L. *Org. Biomol. Chem.* **2017**, *15* (24), 5145–5154.
(123) Xuesi, Chen; Jianxun, Ding; Jiandong, H. CN108003342A, **2018**.
(124) Wilder, R.; Mobashery, S. *J. Org. Chem.* **1992**, *57* (9), 2755–2756.
(125) Skoulas, D.; Stavroulaki, D.; Santorinaios, K.; Iatrou, H. *Polymers (Basel).* **2017**, *9* (11), 564.
(126) Otake, Y.; Nakamura, H.; Fuse, S. *Angew. Chemie - Int. Ed.* **2018**, *57* (35), 11389–11393.
(127) Sugimoto, T.; Kuwahara, T.; Liang, F.; Wang, H.; Tsuda, A. *ACS Omega* **2022**, *7* (43), 39250–39257.
(128) Leuchs, H. *Berichte der Dtsch. Chem. Gesellschaft* **1906**, *39* (1), 857–861.
(129) Ohkawa, K.; Nagai, T.; Nishida, A.; Yamomoto, H. *J. Adhes.* **2009**, *85* (11), 770–791.
(130) Devarayan, K.; Nakagami, S.; Suzuki, S.; Yuki, I.; Ohkawa, K. *Polymers (Basel).* **2020**, *12* (2), 1–15.
(131) Poroshin, K. T. *Bull. Acad. Sci. USSR Div. Chem. Sci.* **1956**, *5* (6), 755–758.
(132) Zalipsky, S.; Albericio, F.; Slomczynka, U.; Barany, G. *Int. J. Pept. Protein Res.* **1987**, *30* (6), 740–783.
(133) Akssira, M.; Boumzebra , M.; Kasmi, H; Dahdouh, A; Roumestant, M L; Viallefont, P. *Tetrahedron* **1994**, *50* (30), 9051–9060.
(134) Semple, J. E.; Sullivan, B.; Sill, K. N. *Synth. Commun.* **2017**, *47* (1), 53–61.
(135) Khair, B. M. A.-E.; Kōmoto, T.; Kawai, T. *Die Makromol. Chemie* **1976**, *177* (8), 2481–2489.
(136) Komoto, T.; Oya, M.; Kawai, T. *Die Makromol. Chemie* **1973**, *175*, 301–310.
(137) Ohnishi, S.; Kamikubo, H.; Onitsuka, M.; Kataoka, M.; Shortle, D. *J. Am. Chem. Soc.* **2006**, *128* (50), 16338–16344.
(138) Sack, I.; Macholl, S.; Wehrmann, F.; Albrecht, J.; Limbach, H. H.; Fillaux, F.; Baron, M. H.; Buntkowsky, G. *Appl. Magn. Reson.* **1999**, *431*, 413–431.
(139) Kuroki, S.; Takahashi, A.; Ando, I.; Shoji, A.; Ozaki, T. *J. Mol. Struct.* **1994**, *323* (C), 197–208.
(140) Weidner, S. M.; Just, U.; Wittke, W.; Rittig, F.; Gruber, F.; Friedrich, J. F. *Int. J. Mass Spectrom.* **2004**, *238* (3 SPEC. ISS.), 235–244.
(141) Xu, Q.; Gao, X.; Zhao, S.; Liu, Y.; Zhang, D.; Zhou, K.; Khanbareh, H.; Chen, W.; Zhang, Y.; Bowen, C. *Adv. Mater.* **2021**, *33* (27), 1.
(142) Wang, T.; Zhang, L.; Zhang, C.; Deng, D.; Wang, D.; Luo, L. *Molecules* **2023**, *28* (3).
(143) Meyer, V. E.; Lowry, G. G. *J. Polym. Sci. Part A Gen. Pap.* **1965**, *3* (8), 2843–2851.
(144) Harrisson, S.; Ercole, F.; Muir, B. W. *Polym. Chem.* **2010**, *1* (3), 326–332.
(145) Van Den Brink, M.; Van Herk, A. M.; German, A. L. *J. Polym. Sci. Part A Polym. Chem.* **1999**, *37* (20), 3793–3803.
(146) Umoren, S. A.; Solomon, M. M.; Saji, V. S. In *Polymeric Materials in Corrosion Inhibition*; Saviour A. Umoren, Moses M. Solomon, V. S. S., Ed.; Elsevier, **2022**; pp 49–81.
(147) Zhao, W.; Lv, Y.; Li, J.; Feng, Z.; Ni, Y.; Hadjichristidis, N. *Nat. Commun.* **2019**, *10* (1), 3590.
(148) Li, K.; Li, Z.; Shen, Y.; Fu, X.; Chen, C.; Li, Z. *Polym. Chem.* **2022**, *13* (5), 586–591.
(149) Chan, B. A.; Xuan, S.; Horton, M.; Zhang, D. *Macromolecules* **2016**, *49* (6), 2002–2012.
(150) Wu, Y.; Chen, K.; Wu, X.; Liu, L.; Zhang, W.; Ding, Y.; Liu, S.; Zhou, M.; Shao, N.; Ji, Z.;

Chen, J.; Zhu, M.; Liu, R. *Angew. Chemie Int. Ed.* **2021**, *60* (50), 26063–26071.
(151) Laconde, G.; Amblard, M.; Martinez, J. *Org. Lett.* **2021**, *23* (16), 6412–6416.

Chapter III: Hysteretic LCST of proline based peptidic polymers

1. Introduction

Smart materials based on polymers are a class of innovative materials that can change their properties in response to external stimuli, such as temperature, light, electrical or magnetic fields, pH, pressure, or other environmental factors. These materials are designed to have dynamic, adaptive, and often reversible behaviours, making them highly useful for a wide range of applications. This interesting property, able to adapt and respond to their environment, is also found in natural proteins. Because they combine the advantages of proteins and synthetic polymers, peptidic polymers (or polypeptides), that show stimuli responsiveness are of great interest for the development of novel materials. One example of this stimuli responsiveness found in polypeptides is the pH responsiveness of poly(L-glutamic acid) or PGA described in the previous chapter. In this chapter, we will focus on another stimulus for polypeptides, the change of temperature.

1.1. Thermoresponsiveness and synthetic polymers

Many water-soluble macromolecules, whether synthetic polymers or biomacromolecules (nucleic acids proteins, etc.) undergo phase separation to form condensates.[1] These processes are based on the same thermodynamic principles: the starting point is a thermodynamically stable solution that is subjected to conditions that induce the demixing, such as a change in temperature or the addition of a non-solvent.[2,3] This includes synthetic polymers that exhibit a lower critical solution temperature (LCST). These polymers are miscible with the solvent at all concentrations below the LCST, while above the LCST aggregation may occur to create two immiscible phases in thermodynamic equilibrium.[4] For any specific concentration, the point at which phase separation occurs is known as the cloud point temperature (T_{CP}, where $T_{CP} \geq$ LCST). The inverse phenomena is also observed in some synthetic polymers and is called upper critical solution temperature (UCST). In this case the polymer-solvent system is immiscible at low temperature and becomes miscible at a critical

point. Both phenomena result in binary polymer-solvent phase diagrams with a characteristic bell shape (**Figure 1**).

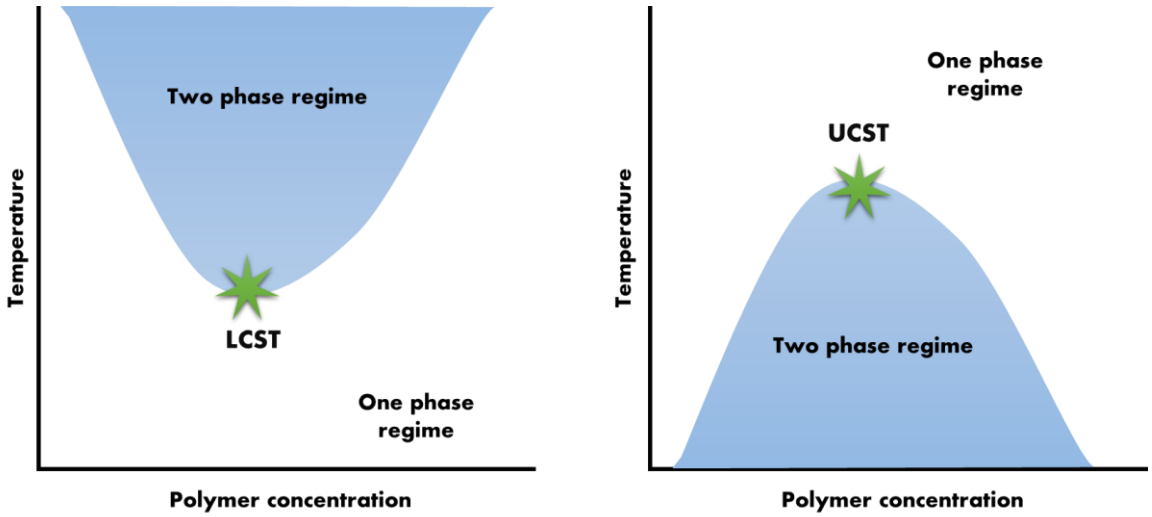


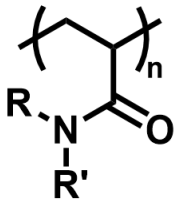
Figure 1 : *Plots of polymer binary solution and typical phase behaviours: on the left for LCST and on the right for UCST*

The LCST behaviour is determined by the thermodynamics of solution: the balance between polymer-solvent, polymer-polymer and solvent-solvent interactions. It is driven by changes in enthalpy and entropy on dissolution. Dissolution is thermodynamically favourable if the Gibbs free energy of dissolution (ΔG_{solv}) is negative (equation 1).

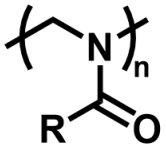
$$\Delta G_{solv} = \Delta H_{solv} - T\Delta S_{solv} \quad (1)$$

If the enthalpy of solution (ΔH_{solv}) is negative, but the entropy of solution (ΔS_{solv}) is positive, the solution will display LCST-type behaviour, becoming soluble when T is lower than $\Delta H_{solv}/\Delta S_{solv}$. This situation is unusual for small molecules, but relatively common for polymers with the right balance of hydrophilic and hydrophobic groups. In this case, the presence of polymer chains causes the water molecules to structure themselves, leading to a reduction in the entropy of the polymer solution. Heating the system increases the entropic term, leading to a positive free energy of dissolution leading the polymer chains to phase separate from the solvent. Two immiscible phases emerge, a polymer-rich phase and a polymer-poor phase. Phase separation is an essential step in most formulation processes specifically in drug delivery, nanotechnology, tissue engineering and biotechnology.[5] Some common examples of

synthetic polymers that exhibit an LCST in water are poly(N-alkyl acrylamide)s[6] and poly(oxazoline)s[7,8] (**Scheme 1**).



poly(*N*-alkyl acrylamide)s



poly(oxazoline)s

Scheme 1: *structures of LCST behaving poly(N-alkyl acrylamide)s and poly(oxazoline)s*

Poly(*N*-isopropylacrylamide) (PNIPAM) is a thermoresponsive *N*-alkyl acrylamide polymer where R=H and R'=isopropyl (**Scheme 1**) and is the most widely-studied LCST polymer. Along with another thermoresponsive class of polymers, poly(oxazoline)s (**Scheme 1**), both are structural isomers of polypeptides. In both cases, LCST behavior is driven by the balance between the polar, hydrogen-bond-forming amide groups and the hydrophobic alkyl chains. Dissolution is enthalpically favourable, but entropically unfavourable due to the structuring effect of the hydrogen bonds and the hydrophobic substituents. The temperature at which phase separation occurs can be modulated by varying the hydrophobic substituents,[9] or by copolymerization of two or more monomers with different substituents. PNIPAM for example has an LCST of 32°C that can easily be tuned in this way to around body temperature. This property, combined with its biodegradability and low toxicity have led to its possible use in biomedical applications.[10–12]

LCST polymers typically exhibit a small thermal hysteresis (< 5°C) in their phase separation behaviour, yet it has been shown in the literature that some LCST polymers, including PNIPAM,[13] can express larger hysteresis.[14] Some disubstituted vinyl polymers with amide moieties have been shown to express such hysteretic LCST.[15] This hysteresis could be explained by the metastability of the aggregates due to interchain interactions.

Similarly, block copolymers and mixtures of poly(methacrylic acid) with polyethylene oxide show hysteretic behaviour, forming stable aggregates on heating that dissociate only when cooled extensively.[16] Aqueous solutions of cellulose ethers are also known

to have hysteretic LCST[17] due to the form of aggregation[18] and are influenced by salt concentrations[19]. These polymers have been used to develop thermosensitive hysteretic hydrogels with shape memory effects, temperature sensors with memory function[14] and actuators[20]. Hysteretic LCST polymers are also found in living systems: such behaviour has recently been established for intrinsically disordered proteins. Specific recombinantly-produced sequences of elastin (ELPs) spanning a wide range of molar mass display hysteretic LCST and show promise for the development of future advanced materials.[21]

In addition to their thermoresponsive properties, synthetic polymers that are biodegradable, biocompatible, and non-toxic could be of interest for most applications. This makes thermoresponsive polypeptide a prime candidate. As described in chapter 1, the ROP of NCAs to produce amino acid-based polymers has gained much attention in the past 2 decades. The versatility of this polymerization and its compatibility with modern synthetic polymerization techniques allows polypeptidic backbones to be incorporated into synthetically engineered polymers. Peptidic polymers have unique biomimetic properties and conjugating polypeptides to classical thermoresponsive polymers is a common approach to prepare thermoresponsive synthetic polymers containing amino acids. Grafted and block copolymers of polypeptides with thermoresponsive polymers such as PNIPAM and PEG have been developed. Several peptidic backbones have been explored such as glutamate,[22–28] tyrosine,[29] cysteine, [30] lysine, [31] threonine,[32,33] valine,[34,35] isoleucine, [36] alanine, [37–39] phenylalanine,[40,41] glycine[36] and allylglycine[42] each bringing their unique properties to the copolymers. In these polymers, the polypeptide brings an additional contribution, notably thanks to its stimuli-responsive conformational rigidity (nanomaterials, gels, etc.). This design seems to be of great interest in many applications, especially in nanomedicine. To go further in the biomimicry, emerging approaches show an interest in replacing classical thermoresponsive synthetic polymers by fully peptidic polymers to produce even more biocompatible scaffolds.

1.2. Proteins and temperature

In living systems, the driving force of protein self-assembly to fold or to phase separate as nano- and micrometric objects involves supramolecular bonding, exploiting amphiphilic properties, secondary structure or coulombic interactions.[43,44] The folding

of proteins depends on numerous interactions between amino acid side chains and between these chains and the solvent, typically water.[45] As discussed in the introductory chapter, the three-dimensional structure of a polypeptide is only stable within a narrow range of physico-chemical conditions. Denaturation can be induced by increasing pressure, extreme pH or the addition of certain salts, alcohol or large quantities of surfactants.[44,46,47] Temperature also affects the behaviour and structure of proteins. Most proteins start to aggregate when the temperature is increased.[45,48] Denaturation first causes the breaking of low energy bonds within the molecule, which leads to a reorganization of the chain and generally to the emergence of the more hydrophobic areas of the protein.[49,50] This reduces protein-solvent interactions in favour of protein-protein interactions leading to the irreversible formation of aggregates (**Figure 2**).[51,52]

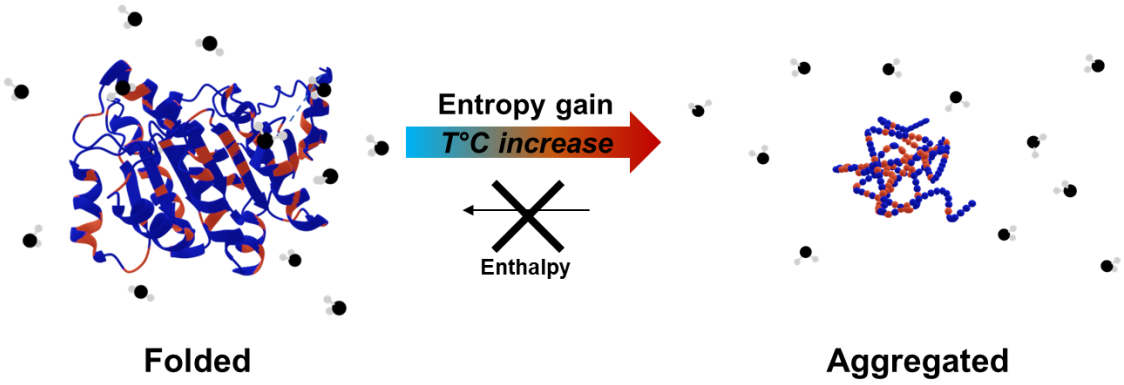


Figure 2 : *On the left, protein in its folded state at certain conditions (temperature, pH, etc…). This form is stabilized due to interaction between water molecules and hydrophilic areas of the macromolecule (blue patches). When temperature is increased, the gain in entropy manifests in the movement of water molecules and the loss of the interaction. The protein thus collapses on itself in an aggregated state (on the right).*

In contrast, some intrinsically disordered proteins (IDPs) undergo reversible phase separations. Their unique behaviour towards temperature changes stems from their solubility in their unfolded state, as discussed in the introduction of this manuscript. IDPs with a large number of hydrophobic amino acids in their sequences can form fused globules (coacervates) and present polymer-like phase separation upon temperature variation.[53] This can manifest itself as a lower critical solution temperature (LCST) behaviour that will be discussed in more detail in the next part. In biology, phase separation of biomacromolecules is at the origin of complex intracellular organization

such as the formation of membraneless organelles.[54] These biomacromolecular condensates have been identified in both the nucleus and cytoplasm of living systems, implying that they offer advantages as a compartmentalization strategy.[55,56] They are usually formed by assemblies of proteins with a high content of intrinsically disordered regions that undergo a demixing transition to form two or more coexisting phases. These phenomena have consequences for (bio)chemical reactivity[1,57] and are also associated with neurodegenerative diseases.[58]

1.3. Thermoresponsive polypeptides

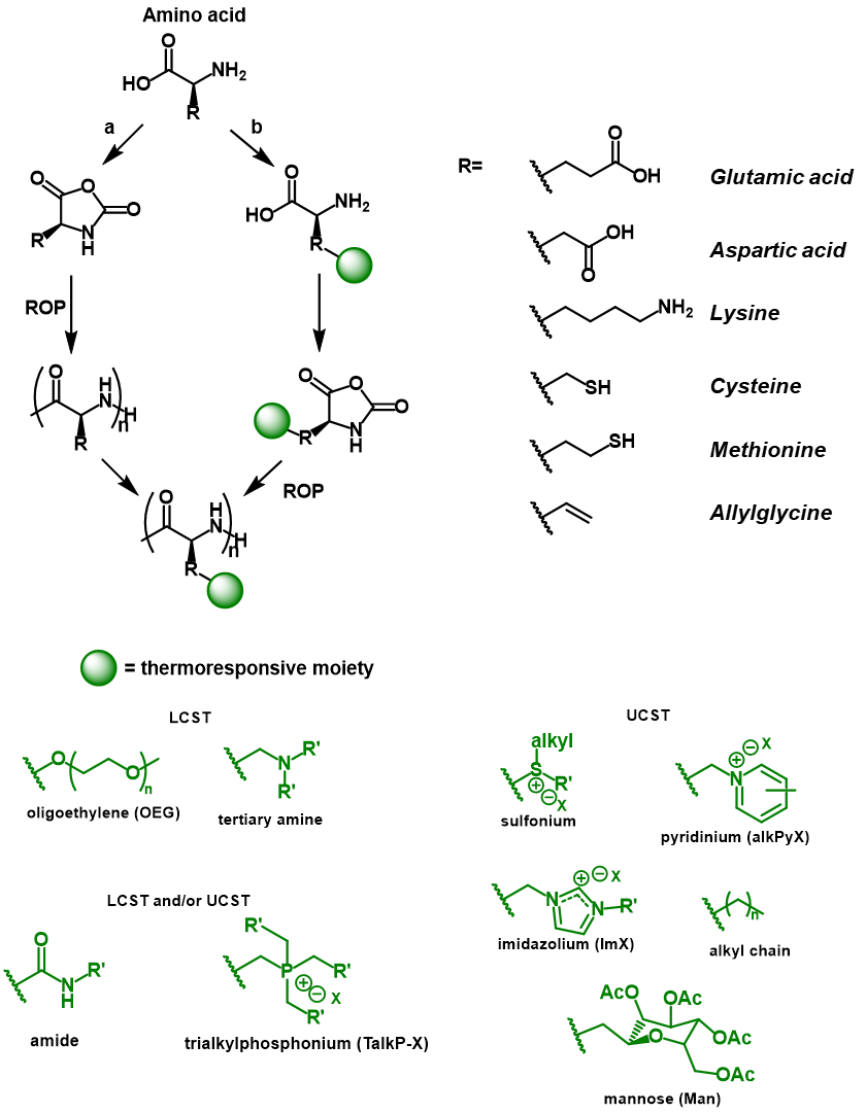
Generally, polypeptides synthesized using ROP of NCAs do not possess thermoresponsive behaviour if the amino-acid side chain is not modified. Thermoresponsive properties can however be introduced by functionalizing these side chains with unnatural moieties. Two different approaches have been employed in this direction (**Scheme 2**):

a) synthesizing amino acids with the desired side chain before making the NCAs and polymerizing them.

b) modifying the polymer using post-polymerization chemistry to introduce the stimuli responsive group.

Overall, NCAs based on amino acids with functional side chains (**Scheme 2**) must be used. As glutamate and lysine are the two most produced amino acids,[59] they have been the most studied in the scope of the ROP of NCA and in the synthesis of thermoresponsive polypeptides. By using both synthesis pathways presented in **Scheme 2**, it was possible to produce polypeptides based on both these amino acids possessing highly tuneable thermoresponsivness.

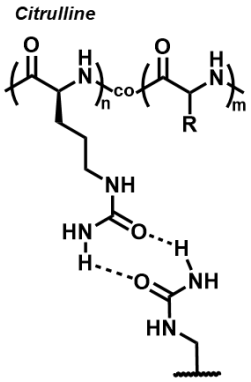
Oligoethylene glycol moieties (**Scheme 2**) are the most used to confer LCST behaviour to the polypeptides[60–69], due to their temperature-responsive dehydration mechanisms. Furthermore, their hydrated or dehydrated state can stabilize certain soluble or insoluble secondary structures of the polypeptide backbones.[70–79]



Scheme 2: *Two synthesis pathways can be used to introduce thermoresponsive behaviour to polypeptide backbones. Pathway* ***a*** *relies on the post-polymerization functionalization of polypeptides with a thermoresponsive moiety. Pathway* ***b*** *corresponds to the addition of a thermoresponsive moiety to an amino acid while pathway. The amino acids used all have side group functionalities. The thermoresponsive moieties confer either LCST, UCST or both behaviours. In the last case the behaviours may be observed on the same polymer or separately on different polymers.*

Many chemistries have been employed in order to insert these thermoresponsive moieties including click chemistry. Other functional amino acids[80–82] (Scheme 2) have also been employed including thiol bearing ones like cysteine[83–89] and methionine[90]. Amides and tertiary amines on side chains have also conferred LCST behaviour to polypeptides[81,91] while ionic moieties such as sulfoniums[90], pyridiniums[92] and imidazoliums[93] confer UCST behaviour.

Citrulline (**Scheme 3**) is a natural non-proteinogenic amino acid with a ureido function on its side chain. These copolymers display UCST behaviour due to their hydrogen bonding abilities that get disrupted when heated and allow solubilization. [94,95]



Scheme 3 : *Citrulline based copolymers and hydrogen bonding of the ureido group allowing for the UCST behaviour.*

These copolymers were generally prepared by modifying poly(*L*-ornithine) prepared through ROP of the corresponding NCAs.[96] Their thermoresponsive behaviour was easily tuned through copolymerization with other monomers bearing different functional groups[97,98] or by modifying the ureido group for example through alkylation.[99] These polypeptides are stable in physiological conditions, are both enzymatically[96] and biodegradable[98] and can also promote cell adhesion coatings[99] and controlled delivery of active molecules[97].

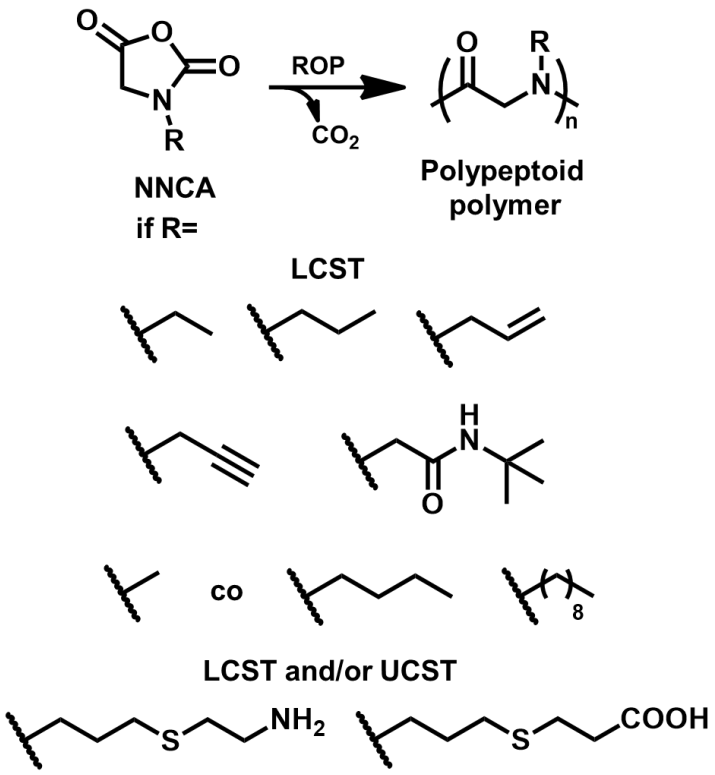
These examples show the possibilities offered by chemical modification of side chains to bring thermoresponsive behaviour to polypeptides. Based on these pioneering works, one can imagine many designs to introduce sufficient hydrophobicity and limited secondary structuring, two parameters also found with IDPs that make them thermoresponsive in living systems. Yet most published methods involve multiple steps and tedious purifications. It would thus be useful to be able to synthesize polypeptides whose main chain actively participates in the phase separation processes. This possibility will be presented in the next part starting with polypeptoids then moving to the peculiar case of polyproline.

1.4. Polypeptoids and thermoresponsivness

Polypeptoids, also called poly(*N*-substituted-glycines), are *N*-alkylated analogues of synthetic polypeptide polymers (**Scheme 4**). The side chain is thus presented on the

amine of the amide bond, and they lack the stereogenic centre of their peptidic counterparts.

The poly(*N*-alkylated glycine) backbone can be prepared by ROP of *N*-alkylated-*N*-carboxyanhydrides (NNCA).[100] The physicochemical properties of polypeptoids can be tailored by N-alkylation, enabling control over the backbone conformation,[101] their solubility,[102] and the thermal and crystallization properties.[103,104] Thus peptoid polymers may exhibit UCST or LCST depending on their structure[105–114] (Scheme 4) and their environment.[115,116] Their thermoresponsive behaviour can be further tuned by varying the molar mass, architecture,[117] concentration, and composition.[118,119]



Scheme 4: *Schematic representation of the ring opening polymerization of N-alkylated N-carboxyanhydrides (NNCA) to obtain polypeptoid polymers. The choice of the side chain (R group) can confer thermoresponsive behaviour innate to the polypeptoid backbone. Other thermoresponsive moieties can be added to confer the property to polypeptoid and are also illustrated in the following section.*

The presence of a tertiary amide in the backbone is the key reason explaining their innate thermoresponsiveness in comparison to their non-alkylated counterparts and this behaviour stems from the cis-trans interplay of the peptoid backbone.[8] For this reason, the only *N*-alkylated natural amino acid, proline, should be considered as a scaffold for thermoresponsive polypeptides.

1.5. Polyproline and thermoresponsiveness

1.5.1. Secondary structuring

Fundamentally, proline bridges the gap between polypeptides and polypeptoids. This amino acid possesses both the tertiary amide of the polypeptoids and the chiral centre of most polypeptides. The peptide bond is a partial double bond due to the delocalization of electrons on several atoms. The O, C_α and N atoms of this bond are coplanar which induces cis/trans isomers. When the nitrogen atom is unsubstituted, only the trans conformation is observed,[44] but when the backbone is *N*-alkylated both conformations (**Figure 3**) are found in an equilibrium which depends on both the nature of the substitution and the environment.[120] In proline-based polymers the combination of chirality and conformation induces unique secondary structuring.

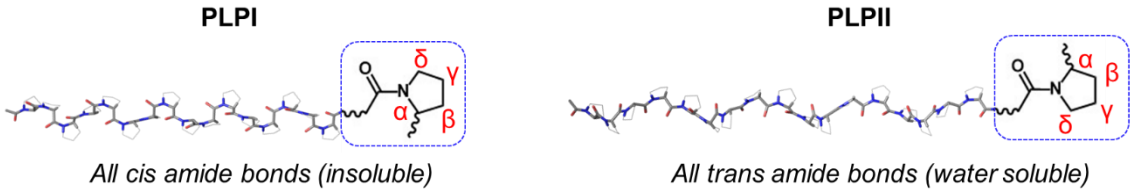


Figure 3 : *A representation of the two secondary structures of proline-based polymers and the conformation of the pyrrolidine ring in both cases. On the left the PLPI structure where the amide bonds are in an all-cis state (10_3 symmetry, 3.33 residues/turn, 1.9Å residues per repeat unit, dihedral angles φ=-75°; ψ=+160°). On the right the PLPII structure where the amide bonds are in an all-trans state (3_1 symmetry, 3 residues/turn, 3.12Å residue repeat, dihedral angles φ=-75°; ψ=+145°).*

The first structure is known as **PLPI** and is entirely composed of cis amide bonds as shown in **Figure 3**.[121] This structure is possible thanks to the *N*-alkylation of proline, which makes the probability of the presence of a cis bond within the chain about 100 times higher than for other amino acids.[44,122] This compact right-handed helix consists of 3.3 residues per turn, and usually forms only in organic media.[122] The second conformation contains all trans amide bonds in the backbone and is called **PLPII** (**Figure 3**). This more extended helix with 3 residues per turn is water-soluble so is naturally more abundant in proteins and plays important roles in many biological processes.[123,124] Its stability in water can be attributed to hydrogen bonding with water molecules that is only possible in the trans conformation.[122,125] Furthermore, most intrinsically disordered proteins contain proline residues which are considered to

disrupt the formation of ordered secondary structures such as alpha helices and beta sheets.[124] PLPII structures resemble random coils[121] and are believed to be adopted in many unfolded parts of the proteome, highlighting their importance to the physicochemical properties of proteins and polypeptides.[126]

1.5.2. Thermoresponsive behaviour

The reversible switch between the cis and trans proline or PLPI and PLPII conformations is biologically relevant and occurs in proteins. For instance passing from the cis form to the more rigid and extended trans form can promote the compaction or the expansion of certain protein regions.[124,127] Furthermore, the naturally slow interconversion between the forms can act as a molecular timer in some physiological and pathological processes.[58,128–130] The interplay between the conformations can also be controlled via enzymatic reactions to steer protein folding thus controlling the amplitude of some cellular behaviours.[124,129]

The trans conformation is slightly more energetically favourable and there exists an energy barrier (~30kcal.mol^{-1}) to transform it to the cis conformation and back again, thus affecting the secondary structure.[129] Experimentally the interconversion between the two states is can be effected by changing solvent polarity. As mentioned earlier, the PLPI conformation is usually observed in organic solvents and is soluble in aliphatic alcohols such as isopropanol[122,131] or butanol[132]. Conversion from PLPI to PLPII occurs when polyprolines are dissolved in trifluoroethanol[132]. It is also possible to promote this configurational change by addition of weak acids like acetic, propionic, and formic acid while strong acids promote both transitions.[133,134] When dispersed in water, PLPI converts slowly to PLPII, but this process usually takes several days.[135–137] Salts, like lithium or calcium salts, can stabilize the trans conformation in organic solvents and also the cis conformation in aqueous solutions, through interactions with the backbone.[136,138] During such conversions intermediates containing both types of amide conformations are formed.[121,122]

Temperature also plays a role in this interconversion. The PLPI conformation can be partially converted to PLPII when heated in isopropanol solutions.[123] The PLPII heat precipitation from aqueous solution has been studied in the past.[139] Solubilized PLPII in aqueous salt solutions can express a reversible LCST behaviour. When these solutions where heated, the PLP precipitated out of solution at temperatures that depend on the nature and concentration of the salt as well as the rate of heating.[139]

The resolubilization of the precipitated polyprolines was not analysed in most of these studies but it was observed to take place at much lower temperatures than the precipitation.[140] It has also been observed that aqueous solutions of PLPII largely retain their PLPII conformation upon heating.[139–141] Highly concentrated aqueous solutions of PLPII can form thermoresponsive gels.[142] These solutions gelled at low temperatures (-15°C and 5°C) and when heated to 35°C, the gels became stiffer due to change in the gels' physical network. In another study, the authors showed that soluble PLPII polymers in D_2O formed aggregated films while changing its conformation into a more highly packed one.[143]

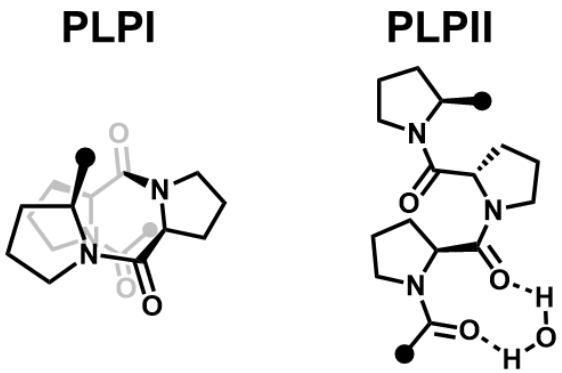
Thus, polyproline is a prime candidate to design thermoresponsive polymers based on natural amino-acid. Even though, there is still no full study on the liquid-liquid phase separation of high molecular weight polyprolines due to the difficulty of obtaining them synthetically. To carry this out, one remaining challenge is still to produce highly controlled and predictable PLP structures. This subject will be discussed in the next section 1.5.3.

1.5.3. ROP of proline NCA

Most of the above-mentioned studies have been conducted on proline-based peptide of small Dp or proline oligomers (lower than 10 units). When NCA ROP was used, polymers had high dispersities and were fractionated before use.[132] This is because ROP of NCAs is usually carried out in organic solvents in which the PLPI conformation is favoured. As mentioned above this insoluble conformation precipitates when polymerization starts and thus hinders the control of the polymerization process. This leads to the formation of highly disperse PLPI that must later be transformed into water soluble PLPII.[142]

The ROP of proline NCA (Pro NCA) has benefited from recent and impressive advances. For example block copolymers can be synthesized by first growing a soluble block (poly(ethylene oxide), polylysine, polyglutamate)[144] then sequentially polymerizing the Pro NCA to ensure that the block copolymer remains soluble, or at least well-dispersed. On another hand, *Kramer* et al. have shown that organometallic catalysts based on nickel or cobalt catalyse the ROP of proline NCA and its derivatives.[137] Using an amido-amidate nickel catalyst, they were able to produce high molecular weight polyprolines, as analysed by ^{1}H NMR. Unfortunately, these polymers were not analysed by SEC which would have allowed their dispersity to be evaluated.

Furthermore, the accepted mechanism of these catalysts (described in chapter **1**) supposes a hydrogen atom on the amine, which is lacking in these *N*,*N*-dialkylated monomers. An important contribution regarding the ROP of Pro-NCA has been published in 2022, *Lu* et al. showed that the ROP of Pro-NCA could be extremely accelerated and enhanced in the presence of water.[125] They showed that the presence of even catalytic quantities of H_2O accelerated the propagation so that the polymerization usually requiring days was fully achieved in minutes. This behaviour was attributed to the formation of the specific interaction displayed by H_2O with the trans conformation of the amide bonds. This allows to form the less sterically hindered and more active PLPII helix that accelerates the polymerization (**Scheme 5**). The authors also showed how interactions with water stabilize intermediates during propagation.



***Scheme 5**: geometric representation of PLPI helix on the left and PLPII helix on the right with water interaction with adjacent carbonyls.*

Thus, when acetonitrile with sufficient water content (above 20%) was used for polymerization, the rate was extremely fast as the PLPII form was also soluble. The polymers obtained this way showed narrow dispersities and controlled molecular weights in agreement with the M/I ratio.

2. Results and discussion

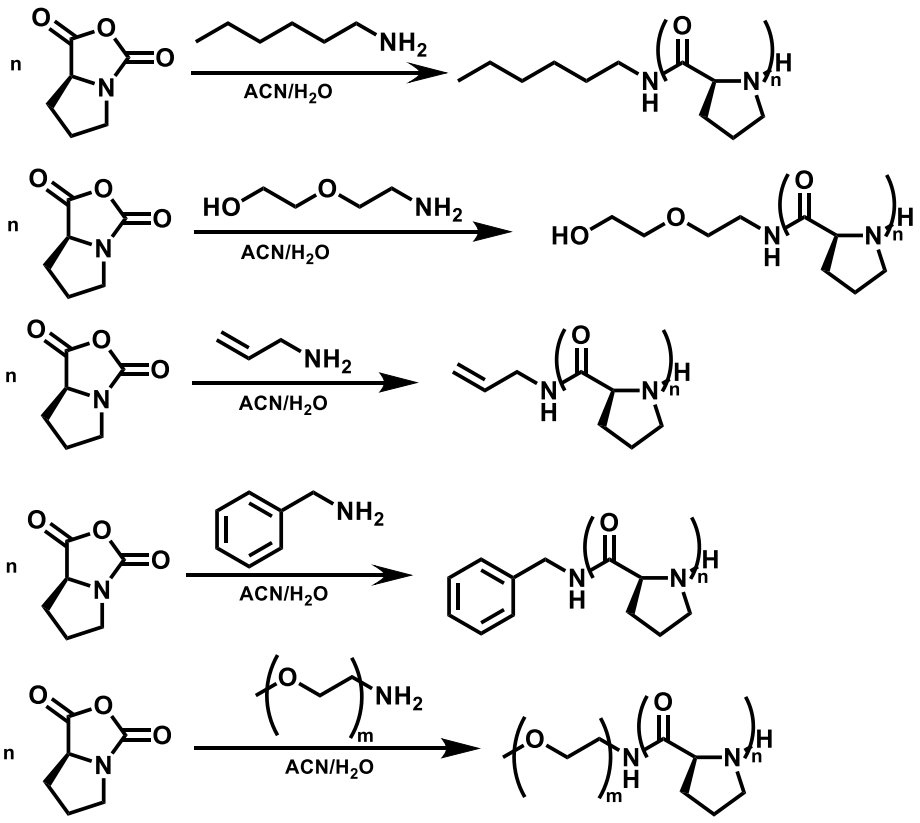
Following the aqueous ROP methodology developed by *Lu* et al., in this research work, the polymerization of Pro-NCA using different initiators was first developed and the resulting polymers thoroughly analysed (section 2.1). Other synthesis routes were then explored so that to prepare thermoresponsive PLPs using new methodological conditions (section 2.2). The thermoresponsive behaviour of these proline-based

polypeptides was then comprehensively studied and the underlying mechanisms related to the secondary structuring were explored (section 2.3 and 2.4). Several factors to fine-tune the thermoresponsive behaviour were considered including salt concentration, macromolecular engineering, and formulations (**2.5**).

2.1. Synthesis and analysis

2.1.1. Water-assisted ROP

Several initiators were used to achieve the ROP of Pro NCA to extend the scope of this mechanism. This synthesis was optimized for the preparation of well controlled PLPs. Pro NCA was polymerized at 0.3 M in a mixture containing 60% acetonitrile and 40% water using different initiators (**Scheme 6**) at 10°C and under vigorous stirring. When hexylamine was used as an initiator, the M/I ratio was also varied to assert a control over the molecular weight.



Scheme 6: Polymerizations of Pro NCA using the water assisted ROP and different initiators as well as M/I.

The resulting polymers were characterized by ^{1}H NMR and aqueous SEC and the results are illustrated in **Table 1**.

***Table 1** : Analysis and results for polymers **1-10** and copolymers **1-2** synthesis.*

Polymer	Targeted structure[a]	Yield (%)	DP by ^{1}H NMR[b]	Theoretical Mn (kDa)	Mn by SEC (kDa)[c] Đ[d]
Polymer **1**	HA-PLP$_5$	65	5	0.6	0.9 1.24
Polymer **2**	HA-PLP$_{20}$	61	18	2.0	2.4 1.03
Polymer **3**	HA-PLP$_{30}$	88	30	3.0	4.0 1.08
Polymer **4**	HA-PLP$_{50}$	90	56	5.0	4.9 1.01
Polymer **5**	HA-PLP$_{70}$	92	68	6.9	7.8 1.03
Polymer **6**	HA-PLP$_{100}$	80	113	9.8	12.4 1.02
Polymer **7**	HA-PLP$_{200}$	75	230	19.5	20.1 1.04
Polymer **8**	BA-PLP$_{70}$	86	71	6.9	7.1 1.01
Polymer **9**	AA-PLP$_{70}$	49	*e*	6.9	*f*
Polymer **10**	EtOHA-PLP$_{70}$	60		6.9	
Polymer **11**	HA-PDP$_{100}$	75	120	9.8	11.2 1.01
Copolymer **1**	PEG$_{5k}$-PLP$_{50}$	65	50	9.9	10.9 1.04
Copolymer **2**	PEG$_{10k}$-PLP$_{100}$	78	100	19.7	19.1 1.04

[a]determined by the initial feed ratio M/I, [b]determined by the integration of the different backbone peaks in comparison to the initiator peaks in D_2O ^{1}H NMR, [c]determined by dn/dc of polyproline (0.1616) or of block copolymers (0.1444) in ACN/acetate buffer SEC and MALS detection, [d]determined by SEC peaks, [e]end-chain groups were not easily identifiable, [f]chromatograms showed double peaks. HA=hexylamine, BA=benzylamine, AA=allylamine, EtOHA=2-(2-aminoethoxy) ethanol, PEG=polyethyleneglycol.

Using the small hydrophobic amine, hexylamine, it was possible to synthesize PLPs with polymerization degrees varying from 5 to 200 (polymers **1**-**7** in **Table 1**). The synthesized polymers were purified by dialyzing against MilliQ water and freeze-drying to obtain polyprolines with good yields in their PLPII conformation directly after polymerization. Furthermore, these polymers were directly water soluble. On analysis

by ^{1}H NMR in D_2O the different peaks were easily assigned, and the DP was calculated by integration relative to the initiator peaks (**Figure 4** and **Figure S1-5**).

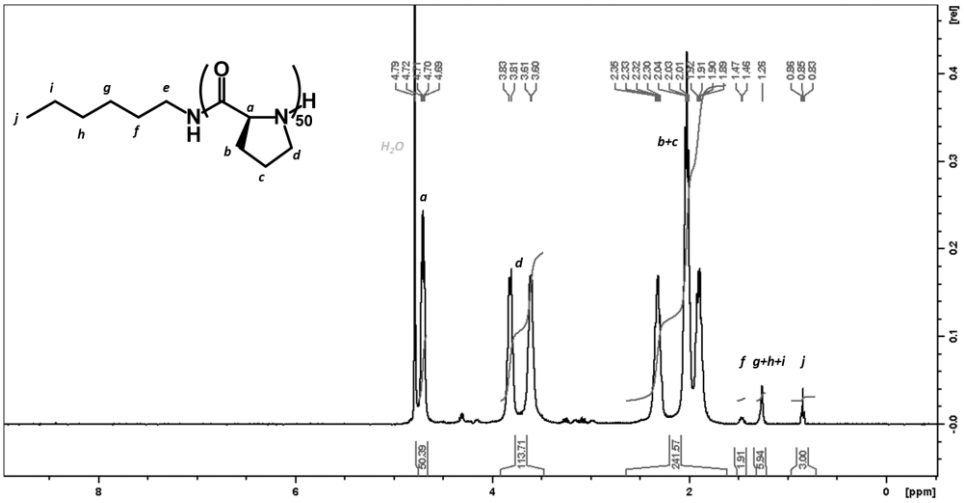


Figure 4 : *^{1}H NMR for polymer **4** in D_2O. The polymer structure of is represented with the peak assignation. The calculation of the Dp is shown by the calculations from the integration of the initiator peaks (j,i,h,g and f)) and the respective backbone peaks of proline repeat units (b,c and d).*

The analysis of these copolymers by aqueous SEC supports the formation of well controlled high molar mass polymers with narrow dispersities (**Figure 5** and **Figure S9**). Furthermore, the dn/dc of polyproline was calculated to be 0.1616 in these conditions to correctly estimate the Mn of the synthesized polymers. The results were in agreement with the theoretical Mn (**Table 1**) where the repeat unit of 97 was multiplied by the DP value. The synthesis of the mirror image of PLP, PDP, using *D*-Pro NCA was also successful and the produced polymer **11** (HA-PDP_{100}) had comparable characteristics to polymer **6** (HA-PLP_{100}). This shows that the polymerization can form helices of either orientation, as PDP forms PPII structures with a right-handed helix.

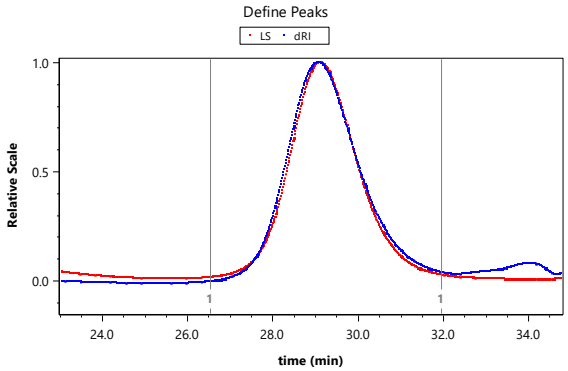


Figure 5 : *SEC chromatogram for polymer **4**. Both MALS and RI detection are shown.*

Other small amines with varying hydrophilicity were also explored as initiators in this ROP technique. While benzylamine (polymer **8** in **Table 1**) gave similar results to hexylamine (**Figure S9-Figure S10**), more hydrophilic amines such as allylamine (polymer **9** in **Table 1**) and 2-(2-aminoethoxy) ethanol (polymer **10** in **Table 1**) did not produce well controlled polyprolines. SEC analysis of these polymers showed the formation of two polymeric populations (**Figure 6**).

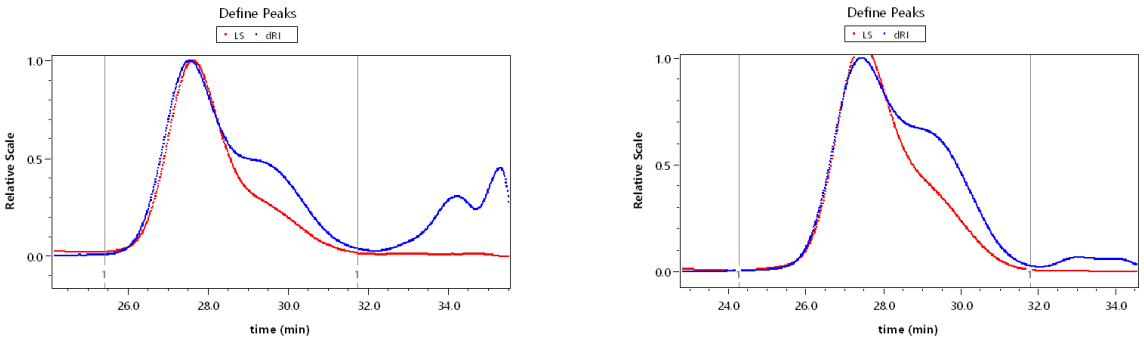


Figure 6 : *SEC chromatogram for polymer **9** (on the left) and polymer **10** (on the right). Both MALS and RI detection are shown.*

Unlike hexylamine and benzylamine, these hydrophilic initiators are water soluble. To better understand the phenomena all the polymers were analysed by MALDI-TOF to confirm their structure and the initiation moiety. The different peaks had a 97.12 g/mol difference confirming that these high molecular weight species have proline repeat units. Furthermore, the calculations confirmed the end-chain initiation by hexylamine (101.19 g/mol for polymers **1-7**) and benzylamine (107.15 g/mol for polymer **8**). Polymers **1-8** thus showed corresponding structures using MALDI-TOF analysis confirming the synthesis (**Figure 7** and **Figure S14**).

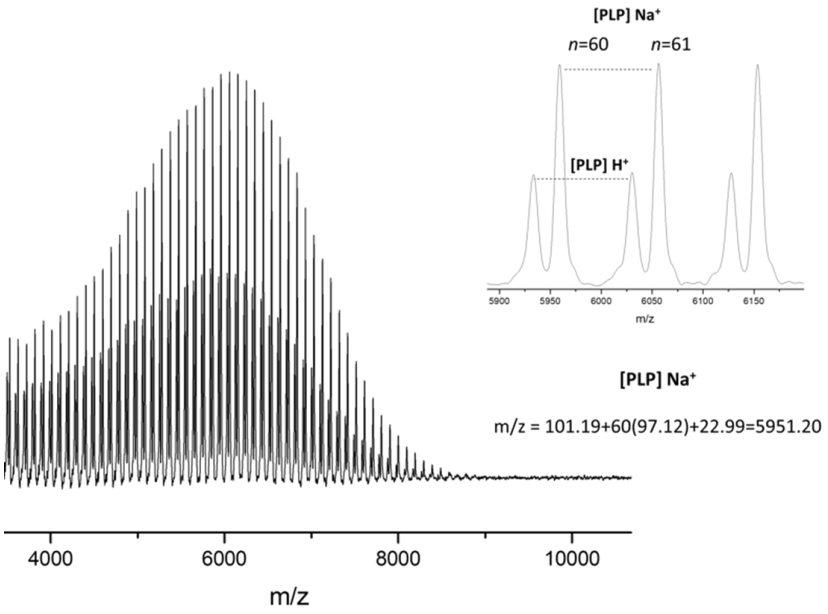


Figure 7: *MALDI-TOF spectra of polymer **4**. (On the right) assignment of peaks to proline units along the chain.*

On the other hand, MALDI-TOF analysis of polymers **9-10** showed two populations as shown previously in SEC analysis. Both populations had proline repeat units (97.12 g/mol) confirming that polyprolines have been synthesized. While the higher molar mass population had the amine initiator at the chain-end, the lower molar mass population in both cases was clearly initiated by water (18.015 g/mol) (**Figure 8**). It is thus hypothesized that since water could also initiate the polymerization, the slightly soluble amines tend to react faster with the NCAs avoiding the hydrolysis. On the other hand, water soluble amines might generate -OH moieties capable of the initiation.

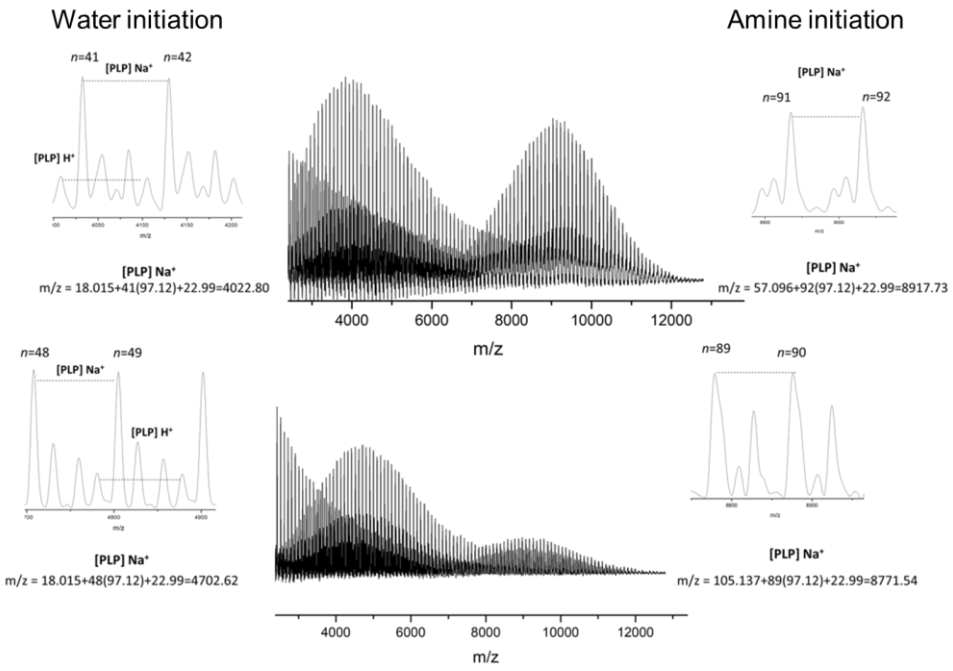
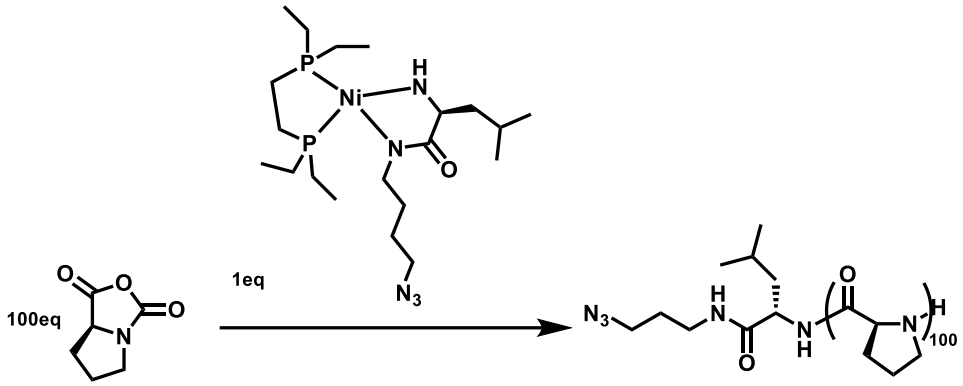


Figure 8 : *MALDI-TOF spectra of polymer **9** (on the top) and **10** (on the bottom). Assignment of peaks to proline units along the chain are shown as well as the calculation for chain-end group. The lower molecular weight peaks (on the left) show water initiation with -OH and -H at each end, while the bigger polymer chains on the right show the amine at the chain end.*

A final series of polymerizations was then developed using PEG-NH_2 macroinitiators. Using PEG5k-NH_2 to ROP Pro NCA with an M/I=50 and PEG10k-NH_2 with an M/I=100, block copolymers with equivalent weight ratios were synthesized (copolymers **1**-**2** in **Table 1**) using similar conditions as for the before mentioned polymers. The formed block copolymers were then dialyzed and lyophilized to obtain the pure block copolymers. Although the PEG initiators are water soluble, ^{1}H NMR (**Figure S3-4**), SEC (**Figure S11**) and MALDI-TOF (**Figure S15**) analysis confirmed the successful synthesis of well controlled PEG-polypeptide block copolymers without any water initiation and with adequate yields excluding hydrolysis reaction to amino acids.

2.2. Organic media polymerization

2.2.1. Organometallic catalysis

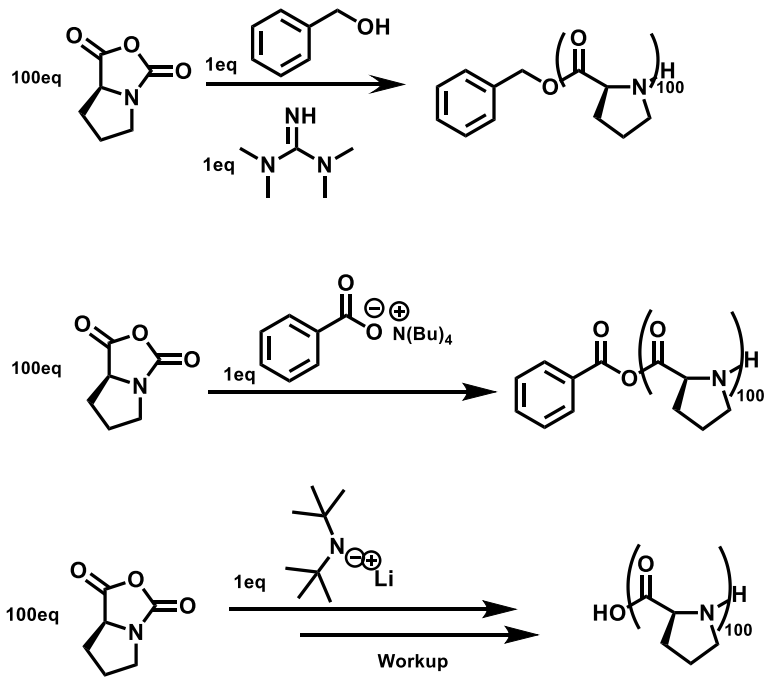


Scheme 7: *ROP of proline NCA using Ni amido amidate initiator catalyst. Theoretical structure of polymer **12** is presented.*

Following the reported protocol for the synthesis of PLP using organometallic catalysis[137], similar experimental conditions were tested. In the context of a collaboration with Professor Deming's laboratory in UCLA, nickel based organometallic catalysts (**Scheme 7**) were obtained to achieve this ROP. In dry DMF, Pro NCA was polymerized using the amido amidate moiety initiation (**Scheme 7,** mechanism is described in the supporting information **Scheme 10**). The monomers were only slightly soluble in DMF and when the polymerization was conducted the medium became more turbid over time. While higher concentration trials formed organogels in DMF, 50 mg/ml was found to be an adequate concentration. Polymer **12** (M/I=100) was polymerized using these conditions and was then purified by water washing to remove any nickel species. After dialysis and lyophilization the polymer was obtained with 90% yield. Although such polymerizations are reported to be rapid, it took 7 days for this polymerization to complete. While NMR analysis showed typical PLP peaks and an adequate DP (**Figure S6**), SEC analysis in TFE showed the formation of polymers with a large dispersity (2.18).

2.2.2. Other polymerization systems

Most of the recent advances in the catalysis of ROP of amino acid NCA discussed in chapter **1** have not been studied on *N*-alkylated monomers. Thus, their use in the ROP of Pro NCA was still not considered. Of the few systems that showed promising results with other *N*-alkylated monomers are tetramethylguanidine (TMG) catalysis with alcohol initiators,[145] TBBA[146] and LiHMDS initiation[147] (**Scheme 8**).



***Scheme 8** : reaction schemes for polymers **13-15** using TMG, TBBA and LiHMDS.*

Polymers **13**, **14** and **15** were synthesized using these three pathways respectively. In the case of TMG and TBBA, the polymerization mechanism is similar to those presented in chapter **1**. The NCA was thus reacted in dry DMSO where it was soluble with the different initiators. The media slowly turned to a milky solution. After completion verified by FTIR, the polymers were dialyzed against MilliQ water and lyophilized. The polymers showed adequate yields (80-90%) and were then analyzed for their structure. NMR analysis of polymers **13-14** showed typical PLP peaks but no initiator peaks were identifiable.

LiHMDS was used to produce cyclic and linear polypetoids via a zwitterionic ring expansion polymerization mechanism (supporting information **Scheme 10**).[147] Polymer **15** was thus prepared following the protocol developed in the literature with Pro NCA dispersed in DMF and LiHMDS added as a THF solution to start the polymerization. The polymerization was concluded quickly (less than 1 day) and polymer **15** was obtained at 85% yield after precipitation in ether and evaporation of the solvents. NMR showed typical PLP peaks (**Figure S7**).

Polymers **12-15** were insoluble in water as they were formed in PLPI conformation. When their aqueous suspension was left to stir for 1 week at 5°C, they were readily soluble as they converted to PLPII conformation. The polymers were thus solubilized in ACN/buffer for SEC analysis. SEC analysis of polymers **13-14** in aqueous conditions

revealed lower molecular weight, multimodal and broad peaks with high dispersities (**Figure S12**). Polymer **15**'s chromatograms of high molecular weight polymers (7.2 kg/mol) and adequate dispersity of 1.26 (**Figure S13**). Further investigation into the use of LiHMDS in the ROP of Pro NCA is needed to evaluate the architecture of the polymers as they can be cyclic or linear.[147]

2.3. Thermoresponsive behaviour

Since only the polymers prepared via the aqueous ROP had controlled structures, only these polymers were thoroughly analysed to evaluate their thermoresponsiveness. Many techniques can be used to assess the LCST behaviour of polymers including various turbidimetric techniques and spectroscopy (UV, light scattering, etc.) along with calorimetric measurements.[148]

2.3.1. UV-Vis turbidimetric study

LCST is characterized by a difference between the cloud point (T_{CP}) being the temperature at which the aggregates appear upon heating and the clearing point (T_{CL}) the temperature at which solubilization of the chains occurs on cooling.[149] When the values of these temperatures are not equal, the LCST is considered to be hysteretic and the difference between these two values is the extent of hysteresis (X_H). It should be noted that the value assigned as the transition temperature depends on the technique used to measure it. The dispersity of the polymer may also play a role, as highly disperse polymers will undergo fractionation as they precipitate, leading to broader transitions. For these reasons, it is sometimes difficult to compare cloud point values between different studies.[6] To be able to compare the effect of parameters such as concentration, chain length, and solvent composition, a precise method was employed using UV-Vis spectroscopy to monitor the thermoresponsive behaviour. This method was selected as it is precise and easy to use. *Zhang* and coauthors have recently provided recommendations in a recent tutorial review on using this technique which were followed in this study.[148]

Hysteresis analysis

PLPs synthesized by water assisted ROP (**Table 1**) were characterized by turbidimetric analysis using UV-Vis spectroscopy. Polymer **4** (HA-PLP$_{50}$) was dissolved in water at a concentration of 5 mg/mL. This solution was then heated at a rate of 1°C/min and

monitored by UV-vis spectroscopy at a wavelength of 550 nm, since polyproline moieties do not absorb at this wavelength.

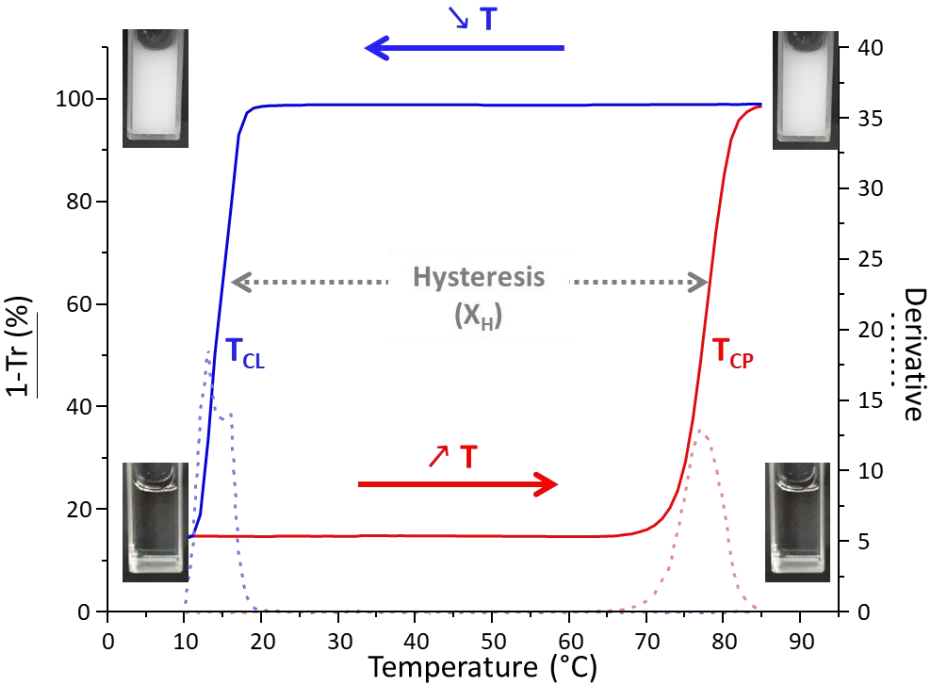


Figure 9: *Characterization the thermoresponsive behaviour of polymer **4** at 5mg/ml in an aqueous solution. The absorbance measured by the machine was converted to transmittance and the 1-Transmission is plotted for easier relating the soluble state at low values (pictures on the bottom of the plot) and insoluble state at higher values (pictures at the top). The derivatives of both heating and cooling ramps are presented, and the T_{CP} and T_{CL} are determined at the inflection points respectively. These values allow us to calculate the extent of hysteresis X_H.*

Aggregation was observed at 78°C confirming the LCST aggregation properties of this homopolymer. After a 10-minute isotherm at 85°C, the solution was cooled down at a rate of 1°C/min to verify the resolubilization of polymer **4** at lower temperatures. The solubilization was found at the much lower value of 10°C thus revealing the hysteresis phenomenon of the aggregation with a temperature range of more than 60 K (**Figure 9**). The T_{CP} and T_{CL} were determined as the inflection points of the heating and cooling curves, respectively. In addition, the width of the hysteresis, denoted X_H, was calculated as the difference of these two temperatures. Turbidimetric curves of all polymers from **Table 1** resembled the one presented in **Figure 9** (**Table S2**). These results show that PLP with high molecular weight prepared via the water assisted ROP express a large hysteretic LCST.

Effect of concentration

The hysteretic LCST behaviour of PLP solutions was then assessed in different conditions to better understand this behaviour. The effect of concentration was first explored. Several solutions of polymer **4** were prepared with concentrations ranging from 0.5 mg/mL to 5 mg/mL, i.e., from 5.2 to 52 µM in monomer units and analysed by the same technique. The T_{CP} decreases as the concentration of polyproline in solution increases (**Figure 10**). The T_{CL}, for the range of concentration studied, remains approximately constant. This causes the width of the hysteresis to decrease as the concentration increases. A similar trend was observed for all other chain lengths (**Table S2**).

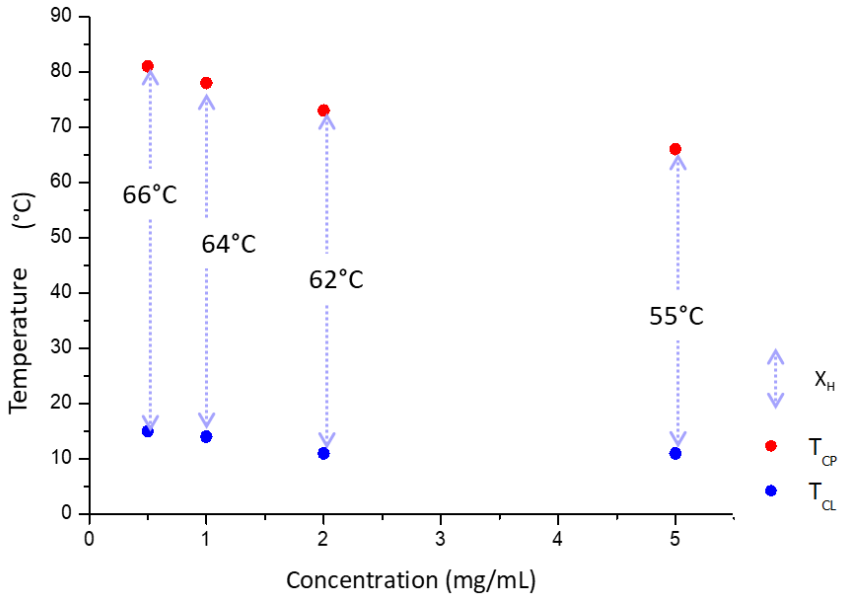


Figure 10 : *Effect of concentration on the T_{CP} and the T_{CL}, X_H obtained for polymer* ***4****.*

Effect of chain length

The effect of chain length on T_{CP}, T_{CL} and X_H was explored next. These analyses indicated that the hysteresis was increased up to an optimum degree of polymerization (DP) of 200 (Polymer **2**: hysteresis width of 35°C, polymer **7**: hysteresis width of 60°C) (**Figure 11**). While T_{CP} was relatively unaffected by polymer length, the observed differences in hysteresis width were due to the higher clearing temperatures that were observed for shorter polymers (Polymer **2**: clearing point temperature of 41°C, **Figure 11**).

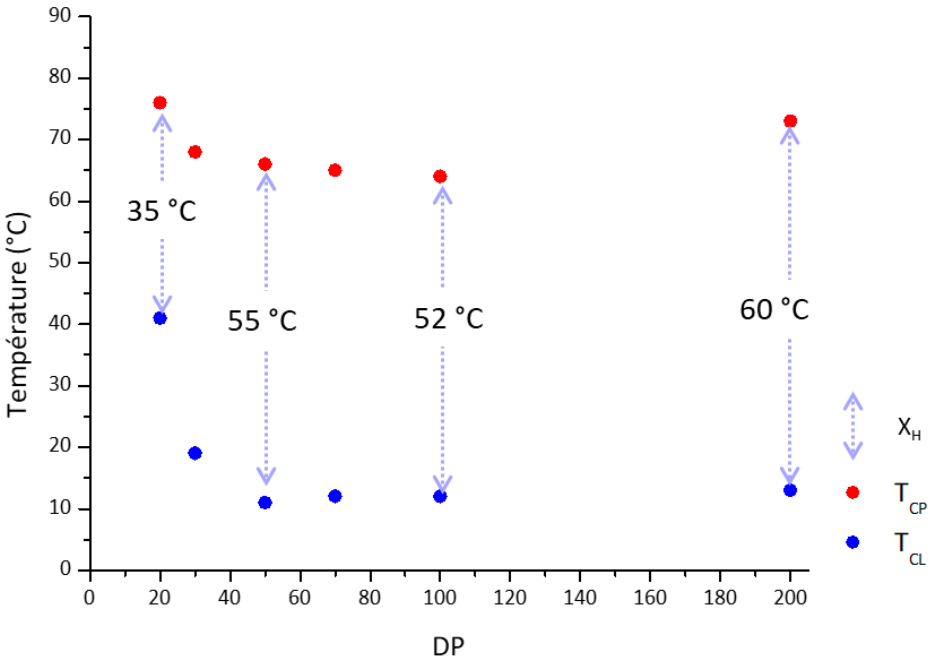


Figure 11 *: Influence of the polyproline length (polymers **2-7**) on the hysteretic LCST behaviour. Shorter chains show smaller hysteresis that seems to be mildly affected by chain length at higher DP.*

Stability of the aggregation and robustness of the hysteresis

To evaluate the robustness of this hysteresis phenomenon, the stability of the aggregation above T_{CL} was evaluated. Polymer **4** was analysed in water at a concentration of 5 mg/mL with differing temperature ramps (**Figure S16**). For all solutions, we first monitored the solution absorbance by UV spectroscopy while increasing the temperature at a rate of 1 °C/min up to a temperature of 85°C, that was then stabilized for 10 minutes (**Figure 12**). Subsequently, each aqueous solution was cooled to a different temperature (between T_{CP} and T_{CL}) and held at that temperature for up to 4 h to determine whether the aggregates were stable over time (**Figure 12**).

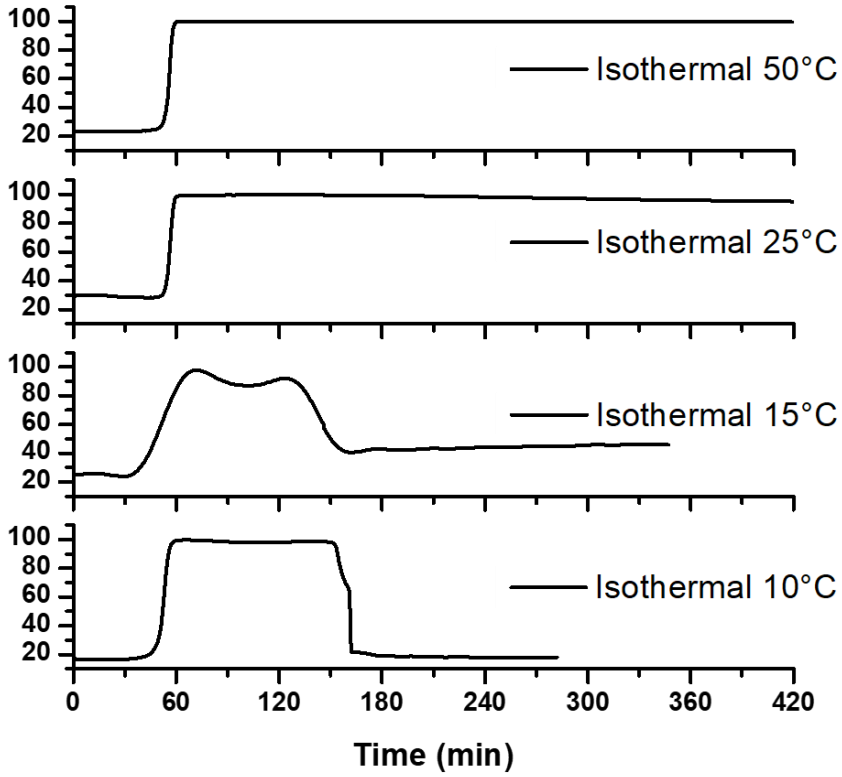


Figure 12: *Stability of the aggregation/hysteresis phenomena of polymer **4** solutions at 5mg/ml at different isothermal temperatures. The samples were heated to 85°C and held for 10 min, during which time aggregation occurred. Subsequently the samples were cooled at 1°C/min to the indicated temperatures and held at these temperatures with stirring for up to 4 hours. Temperature profiles can be consulted in as well as other isothermal experimental results (Figure S16).*

At temperatures of 25°C and above, the turbidity was observed for several hours and even 1 week at 25°C. This shows that the “aggregation memory” was sufficiently long and stable to be studied and/or used in a formulation process. On the other hand, on cooling to temperatures near (15°C) or lower than the T_{CL} (10°C), resolubilization was observed. Similarly, solubilized samples stayed soluble for up to 3 months at room temperature, meaning that the soluble and insoluble states are both thermodynamically and kinetically stable between T_{CP} and T_{CL}. Since the hysteretic thermoresponsive behaviour of some polymers is also highly dependent on the speed of heating and cooling,[150] a solution at 5 mg/ml of polymer **4** was exposed to the same thermal treatment but the rate of heating and cooling was changed to 5 °C/min. Not only was the hysteresis retained, but the heat treatment was also repeated 5 times keeping the exact same properties (**Figure S17**).

Salts effect

The LCST behaviour of thermoresponsive polymers is known to be affected by the ionic strength of the solution.[151] The influence of different concentrations of NaCl on the thermally induced aggregation of polymer **4** at a concentration of 5 mg/mL in water is shown in **Figure 13.** Increasing the ionic strength by adding up to 1 M NaCl significantly reduced T_{CP}. The addition of salt also seemed to reduce T_{CL} slightly, although at higher NaCl concentrations resolubilization happened after an extended amount of time (**Figure S19**). Beyond 1.5M, the PLP was insoluble at room temperature due to the salting-out effect.

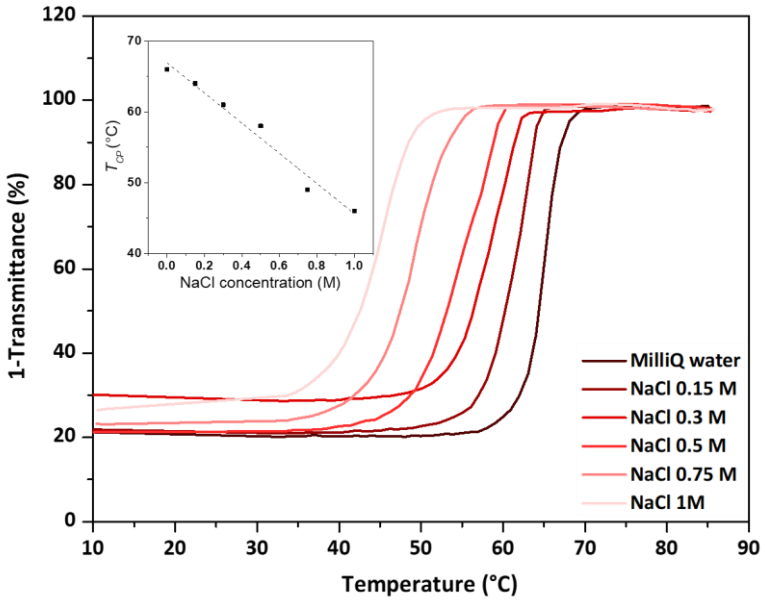


Figure 13: *UV-Vis turbidimetry assays showing the effect of salt concentration on the TCP of different solutions of polymer **4** (5mg/ml). Plotted in the square is the TCP values depending on the concentration of salt.*

These solutions containing different salt concentrations and polymer **4** at a concentration of 5 mg/mL were exposed to heating cycles at different temperatures. After each heating cycle, the temperature was lowered to room temperature and a picture was taken; as shown in **Figure 14**. These heating cycles revealed the memory kept by the polymer in solution of heating at a temperature above T_{CP} with an accuracy of a few °C. This result is remarkable given the simplicity of the system: it may allow cheap temperature monitoring of many systems or processes in real life applications

(pharmaceutical, food processing, etc.) that could be ultimately reprogrammed under the T_{CL}.

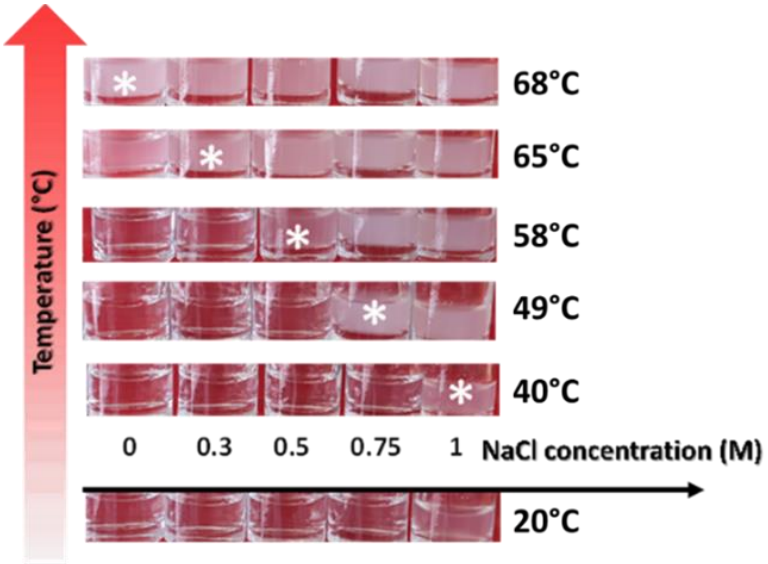


Figure 14 : *temperature sensor model using different solutions of polymer* ***4*** *5mg/ml varying only the NaCl concentration. All samples were heated to the temperatures indicated, stabilized for 5 minutes and a picture of the whole series taken at room temperature. The stars highlight the working range of this simple temperature sensor.*

To add to the versatility of this system other salts were used to expand the range and accuracy of the temperature probing. Replacing Na^+ with the more kosmotropic cation Ca^+, the temperature sensor model was not viable as the polymer solutions pre-aggregated at low temperatures (55°C) before expressing the LCST transition at their corresponding T_{CP} (**Figure S20**). When aqueous solutions of $(NH_4)_2SO_4$ were added, a similar model was developed at much lower concentration, allowing a broader range of usable temperatures (**Figure 15**).

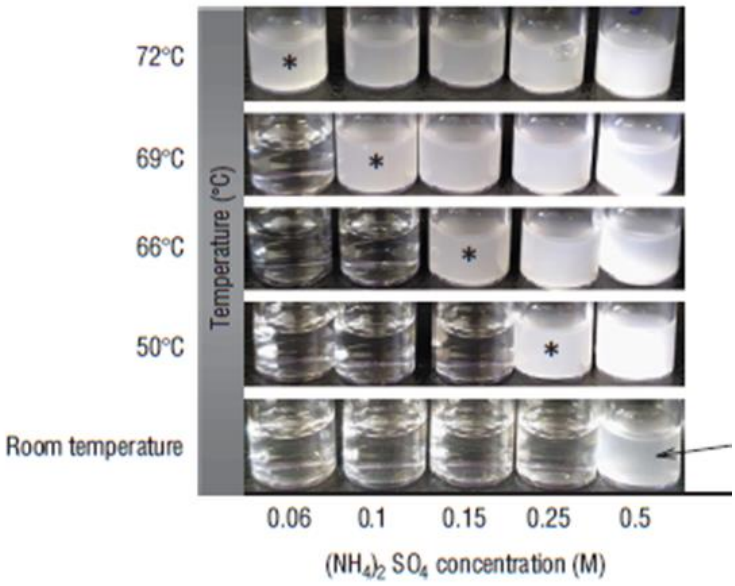


Figure 15: *temperature sensor model using different solutions of polymer **4** 5mg/ml varying only the $(NH_4)_2SO_4$ concentration. All samples were heated to the temperatures indicated, stabilized for 5 minutes and a picture of the whole series taken at room temperature. The stars highlight the working range of this simple temperature sensor.*

These preliminary results show further possibilities in preparing simple temperature sensors based on polyproline saline solutions. Further ions from the Hoffmeister series should be explored to understand the overall mechanism and expand the possibilities of applications.[152]

2.4. Secondary structuring and thermoresponsive behaviour

As presented in the introductory part of this chapter, the equilibrium between the two secondary structures of polyproline is influenced by the temperature. We thus emitted the hypothesis that the hysteresis phenomenon had to be related to secondary structure changes between PLPII and PLPI driven by the aqueous solvent. Before heating, the soluble state in aqueous solution is PLPII. Upon heating, sufficient energy would be given to the bonds to change the conformation of the amide bonds from trans to cis, resulting in the formation of PLPI aggregates. A new polymer, polymer **S1** (HA-PLP20), was synthesized in dry ACN to obtain a sample in PLPI conformation to be compared with the other polymers (Supplementary information for synthesis and analysis).

Polymer **1**, with a DP equal to 5, did not show thermoresponsive behaviour. This may be because the concentration of HA-PLP_5 was too low, or the chains were too short to form an insoluble secondary structure. This may indicate that the secondary structure, achievable with higher molecular weights, is crucial in this phenomenon. In the next part, the thermoresponsive behaviour of PLP and its correlation with the conformation will be presented using FTIR spectroscopy and circular dichroism.

2.4.1. FTIR spectroscopy

Secondary structuring can be determined by FTIR spectroscopy. This technique can detect the amide I and II bands corresponding to CO stretching and N-H bending/C-N stretching of amide groups respectively.[153] Polymer **4** was thus analysed by FTIR at 5mg/ml in milliQ water at room temperature (PLPII) and after heating to aggregation.

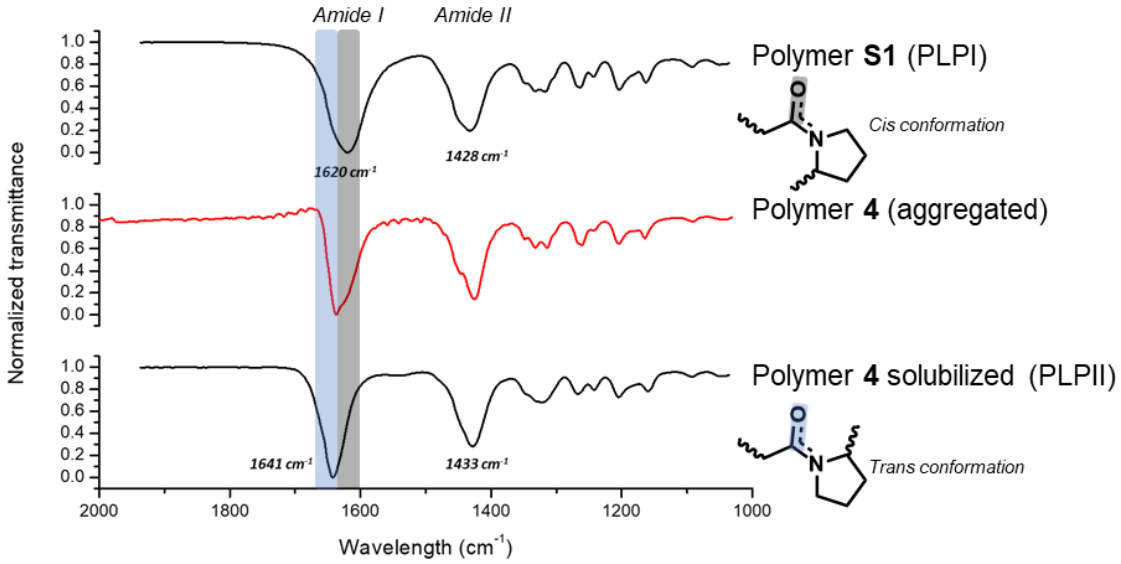


Figure 16 : *FTIR spectra of polymer solutions in H_2O. polymer **S1** hexyl-PLP20 (polymerized in dry ACN) is in PLPI conformation and polymer **4** hexyl-PLP50 in PLPII conformation. The amide I bond stretching wavelengths of the cis and trans conformation (deduced from the two first spectra respectively) are highlighted. Polymer **4** was then precipitated by heating and brought to room temperature (spectra is shown in red).*

The spectra were compared to those of polymer **S1** (PLPI) (**Figure 16**). The amide I bands for the cis and trans conformations are at 1620 cm^{-1} and 1641 cm^{-1} respectively, while the amide II bands are closer together (1428 cm^{-1} and 1433 cm^{-1} respectively) (**Figure 16**).[154] After heating-induced precipitation, polymer **4**, originally in the PLPII conformation, gave stretching bands in FTIR corresponding to both the cis and trans

conformations for amide I (**Figure 16**). For amide II a double peak was also observed. These results suggest that when PLPII polymers are heated, some but not all trans amide bonds are converted to cis amide bonds causing the formation of an insoluble conformation and collapse of the soluble helices followed by precipitation from solution.

2.4.2. CD spectroscopy

To verify that the change in amide conformation (observed by FTIR) is accompanied by a change in secondary structure, CD spectroscopy was employed. As mentioned in the introduction, each conformation of polyproline has a characteristic curve due to its compactness and helix orientation.

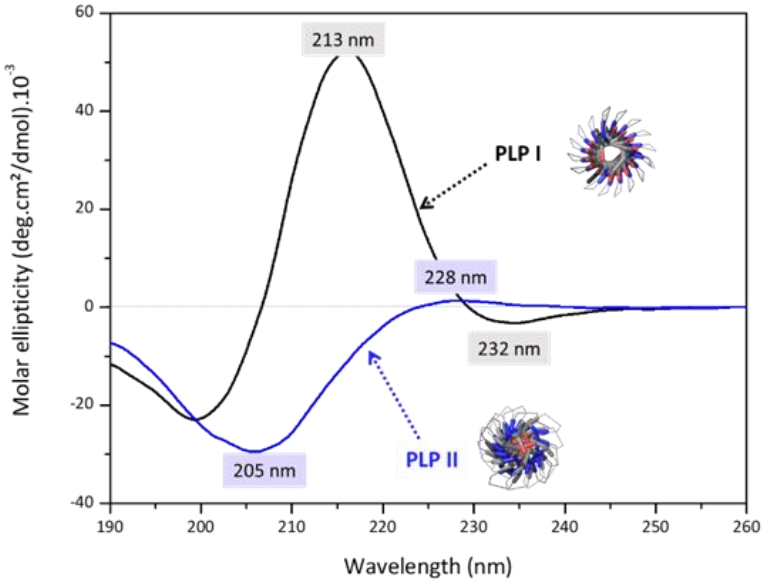


Figure 17: *CD spectra of PLPI, polymer **S1** and its modelized structure (top in black) and PLPII, polymer **4** and its modelized structure (bottom blue). the characteristic peaks are presented.*

Since most CD measurements are recorded at low concentrations where thermoresponsive behaviour is not observed, polymer concentrations in this study were kept higher (at 0.25mg/ml) to observe both the secondary structures and the thermoresponsiveness. The CD spectra of both conformations was obtained experimentally on samples of polymer **S1** and polymer **4**. The shape of the obtained spectra (**Figure 17**) is in good agreement with the literature.[137] For polymer **S1**, the PLPI helix is recognized by its strong maximum (as it is right-handed) at 213nm and a weak negative peak at 232nm. As for PLP II conformation (polymer **4**), a minimum at 205 nm and a weak positive peak at 228nm can be observed.

Polymer **4** in its PLPII conformation when heated, evolved slightly with a decrease of the intensity of the peak and its slight red shift (**Figure 18**). This was previously documented in the literature and suggests a loosening of the helical structure.[141] Yet after heating to above 80°C (>T_{CP}), when precipitation occurred, it changed abruptly to a new CD signal that corresponded to neither PLPII nor PLPI conformation. The strong minimum at 205 nm characteristic of PLP II at room temperature was shifted to a minimum at 219 nm (**Figure 18**).

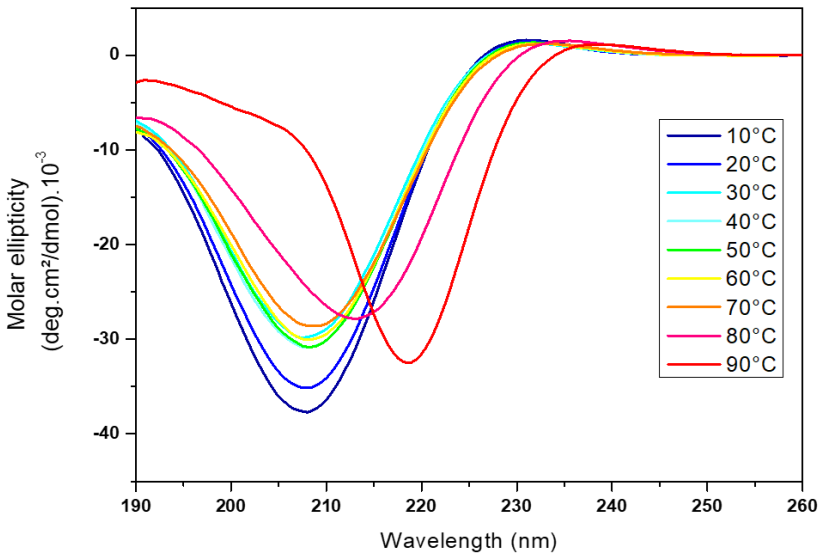


Figure 18: *CD spectra of polymer **4** at 0.25mg/ml at different temperatures. The polymer was left to stabilize for 5 min between heating ramps.*

Figure 19 shows the molar ellipticity values of both the PLPII peak at 208nm and the aggregated state peak at 219nm during the heating/cooling cycle. The new conformation was stable when the sample was cooled back down to room temperature as the peak values remained the same. When the sample was cooled to below 20°C, nearing the T_{CL}, the values of both peaks shifted to their initial values around 10°C, regaining the PLPII conformation.

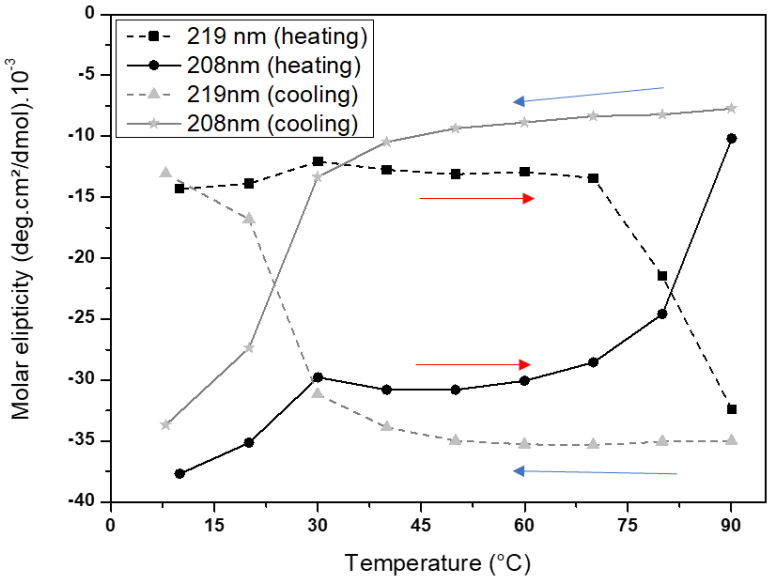


Figure 19: *Molar ellipticity of peaks at 208nm (PLPII) and 219nm (aggregated state) in function of temperature during heating (red arrows) and cooling (blue arrows).*

These results corroborate earlier results obtained by FTIR and past hypothesis of the formation of a new, intermediate, conformation, between PLPI and PLPII. Both analyses support an incomplete transition to PLPI: indeed, it has been shown that PLPI requires several days to be converted to PLPII in water at 4°C and not instantaneously upon cooling as seen in these results. This also suggests that the new conformation contains some cis amide bonds that can be readily isomerized to trans, solubilizing the polymer in its PLPII state. In a scope of a collaboration of the DISCO laboratory from SOLEIL, it was possible to obtain similar results at 5 mg/ml (**Figure S18**). This was possible thanks to CaF_2 cells with smaller path length (27µm) and a synchrotron radiation CD setup.

2.4.3. Chirality, secondary structuring and thermoresponsivness

A new series of copolymers was then synthesized based on a mixture of both Pro NCA enantiomers. Copolymerizing *L*-Pro NCA with *D*-Pro NCA was used to disrupt the long-distance secondary structuring of the polymers. Copolymers **3-7** having increasing *D*-Pro content (from 10-50%), as shown in **Table S2**, were synthesized and analysed as described above. All copolymers had an M/I of 100 and were initiated using hexylamine. The analysis showed DP close to each other, yet molecular weights

calculated using the dn/dC of homoPLP only gave lower than expected values (**Table S2**). This could be due to a decrease in hydrodynamic volume of the polymer chains due to a change in their secondary structure.

While polymer **11** HA-PDP$_{100}$ showed activity comparable to that of HA-PLP$_{100}$ (polymer **6**) (**Table S2**), statistical copolymers **3-7** of *L*- and *D*-proline showed no thermoresponsive behaviour. **Figure 20** shows the CD spectra of polymers **6** and **11,** homopolymers of *L*- and *D*-proline respectively.

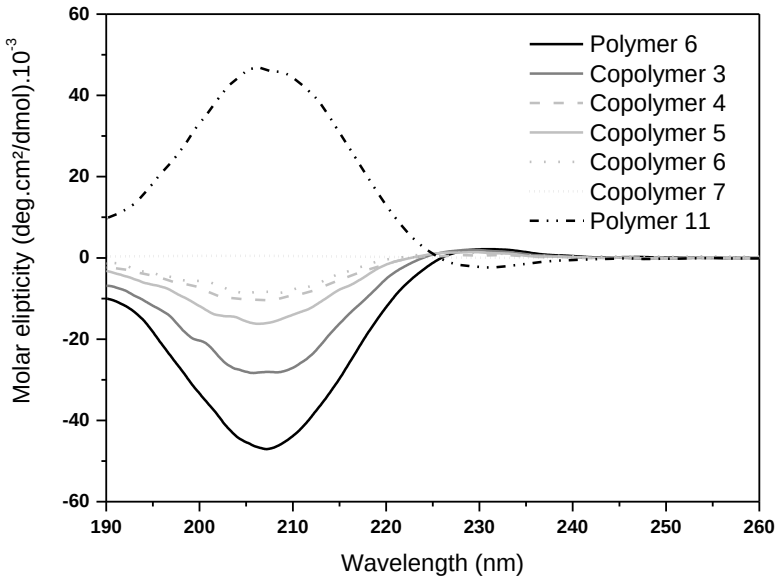


Figure 20: *CD spectra of copolymers **3-7** compared to polymer **6** and **11** (homopolyprolines).*

The mirror image PPII structure can be easily identified meaning that both these polymers retain the same helical conformation but with an inverse orientation. The statistical copolymers **3-7** also retained their PLPII conformation as can be assigned by the value of the peak at 208nm. The signal of the peak decreased with the increasing *D* content as the enantiomers' signal cancel each other out; for copolymer **7** (HA-PLP50-PDP50) where no signal can be detected.

NMR analysis showed increase of cis conformation with increasing content of the *D*-proline (**Figure S8**).[155] These peaks were not present for polymer **11** that had a spectra similar to that of homoPLP (**Figure S5**). These results suggest that both the secondary structure and the cis-trans bond conformation play a role in the phase transition.

As a control, a series of mixtures of polymers **6** and **11,** homopolymers of *L*- and *D*-proline respectively, were prepared at a total concentration of 5mg/ml. The different ratios were chosen to correspond to the ratios of enantiomers in copolymers **3-7** (**Table 2**).

Table 2 : *Turbidimetric results for different mixtures of polymer **6** and **11** at 5mg/ml.*

Weight ratio PLP/PDP	T_{CP} (°C)	T_{CL} (°C)
100/0	64	12
90/10	57	12
80/20	54	11
70/30	53	11
60/40	52	11
50/50	53	11
0/100	64	13

All mixtures retained hysteretic LCST behaviour confirming that both the overall secondary structure and the trans-cis interplay are important for the thermoresponsive behaviour. T_{CP} values were lower than pure solutions suggesting a cooperative aggregation mechanism when enantiomeric helices are mixed.

2.5. Thermoresponsive materials based on PLP

Block copolymers based on PEG and PLP (copolymers **1**-**2** in **Table 1**) could prove interesting to design materials with hysteretic thermoresponsive behaviour. These copolymers showed thermoresponsive behaviour with a high T_{CP} due to the increased solubility of PEG (**Figure S21Figure S22**). A considerable amount of salt was added to a concentrated PLP solution (10mg/ml), to achieve a concentration of NaCl of 3M, as to mimic heating above the T_{CP}. Under these conditions the block copolymers thus formed nanoparticles that were analysed by DLS (**Figure 21**).

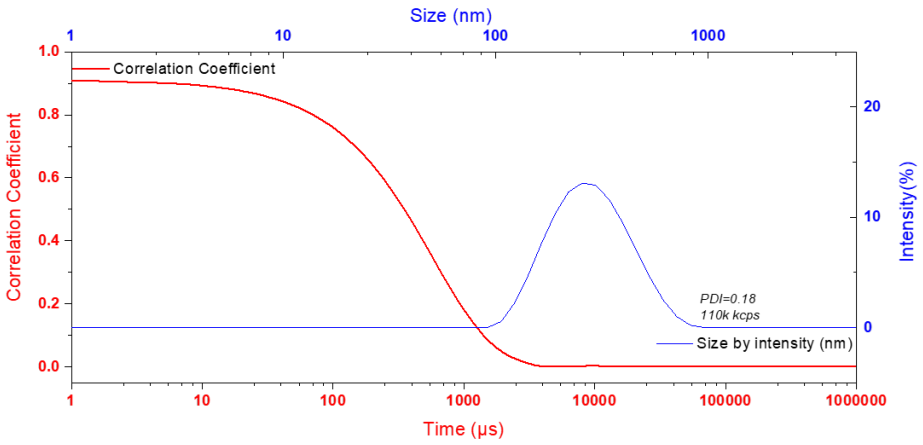


Figure 21 : *DLS data of copolymer **1** at 10mg/ml in 3M NaCl solution at room temperature. The size by intensity as well as the correlation factor are shown.*

These nanoparticles had an average diameter of 128nm and a monomodal size distribution with a dispersity of 0.18. Further imaging is needed to ascertain the shape of the nanoparticles as the models used by the DLS software are optimized for spherical particles. When heated the copolymers showed a T_{CP} around 50°C and precipitated out of solution. The precipitates were stable on cooling to room temperature, but dispersed nanoparticles were reformed when the solution was cooled to 5°C (**Figure S23**). The broad hysteresis indicates that the thermoresponsive behaviour is due to the polyproline block. As polyproline tends to precipitate in NaCl solutions of 1.5M or higher, it is hypothesized that the proline blocks nano-aggregate together and are surrounded by the soluble PEG shell. When heated the bundled proline chains still expressed thermoresponsiveness meaning that the salt did not affect the conformation. PEG solutions did show LCST behaviour, in these same conditions, that was not hysteretic with a T_{CP} and T_{CL} around 75°C (**Figure S24**). This confirms that the behaviour of these nanoparticles stems from the PLP block.

Since materials with shape memory are also of great interest, hydrogels based on these copolymers were also explored. Only with high polymer concentration (100mg/ml) and addition of NaCl (2M) were thermogelling solutions obtained (**Figure 22**).

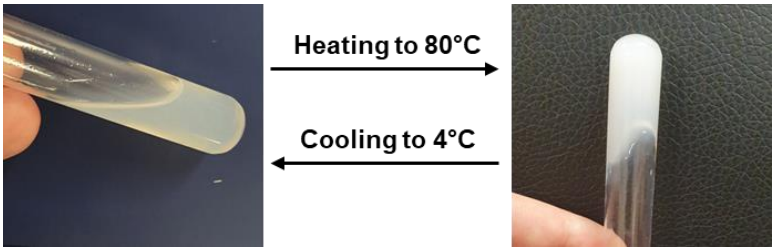


Figure 22 : *thermogelling solution of copolymer* ***2*** *and its reversible thermoresponsive behaviour.*

Such programmable materials are of interest as the forward and reverse transitions could be triggered at different temperatures. These promising results will be followed by further analysis of the gels' stability, mechanical properties and other macromolecular engineered copolymers based on PLP will be explored.

3. Conclusions

Hysteretic temperature response could be a useful feature in designing or programming phase separation driven assemblies (coacervation, etc.). Most known LCST polymers exhibit small thermal hysteresis in their phase separation behaviour. It was shown that fully peptidic polymers based on proline prepared via aqueous NCA ROP with more than 20 repeating units displayed hysteretic LCST. Polymers synthesized via this pathway showed controlled molecular weights and narrow dispersities while other ROP techniques yielded poorly controlled polymers. The unprecedentedly wide hysteresis obtained in the thermoresponsiveness of PLP was only slightly affected by concentration and molecular weight. This study proved both kinetic and thermodynamic stability of both soluble and aggregated polymers. This stability was attributed to the conformation of the peptide bonds and the secondary structure of the polypeptides. CD measurements proved the shift from the highly soluble PLPII conformation to an intermediate insoluble conformation. Further analyses via FTIR proved the presence of both cis and trans amide conformations providing evidence that temperature-induced changes in the secondary structure were responsible for the reversibly hysteretic LCST.[156] Functional materials based on the PLP scaffold have been briefly investigated including programmable nanoparticles and thermogelling solutions. The results reported in this chapter have already promoted development of other proline-based thermoresponsive polypeptides.[157,158] PLPs also showed interesting properties as antifreeze polymers as they can inhibit ice

recrystallization at lower temperatures[159] allowing for use in cryopreservation of cells[160]. This property can be attributed to both their peculiar secondary structure and to their thermal hysteresis behaviour.[161]

4. Supporting information

4.1. Materials and methods

4.1.1. Materials

All chemicals and solvents in this work were purchased from Sigma Aldrich, Fluorochem, Acros, TCI, Strem and VWR, and were used without any purification unless otherwise described. Acetonitrile (ACN), was obtained from a solvent system purifier (PureSolv, Innovative Technology), kept under argon atmosphere and freshly used. Milli-Q water was obtained from a Purelab Prima purification system (ELGA) with a resistivity of 18.2 MΩ cm-1. The N-carboxyanhydride monomers were purchased from PMC Isochem, stored at 20 °C under argon atmosphere and weighed in the glove box Jacomex GP13 no. 2675 at the Laboratoire de Chimie des Polymères Organiques (LCPO, Bordeaux, France). Dialysis membranes were purchased from SpectrumLabs. Hexylamine, benzylamine, allylamine and 2-(2-aminoethoxy)ethanol were all freeze-thawed and cryo-distilled on the Schlenk line prior to use.

4.1.2. Ring opening polymerization (ROP) of N-carboxyanhydrides (NCA)

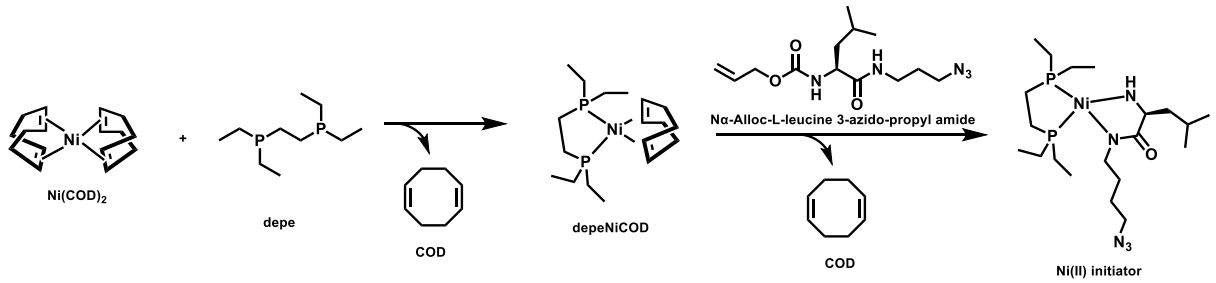
Typical polymerization was adapted from a recent study[125] and goes as follows: *L*-Pro-NCA (100 mg, 71*10-2 mmol, 50 eq) was introduced in a flame-dried Schlenk equipped with a stirrer and suspended in 0.5 ml of dry ACN. Hexylamine (1.43 mg, 1.4*10-2 mmol, 1eq) dissolved in 1.5 ml H2O: ACN (1:1) at 10°C was added to the first solution to start the ROP under vigorous stirring. The solution quickly became clear and bubbling was observed. The disappearance of the bands at 1859 cm-1 and 1787 cm-1 characteristic of the C=O stretching of NCAs was monitored by FTIR and confirmed the reaction completion after 10 minutes. The reaction mixture was dialyzed against Milli-Q water in a 100-500 Da dialysis membrane to remove all unreacted monomers and solvent before being lyophilized. Yield: 90%, white powder. For polymers prepared in dry acetonitrile, the NCA was dissolved in ACN at 0.1 M and hexylamine was added to start the reaction under inert atmosphere, either under Schlenk line or in glovebox, the solution became milky. Reaction was monitored by FTIR, and at full conversion,

acetonitrile was evaporated, and polymers were used without further purification, Yield: 67%, white powder.

For polymer **S1** prepared in dry organic solvents, the NCA was dispersed in ACN at 0.1M and hexylamine was added to start the reaction under inert atmosphere using a Schlenk line. The solution became milky over time. At the end of the reaction acetonitrile was evaporated under vacuum. The white powders recovered with a yield about 60% were used as such without further purification.

For polymers **13-14** prepared in dry DMSO, the NCA was dissolved in DMSO at 0.1M and the initiator was added as a concentrated solution of DMSO to start the reaction under inert atmosphere using a Schlenk line. The solution became milky over time. At the end of the reaction the polymers were dialyzed against DMSO and lyophilized. The white powders recovered with a yield about 80% and 91% respectively.

4.1.3. Preparation of Ni(II) amido amidate initiator



Scheme 9 : *Preparation of nickel based amido amidate initiator species*

$Ni(COD)_2$ (20 mg, 0.073 mmol, 1.0 eq) was suspended in 555 µL THF in a pressure vessel. 1,2-Bis(diethylphosphino)ethane (depe) (37.2 µL, 0.159 mmol, 2.2 eq) was added to the bright yellow $Ni(COD)_2$ solution followed by stirring for 10 minutes. The solution turned orange with the formation of depeNi(COD). Nα-Alloc-L-leucine 3-azido-propyl amide (23.8 mg, 0.08 mmol, 1.1 eq) was added as a solution in DMF (1.1 ml) (**Scheme 9**). The pressure vessel was sealed, removed from the glovebox, and placed in an oil bath at 80°C overnight. The orange coloured amido-amidate nickelacyle solution (44mM) was used for the polymerization of polymer **12**.

4.1.4. Nuclear Magnetic Resonance (NMR)

^1H NMR 400 MHz spectra were obtained using a Bruker Avance I (Liquid-state 400 MHz NMR spectrometer with 5 mm BBFO probe). Deuterated water (D_2O, Euriso-top, 99.8%) was used as solvent and reference for the lock. The spectra obtained were

calibrated using the residual solvent signals (H2O 4.79 ppm). The signals were categorized as follows: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m) and broad (br).

4.1.5. Size Exclusion Chromatography (SEC) Analyses

Polymer molar masses were determined by Size Exclusion Chromatography (SEC) using an aqueous solvent as the eluent. Measurements were performed on an Ultimate 3000 system from Thermoscientific equipped with diode array detector DAD. The system also includes a multi-angle light scattering detector (MALS) and differential refractive index detector (dRI) from Wyatt technology. Polymers were separated on two TOSOH TSK Gel G4000PWXL and G3000PWXL columns (300 × 7.8 mm) (exclusion limits from 200 Da to 300 000 Da) at a flowrate of 0.6 mL/min. Aqueous solvent composed of acetic acid (AcOH) 0.3 M, ammonium acetate 0.2 M and ACN (H2O/ACN: 6.5/3.5, v/v) was used as the eluent. The column temperature was held at 25°C. Measurements of dn/dc were performed on a differential refractive index detector dRI from Wyatt technology by injecting 500 µL of each sample dissolved in the aqueous phosphate buffer at 0.5-5 mg mL−1. The chromatograms were recorded with Chromeleon 7.2 software and and analyzed using Astra 7.1.0 software. Calculated dn/dC values were 0.1616 mg/ml for L-Proline homopolymers and copolymers with D-Proline and 0.1444 mg/ml for block copolymers with PEG.

4.1.6. UV−Vis Spectroscopy

The cloud point temperature (T_{CP}) and the clearing point temperature (T_{CL}) of PLP solutions in pure water were determined by measuring the turbidity at 550 nm between 10°C and 85°C at a 1 °C.min^{-1} scan rate at different concentrations of polymer and salt. Data were collected on a Cary 100 UV−Vis spectrophotometer equipped with a multicell thermoelectric temperature controller from Agilent Technologies. T_{CP} and T_{CL} were defined as the temperatures at the inflection point of the transmittance-temperature curves during heating ramps (T_{CP}) or cooling ramps (T_{CL}). A typical experiment followed the following heating program:

1- 10 min isotherm at 10°C

2- 1°C/min ramp to 85°C

3- 10 min isotherm at 85°C

4- 1°C/min ramp to 10°C

Absorbance was plotted as a function of temperature, with isotherms' data not collected. Subsequently, absorbance was converted to transmittance for a better visualization of the degree of turbidity and to discriminate between complete and incomplete aggregations. The y-axis in the results is displayed as 1-Transmittance (%) to help the reader relate the soluble and aggregated forms to low and high values respectively.

4.1.7. FTIR spectroscopy

The IR spectra were recorded using the FTIR spectrometer (Vertex 70, Bruker) equipped with a temperature control cell, and the samples were measured with the ATR (GladiATR, Pike Technologies) from Fisher technologies performing 64 scans with a resolution of 2 cm^{-1} in solution at the LCPO (Bordeaux, France). The plate and pressure system of the ATR were equipped with a liquid sample set (Pike Technologies) bought from eurolabo to keep the solvent from evaporating during the heat treatment and thus analyse the sample in the liquid state as other experiments. The raw data were obtained with the Opus7.5 software and processed using the Originlab 2016 software. Samples were dissolved in water and analysed as such on the ATR plate. A background with pure water was run before each sample in the same manner so as not to mask the vibration bands of the C=O bonds between 1500 cm^{-1} and 1700 cm^{-1}.

4.1.8. MALDI-MS measurements

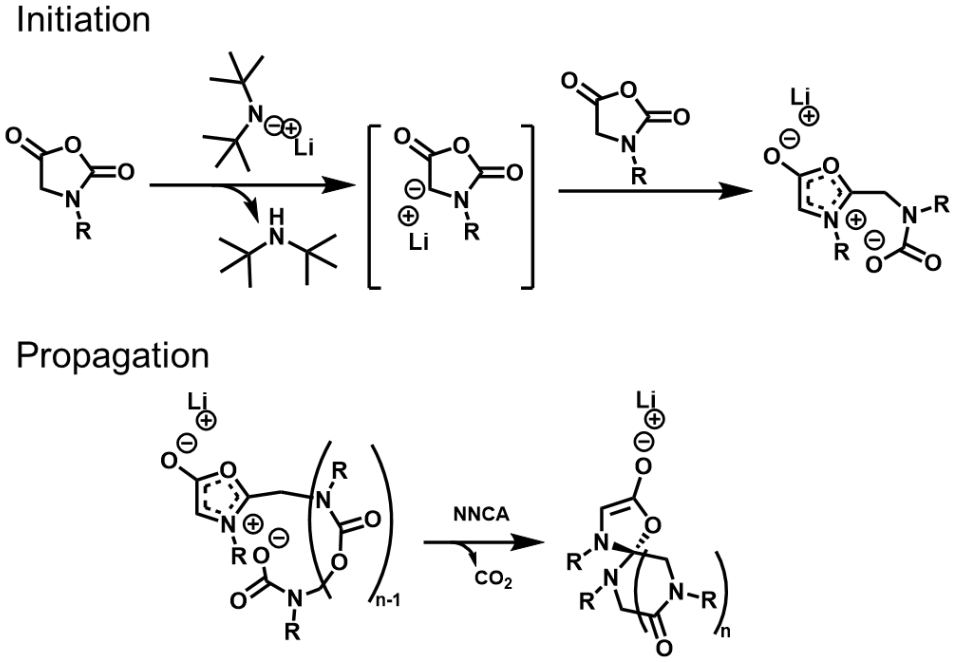
MALDI-MS spectra were performed at the CESAMO facility (Bordeaux, France) on an Autoflex maX TOF mass spectrometer equipped with a frequency tripled Nd:YAG laser emitting at 355 nm. Spectra were recorded in the positive-ion mode with an accelerating voltage of 19 kV. For MALDI-MS analyses, polymers' deposits were prepared by dissolving 10 mg of polymer in 1 ml of Milli-Q water, 10 mg of NaI in 1 ml methanol and 22 mg of sinapinic acid in methanol. Subsequently 2 µL of polymer solution was mixed with 2 µL of salt solution and added to 20 µL of matrix solution. 1 µL of this mixture was deposited on the MALDI plate and dried.

4.1.9. CD spectroscopy

CD measurements were performed on a JASCO J-815 spectropolarimeter between 190 nm and 260 nm (far-UV). A quartz cell of 1 mm path length (type: 21/10/Q/1) was purchased from Starna Scientific, Ltd. Spectra were recorded at desired temperatures:

20 °C for standard measurements, or a temperature gradient between 10 °C and 90 °C. The measurement parameters were optimized as follows: sensitivity between 5 and 200 mdeg, 0.01 mdeg resolution, 8 seconds response time (Digital Integration Time), 1 nm bandwidth and 10 nm.min^{-1} scanning rate. Polymer solutions at a monomer unit concentration of 2.5 µM in Milli-Q water were used for the measurements. For the turbidity tests, the concentration of the samples had to be adapted for this analysis to a concentration of 0.25 mg/mL to avoid flocculation of the polymers. Thus, an optimum concentration was found that neither saturated the signal nor brought the cloud point temperature above reaction conditions.

4.2. Further synthesis



Scheme 10: *Mechanism of the initiation and propagation steps of the REP of NNCAs by LiHMDS.*

Table S1 : *Synthesis and analysis for copolymers* ***3-8***

Copolymer		Theoretical DP[a]	DP[b]	Mn (kDa)[c]	Đ[d]	Yield (%)
3	HA-PLP$_{90}$-PDP$_{10}$	100	150	6.1	1.18	81
4	HA-PLP$_{80}$-PDP$_{20}$		95	6.2	1.14	73
5	HA-PLP$_{70}$-PDP$_{30}$		106	6.0	1.14	85
6	HA-PLP$_{60}$-PDP$_{40}$		130	6.6	1.13	92
7	HA-PLP$_{50}$-PDP$_{50}$		160	6.9	1.17	51

a. determined by the initial feed ratio M/I, b. determined by the integration of the different backbone peaks in comparison to the initiator peaks in D_2O ^1H NMR, c. determined by dn/dc of polyproline (0.1616) in ACN/acetate buffer SEC and MALS detection, d. determined by SEC peaks.

4.3. NMR

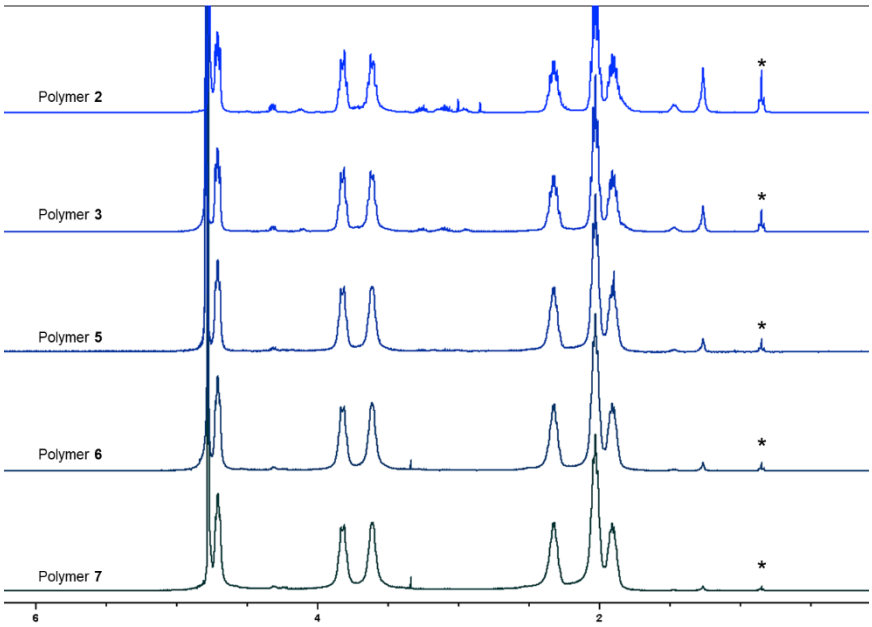


Figure S1 : *^{1}H NMR spectra in D_2O of polymer **2-7** with increasing M/I ratio. Backbone peaks intensity has been normalized for all spectra to show the decrease of the initiator peaks intensity (shown with an *).*

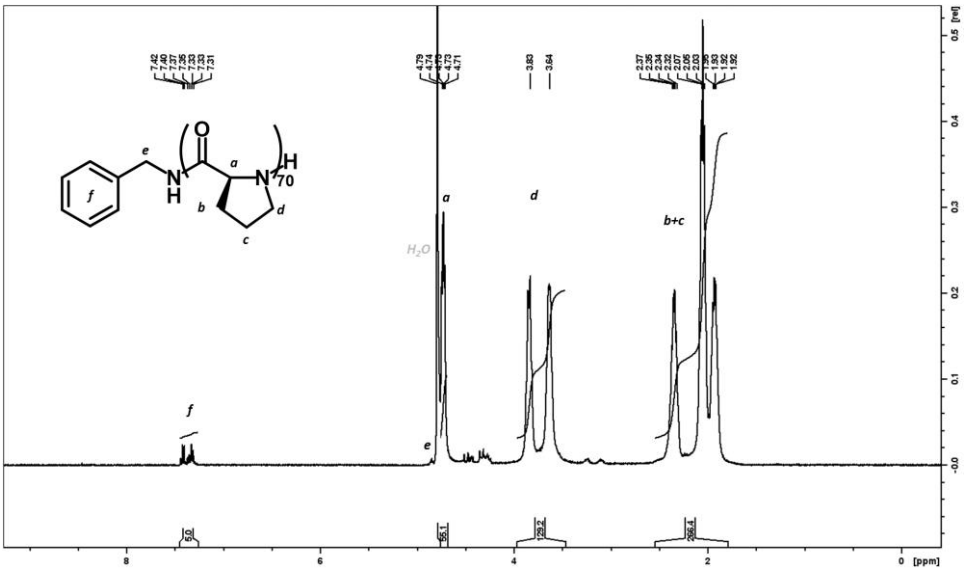


Figure S2 : *^{1}H NMR for polymer **8** in D_2O. The polymer structure of is represented with the peak assignation. The calculation of the Dp is shown by the calculations from the integration of the initiator peaks (e and f)) and the respective backbone peaks of proline repeat units (b,c and d).*

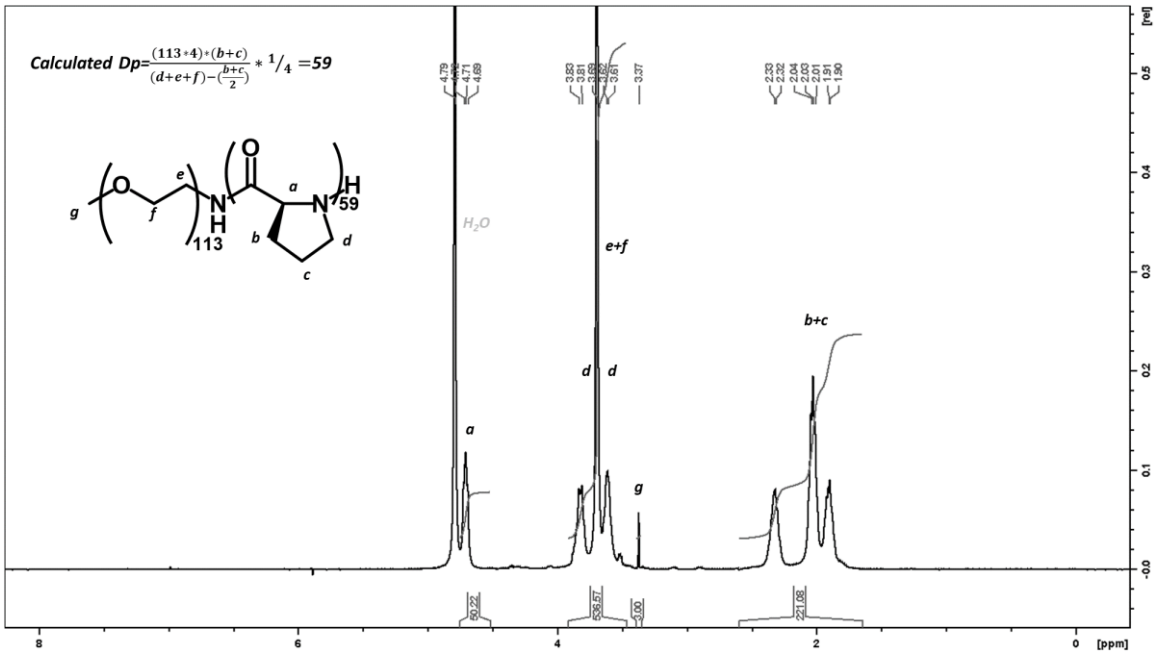


Figure S3: *^{1}H NMR for copolymer **1** in D_2O. The polymer structure is represented with the peak assignation. The calculation of the DP is represented.*

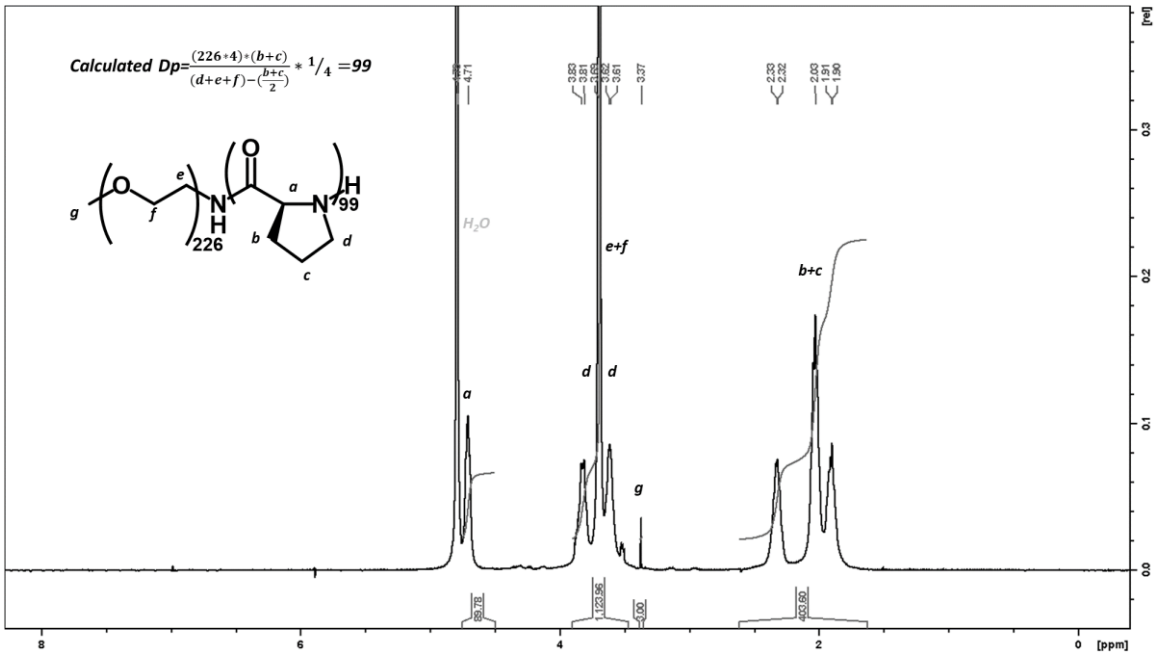


Figure S4 : *^{1}H NMR for copolymer **2** in D_2O. The polymer structure of is represented with the peak assignation. The calculation of the Dp is represented.*

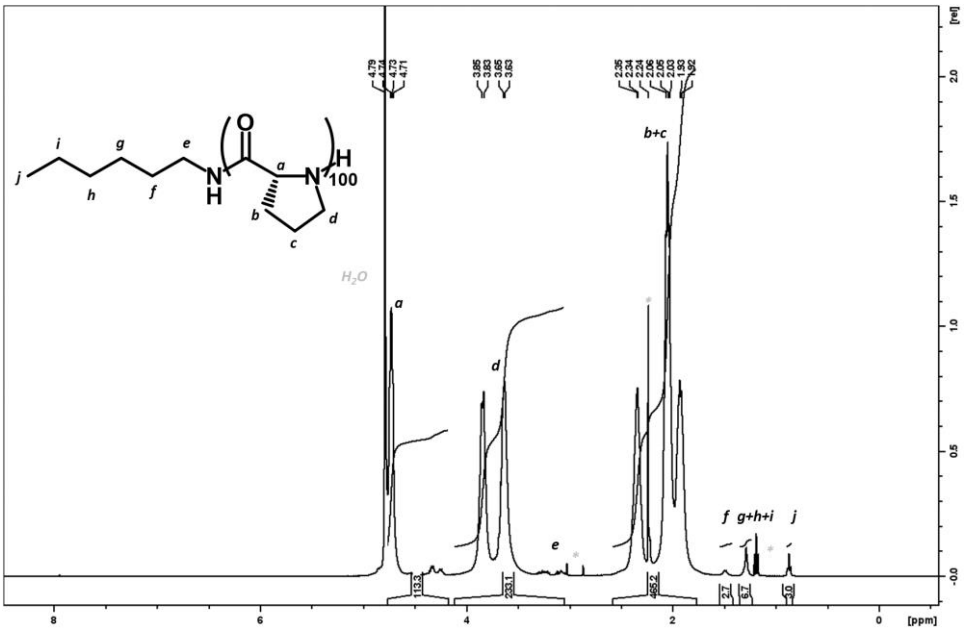


Figure S5 : *^{1}H NMR of polymer **11** HA-PDP$_{100}$ in D_2O and polymer scheme. Letters indicate the attribution of peaks to protons. Peaks labelled with * correspond to impurities in the NMR sample.*

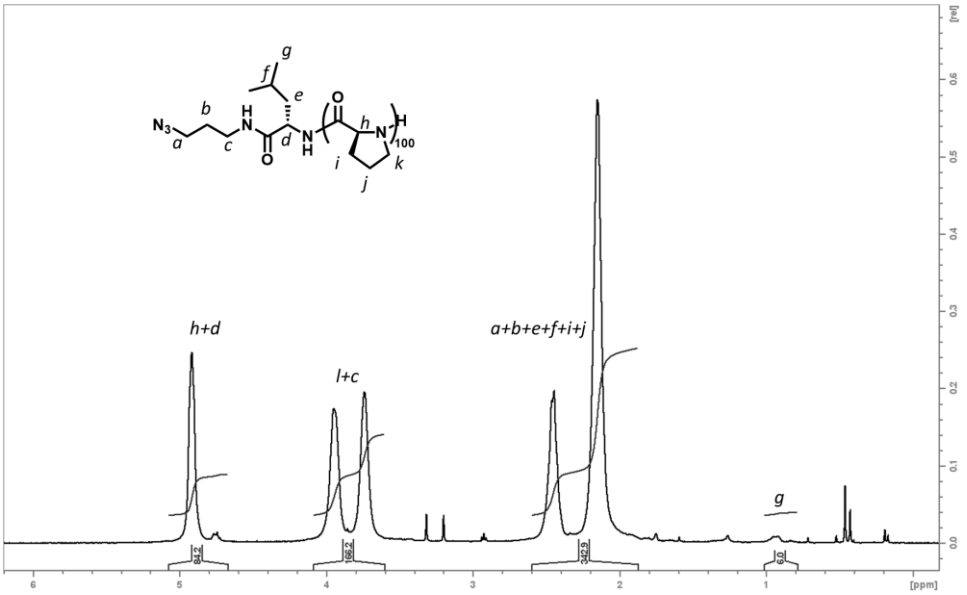


Figure S6 : *^{1}H NMR of polymer **12** in TFA-d and polymer scheme. Letters indicate the attribution of peaks to protons. Some impurities in the NMR sample can be detected.*

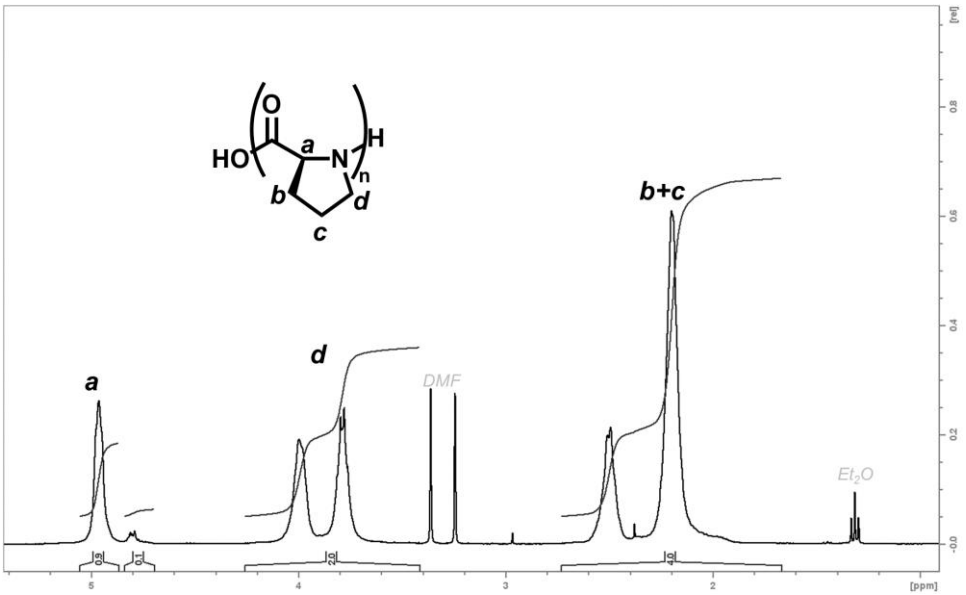


Figure S7 *: ^{1}H NMR of polymer **15** in TFA-d and polymer scheme. Letters indicate the attribution of peaks to protons. Some impurities in the NMR sample can be detected.*

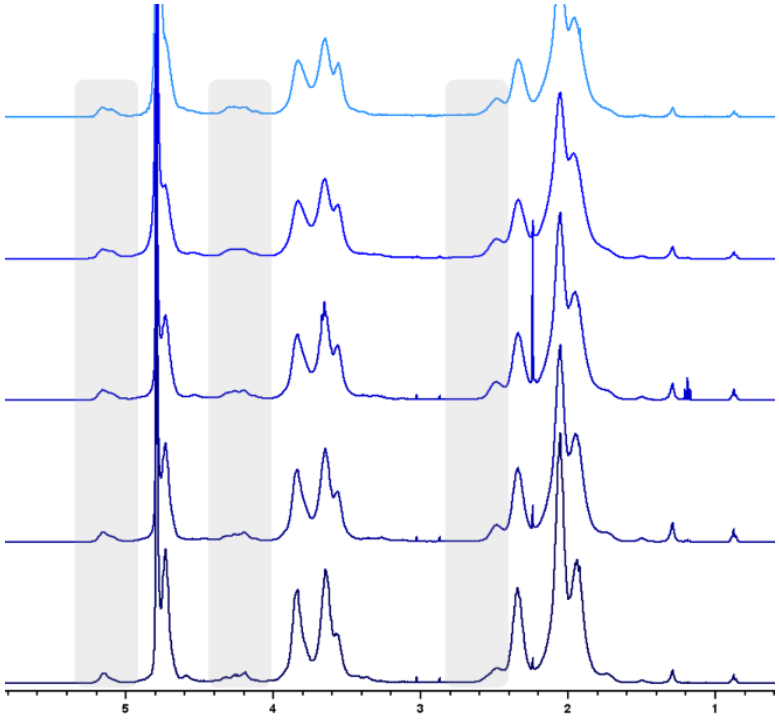


Figure S8 : *^{1}H NMR of copolymers **3-7** (bottom to top) in D_2O. Spectra showed all PLP peaks with some broadening and peaks different from the homopolymer were distinguished (highlighted in grey) and could be attributed to the cis conformation.*

4.4. SEC

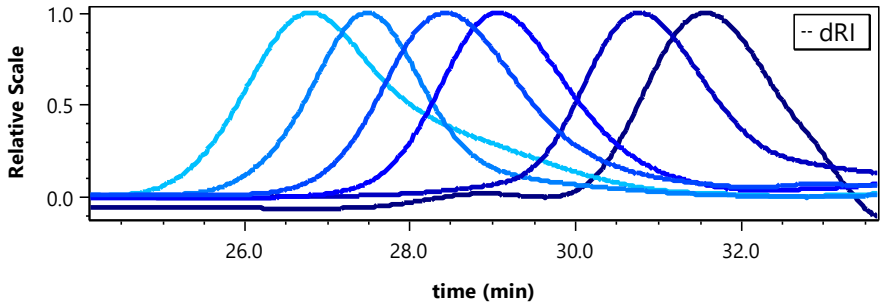


Figure S9 : *Aqueous SEC chromatograms of polymers **1,2,3,4,6** and **7**. RI detection is shown.*

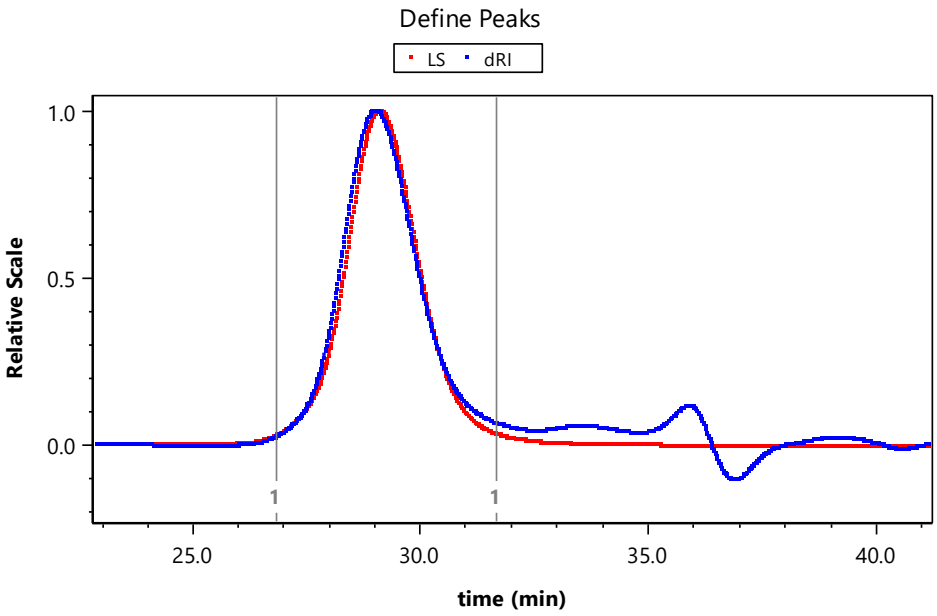


Figure S10 : *SEC chromatogram for polymer **8**. Both MALS and RI detection are shown.*

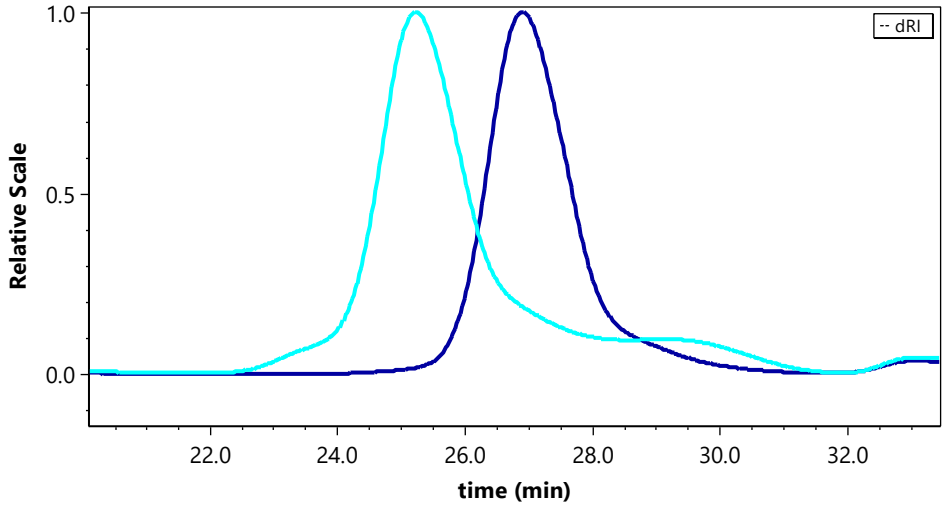


Figure S11 : *SEC chromatograms of copolymers **1**-**2** in aqueous buffer/ACN (RI detection).*

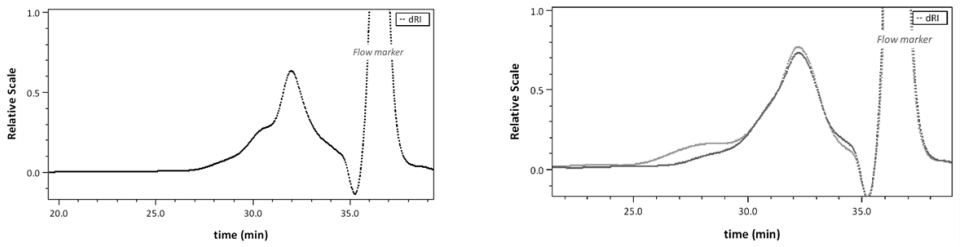


Figure S12 : *SEC chromatograms of polymers **13** (on the left) and **14** (on the right) in aqueous buffer/ACN (RI detection).*

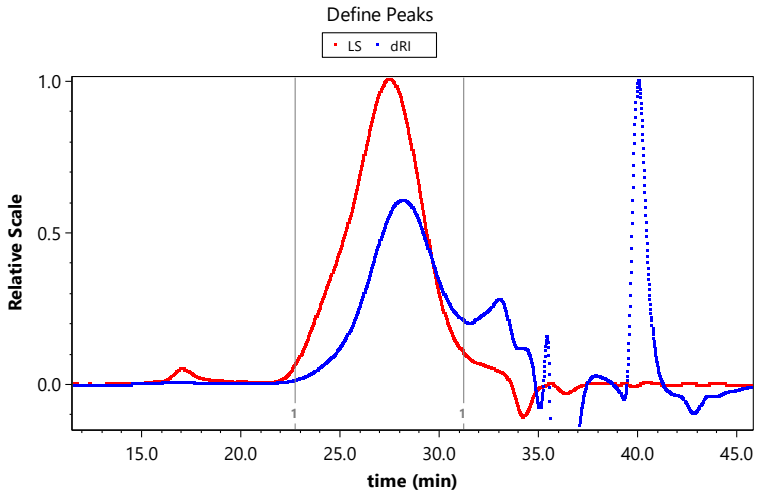


Figure S13 : *SEC chromatograms of polymer **15** in aqueous buffer/ACN. MALS and RI detection are shown.*

4.5. MALDI-TOF

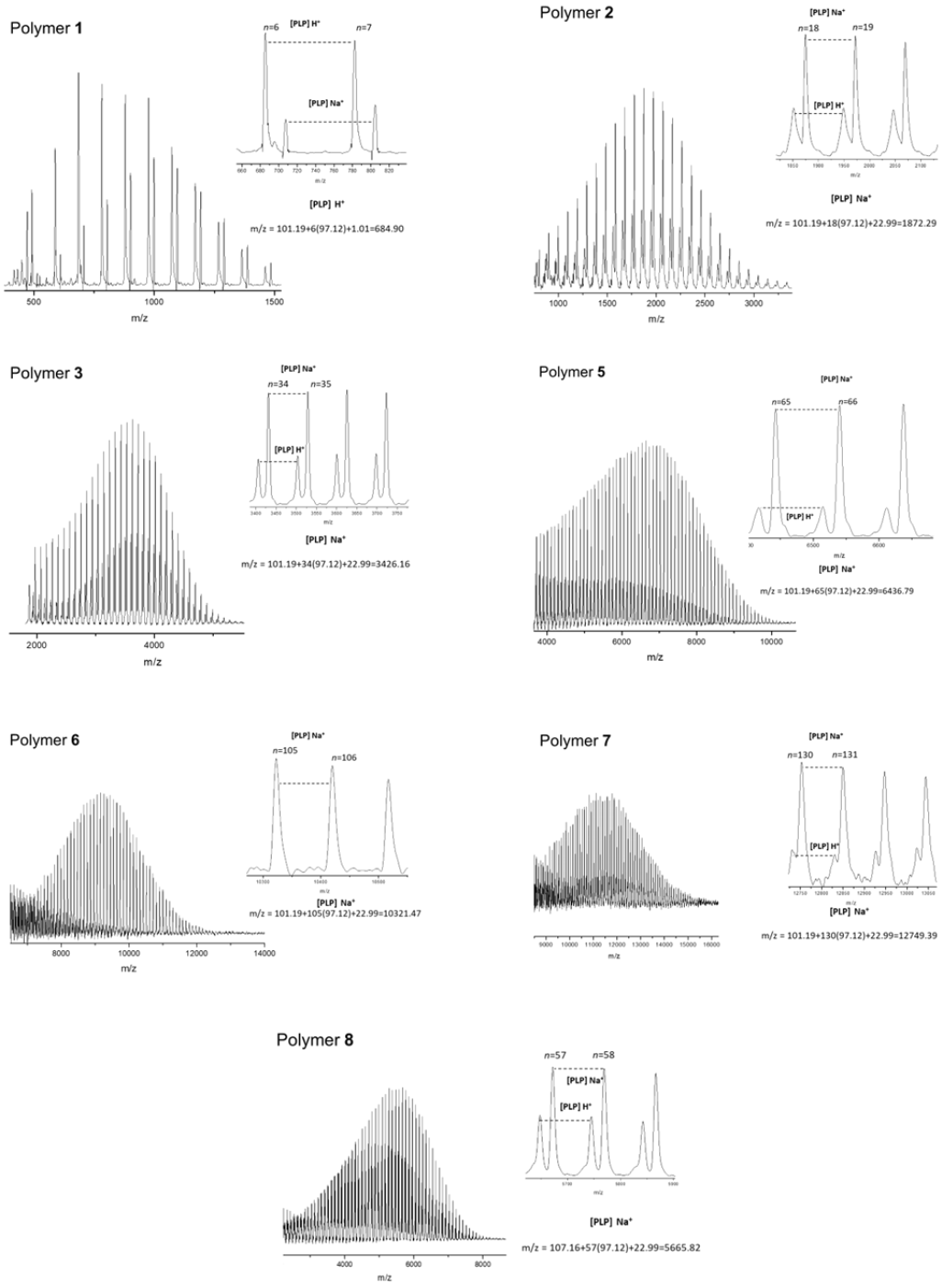


Figure S14: *MALDI-TOF spectra of polymers* ***1,2,3,5,6,7*** *and* ***8*** *accompanied by the assignment of peaks to proline units along the chain and initiator chain end verification.*

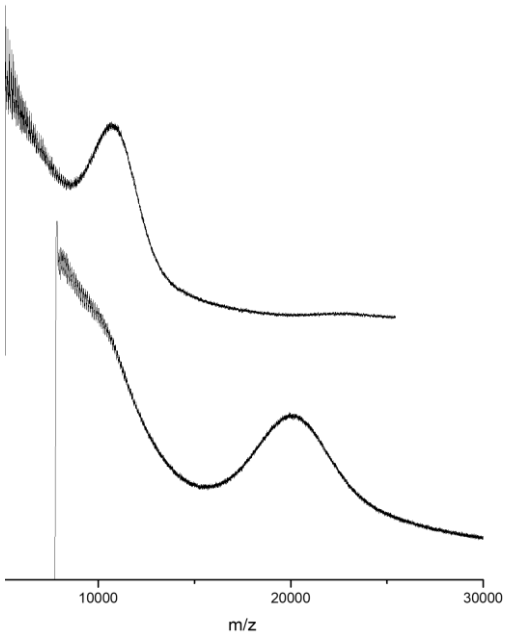


Figure S15 : *MALDI-TOF spectra of copolymer **1** (top) and **2** (bottom). The resolution of the machine is not sharp enough to detect specific peaks for such block copolymers, but polymer species with the expected molecular weight are detected.*

4.6. Thermoresponsivness

Table S2: *All results for the UV-Vis analysis of PLPs*

Polymer	Concentration (mg/mL)	Molarity[a] (µM)	T_{CP} (°C) [b]	T_{CL} (°C) [b]	X_H (K) [c]
1	5	52	-	-	-
2	10	104	69	41	28
	5	52	76	41	35
3	5	52	68,0	19,0	49
	2	21	74	20	54
	1	10	81	21	60
	0.5	5,2	> 85	22	> 63
4	5	52	66	11	55
	2	21	73	11	62
	1	10	78	14	64
	0.5	5.2	81	15	66
5	5	52	65	12	53
	2	21	73	11	62
5	1	10	79	16	63
	0.5	5.2	84	16	68

Polymer	Concentration (mg/mL)	Molarity[a] (µM)	T_{CP} (°C) [b]	T_{CL} (°C) [b]	X_H (K) [c]
6	5	52	64	12	52
	2	21	74	14	60
	1	10	77	15	62
	0.5	5.2	84	16	68
7	5	52	73	13	60
	2	21	77	14	63
	1	10	> 85	15	> 70
	0.5	5,2	> 85	14	> 71
8	5	52	65	11	54
	2	21	70	11	59
	1	10	79	13	66
	0,5	5,2	79	15	64
9	5	52	70	14	56
	2	21	75	16	59
	1	10	> 85	16	> 69
	0.5	5.2	> 85	24	> 61
10	5	52	74	16	58
	2	21	77	16	61
	1	10	82	18	64
	0.5	5.2	84	15	69
11	5	52	64	13	51

a. considering proline units; b. calculated at the inflexion points of heating and cooling curves; c. $X_H = T_{CP} - T_{CL}$.

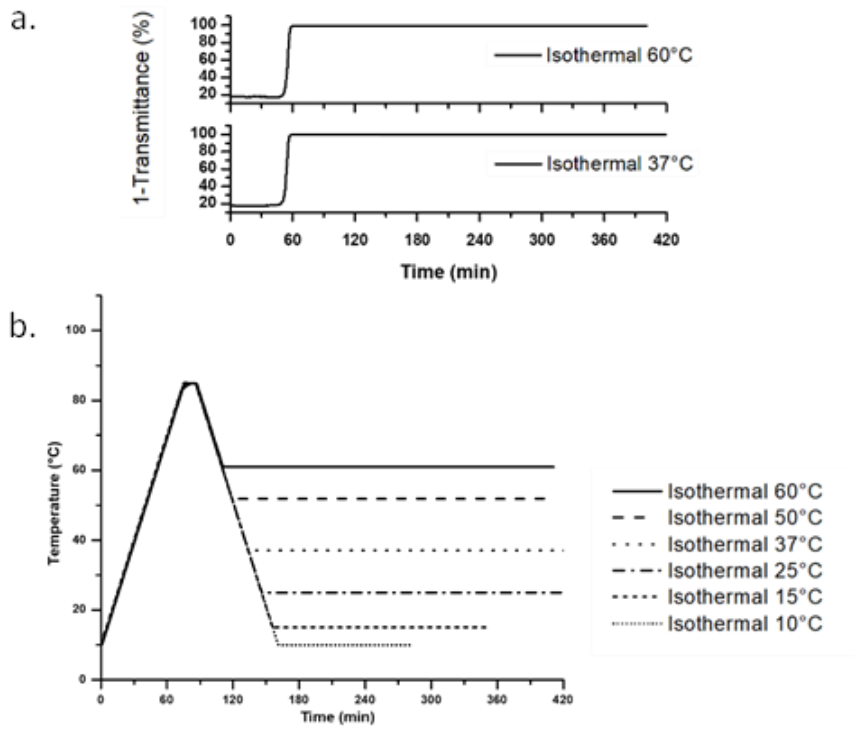


Figure S16 : *a. Stability of the aggregation/hysteresis phenomena at 60°C and 37°C (isothermal temperatures). b. Thermal profiles of all aggregation stability experiments. All samples were heated from 10°C to 85°C at 1°C/min. After aggregation and a 10 min isotherm at 85°C, samples were cooled down by 1°C/min to indicated temperatures (60°C, 50°C, 37°C, 25°C, 15°C and 10°C respectively). Acquisition was continued isothermally for up to 4 hours. All tests were done with aqueous solutions of polymer 4 at 5mg/ml.*

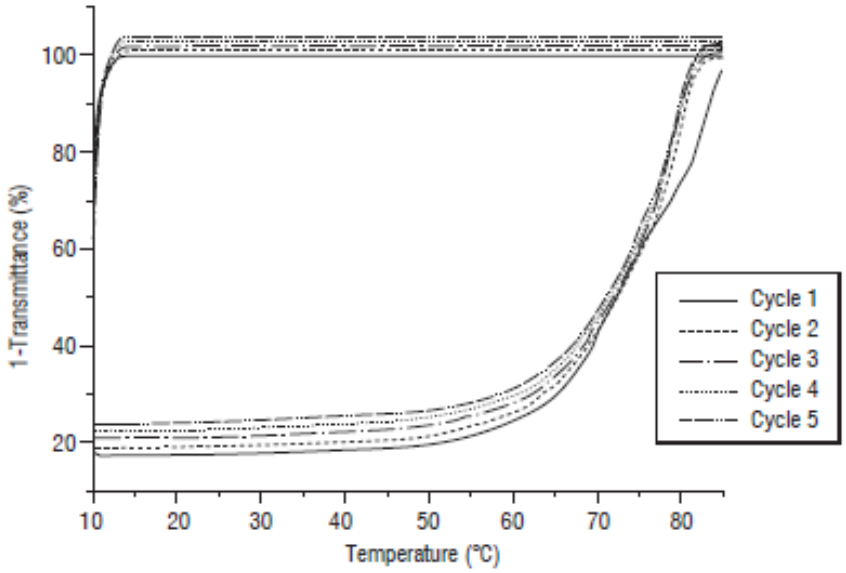


Figure S17 : *Polymer **4** at 5mg/ml in MilliQ water as measured by UV-Vis turbidimetry at 5°C/min ramps Spectra have been slightly offset towards higher 1-Tr values for better visibility.*

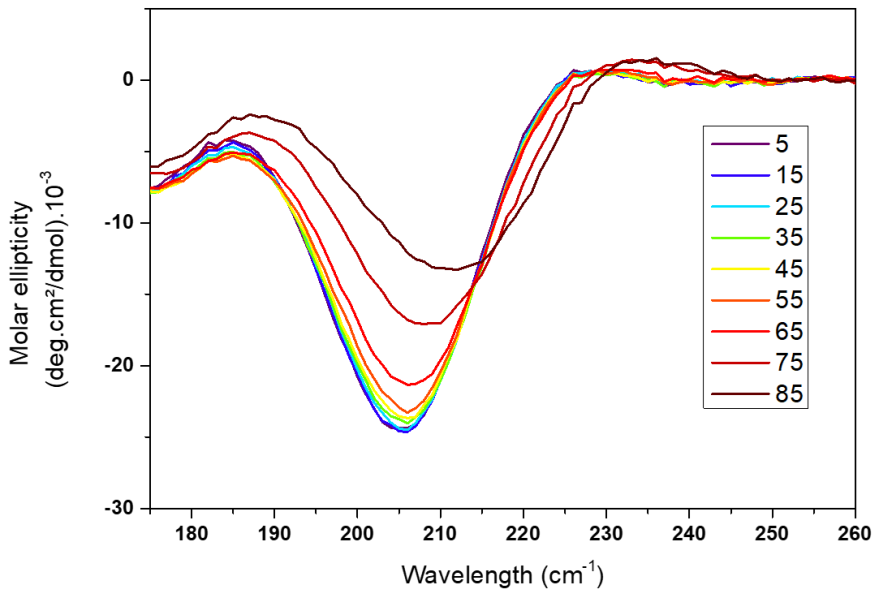


Figure S18 : *srCD spectra of polymer **4** at 5 mg/ml at different temperatures. The polymer was left to stabilize for 5min between heating ramps.*

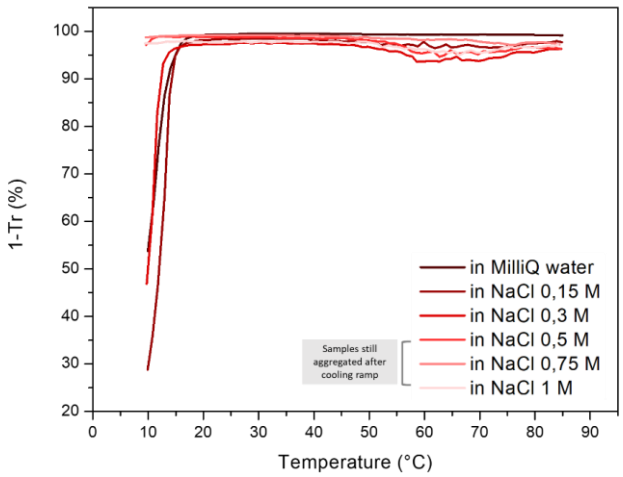


Figure S19 : *UV-Vis turbidimetry showcasing the effect of salt concentration on the T_{CL} of different solutions of polymer **4** (5 mg/ml). Completes the temperature cycle of **Figure 13**. At high concentrations of salt, the T_{CL} was not observed in the time scope of the experiments, but polymers were again soluble after a few hours' storage in the fridge at 4°C.*

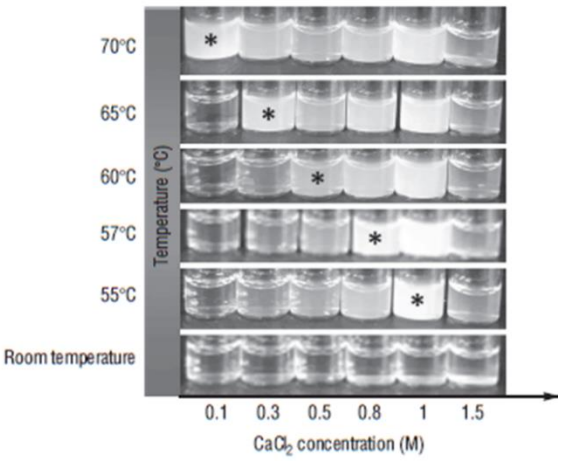


Figure S20: *temperature sensor model using different solutions of polymer **4** at 5mg/ml varying only the $CaCl_2$ concentration. All samples were heated to the temperatures indicated, stabilized for 5 minutes and a picture of the whole series taken at room temperature. The stars highlight the working range of this simple temperature sensor. Nonetheless, most polymer solutions express a pre-aggregation starting at 55°C.*

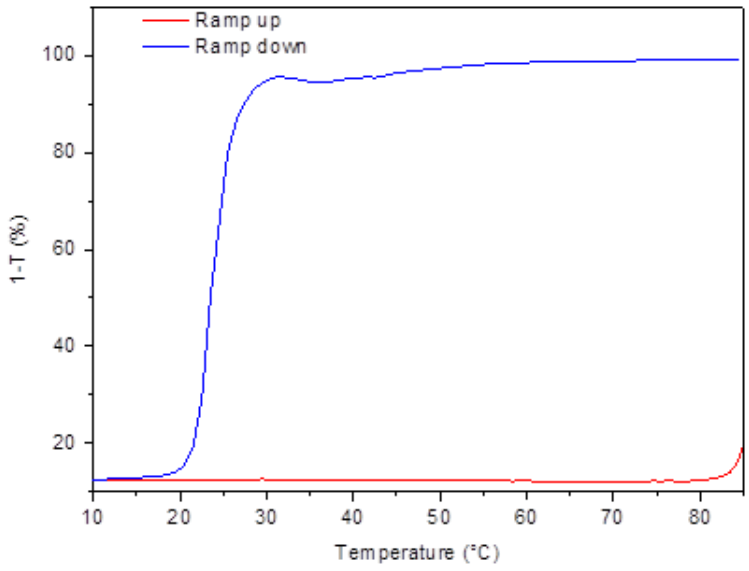


Figure S21: *UV-Vis turbidimetry curves for copolymer* ***1*** *PEG5K-PLP50.*

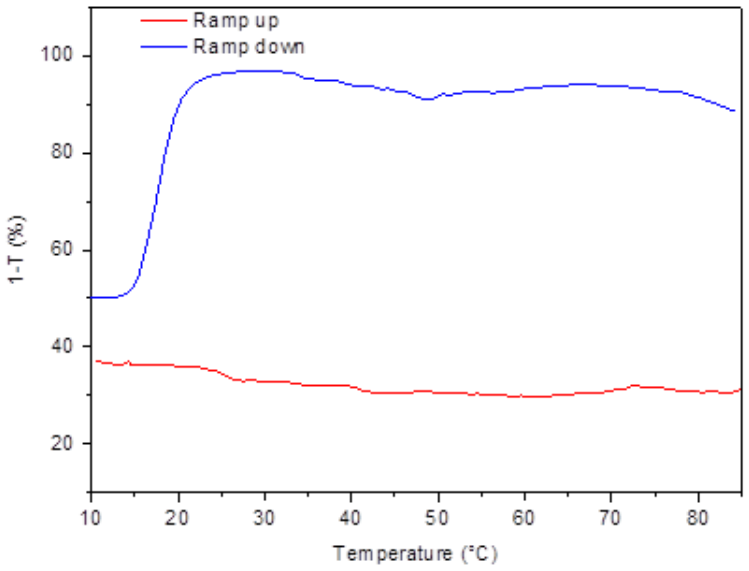


Figure S22 : *UV-Vis turbidimetry curves for copolymer* ***2*** *PEG10K-PLP100.*

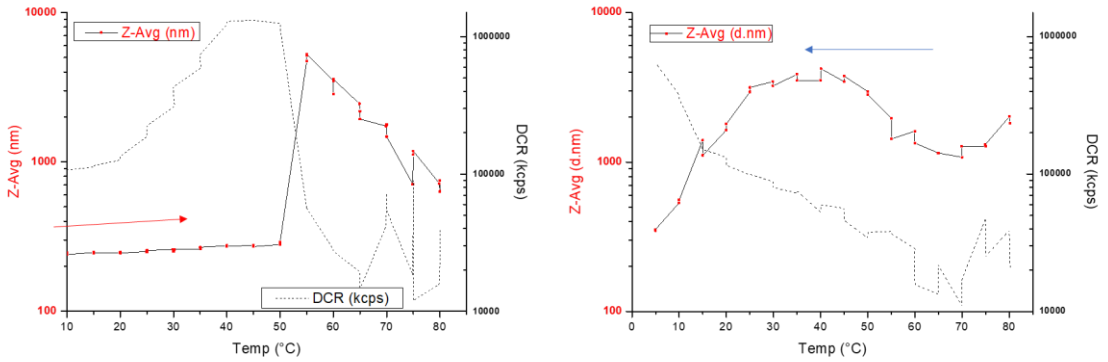


Figure S23 : *DLS data of heating (left) and cooling (right) ramps of copolymer **1** at 10mg/ml in a 3M NaCl concentration.*

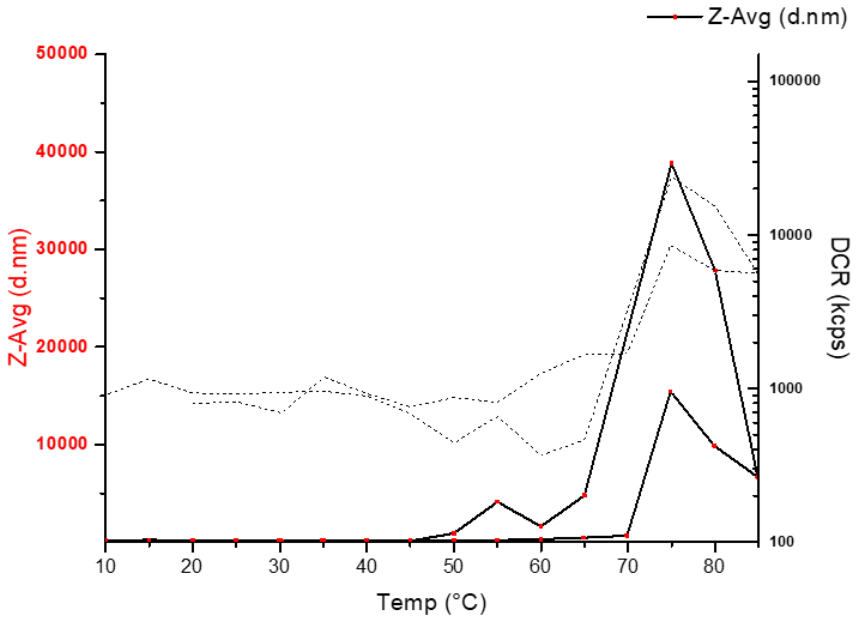


Figure S24 : *DLS of heating and cooling ramps of PEG10k at 10mg/ml and 3M NaCl concentration.*

5. References

(1) Keating, C. D. *Acc. Chem. Res.* **2012**, *45* (12), 2114–2124.

(2) Walter, H.; Johansson, G. *Anal. Biochem.* **1986**, *155* (2), 215–242.

(3) Albertsson, P.-Å. **1970**; pp 309–341.

(4) Paricaud, P.; Galindo, A.; Jackson, G. *Mol. Phys.* **2003**, *101* (16), 2575–2600.

(5) Frazar, E. M.; Shah, R. A.; Dziubla, T. D.; Hilt, J. Z. *J. Appl. Polym. Sci.* **2020**, *137* (25), 1–14.

(6) Halperin, A.; Kröger, M.; Winnik, F. M. *Angew. Chemie Int. Ed.* **2015**, *54* (51), 15342–15367.

(7) Hoogenboom, R.; Thijs, H. M. L.; Jochems, M. J. H. C.; van Lankvelt, B. M.; Fijten, M. W. M.; Schubert, U. S. *Chem. Commun.* **2008**, No. 44, 5758.

(8) Hoogenboom, R.; Schlaad, H. *Polym. Chem.* **2017**, *8* (1), 24–40.

(9) Liu, H. Y.; Zhu, X. X. *Polymer (Guildf).* **1999**, *40* (25), 6985–6990.

(10) Ansari, M. J.; Rajendran, R. R.; Mohanto, S.; Agarwal, U.; Panda, K.; Dhotre, K.; Manne, R.; Deepak, A.; Zafar, A.; Yasir, M.; Pramanik, S. *Gels* **2022**, *8* (7), 454.

(11) Haladjova, E.; Toncheva-Moncheva, N.; Apostolova, M. D.; Trzebicka, B.; Dworak, A.; Petrov, P.; Dimitrov, I.; Rangelov, S.; Tsvetanov, C. B. *Biomacromolecules* **2014**, *15* (12), 4377–4395.

(12) Guan, Y.; Zhang, Y. *Soft Matter* **2011**, *7* (14), 6375.

(13) Michalska-Walkowiak, J.; Förster, B.; Hauschild, S.; Förster, S. *Adv. Mater.* **2022**, *34* (13), 2108833.

(14) Sambe, L.; Delarosa, V. R.; Belal, K.; Stoffelbach, F.; Lyskawa, J.; Delattre, F.; Bria, M.; Cooke, G.; Hoogenboom, R.; Woisel, P. *Angew. Chemie - Int. Ed.* **2014**, *53* (20), 5044–5048.

(15) Mori, T.; Beppu, S.; Berber, M. R.; Mori, H.; Makimura, T.; Tsukamoto, A.; Minagawa, K.; Hirano, T.; Tanaka, M.; Niidome, T.; Katayama, Y.; Hirano, T.; Maeda, Y. *Langmuir* **2010**, *26* (12), 9224–9232.

(16) Matsuda, Y.; Morishima, S.; Takahara, A.; Tasaka, S. *Polym. J.* **2021**, *53* (10), 1101–1109.

(17) Li, L. .; Shan, H.; Yue, C. Y.; Lam, Y. C.; Tam, K. C.; Hu, X. *Langmuir* **2002**, *18* (20), 7291–7298.

(18) Bizmark, N.; Caggiano, N. J.; Liu, J. X.; Arnold, C. B.; Prud'homme, R. K.; Datta, S. S.; Priestley, R. D. *Soft Matter* **2022**, 6254–6263.

(19) Xu, Y.; Li, L. *Polymer (Guildf).* **2005**, *46* (18), 7410–7417.

(20) Flemming, P.; Münch, A. S.; Fery, A.; Uhlmann, P. *Beilstein J. Org. Chem.* **2021**, *17*, 2123–2163.

(21) Quiroz, F. G.; Li, N. K.; Roberts, S.; Weber, P.; Dzuricky, M.; Weitzhandler, I.; Yingling, Y. G.; Chilkoti, A. *Sci. Adv.* **2019**, *5* (10).

(22) He, C.; Zhao, C.; Guo, X.; Guo, Z.; Chen, X.; Zhuang, X.; Liu, S.; Jing, X. *J. Polym. Sci. Part A Polym. Chem.* **2008**, *46* (12), 4140–4150.

(23) Wei, Z.; Zhu, S.; Zhao, H. *Polym. Chem.* **2015**, *6* (8), 1316–1324.

(24) Zhang, X.; Li, J.; Li, W.; Zhang, A. *Biomacromolecules* **2007**, *8* (11), 3557–3567.
(25) Mokrus, M.; Menzel, H. *Macromol. Biosci.* **2022**, *22* (6), 1.
(26) Agut, W.; Brûlet, A.; Schatz, C.; Taton, D.; Lecommandoux, S. *Langmuir* **2010**, *26* (13), 10546–10554.
(27) Wu, Q.; Zhou, D.; Kang, R.; Tang, X.; Yang, Q.; Song, X.; Zhang, G. *Chem. - An Asian J.* **2014**, *9* (10), 2850–2858.
(28) Le Fer, G.; Wirotius, A. L.; Brûlet, A.; Garanger, E.; Lecommandoux, S. *Biomacromolecules* **2019**, *20* (1), 254–272.
(29) Huang, J.; Hastings, C. L.; Duffy, G. P.; Kelly, H. M.; Raeburn, J.; Adams, D. J.; Heise, A. *Biomacromolecules* **2013**, *14* (1), 200–206.
(30) Anas, M.; Jana, S.; Mandal, T. K. *Polym. Chem.* **2020**, *11* (16), 2889–2903.
(31) Zhao, C.; Zhuang, X.; He, C.; Chen, X.; Jing, X. *Macromol. Rapid Commun.* **2008**, *29* (22), 1810–1816.
(32) Zhu, H.; Liu, Y.; Gu, D.; Rao, Z.; Li, Y.; Hao, J. *Polymer (Guildf).* **2021**, *226* (February), 123841.
(33) Zhu, H.; Gu, D.; Rao, Z.; Li, Y.; Liu, Y.; Hao, J. *Polym. J.* **2021**, *53* (8), 943–949.
(34) Zhu, H.; Zhu, M.; Shuai, S.; Zhao, C.; Liu, Y.; Liu, Y.; Li, X.; Rao, Z.; Li, Y.; Hao, J. *J. Sol-Gel Sci. Technol.* **2019**, *92* (3), 618–627.
(35) Zhao, S.; Zhu, H.; Chen, Z.; Shuai, S.; Zhang, N.; Liu, Y.; Rao, Z.; Li, Y.; Zhao, C.; Zhou, K.; Ge, W.; Hao, J. *Mater. Res. Express* **2019**, *6* (8), 085711.
(36) Fan, J.; Li, R.; Wang, H.; He, X.; Nguyen, T. P.; Letteri, R. A.; Zou, J.; Wooley, K. L. *Org. Biomol. Chem.* **2017**, *15* (24), 5145–5154.
(37) Zhang, Y.; Song, H.; Zhang, H.; Huang, P.; Liu, J.; Chu, L.; Liu, J.; Wang, W.; Cheng, Z.; Kong, D. *J. Polym. Sci. Part A Polym. Chem.* **2017**, *55* (9), 1512–1523.
(38) Choi, Y. Y.; Jang, J. H.; Park, M. H.; Choi, B. G.; Chi, B.; Jeong, B. *J. Mater. Chem.* **2010**, *20* (17), 3416–3421.
(39) Kim, J. Y.; Park, M. H.; Joo, M. K.; Lee, S. Y.; Jeong, B. *Macromolecules* **2009**, *42* (8), 3147–3151.
(40) Jeong, Y.; Joo, M. K.; Bahk, K. H.; Choi, Y. Y.; Kim, H. T.; Kim, W. K.; Jeong Lee, H.; Sohn, Y. S.; Jeong, B. *J. Control. Release* **2009**, *137* (1), 25–30.
(41) Joo, M. K.; Ko, D. Y.; Jeong, S. J.; Park, M. H.; Shinde, U. P.; Jeong, B. *Soft Matter* **2013**, *9* (33), 8014–8022.
(42) He, X.; Fan, J.; Zhang, F.; Li, R.; Pollack, K. A.; Raymond, J. E.; Zou, J.; Wooley,

K. L. *J. Mater. Chem. B* **2014**, *2* (46), 8123–8130.
(43) Musacchio, A. *EMBO J.* **2022**, *41* (5), 1.
(44) Buxbaum, E. ; "*Fundamentals of Protein Structure and Function*"; ; Springer International Publishing: Cham, **2015**.
(45) Anfinsen, C. B. *Science (80-.).* **1973**, *181* (4096), 223–230.
(46) Pace, C. N.; Shirley, B. A.; McNutt, M.; Gajiwala, K. *FASEB J.* **1996**, *10* (1), 75–83.
(47) Shortle, D. *FASEB J.* **1996**, *10* (1), 27–34.
(48) Anfinsen, C. B. *Biochem. J.* **1972**, *128* (4), 737–749.
(49) Seckler, R.; Jaenicke, R. *FASEB J.* **1992**, *6* (8), 2545–2552.
(50) England, J. L.; Haran, G. *Annu. Rev. Phys. Chem.* **2011**, *62* (1), 257–277.
(51) Erik van der Linden; Foegeding, E. A. In *Modern Biopolymer Science*; Elsevier, **2009**; pp 29–91.
(52) Renkema, J. M. .; Lakemond, C. M. .; de Jongh, H. H. .; Gruppen, H.; van Vliet, T. *J. Biotechnol.* **2000**, *79* (3), 223–230.
(53) Garanger, E.; MacEwan, S. R.; Sandre, O.; Brûlet, A.; Bataille, L.; Chilkoti, A.; Lecommandoux, S. *Macromolecules* **2015**, *48* (18), 6617–6627.
(54) Ibrahimova, V.; Zhao, H.; Ibarboure, E.; Garanger, E.; Lecommandoux, S. *Angew. Chemie - Int. Ed.* **2021**, *60* (27), 15036–15040.
(55) Hyman, A. A.; Weber, C. A.; Jülicher, F. *Annu. Rev. Cell Dev. Biol.* **2014**, *30* (1), 39–58.
(56) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (5), 285–298.
(57) Perry, S. L. *Curr. Opin. Colloid Interface Sci.* **2019**, *39*, 86–97.
(58) Kanaan, N. M.; Hamel, C.; Grabinski, T.; Combs, B. *Nat. Commun.* **2020**, *11* (1), 2809.
(59) Wendisch, V. F. *Metab. Eng.* **2020**, *58* (March 2019), 17–34.
(60) Chen, C.; Wang, Z.; Li, Z. *Biomacromolecules* **2011**, *12* (8), 2859–2863.
(61) Zhang, S.; Chen, C.; Li, Z. *Chinese J. Polym. Sci. (English Ed.* **2013**, *31* (2), 201–210.
(62) Liao, Y.; Dong, C. M. *J. Polym. Sci. Part A Polym. Chem.* **2012**, *50* (9), 1834–1843.
(63) Yu, L.; Fu, W.; Li, Z. *Soft Matter* **2015**, *11* (3), 545–550.
(64) Zhang, S.; Alvarez, D. J.; Sofroniew, M. V.; Deming, T. J. *Biomacromolecules*

2015, *16* (4), 1331–1340.

(65) Cheng, Y.; He, C.; Xiao, C.; Ding, J.; Zhuang, X.; Chen, X. *Polym. Chem.* **2011**, *2* (11), 2627.

(66) Kapetanakis, A.; Heise, A. *Eur. Polym. J.* **2015**, *69*, 483–489.

(67) Yang, Y.; Wu, Y.; Li, R.; Ling, Y.; Tang, H. *Aust. J. Chem.* **2016**, *69* (1), 112–118.

(68) Ge, C.; Liu, W.; Ling, Y.; Tang, H. *J. Polym. Sci. Part A Polym. Chem.* **2018**, *56* (2), 163–173.

(69) Fu, X.; Ma, Y.; Sun, J.; Li, Z. *RSC Adv.* **2016**, *6* (74), 70243–70250.

(70) Meng, F.; Fu, X.; Ni, Y.; Sun, J.; Li, Z. *Polymer (Guildf).* **2017**, *118*, 173–179.

(71) Lavilla, C.; Byrne, M.; Heise, A. *Macromolecules* **2016**, *49* (8), 2942–2947.

(72) Zhou, C.; Shi, Z.; Xu, F.; Ling, Y.; Tang, H. *Colloid Polym. Sci.* **2020**, *298* (10), 1293–1302.

(73) He, X.; Zhou, R.; Ge, C.; Ling, Y.; Luan, S.; Tang, H. *Eur. Polym. J.* **2019**, *112* (September 2018), 547–554.

(74) He, X.; Wang, X.; Zhou, R.; Lu, H.; Luan, S.; Tang, H. *Macromol. Chem. Phys.* **2020**, *221* (3), 1–11.

(75) Fu, X. hui; Ma, Y. nan; Sun, J.; Li, Z. bo. *Chinese J. Polym. Sci. (English Ed.* **2016**, *34* (12), 1436–1447.

(76) Liang, C.; Wang, X.; Zhou, R.; Shi, H.; Yan, S.; Ling, Y.; Luan, S.; Tang, H. *Polym. Chem.* **2019**, *10* (17), 2190–2202.

(77) Zhang, L.; Zhao, L.; Ling, Y.; Tang, H. *Eur. Polym. J.* **2019**, *111* (December 2018), 38–42.

(78) Jiang, Y.; Wang, S.; Zhang, X.; Tao, Y.; Wang, X. *J. Polym. Sci. Part A Polym. Chem.* **2016**, *54* (16), 2618–2624.

(79) Ding, J.; Zhao, L.; Li, D.; Xiao, C.; Zhuang, X.; Chen, X. *Polym. Chem.* **2013**, *4* (11), 3345–3356.

(80) Perlin, P.; Scott, W. A.; Deming, T. J. *Biomacromolecules* **2020**, *21* (1), 126–132.

(81) Ohya, Y.; Toyohara, M.; Sasakawa, M.; Arimura, H.; Ouchi, T. *Macromol. Biosci.* **2005**, *5* (4), 273–276.

(82) Ishido, Y.; Kanbayashi, N.; Okamura, T. aki; Onitsuka, K. *Macromol. Rapid Commun.* **2021**, *42* (14), 1–8.

(83) Ma, Y.; Fu, X.; Shen, Y.; Fu, W.; Li, Z. *Macromolecules* **2014**, *47* (14), 4684–4689.

(84) Fu, X.; Shen, Y.; Fu, W.; Li, Z. *Macromolecules* **2013**, *46* (10), 3753–3760.
(85) Fu, X.; Ma, Y.; Shen, Y.; Fu, W.; Li, Z. *Biomacromolecules* **2014**, *15* (3), 1055–1061.
(86) Kramer, J. R.; Deming, T. J. *J. Am. Chem. Soc.* **2014**, *136* (15), 5547–5550.
(87) Gharakhanian, E. G.; Deming, T. J. *J. Phys. Chem. B* **2016**, *120* (26), 6096–6101.
(88) Xiao, J.; Li, M.; Liu, W.; Li, Y.; Ling, Y.; Tang, H. *Eur. Polym. J.* **2017**, *88*, 340–348.
(89) Anas, M.; Dinda, P.; Kar, M.; Mandal, T. K. *Polym. Chem.* **2021**, *12* (43), 6329–6343.
(90) Jana, S.; Biswas, Y.; Mandal, T. K. *Polym. Chem.* **2018**, *9* (14), 1869–1884.
(91) Stamou, A.; Iatrou, H.; Tsitsilianis, C. *Polymers (Basel).* **2022**, *14* (4), 802.
(92) Liu, S.; Ge, C.; Ling, Y.; Tang, H. *Aust. J. Chem.* **2017**, *70* (3), 245–251.
(93) Shi, Z.; Zhang, X.; Wang, X.; Yang, F.; Yu, Z.; Ling, Y.; Lu, H.; Luan, S.; Tang, H. *Biomacromolecules* **2020**, *21* (8), 3468–3478.
(94) Fujii, S.; Kuroyanagi, S.; Shimada, N.; Matsuno, J.; Lee, J. H.; Takahashi, R.; Maruyama, A.; Sakurai, K. *J. Phys. Chem. B* **2020**, *124* (20), 4036–4043.
(95) Seuring, J.; Agarwal, S. *Macromol. Rapid Commun.* **2012**, *33* (22), 1898–1920.
(96) Zhou, Q.; Palanisamy, A.; Albright, V.; Sukhishvili, S. A. *Polym. Chem.* **2018**, *9* (40), 4979–4983.
(97) Palanisamy, A.; Albright, V.; Sukhishvili, S. A. *Chem. Mater.* **2017**, *29* (21), 9084–9094.
(98) Kuroyanagi, S.; Shimada, N.; Fujii, S.; Furuta, T.; Harada, A.; Sakurai, K.; Maruyama, A. *J. Am. Chem. Soc.* **2019**, *141* (3), 1261–1268.
(99) Xue, X.; Thiagarajan, L.; Dixon, J. E.; Saunders, B. R.; Shakesheff, K. M.; Alexander, C. *Materials (Basel).* **2018**, *11* (1), 95.
(100) Chan, B. A.; Xuan, S.; Li, A.; Simpson, J. M.; Sternhagen, G. L.; Yu, T.; Darvish, O. A.; Jiang, N.; Zhang, D. *Biopolymers* **2018**, *109* (1), 1.
(101) Guo, L.; Li, J.; Brown, Z.; Ghale, K.; Zhang, D. *Biopolymers* **2011**, *96* (5), 596–603.
(102) Fetsch, C.; Grossmann, A.; Holz, L.; Nawroth, J. F.; Luxenhofer, R. *Macromolecules* **2011**, *44* (17), 6746–6758.
(103) Fetsch, C.; Luxenhofer, R. *Polymers (Basel).* **2013**, *5* (1), 112–127.
(104) Lee, C.-U.; Li, A.; Ghale, K.; Zhang, D. *Macromolecules* **2013**, *46* (20), 8213–

8223.
(105) Robinson, J. W.; Secker, C.; Weidner, S.; Schlaad, H. *Macromolecules* **2013**, *46* (3), 580–587.
(106) Tian, J.; Sun, J.; Li, Z. *Polymer (Guildf).* **2018**, *138*, 132–138.
(107) Fu, X.; Tian, J.; Li, Z.; Sun, J.; Li, Z. *Biopolymers* **2019**, *110* (4), e23243.
(108) Xing, C.; Shi, Z.; Tian, J.; Sun, J.; Li, Z. *Biomacromolecules* **2018**, *19* (6), 2109–2116.
(109) Tao, X.; Du, J.; Wang, Y.; Ling, J. *Polym. Chem.* **2015**, *6* (16), 3164–3174.
(110) Fu, X.; Xing, C.; Sun, J. *Biomacromolecules* **2020**, *21* (12), 4980–4988.
(111) Xuan, S.; Lee, C. U.; Chen, C.; Doyle, A. B.; Zhang, Y.; Guo, L.; John, V. T.; Hayes, D.; Zhang, D. *Chem. Mater.* **2016**, *28* (3), 727–737.
(112) Lahasky, S. H.; Hu, X.; Zhang, D. *ACS Macro Lett.* **2012**, *1* (5), 580–584.
(113) Li, R.; Ding, X.; Ding, J.; Sun, J. *ACS Appl. Polym. Mater.* **2022**, 4–10.
(114) Li, R.; Ni, Y.; Liu, D.; Li, Z.; Cheng, Y.; Xia, M.; Fu, X. *Gaodeng Xuexiao Huaxue Xuebao/Chemical J. Chinese Univ.* **2021**, *42* (3), 850–856.
(115) Guo, L.; Lahasky, S. H.; Ghale, K.; Zhang, D. *J. Am. Chem. Soc.* **2012**, *134* (22), 9163–9171.
(116) Okuno, Y.; Nishimura, T.; Sasaki, Y.; Akiyoshi, K. *Biomacromolecules* **2021**, *22* (7), 3099–3106.
(117) Lahasky, S. H.; Lu, L.; Huberty, W. A.; Cao, J.; Guo, L.; Garno, J. C.; Zhang, D. *Polym. Chem.* **2014**, *5* (4), 1418–1426.
(118) Kurzhals, S.; Pretzner, B.; Reimhult, E.; Zirbs, R. *Macromol. Chem. Phys.* **2017**, *218* (13), 1700116.
(119) Lu, L.; Lahasky, S. H.; McCandless, G. T.; Zhang, D.; Garno, J. C. *ACS Appl. Nano Mater.* **2019**, *2* (12), 7617–7625.
(120) Gregory A. Petsko, D. R. ; "*Protein Structure and Function*"; ; Wiley-Blackwell, **2004**.
(121) Dukor, R. K.; Keiderling, T. A. *Biospectroscopy* **1996**, *2* (2), 83–100.
(122) Shi, L.; Holliday, A. E.; Shi, H.; Zhu, F.; Ewing, M. A.; Russell, D. H.; Clemmer, D. E. *J. Am. Chem. Soc.* **2014**, *136* (36), 12702–12711.
(123) Kuemin, M.; Engel, J.; Wennemers, H. *J. Pept. Sci.* **2010**, *16* (10), 596–600.
(124) Theillet, F.-X.; Kalmar, L.; Tompa, P.; Han, K.-H.; Selenko, P.; Dunker, A. K.; Daughdrill, G. W.; Uversky, V. N. *Intrinsically Disord. Proteins* **2013**, *1* (1), e24360.

(125) Hu, Y.; Tian, Z.-Y.; Xiong, W.; Wang, D.; Zhao, R.; Xie, Y.; Song, Y.-Q.; Zhu, J.; Lu, H. *Natl. Sci. Rev.* **2022**, *9* (8), nwac033.

(126) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106* (5), 1877–1897.

(127) Dobitz, S.; Aronoff, M. R.; Wennemers, H. *Acc. Chem. Res.* **2017**, *50* (10), 2420–2428.

(128) Barron, L. D. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2023**, *300* (300), 122959.

(129) Lu, K. P.; Finn, G.; Lee, T. H.; Nicholson, L. K. *Nat. Chem. Biol.* **2007**, *3* (10), 619–629.

(130) Adzhubei, A. A.; Sternberg, M. J. E.; Makarov, A. A. *J. Mol. Biol.* **2013**, *425* (12), 2100–2132.

(131) Proteaset, P.; Lin, L.; Brandts, J. F. **1980**, 3055–3059.

(132) Wixklaiair, D.; Krause, G. **1971**, *737*, 721–737.

(133) Steinberg, I. Z.; Harrington, W. F.; Berger, A.; Sela, M.; Katchalski, E. *J. Am. Chem. Soc.* **1960**, *82* (20), 5263–5279.

(134) Kurtz, J.; Berger, A.; Katchalski, E. *Nature* **1956**, *178* (4541), 1066–1067.

(135) Shi, L.; Holliday, A. E.; Bohrer, B. C.; Kim, D.; Servage, K. A.; Russell, D. H.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (6), 1037–1047.

(136) Torchia, D. A.; Bovey, F. A. *Macromolecules* **1971**, *4* (2), 246–251.

(137) Detwiler, R. E.; Schlirf, A. E.; Kramer, J. R. *J. Am. Chem. Soc.* **2021**, *143* (30), 11482–11489.

(138) Pulla Rao, C.; Balaram, P.; Rao, C. N. . *Int. J. Biol. Macromol.* **1983**, *5* (5), 289–295.

(139) Ciferri, A.; Orofino, T. A. *J. Phys. Chem.* **1966**, *70* (10), 3277–3285.

(140) Lorusso, M.; Pepe, A.; Ibris, N.; Bochicchio, B. *Soft Matter* **2011**, *7* (13), 6327–6336.

(141) Tifany, M. L.; Krimm, S. *Biopolymers* **1972**, *11* (11), 2309–2316.

(142) Gkikas, M.; Avery, R. K.; Olsen, B. D. *Biomacromolecules* **2016**, *17* (2), 399–406.

(143) Tooke, L.; Duitch, L.; Measey, T. J.; Schweitzer-Stenner, R. *Biopolymers* **2010**, *93* (5), 451–457.

(144) Gkikas, M.; Iatrou, H.; Thomaidis, N. S.; Alexandridis, P.; Hadjichristidis, N. *Biomacromolecules* **2011**, *12* (6), 2396–2406.

(145) Chan, B. A.; Xuan, S.; Horton, M.; Zhang, D. *Macromolecules* **2016**, *49* (6), 2002–2012.

(146) Wu, Y.; Chen, K.; Wu, X.; Liu, L.; Zhang, W.; Ding, Y.; Liu, S.; Zhou, M.; Shao, N.; Ji, Z.; Chen, J.; Zhu, M.; Liu, R. *Angew. Chemie Int. Ed.* **2021**, *60* (50), 26063–26071.

(147) Salas-Ambrosio, P.; Tronnet, A.; Since, M.; Bourgeade-Delmas, S.; Stigliani, J. L.; Vax, A.; Lecommandoux, S.; Dupuy, B.; Verhaeghe, P.; Bonduelle, C. *J. Am. Chem. Soc.* **2021**, *143* (10), 3697–3702.

(148) Zhang, Q.; Weber, C.; Schubert, U. S.; Hoogenboom, R. *Mater. Horizons* **2017**, *4* (2), 109–116.

(149) Osváth, Z.; Iván, B. *Macromol. Chem. Phys.* **2017**, *218* (4), 1600470.

(150) Chittari, S. S.; Obermeyer, A. C.; Knight, A. S. *J. Am. Chem. Soc.* **2023**, *145* (11), 6554–6561.

(151) Dorman, D. E.; Torchia, D. A.; Bovey, F. A. *Macromolecules* **1973**, *6* (1), 80–82.

(152) Wenisch, S. E.; Schaffer, A.; Rieger, B. *Macromol. Chem. Phys.* **2022**, *223* (15), 2200063.

(153) Bonduelle, C. *Polym. Chem.* **2018**, *9* (13), 1517–1529.

(154) Gkikas, M.; Haataja, J. S.; Ruokolainen, J.; Iatrou, H.; Houbenov, N. *Biomacromolecules* **2015**, *16* (11), 3686–3693.

(155) Chao, Y.-Y. H.; Bersohn, R. *Biopolymers* **1978**, *17* (12), 2761–2767.

(156) Badreldin, M.; Le Scouarnec, R.; Lecommandoux, S.; Harrisson, S.; Bonduelle, C. *Angew. Chemie Int. Ed.* **2022**, *61* (46), e202209530.

(157) Detwiler, R. E.; Mcpartlon, T. J.; Coffey, C. S.; Kramer, J. R. *ACS Polym. Au* **2023**, No. "2023 Rising Stars".

(158) Bisht, A. S.; Maity, P.; Roy, R. K. *Macromolecules* **2023**, *56* (11), 3922–3930.

(159) Judge, N.; Georgiou, P. G.; Bissoyi, A.; Ahmad, A.; Heise, A.; Gibson, M. I. *Biomacromolecules* **2023**, *24* (6), 2459–2468.

(160) Graham, B.; Bailey, T. L.; Healey, J. R. J.; Marcellini, M.; Deville, S.; Gibson, M. I. *Angew. Chemie - Int. Ed.* **2017**, *56* (50), 15941–15944.

(161) Rojas, R.; Aróstica, M.; Carvajal-Rondanelli, P.; Albericio, F.; Guzmán, F.; Cárdenas, C. *Electron. J. Biotechnol.* **2022**, *59*, 62–73.

Chapter IV: Towards synthetic mimics of elastin-like polypeptides

1. Introduction

The primary sequences of intrinsically disordered polypeptides (IDPs) usually possess secondary structure-breaking amino acids like glycine and proline.[1,2] Chapter **2** introduced the possibility of copolymerizing glycine to disrupt α-helical structures while chapter **3** focused on the thermoresponsive behaviour of homopolymers based on proline. This protein-like stimuli-responsiveness is also found in some IDPs. Indeed, adding hydrophobic amino acids to the sequences of IDPs results in considerable entropy gains upon binding to a partner; these segments often form fused globules (coacervates) and present a characteristic temperature response. Indeed, the more hydrophobic the IDP, the lower the temperature at which phase separation occurs.[3] The sequence of hydrophobicity (determined by both aromatic and aliphatic residues) is very important to maintain the temperature induced phase separation.[4] This has been particularly studied in the context of elastin-like polypeptides (ELPs).

1.1. Elastin Like Polypeptides (ELPs)

ELPs are a class of recombinant IDP based on -Val-Pro-Gly-Xaa-Gly- pentapeptide repeats (**Figure 1**) that were originally identified in tropoelastin and in α-elastin (Xaa being Ala, Gly, or Val). These proteins possess a thermoresponse, usually with low hysteresis (**Figure 1**), that can be tuned by varying the "guest" residue (Xaa).[5–9] Xaa can be any natural or non-natural alpha amino acid except for proline, which disrupts the thermoresponsive behavior.[10,11] As presented in chapter **3**, our studies on fully synthetic poly(L-proline) (PLP) backbone tend to contradict this “proline” statement in ELPs. ELPs are generally produced by recombinant synthesis[12], an approach which allows researchers to explore an expansive array of sequences and lengths, and consequently, a diverse spectrum of thermoresponsive behaviors.[13] This has recently included some genetically encoded ELPs with hysteretic thermoresponsivess as we observed with PLP.[14]

The LCST behaviour exhibited by ELPs aligns with that of synthetic LCST polymers (**Figure 1**). The dissolution dynamics of ELPs are driven by enthalpically favourable interactions with solvent, and counteracted by entropic considerations due to structuring effect of hydrogen bonding amide groups and the voluminous hydrophobic substituents of the valine and proline units.[12,15] The thermoresponsive behaviour of ELPs is generally accompanied by a change in conformation, much like that of PLP.[16] The soluble state is found in a mostly random coil structure while a β-turn CD signature arises above its T_{CP}. This signature, that is due to the -PGGV- motifs,[17] has a negative band at 224 nm, a positive band at 205 nm and a second negative band below 190 nm.[18–20]

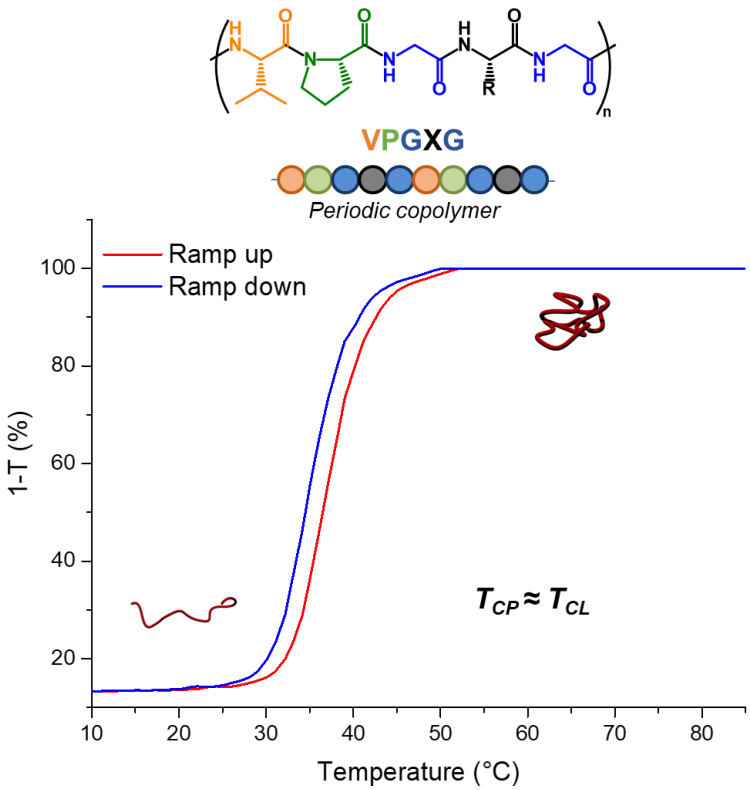


Figure 1 : *Generic structure of ELPs' repeat unit $(VPGXG)_n$. Representation of LCST behaviour of ELP with both transition temperatures around 37°C.*

The thermoresponsive behaviour of ELPs thus spans a large temperature window.[12] Furthermore, their properties can be tuned by chemical modification [21–24] or functionalization with saccharides and oligosaccharides,[25–28] or by including them in block copolymer structures[29–32] or even artificial organelles[33,34]. Such

biomacromolecules that respond to biologically relevant factors by changing their conformation or mechanical properties are promising building blocks for the design of smart biomaterials.[35–38] The use of synthetic protein-like macromolecules may also shed light on biological phase separations and pave the way to new biomimetic materials.[39]

Inspired by the knowledge surrounding ELPs and building on the previous synthetic effort, this chapter will focus first on the copolymerization of proline NCA with glycine NCA to observe the impact of the secondary structure change on the thermoresponsive behaviour. As a means to produce polypeptides with tuneable thermoresponsiveness, valine units will then be introduced to copolymerization and terpolymerization systems to add hydrophobicity. The goal of this copolymerization is to produce statistical copolymers with a microstructure that best matches the periodic sequence of ELPs (**Figure 2**). The kinetic studies developed in chapter **2** can prove useful in exerting stochastic control over the monomer insertions to increase the probability of finding the desired sequences along the copolymer chains. The overarching goal of the chapter is thus to synthesize intrinsically disordered polypeptides with thermoresponsive behaviour comparable to that of ELPs (**Figure 1**).

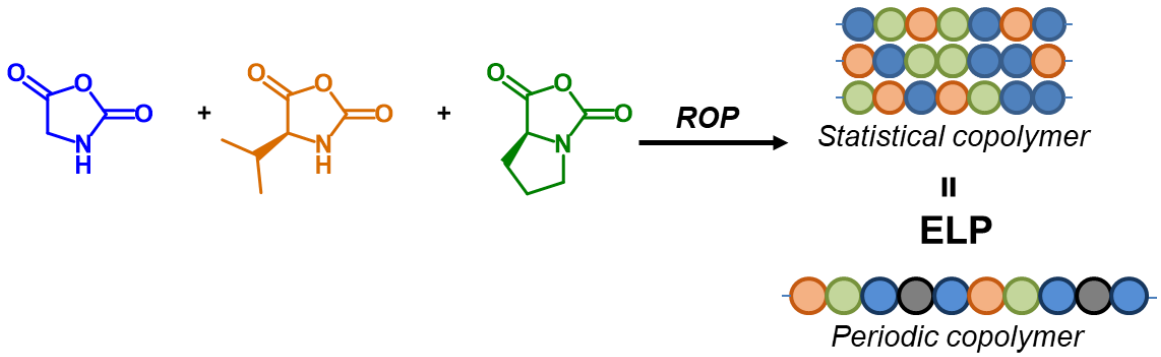
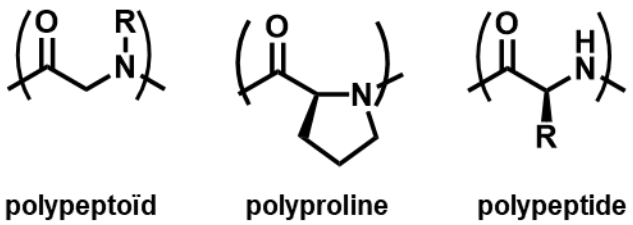


Figure 2 : *Copolymerization of Gly NCA, Val NCA and Pro NCA to produce statistical copolypeptides with the goal to mimic the periodic primary sequence of ELPs and their thermoresponsive behaviour.*

1.1. Copolymerization

Proline NCA, is an example of an N-alkylated NCA (NNCA). When unsubstituted on the C atom, these NCAs generally give access to polypeptoids[40] (**Scheme 1**) which are known to have thermoresponsive behaviours.[41]

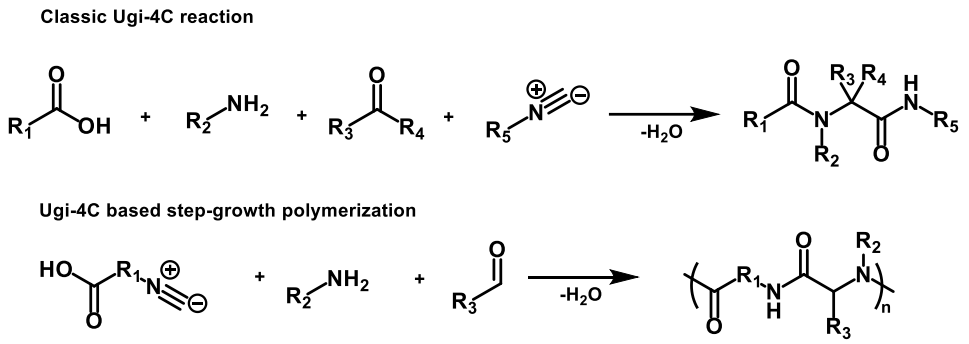


Scheme 1 : *Generic structures of polypeptoïds (obtained from ROP of NNCAs), polyproline (from ROP of Pro NCA) and polypeptide (from ROP of NCAs of α-amino acids).*

Examples, of statistic copolymerizations of NCA with their *N*-alkylated counterparts are scarce in the literature[42] and mainly include copolymerizations of sarcosine NCA, the simplest, *N*-methylated NNCA.[43–48] The more sterically hindered NNCAs are less reactive than NCAs and propagate via secondary amines rather than primary amines when the NAM mechanism is involved. This generally results in the formation of block-like copolymers. To overcome the difference in reactivity of NCAs and NNCAs it is possible to substitute the latter with more reactive NNTAs. In this manner, using the classical ROP mechanism, nearly random copolymers containing both amino acids and N-alkylated repeat units have been produced.[49] The recent advances in catalysis for the ROP of both NCAs and NNCAs that were discussed in the previous chapters have scarcely been applied to statistical copolymerization. To the best of our knowledge, only one example using Ni-based organometallic catalysts to statistically copolymerize proline NCA and glutamate NCA has been reported.[50] In the following, efforts on the copolymerization of Pro NCA (an N-alkylated amino acid) with NCAs of glycine and valine will be discussed using some of the catalyst and processes recently developed.[51,52] The highly successful water-assisted ROP exploited for the polymerization of Pro NCA has yet to be explored in copolymerization and will be the starting point of the experimental part building on the results from chapter **3** (section **2.1** of this chapter **4**). It will be followed by copolymerizations using other catalytic mechanisms including organocatalysis and organometallic catalysis (section 2.2 of this chapter **4**).

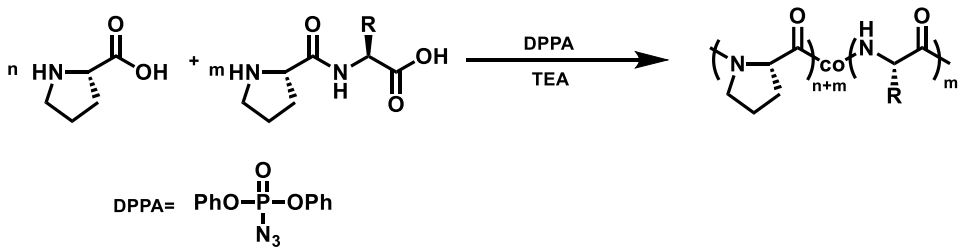
It should be noted that other ways to obtain polymers having both amino acids and *N*-alkylated units in a single backbone do exist, including copolymerizations based on the Ugi multi-component reaction (Ugi 4CC). This reaction of equimolar amounts of an aldehyde or ketone, an amine, an isonitrile and a carboxylic acid produces a dipeptide.

By combining isonitrile and carboxylic acid components in the same molecule, it is possible to produce alternating peptides via polycondensation (**Scheme 2**).[53]



Scheme 2: *Ugi 4 component reaction (on the top) producing a dipeptide and the step-growth polycondensation based on the reaction (on the bottom) producing alternating polypeptides or alternating peptide/peptoid hybrids.*

The Ugi 4CC polymerization process gives access to perfectly alternating copolymers of amino acids and N-alkylated amino acids. Nonetheless, the polymers produced via Ugi 4CC polymerization are usually of low molecular weight and high dispersity, as is usual for step growth processes.[54] Closer to solid-phase peptide synthesis, classical coupling agents such as diphenyl phosphoryl azide (DPPA) in the presence of triethylamine (TEA) may be used to produce peptides containing proline and amino acids like glycine[55], alanine[55,56] or other proline derivatives[57] (**Scheme 3**). Using this approach, backbones with non-alkylated amino acids were produced by first preparing dipeptides of Pro-Xaa that were later condensed with proline amino acids. This process produced small peptides with no Xaa-Xaa bonds, ensuring an optimized PLPII secondary structuring.[55]



Scheme 3 *: DPPA/TEA coupling of proline and proline-based dipeptides to produce the desired copolymers*

The Ugi polymerization approach and the DPPA coupling approach have both been used to produce thermoresponsive peptides.

1.2. Thermoresponsive copolymers based on proline

Copolymers containing proline produced by the DPPA technique described above showed LCST behaviour. The T_{CP} of these peptides increased with increasing glycine and alanine content (from 50°C to 68°C at 20 mg/ml) as these are more hydrophilic than proline, and decreased (from 65°C to 48°C at 2.5 mg/ml) when the more hydrophobic comonomer *O*-benzylhydroxyproline was introduced.[55–57] Nonetheless, these studies did not focus on the resolubilization process (T_{CL}) nor the hysteresis of these low molecular weight copolymers (up to 2.2 kg/mol). The results did confirm that PLP thermoresponsiveness could be tuned by introducing units of higher or lower hydrophobicity. On another hand, using the Ugi four-component polymerization reaction, alternating copolymers made of glycine and *N*-alkylated-glycines have already expressed LCST behavior.[54,58] When glycine was exchanged with valine, the alternating copolymers also showed UCST behaviour.[59]

The same Ugi 4CC polymerization technique was used to produce alternating copolymers of glycine and modified *N* and *C* alkylated glycines.[60] By introducing an alkane chain on the C_α and either a propyl or a hydroxyethyl substituent on the amine, the process gives access to different copolymers exhibiting UCST, LCST or UCST and LCST simultaneously (**Figure 3**). This thermoresponsiveness was explained by the alternating hydrophilic (native glycine) and hydrophobic (*N* and *C* alkylated) units. Overall, these results show that copolymers based on valine and glycine may express thermoresponsivness.

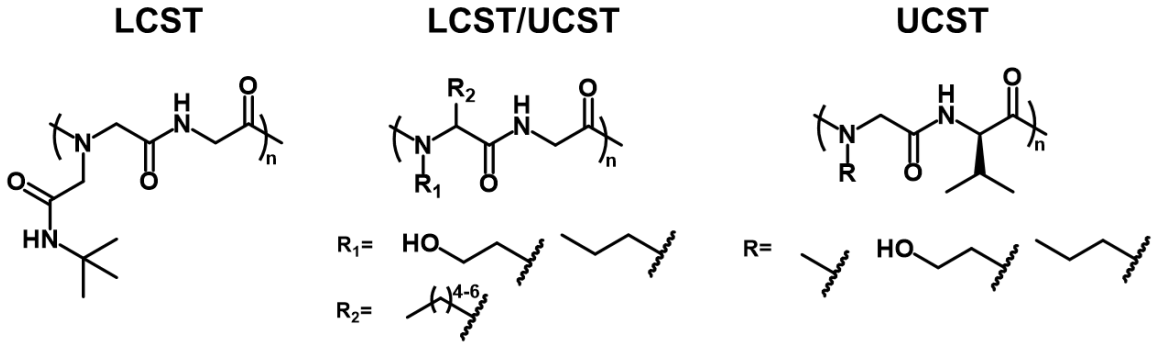


Figure 3 : *Alternating peptide/peptoid hybrids produced via Ugi 4CC polymerization and their thermoresponsive behaviours.*

Koyama and coauthors published a series of articles on thermoresponsive copolymers prepared by ROP of NCAs containing valine and glycine units.[61] Using appropriate

mixtures of NCAs, they produced both block and random polypeptides with molar masses obtained by DOSY analysis of 3.4 kg/mol and 3.3 kg/mol respectively. Both copolymers exhibited similar LCST at a temperature around 80°C. To better understand the effect of the sequence on the thermoresponsiveness, they also prepared an alternating peptide using the Ugi polycondensation discussed earlier. Unexpectedly, this new copolymer expressed a UCST at low temperatures that was further tuned by *N*-alkylation of the valine units.[59] These examples highlight the importance of the primary sequence of the peptide backbone on the thermoresponsiveness. Nonetheless, the copolymers produced were not analysed by SEC techniques and the thermoresponsive behaviour was not extensively studied. The random and block copolymers showed very small changes in turbidity when heated above their T_{CP}. Further investigation into the copolymer structures and aggregation mechanism would thus be needed.

The various ROP mechanisms developed in chapter **3** for Pro NCA will be used in the following part to copolymerize Pro NCA with Gly and Val NCA. First binary systems based on proline with glycine or valine will be explored. The resulting copolymers and their thermoresponsive behaviour will be characterized. More complexity to the system will be added by terpolymerizing the three NCAs. By modulating the feed ratio interesting microstructures and thermoresponsive behaviour will be targeted. This analysis will be discussed for the water assisted mechanism and for other synthesis procedures in organic media.

2. Results and discussion

2.1. Water assisted ROP of proline: towards copolymerization

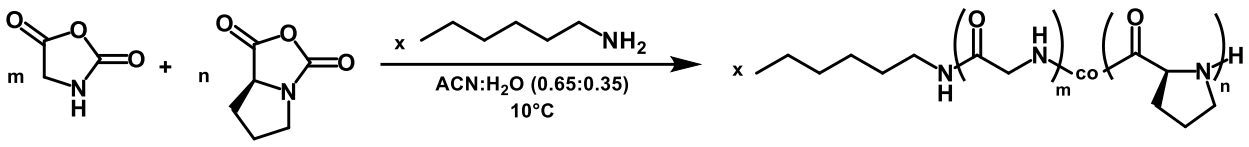
The water assisted ROP presented in the previous chapter was explored in copolymerization systems to obtain thermoresponsive copolypeptides. The procedure has been kept the same as previously explained for initial monomer concentration, temperature, and overall process.

2.1.1. Proline-Glycine copolymers

Synthesis and analysis

Although the water assisted mechanism has been proven successful for the ROP of Pro NCA,[51] its use for statistical copolymerizations has yet to be developed. To prove

this possibility, statistical copolymerization of Pro-NCA with Gly-NCA using the water assisted ROP in H_2O/ACN mixture was first explored (**Scheme 4**).



Scheme 4: *Copolymerization of glycine NCA and proline NCA using hexylamine as initiator via the water assisted ROCOP. The first series of copolymers was explored by changing the m/n ratio. A second series of copolymers was studied by changing the m+n/x ratio.*

Both NCAs were first suspended in acetonitrile, then the initiator (hexylamine) was added in a water/acetonitrile mixture at 10°C and under stirring. The reaction medium first became homogeneous, then became opalescent. Two parameters were studied by synthesizing a set of copolymers (**1-8** in **Table 1**)**.** First the glycine content was increased from 10% for copolymer **1** to 78% for copolymer **5** (**Table 1**). Second, copolymers having a targeted glycine content of 50% were synthesized with increasing M/I ratio, 50 for copolymer **6** and 130 for copolymer **8** (**Table 1**). All copolymers were purified by dialysis against MilliQ water followed by lyophilization. The copolymers' structure was analysed by ^{1}H NMR and SEC. The solubility/thermoresponsive behaviour of the obtained copolymers was also investigated.

Table 1: *Analysis and results for copolymers **1-8** synthesis*

	Targeted structure (Theoretical Mn g/mol)	^{1}H NMR	Gly%	Mn by SEC (g/mol) Đ[e]
Copolymer **1**	HA-Pro$_{90}$-Gly$_{10}$ (9300)	HA-Pro$_{66}$-Gly$_{7}$[a]	10	8800[c] 1.01
Copolymer **2**	HA-Pro$_{70}$-Gly$_{30}$ (8400)	HA-Pro$_{75}$-Gly$_{37}$ [a]	33	9700 [c] 1.03
Copolymer **3**	HA-Pro$_{50}$-Gly$_{50}$ (7700)	HA-Pro$_{66}$-Gly$_{52}$ [a]	44	7800 [c] 1.02 15900 [d] 1.02
Copolymer **4**	HA-Pro$_{40}$-Gly$_{60}$ (7300)	HA-Pro$_{38}$-Gly$_{53}$ [b]	58	19200 [d] 1.18
Copolymer **5**	HA-Pro$_{20}$-Gly$_{80}$ (6500)	HA-Pro$_{16}$-Gly$_{69}$ [b]	78	15500 [d] 1.08

	Targeted structure (Theoretical Mn g/mol)	^{1}H NMR	Gly%	Mn by SEC (g/mol) Đ[e]
Copolymer **6**	HA-Pro$_{25}$-Gly$_{25}$ (3850)	HA-Pro$_{30}$-Gly$_{27}$	47	3800 [c] 1.11
Copolymer **7**	HA-Pro$_{38}$-Gly$_{38}$ (5800)	HA-Pro$_{30}$-Gly$_{27}$	47	4100 [c] 1.10
Copolymer **8**	HA-Pro$_{65}$-Gly$_{65}$ (10000)	HA-Pro$_{70}$-Gly$_{61}$	47	16000 [f] 1.10

[a]determined by integration of the proline and glycine backbone peaks and hexylamine peaks in ^{1}H NMR in D_2O; [b]determined by integration of the proline and glycine backbone peaks and hexylamine peaks in ^{1}H NMR in TFA-d; [c]determined by dn/dc of polyproline (0.1414) in ACN/acetate buffer SEC and MALS detection; [d]determined by PMMA calibration in HFIP SEC and RI detection; [e]determined by SEC peaks; [f]determined by PEG calibration in ACN/acetate buffer and RI detection as the polymer caused high aggregation in aqueous SEC and Mn determination using MALS was not possible.

Copolymers were analysed by ^{1}H NMR in either D_2O or TFA-d depending on their solubilities (**Table 1**). Copolymers containing less than 50% glycine (Copolymer **1**-**3** and **6-8**) were soluble in water thanks to their high PLP content. These copolymers were thus analysed in aqueous conditions and their molecular weights were determined by dn/dC from MALS detection (**Table 1**). Copolymers with higher glycine contents (copolymer **4-5**) were insoluble in water and were thus analysed in SEC HFIP. Their molar mass was estimated using calibration with PMMA standards (**Table 1**). Copolymer **3,** with a glycine content of 44%, was soluble in both media and was analysed in both to obtain a comparison of the whole series.

All copolymers showed adequate glycine content to theoretical values as determined by ^{1}H NMR. **Figures 5-6** show the spectra obtained for copolymer **3** and copolymer **4** with adequate glycine content (44% and 58%) in D_2O and TFA-d respectively. The methylene peaks (i and j in **Figures 5-6**) from the proline side chain were integrated separately from all other backbone and side chain peaks (g for glycine and h and k for proline). This allowed for easier calculation of the glycine content and Dp by comparing to the initiator peaks (a).

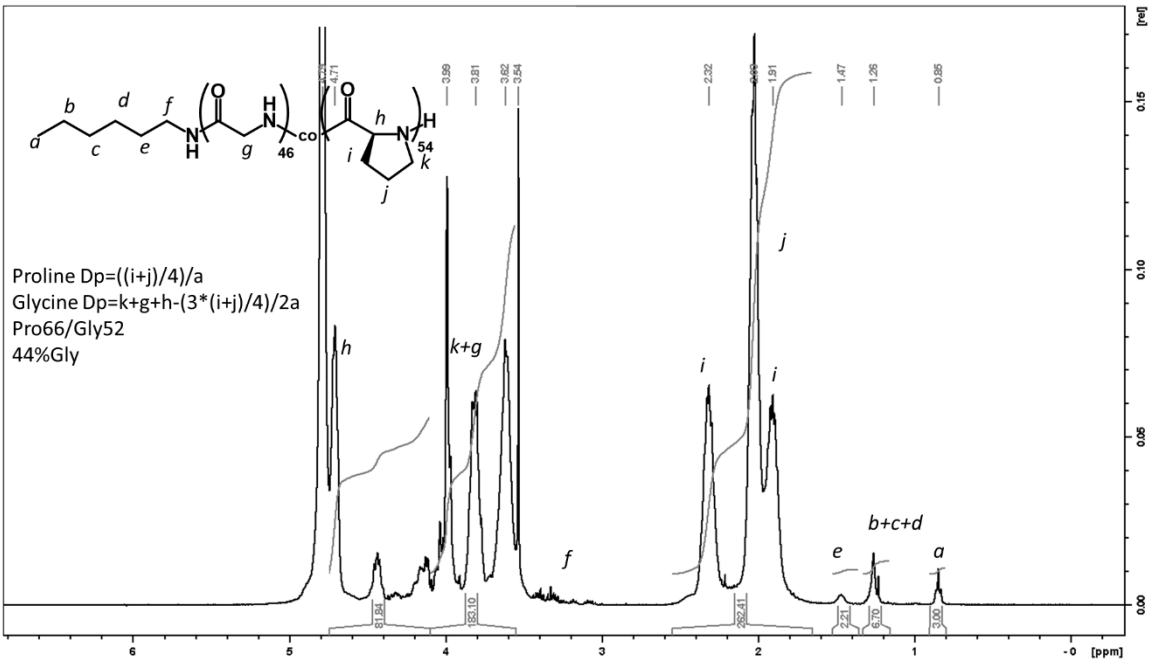


***Figure 4**: [1]H NMR of copolymer **3** in D_2O The copolymer structure is drawn with the peak assignation. The calculation of the Dp and glycine content is shown by the calculations from the integration of the initiator peak (a) and the respective backbone peaks of glycine (g) and proline (h,I,j,k).*

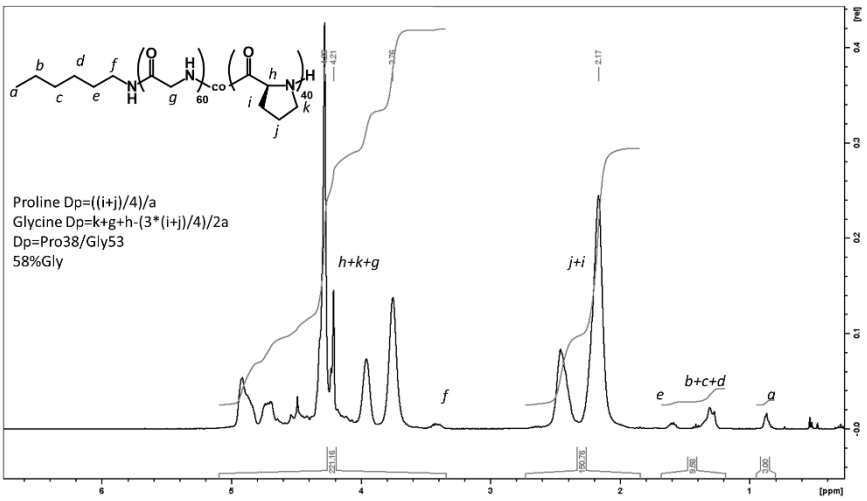


***Figure 5:** [1]H NMR of copolymer **4** in TFA-d. The copolymer structures is drawn with the peak assignation. The calculation of the Dp and glycine content is shown by the calculations from the integration of the initiator peak (a) and the respective backbone peaks of glycine (g) and proline (h,I,j,k).*

To confirm peak attribution, 2D NMR (HSQC and HMBC) analyses were also conducted on copolymer **6** (**Figure 6**) as this copolymer has the lowest M/I ratio, and thus the strongest initiator peaks. NMR data for all the copolymers is presented in **Table 1** and the spectra in the **Figure S1-Figure S2**.

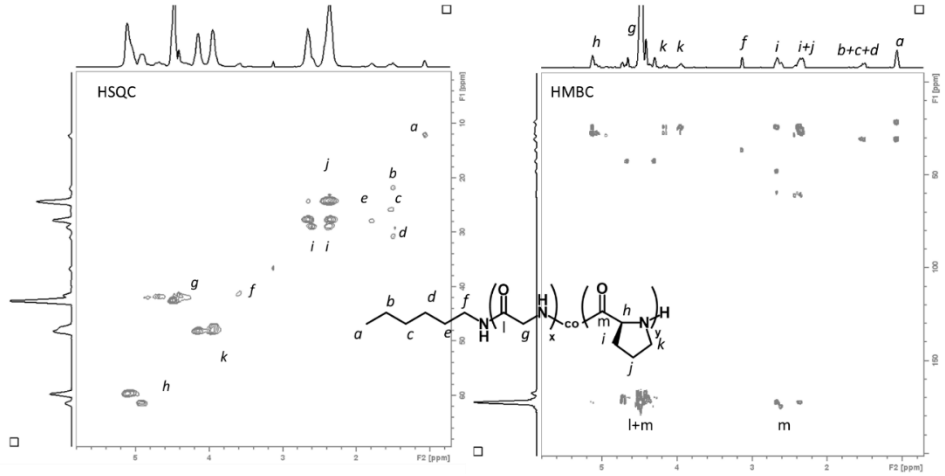


Figure 6: *H-C 2D NMR of copolymer* ***6*** *in D_2O. On the left HSQC spectra allows for the attribution of different carbon atoms attached to the protons previously determined. On the right HMBC spectra to confirm the correlation between neighbouring atoms to confirm the proline backbone peaks form the glycine ones. The latter also allows to detect carbon atoms from the corresponding carbonyls.*

SEC analyses of all the synthesized copolymers revealed the presence of high molecular weight species with narrow dispersities in each case. In both HFIP and aqueous SEC the Mn calculation was relative as the dn/dC used for the aqueous SEC conditions was that of PLP and the PMMA calibration gives relative values for the organic SEC conditions. Copolymer **3** was soluble in both eluents and gave comparable results in both SEC systems (**Figure 7**).

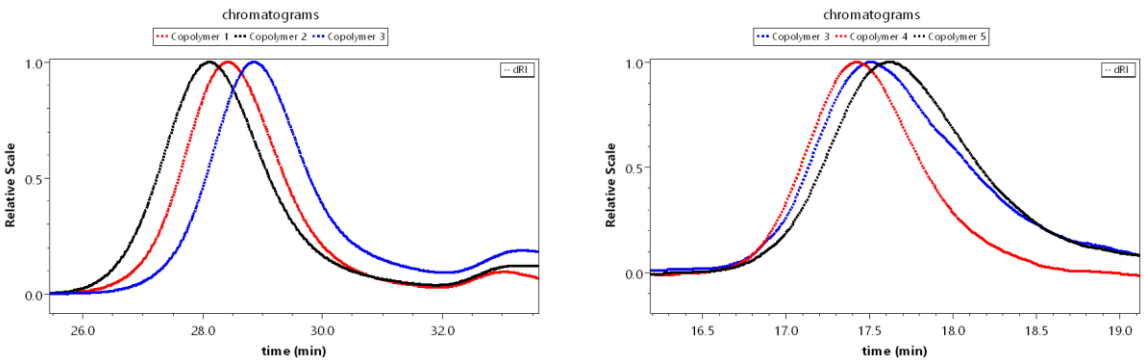


Figure 7: *SEC chromatograms for Copolymers* ***1-5****. On the left MALS chromatograms of copolymers* ***1-3*** *in ACN/acetate buffer SEC and on the right RI chromatograms of copolymers* ***3-5*** *in HFIP SEC.*

In SEC the elution time should decrease as the molecular weight of the copolymer chains increases. Nonetheless, chromatograms from **Figure 7** show that although copolymer **2** has lower molar mass than copolymer **1** (higher glycine content) it eluted

at a shorter time. Similar behavior was found for copolymer **3** and **4** (**Figure 7**). As the glycine content varies between these different copolymers, the hydrodynamic volume may be affected by the incorporation of glycine and not just be reflected by the change in molar mass. Nonetheless, the Mn values calculated for all these copolymers are similar (**Table 1**) When the glycine content was kept constant, the copolymers (**3, 6-8**) behaved as expected, where higher Mn copolymer eluted before lower Mn species (**Figure 8**).

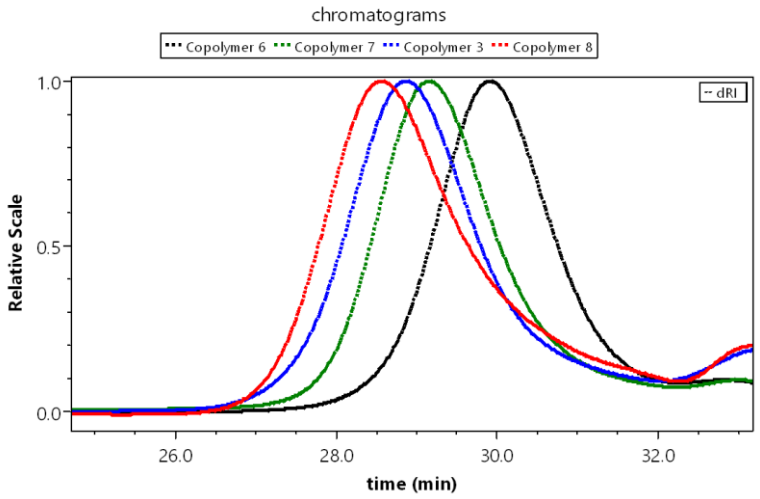


Figure 8: *SEC chromatograms for Copolymers **3** and **6-8**. RI chromatograms in ACN/acetate buffer SEC are shown.*

Calculating the absolute average molar masses would require values of dn/dC for each copolymer in its corresponding SEC media. Additionally, the HFIP system is not equipped with a MALS detector. Nonetheless, this copolymerization system was able to produce high molecular weight copolymers with low dispersity values and controlled monomer composition (**Table 1**).

Thermoresponsive behaviour

The thermoresponsive behaviour of copolymers having M/I=100 and increasing glycine content (copolymer **1-5**) were first examined to observe the influence of glycine on the solubility and the thermal properties. The copolymers were dissolved/dispersed in MilliQ water and analysed by UV-Vis turbidimetry as described in chapter **3**.

Only copolymers **1-2** with the lowest glycine contents (10% and 33%) showed hysteretic LCST behaviour (**Figure 9**). The calculated values of T_{CP} were 52°C, 48°C and 52°C for homoPLP, copolymer **1** and copolymer **2** respectively. The addition of

glycine units increased both the T_{CP} and the T_{CL} (**Figure 9**) while retaining the hysteretic behaviour of the homopolymer.

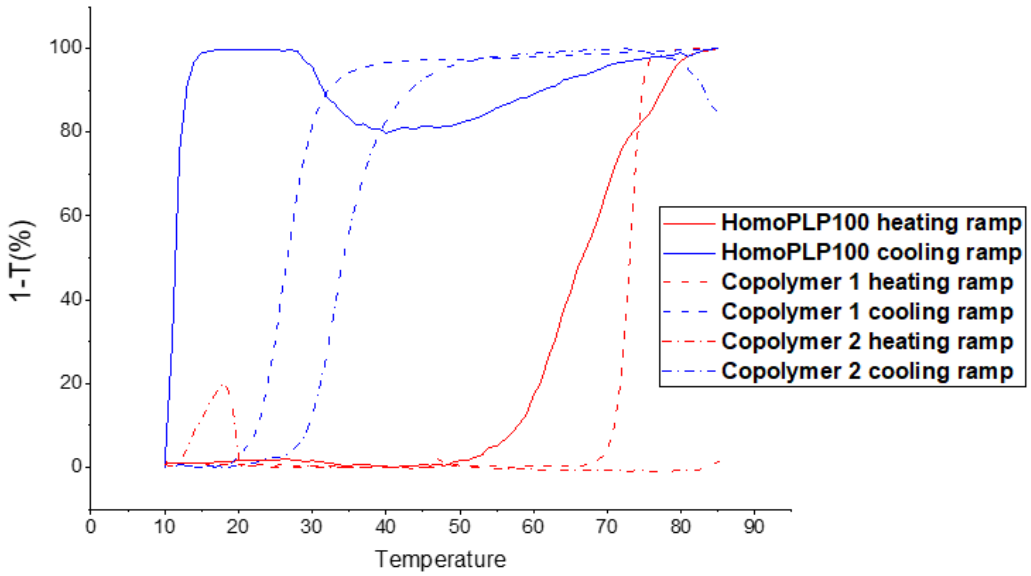


Figure 9: *UV-Vis turbidimetry curves for copolymers **1-2** and a homoPLP$_{100}$ for comparison as function of temperature at 5mg/ml in MilliQ water and 1°C/min temperature ramps. The samples were heated to 85°C and aggregated above their respective T_{CP} (sharp increase in absorption, red curves). After a 10 min isotherm, the samples were cooled down and the T_{CL} was only observed at a lower temperature (sharp decrease in absorbance, blue curves) during the cooling process. Absorbance values collected by the machine were converted to transmittance and normalized for a better visualization of the degree of turbidity.*

This is attributed to the replacement of the hydrophobic side chains of proline units by hydrogen atoms. The incorporation of more hydrophilic units increased the transition temperatures as the copolymers become more soluble in water.[55] The observed width of hysteresis did not change considerably as both T_{CL} and T_{CP} increased. Increasing the glycine content to 50% (copolymer **3**) increased the T_{CP} above the experimental temperature window and the copolymer was completely soluble without any thermoresponsiveness (**Figure 10**). Further increases in glycine content resulted in completely insoluble copolymers (**4**-**5**) as the amount of soluble PLP did not counterbalance the amount of insoluble PGly sequences. High glycine contents resulted in lower soluble PLPII content that is also responsible for thermoresponsive behaviour.

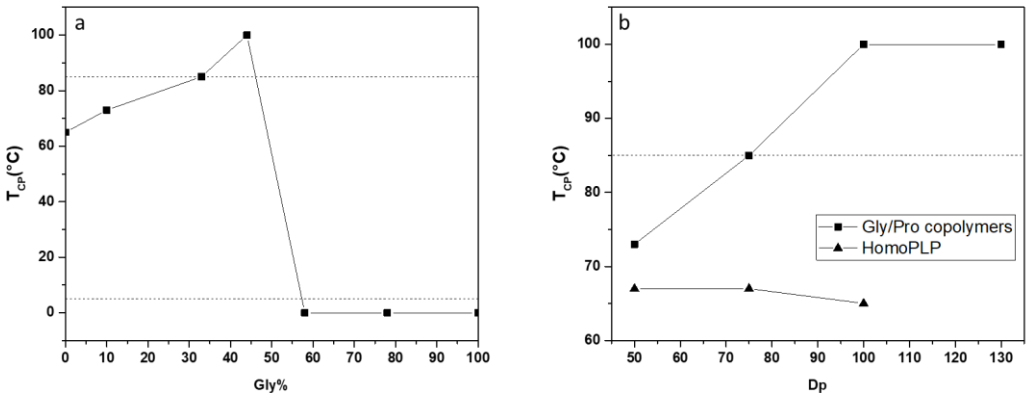


Figure 10: *a) T_{CP} values of homoPLP and copolymers with glycine in function of the glycine content. b) T_{CP} values of homoPLP (triangles) and copolymers having 50% glycine content (squares) in function of their overall Dp. Dashed lines correspond to the experimental temperature window. Points above (at 100°C) and below (at 0°C) this window were soluble and insoluble respectively at all data points. Data was determined using UV-Vis turbidimetry at 5mg/ml in milliQ water.*

Changing the overall Dp may also affect the thermoresponsive behaviour. As shown in **Figure 10b**. decreasing the Dp (from a value of 75 to a value of 50 for copolymers **6-7**) allowed for the observation of a T_{CP} that decreased with molecular weight (from a value of 85 to a value of 73). This behaviour was not observed for homopolymers of proline in this window (**Figure 10b**). Turbidimetric analysis of copolymer **6** and **7** by UV-Vis showed that they were slightly soluble in aqueous solution at 5 mg/ml. This was further confirmed by the lower transmittance values at lower temperatures (1-T= 40-60%, **Figure 11**). Nonetheless, the suspended particles still showed thermoresponsive behaviour by aggregating above their $T_{CP,}$ 55°C and 75°C respectively (**Figure 11**).

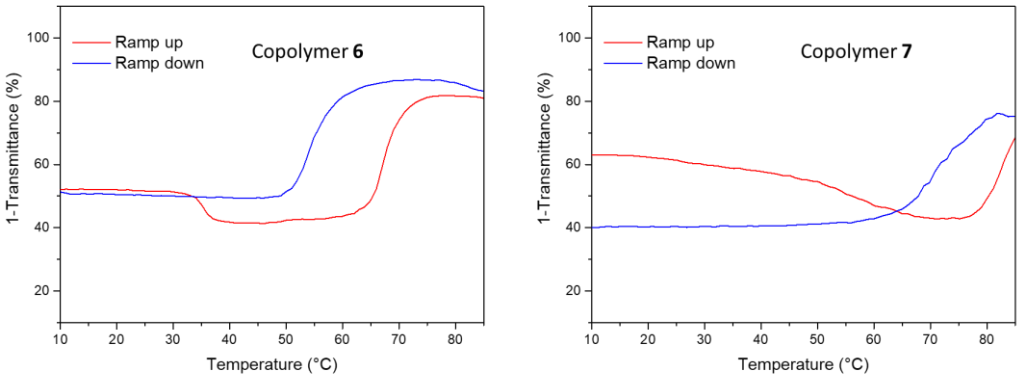


Figure 11: *UV-Vis turbidimetry curves for copolymers 6-7 as function of temperature at 5 mg/ml in MilliQ water and 1°C/min temperature ramps. The samples were heated to 85°C and aggregated above their respective T_{CP} (sharp increase in absorption, red curves). After a 10 min isotherm, the samples were cooled down and the T_{CL} was only observed at a slightly lower temperature (sharp decrease in absorbance, blue curves) during the cooling process. Absorbance values collected by the machine were converted to transmittance without further normalization.*

When cooled back down, a T_{CL} could be observed at a close temperature to the measured T_{CP} and the copolymers regained their original suspended state (**Figure 11**). To confirm these results, DLS measurements were done on the same solutions to better analyse the state of aggregation at both low and high temperatures and presented in **Figure 12**. **Figure 12** shows results obtained by DLS that confirm the LCST behaviour of copolymers **6-7** observed by other turbidimetric techniques (eye and UV-Vis). The transition temperatures T_{CP} *and* T_{CL} are of the same value for each copolymer (55°C and 75°C respectively) meaning that there was very little hysteresis between the aggregation and the solubilization step with those copolymers. Aqueous solutions of the copolymers **6-7** at temperatures lower than T_{CP} and T_{CL} showed the presence of nanoparticles with a hydrodynamic diameter of around 100 nm and low PDI values (**Figure 13**). Above these temperatures, the nanoparticles aggregated at the micrometric scale (and therefore were not easily analysed by LS).

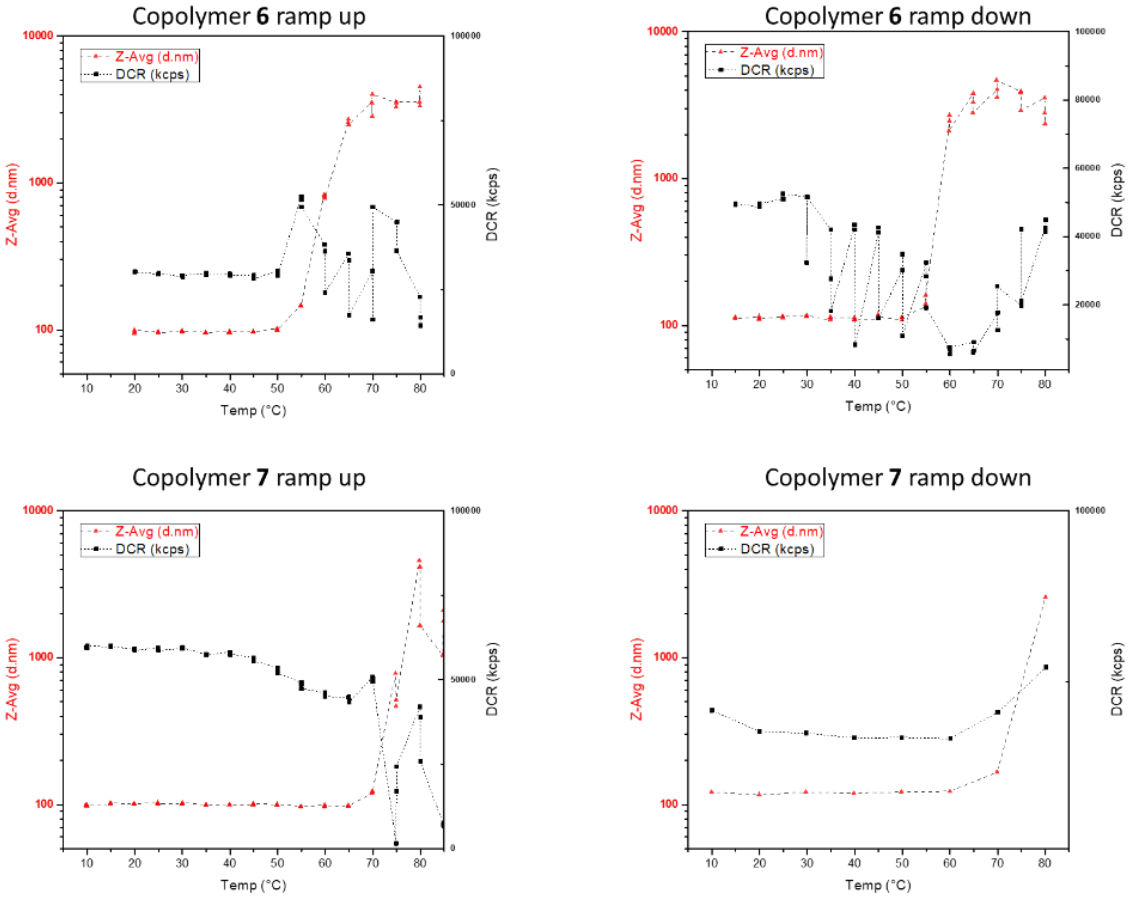


Figure 12: *DLS temperature ramps for copolymers* ***6-7****. 3 measurements were done for each point and the spacing of the overall ramp point was adjusted to have similar rime to 1°C/min in UV-Vis measurements. On the right y-axis the Z-average diameter values is shown (nm) in a logarithmic scale and on the left y-axis the derived count rate.*

TEM analysis of copolymer solutions at low temperatures showed the formation of wormlike nanoparticles (**Figure S11**). These wormlike structures presented diameters lower than that shown by DLS since the DLS model assumes spherical objects (Malvern analysis at an angle of 173°). This behaviour was also observed in the laboratory for block copolymers of PLP-*b*-PGly in the scope of other projects (Hannah Beauseroy and ROPISA).

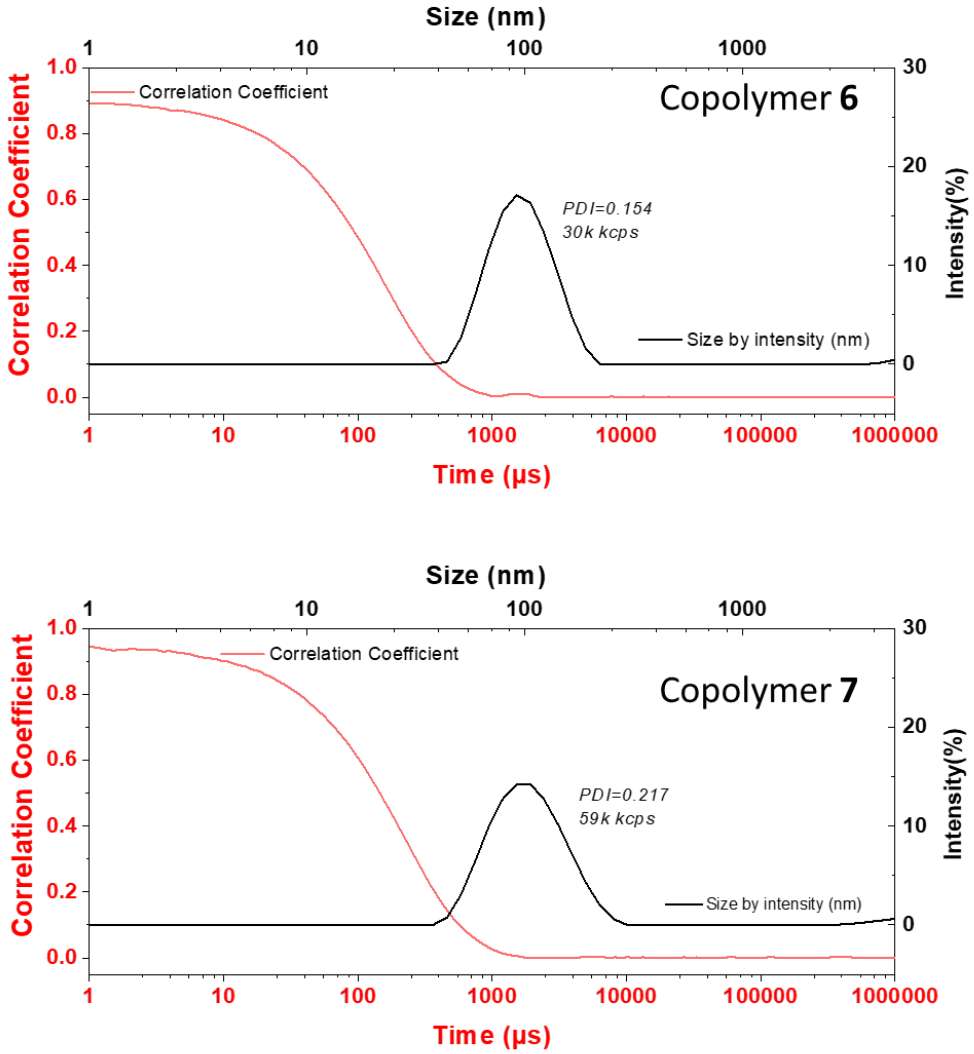


Figure 13: *DLS measurements of copolymers* ***6-7*** *at 5 mg/ml in MilliQ water at room temperature before temperature ramps. The correlation coefficient is shown in red and the signal intensity by size is shown in black.*

Copolymerization kinetics

To shed light on the microstructure of the copolymer, the kinetics of a copolymerization system with 50% glycine content was monitored by ^{1}H NMR (**Figure 14**). The water assisted ROP does not allow to follow the kinetics using the method developed earlier in Chapter **2**, as the ROP is much faster and proceeds in a heterogenous manner. A copolymerization using a PEG-NH_2 (2kg/mol) macroinitiator was conducted and aliquots from this reaction were quenched at different times using HCl to stop the propagation. The use of acid quenching ensures that the polymerization was stopped by producing inactive ammonium chain ends and concomitantly, the hydrolysis of the unreacted NCAs to their corresponding amino acids.[62] Each aliquot was then dialyzed

to remove unreacted amino acids and the remaining copolymer was analysed by ^{1}H NMR in TFA-d to determine the glycine incorporation at different reaction times (**Figure 14**).

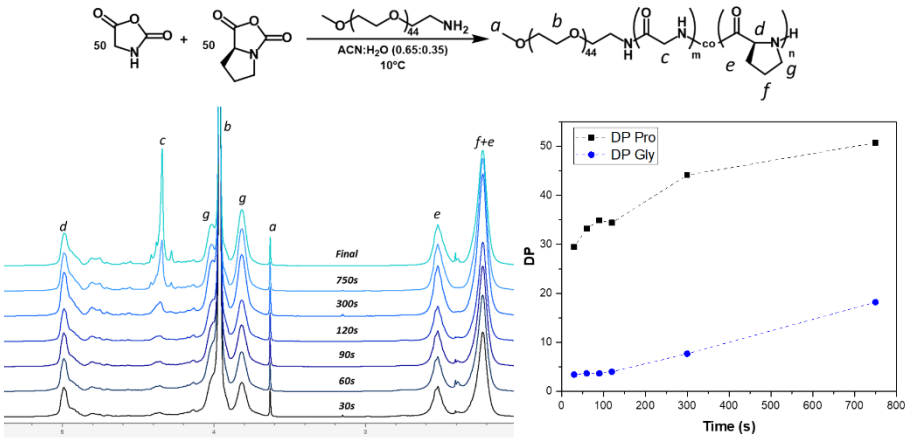
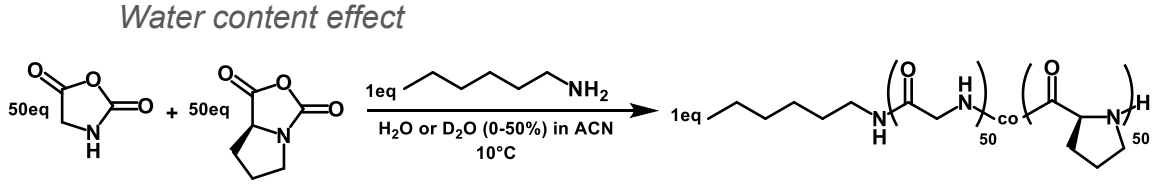


Figure 14: *(Top) the reaction scheme of the copolymerization system of equimolar amounts of glycine NCA and proline NCA initiated by PEG-NH_2 (2kg/mol). (Bottom left) the staggered ^{1}H NMR spectra of quenched aliquots at different reaction times from 30s to 750s as well as the final copolymer are shown. (Bottom right) the DP of each comonomer in function of time of reaction is represented. This data is based on the integration from ^{1}H NMR spectra.*

These results confirmed the formation of block-like copolymers of proline and glycine. In this case proline polymerised much faster than glycine to generate a copolymer with 30 units of proline with only one or two units of glycine. The copolymerization then proceeded by incorporating mostly proline units reaching close to 70% proline conversion before the glycine peak started to increase significantly in intensity. As discussed in the previous chapter, water acts as a catalyst during ROP and drives the formation of the PLPII helix that increased the proline homopolymerization.[52] This same behaviour was thus present during the copolymerization and was certainly at the origin of the result we observed. As the polymerization begins, the initiator has equal chances to react with a glycine or a proline monomer as they are present in equimolar quantities. But when proline units start to polymerise, the water-induced acceleration kicks in and privileges the formation of the PLP homosequence. Only when the feed ratio of proline drops, do glycine units start to be incorporated in the copolymer. This effect resulted in highly gradient, block-like copolymers that formed nanoparticles at smaller DP since PGly sequences are insoluble at lower DP. At higher molecular

weights, the long soluble PLP block renders the overall copolymer soluble. These conclusions can also be confirmed by the observation of the copolymerization process that proceeds in a homogeneous manner once the water and the initiator are added but then becomes turbid after a prolonged time due to the formation of the polyglycine block (**Figure S12**).



Scheme 5 : *Copolymerization of glycine NCA and proline NCA using hexylamine as initiator via the water assisted ROCOP. This series of copolymers (**9-14**) was tested keeping a equimolar ratio of both comonomers and an M/I=100. The water content and nature were changed during the copolymerization process to analyse the effect water would have on the copolymerization kinetics.*

As discussed in the previous chapter **3**, the ROP of Pro-NCA is catalysed by the presence of water due to the formation of the PLPII helix and its specific affinity towards this solvent. In the previous part it was shown how the same catalytic effect accelerates the polymerization of homopolyproline sequences due to the driving force of its secondary structure but had less effect on the polymerization of glycine producing gradient copolymers. To further study the influence of this catalytic effect during copolymerization with glycine, a series of copolymers were synthesized varying both the content and the isotopic content of the water (**Scheme 5**). Copolymers **9-14** were synthesized using the same procedure described above, having a Gly% of 50 and a M/I=100. The water content was decreased from 50% for copolymer **9** to 0 for copolymer **14** (**Table 2**). Copolymer **3** with water content of 35% can be compared to these copolymers. Finally, copolymer **10** was synthesized in an acetonitrile/D2O mixture (1:1) to analyse the hydrogen bonding effect on the copolymerization.

Table 2 : *50/50 Gly-Pro copolymers* ***9-14*** *synthesis conditions and analysis. All copolymers were initiated by hexylamine at M/I ratio of 100*

	^{1}H NMR [a]	H_2O%
Copolymer **9**	HA-Pro_{56}-Gly_{55}	50
Copolymer **10**	HA-Pro_{50}-Gly_{49}	50 [b]
Copolymer **11**	HA-Pro_{54}-Gly_{52}	20
Copolymer **12**	HA-Pro_{49}-Gly_{48}	10
Copolymer **13**	HA-Pro_{45}-Gly_{32}	2
Copolymer **14**	HA-Pro_{44}-Gly_{37}	0

a. as determined by the integration of the proline and the glycine backbone peaks and hexylamine peaks in ^{1}H NMR in TFA-d. b. D_2O was used.

Both copolymers **9** and **10** had comparable glycine content to theoretical feed and to copolymer **3** (**Table 2**). Nonetheless, copolymer **9** (H_2O) had slightly lower molar mass and higher dispersity when analysed with SEC (**Figure 15**), probably due to increased hydrolysis.[52] Copolymer **10** (50% D_2O) showed a molar mass closer to copolymer **3** and lower dispersity (**Figure 15**)**.** This is due to different interactions with the deuterium isotope.

The water content was then reduced to 20% and 10% for copolymers **11** and **12** respectively (**Table 2**). Both copolymer dispersities close to copolymer **3** (**Figure 15**) but copolymer **11** had slightly higher Mn value. Both these copolymers formed aggregating gels during the copolymerization process as the copolymers were not easily dispersible in organic media compared to copolymers **3**, **9** and **10** (**Figure S13**) which led to decreased final yield. When the water content was further decreased, the copolymerization proceeded in a heterogenous manner (**Figure S13**) and the resulting copolymers had lower glycine content than targeted (**Table 2**).

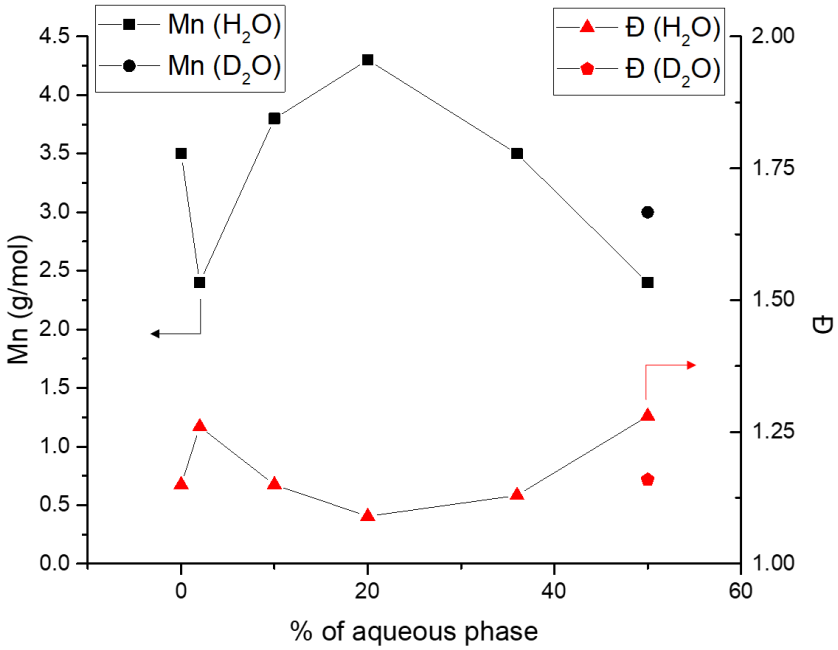


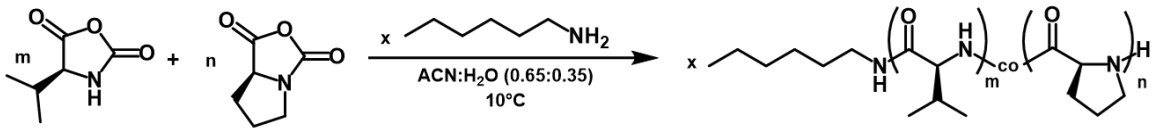
Figure 15 : *SEC results of copolymer **3** and **9-14** in function of water content during the synthesis. On the left-hand side y-axis, the Mn was determined by* PEG *calibration in ACN/acetate buffer and RI detection as some polymers caused high aggregation in and Mn determination using MALS was not possible. On the right-hand y-axis, the dispersities of the polymer peaks as determined by the RI detection are represented.*

The SEC results for copolymers **13-14** (2% and 0% H_2O respectively) showed lower molar mass species that may be due to water initiation as the polymerization proceeds more slowly (**Figure S14**). The resulting copolymers had similar final structures when the water content was kept above 10% as analysed by NMR and differed slightly using SEC analyses. Nonetheless, all copolymers were completely soluble in MilliQ water and had no thermoresponsiveness similarly to copolymer **3**. As all copolymers had an M/I of 100, similar conclusions can be made: when high glycine content and overall molar mass increased the T_{CP} above the experimental window, no LCST could be observed. Furthermore, this showed that the water content did not influence the microstructure of the copolymers and the catalytic effect of the water was the same in all cases.

2.1.2. Proline-Valine copolymers

Synthesis and analysis

Proline NCA was next copolymerized with valine NCA (**Scheme 6**), the third amino acid of ELPs. Valine has a hydrophobic side chain, and it is expected that by incorporating it in a PLP chain the transition temperatures might decrease due to the increased local hydrophobicity. For this purpose, Val NCA and Pro NCA were copolymerised similarly to the procedure explained earlier for the copolymerization with Gly NCA.



Scheme 6 : *Copolymerization of valine NCA and proline NCA using hexylamine as initiator via the water assisted ROCOP. A series of copolymers was explored by changing the m/n ratio with the same m+n/x ratio (100).*

Two copolymers having 10 and 20% valine content were first synthesized and analysed. As shown in **Table 3**, both copolymers had lower valine content than theoretical feed as calculated by ^{1}H NMR. SEC analysis in aqueous conditions show high molecular weight polymers with low dispersities. The analysis was done with the dn/dC value of homoPLP and thus the molar masses are only indicative.

Table 3 *: Analysis and results for copolymers* ***15-16*** *synthesis*

	Targeted structure (Theoretical Mn in g/mol)	Yield	^{1}H NMR [a]	Val%	Mn by SEC[b] Đ[c]
Copolymer **15**	HA-Pro$_{90}$-Val$_{10}$ (9830)	78%	HA-Pro$_{99}$-Val$_{8}$	8	7500 g/mol 1.03
Copolymer **16**	HA-Pro$_{80}$-Gly$_{20}$ (9850)	70%	HA-Pro$_{93}$-Val$_{13}$	12.5	7900 g/mol 1.02

a. determined by the integration of the proline and the valine backbone peaks and hexylamine peaks in ^{1}H b. determined by dn/dc of polyproline (0.1414) in ACN/acetate buffer SEC and MALS detection c. determined by SEC peaks.

Furthermore, peaks from valine backbone (peak e in **Figure 16**) were not easily identified in NMR analysis and had high multiplicity. This could be indicative of small valine sequences with different electronic environments (beginning of the chain, chain end…).

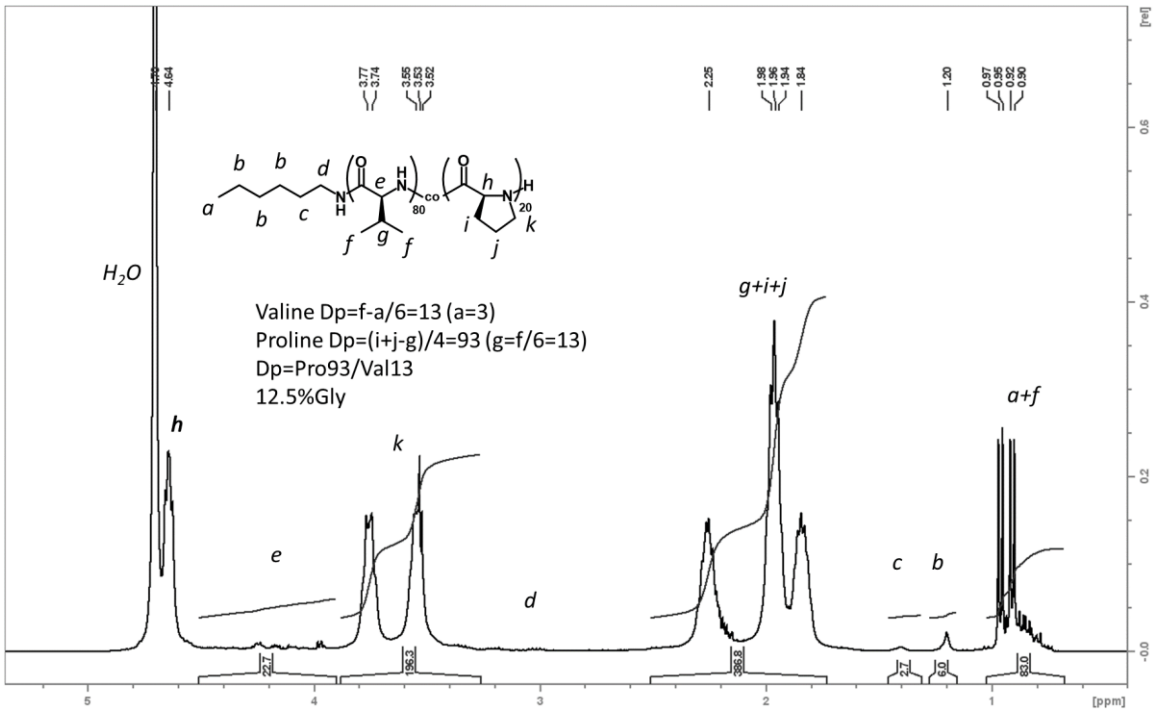


Figure 16 *: [1]H NMR for copolymer **16** in D_2O and its theoretical structures is drawn with the peak assignation. The calculation of the Dp and valine content is shown by the calculations from the integration of the initiator peak (b) and the respective backbone peaks of valine (f,g) and proline (i,j).*

The conversion of valine was not always complete as less valine was incorporated than expected. This is also reflected in the lower yields and lower molar masses as calculated by SEC. The formation of insoluble β-sheet from homopolyvaline sequences makes the ROP of valine highly difficult.[63–65] The thermoresponsive copolypeptides based on the copolymerization of valine NCA and glycine NCA mentioned in the introduction,[61] were of smaller DP (40) and were not characterized by SEC. Furthermore, their LCST behaviour showed very slight decrease in transmittance, which could mean a high interchain heterogeneity. It was thus concluded that this copolymerization, if proceeding in a highly gradient manner like in the case of the Pro/Gly system presented before, will not easily yield copolymers with higher valine content.

Thermoresponsive behaviour

Copolymers **15** and **16** having 8 and 12.5% valine content were analysed by turbidimetry to study their LCST behaviour.

Table 4 : *Results from turbidimetric UV-Vis analysis of Pro/Val copolymers compared to homoPLP. All polymers have a theoretical M/I of 100.*

	Val%	T_{CP}	T_{CL}	X_H
$HomoPLP_{100}$	0	63°C	13°C	50°C
Copolymer **15**	8	73.5°C	10.5°C	63°C
Copolymer **16**	12.5	75.5°C	18.5°C	57°C

Table 4 shows the transition temperatures (T_{CP} and T_{CL}) as well as the hysteresis width (X_H) obtained using UV-Vis. With comparison to the corresponding homopolyproline, there is no observable trend except for the increase in T_{CP} with increasing the valine content. This phenomenon is like the one observed with increasing glycine content. Although adding a hydrophobic monomer the polymers are more soluble since the transition temperatures are increasing. Furthermore, the hysteresis width is larger than that of the homopolymer but seems to decrease with more added valine.

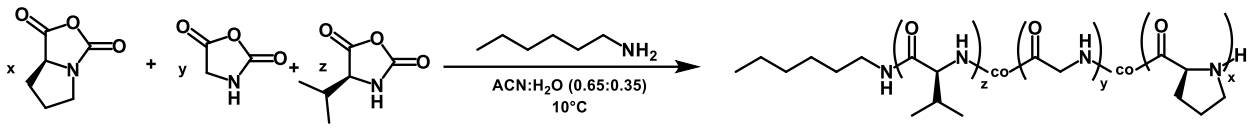
These results indicate that proline NCA polymerizes much more rapidly than the valine comonomer which leads to degradation of the latter and lower yields as well as lower final valine content. The small change in thermoresponsive behaviour also confirms similar results to the copolymerization with glycine producing copolymers of high gradient nature.

2.1.3. Terpolymerization

ELPs comprise the three amino acids described earlier: glycine, valine and proline. As presented, proline's unique structure and peptide bond gives rise to the LCST behaviour. The aqueous ROP developed for the synthesis of high molecular weight polyprolines with a controlled structure was also successful in producing high molecular weight copolymers based on proline with either glycine or valine as shown in sections **2.1.1** and 2.1.2 respectively. The analysis of the structures, kinetics and thermoresponsiveness of binary systems comprising of proline on one hand and glycine or valine on the other hand showed high gradient structures as proline polymerized much faster in these conditions. For this reason, the thermoresponsive behaviour of synthesized copolymers with high proline content and high molecular weight showed large hysteresis (>40°C). When the glycine content was closer to 50%, glycine/proline copolymers were either completely soluble or completely insoluble. In this next part, tertiary systems based on these results and made in aqueous conditions will be explored.

Synthesis and NMR analysis

As discussed in the introduction to this chapter, a generic ELP consists of a repeated VPGXG sequence. Thus, in ELPs, proline represents only 20% of the whole polypeptide, as the X residue is never a proline. Based on this and the previous results a series of terpolymers were synthesized using water-assisted ROP (**Scheme 7**). The aim was to produce statistical copolymers with similar overall composition as ELP sequences to mimic their thermoresponsive behaviours.



Scheme 7 : *Terpolymerization of proline NCA, glycine NCA and valine NCA using hexylamine as initiator via the water assisted ROP. A series of copolymers was explored by changing the different feed ratios of comonomers (x,y and z respectively).*

Different proline/glycine/valine ratios were tested keeping the proline content lower than 50%. Terpolymers **1** and **2** were first synthesized with the generic structure of ELPs (VPGGG) in mind (20% valine and 20% proline), having an M/I of 50 and 100 respectively. Synthesis of a series of terpolymers (**Table 5**) all having an M/I=100 (except Terpolymer **1**) with varying monomer ratios were done. The terpolymerizations were conducted in a similar manner to the copolymerization described above. After purification, resulting yields were satisfactory and all terpolymers had overall DP comparable to theoretical values as characterized by ^{1}H NMR. (**Table 5**).

Table 5*: Analysis and NMR results for terpolymers **1-8** synthesis.*

	Targeted structure	Pro%[a]	Gly%[a]	Val%[a]	Yield
Terpolymer **1**	HA-Pro$_{10}$-Gly$_{30}$-Val$_{10}$	20	60	20	88
Terpolymer **2**	HA-Pro$_{20}$-Gly$_{60}$-Val$_{20}$	20	63	17	84
Terpolymer **3**	HA-Pro$_{20}$-Gly$_{50}$-Val$_{30}$	21	52	27	79
Terpolymer **4**	HA-Pro$_{30}$-Gly$_{50}$-Val$_{20}$	30	51	19	84
Terpolymer **5**	HA-Pro$_{40}$-Gly$_{50}$-Val$_{10}$	40	53	7	79
Terpolymer **6**	HA-Pro$_{47}$-Gly$_{46}$-Val$_{7}$	48	48	4	80
Terpolymer **7**	HA-Pro$_{43}$-Gly$_{42}$-Val$_{15}$	45	44	11	81
Terpolymer **8**	HA-Pro$_{40}$-Gly$_{40}$-Val$_{20}$	41	43	16	78

a. determined by the relative integration of the different backbone peaks in ^{1}H NMR in TFA-d.

To correctly determine the DP and the terpolymer composition, 2D NMR was conducted in order to segregate the different peaks building on the peak assignments

already obtained for binary systems. As all copolymers had the same structural basis (HA-Pro$_X$-Gly$_Y$-Val$_Z$), terpolymer **3,** having sufficient content from each monomer (20/50/30), was analysed by combining HMBC and HSQC (**Figure 18**).

In **Figure 18,** it is possible to identify ^{1}H NMR peaks for each repat unit in the backbone and their side chains' peaks. For example, the α-carbon region (4.3ppm to 5.5ppm) shows the overlapping of the three peaks i, e and h (Pro, Val and Gly respectively). Nonetheless, side chain peaks for proline (l) and valine (f) has chemical shifts in secluded regions close to the initiator peaks. As presented in **Figure 17**, by separating the valine integration around 1ppm, it was possible to determine the ratio of valine in comparison to the initiator peaks (when visible). The different proline side chain peaks (j and k) were then integrated easily, and the valine subtracted (g). Finally, a similar calculation was done for glycine's peak (h) as it falls in a zone with several peaks from the other monomers (e, i and l).

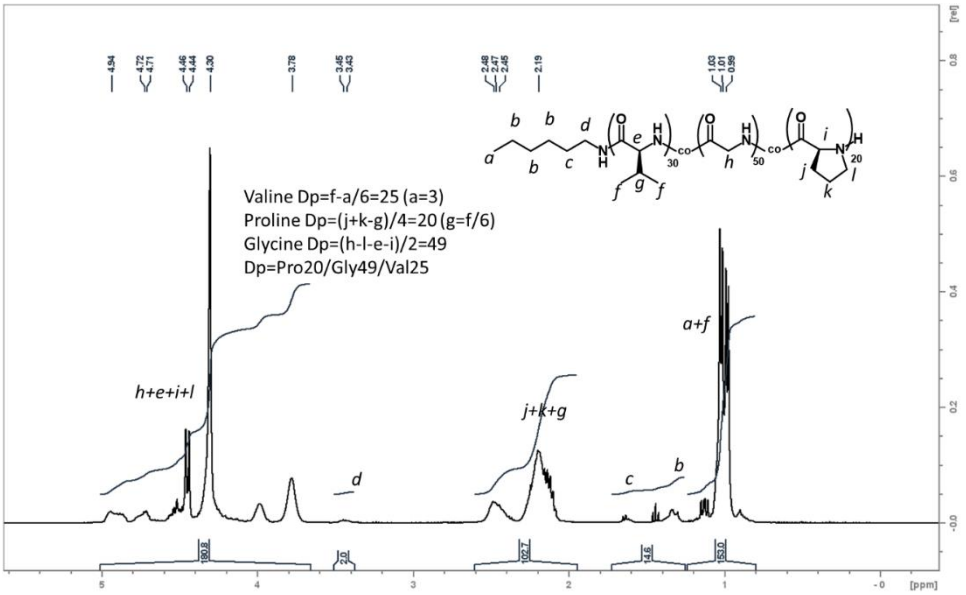


Figure 17 : *^{1}H NMR of terpolymer **3** in TFA-d. Its theoretical structures is drawn with the peak assignation. The calculation of the different Dp is shown by the calculations from the integration of the initiator peak (d) and the respective backbone peaks of valine (f), proline (j, k) and glycine (h).*

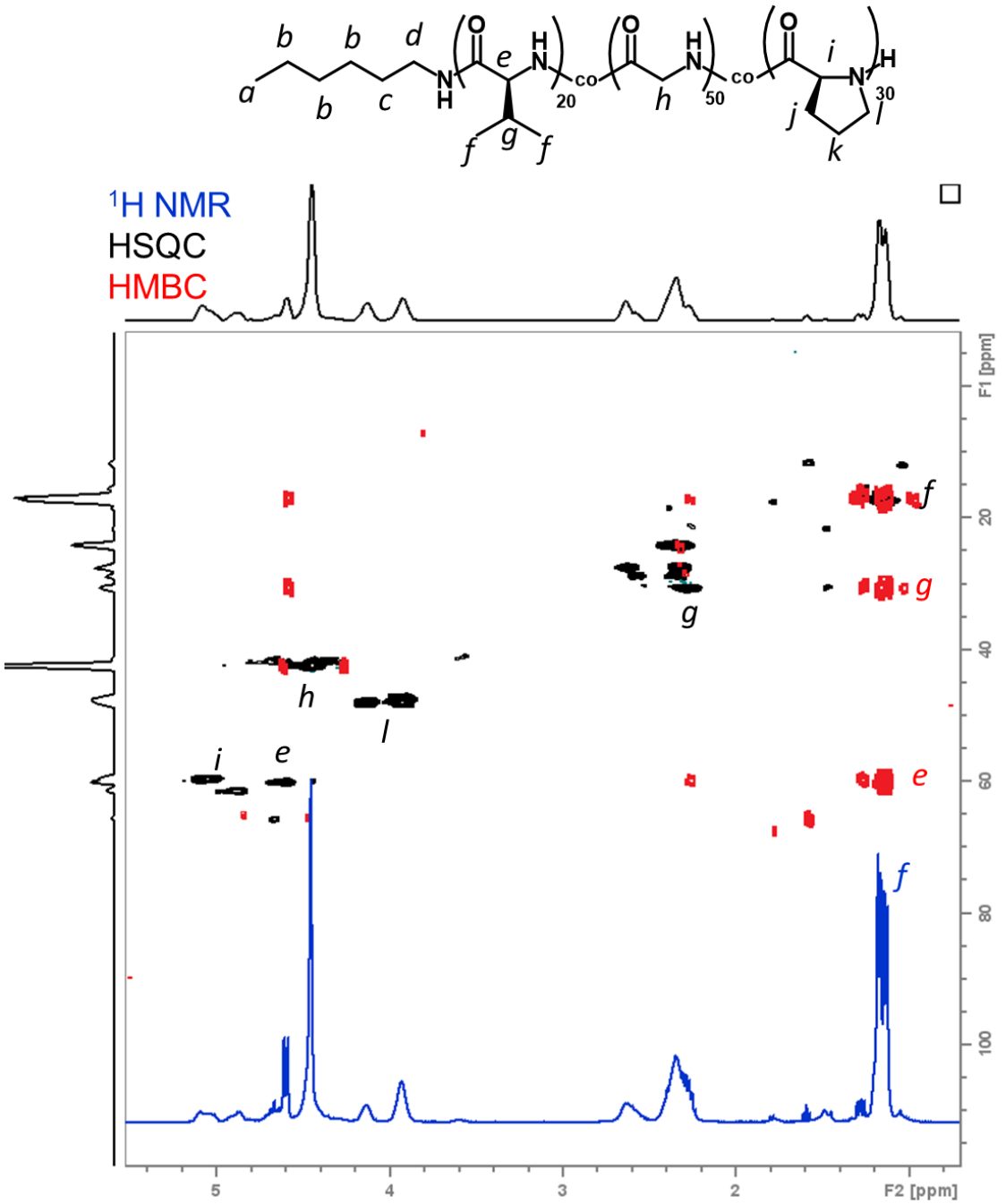


Figure 18 : *On the top, scheme of terpolymer **3**, on the bottom ^{1}H NMR (in blue), HSQC (in black) and HMBC (in red) spectra in TFA-d. Different peak assignments are noted on the spectra in the corresponding colours.*

All terpolymers were thus analysed by ^{1}H NMR (**Figure S4-6**) and the results are presented in **Table 5**. The experimental values of the terpolymers content were in good agreement with the theoretical ones.

SEC analysis

SEC analysis of the different terpolymers were conducted in HFIP at 5 mg/ml.

Table 6: *SEC results for terpolymers* ***1-8.***

	Structure by NMR	Mn by SEC (kg/mol)[a]	Đ[b]
Terpolymer **1**	HA-Pro$_{10}$-Gly$_{30}$-Val$_{10}$	13.2[c]	1.15
Terpolymer **2**	HA-Pro$_{20}$-Gly$_{63}$-Val$_{17}$	18.2[c]	1.13
Terpolymer **3**	HA-Pro$_{21}$-Gly$_{52}$-Val$_{27}$	17.6[c]	1.14
Terpolymer **4**	HA-Pro$_{30}$-Gly$_{51}$-Val$_{19}$	17.8[c]	1.12
Terpolymer **5**	HA-Pro$_{40}$-Gly$_{53}$-Val$_{7}$	17.4	1.18
Terpolymer **6**	HA-Pro$_{48}$-Gly$_{48}$-Val$_{4}$	15.7	1.17
Terpolymer **7**	HA-Pro$_{45}$-Gly$_{44}$-Val$_{11}$	16.1	1.14
Terpolymer **8**	HA-Pro$_{41}$-Gly$_{43}$-Val$_{16}$	19.7	1.10

[a]determined by PMMA calibration in HFIP SEC and RI detection; [b]determined by SEC peaks; [c]presence of a second peak with higher molecular weight.

As shown in **Table 6**, all terpolymers presented high molecular weight peaks with low dispersity. The calculated molecular weights were relative to the PMMA calibration and permitted to conclude that the terpolymers had overall comparable Mn values. Nonetheless, most terpolymers showed a second peak with much higher molecular weight even though they were all visibly soluble and passed through the syringe filter.

For instance, both terpolymers **1** and **2** showed a second peak when passed through the SEC in HFIP (**Figure 19**). The second peak at lower elution time, was still visible after diluting the samples by a factor of 2. The relative heights of both peaks did not change meaning that the higher molecular weight peak was not due to aggregation but was a second polymer species in the sample.

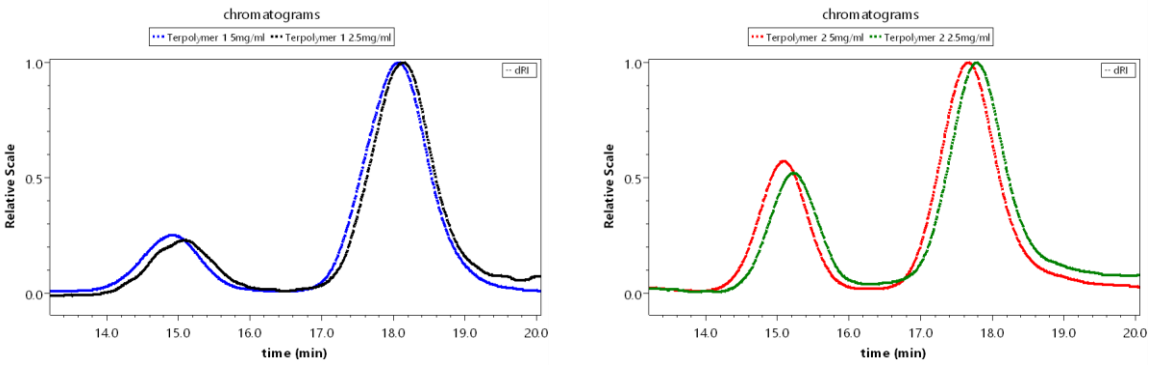


Figure 19: *HFIP SEC chromatograms of terpolymer* ***1*** *(on the left) and terpolymer* ***2*** *at 5mg/ml and 2.5mg/ml. RI detection is shown, and peaks are normalized for relative scaling.*

This suggests either physically or temporally separate polymerizations. Many hypotheses can be made: 1- slowly reactive valine NCAs that get hydrolysed and

initiate new polymer chains; 2- long valine sequences in some copolymer chains stop the reaction as they precipitate completely. Both these hypotheses need further investigation to be verified, but other characterizations of these complex terpolymers might be time consuming (MALDI-TOF…).

Valine to proline ratio was then decreased hoping at obtaining terpolymers one polymeric specie. Terpolymers **3**, **4** and **5** were then synthesized by keeping the glycine content at 50% and changing the valine to proline ratio progressively (**Table 6**). Valine content was decreased from 27 to 7 and the composition was verified by ^{1}H NMR (**Table 5**). SEC results show that while terpolymer **3** (27%Val and 21%Pro) and **4** (19%Val and 30%Pro) had a higher molecular weight species at the same ratio for two different concentrations (**Figure S16**), terpolymer **5** (7%Val and 40%Pro) had no aggregated species (**Figure 20**). This result showed promise in the possibility of producing terpolymers based on Pro, Gly and Val NCA with controlled structure.

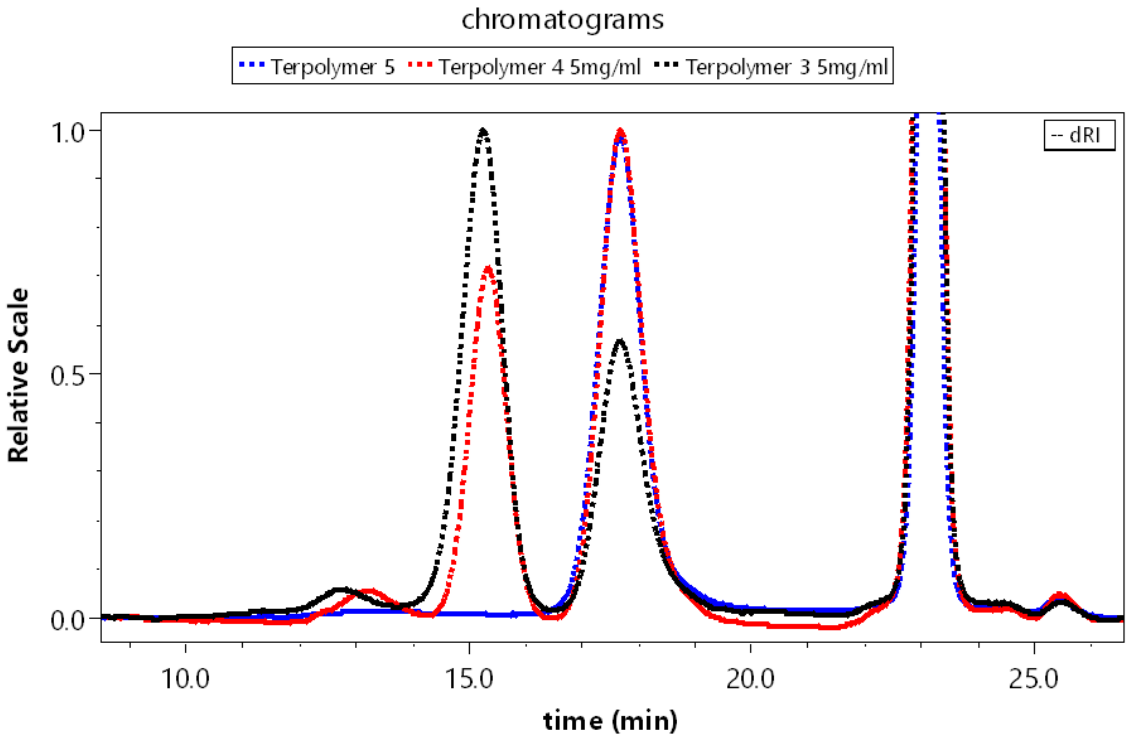


Figure 20: *HFIP SEC chromatograms of terpolymers **3-5** at 5mg/ml. RI detection is shown and peaks are normalized for relative scaling.*

These results showed that high valine content and low proline content lead to the formation of HFIP soluble high molecular weight species. It is also possible that these species are polymer chains with concentrated valine sequences leading to slightly

soluble beta sheet structures. Such interchain heterogeneity was previously observed for copolymers containing valine[64] and needs further investigation.

As terpolymers **1**-**5** were insoluble in water, terpolymers **6-8** were synthesized having equimolar amounts of glycine and proline and increasing valine content up to 16%. This series could be compared to water soluble copolymer **3** (HA-Pro_{50}-Gly_{50}-Val_0) where valine could introduce the hydrophobicity necessary for the LCST. Their structure was verified using ^{1}H NMR to confirm the terpolymers' composition (**Figure S6**-**Figure S7**-**Figure S8**) and the results obtained corresponded to theoretical values (**Table 5**). All three terpolymers showed high molecular weight and controlled structures without detectable aggregation when analysed by HFIP SEC (**Table 6** and **Figure S17**). Therefore, terpolymerization of glycine NCA, valine NCA and proline NCA successfully yielded high molecular weight polymers with low dispersities and controlled composition using the water assisted ROP (terpolymers **6-8**).

Solubility and thermoresponsiveness

Terpolymers **1-8** were further analysed to assess their water solubility and potential thermoresponsive behaviour. The terpolymers all had a proline content lower than 50% and were all insoluble in water without any thermoresponsive behaviour. A phase diagram of the studied system comprising of the two studied binary systems (Pro-Gly and Pro-Val) copolymers as well as terpolymers **1-8** is represented in **Figure 21**. The copolymerization of proline NCA with either valine NCA or glycine NCA increases the T_{CP}. This suggests that both monomers are less hydrophobic than proline. Although valine has a hydrophobic side-chain (isopropyl) the pyrrolidine side chain of proline with similar alkyl content proved to be more hydrophobic. When the glycine content was increased the system goes from completely soluble to completely insoluble as polyglycine is insoluble in its nature due to its structuration. No terpolymer system with proline content below 50% was soluble or had thermoresponsivness. Nonetheless, an area with relatively high valine content and relatively low proline content showed a second high molecular weight species. For future works, terpolymers system with higher proline content than 50% could be explored, nonetheless, it is possible that any LCST behaviour would be hysteretic like the copolymers presented thus far having a DP of 100 or higher. This is due to the higher proline reactivity in water assisted conditions. This high reactivity was not counterbalanced with water content nor water nature. Decreasing the feed ratio of proline did not seem to introduce better distribution

but unsatisfactory synthesis. This was shown for terpolymers with low proline content that showed poor SEC profiles. It is also possible to explore lower molecular weight polymers as to compare their behaviour to copolymers **6** and **7** (Gly/Pro copolymer with M/I=50 and 75 respectively) presented earlier in this chapter.

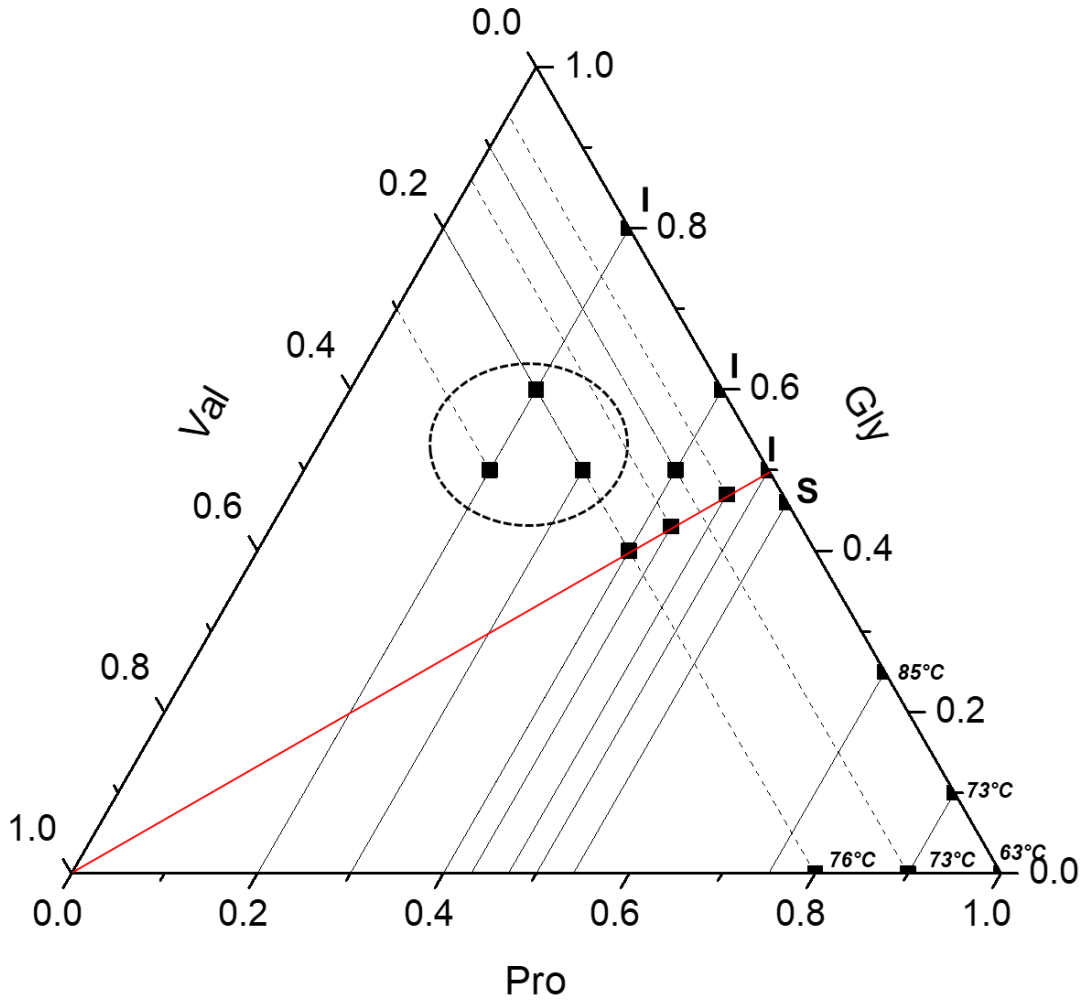


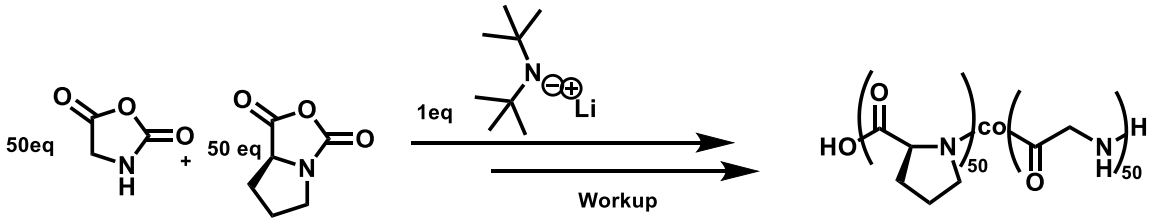
Figure 21 : *Ternary system of valine, proline, and glycine terpolymers phase diagram in water at 5 mg/ml. Each square represents a different synthesized copolymer. All points intersecting and above the red line are systems that are insoluble in water at all temperatures. The dashed circle represents systems that had two soluble species in SEC HFIP. The T_{CP} of homoPLP and binary systems is represented next to the corresponding dots. I=insoluble at all temperatures, S=soluble at all temperatures.*

2.2. Organic media copolymerization

With the limitations of the water assisted ROP to control the distribution of comonomers (presented in paragraph 2.1), other polymer synthesis pathways involving anhydrous

conditions have also been explored. Systems and mechanisms discussed or experimented in the previous chapter **2** and **3** able to both polymerize proline NCA and amino acid NCAs have been explored in this part of the research project. These other ROP systems first include the normal amine mechanism (NAM) or the use of tetrabutylammonium benzoate (TBBA) to initiate/catalyze the ring-opening. Both NAM and TBBA failed to produce the intended copolymers and the corresponding results are fully detailed in the **Table S1**. Alternative systems also include LiHMDS-mediated ROP (section **2.2.1**), TMG-mediated ROP with different initiators (part 2.2.2) and organometallic nickel-based initiation/catalysis (part 2.2.3). These 3 different systems will be discussed each in a separate part for easier understanding. Readers can refer to chapters **1** and **3** for detailed mechanisms and experimental procedures if they are not presented in this chapter.

2.2.1. LiHMDS



Scheme 8 : *Proline NCA and glycine NCA ROP using LiHMDS as initiator. The resulting copolymer after purification is presented.*

In the literature, the LiHMDS-mediated ROP involves an anionic mechanism in the case of NCAs[66] while a ring-expansion polymerization mechanism is suspected in the case of NNCAs[67]. No examples of statistical copolymerizations of these two types of monomers have been published so far. To this end, both polymerization conditions were conducted in parallel (**Table 7**): copolymerization of equimolar amounts of glycine NCA and proline NCA were conducted to assess the potential of this pathway to form polypeptides. The copolymerization was conducted either in THF (conditions inspired by LiHMDS-mediated ROP of NCAs) or DMF/THF mixture (conditions inspired by LiHMDS-mediated ROP of NNCAs).

Table 7 : *reaction conditions and synthesis results for copolymers* ***17-18****.*

	Gly%[a]	M/B[b]	Reaction solvent Concentration	Yield	Gly%[c]	Mn by SEC[f] Đ[g]
Copolymer **17**	50	100	THF 0.4M	*d*	-	-
Copolymer **18**	50	100	DMF[e] 0.4M	22	49	32kg/mol 1.11

[a]Glycine content determined by the initial feed ratio; [b]Monomer to LiHMDS ratio (B=base); [c]Glycine content determined by the integration of the proline and the glycine backbone peaks in [1]H NMR analysis of the final copolymer; [d]No polymerization was observed; [e]LiHMDS was introduced in a THF stock solution as it is insoluble in DMF as reported in the literature[67]; [f]determined by PMMA calibration in HFIP SEC and RI detection; [g]determined by SEC peaks.

Table 7 shows that the copolymerization was only successful in DMF (copolymer **18**), conditions previously developed for the polymerization of other NNCAs.[67] In this solvent, the polymerization proceeded extremely rapidly, consistently with the anionic mechanism. After purification, ^{1}H NMR analysis confirmed the presence of both monomer backbones in equimolar amounts (**Figure S9**). The SEC analysis in HFIP of copolymer **18** showed high molecular weight polymers with narrow dispersity (**Figure 22**), but the Mn value was much higher than expected (7.7 kg/mol). The Mn value corresponds to a DP≈400. Such deviation is common for the ROP of NCAs using LiHMDS[66] but has not been observed for NNCAs[67].

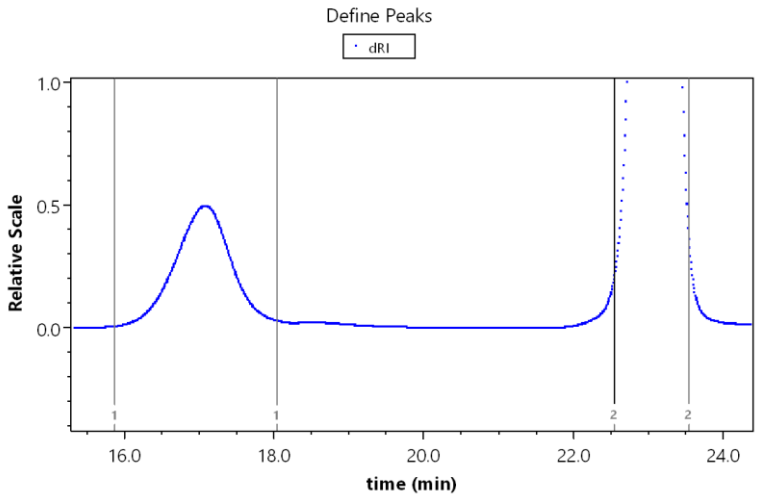
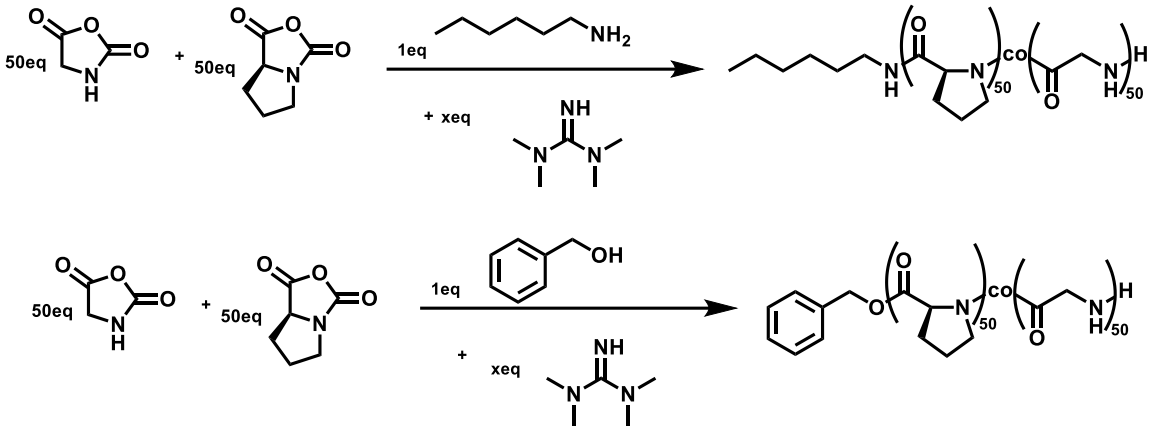


Figure 22 : *HFIP SEC chromatogram of copolymer* ***18****. Copolymer's high molecular weight is shown on the left and the flow marker in the right.*

This high molecular weight copolymer **18** was insoluble in water and showed no thermoresponsive behaviour. Further investigation into the mechanism of the ROP is needed to be able to better control the copolymer properties. If the anionic mechanism is considered, other copolymers with a lower M/B ratio must first be synthesized to produce lower molecular weight copolymers that are water soluble.

2.2.2. TMG catalysis



Scheme 9 : Proline NCA and glycine NCA ROCOP using different initiators and TMG as catalyst. The resulting copolymers after purification are presented.

Tetramethylguanidine (TMG) has also been successfully used to catalyse the ROP of NCAs[68] and NNCAs[69] separately using a variety of different initiators and conditions. Copolymerizing both monomers using this methodology is still unreported. Several experimental procedures have been tested in this research project (**Scheme 12**), all having equimolar amounts of glycine and proline to be compared to the other systems. The results are illustrated in **Table 8**.

***Table 8** : reaction conditions and synthesis results for copolymers **19-22**.*

	Targeted structure	Reaction solvent Concentration	Temp. (°C)	TMG	Yield
Copolymer **19**	HA-Pro$_{50}$-Gly$_{50}$	THF 0.1M	RT	1eq	*a*
Copolymer **20**	HA-Pro$_{50}$-Gly$_{50}$	THF 1M	50°C	0.01eq	82
Copolymer **21**	BOH-Pro$_{50}$-Gly$_{50}$	THF 1M	50°C	0.01eq	85
Copolymer **22**	HA-Pro$_{50}$-Gly$_{50}$	DMSO 1M	50°C	0.01eq	83

a. no polymerization was observed. HA=hexylamine, BOH=benzylalcohol.

At room temperature, even with high catalyst loading, no copolymers were formed (copolymer **19**). Only at 50°C were copolymers formed using low catalyst loading (0.01 eq), a procedure inspired by the litterature.[69] After purification, ^{1}H NMR analysis of copolymers **20-21** showed the formation of copolypeptides containing both proline and glycine units (**Figure 23**). The structure of the copolymers was confirmed with the presence of the initiator peaks (a and b). The analysis of the spectra showed a glycine content of around 42% for both copolymers.

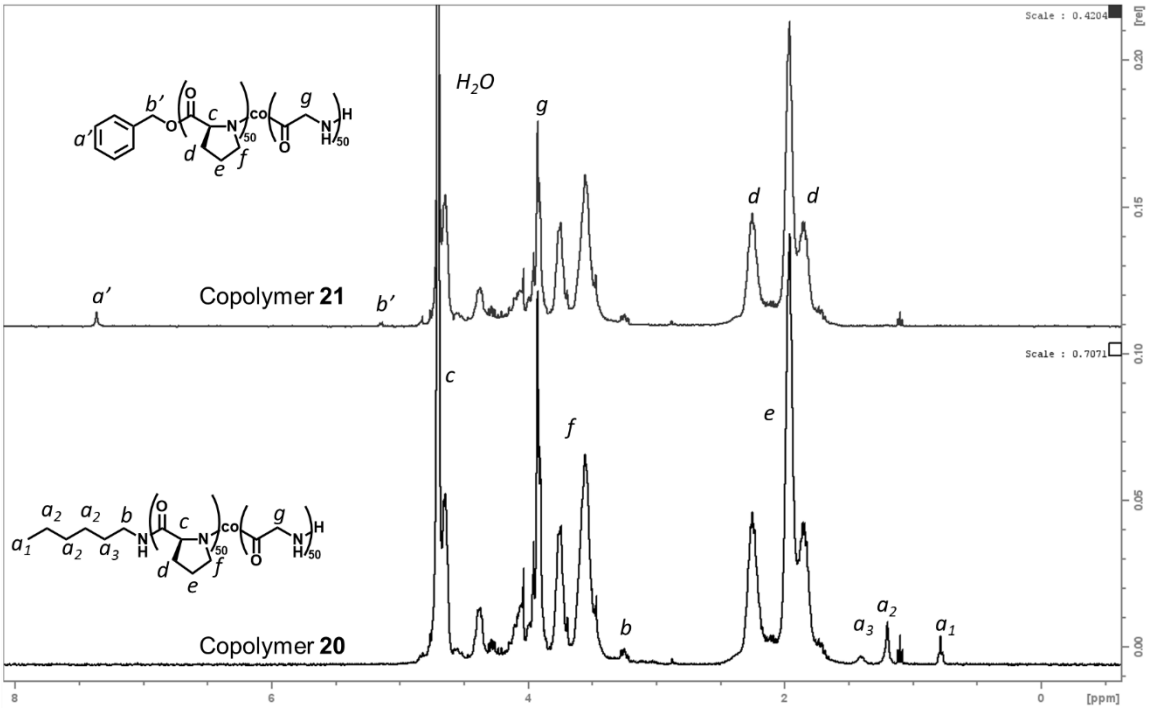


Figure 23 : *^{1}H NMR for copolymers **20-21** in D_2O. The respective structures of the copolymers are drawn with the peak assignation.*

Aqueous SEC analysis of both copolymers show the formation of high molecular weight copolymers with dispersity around 1.15 (**Figure 24**, SEC in HFIP is presented in **Figure S15**). The calculated molecular weights were 4.6 kg/mol and 5.0 kg/mol respectively using dn/dC of homoPLP. These values are close to theoretical ones (7.8kg/mol) but slightly lower than those produced by water assisted ROP (**Table 1**).

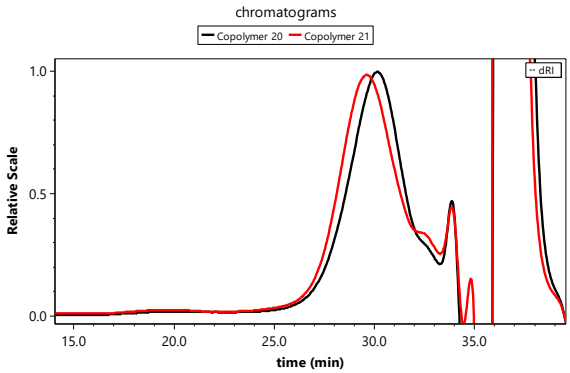


Figure 24 : *Aqueous SEC of copolymers* ***20-21****. RI detection is shown.*

These promising results using TMG prompted us to study the kinetics of the copolymerization. Nonetheless, Pro NCA and the copolymers are insoluble in THF, not allowing for *in-situ* NMR monitoring. For this purpose, copolymer **22** was synthesized in DMSO where Pro NCA is slightly more soluble at a stoichiometric comonomer ratio. Upon purification, NMR analysis showed a glycine content of 33% and high residual TMG (peaks a' and b') even after extensive purification (**Figure 25**). Chain initiation through nucleophilic attack of TMG has been previously reported in the literature,[68] but was not solvent dependant.

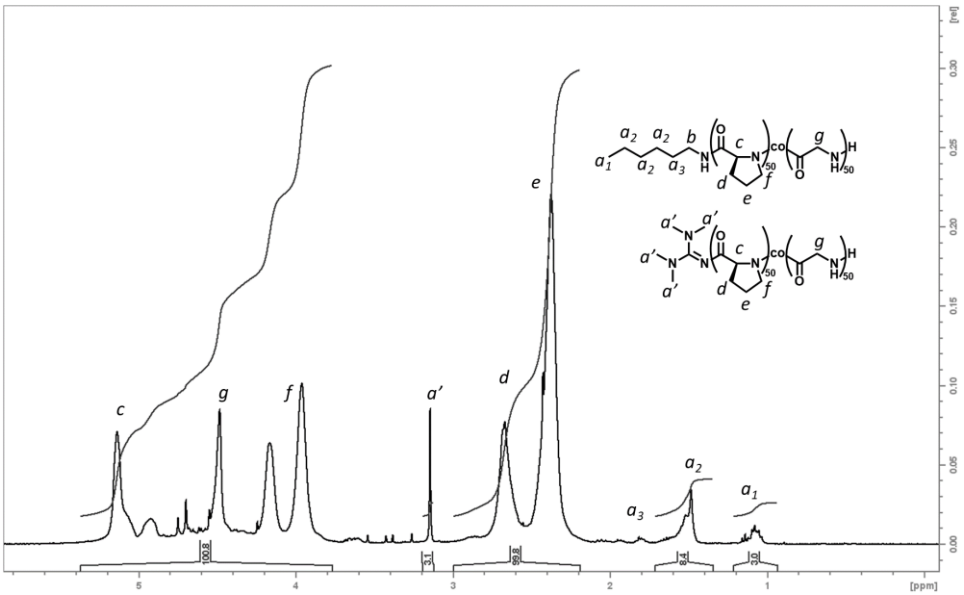


Figure 25 : *a.* ^{1}*H NMR for copolymer* ***22*** *in TFA-d. Two possible structures of the copolymers are drawn with their peak assignation.*

SEC analysis (**Figure 26**) confirmed the coexistence of two polymeric species by the presence of two peaks. This may indicate the coexistence of two initiation mechanisms by hexylamine and TMG.

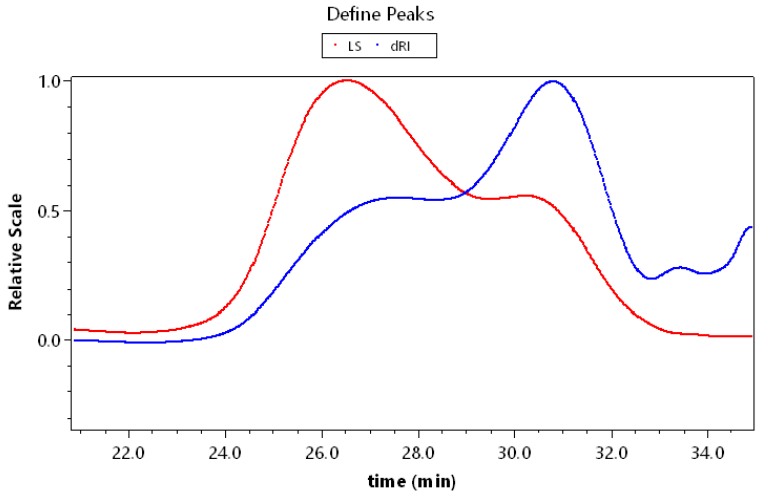


***Figure 26** : Aqueous SEC of copolymer **22**. RI detection is shown in blue while MALS detection is shown in red.*

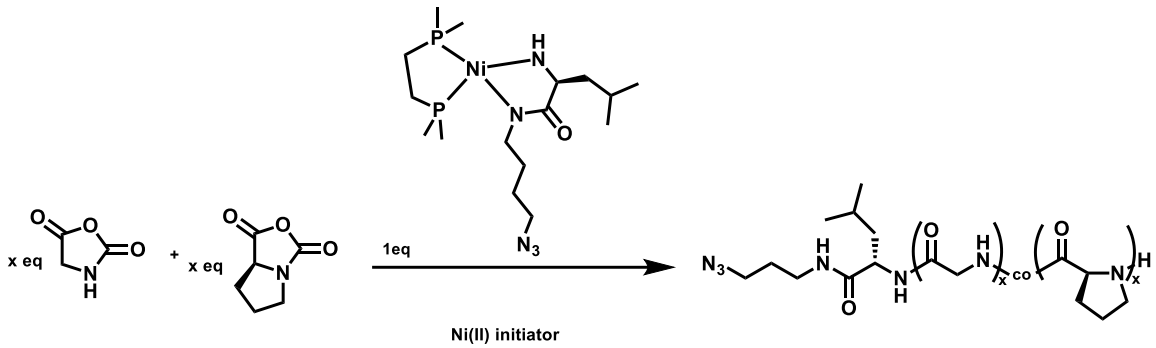
The choice of solvent is thus key for the stability of the initiator/catalyst complex and the stabilization of the propagating chain-end carbamate specie with TMG. Both copolymer **20** and **21** were insoluble in MilliQ water at 5 mg/ml and showed no thermoresponsive behaviour. The insolubility of Pro NCA and copolymers hindered the possible kinetic study of this system by *in-situ* NMR analysis. Nonetheless, since this reaction takes several days to complete, it would be possible to conduct the quenching technique monitoring described earlier in this chapter (section 2.1.1) to calculate the reactivity ratios in future work.

2.2.3. Ni-based organometallic catalysis

As described in the previous chapter, polymerization of proline NCA using amido-amidate nickel initiator resulted in slower than reported kinetics (see chapter **3**).[51] PLP prepared via this route were also highly disperse. Nonetheless, PLPs prepared via this pathway were thermoresponsive, and copolymerization was thus studied using Ni(II) based initiators.

Proline NCA and Glycine NCA ROCOP

Our early copolymerization efforts showed that proline, in these conditions, tends to react faster than glycine (data not shown).



Scheme 10: *Proline NCA and glycine NCA ROCOP using Ni(II) amido amidate initiator. The resulting copolymers after purification are presented.*

Three copolymers (**23-25**) with a 50% glycine monomer feed and increasing M/I ratio were thus synthesized using Ni(II) amido amidate catalyst (**Scheme 10**). After purification by precipitation followed by dialysis, the polymers were analysed by ^{1}H NMR and SEC in TFE. (**Table 9**).

Table 9 : *Analysis and results for copolymers **23-25** synthesis*

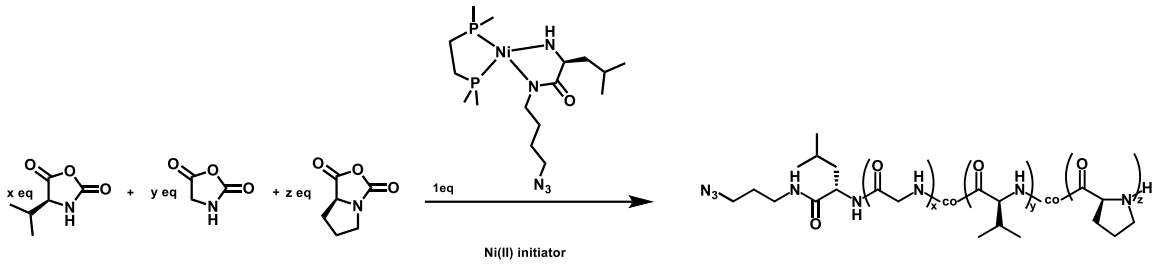
	Targeted structure	Gly%[a]	Mn by SEC[b] Đ[c]	T_{CP}[d]	T_{CL}[d]
Copolymer **23**	AzPrLeu-Pro_{25}-Gly_{25}	34	14kg/mol 1.88	>80°C	RT
Copolymer **24**	AzPrLeu-Pro_{38}-Gly_{38}	40	12kg/mol 2.14	>80°C	RT
Copolymer **25**	AzPrLeu-Pro_{50}-Gly_{50}	35	12kg/mol 2.14	Insoluble	

*[a]Glycine content determined by the integration of the proline and the glycine backbone peaks in ^{1}H NMR analysis of the final copolymer in TFA-d; [b]Determined by PMMA calibration in TFE SEC and RI detection; [c]determined by SEC peaks d. determined using benchtop heating plate and turbidimetric assays by eye. AzPrLeu=azido-propyl amide L-leucine, as shown in **Scheme 10**.*

^{1}H NMR analysis (**Figure S10**) of the copolymers showed lower than expected glycine content (**Table 9**). SEC analysis in TFE showed broad peaks with high dispersity (**Table 9**). Copolymers **23-24** (to be compared to copolymers **6-7**) were completely soluble at room temperature. They also expressed an LCST behaviour consistent with that of low molecular weight PLP (**Table 9**). These results suggest that in these conditions, proline NCA tends to react faster than glycine NCA producing highly gradient copolymers. These soluble copolymers expressing an LCST behaviour are also clickable due to the azide group present on the initiator (**Scheme 10**).

Terpolymerization

As these copolymerizations are done in DMF at 50 mg/ml, they are usually heterogenous as Pro NCA, Gly NCA and the copolymers are mostly insoluble. Furthermore, copolymerization of proline NCA with valine NCA resulted in copolymers with much higher proline content than initial feed. Other copolymerization efforts with higher valine monomer feed resulted in very low yields possibly due to the secondary structuring of valine sequences discussed earlier that hinders the polymerization. This was similar in most terpolymerization (**Scheme 11**) systems tested.



Scheme 11 : *Terpolymerization of proline NCA, glycine NCA and valine NCA using Ni(II) amido amidate initiator. A series of copolymers was explored by changing the different feed ratios of comonomers (x,y and z respectively).*

Since the proline reactivity was found to be higher in these conditions as for the water assisted ROP, the proline content was kept at 50% and the valine content was slightly increased. At M/I=50, two terpolymerization systems were soluble in the reaction medium (DMF) (**Table 10**). After purification by precipitation in ether, the structures of these copolymers were analysed by ^{1}H NMR. Their structure was verified, and the results are in accordance with the theoretical values (**Table 10**). Furthermore, the SEC analysis of these terpolymers (**Table 10)**, showed much lower dispersities than PLP (chapter **3**) and Gly-Pro copolymers (earlier in this section) prepared via this pathway. Terpolymers **10-11** were completely soluble at room temperature in MilliQ water. They also expressed an LCST behaviour consistent with that of low molecular weight PLP as (**Table 10**) similar to copolymer **23**, suggesting that incorporating valine did not have a big effect on the thermoresponsive behaviour.

Table 10 : *Analysis and results for terpolymers **10-11** synthesis.*

	Targeted structure	Structure by NMR	Mn by SEC[b] Đ[c]	T_{CP}[d]	T_{CL}[d]
Terpolymer **10**	AzPrLeu-Pro_{25}-Gly_{20}-Val_5	AzPrLeu-Pro_{29}-Gly_{17}-Val_4	8.5kg/mol 1.68	>80°C	RT
Terpolymer **11**	AzPrLeu-Pro_{25}-Gly_{15}-Val_{10}	AzPrLeu-Pro_{31}-Gly_{14}-Val_6	8.1kg/mol 1.58	>80°C	RT

*[a]Determined by the integration of the different backbone peaks in ^{1}H NMR analysis of the final copolymer in TFA-d; [b]Determined by PMMA calibration in TFE SEC and RI detection; [c]Determined by SEC peaks; [d]Determined using benchtop heating plate and turbidimetric assays by eye. AzPrLeu=azido-propyl amide L-leucine, as shown in **Scheme 10**.*

As only these polymerizations advanced in a soluble manner in DMF, we hypothesize that the synthesis of PLP and copolymers of Pro NCA using Ni(II) initiator in organic solvent is highly process dependant. For this purpose, the Pro NCA initial concentration should be kept around 20 mg/ml (0.15 M) instead of 50 mg/ml for it to be more soluble. Furthermore, other terpolymerization with Pro% less than 50 should be tested to assess their solubility. When copolymerizing Pro NCA with Gly NCA (copolymers **23-25**), the growing copolymers were insoluble in DMF, yet by adding little valine (terpolymers **10-11**) the polymerization advanced in a homogeneous manner. This phenomenon can be explained by disruption of insoluble secondary structuring during polymerization: PLPI for proline, PGly for glycine and β-sheets for valine.

These promising results in the copolymerization of Pro NCA with NCAs (glycine and valine) using catalysis in organic media are preliminary results paving the way to terpolymerization system producing ELP-like thermoresponsiveness. Copolymerization of proline and glycine using LiHMDS are particularly promising because it enables access to high molecular weight copolymers with a controlled structure. Other copolymers with different compositions and molecular weights must be synthesized to conclude on the feasibility of this approach towards ELP mimics. Ideal experimental conditions were also determined for the use of TMG catalysis to produce copolymers based on proline and glycine with controlled structures. The study of the copolymerization kinetics of this system could also prove useful in steering the primary sequence of the produced copolymers.

3. Conclusion

As a means to mimic the thermoresponsive behaviour of ELPs, copolymerization of Pro NCA with Gly-NCA and/or Val-NCA was explored in many different conditions. The knowledge accumulated from the two previous studies developed in chapter **2** and **3** had to be combined. On one hand the understanding of the kinetics of copolymerization of NCAs and its impact on the final copolymer microstructure and on the other the thermoresponsive behaviour of polypeptides based on natural amino acids. Since the statistical copolymerization of NCAs with NNCAs is still an unexplored field, many copolymerization systems had to be optimized. Firstly, the water-assisted ROP mechanism produced copolymers with a strong gradient structure that showed similar LCST behaviour to homopolyprolines when having high proline content. When less proline was introduced, the copolymers either became completely soluble or completely insoluble. Lower molecular weight copolymers of Gly-Pro (50:50) showed block-like structures and formed thermoresponsive nanoparticles. Second, classic copolymerization in organic solvent using NAM was explored. Initial results showed poor results and thus organocatalysis using TMG was employed. Using this catalysis mechanism, it was possible to produce Pro-Gly copolymers with controlled structures and dispersities. Optimal reaction conditions were found for further exploration of this pathway to produce not yet attained thermoresponsive copolymers. Next, the exploration of another mechanism using LiHMDS (AMM) was explored. This synthesis technique was previously developed separately for NCAs and NNCAs and copolymerization combined challenges from both. Nonetheless, a copolymer with a well determined structure has been synthesized also proving another promising path towards thermoresponsive copolymers based on proline. Finally, organometallic catalysis using Ni(II) initiator was studied to copolymerise Pro NCA. This synthesis pathway proved to be highly process-dependent as it proceeded in a heterogenous manner and copolymers produced were highly disperse. Nonetheless, reaction conditions were developed to produce copolymerization that were soluble and produced better controlled copolymers. This final pathway also produced thermoresponsive copolymers based on proline repeating unit that are clickable on one chain end so they can be coupled with other functional moieties. These final three copolymerization pathways will be further explored, and their kinetics studied to establish a new base for producing desirable copolymers.

4. Supporting information

4.1. Materials

All chemicals and solvents in this work were purchased from Sigma Aldrich, Fluorochem, Acros, TCI, Strem and VWR, and were used without any purification unless otherwise described. DMF, THF and acetonitrile (ACN), were obtained from a solvent system purifier (PureSolv, Innovative Technology), kept under argon atmosphere and freshly used. Milli-Q water was obtained from a Purelab Prima purification system (ELGA) with a resistivity of 18.2 MΩ cm-1. The N-carboxyanhydride monomers were purchased from PMC Isochem, stored at 20 °C under argon atmosphere and weighed in the glove box Jacomex GP13 no. 2675 at the Laboratoire de Chimie des Polymères Organiques (LCPO, Bordeaux, France). Dialysis membranes were purchased from SpectrumLabs. Hexylamine and TMG were freeze-thawed and cryo-distilled on the Schlenk line prior to use. Gly NCA was synthesized and used as shown in chapter **2**. Ni initiator was obtained from collaborators from Deming's group in UCLA and used as received. The experiments involving its use (section **2.2.3**) were conducted in their laboratory settings.

4.1.1. Nuclear Magnetic Resonance (NMR)

^{1}H NMR 400 MHz spectra were obtained using a Bruker Avance I (Liquid-state 400 MHz NMR spectrometer with 5 mm BBFO probe). Deuterated water (D_2O, Euriso-top, 99.8%) was used as solvent and reference for the lock. The spectra obtained were calibrated using the residual solvent signals (H2O 4.79 ppm). The signals were categorized as follows: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m) and broad (br).

4.1.2. Size Exclusion Chromatography (SEC) Analyses

Polymer molar masses were determined by Size Exclusion Chromatography (SEC) using an aqueous solvent as the eluent. Measurements were performed on an Ultimate 3000 system from Thermoscientific equipped with diode array detector DAD. The system also includes a multi-angle light scattering detector (MALS) and differential refractive index detector (dRI) from Wyatt technology. Polymers were separated on two TOSOH TSK Gel G4000PWXL and G3000PWXL columns (300 × 7.8 mm) (exclusion limits from 200 Da to 300 000 Da) at a flowrate of 0.6 mL/min. Aqueous solvent composed of acetic acid (AcOH) 0.3 M, ammonium acetate 0.2 M and ACN (H2O/ACN:

6.5/3.5, v/v) was used as the eluent. The column temperature was held at 25°C. Measurements of dn/dc were performed on a differential refractive index detector dRI from Wyatt technology by injecting 500 µL of each sample dissolved in the aqueous phosphate buffer at 0.5-5 mg mL−1. The chromatograms were recorded with Chromeleon 7.2 software and and analyzed using Astra 7.1.0 software. Calculated dn/dC values were 0.1616 mg/ml for L-Proline homopolymers and copolymers with D-Proline and 0.1444 mg/ml for block copolymers with PEG.

4.1.3. UV−Vis Spectroscopy

The cloud point temperature (T_{CP}) and the clearing point temperature (T_{CL}) of PLP solutions in pure water were determined by measuring the turbidity at 550 nm between 10°C and 85°C at a 1 °C.min^{-1} scan rate at different concentrations of polymer and salt. Data were collected on a Cary 100 UV−Vis spectrophotometer equipped with a multicell thermoelectric temperature controller from Agilent Technologies. T_{CP} and T_{CL} were defined as the temperatures at the inflection point of the transmittance-temperature curves during heating ramps (T_{CP}) or cooling ramps (T_{CL}). A typical experiment followed the following heating program:

1- 10 min isotherm at 10°C

2- 1°C/min ramp to 85°C

3- 10 min isotherm at 85°C

4- 1°C/min ramp to 10°C

Absorbance was plotted as a function of temperature, with isotherms' data not collected. Subsequently, absorbance was converted to transmittance for a better visualization of the degree of turbidity and to discriminate between complete and incomplete aggregations. The y-axis in the results is displayed as 1-Transmittance (%) to help the reader relate the soluble and aggregated forms to low and high values respectively.

4.2. NMR

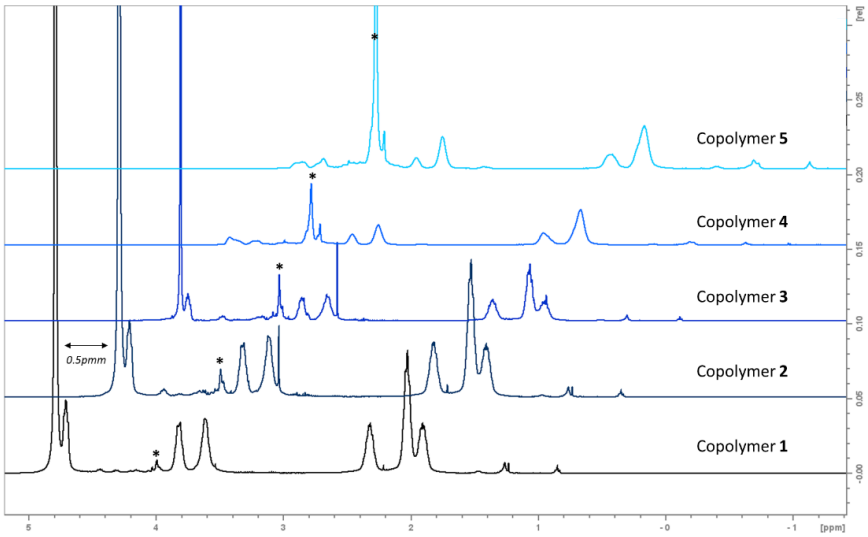


Figure S1: *^{1}H NMR spectra of copolymers **1-5** with increasing glycine content (shown with an * is the glycine peak). The spectra are staggered by 0.5 ppm for easier visualization. Copolymer **1-3** were analyzed in D_2O and copolymers **4-5** in TFA-d.*

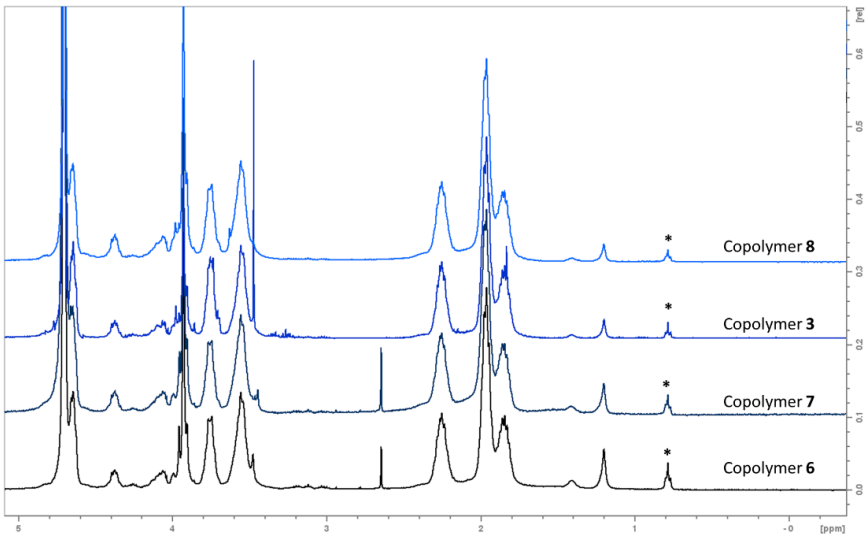


Figure S2: *^{1}H NMR spectra in D_2O of copolymers **3,6,7,8** with increasing M/I ratio. Backbone peaks intensity has been normalized for all 4 spectra to show the decrease of the initiator peaks intensity (shown with an *).*

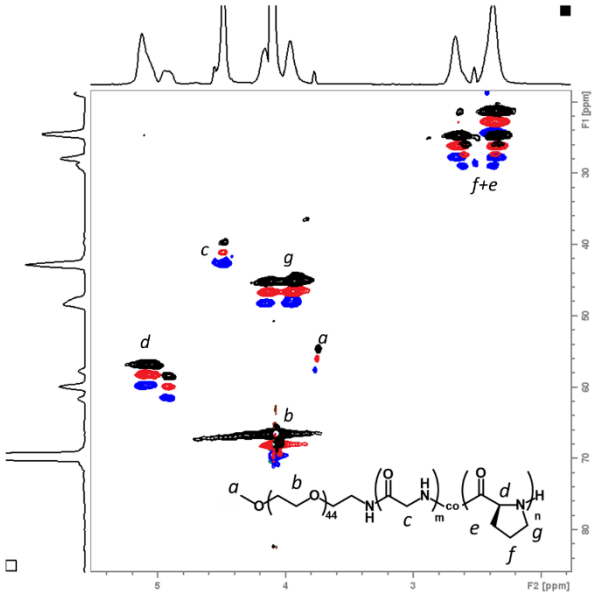


Figure S3: *HSQC H-C 2D NMR of different kinetic points of an equimolar copolymerization of glycine NCA and proline NCA initiated by PEG-NH_2 (2kg/mol) in TFA-d. The spectra allow for the attribution of different carbon atoms attached to the protons as attributed by in the scheme. In black, the spectrum of the aliquot after 30 seconds of reaction, in red after 120s and in black is the final copolymer. The spectra were staggered in the y-axis for easier visualization.*

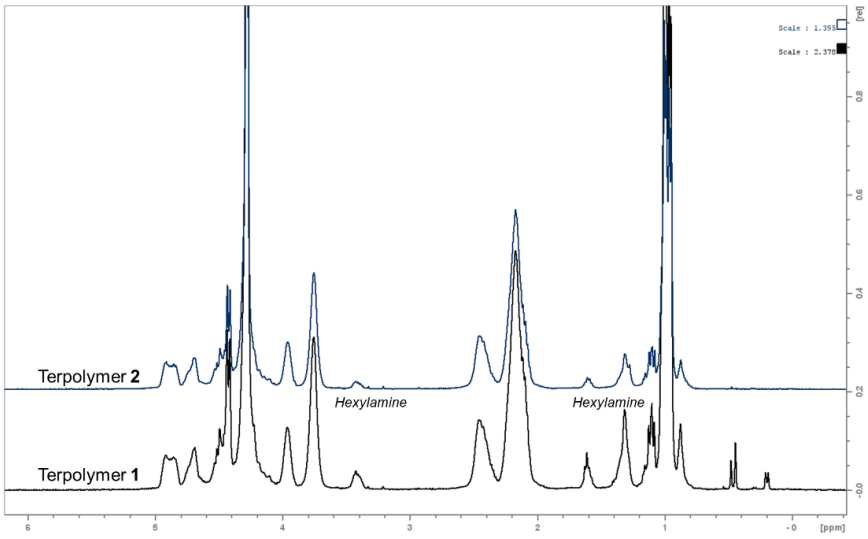


Figure S4 : *^{1}H NMR spectra in TFA-d of terpolymers **1** and **2** with increasing M/I ratio. Backbone peaks intensity has been normalized for both spectra to show the decrease of the initiator peaks intensity indicated by with hexylamine.*

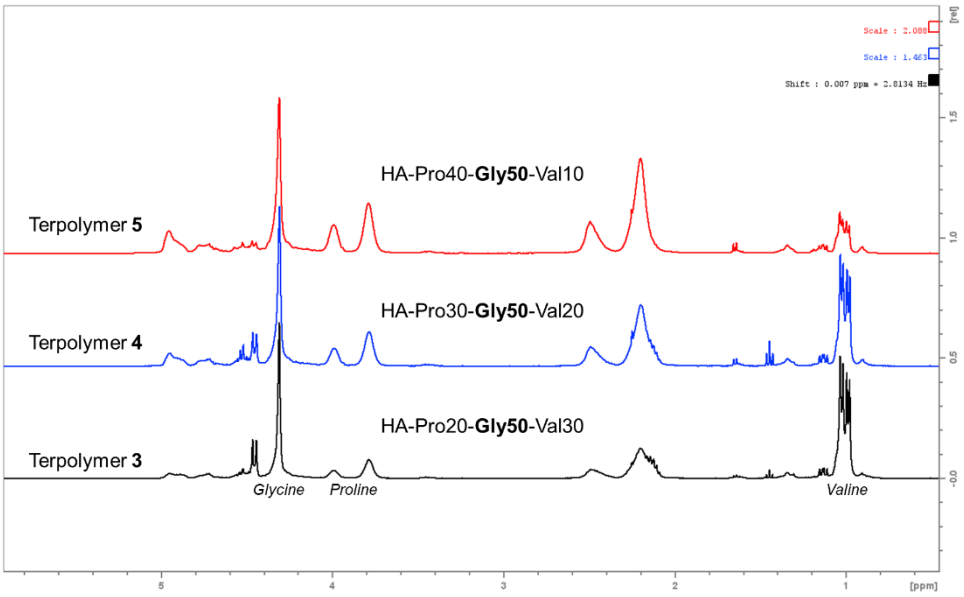


Figure S5 : *^{1}H NMR spectra in TFA-d of terpolymers **3-5** varying valine and proline content. Glycine peaks' intensity has been normalized for all spectra to show the change in the relative peaks of proline and valine as indicated on the spectra.*

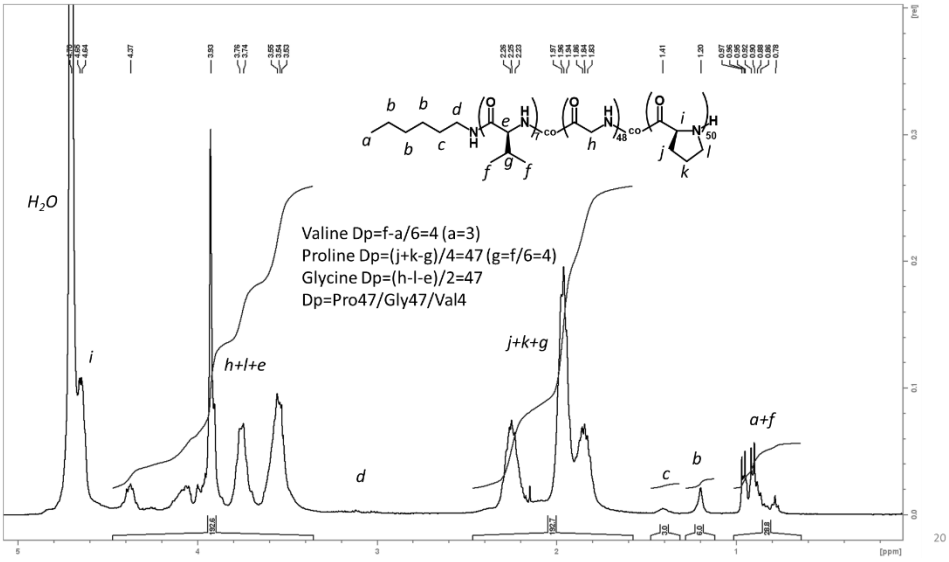


Figure S6 : *^{1}H NMR of terpolymer **6** in D_2O. Its theoretical structures is drawn with the peak assignation. The calculation of the different Dp is shown by the calculations from the integration of the initiator peak (d) and the respective backbone peaks of valine (f), proline (j,k) and glycine (h).*

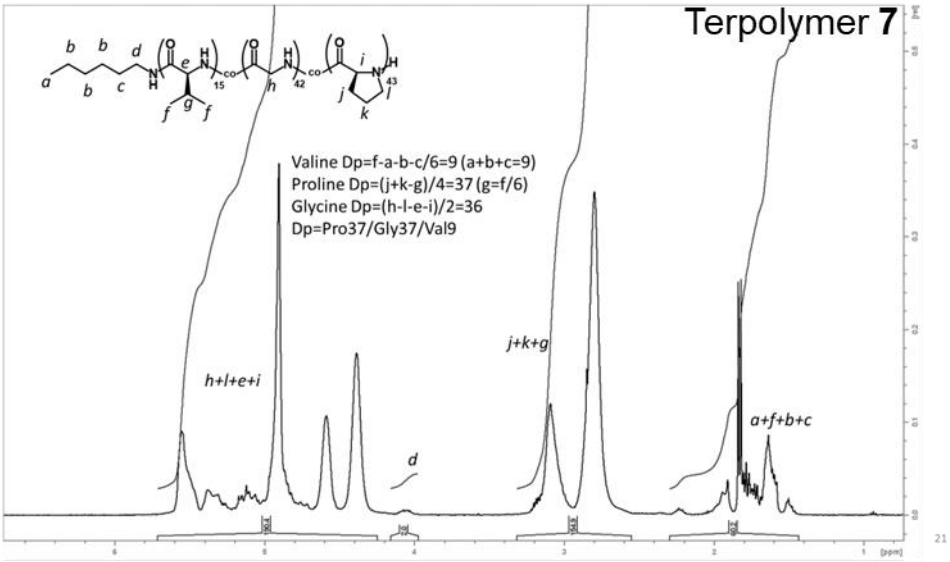


Figure S7 *: [1]H NMR of terpolymer **7** in TFA-d. Their theoretical structures are drawn with the peak assignation. The calculation of the different Dp is shown by the calculations from the integration of the initiator peak (d) and the respective backbone peaks of valine (f), proline (j,k) and glycine (h).*

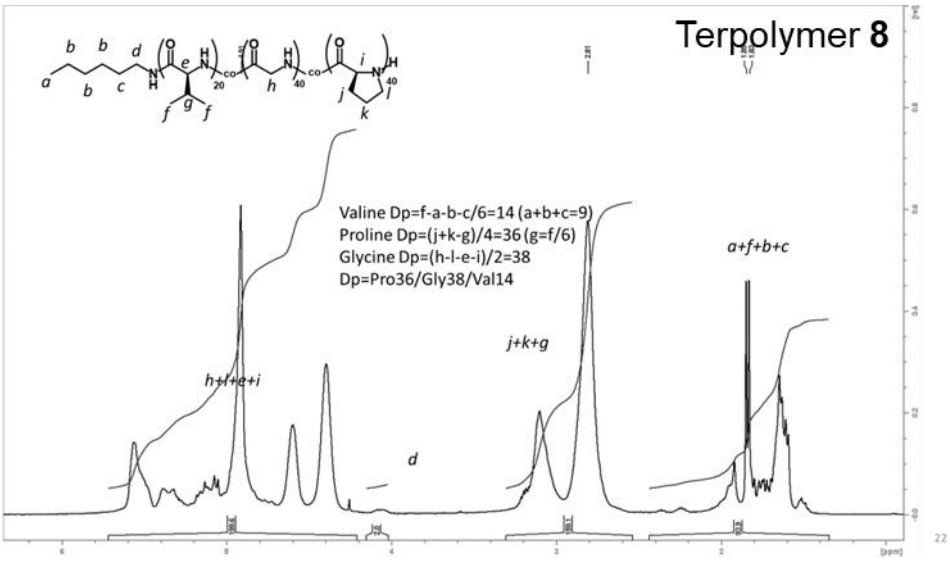


Figure S8 : *[1]H NMR of terpolymer **8** in TFA-d. Their theoretical structures are drawn with the peak assignation. The calculation of the different Dp is shown by the calculations from the integration of the initiator peak (d) and the respective backbone peaks of valine (f), proline (j,k) and glycine (h).*

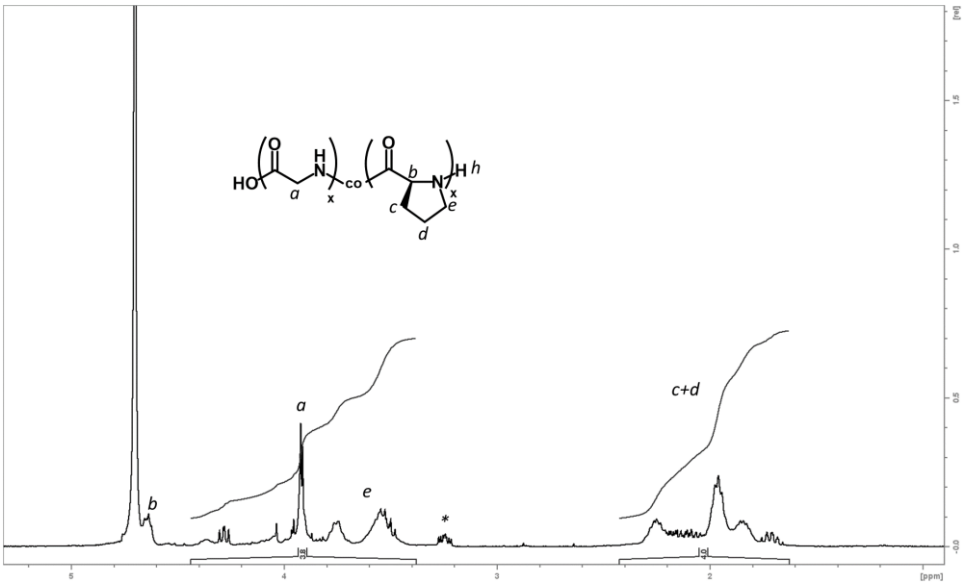


Figure S9 : *^{1}H NMR of copolymer **18** in D_2O. Its theoretical structures is drawn with the peak assignation. The calculation of the Dp is not possible as the polymer does not possess identifiable chain ends when polymerized with LiHMDS. The analysis of the peaks of the backbones shows that Pro and Gly are present in equimolar amounts. *unidentified*

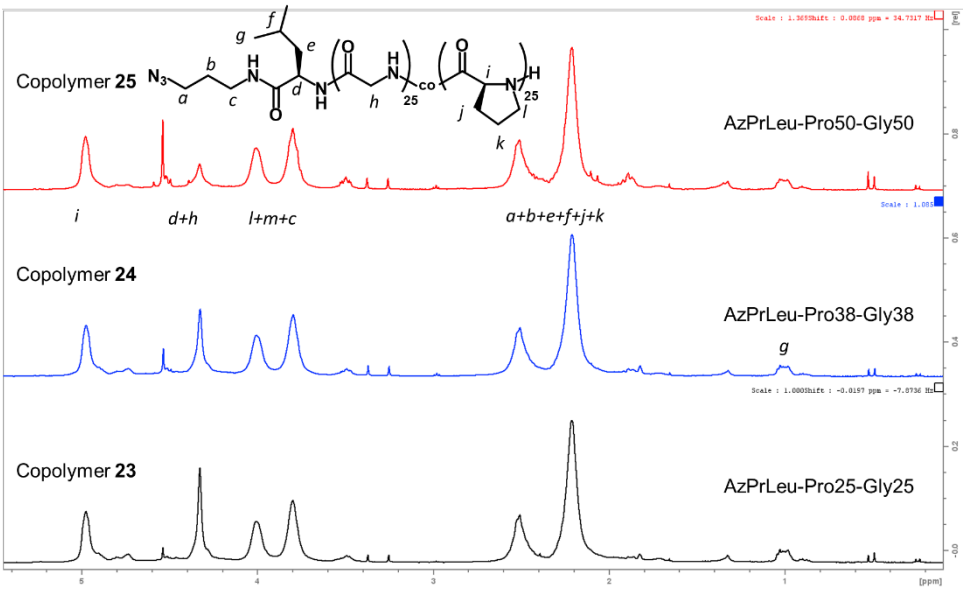


Figure S10 *: ^{1}H NMR spectra in TFA-d of copolymers **23-25** varying valine and proline content. Proline peaks' (i) intensity has been normalized for all spectra to show the change in the relative peaks of proline and glycine as indicated on the spectra.*

4.3. Microscopy

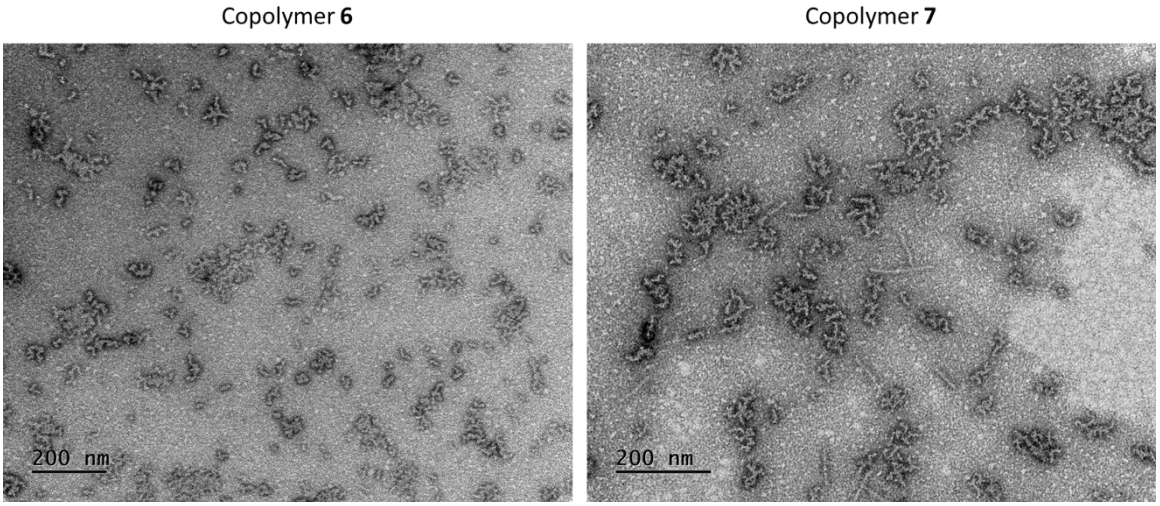


Figure S11: *TEM images of copolymers* ***6-7****. The grids were prepared from solutions at 0.1mg/ml at room temperature keeping the low temperature nano-aggregates. Images show the formation of wormlike structures.*

4.4. Synthesis

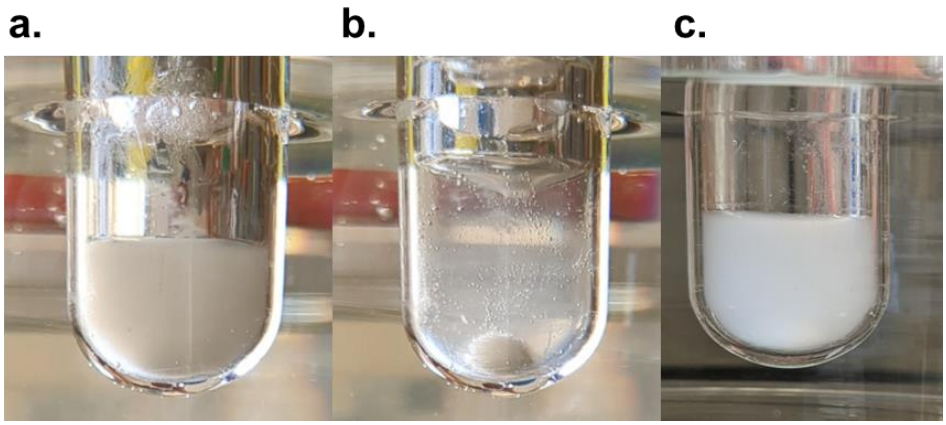


Figure S12: *Pictures from the synthesis of copolymerization kinetics system. a. at the beginning of the copolymerization with NCAs suspended in ACN, b. after adding the water and the initiator (bubbles can observed), c. after 1 hour of reaction.*

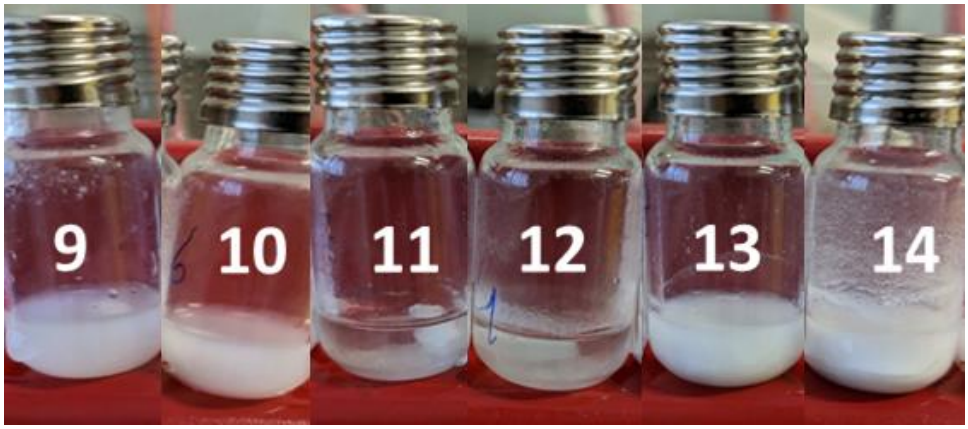
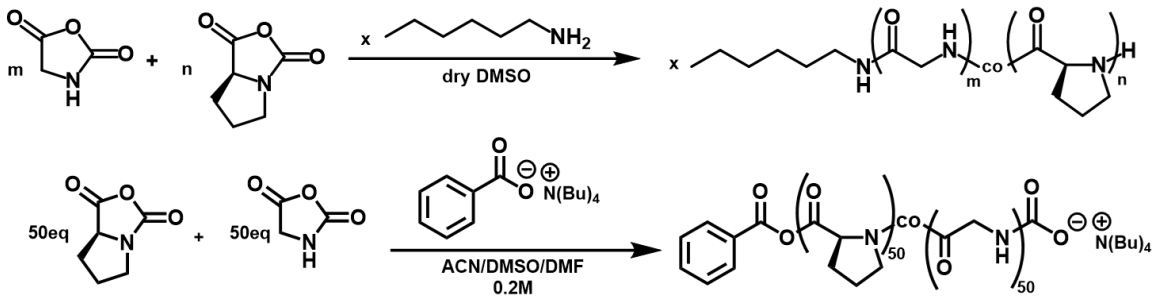


Figure S13 *: Copolymers* ***9-14*** *at the end of the synthesis. The aqueous phase % highly affected the copolymerization process due to the solubility of the active species.*

4.4.1. Glycine NCA and Proline NCA's other copolymerization systems

Copolymerizations using NAM or TBBA (produced copolypeptides with very low yields. Some copolypeptides had lower glycine contents than theoretical feed.



Scheme 12 : *ROP of glycine NCA and proline NCA by hexylamine initiator (top) or by TBBA initiator (bottom)*

Table S1 : *results for copolymerization systems using NAM and TBBA*

	Targeted structure	Yield	Initiator	Reaction solvent	Gly%[a]
Copolymer S**1**	HA-Pro$_{15}$-Gly$_{15}$	26	Hexylamine	DMSO	23
Copolymer S**2**	HA-Pro$_{25}$-Gly$_{25}$	22	Hexylamine	DMSO	41
Copolymer S**3**	Pro$_{50}$-Gly$_{50}$	*b*	TBBA[c]	ACN	-
Copolymer S**4**	Pro$_{50}$-Gly$_{50}$	12	TBBA[c]	DMSO	40
Copolymer S**5**	Pro$_{50}$-Gly$_{50}$	14	TBBA[c]	DMF	36

a. glycine content determined by the integration of the proline and the glycine backbone peaks and hexylamine peaks by ^{1}H NMR in the final copolymer. b. no polymerization was observed. c. tetrabutylammonium benzoate.

4.5. Size exclusion chromatography

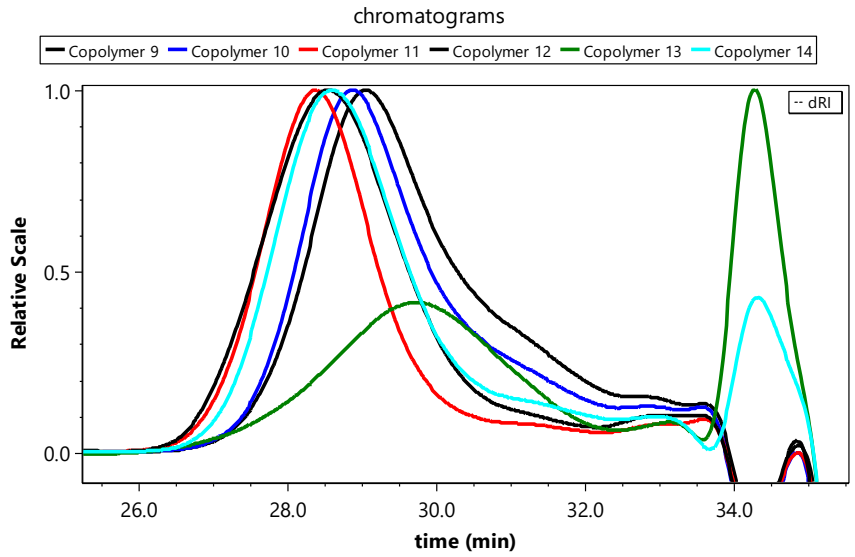


Figure S14: *Aqueous SEC chromatograms of copolymers* ***9-14*** *results. RI detection is shown, lower molar mass species can be observed for copolymers* ***13*** *and* ***14****.*

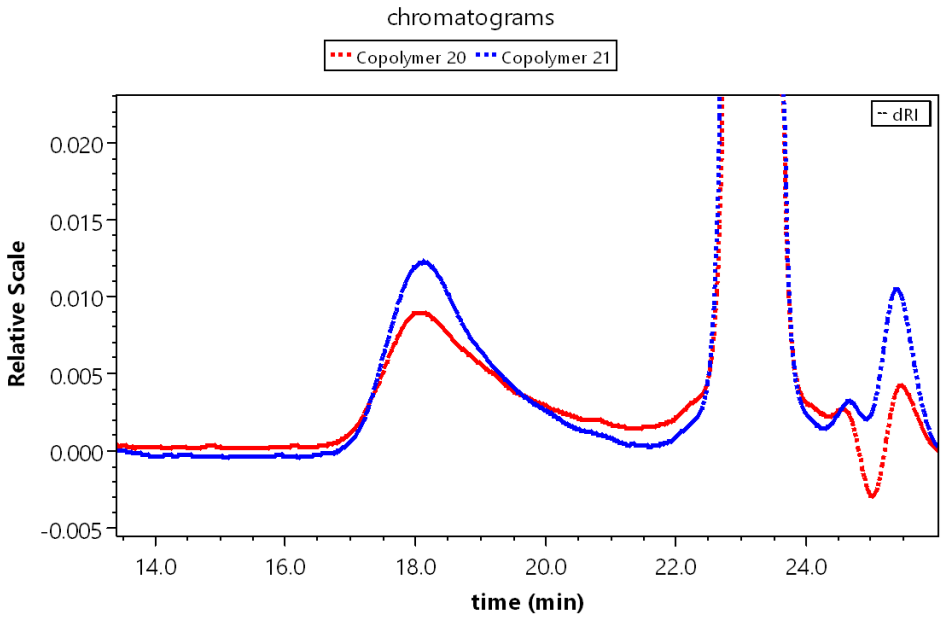


Figure S15 : *HFIP SEC chromatograms of copolymers* ***20-21*** *at 5mg/mll. RI detection is shown. Mn calculation using PMMA calibration was used, and the resulting values were 10.1 kg/mol and 10.0 kg/mol respectively. Their dispersities were 1.34 and 1.33 respectively.*

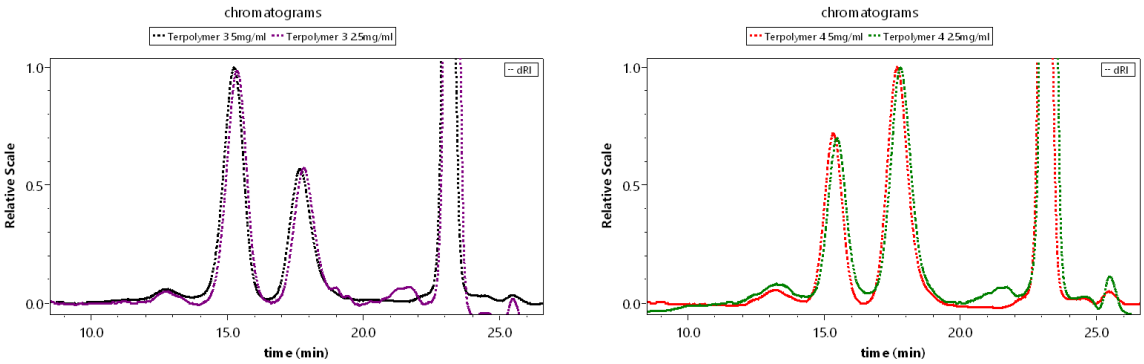


Figure S16 *: HFIP SEC chromatograms of terpolymer* ***3*** *(on the left) and terpolymer* ***4*** *at 5mg/ml and 2.5mg/ml. RI detection is shown and peaks are normalized for relative scaling.*

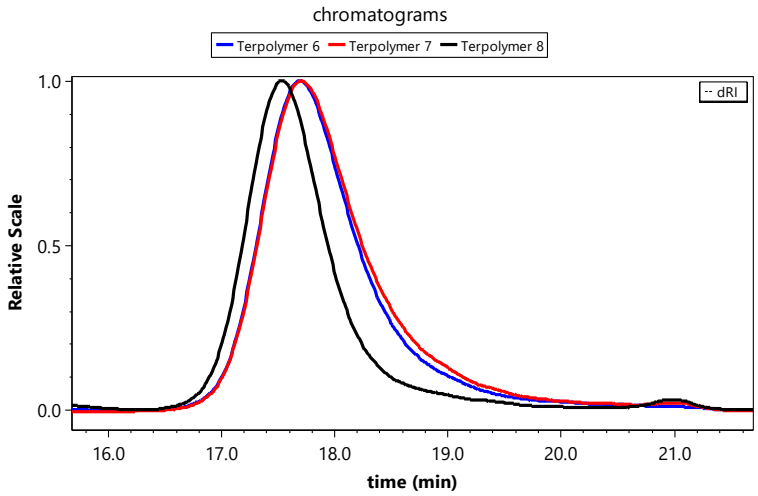


Figure S17 *: HFIP SEC chromatograms of terpolymers* ***6-8****. RI detection is shown.*

5. References

(1) Imai, K.; Mitaku, S. *Biophysics (Oxf).* **2005**, *1*, 55–65.

(2) Radivojac, P.; Iakoucheva, L. M.; Oldfield, C. J.; Obradovic, Z.; Uversky, V. N.; Dunker, A. K. *Biophys. J.* **2007**, *92* (5), 1439–1456.

(3) Garanger, E.; MacEwan, S. R.; Sandre, O.; Brûlet, A.; Bataille, L.; Chilkoti, A.; Lecommandoux, S. *Macromolecules* **2015**, *48* (18), 6617–6627.

(4) Pak, C. W.; Kosno, M.; Holehouse, A. S.; Padrick, S. B.; Mittal, A.; Ali, R.; Yunus, A. A.; Liu, D. R.; Pappu, R. V.; Rosen, M. K. *Mol. Cell* **2016**, *63* (1), 72–85.

(5) Vrhovski, B.; Weiss, A. S. *Eur. J. Biochem.* **1998**, *258* (1), 1–18.

(6) Zhao, B.; Li, N. K.; Yingling, Y. G.; Hall, C. K. *Biomacromolecules* **2016**, *17* (1), 111–118.

(7) Dzuricky, M.; Rogers, B. A.; Shahid, A.; Cremer, P. S.; Chilkoti, A. *Nat. Chem.* **2020**, *12* (9), 814–825.

(8) Garcia Quiroz, F.; Li, N. K.; Roberts, S.; Weber, P.; Dzuricky, M.; Weitzhandler, I.; Yingling, Y. G.; Chilkoti, A. *Sci. Adv.* **2019**, *5* (10).

(9) Quiroz, F. G.; Chilkoti, A. *Nat. Mater.* **2015**, *14* (11), 1164–1171.

(10) Chow, D.; Nunalee, M. L.; Lim, D. W.; Simnick, A. J.; Chilkoti, A. *Mater. Sci. Eng. R Reports* **2008**, *62* (4), 125–155.
(11) Urry, D. W.; Gowda, D. C.; Parker, T. M.; Luan, C.-H.; Reid, M. C.; Harris, C. M.; Pattanaik, A.; Harris, R. D. *Biopolymers* **1992**, *32* (9), 1243–1250.
(12) Garanger, E.; Lecommandoux, S. *Adv. Drug Deliv. Rev.* **2022**, *191*, 114589.
(13) Meyer, D. E.; Chilkoti, A. *Biomacromolecules* **2004**, *5* (3), 846–851.
(14) Garcia Quiroz, F.; Li, N. K.; Roberts, S.; Weber, P.; Dzuricky, M.; Weitzhandler, I.; Yingling, Y. G.; Chilkoti, A. *Sci. Adv.* **2019**, *5* (10), eaax5177.
(15) Roberts, S.; Dzuricky, M.; Chilkoti, A. *FEBS Lett.* **2015**, *589* (19PartA), 2477–2486.
(16) Badreldin, M.; Le Scouarnec, R.; Lecommandoux, S.; Harrisson, S.; Bonduelle, C. *Angew. Chemie Int. Ed.* **2022**, *61* (46), e202209530.
(17) GRAY, W. R.; SANDBERG, L. B.; FOSTER, J. A. *Nature* **1973**, *246* (5434), 461–466.
(18) Urry, D. W.; Long, M. M.; Ohnishi, T.; Jacobs, M. *Biochem. Biophys. Res. Commun.* **1974**, *61* (4), 1427–1433.
(19) Cook, W. J.; Einspahr, H.; Trapane, T. L.; Urry, D. W.; Bugg, C. E. *J. Am. Chem. Soc.* **1980**, *102* (17), 5502–5505.
(20) Urry, D. W.; Shaw, R. G.; Prasad, K. U. *Biochem. Biophys. Res. Commun.* **1985**, *130* (1), 50–57.
(21) Petitdemange, R.; Garanger, E.; Bataille, L.; Bathany, K.; Garbay, B.; Deming, T. J.; Lecommandoux, S. *Bioconjug. Chem.* **2017**, *28* (5), 1403–1412.
(22) Petitdemange, R.; Garanger, E.; Bataille, L.; Dieryck, W.; Bathany, K.; Garbay, B.; Deming, T. J.; Lecommandoux, S. *Biomacromolecules* **2017**, *18* (2), 544–550.
(23) Rosselin, M.; Xiao, Y.; Belhomme, L.; Lecommandoux, S.; Garanger, E. *ACS Macro Lett.* **2019**, *8* (12), 1648–1653.
(24) Ibrahimova, V.; Zhao, H.; Ibarboure, E.; Garanger, E.; Lecommandoux, S. *Angew. Chemie - Int. Ed.* **2021**, *60* (27), 15036–15040.
(25) Xiao, Y.; Chinoy, Z. S.; Pecastaings, G.; Bathany, K.; Garanger, E.; Lecommandoux, S. *Biomacromolecules* **2020**, *21* (1), 114–125.
(26) Rosselin, M.; Chinoy, Z. S.; Bravo-Anaya, L. M.; Lecommandoux, S.; Garanger, E. *ACS Macro Lett.* **2021**, *10* (1), 65–70.
(27) Anaya, L. M. B.; Petitdemange, R.; Rosselin, M.; Ibarboure, E.; Garbay, B.; Garanger, E.; Deming, T. J.; Lecommandoux, S. *Biomacromolecules* **2021**, *22* (1), 76–85.
(28) Levêque, M.; Xiao, Y.; Durand, L.; Massé, L.; Garanger, E.; Lecommandoux, S. *Biomater. Sci.* **2022**, *10* (22), 6365–6376.
(29) Le Fer, G.; Portes, D.; Goudounet, G.; Guigner, J. M.; Garanger, E.; Lecommandoux, S. *Org. Biomol. Chem.* **2017**, *15* (47), 10095–10104.
(30) Dai, M.; Georgilis, E.; Goudounet, G.; Garbay, B.; Pille, J.; Van Hest, J. C. M.; Schultze, X.; Garanger, E.; Lecommandoux, S. *Polymers (Basel).* **2021**, *13* (9), 1–17.
(31) Dai, M.; Goudounet, G.; Zhao, H.; Garbay, B.; Garanger, E.; Pecastaings, G.; Schultze, X.; Lecommandoux, S. *Macromolecules* **2021**, *54* (1), 327–340.
(32) Bravo-Anaya, L. M.; Rosselgong, J.; Fernández-Solís, K. G.; Xiao, Y.; Vax, A.; Ibarboure, E.; Ruban, A.; Lebleu, C.; Joucla, G.; Garbay, B.; Garanger, E.; Lecommandoux, S. *Polym. Chem.* **2021**, *12* (2), 226–241.
(33) Zhao, H.; Ibrahimova, V.; Garanger, E.; Lecommandoux, S. *Angew. Chemie - Int. Ed.* **2020**, *59* (27), 11028–11036.

(34) Zhao, H.; Ibarboure, E.; Ibrahimova, V.; Xiao, Y.; Garanger, E.; Lecommandoux, S. *Adv. Sci.* **2021**, *8* (24), 2102508.
(35) Bian, Q.; Fu, L.; Li, H. *Nat. Commun.* **2022**, *13* (1), 137.
(36) Wei, G.; Su, Z.; Reynolds, N. P.; Arosio, P.; Hamley, I. W.; Gazit, E.; Mezzenga, R. *Chem. Soc. Rev.* **2017**, *46* (15), 4661–4708.
(37) Giaveri, S.; Schmitt, A. M.; Roset Julià, L.; Scamarcio, V.; Murello, A.; Cheng, S.; Menin, L.; Ortiz, D.; Patiny, L.; Bolisetty, S.; Mezzenga, R.; Maerkl, S. J.; Stellacci, F. *Adv. Mater.* **2021**, *33* (44), 2104581.
(38) Silva, N. H. C. S.; Vilela, C.; Marrucho, I. M.; Freire, C. S. R.; Pascoal Neto, C.; Silvestre, A. J. D. *J. Mater. Chem. B* **2014**, *2* (24), 3715.
(39) Abbas, M.; Lipiński, W. P.; Wang, J.; Spruijt, E. *Chem. Soc. Rev.* **2021**, *50* (6), 3690–3705.
(40) Salas-Ambrosio, P.; Tronnet, A.; Badreldin, M.; Ji, S.; Lecommandoux, S.; Harrisson, S.; Verhaeghe, P.; Bonduelle, C. *Polym. Chem.* **2022**, *13* (43), 6149–6161.
(41) Robinson, J. W.; Secker, C.; Weidner, S.; Schlaad, H. *Macromolecules* **2013**, *46* (3), 580–587.
(42) Klinker, K.; Barz, M. *Macromol. Rapid Commun.* **2015**, *36* (22), 1943–1957.
(43) Volpe, R. A.; Frisch, H. L. *Macromolecules* **1987**, *20* (8), 1747–1752.
(44) Ballistreri, A.; Giuffrida, M.; Maravigna, P.; Montaudo, G. *J. Polym. Sci. A1.* **1985**, *23* (6), 1731–1747.
(45) Fasman, G. D.; Blout, E. R. *Biopolymers* **1963**, *1* (2), 99–109.
(46) Xiao, Y.; Wang, J.; Zhang, J.; Heise, A.; Lang, M. *Biopolymers* **2017**, *107* (10), 1–9.
(47) Kricheldorf, H. R.; Tönnes, K.-U. *Makromol. Chemie. Macromol. Symp.* **1991**, *42–43* (1), 313–328.
(48) Kricheldorf, H.; Schilling, G. *Die Makromol. Chemie* **1978**, *179* (5), 1175–1191.
(49) Zou, J.; Zhou, M.; Ji, Z.; Xiao, X.; Wu, Y.; Cui, R.; deng, shuai; Liu, R. *Polym. Chem.* **2021**, 388–394.
(50) Detwiler, R. E.; Mcpartlon, T. J.; Coffey, C. S.; Kramer, J. R. *ACS Polym. Au* **2023**, No. "2023 Rising Stars".
(51) Detwiler, R. E.; Schlirf, A. E.; Kramer, J. R. *J. Am. Chem. Soc.* **2021**, *143* (30), 11482–11489.
(52) Hu, Y.; Tian, Z.-Y.; Xiong, W.; Wang, D.; Zhao, R.; Xie, Y.; Song, Y.-Q.; Zhu, J.; Lu, H. *Natl. Sci. Rev.* **2022**, *9* (8), nwac033.
(53) Koyama, Y.; Gudeangadi, P. G. *Chem. Commun.* **2017**, *53* (27), 3846–3849.
(54) Stiernet, P.; Couturaud, B.; Bertrand, V.; Eppe, G.; De Winter, J.; Debuigne, A. *Polym. Chem.* **2021**, *12* (14), 2141–2151.
(55) Kitamura, M.; Kakinoki, S.; Hirano, Y.; Oka, M. *Polym. Bull.* **2005**, *54* (4–5), 303–310.
(56) Kakinoki, S.; Kitamura, M.; Yuge, M.; Furuta, M.; Oka, M.; Hirano, Y.; Kono, K.; Kaetsu, I. *Polym. Bull.* **2007**, *58* (2), 393–400.
(57) Kitamura, M.; Yamauchi, T.; Oka, M.; Hayashi, T. *Polym. Bull.* **2003**, *51* (2), 143–150.
(58) Al Samad, A.; De Winter, J.; Gerbaux, P.; Jérôme, C.; Debuigne, A. *Chem. Commun.* **2017**, *53* (90), 12240–12243.
(59) Ihsan, A. Bin; Nargis, M.; Koyama, Y. *Int. J. Mol. Sci.* **2019**, *20* (18), 1–11.
(60) Komuro, N.; Nakajima, N.; Hamada, M.; Koyama, Y. *Polym. J.* **2022**, *54* (7), 903–912.
(61) Ihsan, A. Bin; Koyama, Y. *Polymer (Guildf).* **2019**, *161* (September 2018), 197–

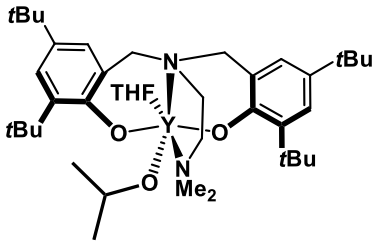
204.
(62) Zelzer, M.; Heise, A. *J. Polym. Sci. Part A Polym. Chem.* **2014**, *52* (9), 1228–1236.
(63) Zhu, H.; Zhu, M.; Shuai, S.; Zhao, C.; Liu, Y.; Liu, Y.; Li, X.; Rao, Z.; Li, Y.; Hao, J. *J. Sol-Gel Sci. Technol.* **2019**, *92* (3), 618–627.
(64) Mitra, S. B.; Patel, N. K.; Anderson, J. M. *Int. J. Biol. Macromol.* **1979**, *1* (2), 55–60.
(65) Kumar, A. *J. Macromol. Sci. Part A - Chem.* **1987**, *24* (6), 707–710.
(66) Wu, Y.; Zhang, D.; Ma, P.; Zhou, R.; Hua, L.; Liu, R. *Nat. Commun.* **2018**, *9* (1), 1–10.
(67) Salas-Ambrosio, P.; Tronnet, A.; Since, M.; Bourgeade-Delmas, S.; Stigliani, J. L.; Vax, A.; Lecommandoux, S.; Dupuy, B.; Verhaeghe, P.; Bonduelle, C. *J. Am. Chem. Soc.* **2021**, *143* (10), 3697–3702.
(68) Li, K.; Li, Z.; Shen, Y.; Fu, X.; Chen, C.; Li, Z. *Polym. Chem.* **2022**, *13* (5), 586–591.
(69) Chan, B. A.; Xuan, S.; Horton, M.; Zhang, D. *Macromolecules* **2016**, *49* (6), 2002–2012.

Perspectives: towards stereoselective ROP of NCAs and ROCOP systems for polyesteramide synthesis

1. Yttrium mediated ring opening polymerization of NCAs

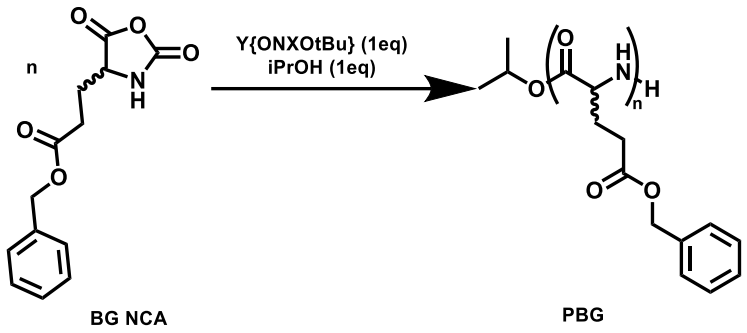
The many advances in the ROP of NCAs using the catalytic systems mentioned in Chapter **1** of this manuscript are largely inspired by previous ROP systems for instance developed with cyclic esters.[1] Among those catalytic systems, organometallic species have provided exquisite control over the tacticity or even the sequence of many polymers.[2] The use of organometallic species has shown many promise for the ROP of NCAs as already discussed in this manuscript (section **2.2.3** in chapter **1**, section **2.2.1** in chapter **3** and section **2.2.3** in chapter **4**) and this includes the use of catalysts based on nickel,[3] cobalt, iron,[4] ruthenium, iridium,[5] and rare earth metals such as yttrium[6]. Inspired by these early works, and in the framework of a collaborative effort with Prof. S. Guillaume's team from ISCR, we explored the use of alkoxy-amino-bis(phenolate) yttrium amide based catalytic complexes to instigate the ROP of NCAs. These catalysts have been used previously for the ROP of several cyclic monomers including lactide, β-lactones and morpholinediones.[7–9] The latter producing polyesteramides is particularly of interest to open the scope of this PhD study. Furthermore, yttrium amide based catalytic complexes have been used to perform stereoselective ROP of chiral monomers by choosing the right ligand.[7] This unexplored stereocontrol in the ROP of NCA can prove useful in accessing novel microstructures such as alternating copolymers (L/D) or new conformations.[10]

The catalytic complex Y{ONNO^{tBu}}/*i*PrOH (**Scheme 1**) has been used in this research project as it successfully achieved the ROP of morpholinediones[9] and similar complexes achieved the stereoselective ROP of cyclic esters[8]. The catalytic complex was prepared *in situ* by mixing tris(amido) precursor [Y(N(SiHMe$_2$)$_2$)$_3$(THF)$_2$] with the ligand {ONNO^{tBu,tBu}}, both received as such from our collaborators. This catalytic complex was then reacted with equimolar amount of isopropanol to produce the active species, Y{ONNO^{tBu}}/*i*PrOH (**Scheme 1**).



Scheme 1 : *Y{ONNO^{tBu}}/iPrOH catalyst/initiator complex*

Our goal was to establish i) that the alcoholate species could carry out an initiating nucleophilic attack on the NCA and ii) the catalytic efficacy of the yttrium complex towards the ROP of NCA. For this purpose, the ROP of BLG-NCA and its enantiomer BDG-NCA has been investigated using 2% of Y{ONNO^{tBu}}/*i*PrOH complex. This was done in dry THF with an M/I=50 for each enantiomer, independently, to shed light on the ROP and its mechanism (**Scheme 2**).



Scheme 2 : *ROP of benzyl glutamate NCA using Y{ONNO^{tBu}}/iPrOH to obtain polyglutamates with differing stereoregularity..*

Table 1 : *Glutamate polymers prepared via the ROP of NCAs initiated by Y{ONNO^{tBu}}/iPrOH and characterization results.*

Polymer	Structure[a]	Mn[b] (kg/mol)	Đ[c]	Mn[d] (kg/mol)	Đ[e]	End chain[f]
Polymer **1**	$PBLG_{50}$	11.5	1.24	4.3	1.26	isopropoxy
Polymer **2**	$PBDG_{50}$	13.9	1.05	5.0	1.01	isopropoxy
Polymer **3**	$PB(L/D)G_{100}$	15.2	1.14	9.3	1.01	isopropoxy

[a]As determined by the initial M/I ratio; [b]Calculated by dn/dC of PBLG in SEC DMF; [c]As measured by polymer peaks in MALS detection; [d]Calculated by dn/dC of PGA (deprotected form) in aqueous SEC; [e]As measured by polymer peaks in MALS detection; [f]as detected by MALDI-TOF analysis. All data can be consulted in the ***Supporting Information****.*

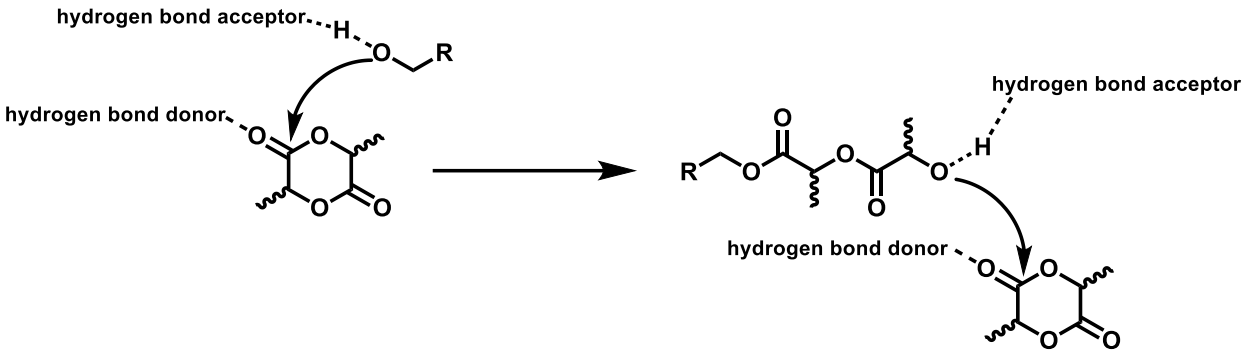
After purification by precipitation, ^{1}H NMR analysis (**Supporting Information**) and SEC analysis (**Table 1**) showed that the Y{ONNO^{tBu}}/*i*PrOH complex was able to promote the ROP of NCAs to produce polypeptides with predictable and controlled

molecular weights. To confirm the end-chain groups and that the alcoholate specie initiated the ROP, MALDI-TOF analyses were performed. Both polymer **1** and **2** showed a set of peaks corresponding to isopropoxy capped polypeptides (**Table 1** and **Figure S3**). To assess the catalyst's contribution to stereocontrol during propagation, a copolymerization with equal amounts of *L* and *D* enantiomers was also conducted in similar experimental conditions with an M/I=100 (polymer **3**, 1% of Y{ONNO^{tBu}}/*i*PrOH complex). The produced copolymer showed similar results as the homopolymers (**1-2**), with predictable Mn, controlled structure, and the isopropanol initiation. Unlike polylactide (PLA), the tacticity of the polypeptide is difficult to determine by ^{13}C NMR, as already described in chapter **2**. Nevertheless, the helicity of the polymer, which can be established by CD analysis, could be another way of defining the stereocontrol, particularly for syndiotactic polypeptide which should give access to peculiar ββ helix secondary structuring.[11] Nonetheless, the CD analysis of the deprotected form of polymer **3** was not conclusive as signals from enantiomers cancelled each other out (**Figure S4**) but this applies regardless of the microstructure (homochiral block blend or random copolymer). Non-racemic mixtures (BLG/BDG) were thus achieved in experimental conditions that were similar to polymer **3** and studied by CD at different times during propagation ("ROP" has been quenched by addition aqueous acidic solutions). If the catalytic complex is stereoselective, the polymerization will advance on the racemic mixture first leading to no change in conformation along the polymerization (as CD signals from enantiomers cancelled each other out). When the less concentrated monomer is totally consumed, the CD signal will increase with the advancement of the polymerization until completion. These experiments are still in progress. To further investigate the stereocontrol that Y{ONNO^{tBu}}/*i*PrOH complex) could bring during the ROP of NCAs, copolymerizing monomers with different side chains and chirality will also be contemplated. Initial results on the copolymerization of BDG-NCA with *L*-lysine based NCAs are being investigated. The kinetic study of these systems may provide direct information on the activity of the catalyst towards copolymerizing racemic mixtures without using CD by monitoring the *in-situ* NMR technique developed in chapter **2**. Furthermore, syndiotactic copolymers based on NCA monomers have never been achieved by ROP and may provide original structuring properties.[10,11] The results obtained until now confirmed the successful ROP of NCAs using yttrium-based catalysts. And the use of isopropanol as initiator was

possible using this catalytic system as confirmed by end-chain analysis in MALDI-TOF analysis (**Scheme 2**).

2. ROCOP of NCAs to produce polyesteramides

Many catalytic systems with alcohol initiators are at the heart of the ROP of lactone monomers. These are usually coupled with the use of organocatalysts with hydrogen bonding ability have also been used for the ROP on NCAs.[12–14] From the large selection of organocatalysts fluorinated alcohols[15], thioureas[16–18] and guanidines[19,20] can be mentioned as they have been efficient in the ROP of NCAs along with lactone monomers. In this PhD project, we have tried to use different catalytic systems to achieve statistical copolymerization of NCAs with cyclic esters. The incorporation of such repeat units in polypeptide backbone can improve their mechanical abilities, their degradability and their self-assembly although this has been proven only for diblock copolymers.[21–29] Our first attempt was to use Y{ONNO^{tBu}}/*i*PrOH complex (see previous paragraph) to achieve a ROCOP process in which NCAs would be copolymerize with other cyclic monomers to produce new classes of polyesteramides. In this direction, investigations on the copolymerization of NCAs with β-lactones[8] are ongoing in a collaborative effort with Prof. S. Guillaume's team from ISCR. Indeed, synthesis of polyesteramides containing amino acid units through ROCOP is still challenging, even though this methodology would be the best to tune the composition profile and microstructure. This could lead to block-like, gradient, or random copolymers, each with different properties. These polymers could be of significant interest for medical applications where degradable materials are required, as they will decompose into innocuous hydroxyacids (lactic acid, etc.) and amino acid degradation products. Beside the use of organometallic species, several syntheses using organocatalytic systems have also been tried to promote the ROCOP of NCAs in combination with racemic lactide (*rac*-LA). First the polymerization of *rac*-LA alone to produce polylactic acid (PLA) was achieved to benchmark the ROP reactivity (**Scheme 3**) with 7,8-diazabicyclo[5.4.0]undec-7-ene (DBU), DBU with thiourea cocatalyst, triazabicyclodecene (TBD) and zinc acetate with different alcohol initiators.[30–32]



Scheme 3 : *Simplified mechanism of the initiation/propagation steps of the ROP of rac-LA using alcohol initiator with a catalyst. Two catalytic mechanisms are proposed and can coexist; 1. activation of the monomer by a hydrogen-bond donor and 2. activation of the initiating/propagating species by hydrogen-bond acceptor.*[14,33]

DBU alone or coupled with thiourea were mixed in THF with *rac*-LA and isopropanol. TBD was used in both THF and DCM with benzylalcohol, as this catalyst works best with primary alcohols.[30] On another hand, zinc acetate with and without initiator was also used for the ROP of *rac*-LA in bulk. The best results obtained are shown in **Table 2**.

Table 2 : *Successful ROP conditions for rac-LA to produce PLA and characterization results. Polymerizations were run for 5 hours then quenched with acetic acid and precipitated from cold methanol.*

Polymer	Structure[a]	Initiator	Catalytic system	Solvent	Conversion after 5h[b]	Mn[c] (kg/mol) Đ[d]
PLA**1**	PLA_{200}	IsoPrOH	DBU 2eq	THF	Incomplete	8.7 (1.10)
PLA**2**	PLA_{200}	IsoPrOH	DBU 2eq + TU 2eq	THF	Incomplete	10.5 (1.09)
PLA**3**	PLA_{400}	BnOH	TBD 1eq	DCM	Complete[f]	30.0 (1.40)
PLA**4**	PLA_{400}	BnOH	TBD 1eq	THF	Complete[f]	34.5 (1.50)
PLA**5**	PLA_{400}	BnOH	$Zn(AcOH)_2.2H_2O$ 1eq	Bulk (145°C)	Complete[f]	15.9 (1.35)
PLA**6**	PLA_{400}	- [e]	$Zn(AcOH)_2.2H_2O$ 1eq	Bulk (145°C)	Complete[f]	19.8 (1.39)

[a]As determined by the initial M/I ratio; [b]As determined by ^1H NMR of a quenched aliquot in toluene-d by comparing residual monomer peaks to polymer peaks' integrations; [c]Calculated by PS calibration in THF SEC; [d]As measured by polymer peaks in RI detection; [e]Water is considered the initiator[32]*; [f]As determined by ^1H NMR of a quenched aliquot in CD_2Cl_2 by comparing residual monomer peaks to polymer peaks' integrations. IsoPrOH=isopropanol; BnOH=benzylalcohol. All data can be consulted in the* ***Supporting Information****.*

The results shown in **Table 2** are in line with the literature and PLA was successfully produced in different reaction conditions, solvents and bulk, and with different catalytic mechanisms (**Scheme 3**). Although the catalytic loading in most these cases was high, the use of DBU alone or with a cocatalyst did not proceed to completion as it would have been expected after 300 min (PLA**1-2** in **Table 2**).[30] The use of the more basic TBD (PLA**3-4**) allowed to obtain complete polymerization of *rac*-LA that showed higher dispersities when analysed with SEC (**Table 2**). In bulk conditions using $Zn(AcOH)_2.2H_2O$ as a catalyst for the ROP using BnOH as initiator (PLA**5**) or as a catalyst/initiator complex (PLA**6**). The polymerization proved complete by ^1H NMR and SEC analysis showed the formation of polymers with adequate dispersities (**Table 2**).

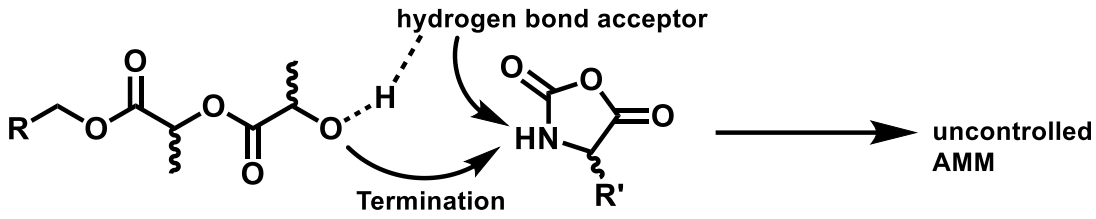
In a subsequent step, copolymerizations with BLG-NCAs have been investigated (as already seen in the first part of this perspective, this monomer is a gold standard to study the ROP of polypeptides). A series of copolymers of *rac*-LA and BLG-NCA have thus been synthesized using TBD and zinc acetate systems. TBD-based ROCOP was achieved using benzylalcohol as initiator in both THF and DCM. Zinc acetate was used in bulk with or without catalyst by solubilizing the NCA in the molten lactide at 145°C. Zinc acetate is also a known catalyst for the ROP of BLG-NCA in solution (DCM)[32] and this condition was also tested for the copolymerization (**Table 3**).

Table 3 *: Experimental ROP conditions for the copolymerization of rac-LA with BLG NCA. to produce PLA and characterization results. Polymerizations were run for adequate amount of time then quenched with acetic acid and precipitated from cold methanol.*

Polymer	Structure[a]	Initiator	Catalytic system	Solvent	Conversion[b]	Mn[c] (kg/mol) Đ[d]
Copolymer **1**	$PBLG_{100}/PLA_{200}$	BnOH	TBD 1eq	THF	No conversion	
Copolymer **2**	$PBLG_{100}/PLA_{200}$	BnOH	TBD 1eq	DCM	Polypeptide only	50 (1.48)
PBLG**1**	$PBLG_{200}$	BnOH	TBD 1eq	DCM	Complete	9.7 (1.07)/ 50 (1.48)
Copolymer **3**	$PBLG_{100}/PLA_{200}$	BnOH	$Zn(AcOH)_2.2H_2O$ 1eq	DCM	No conversion	
Copolymer **4**	$PBLG_{100}/PLA_{200}$	BnOH	$Zn(AcOH)_2.2H_2O$ 1eq	Bulk (145°C)	Polypeptide only	Oligomers[f]
Copolymer **5**	$PBLG_{100}/PLA_{200}$	-[e]	$Zn(AcOH)_2.2H_2O$ 1eq	Bulk (145°C)	Polypeptide only	Oligomers[f]

[a]As determined by the initial M/I ratio; [b]As determined by [1]H NMR of a quenched aliquot in dichloromethane by comparing residual monomer peaks to polymer peaks' integrations and FTIR spectra for amide bonds; [c] Calculated by dn/dC of PBLG in DMF SEC; [d]As measured by polymer peaks in MALS detection; [e]Water is considered the initiator[32]; [f]As detected by RI in THF SEC; BnOH=benzylalcohol; IsoPrOH=isopropanol; Y{ONNOtBu}= yttrium based catalyst. All data can be consulted in the ***Supporting Information****.*

All the copolymers were analysed by FTIR, ^{1}H NMR and SEC. First, the NCA's carbonyl bands were monitored by IR spectroscopy to assess the conversion and the formation of peptide bonds. Later, using ^{1}H NMR in either toluene-d or CD_2Cl_2 allowed to monitor the disappearance of *rac*-LA and the formation of PLA. These both analyses (**Table 3**) showed that TBD only worked in DCM for both the copolymer **2** and the homopolymer of BLG used as a control experiment (PBLG **1**). Nonetheless, the copolymerization of BLG-NCA with *rac*-LA did not show the formation of polyester and only polypeptide was formed. Furthermore, the homopolymer (PBLG **1**) showed two populations in SEC analysis suggesting two initiation systems or transesterification occurring (which is possible considering the side chain of the BLG monomer units). These results need more investigation to understand the mechanism involved in the ROCOP. The use of $Zn(AcOH)_2.2H_2O$ in DCM showed no conversion of both the cyclic ester or the NCA (copolymer **4**). When this ROCOP was underwent in bulk at 145°C, polypeptides were formed but their SEC analysis showed only small oligomeric species suggesting that the ROP of NCAs might not be efficient at such high temperatures. Overall, these results show that although the organocatalytic systems may work for the ROP of cyclic esters and NCAs separately, they failed to copolymerize these monomers together. When copolymerizations were undergone, no polyesters were produced and only polyamides were identified. This could be due to the termination of the propagating species from rac-LA (**Scheme 4**). Furthermore, the produced polypeptides have uncontrolled structure and dispersities hinting to the fact that the deprotonation of the amine produces the activated monomer mechanism (**Scheme 4**). Further investigation in the mechanisms of copolymerization and possible interactions between the different propagating species is needed. Notably the interplay between alcoholate, the protons on the NCA and the propagating primary amine must be considered to be able to ROCOP NCAs with cyclic esters. Furthermore, the ester side chain of BLG NCA may undergo transesterification hindering this mechanism.



Scheme 4 : *Possible undesired reactions during the ROCOP of rac-LA with NCAs using organobases as catalysts. In both cases the activated monomer mechanism (AMM) may proceed with unctrolled ROP of NCAs.*

Pro-NCA was thus used for copolymerization with *rac*-LA. The lack of hydrogen on the amine in this NCA motivated this choice. Although the DBU system has been shown to polymerize efficiently NNCAs[34], this system failed to copolymerize *rac*-LA with Pro-NCA in already described conditions (data not presented). Another catalytic system proven for both the ROP of NNCAs[20] and lactide[35] was TMG, but our efforts failed to produce either PLA or copolymers based on *rac*-LA and Pro-NCA (data not presented). The issues raised in Chapter **3** about the polymerization of Pro-NCA in organic media has led us to put off further investigation into this copolymerization system. Simpler NNCAs can also be used, such as sarcosine to assess the possibility of such a ROCOP. This is certainly an interesting perspective to develop towards polyesteramides.

Supporting information

Typical ROP of BLG NCA using Y{ONNO^{tBu}}/*i*PrOH

For a typical polymerization, the NCA monomer of γ-benzyl-L-glutamate (BLG NCA in Scheme 1, 100mg, 0.38mmol, 50 equivalent) was weighed in a flame dried Schlenk equipped with a magnetic stirrer in a glovebox under pure argon. Then 1 mL of anhydrous THF (0.36 M) were added. To this solution, 41µL of a 1% v/v solution of isopropanol in dry THF (7.6*10-3 mmol, 1 equivalent) were added. Simultaneously, in a dry, clean flask, tris(amido) precursor [Y(N(SiHMe$_2$)$_2$)$_3$(THF)$_2$] (20mg, 0.032 mmol) and ligand {ONNO^{tBu,tBu}} (16.7 mg, 0.032 mmol) (both received and used as are from collaborators) were dissolved in 250µL of dry THF obtaining a 0.13M solution of the catalytic form Y{ONNO^{tBu}}. 60µL of this solution was added to the reaction mixture (7.6*10^{-3} mmol, 1 equivalent). The Schlenk was capped with a septum pierced with a syringe to let out CO2 and left to stir overnight in the glovebox. FTIR allowed to confirm

the completion of the reaction by monitoring the disappearance of the bands at 1857cm-1 and 1785cm-1 characteristic of the C=O stretching of NCAs. At the end of the ROP, the reaction mixture had highly increased viscosity in THF. The reaction was quenched with 0.5ml of acetic acid 1.6M solution (0.8mmol) then the polymer was recovered by precipitation in cold diethyl ether and dried under high vacuum. Sample was then analysed by ^{1}H NMR (5 mg/ml in $CDCl_3$ containing 15% trifluoroacetic acid). ^{1}H-NMR of the polypeptide backbone (Annex 1 and 2, 400MHz, δ, ppm): 1.84-2.19 (m, 2H, CH2), 2.44 (t, 2H, CH2), 4.59 (m, 1H, CH), 5.03-5.11 (m, 2H, CH2O), 7.25-7.29 (m, 5H, ArH), 7.85 (d, 1H, NH) (**Figure S1**).

Polymer 1-3 characterizations

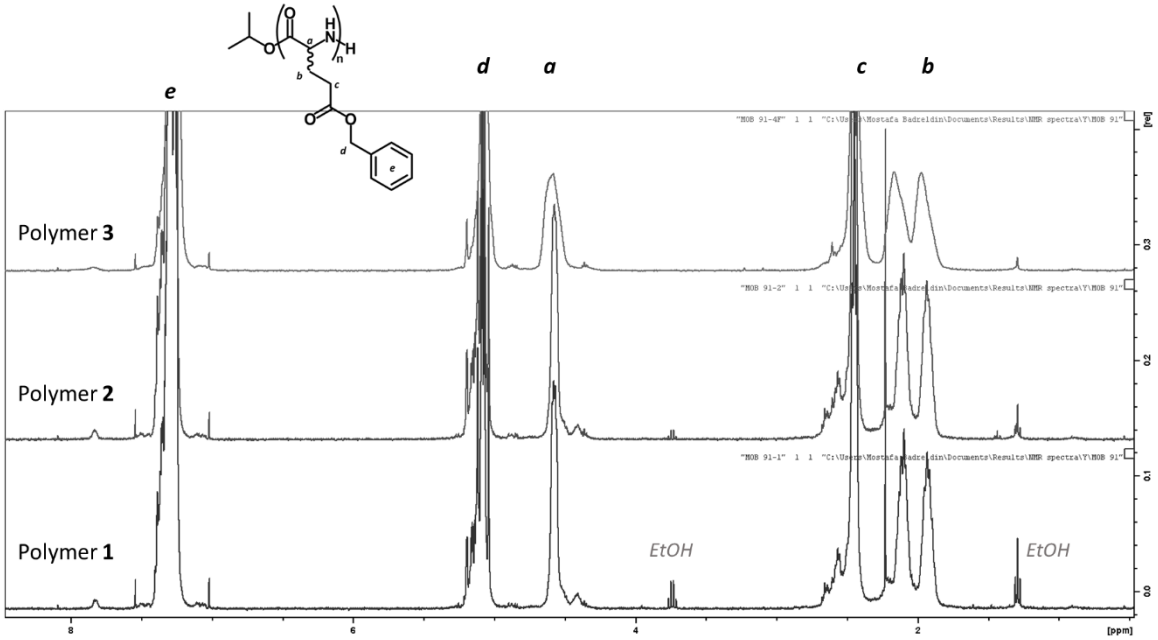


Figure S1 *: ^{1}H NMR for polymer **1-3** in $CDCl_3$+15%TFA. The structures of the polymers is drawn with the peak assignation. The isopropoxy end chain could not be identified by this technique.*

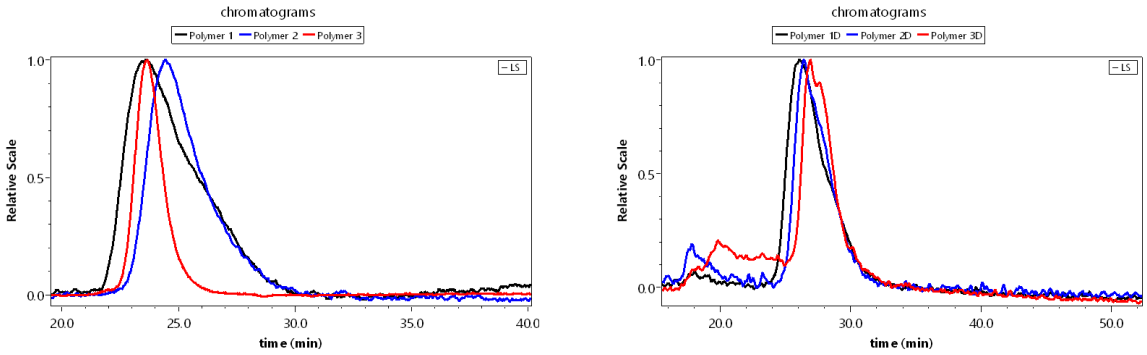


Figure S2 : *(on the left) DMF SEC of polymer **1-3**, (on the right) aqueous SEC of polymers **1D-3D**. MALS detection is shown.*

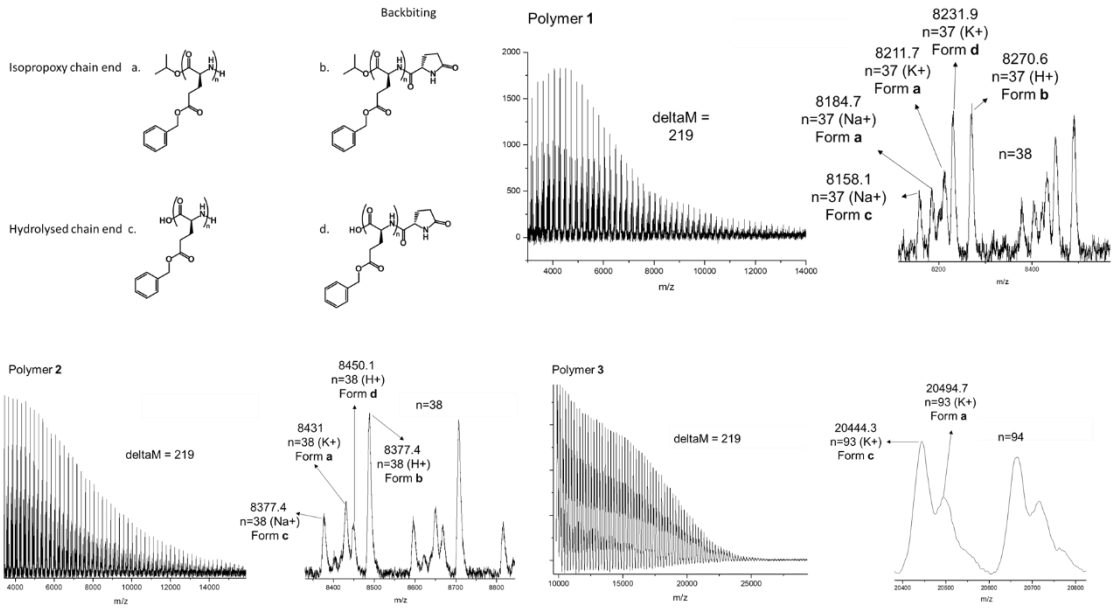


Figure S3 : *MALDI-TOF analysis using DCTB of polymer **1-3** results. The different detected structures are drawn (forms a-d). Forms a and b prove the presence of the isopropoxy chain end confirming the initiations using the catalytic complex Typical ROP of BLG NCA using Y{ONNO^{tBu}}/iPrOH. Forms c and d can be the result of hydrolysis during the ionization of the analysis. Forms b and d are common for benzyl glutamate polymers in solution at room temperature.*

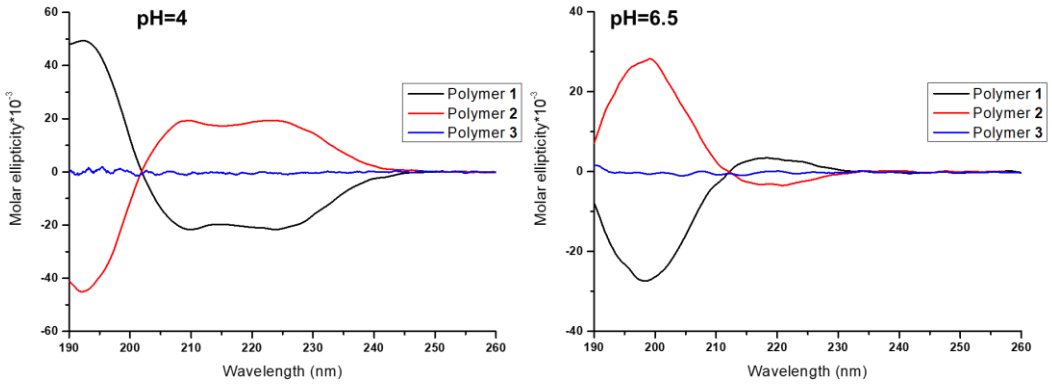


Figure S4 : *CD spectra of deprotected polymers (**1D-3D**) at different pH*

PLA1-6 characterization

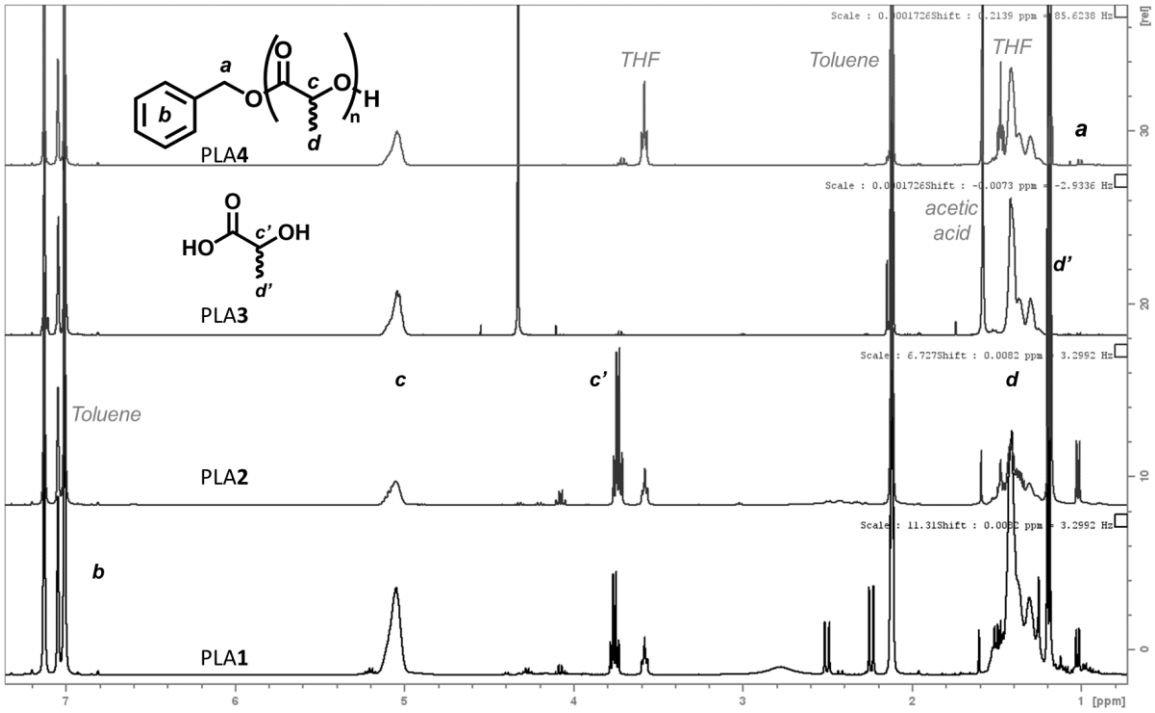


Figure S5 : *^{1}H NMR for PLA**1-4** in toluene-d. Aliquots after 5 hours of reaction were quenched with acetic acid and the solvents evaporated under vacuum before characterization. The structures of the polymer and the unreacted monomers is drawn with the peak assignation.*

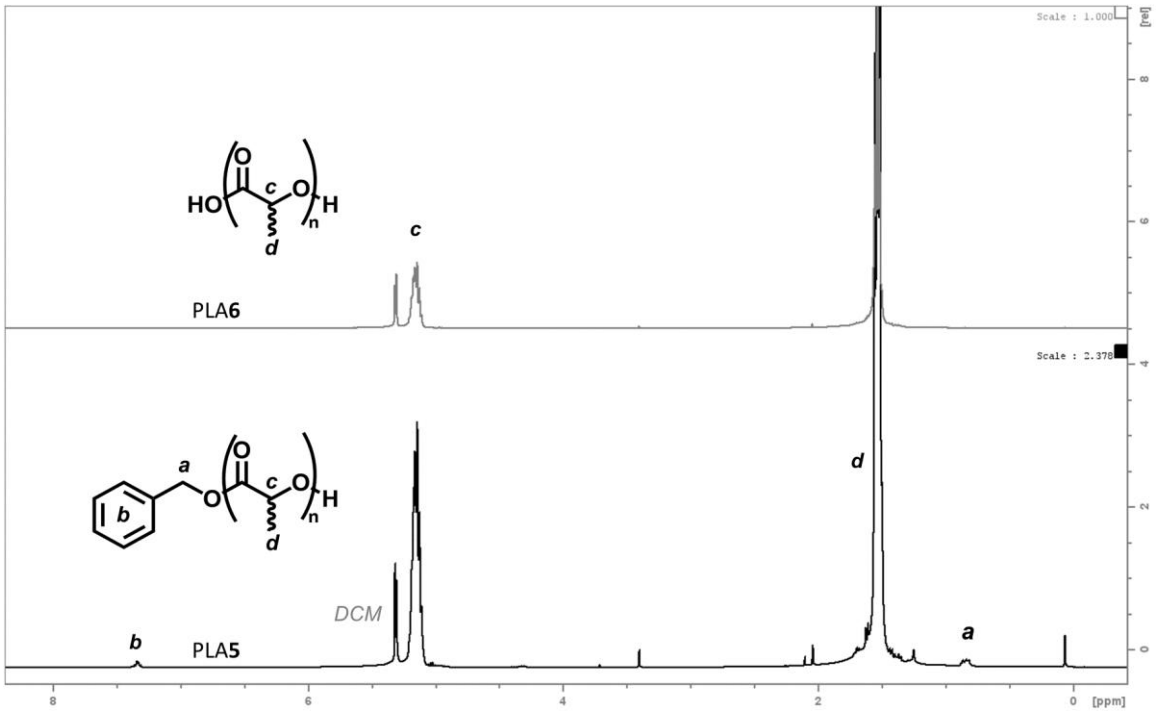


Figure S6 : *^{1}H NMR for PLA**5-6** in CD_2Cl_2. Aliquots after 12 hours of reaction were quenched with acetic acid and the solvents evaporated under vacuum before characterization. The structures of the polymers are drawn with the peak assignation.*

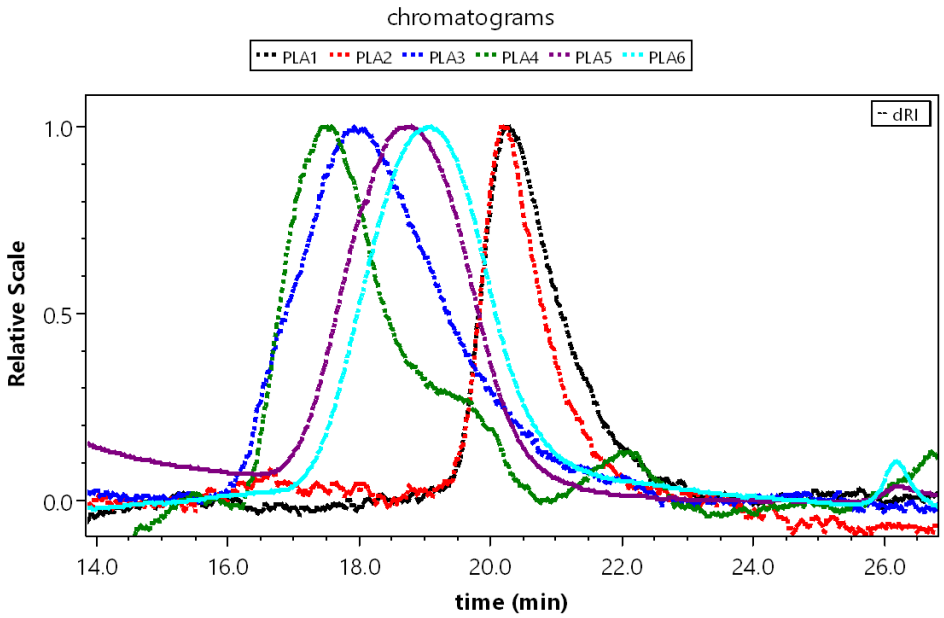


Figure S7 : *THF SEC chromatograms of PLA**1-6**. RI detection is shown.*

Copolymers 1-5 characterizations

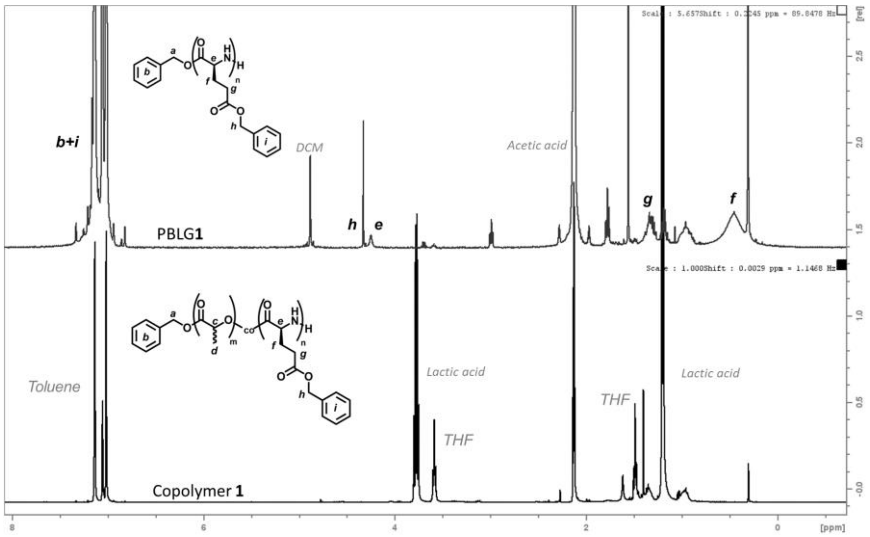


Figure S8 : *^{1}H NMR for copolymer **1** and PBLG **1** (only slightly soluble) in toluene-d. Aliquots after 5 hours of reaction were quenched with acetic acid and the solvents evaporated under vacuum before characterization. The structures of the polymers are drawn with the peak assignation.*

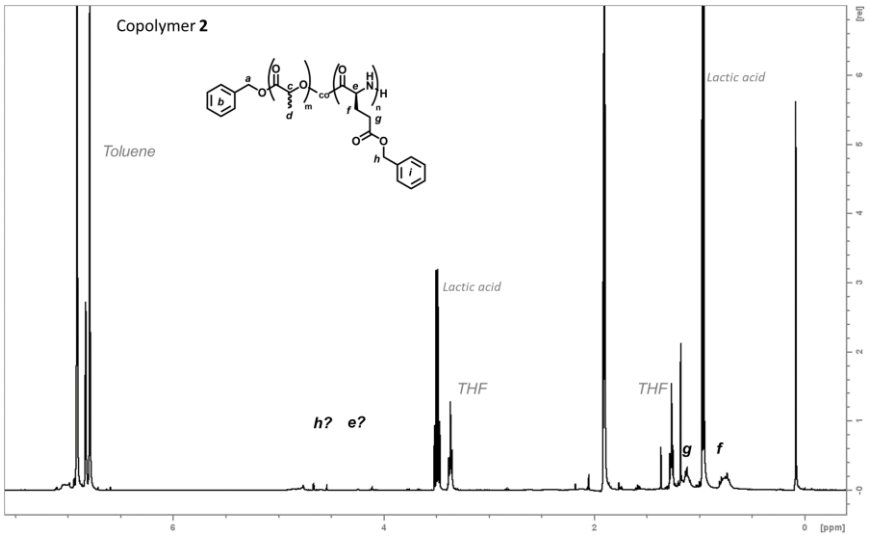


Figure S9 : *^{1}H NMR for copolymer **2** (only slightly soluble) in toluene-d. Aliquot after 12 hours of reaction were quenched with acetic acid and the solvents evaporated under vacuum before characterization. The structures of the polymers are drawn with the peak assignation. Polypeptide peaks were not easily identifiable, and no PLA peaks were visible.*

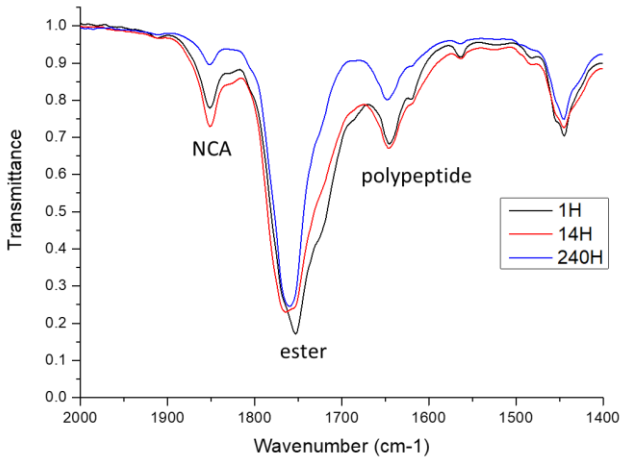


Figure S10 : *FTIR monitoring of copolymer **1** shows small evolution of NCA and amide bonds over a long period of reaction time.*

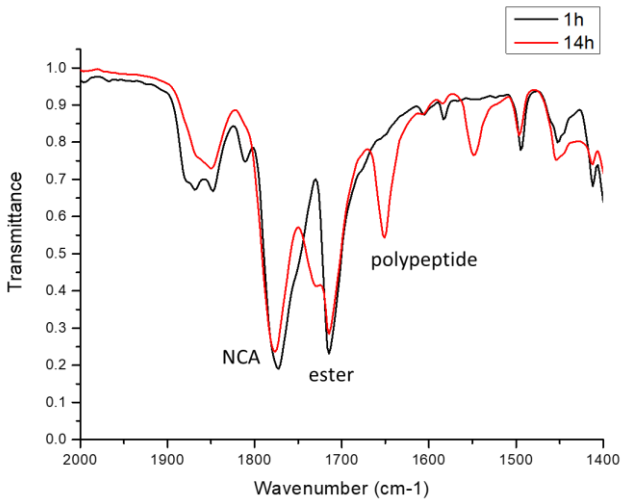


Figure S11: *FTIR monitoring of PBLG **1** shows the formation of amide bonds.*

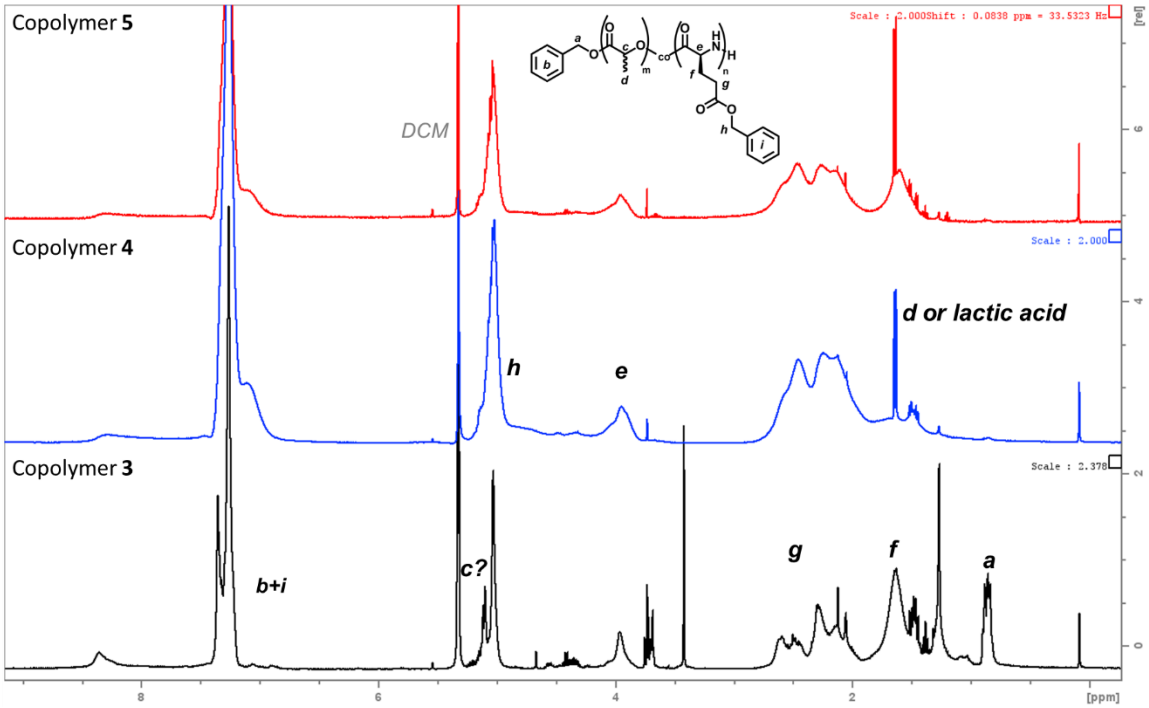


Figure S12 *: ^{1}H NMR for copolymers **3-5** in CD_2Cl_2. Aliquots after 12 hours of reaction were quenched with acetic acid and the solvents evaporated under vacuum before characterization. The structures of the polymers are drawn with the peak assignation.*

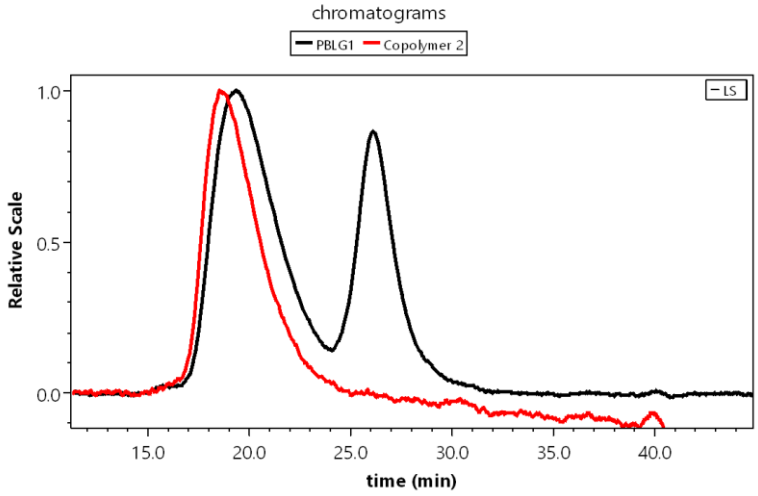


Figure S13 *: DMF SEC of PBLG **1** and copolymer **2**. MALS detection is shown. These polymers were not soluble in THF.*

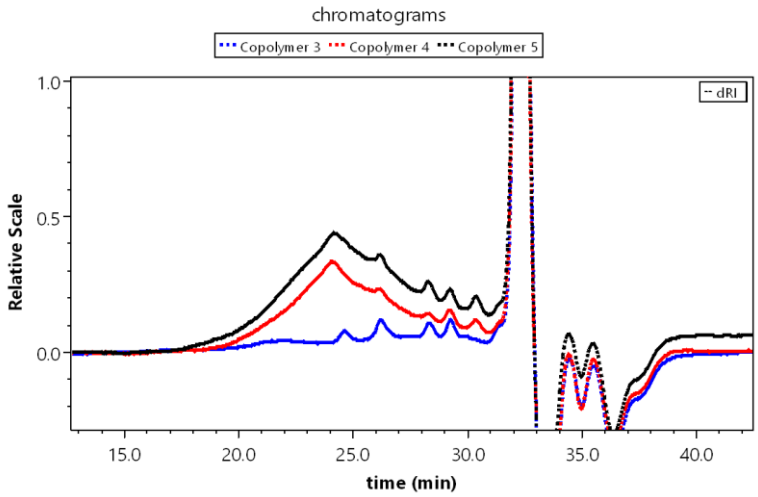


***Figure S14** : THF SEC of copolymers **3-5**, RI detection is shown. Only small molecular weight species can be detected.*

References

(1) Rasines Mazo, A.; Allison-Logan, S.; Karimi, F.; Chan, N. J. A.; Qiu, W.; Duan, W.; O'Brien-Simpson, N. M.; Qiao, G. G. *Chem. Soc. Rev.* **2020**, *49* (14), 4737–4834.

(2) Williams, C. K. *Chem. Soc. Rev.* **2007**, *36* (10), 1573.

(3) Deming, T. J. *Nature* **1997**, *390* (6658), 386–389.

(4) Deming, T. J. *Macromolecules* **1999**, *32* (13), 4500–4502.

(5) Seidel, S. W.; Deming, T. J. **2003**, *1* (Table 1), 969–972.

(6) Peng, H.; Ling, J.; Shen, Z. *J. Polym. Sci. Part A Polym. Chem.* **2012**, *50* (6), 1076–1085.

(7) Carpentier, J. F. *Organometallics* **2015**, *34* (17), 4175–4189.

(8) Ligny, R.; Guillaume, S. M.; Carpentier, J. F. *Chem. - A Eur. J.* **2019**, *25* (25), 6412–6424.

(9) Castro, P. M.; Zhao, G.; Amgoune, A.; Thomas, C. M.; Carpentier, J. F. *Chem. Commun.* **2006**, *1* (43), 4509–4511.

(10) Kessler, N.; Schuhmann, H.; Morneweg, S.; Linne, U.; Marahiel, M. A. *J. Biol. Chem.* **2004**, *279* (9), 7413–7419.

(11) Zhang, X.; Fevre, M.; Jones, G. O.; Waymouth, R. M. *Chem. Rev.* **2018**, *118* (2), 839–885.

(12) Bourissou, D.; Moebs-Sanchez, S.; Martín-Vaca, B. *Comptes Rendus Chim.* **2007**, *10* (9), 775–794.

(13) Thomas, C.; Bibal, B. *Green Chem.* **2014**, *16* (4), 1687–1699.

(14) Zhao, W.; Lv, Y.; Li, J.; Feng, Z.; Ni, Y.; Hadjichristidis, N. *Nat. Commun.* **2019**,

10 (1), 3590.

(15) Wang, K.-Y.; Li, Z.-Q.; Xu, M.-N.; Li, B. *Macromolecules* **2023**, *56* (14), 5599–5609.

(16) Lv, W.; Wang, Y.; Li, M.; Wang, X.; Tao, Y. *J. Am. Chem. Soc.* **2022**, *144* (51), 23622–23632.

(17) Zhao, W.; Gnanou, Y.; Hadjichristidis, N. *Polym. Chem.* **2015**, *6* (34), 6193–6201.

(18) Li, K.; Li, Z.; Shen, Y.; Fu, X.; Chen, C.; Li, Z. *Polym. Chem.* **2022**, *13* (5), 586–591.

(19) Chan, B. A.; Xuan, S.; Horton, M.; Zhang, D. *Macromolecules* **2016**, *49* (6), 2002–2012.

(20) Planellas, M.; Puiggalí, J. *Int. J. Mol. Sci.* **2014**, *15* (8), 13247–13266.

(21) Cui, S.; Wang, X.; Li, Z.; Zhang, Q.; Wu, W.; Liu, J.; Wu, H.; Chen, C.; Guo, K. *Macromol. Rapid Commun.* **2014**, *35* (22), 1954–1959.

(22) Cui, S.; Pan, X.; Gebru, H.; Wang, X.; Liu, J.; Liu, J.; Li, Z.; Guo, K. *J. Mater. Chem. B* **2017**, *5* (4), 679–690.

(23) Deng, C.; Rong, G.; Tian, H.; Tang, Z.; Chen, X.; Jing, X. *Polymer (Guildf).* **2005**, *46* (3), 653–659.

(24) Price, D. J.; Khuphe, M.; Davies, R. P. W.; McLaughlan, J. R.; Ingram, N.; Thornton, P. D. *Chem. Commun.* **2017**, *53* (62), 8687–8690.

(25) Peng, H.; Xiao, Y.; Mao, X.; Chen, L.; Crawford, R.; Whittaker, A. K. *Biomacromolecules* **2009**, *10* (1), 95–104.

(26) Guo, A.; Yang, W.; Yang, F.; Yu, R.; Wu, Y. *Macromolecules* **2014**, *47* (16), 5450–5461.

(27) Caillol, S.; Lecommandoux, S.; Mingotaud, A.-F.; Schappacher, M.; Soum, A.; Bryson, N.; Meyrueix, R. *Macromolecules* **2003**, *36* (4), 1118–1124.

(28) Luppi, L.; Babut, T.; Petit, E.; Rolland, M.; Quemener, D.; Soussan, L.; Moradi, M. A.; Semsarilar, M. *Polym. Chem.* **2019**, *10* (3), 336–344.

(29) Lohmeijer, B. G. G.; Pratt, R. C.; Leibfarth, F.; Logan, J. W.; Long, D. A.; Dove, A. P.; Nederberg, F.; Choi, J.; Wade, C.; Waymouth, R. M.; Hedrick, J. L. *Macromolecules* **2006**, *39* (25), 8574–8583.

(30) Dove, A. P. *ACS Macro Lett.* **2012**, *1* (12), 1409–1412.

(31) Gowda, R. R.; Chakraborty, D. *J. Mol. Catal. A Chem.* **2010**, *333* (1–2), 167–172.

(32) Bonduelle, C.; Martín-Vaca, B.; Cossío, F. P.; Bourissou, D. *Chem. - A Eur. J.* **2008**, *14* (17), 5304–5312.

(33) Li, A.; Lu, L.; Li, X.; He, L. L.; Do, C.; Garno, J. C.; Zhang, D. *Macromolecules* **2016**, *49* (4), 1163–1171.

(34) Kuroishi, P. K.; Dove, A. P. *Chem. Commun.* **2018**, *54* (49), 6264–6267.

General conclusion and perspectives

In this study, we used the ring-opening polymerization (ROP) of N-carboxyanhydrides (NCA) to synthesize polymers composed of amino acids with structural and functional characteristics similar to those of intrinsically disordered proteins (IDPs). After describing the state of the art regarding this polymerization methodology, we first explored the distribution of glycine units within a poly(glutamic acid-*co*-glycine) P(GA-*co*-Gly) chain obtained through copolymerization. The results confirmed that the statistical distribution of these glycine units could be controlled by modifying the polymerization conditions (such as solvent or reaction temperature, initial monomer ratio, or reaction mechanism). This "stochastic" control over the sequence of the primary structure of a copolypeptide that includes glycine allows the preparation of copolymers with a strong composition gradient that are devoid of secondary structure, similar to the disordered regions of IDPs. In the second part, we investigated another amino acid commonly found in IDP sequences, proline. We first studied the hysteretic behaviour of the lower critical solution temperature (LCST) of polyproline (PLP) homopolymer. We observed that this behaviour depended on the secondary structure and molecular weight of the polymers. We also explored potential applications of this unique behaviour, including designing temperature memory sensors, thermosensitive nanoparticles, or hydrogel solutions with memory effects. Inspired by the prevalence of proline in thermosensitive IDPs like elastin-like polypeptides (ELP), we copolymerized proline with the aim of creating periodic sequences similar to ELPs that could result in an LCST. We successfully copolymerized proline NCA with glycine NCA using water-assisted mechanisms. This synthesis produced copolymers with a very strong gradient. By controlling the overall molecular weight of these copolymers, we obtained thermosensitive nanoparticles with an LCST behaviour devoid of hysteresis. This demonstrates the potential of using glycine to modulate the thermosensitivity of polyproline. We also explored other ROP mechanisms, such as catalysis by LiHMDS and TMG in THF, which allowed us to obtain both PLPs and copolymers with the desired structures. Further research on copolymerization mechanisms, copolymer structures, and thermosensitivity is planned to gain a complete understanding of this system. Lastly, we studied the terpolymerization of proline, glycine, and valine, which make up the periodic sequence of ELPs and are responsible for their thermosensitive

behaviour. We successfully developed optimal conditions for producing polymers with controlled structures (water-assisted ROP and organometallic catalysis), paving the way for the synthesis of thermosensitive LCST polymers composed entirely of natural amino acids. The knowledge gained during this study on the copolymerization of NCAs and the catalysis of their ROP has also inspired future research in the laboratory and yielded preliminary results. These future investigations will focus on the stereocontrolled ROP of NCAs and the ring-opening copolymerization (ROCOP) of NCAs with cyclic esters to prepare polyesteramides.

Conclusions générales et perspectives

Dans cette étude, nous avons utilisé la polymérisation par ouverture de cycle (ROP) des *N*-carboxyanhydrides (NCA) pour synthétiser des polymères constitués d'acides aminés présentant des caractéristiques structurales et fonctionnelles similaires aux protéines intrinsèquement désordonnées (IDP). Après avoir décrit l'état de l'art concernant cette méthodologie de polymérisation, nous avons exploré dans une première partie la distribution d'unités de glycine au sein d'une chaîne de poly(acide glutamique-*co*-glycine) P(GA-*co*-Gly) obtenue par copolymérisation. Les résultats obtenus nous confirment la possibilité de contrôler la distribution statistique de ces unités glycine en modifiant les conditions de polymérisation (le solvant ou la température de réaction, le ratio initial des monomères ou le mécanisme de réaction). Ce contrôle de la séquence « stochastique » de la structure primaire d'un polymère peptidique incluant de la glycine permet de préparer des copolymères à haut gradient, dépourvus de structure secondaire, une propriété structurale similaire de régions désordonnées d'IDP riches en glycine. Dans une deuxième partie, nous nous sommes intéressés à un autre acide aminé que l'on retrouve dans la séquence des IDPs, la proline. Nous avons d'abord étudié le comportement hystérétique de la température de solution critique inférieure (LCST) de l'homopolymère de la proline, la polyproline (PLP). Nous avons pu observer que ce comportement dépend de la structuration secondaire et de la masse moléculaire du polymère. Nous avons aussi exploré les applications potentielles de ce comportement unique, notamment pour concevoir des capteurs de température à mémoire, des nanoparticules thermosensibles ou des solutions hydrogélifiantes possédant un effet mémoire. Nous nous sommes enfin inspirés de la prévalence de la proline dans des IDP thermosensibles de type élastine (ELP), afin de mimer avec des séquences statistiques la séquence périodique de ces protéines possédant une LCST. Nous avons par exemple réussi à copolymériser le NCA de proline avec le NCA de glycine en utilisant un mécanisme réactionnel assisté par l'eau. Cette synthèse produit des copolymères à très fort gradient. En contrôlant leur masse moléculaire, nous avons obtenu des nanoparticules thermosensibles avec un comportement LCST original démontrant le potentiel de la glycine pour moduler la thermosensibilité d'un bloc de PLP. Nous avons également exploré d'autres mécanismes de ROP, tels que la catalyse par LiHMDS et TMG dans le THF, qui ont permis d'avoir à la fois des PLP et des copolymères avec des microstructures variées

notamment incluant de la valine. Finalement, nous avons étudié la terpolymérisation de la proline, de la glycine et de la valine, qui constituent la séquence périodique des ELP et qui est à l'origine de leur comportement thermosensible. Nous avons développé avec succès des conditions optimales pour produire des polymères à structures contrôlées (ROP assistée par l'eau et catalyse organométallique) en ouvrant ainsi la voie à la synthèse de polymères LCST thermosensibles entièrement constitués d'acides aminés naturels. Les connaissances acquises au cours de cette étude sur la copolymérisation des NCA et la catalyse de leur ROP ont aussi inspiré les futures recherches du laboratoire et donné lieu à des résultats préliminaires. Ces recherches se focaliseront sur la ROP stéréo-contrôlée des NCA et la copolymérisation par ouverture de cycle (ROCOP) des NCA avec des esters cycliques pour préparer des polyesteramides.

Les protéines sont une classe très importante de biomacromoléules naturelles. Elles remplissent de nombreuses fonctions dans les systèmes vivants mais sont composés de seulement 20 types d'acides aminés. Mimer la structure chimique des protéines en produisant des polypeptides synthétiques est d'un grand intérêt pour produire des matériaux fonctionnels, biodégradables et biocompatibles. Le moyen le plus économique et le plus efficace pour produire des polymères à base d'acides aminés est la polymérisation à ouverture de cycle des N-carboxyanhydrides (NCA) dérivés de ces acides aminés. L'objectif de cette thèse de doctorat a été d'imiter la structure et les propriétés de protéines intrinsèquement désordonnées en utilisant cette polymérisation. La séquence des acides aminés dans ces protéines est relativement simple si on la compare à d'autres régions du protéome et se caractérise par des séquences répétitives enrichies en glycine et en proline. En utilisant les monomères NCA dérivés de ces deux acides aminés, ce travail de doctorat développe la préparation de copolymères statistiques par copolymérisation pour produire des polypeptides avec une séquence monomère contrôlée stochastiquement. Un tel contrôle permet d'induire de manière statistique certaines séquences (AA, AB, BB) dans la distribution des monomères le long de la macromolécule. Les études cinétiques d'un modèle de copolymérisation comprenant de la glycine et un glutamate protégé ont d'abord permis de calculer leurs rapports de réactivité respectifs. Le contrôle des différentes conditions de réaction (solvant, température, catalyse) a permis d'orienter stochastiquement le système de copolymérisation vers différentes séquences en monomères. L'influence de cette séquence sur les propriétés macromoléculaires, notamment la solubilité et la structure secondaire, a ensuite été élucidé. Dans un deuxième temps, l'approche biomimétique basée sur la chimie des polymères a été utilisé pour reproduire la séquence de protéines thermosensibles riches en glycine et en proline. En biologie, les polypeptides de type élastine (ELP), sont des protéines thermosensibles qui présentent un caractère thermosensible proche de la température corporelle (37 °C) et sont donc utilisés dans de nombreuses applications biomédicales. Cette classe de protéine a donc été choisie comme modèle de séquence riche en glycine et en proline à mimer. Des homopolymères de polyproline (PLP) de poids moléculaires élevés ont d'abord été préparés en milieu aqueux et étudiés pour comprendre leur comportement thermosensible peu exploré et révélant une hystérèse exceptionnellement large. Avec ce polymère, la structure, la chiralité et la structuration secondaire sont des paramètres importants qui explique un effet mémoire qui a permis de produire des formulations originales telles que des sondes de température ou des hydrogels programmables. Dans des réactions de copolymérisations en milieu aqueux, introduire de la proline avec de la glycine permet d'affiner son comportement thermosensible. En réalisant la polymérisation dans l'eau, cette copolymérisation produit des macromolécules à gradient élevé dont on peut ajuster le comportement thermosensible. Ces copolymères s'auto-assemblent spontanément en nanoparticules thermosensibles. Des réactions de polymérisation plus complexes, des terpolymérisations entre la glycine, la proline et la valine ont ensuite été explorées pour mieux imiter le comportement des ELP, mais cette réaction a montré certaines limites et des problèmes de solubilités. Finalement, d'autres approches de polymérisation ont été explorés, y compris l'organocatalyse et la catalyse organométallique. Ces systèmes permettront d'obtenir des copolymères contenant de la proline avec une thermosensibilité plus modulable.

Mots clés :
Polymérisation à ouverture de cycle, Monomères N-carboxyanhydride, Contrôle de séquence en copolymérisation, Polypeptides thermosensibles, Copolymères biomimétiques

Unité de recherche

Laboratoire de Chimie des Polymères Organiques, UMR 5629

Ecole Nationale Supérieure de Matériaux, d'Agroalimentaire et de Chimie, ENSMAC (ex ENSCBP)

16 Avenue Pey-Berland

33607 Pessac Cedex

France