Toward Synthetic Intrinsically Disordered Polypeptides (IDPs): Controlled Incorporation of Glycine in the Ring-Opening Polymerization of N -Carboxyanhydrides

Mostafa Badreldin, Pedro Salas-Ambrosio, Sylvain Bourasseau, Sebastien Lecommandoux, Simon Harrisson, Colin Bonduelle

Towards Synthetic Intrinsically Disordered Polypeptides (IDPs): Controlled Incorporation of Glycine in Ring-Opening Polymerization of *N*-carboxyanhydrides.

Mostafa Badreldin,[a] Pedro Salas-Ambrosio,[a,b] Sylvain Bourasseau,[a] Sebastien Lecommandoux,[a] Simon Harrisson,[a] Colin Bonduelle[a]**

[a] Université Bordeaux, CNRS, Bordeaux INP, LCPO, UMR 5629, F-33600 Pessac, France.

[b] Department of Chemistry and Biochemistry, University of California Los Angeles. 607 Charles E. Young Drive East, Los Angeles, CA 90095-1569, United States

ABSTRACT: Intrinsically disordered proteins (IDPs) do not have a well-defined folded structure, but instead behave as extended polymer chains in solution. Many IDPs are rich in glycine residues, which create steric barriers to secondary structuring and protein folding. Inspired by this feature, we have studied how the introduction of glycine residues influences the secondary structure of a model polypeptide, poly(*L*-glutamic acid), a helical polymer. For this purpose, we carried out ring opening copolymerization with γ-benzyl-*L*-glutamate and glycine N-carboxyanhydride (NCA) monomers. We aimed to control the glycine distribution within PBLG by adjusting the reactivity ratios of the two NCAs using different reaction conditions (temperature, solvent). The relationship between those conditions, the monomer distributions and the secondary structure enabled the design of intrinsically disordered polypeptides when a highly gradient microstructure was achieved in DMSO.

KEYWORDS: Polypeptides polymers, N-carboxyanhydride, copolymerization, reactivity ratios, intrinsically disordered

INTRODUCTION:

Mimicking the structural and functional abilities of proteins, nature's most versatile biomacromolecules, is a long-standing goal of polymer chemists. The extreme structural complexity of most proteins, however, is a near-insurmountable obstacle. Intrinsically disordered proteins (IDPs) lack higher order structure and behave as free polymer chains in solution.[1] Nonetheless, they play major structural and functional roles in living organisms,[2] as well as in some notable human diseases.[3] IDPs mainly consist of random coil segments whose amino acids do not have the fixed torsion angles characteristic of common secondary structures.[4,5] They are characterized by low complexity sequences[2] of repeating patterns of amino acids that promote extended coil conformations and disfavor tertiary folding. These features lead to highly soluble and non-aggregated proteins. The

prevalence of polar amino acids promotes the solvation of the extended chains while charged residues generate repulsion between chain segments.[6] Proline and glycine, commonly found in IDPs, create steric barriers to the appearance of secondary structure.[7]

Peptidic polymers, in which amino acids are randomly distributed, are much simpler macromolecules than proteins.[8] They combine the advantageous features of synthetic polymers (solubility, process, elasticity, etc.) with many of those of natural proteins (secondary structure, functionality, biocompatibility, etc.). These synthetic polypeptides are efficiently and economically prepared by ring-opening polymerization (ROP) of amino acid N-carboxyanhydrides (NCA) and the monomer distribution, like that of proteins, strongly influences their functional properties.[9,10] However, relatively few studies have focused on studying the copolymerization kinetics of NCAs to control the primary sequence of synthetic polypeptides and its effect on the secondary structuring.[11,12] The ability to control monomer sequence in synthetic macromolecules is a powerful tool for tailoring copolymer properties and functions.[13] Techniques for sequence control have been historically developed in other fields before being applied to synthetic polymer chemistry, including iterative synthesis (biochemistry, chemical biology) and template polymerization (biology).[14] In classical polymerization processes (step- and chain-growth), macromolecules with statistical distributions of molecular weight, structure, and/or composition are produced.[15] These distributions are determined by the reaction conditions and reactivities of the monomers involved,[16] and can be modified by controlling the reaction conditions.[17] The result is a degree of control over the monomer sequence, albeit imperfect and statistical in nature.

In polymer chemistry, the probability of obtaining the desired monomer distribution can be maximized by choosing appropriate monomer feed ratios. In this direction, tandem repeat proteins have already inspired studies of terpolymerization systems based on ROP of NCA monomers. For example, the proportion of Leu-Asp-Val sequences can be maximized by adapting the feed ratio of leucine, aspartate and valine NCA monomers.[10] Similarly, Arg-Gly-Asp sequences can be obtained from the terpolymerization of arginine, glycine and aspartate NCAs.[9] In both examples, controlling the monomer distribution led to an increase in the desired cell adhesion property. With NCAs, the monomer distributions are also sensitive to the reaction conditions, and to the secondary structure of the polypeptide.[18] For instance the solvent significantly affects secondary structuring in copolymerizations of γ-benzyl-L-glutamate NCA (BLG-NCA) and valine NCA,[19] which in turn affects the reactivity of the different comonomers.[20] Nevertheless, studies on the relationship between the monomer distribution and the secondary structure are scarce. The initiator type, the solvent and temperature affected the copolymerization kinetics and, to an extent, the secondary structure of a binary system comprising of valine NCA and leucine NCA.[21] Similarly, when BLG-NCA was copolymerized with valine NCA the resulting gradient copolymers showed β-sheet-like structures consistent with short BLG sequences and valine sequences of variable length.[22] It was also shown that adding glycine units in a random copolymerization system with BLG, produces similar secondary structure trends.[23]

In this work, we study the kinetics, the reactivity ratios, the monomer distribution and the secondary structure of various polypeptides obtained through copolymerization of BLG-NCA and glycine-NCA, the main disorder inducing amino-acid in IDPs. We found that the helical conformation can be disrupted by controlling the reactivity and the addition of glycine, producing an unprecedented homochiral poly(γ-benzyl-L-glutamate) with a coil disorder state, an intrinsically disordered polypeptide. This was optimal when the synthesis was conducted in DMSO, via the formation of highly gradient copolymer microstructures.

EXPERIMENTAL SECTION:

Materials and Methods. All chemicals and solvents were purchased from Sigma Aldrich, Fluorochem, Acros, TCI, Strem and VWR, and were used without further purification unless otherwise described. Dimethylformamide (DMF), tetrahydrofuran (THF), dichloromethane (DCM), ethyl acetate and cyclohexane were obtained from a solvent system purifier (PureSolv, Innovative Technology), kept under argon atmosphere and freshly used. Diethylether (Et_2O) was dried over 3Å molecular sieves before use in NCA purification. Milli-Q water was obtained from a Purelab Prima purification system (ELGA) with a resistivity of 18.2 MΩ cm^{-1}. BLG N-carboxyanhydride was purchased from PMC Isochem, stored at -20 °C under argon atmosphere and weighed in a glove box (Jacomex GP13 no. 2675). Dialysis membranes were purchased from SpectrumLabs. Hexylamine was 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 3Å molecular sieves for at least one week and cryodistilled before use. DMSO-d was stored on activated 3Å molecular sieves for at least one week before use.

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 recorded on a 500 MHz Prodigy NMR AVANCE NEO 500. Solvent peaks were used as reference for the locking. The spectra obtained were calibrated using residual solvent signals. Signals were categorized as follows: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m) and broad (br).

Procedure for in-situ ^{1}H NMR monitoring

Measurements were carried out in Young NMR tubes using dried deuterated solvents. Samples were prepared inside the glovebox and the tubes sealed before inserting them into the spectrometer at the specified temperature. The tubes were spun inside the magnet to induce stirring in the samples. The magnet was auto-shimmed and the first spectrum collected served as a reference. The experiments were conducted directly inside the spectrometer with controlled temperature depending on the system (25°C for systems **X2** and **X6**, 0°C for systems **X1** and **X3** and 50°C for system **X5**). Subsequent spectra were collected at different time periods until high conversion was obtained. The NMR data were analyzed using Topspin software with automatized baseline and phase correction and integration. Manual adjustments were made as needed and solvent peaks were used as

references and internal standards. In each case, the disappearing peaks of the different NCAs were integrated in comparison to the unchanged peaks of the solvent peaks of the solvent.

Size exclusion chromatography (SEC)

1. DMF + LiBr 1g/L or DMSO as eluent. Measurements 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. Column temperature was held at 50°C. M_n values were calculated from the MALS signal using dn/dC of 0.1197 calculated using homo-PBLG samples. Pullulan from Agilent was used as standards for calibration curve calculations.

2. 1,1,1,3,3,3-Hexafluoroisopropanol (HFIP) + 0,05% KTFA as eluent. Measurements were performed on an Ultimate 3000 system from Thermoscientific equipped with UV detector and dRI detector 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. Column temperature was held at 40°C. An Easivial kit of poly(methyl methacrylate) (PMMA) from Agilent was used as standards for calibration curve calculations.

3. aqueous buffer as the eluent. Measurements were performed on an Ultimate 3000 system from Thermoscientific equipped with diode array detector DAD, MALS and dRI detectors 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 as the eluent. Column temperatures was held at 25°C. M_n values were calculated using dn/dC value of 0.1489 calculated using homo-PGA samples. Easivial kit of polyethylene glycol (PEG) from Agilent was used as standards for calibration curve calculations.

Fourier-transform infrared (FTIR) spectrometery

Infrared (IR) spectra were recorded using an FTIR spectrometer (Vertex 70, Bruker), and an attenuated total reflectance (ATR) (GladiATR, Pike Technologies) apparatus from Fisher technologies performing 32 scans. The raw data were analyzed with Opus7.5 software. The analyzed compounds were dissolved in anhydrous solvent 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 the corresponding solvent 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). Solid compounds were crushed on the ATR before recording the spectra. Data analysis was carried out using Microsoft Excel.

Circular Dichroism (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: 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 2.5 mM. For pH=4 sodium acetate/acetic acid was used, while for pH=6.5 sodium acetate was used. Unprotected polymers were analyzed as 0.57mM solutions in HFIP.

Synthesis of glycine N-carboxyanhydride (Gly NCA)

Gly NCA was synthesized using Leuch's method as previously described:[24]

In a clean and dried Schlenk flask *N*-Boc-glycine (2 g, 11.4 mmol, 1 equiv.) was dried for 15 min in the Schlenk line, then 32 mL of anhydrous THF was added, followed by thionyl chloride (1 mL, 13.7 mmol, 1.2 equiv.). The reaction was stirred for 4 h at RT and its color turned from yellow to dark brown. The solvent was evaporated under vacuum, and the residue was 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 dry celite. The system was washed with 60 mL of anhydrous ethyl acetate and the solution concentrated. The solution was filtered again over dry celite and the system washed with 60 ml of anhydrous ethyl acetate. The solid was washed twice with 10 mL of dry diethyl ether before drying. This last step removed any remaining Boc-containing side products. 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 $C_3H_2NO_3$.

Synthesis of poly(γ-benzyl-L-glutamate) (PBLG)

In a typical ROP of BLG NCA, 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 to be the starting time of the reaction. The solution was stirred for 10 days at 4°C under argon. The reaction completion was monitored by following the depletion of the bands at 1859 cm^{-1} and 1787 cm^{-1} characteristic of the C=O stretching of NCAs. The polymer was precipitated twice from diethyl ether and dried under high vacuum. Yield: 1690 mg; 67%, white powder. ^{1}H-NMR (10 mg/ml, $CDCl_3$ containing 15% trifluoroacetic acid, 400 MHz, 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). ^{13}C-NMR, (125 MHz, δ,

ppm): 14.7 (1C, CH3- hexylamine), 21.0 (1C, CH3-CH2- hexylamine), 26.5 (1C, -CH2-CH2-CH2- hexylamine), 28.5 (6C, -CH2-CH2-CH2- hexylamine), 31.0 (1C, -CH2-CH2-NH- hexylamine), 40.9 (1C, -CH2-CH2-NH- hexylamine), 26.5(1C, CH-CH2- BLG), 30.0 (2C, -CH2-CH2- BLG), 53.2 (1C, -CH- BLG), 68.2 (1C, -COO-CH2-, BLG), 128-129 (m, 6C, Ar), 173 (1C, -CH-COO- BLG), 175 (1C, -CH2-COO- BLG). SEC analysis using dn/dC calucations gave an Mn=11.1 kg/mol (DP=50) and Đ=1.03 (**Figure S6**).

Synthesis of copolymers (HA-P(Gly_x-BLG_y))

In a typical copolymerization of Gly NCA and BLG NCA: Gly NCA (xeq) and BLG NCA (yeq) were weighed in a flame-dried Schlenk in a glovebox under argon. The powders were then dissolved in anhydrous solvent at a concentration of 0.1M. Hexylamine (x+y/50eq) was added and this was considered starting time of the reaction. The solution was stirred until all NCAs were consumed as determined by FTIR. The polymer was precipitated twice from diethyl ether and dried under high vacuum. ^{1}H-NMR (10 mg/ml TFA-d, 400MHz, δ, ppm): 0.90 (t, 3H, CH3 hexylamine), 1.31 (br, 6H, CH3-(CH2)3- hexylamine), 1.54 (m, 2H, -CH2-CH2-NH- hexylamine), 1.87-2.31 (m, 2yH, CH2 BLG), 2.40-2.70 (m, 2yH, CH2 BLG), 3.30 (m,2H, -CH2-NH- hexylamine), 4.13 (s,br, 2xH, CH2 Gly), 4.63 (m, 1yH, CH BLG), 5.17 (m, 2yH, CH2O BLG), 7.34 (m, 5yH, ArH BLG).

Synthesis of polyglutamate (PGA)/ deprotection from PBLG

HA-$PBLG_{50}$ (100 mg, 0.45 mmol of BLG units, 50eq) was dissolved in TFA (7.7 mL) then HBr 48% aqueous solution (0.18 ml, 1.6 mmol, 175eq) were added under vigorous stirring. The reaction mixture was stirred at room temperature overnight. Solution was precipitated twice from cold Et_2O and centrifuged. The obtained white powder was precipitated in 5ml MilliQ water then 10 ml of a saturated sodium bicarbonate solution was added so that a pH of 8 was obtained. The resulting mixture was dialyzed against milliQ water (MWCO 3.5 kDa) and lyophilized to give the deprotected form; PGA as a white solid (Yield: 72%, 50 mg). Sample was then analyzed by ^{1}H NMR with a small amount of NaOD at 10 mg/ml in D_2O. ^{1}H-NMR (400MHz, δ, ppm): 0.75 (t, 3H, CH3 hexylamine), 1.16 (br, 6H, CH3-(CH2)3- hexylamine), 1.39 (m, 2H, -CH2-CH2-NH- hexylamine), 1.78-2.07 (m, 2H, CH2 BLG), 2.21 (m, 2H, CH2 BLG), 3.10 (t ,2H, -CH2-NH- hexylamine), 4.63 (m, 1H, CH BLG), 4.23 (m, 2H, -CH2O- BLG). ^{13}C-NMR, (500MHz, δ, ppm): 13.4 (1C, CH3- hexylamine), 22.0 (1C, CH3-CH2- hexylamine), 25.7 (1C, -CH2-CH2-CH2- hexylamine), 28.1 (1C, -CH2-CH2-CH2- hexylamine), 30.7 (1C, -CH2-CH2-NH- hexylamine), 39.5 (1C, -CH2-CH2-NH- hexylamine), 25.5 (2C, CH-CH2- BLG), 33.0 (2C, -CH2-CH2- BLG), 42.6 (2C, CH2 Gly), 53.8 (1C, -CH- BLG), 174 (1C, -CH-COO- BLG), 181 (1C, -CH2-COO- BLG). SEC analysis using dn/dC calucations gave an Mn=7.9 kg/mol (DP=52) and Đ=1.03 (**Figure S8**).

Synthesis of poly(glycine-co-glutamate) (HA-P(Gly_x-co-GA_y)/ deprotection from HA-P(Gly_x-BLG_y)

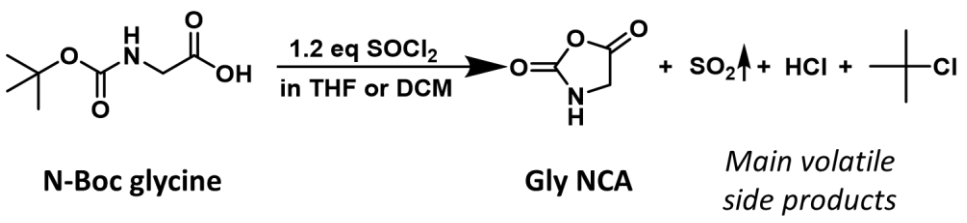
HA-P(Gly_x-co-$PBLG_y$) (yeq) was dissolved in TFA at 13mg/ml then HBr 48% aqueous solution (yeq) were added under vigorous stirring. The reaction mixture was stirred at room temperature overnight. Solution was precipitated twice in cold Et_2O and centrifuged. The obtained white powder was precipitated in 5ml MilliQ water then 10 ml of a saturated

sodium bicarbonate solution was added so that a pH of 8 was obtained. The resulting mixture was dialyzed against milliQ water (MWCO 3.5 kDa) and lyophilized to give the deprotected form HA-P(Gly_x-co-GA_y) as a white solid. Sample was then analyzed by ^{1}H NMR at 10 mg/ml in D_2O with a small amount of NaOD. ^{1}H-NMR (400MHz, δ, ppm): 0.75 (t, 3H, CH3 hexylamine), 1.16 (br, 6H, CH3-(CH2)3- hexylamine), 1.39 (m, 2H, -CH2-CH2-NH- hexylamine), 1.78-2.07 (m, 2yH, CH2 BLG), 2.21 (m, 2yH, CH2 BLG), 3.10 (t ,2H, -CH2-NH- hexylamine), 3.91 (s, 2xH, CH2 Gly), 4.63 (m, 1yH, CH BLG), 4.23 (m, 2yH, CH2O BLG). ^{13}C-NMR, (500MHz, δ, ppm): 13.4 (1C, CH3- hexylamine), 22.0 (1C, CH3-CH2- hexylamine), 25.7 (1C, -CH2-CH2-CH2- hexylamine), 28.1 (1C, -CH2-CH2-CH2- hexylamine), 30.7 (1C, -CH2-CH2-NH- hexylamine), 39.5 (1C, -CH2-CH2-NH- hexylamine), 25.5 (2yC, CH-CH2- BLG), 33.0 (2yC, -CH2-CH2- BLG), 42.6 (2xC, CH2 Gly), 53.8 (1yC, -CH- BLG), 170-172 (1xC, -COO- Gly), 173-175 (1yC, -CH-COO- BLG), 180-182 (1yC, -CH2-COO- BLG).

RESULTS AND DISCUSSION

Glutamate is the highest produced amino acid industrially[25] due to its use in the food sector.[26] Its derivatives, including γ-benzyl-L-glutamate NCA (BLG-NCA), are readily available. BLG-NCA affords poly(γ-benzyl-L-glutamate) (PBLG) which is a helical polymer with biomimetic secondary structuring properties,[27] while its deprotected form, polyglutamic acid (PGA), is water soluble and pH-responsive.[28] Polymers[18,26,29,30] and copolymers[31,32] of PBLG or PGA derivatives have potential applications in bioactive substance delivery,[33–35] and the preparation of antibacterial agents[36] or biomaterials[37].

The objective of our study is to synthesize intrinsically disordered PBLG polymers, meaning that they do not adopt helical conformations. For this purpose, glycine units were introduced into the glutamate backbones *via* copolymerization of BLG NCA and Gly NCA. The less sterically hindered monomer, Gly NCA, is more reactive in ROP than BLG-NCA but the difference in reactivity has not been quantified.[24] We synthesized Gly-NCA using Leuch's method using N-Boc glycine and $SOCl_2$ in THF followed by celite filtration (**Scheme 1**) which proved to be convenient for scale up.



Scheme 1. 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.

Homopolymerization of Gly-NCA is notoriously difficult as the reaction mixture becomes heterogeneous, slowing down the reaction speed and risking the degradation of unreacted monomers.[38–40] Nonetheless, the copolymerization of Gly-NCA with other types of NCAs is possible,[41–43]while being highly more reactive in most cases[12]. The copolymerization of BLG and Gly NCAs was followed using the *in-situ* NMR technique which allows to monitor the consumption of each monomer independently and the corresponding reactivity ratio. The absence of a side chain on glycine makes it easier to

distinguish from other amino acids by NMR,[44] simplifying the characterization of the copolymers.

The monomer distribution can be estimated from the reactivity ratios of the comonomers. In the terminal model of copolymerization,[45] it is assumed that only the terminal group of the polymer chain affects the reaction kinetics, and these kinetics are assumed to be independent of chain length. In a copolymerization between monomer A and monomer B, the reactivity ratio of monomer A ($\mathbf{r_A}$) is defined as the ratio between the propagation constants $\mathbf{k_{AA}}$ and $\mathbf{k_{AB}}$, corresponding to the reaction of growing polymer chain with A as the terminal monomer unit with monomers A (homopropagation) or B (cross-propagation), respectively. The Meyer-Lowry equation (**Equation 1**) can be used to calculate the reactivity ratios by following the changing monomer composition of the reaction mixture as the polymerization proceeds.[46] This method allows reactivity ratios to be determined at moderate to high conversions, the region where high molecular weight polymer is produced in a controlled polymerization. Frequently used low conversion methods of reactivity ratio determination are inappropriate for these polymerizations as the assumption of chain-length independent kinetics is unlikely to be valid when only oligomers are present.[47] All copolymerizations were therefore monitored to high conversions to also consider the effect that secondary structure can have on the copolymerization kinetics and possible deviation from the model.

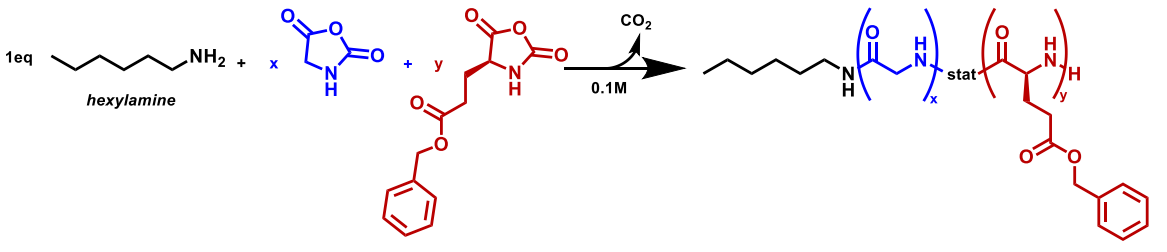
$$x = 1 - \left(\frac{f_1}{f_{1,0}}\right)^{\alpha}\left(\frac{f_2}{f_{2,0}}\right)^{\beta}\left(\frac{f_{1,0}-\delta}{f_1-\delta}\right)^{\gamma}$$

$$\alpha = \frac{r_2}{1-r_2} \qquad \beta = \frac{r_1}{1-r_1}$$

$$\gamma = \frac{1-r_1r_2}{(1-r_1)(1-r_1)} \qquad \delta = \frac{1-r_1r_2}{2-r_1-r_2}$$

Equation 1. Integrated copolymerization equation and the values of the constants. x is reaction conversion, f_i is instantaneous mole fraction of monomer i, $f_{i,0}$ is initial mole fraction of monomer i.[48]

The copolymerization systems comprised BLG and Gly-NCA at a total concentration of 0.1 M. Hexylamine (0.02 eq, targeted DP of 50) was used to initiate polymerizations (**Scheme 2**).



Scheme 2. Copolymerizations' general scheme. x+y is kept at 50.

Gly NCA and BLG NCA were first copolymerized in DMF at 0°C. In total, five copolymerizations were performed with initial BLG-NCA fractions ($\boldsymbol{f_{B,0}}$) of 0.90, 0.80, 0.50, 0.36 and 0.20 (**Figure 1.A**.). The experimental data was fitted to equation 1 to obtain the reactivity ratios. Since both conversion and BLG fraction are experimentally determined it is necessary to consider the induced error in each. For this purpose, a previously developed[48] nonlinear least squares method was used assuming non-negligible error in all variables.[49] Using the visualization of the sum of squares space method, point

estimates of r_{Gly} and r_{BLG} of 1.75 and 0.47 respectively were obtained (**Figure 1.B.**). This result indicated that the copolymer adopts a gradient like microstructure with a higher concentration of glycine units at one end of the chain. Reasonable agreement between experimental results and the model was obtained, although some deviations can sometimes be observed at higher conversion (**Figure 1.A.**).

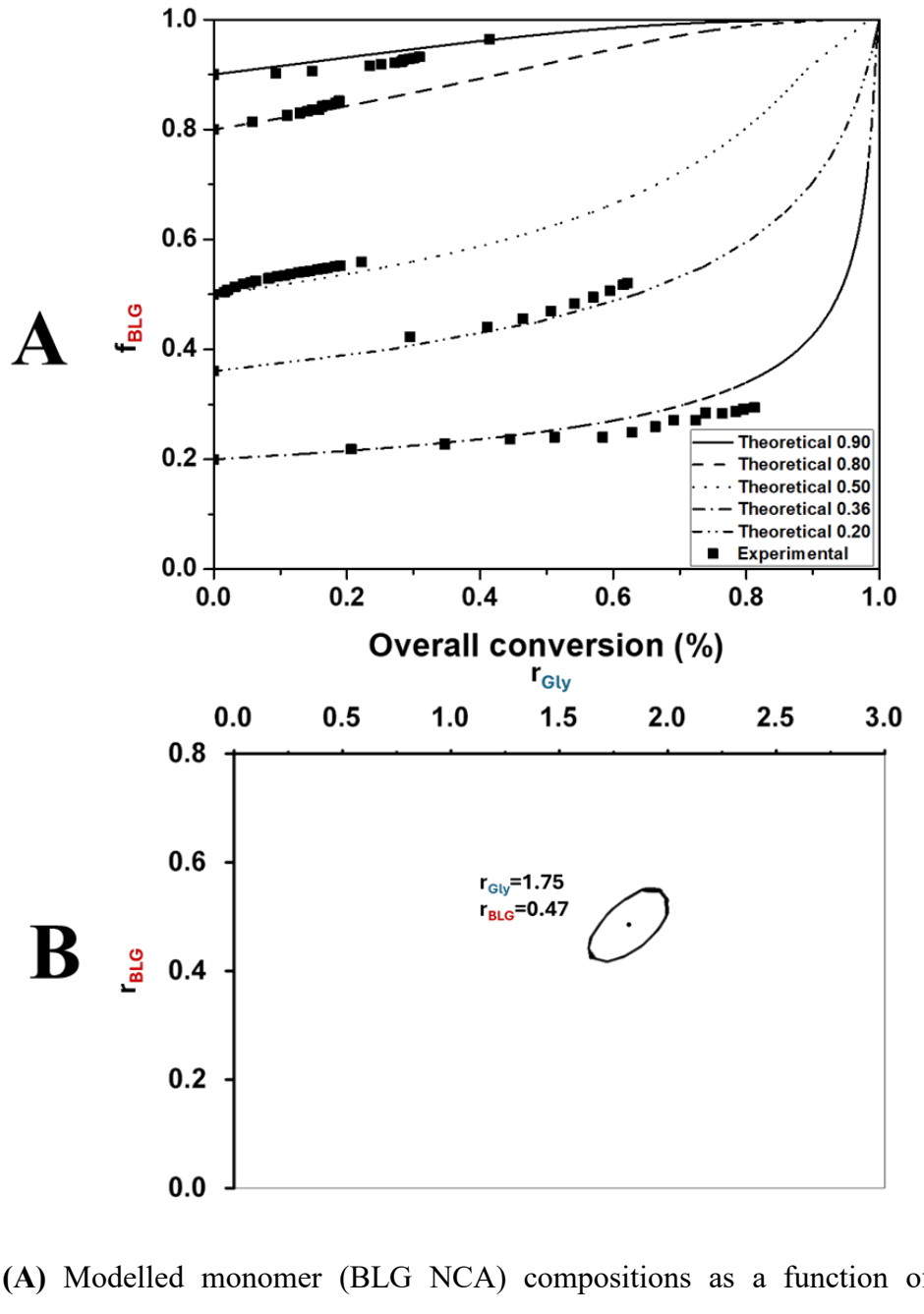


Figure 1. (A) Modelled monomer (BLG NCA) compositions as a function of the copolymerization conversion (lines) in the case of the five experiments listed above with varying initial monomer feeds (copolymerization system **X1**). These plots are compared with the experimental data from all the experiments (points). **(B)** Point estimate and 95% joint confidence interval (contour) of reactivity ratios of copolymerization system **X1**. r_{Gly} = 1.75 and r_{BLG} = 0.47.

The gradient copolymer obtained from this system was schematically depicted 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 2**, 20 chains are presented, with an average number of units equal to the number-average degree of polymerization (50) of the polymer and a length distribution with dispersity of 1.1. 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.[48,50]

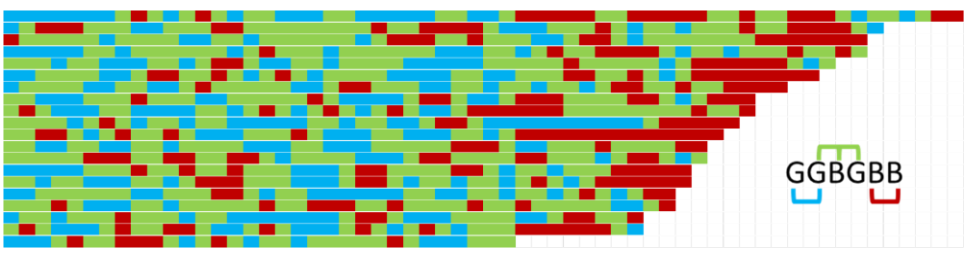


Figure 2. Simulated 20 polymer sequences of copolymerization system **X1** with $f_{G,0} = f_{B,0}$ = 0.5, DP_n of 50, dispersity of 1.1 and a conversion of 99%. The direction of polymerization is from left to right. Gly-Gly dyads are shown in blue (28%), Gly-BLG and BLG-Gly dyads in green (51%) and BLG-BLG dyads in red (28%).

Most radical copolymerizations are relatively insensitive to reaction conditions, as rate constants for each comonomer vary similarly to changes in solvent polarity or temperature. Nonetheless, in some cases it is possible to modify copolymer microstructure by varying reaction conditions like solvent polarity[51] and temperature[52]. In general, much less data exists for non-radical copolymerizations, which tend to exhibit greater monomer selectivity than their radical counterparts. These copolymerizations, which involve higher activation energies and polarized transition states, are more sensitive to changes in temperature and solvent polarity. Indeed, the solvent has previously been shown to affect NCA copolymerization kinetics.[19,20] To study this effect, copolymerizations were carried out in DMF (copolymerization system **X2**), THF (copolymerization system **X3**) and DMSO (copolymerization system **X4**). As DMSO solidifies at 19°C, copolymerizations were carried out at 25°C. All other parameters were kept constant (0.1M concentration of monomers, monomer:hexylamine = 50). We found that the solvent strongly influenced the reactivity ratios of Gly and BLG NCA (**Figure 3.A.**). Simulations of typical chains for copolymers synthesized at an initial glycine fraction of 50% show clear differences in structure (**Figure 3.B.**).

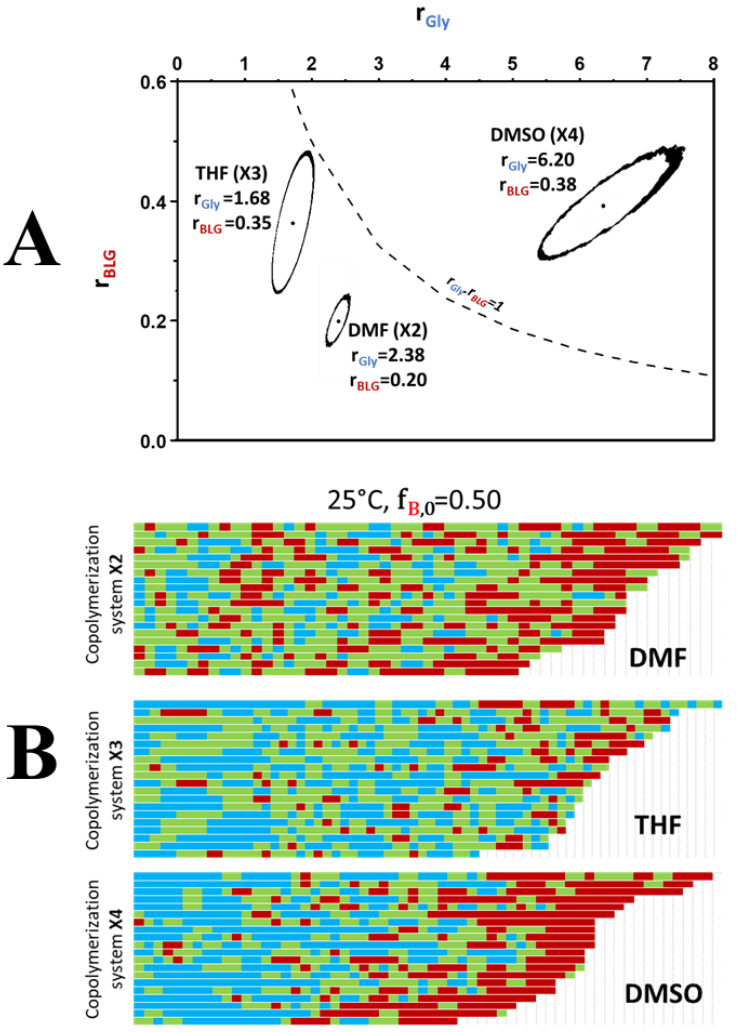


Figure 3. (A) Point estimate and 95% joint confidence interval (contour) of reactivity ratios of copolymerization systems **X2-X4**. The values of system r_{Gly} and r_{BLG} from each solvent system are shown. The dotted line corresponds to the case of an ideal copolymerization where the difference in reactivity between the comonomers is independent of the propagating chain end.[53] (B) Simulated 20 polymer sequences of copolymer systems **X2-X4** with $f_{G,0} = f_{B,0}$ = 0.5, DP_n of 50, dispersity of 1.1 and a conversion of 99%. 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.

Copolymers in both DMF and DMSO show a strong composition gradient, but the DMF copolymer is characterized by short sequences of homo- and alternating dyads, while the homopolymer sequences in the DMSO copolymer are much longer. By contrast, the composition gradient of the THF polymer is much weaker and the polymer has a nearly random structure. This is borne out by the product of the reactivity ratios, $r_{Gly}r_{BLG}$. The lowest $r_{Gly}r_{BLG}$ was obtained in THF (0.59) while in DMF $r_{Gly}r_{BLG}$ was 0.82. These values indicate a tendency to form alternating polymer. In the case of DMSO $r_{Gly}r_{BLG}$ was equal to 2.36 which indicates a tendency to form blocks of homopolymer. We attributed the difference in results between the solvents to the different solubility of the polymer chains, 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.[30] In contrast, formation of secondary structure could be observed in THF, a helicogenic solvent.

At equal initial monomer feeds ($f_{G,0} = f_{B,0} = 0.5$) for Gly NCA and BLG NCA, the copolymer has a gradient nature with a higher concentration of Gly-Gly dyads at the beginning of the chain and of BLG-BLG dyads at the other end with an overall high concentration of alternating dyads. In copolymerization system **X1** (0°C in DMF), the reactivity ratios were calculated to be r_{Gly}=1.75 and r_{BLG}=0.47 (**Figure 1**) while in the same solvent at 25°C (copolymerization system **X2**) the values were 2.38 and 0.2 respectively (**Figure 3**). The difference in reactivity ratios obtained for these two copolymerization systems showed that the temperature can also influence the kinetics. Similar studies were done at 50°C in DMF (copolymerization system **X5**) and at 0°C in THF (copolymerization system **X6**) (**Figure 4**).

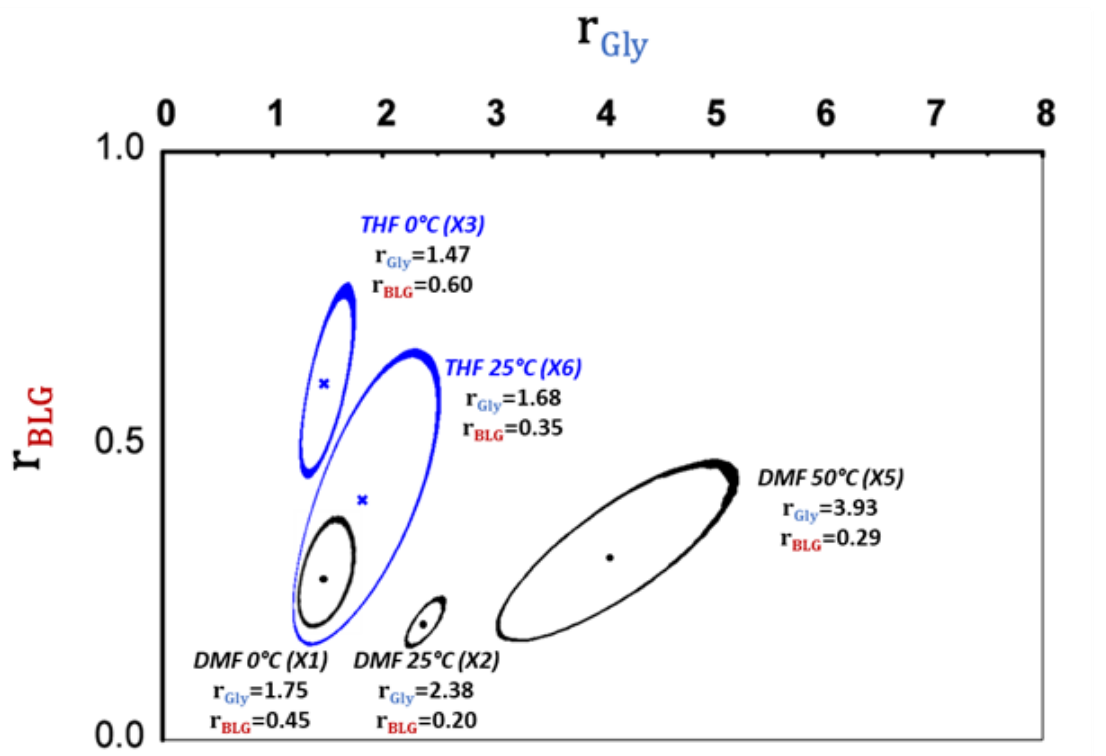
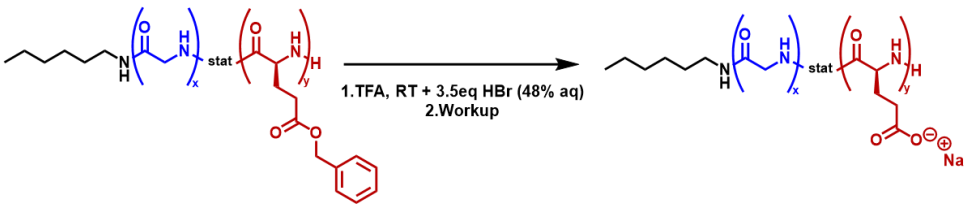


Figure 4: Point estimate (black circles: DMF; blue crosses: THF) and 95% joint confidence interval (contour in black: DMF; blue: THF) of reactivity ratios of copolymerization in DMF and THF at different temperatures.

As shown in **Figure 4**, 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. The increased difference between the two reactivity ratios led to a steeper gradient. Increasing the reaction temperature also increases the overall reactivity of the copolymerization. However, chain end capping due to backbiting of the BLG end-units limits the maximum useful polymerization temperatures.

A PBLG homopolymer (polymer **1**) and a series of copolymers (**2-7**) of Gly NCA and BLG NCA were then synthesized in DMF at 0°C to examine the effect of increasing glycine content. Unlike pure PBLG, the copolymers became insoluble in DMF at different conversion depending on the glycine content (**esi**). Copolymers **2-7** were synthesized with $f_{B,0}$ value spanning a range from 0 to 0.5 (**Table 1**). Copolymers **8-10** were synthesized to establish the effect of the reaction solvent on their macromolecular properties (**Table 1**). Only copolymer **10** was soluble throughout polymerization in DMSO, while copolymers **8** and **9** precipitated at 50% and 40% conversion respectively (**esi**). No single solvent system could solubilize all copolymers, as their solubility depended strongly on the glycine content and the reaction solvent. As a result, characterization by SEC was performed using DMF, DMSO or HFIP as eluent, depending on the copolymer (esi). In order to compare the copolymers, deprotection to water-soluble polyglutamic acid was performed.

Acidic or mildly basic conditions are preferably employed for deprotection in order to avoid epimerization.[28] In this study, deprotection was conducted in TFA using HBr followed by neutralization and dialysis (**Scheme 4**).



Scheme 3. Deprotection of the synthesized copolymers P(Gly-stat-BLG) to obtain their unprotected counterparts (**D**) P(Gly-stat-GA).

The basic conditions used in workup result in the deprotonation of the carboxylic acid side chains,[54] ensuring solubility. Using PEG calibration and RI detection polymer **1D**'s Mn was calculated to be 5.4 kg/mol with a dispersity of 1.08. All synthesized copolymers were recovered in acceptable yields, had expected structures as verified by ^{1}H NMR and eluted at similar volumes in SEC analysis (**Table 1 and esi**).

Table 1. SEC analysis of all synthesized copolymers (**1-10**) and their deprotected, water-soluble form (**1D-10D**) compared to polymer **1**(**D**) where D stands for deprotected.

	$f_{B,0}$ [a]	Reaction solvent	Reaction temperature	Yield (%)[b]	Structure (theoretical)	$f_{B,0}$ by ^{1}H NMR[c]	Overall DP by ^{1}H NMR[d]	Mn (kg/mol) [e]	Đ[f]
Polymer **1D**	1.00	DMF	0°C	86	HA-PGA$_{50}$	1	55	5.4	1.08
Copolymer **2D**	0.95	DMF	0°C	64	HA-PGA$_{47}$-co-PGly$_3$	0.94	57	7.6	1.13
Copolymer **3D**	0.90	DMF	0°C	60	HA-PGA$_{45}$-co-PGly$_5$	0.85	60	7	1.18
Copolymer **4D**	0.80	DMF	0°C	70	HA-PGA$_{40}$-co-PGly$_{10}$	0.80	64	3.7	1.17
Copolymer **5D**	0.70	DMF	0°C	94	HA-PGA$_{35}$-co-PGly$_{15}$	0.65	63	14.2	1.37
Copolymer **6D**	0.60	DMF	0°C	85	HA-PGA$_{30}$-co-PGly$_{20}$	0.60	50	8.4	1.19
Copolymer **7D**	0.50	DMF	0°C	78	HA-PGA$_{25}$-co-PGly$_{25}$	0.50	51	7.7	1.26
Copolymer **8D**	0.50	DMF	25°C	82	HA-PGA$_{25}$-co-PGly$_{25}$	0.50	50	5.1	1.21
Copolymer **9D**	0.50	THF	25°C	90	HA-PGA$_{25}$-co-PGly$_{25}$	0.48	48	6.8	1.12
Copolymer **10D**	0.50	DMSO	25°C	80	HA-PGA$_{25}$-co-PGly$_{25}$	0.50	50	5.3	1.10

[a]initial mole fraction of BLG in monomer feed; [b]yields after deprotection; [c]as determined by integration of glycine and glutamate peaks in ^{1}H NMR in D_2O; [d]as determined by the ratio between the integration of the HA peaks and monomer peaks in ^{1}H NMR in D_2O; [e] as determined by PEG calibration in aqueous SEC using RI detection; [f] as determined by SEC peaks using RI detection.

Copolymers **8-10** were further characterized by quantitative ^{13}C NMR in TFA-d to attribute homo-sequences and hetero-sequences in the carbonyl peak area (**Figure 5**) and compare it to results obtained through theoretical modeling (**Figure 3.B.**). The homo-sequences peaks were identified by analyzing homo-PBLG (polymer **1**) and oligomers of glycine (**esi**).

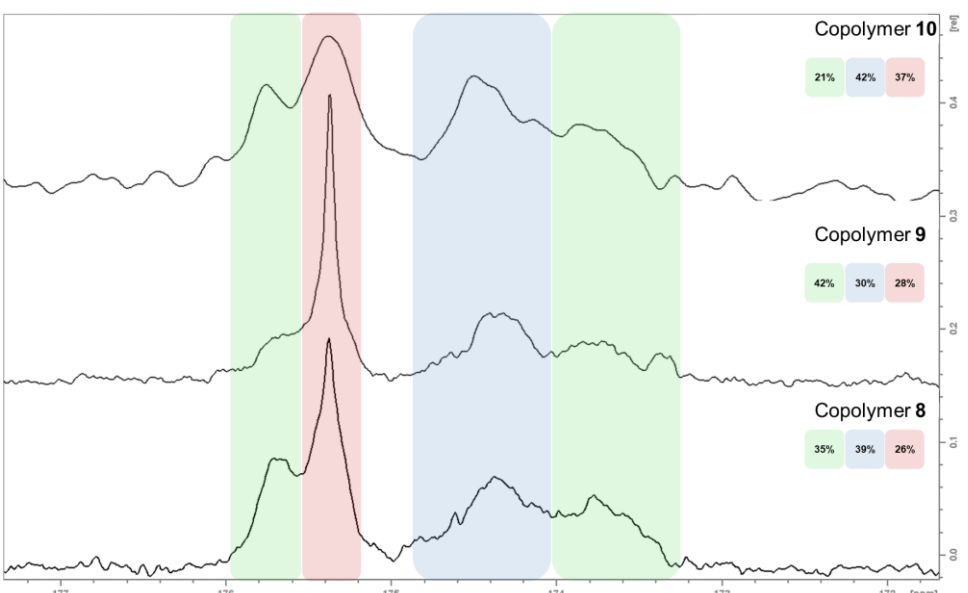


Figure 5. Quantitative ^{13}C NMR of the carbonyl region (170-180 ppm) of copolymers **8-10**. Gly-Gly carbonyl peaks are highlighted in blue, Gly-BLG and BLG-Gly carbonyl peaks in green and BLG-BLG carbonyl peaks in red. The percentages of the different peak regions were determined by peak fitting and deconvolution.

BLG repeat units' carbonyl peaks are deshielded when the previous unit is glycine (blue peaks at 175.4 ppm). 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 deconvolution of the different peaks (esi), it was possible to calculate the percentage of homo-dyads and hetero-dyads (**Figure 5**/**Table 2**). Although the values from ^{13}C NMR and Monte Carlo simulations were not the same, the tendencies were similar. ^{13}C NMR analysis confirmed that copolymer **10**, with steepest composition gradient, had the lowest alternating dyad content and the highest content of homosequences of both monomers. For all copolymers 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 difference in microstructure of copolymers **8-10** confirms the influence of solvent on the reactivity ratios of Gly NCA and BLG NCA. Overall, our characterizations showed that by changing the reaction solvent, the glycine monomer distribution can be controlled to obtain higher alternation.

Table 2: Dyad percentages determined by Monte Carlo simulation and ^{13}C NMR for copolymers **8-10**

		Monte Carlo simulation			^{13}C NMR		
	Solvent	Gly-BLG% BLG-Gly%	BLG-BLG%	Gly-Gly%	Gly-BLG% BLG-Gly%	BLG-BLG%	Gly-Gly%
Copolymer **8**	DMF	51	19	29	35	26	39
Copolymer **9**	THF	54	20	26	42	28	30
Copolymer **10**	DMSO	31	28	41	21	37	42

The goal of our study is to produce glutamate-based copolymers that are intrinsically disordered by breaking their secondary structures using the incorporation of glycine units. PBLG exhibits secondary structuring that depends on its molecular weight and on the solvent. While oligomers of BLG can adopt beta-sheet like structures, above a DP of 10-15, they adopt an α-helical conformation.[55] This conformation reveals itself in helicogenic solvents, such as tetrahydrofuran (THF) or chloroform, and is usually insoluble.[56,57] Circular dichroism spectroscopy (CD) is widely used in structural biology to evaluate the conformation of proteins.[58] The CD spectra of most secondary structures, such as α-helices, β-sheets and random coil, have well documented shapes and magnitudes.[59] This analysis is preferably done in aqueous conditions where proteins can be observed in biological conditions. In its protonated form, the carboxylic acid moieties of PGA side chains form intramolecular hydrogen bonds, stabilizing the α-helical form. In its deprotonated form, electrostatic repulsion between negative charges induces a random coil conformation.[60] The CD spectrum of **Polymer 1D** at pH 4 had two minima at 209 nm and 224 nm and a maximum at 195 nm corresponding to the α-helical form.[60] . 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 6.A.**) attributable to a random coil conformation.

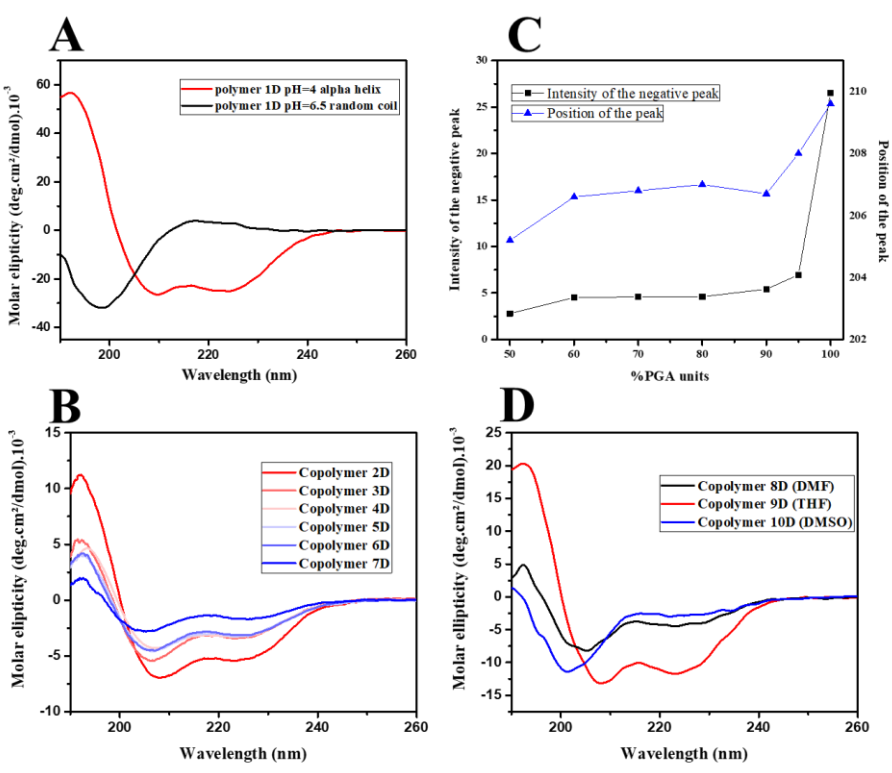


Figure 6. (A) CD spectra of the two forms of PGA (polymer **1D**); (B) CD spectra of copolymers **2D-7D** in helix inducing media (pH=4); (C) data from (A) and (B) was plotted to show the effect of the %PGA units in polymer **1D** and copolymers **2D-7D** on the intensity of the negative peak around 209nm and its position; (D) CD spectra of copolymers **8D-10D** in helix-inducing aqueous media (pH=4).

Since this α-helical structuring requires long sequences of PGA units,[55] we hypothesized that small incorporations of glycine units could disrupt the structure. **Figure 6.B.** shows the CD spectra of the deprotected copolymers (**2D-7D**) at pH=4 where PGA adopts an α-helical structure. The peak intensity is significantly reduced even with only 5% glycine content (**Figure 6.C.**) as in the case of copolymer **2D**. Nonetheless, copolymer **7D** with 50% glycine content retains helical structuring (Figure **6.B.**). The CD spectra also show a shift in the peak maximum towards lower wavelengths (**Figure 6.C.**), suggesting an

incomplete shift towards a random coil conformation, which is characterized by the strong negative peak below 200 nm.

The deprotected copolymers (**8D**-**10D**) containing 50% glycine but prepared in different solvents showed varying behaviors at pH=4. Copolymer **9D**, synthesized in THF showed significant helical conformation despite its high glycine content (50%). In contrast, copolymers **8D** and **10D** showed more random coil behavior with the negative peak shifted towards lower wavelengths. This was more pronounced for copolymer **10D** synthesized in DMSO, which also had the highest gradient nature (**Figure 6.D.**).

These results confirm that the way in which the comonomers are distributed within the copolymer has a strong effect on its secondary structure. Copolymer **9D**, which has the most similar reactivity ratios and the most random monomer distribution, has the most helical nature. We thus hypothesize that the incorporation of evenly distributed glycine units along the PGA chain does not disrupt its helical structure. In other words, helical conformations can still be obtained when isolated chiral glutamate units are replaced by achiral units without side chains. The helicity signal from copolymer **9D** is around half of that of polymer **1D**, corresponding to the reduction in the number of chiral glutamate units due to their replacement by achiral glycine. On the other hand, if the monomer distribution is less even, the conformation is disrupted. A highly gradient copolymer such as copolymer **10D** (**Figure 3.B**), 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 does not form a long enough chain to form a helix structure.

CONCLUSIONS

The copolymerization of BLG-NCA and glycine-NCA was investigated as a model system to design synthetic IDPs using ROP of NCA. Feed ratio, solvent and temperature were changed to observe their effect on the copolymerization kinetics. It was found that by changing these reaction conditions, copolymers with different microstructural properties could be obtained. Differences in primary sequence, analyzed by Monte Carlo simulations and ^{13}C quantitative NMR analysis, correlated with differences in secondary structure after deprotection. As the glycine content of copolymers prepared in DMF increased, the helical nature decreased linearly. At a glycine content of 50%, the polymer had low helicity compared to polypeptide without glycine but still retained a helical CD signature. Synthesis performed in higher polarity solvents accentuated the gradient nature of the copolymers, resulting in the formation of copolymers with significant random coil conformation. Increasing the polymerization temperature had a similar effect. Our experimental conditions thus allow to exert a degree of control over the distribution of comonomers along the polymer chains and the use of Glycine NCA comonomers allow to mimic the properties of IDPs in a synthetic polymer scaffold. Production of intrinsically disordered PBLG was shown to be possible and paves new ways towards the use of a hydrophobic and homochiral polypeptide exhibiting tunable helicity or a coil-disordered state in solution.

ASSOCIATED CONTENT

Supporting Information. NMR spectra, chromatograms, and other results can be consulted.

ACKNOWLEDGMENT

The authors acknowledge Amélie Vax for assistance with size-exclusion chromatography, Anne-Laure Wirotius-Lambert for NMR methodology, Estelle Morvan and Axel Grelard from IECB for high field NMR.

FUNDING SOURCES

M.B. received support from The French Ministry of Higher Education, Research and Innovation (Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation, MESRI). This work was granted access to the TGIR-RMN-THC FR3050 CNRS (high field NMR facilities) at the CBMN - IECB Bordeaux. This work was supported by a grant overseen by the French National Research Agency (ANR): Grant No. ANR-20-CE06-0020-01. P.SA. received support from CONACYT (scholarship holder No. 548662). This work was conducted in the framework of the University of Bordeaux's IdEx "Investments for the Future" program RRI "Frontiers Of Life" that received financial support from the French government.

REFERENCES

(1) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. Conformation of the Backbone in Unfolded Proteins. *Chem. Rev.* **2006**, *106* (5), 1877–1897. https://doi.org/10.1021/cr040433a.

(2) Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. Charge Interactions Can Dominate the Dimensions of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci.* **2010**, *107* (33), 14609–14614. https://doi.org/10.1073/pnas.1001743107.

(3) Theillet, F.-X.; Kalmar, L.; Tompa, P.; Han, K.-H.; Selenko, P.; Dunker, A. K.; Daughdrill, G. W.; Uversky, V. N. The Alphabet of Intrinsic Disorder. *Intrinsically Disord. Proteins* **2013**, *1* (1), e24360. https://doi.org/10.4161/idp.24360.

(4) Radivojac, P.; Iakoucheva, L. M.; Oldfield, C. J.; Obradovic, Z.; Uversky, V. N.; Dunker, A. K. Intrinsic Disorder and Functional Proteomics. *Biophys. J.* **2007**, *92* (5), 1439–1456. https://doi.org/10.1529/biophysj.106.094045.

(5) Serrano, L. Comparison between the φ Distribution of the Amino Acids in the Protein Database and NMR Data Indicates That Amino Acids Have Various φ Propensities in the Random Coil Conformation. *J. Mol. Biol.* **1995**, *254* (2), 322–333. https://doi.org/10.1006/jmbi.1995.0619.

(6) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why Are ?Natively Unfolded? Proteins Unstructured under Physiologic Conditions? *Proteins Struct. Funct. Genet.* **2000**, *41* (3), 415–427. https://doi.org/10.1002/1097-0134(20001115)41:3<415::AID-PROT130>3.0.CO;2-7.

(7) Imai, K.; Mitaku, S. Mechanisms of Secondary Structure Breakers in Soluble Proteins. *Biophysics (Oxf).* **2005**, *1*, 55–65. https://doi.org/10.2142/biophysics.1.55.

(8) Song, Z.; Tan, Z.; Cheng, J. Recent Advances and Future Perspectives of Synthetic Polypeptides from N-Carboxyanhydrides. *Macromolecules* **2019**, *52* (22), 8521–8539. https://doi.org/10.1021/acs.macromol.9b01450.

(9) Mbizana, S.; Hlalele, L.; Pfukwa, R.; Du Toit, A.; Lumkwana, D.; Loos, B.; Klumperman, B. Synthesis and Cell Interaction of Statistical <scp>l</Scp> -Arginine–Glycine– <scp>l</Scp> -Aspartic Acid Terpolypeptides. *Biomacromolecules* **2018**, *19* (7), 3058–3066. https://doi.org/10.1021/acs.biomac.8b00620.

(10) Wamsley, A.; Phiasivongsa, P.; Jasti, B.; Li, X. Microstructural Characterization of Terpolymers of Leucine, β-Benzyl Aspartate, and Valine. *J. Polym. Sci. Part A Polym. Chem.* **2006**, *44* (14), 4328–4337. https://doi.org/10.1002/pola.21536.

(11) Zhou, M.; Zou, J.; Liu, L.; Xiao, X.; Deng, S.; Wu, Y.; Xie, J.; Cong, Z.; Ji, Z.; Liu, R. Synthesis of Poly-α/β-Peptides with Tunable Sequence via the Copolymerization on N-Carboxyanhydride and N-Thiocarboxyanhydride. *iScience* **2021**, *24* (10), 103124. https://doi.org/10.1016/j.isci.2021.103124.

(12) Wang, J.; Zhou, P.; Shen, T.; Xu, S.; Bai, T.; Ling, J. Glycine N -Thiocarboxyanhydride: A Key to Glycine-Rich Protein Mimics. *ACS Macro Lett.* **2023**, *12* (11), 1466–1471. https://doi.org/10.1021/acsmacrolett.3c00491.

(13) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. Sequence-Controlled Polymers. *Science (80-.).* **2013**, *341* (6146), 1238149. https://doi.org/10.1126/science.1238149.

(14) Lutz, J. F.; Schmidt, B. V. K. J.; Pfeifer, S. Tailored Polymer Microstructures Prepared by Atom Transfer Radical Copolymerization of Styrene and N-Substituted Maleimides. *Macromol. Rapid Commun.* **2011**, *32* (2), 127–135. https://doi.org/10.1002/marc.201000664.

(15) Harrisson, S. The Downside of Dispersity: Why the Standard Deviation Is a Better Measure of Dispersion in Precision Polymerization. *Polym. Chem.* **2018**, *9* (12), 1366–1370. https://doi.org/10.1039/C8PY00138C.

(16) Odian, G. *Principles of Polymerization: Chap. 2 Step Polymerization*, 4th ed.; Wiley, Ed.; 2004.

(17) Zhang, J.; Farias-Mancilla, B.; Kulai, I.; Hoeppener, S.; Lonetti, B.; Prévost, S.; Ulbrich, J.; Destarac, M.; Colombani, O.; Schubert, U. S.; Guerrero-Sanchez, C.; Harrisson, S. Effect of Hydrophilic Monomer Distribution on Self-Assembly of a PH-Responsive Copolymer: Spheres, Worms and Vesicles from a Single Copolymer Composition. *Angew. Chemie Int. Ed.* **2021**, *60* (9), 4925–4930. https://doi.org/10.1002/anie.202010501.

(18) Baumgartner, R.; Fu, H.; Song, Z.; Lin, Y.; Cheng, J. Cooperative Polymerization of α-Helices Induced by Macromolecular Architecture. *Nat. Chem.* **2017**, *9* (7), 614–622. https://doi.org/10.1038/nchem.2712.

(19) Sederel.W, Deshmane.S, Hayashi.T, A. J. M. The Random Copolymerization of Y-

Benzyl-L- Glutamate and L-Valine N-Carboxyanhydrides : Reactivity Ratio and Heterogeneity Studies. **1978**, *17*, 2835–2849.

(20) Atreyi, M.; Rao, M. V. R.; Kumar, S. Copolymerization Kinetics of N-carboxyanhydrides of Some A-amino Acids: Influence of Nature of Amino Acids. *Biopolymers* **1983**, *22* (2), 747–753. https://doi.org/10.1002/bip.360220213.

(21) Kricheldorf, H. R.; Müller, D.; Hull, W. E. Secondary Structure of Peptides: 19. Primary/Secondary Structure Relationships of Leucine/Valine Copolypeptides Prepared from α-Amino Acid N-Carboxyanhydrides. *Int. J. Biol. Macromol.* **1986**, *8* (1), 20–26. https://doi.org/10.1016/0141-8130(86)90067-X.

(22) Saudek, V.; Stejskal, J.; Schmidt, P.; Zimmermann, K.; Škarda, V.; Kratochvíl, P.; Drobník, J. Correlation between the Sequence-length Distribution and Structure of Statistical Copolymers of L-valine and Γ-benzyl-L-glutamate. *Biopolymers* **1987**, *26* (5), 705–725. https://doi.org/10.1002/bip.360260511.

(23) Fan, J.; Zou, J.; He, X.; Zhang, F.; Zhang, S.; Raymond, J. E.; Wooley, K. L. Tunable Mechano-Responsive Organogels by Ring-Opening Copolymerizations of N-Carboxyanhydrides. *Chem. Sci.* **2014**, *5* (1), 141–150. https://doi.org/10.1039/c3sc52504j.

(24) Salas-Ambrosio, P.; Tronnet, A.; Badreldin, M.; Ji, S.; Lecommandoux, S.; Harrisson, S.; Verhaeghe, P.; Bonduelle, C. Effect of N -Alkylation in N -Carboxyanhydride (NCA) Ring-Opening Polymerization Kinetics. *Polym. Chem.* **2022**, *13* (43), 6149–6161. https://doi.org/10.1039/D2PY00985D.

(25) Wendisch, V. F. Metabolic Engineering Advances and Prospects for Amino Acid Production. *Metab. Eng.* **2020**, *58* (March 2019), 17–34. https://doi.org/10.1016/j.ymben.2019.03.008.

(26) Johnson, L. C.; Akinmola, A. T.; Scholz, C. Poly(Glutamic Acid): From Natto to Drug Delivery Systems. *Biocatal. Agric. Biotechnol.* **2022**, *40* (September 2021), 102292. https://doi.org/10.1016/j.bcab.2022.102292.

(27) Bonduelle, C. Secondary Structures of Synthetic Polypeptide Polymers. *Polym. Chem.* **2018**, *9* (13), 1517–1529. https://doi.org/10.1039/C7PY01725A.

(28) Conejos-Sánchez, I.; Duro-Castano, A.; Birke, A.; Barz, M.; Vicent, M. J. A Controlled and Versatile NCA Polymerization Method for the Synthesis of Polypeptides. *Polym. Chem.* **2013**, *4* (11), 3182–3186. https://doi.org/10.1039/c3py00347g.

(29) Habraken, G. J. M.; Peeters, M.; Dietz, C. H. J. T.; Koning, C. E.; Heise, A. How Controlled and Versatile Is N-Carboxy Anhydride (NCA) Polymerization at 0 °C? Effect of Temperature on Homo-, Block- and Graft (Co)Polymerization. *Polym. Chem.* **2010**, *1* (4), 514–524. https://doi.org/10.1039/b9py00337a.

(30) Song, Z.; Fu, H.; Baumgartner, R.; Zhu, L.; Shih, K. C.; Xia, Y.; Zheng, X.; Yin, L.; Chipot, C.; Lin, Y.; Cheng, J. Enzyme-Mimetic Self-Catalyzed Polymerization of Polypeptide Helices. *Nat. Commun.* **2019**, *10* (1), 1–7. https://doi.org/10.1038/s41467-019-13502-w.

(31) Baumgartner, R.; Kuai, D.; Cheng, J. Synthesis of Controlled, High-Molecular

Weight Poly(l-Glutamic Acid) Brush Polymers. *Biomater. Sci.* **2017**, *5* (9), 1836–1844. https://doi.org/10.1039/c7bm00339k.

(32) Tian, H. Y.; Deng, C.; Lin, H.; Sun, J.; Deng, M.; Chen, X.; Jing, X. Biodegradable Cationic PEG-PEI-PBLG Hyperbranched Block Copolymer: Synthesis and Micelle Characterization. *Biomaterials* **2005**, *26* (20), 4209–4217. https://doi.org/10.1016/j.biomaterials.2004.11.002.

(33) Sugimoto, H.; Nakanishi, E.; Hanai, T.; Yasumura, T.; Inomata, K. Aggregate Formation and Release Behaviour of Hydrophobic Drugs with Graft Copolypeptide-Containing Tryptophan. *Polym. Int.* **2004**, *53* (7), 972–983. https://doi.org/10.1002/pi.1485.

(34) 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. Lipid-Polyglutamate Nanoparticle Vaccine Platform. *ACS Appl. Mater. Interfaces* **2021**, *13* (5), 6011–6022. https://doi.org/10.1021/acsami.0c20607.

(35) Mahi, B.; Gauthier, M.; Hadjichristidis, N. Hybrid Arborescent Polypeptide-Based Unimolecular Micelles: Synthesis, Characterization, and Drug Encapsulation. *Biomacromolecules* **2022**, *23* (6), 2441–2458. https://doi.org/10.1021/acs.biomac.2c00202.

(36) Ijadi Bajestani, M.; Mousavi, S. M.; Mousavi, S. B.; Jafari, A.; Shojaosadati, S. A. Purification of Extra Cellular Poly-γ-Glutamic Acid as an Antibacterial Agent Using Anion Exchange Chromatography. *Int. J. Biol. Macromol.* **2018**, *113*, 142–149. https://doi.org/10.1016/j.ijbiomac.2018.02.082.

(37) Ikeda, M.; Akagi, T.; Yasuoka, T.; Nagao, M.; Akashi, M. Characterization and Analytical Development for Amphiphilic Poly(γ-Glutamic Acid) as Raw Material of Nanoparticle Adjuvants. *J. Pharm. Biomed. Anal.* **2018**, *150*, 460–468. https://doi.org/10.1016/j.jpba.2017.12.034.

(38) Khair, B. M. A.-E.; Kōmoto, T.; Kawai, T. Formation of Polyglycine Crystals in the Course of Polymerization of Glycine N-carboxy Anhydride Initiated by Primary Diamines. *Die Makromol. Chemie* **1976**, *177* (8), 2481–2489. https://doi.org/10.1002/macp.1976.021770818.

(39) Komoto, T.; Oya, M.; Kawai, T. Crystallization of Polypeptides in the Course of Polymerisation 6. *Die Makromol. Chemie* **1973**, *175*, 301–310.

(40) Kuroki, S.; Takahashi, A.; Ando, I.; Shoji, A.; Ozaki, T. Hydrogen-Bonding Structural Study of Solid Peptides and Polypeptides Containing a Glycine Residue by 17O NMR Spectroscopy. *J. Mol. Struct.* **1994**, *323* (C), 197–208. https://doi.org/10.1016/0022-2860(94)07993-5.

(41) Lamy, C.; Lemoine, J.; Bouchu, D.; Goekjian, P.; Strazewski, P. Glutamate-Glycine and Histidine-Glycine Co-Oligopeptides: Batch Co-Oligomerization versus Pulsed Addition of N-Carboxyanhydrides. *ChemBioChem* **2008**, *9* (5), 710–713. https://doi.org/10.1002/cbic.200700720.

(42) Papadopoulos, P.; Floudas, G.; Schnell, I.; Aliferis, T.; Iatrou, H.; Hadjichristidis,

N. Nanodomain-Induced Chain Folding in Poly(γ-Benzyl-L-Glutamate)-b-Polyglycine Diblock Copolymers. *Biomacromolecules* **2005**, *6* (4), 2352–2361. https://doi.org/10.1021/bm0501860.

(43) Skoulas, D.; Stavroulaki, D.; Santorinaios, K.; Iatrou, H. Synthesis of Hybrid-Polypeptides m-PEO-b-Poly(His-Co-Gly) and m-PEO-b-Poly(His-Co-Ala) and Study of Their Structure and Aggregation. Influence of Hydrophobic Copolypeptides on the Properties of Poly(L-Histidine). *Polymers (Basel).* **2017**, *9* (11), 564. https://doi.org/10.3390/polym9110564.

(44) Schilling, G.; Kricheldorf, H. 13C NMR Sequence Analysis, 6. Sequence Polypeptides of Glycine and B-alanine. *Die Makromol. Chemie* **1977**, *178* (3), 885–892. https://doi.org/10.1002/macp.1977.021780324.

(45) Mayo, F. R.; Lewis, F. M. Copolymerization. I. A Basis for Comparing the Behavior of Monomers in Copolymerization; The Copolymerization of Styrene and Methyl Methacrylate. *J. Am. Chem. Soc.* **1944**, *66* (9), 1594–1601. https://doi.org/10.1021/ja01237a052.

(46) Meyer, V. E.; Lowry, G. G. Integral and Differential Binary Copolymerization Equations. *J. Polym. Sci. Part A Gen. Pap.* **1965**, *3* (8), 2843–2851. https://doi.org/10.1002/pol.1965.100030811.

(47) Konkolewicz, D.; Krys, P.; Matyjaszewski, K. Explaining Unexpected Data via Competitive Equilibria and Processes in Radical Reactions with Reversible Deactivation. *Acc. Chem. Res.* **2014**, *47* (10), 3028–3036. https://doi.org/10.1021/ar500199u.

(48) Harrisson, S.; Ercole, F.; Muir, B. W. Living Spontaneous Gradient Copolymers of Acrylic Acid and Styrene: One-Pot Synthesis of PH-Responsive Amphiphiles. *Polym. Chem.* **2010**, *1* (3), 326–332. https://doi.org/10.1039/b9py00301k.

(49) Van Den Brink, M.; Van Herk, A. M.; German, A. L. Nonlinear Regression by Visualization of the Sum of Residual Space Applied to the Integrated Copolymerization Equation with Errors in All Variables. I. Introduction of the Model, Simulations and Design of Experiments. *J. Polym. Sci. Part A Polym. Chem.* **1999**, *37* (20), 3793–3803. https://doi.org/10.1002/(SICI)1099-0518(19991015)37:20<3793::AID-POLA8>3.0.CO;2-Q.

(50) Bradford, K. G. E.; Gilbert, R. D.; Weerasinghe, M. A. S. N.; Harrisson, S.; Konkolewicz, D. Spontaneous Gradients by ATRP and RAFT: Interchangeable Polymerization Methods? *Macromolecules* **2023**, *56* (21), 8784–8795. https://doi.org/10.1021/acs.macromol.3c01426.

(51) Yamada, K.; Nakano, T.; Okamoto, Y. Free-Radical Copolymerization of Vinyl Esters Using Fluoroalcohols as Solvents: The Solvent Effect on the Monomer Reactivity Ratio. *J. Polym. Sci. Part A Polym. Chem.* **2000**, *38* (1), 220–228. https://doi.org/10.1002/(SICI)1099-0518(20000101)38:1<220::AID-POLA27>3.0.CO;2-P.

(52) Barson, C. A.; Turner, M. J. The Temperature Dependence of the Monomer Reactivity Ratios in the Copolymerization of Styrene with Methyl and Ethyl

Cinnamates. *Eur. Polym. J.* **1973**, *9* (8), 789–793. https://doi.org/10.1016/0014-3057(73)90011-6.

(53) Umoren, S. A.; Solomon, M. M.; Saji, V. S. Basic Polymer Concepts II. In *Polymeric Materials in Corrosion Inhibition*; Saviour A. Umoren, Moses M. Solomon, V. S. S., Ed.; Elsevier, 2022; pp 49–81. https://doi.org/10.1016/b978-0-12-823854-7.00025-4.

(54) Rinaudo, M.; Domard, A. Etude Potentiométrique de l'alpha-Poly(Acide-L-Glutamique) et de Ses Oligomères. *Biopolymers* **1973**, *12*, 2211–2224.

(55) Alekperov, D.; Shirosaki, T.; Sakurai, T.; Popova, G.; Kireev, V.; Ihara, H. Synthesis and Conformational Characterization of Oligopeptide-Cyclotriphosphazene Hybrids. *Polym. J.* **2003**, *35* (5), 417–421. https://doi.org/10.1295/polymj.35.417.

(56) Kratohvil, J. P. On Solution Properties of Poly(γ-Benzyl-L-Glutamate). *Kolloid-Zeitschrift Zeitschrift für Polym.* **1970**, *238* (1–2), 455–459. https://doi.org/10.1007/BF02085572.

(57) 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. https://doi.org/10.1002/pol.1985.130231115.

(58) Miles, A. J.; Janes, R. W.; Wallace, B. A. Tools and Methods for Circular Dichroism Spectroscopy of Proteins: A Tutorial Review. *Chem. Soc. Rev.* **2021**. https://doi.org/10.1039/d0cs00558d.

(59) Bonduelle, C. Secondary Structures of Synthetic Polypeptide Polymers. *Polym. Chem.* **2018**, *9* (13), 1517–1529. https://doi.org/10.1039/c7py01725a.

(60) Aujard-Catot, J.; Nguyen, M.; Bijani, C.; Pratviel, G.; Bonduelle, C. Cd2+ Coordination: An Efficient Structuring Switch for Polypeptide Polymers. *Polym. Chem.* **2018**, *9* (30), 4100–4107. https://doi.org/10.1039/c8py00810h.

TABLE OF CONTENTS GRAPHIC