Macromolecules

Article

pubs.acs.org/Macromolecules

# Revisiting Secondary Structures in NCA Polymerization: Influences on the Analysis of Protected Polylysines

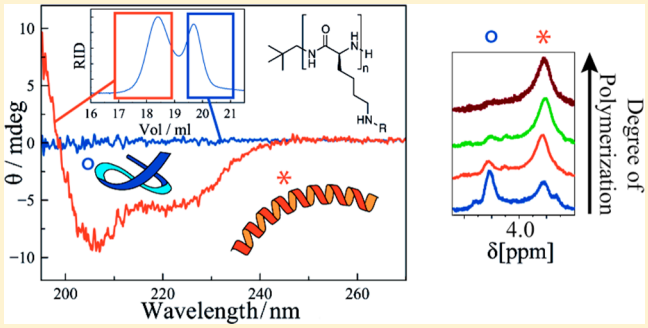
David Huesmann,[†] Alexander Birke,[†] Kristina Klinker,[†] Stephan Türk,[‡] Hans Joachim Räder,[‡] and Matthias Barz*[,†]

[†]Institute of Organic Chemistry, Johannes Gutenberg-Universität Mainz, Duesbergweg 10-14, 55128 Mainz, Germany

[‡]Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

Supporting Information

**ABSTRACT:** Two series (degree of polymerization: 20−200) of polylysines with Z and TFA protecting groups were synthesized, and their behavior in a range of analytical methods was investigated. Gel permeation chromatography of the smaller polypeptides reveals a bimodal distribution, which is lost in larger polymers. With the help of GPC, NMR, circular dichroism (CD), and MALDI-TOF, it was demonstrated that the bimodal distribution is not due to terminated chains or other side reactions. Our results indicate that the bimodality is caused by a change in secondary structure of the growing peptide chain that occurs around a degree of polymerization of about 15. This change in secondary structure interferes strongly with the most used analysis method for polymers—GPC—by producing a bimodal distribution as an artifact. After deprotection, the polypeptides were found to exhibit exclusively random coil conformation, and thus a monomodal GPC elugram was obtained. The effect can be explained by a 1.6-fold increase in the hydrodynamic volume at the coil−helix transition. This work demostrates that secondary structures need to be carefully considered when performing standard analysis on polypeptidic systems.




## INTRODUCTION

The synthesis of polypeptides based on the ring-opening polymerization of $\alpha$-amino acid-*N*-carboxyanhydrides (NCAs) has attracted synthetic chemists since the early 20th century, when Leuchs discovered this class of monomers.[1−3] Since the end of the 20th and the beginning of the 21st century, various techniques to control the polymerization have emerged.[4−11] Even nowadays, the synthesis of polypeptides under controlled or even living conditions seems demanding, which is usually attributed due to the presence of side reactions.[12] While it appears reasonable that at higher degrees of polymerization side reactions like the activated monomer mechanism (AMM) etc. occur, it is interesting that also at low molecular weight polymer side reactions are a matter of debate.[13,14] Side reactions due to the cleavage of relatively labile protecting groups, such as benzyl esters or trifluoroacetic acid amide, might occur under certain conditions due to a nucleophilic attack of the amino terminus at the carbonyl carbon.[15] But in the case of robust protecting groups such as benzyl groups, or even trifluoroacetic acid amide at ambient or low temperatures, these side reactions are very unlikely to occur. On the other hand, also conformational effects are known to influence the propagation speed of the NCA polymerization and therefore influence the molecular weight distribution.[2,5]

In all cases precise sample characterization becomes rather demanding when polymers undergo a change in superstructure at a certain chain length—for example, a coil-to-helix transition. In the case of polypeptides three major secondary structures are observed in nature: $\alpha$-helix, $\beta$-sheet, and random coil. Since $\beta$-sheets often precipitate when formed during the NCA polymerization, polymer chemists usually deal with helical and random coil structures in solution.

A number of methods can be used to characterize the different secondary structures. X-ray crystallography is the most used method when it comes to the study of protein secondary and tertiary structures in the solid state.[16] However, high-quality single crystals are required for this technique. In solution, NMR techniques are also often used to determine protein structures and the chemical shift alone can be used to identify secondary structures.[17,18] Further, circular dichroism (CD)[19,20] and IR[21,22] spectroscopy are useful tools for the determination of secondary structures.

It is known since several decades that there is a change in secondary structure as the polypeptide chain grows. Oligo($\gamma$-ethyl glutamate), for example, was found by CD spectroscopy and NMR in trifluoroethanol to change from a random coil conformation to an $\alpha$-helix at a degree of polymerization around seven.[23] Kricheldorf and Müller also investigated the

Received: January 6, 2014
Revised: January 9, 2014
Published: January 22, 2014

secondary structure of polypeptides in the solid state with the help of IR spectroscopy and $^{13}C$ NMR CP/MAS.[24,25]

The different techniques that have been developed to control NCA polymerization can lead to different end groups,[4,8] which may also have an effect on the secondary structures. In this work we have used primary amines as well as their $HBF_4$ salts to initiate the polymerization of lysine NCAs giving us an alkylamide at the C-terminus and a free amine at the N-terminus.[26] Thus, the effects reported in this work may not be present in all controlled NCA polymerization techniques, although it appears likely that these effects are much more related to the individual monomer than to a specific end group.

In this work we report the synthesis of poly(*N*-ε-benzyloxycarbonyl-L-lysine) (PLys(Z)) and poly(*N*-ε-trifluoroacetyl-L-lysine) (PLys(TFA)) with a degree of polymerization varying from 20 to 200. Polylysine has attracted much attention and is used in a variety of applications like coatings for improved cell attachment[27] or as polycations for gene delivery.[28] Its use in these and other important industrial applications, e.g. detergent formulations, water treatment, or coatings, makes polylysine an industrially relevant polypeptide, and a good insight into its polymerization behavior is indispensable. The ring-opening polymerizations (ROP) were initiated by primary amines as well as their tetrafluoroborane ammonium salts.[9,25] By combining different methods like GPC, CD spectroscopy, NMR, and MALDI-TOF, we show that in the case of the protected polylysines PLys(Z) and PLys(TFA) termination reactions are absent under the applied conditions. Most importantly, we report on a change in secondary structure around a degree of polymerization of 15 that interferes with standard GPC analysis. Finally, we would like to provide a model that explains the observed effects by taking changes in the hydrodynamic volume during the coil-to-helix transition into account.

## EXPERIMENTAL SECTION

**Materials and Methods.** DMF was dried by stirring over molecular sieve (3 Å) and BaO. It was then distilled *in vacuo* at low to ambient temperature onto molecular sieve to remove dimethylamine impurities, was protected from light, and was stored at −80 °C. Ethyl acetate, THF, hexane, and cyclohexane were distilled from Na/K; other solvents were used as received unless otherwise stated. Neopentylamine was dried over $CaH_2$ and distilled before use. Protected L-lysines were purchased from Orpegen.

$^1H$ and $^{13}C$ NMR spectra were recorded on a Bruker AC 300 or AV 400 at room temperature. The spectra were calibrated using the solvent signals.[29]

Infrared spectroscopy was measured on a Jasco FT/IR-4100 with an ATR sampling accessory (MIRacle, Pike Technologies). IR spectra were analyzed using Spectra Manager 2.0 (Jasco).

Gel permeation chromatography (GPC) was performed with DMF containing 0.25 g/L LiBr as eluent at 50 °C. The column was packed with HEMA 300/100/40. A refractive index detector (G 1362A RID) was used to detect the polymer. GPC in HFIP was performed with 3 g/L $K^+TFA^-$ at 40 °C. The columns were packed with modified silica (PFG columns, particle size: 7 μm, porosity: 100 and 1000 Å). A refractive index detector (G 1362A RID, Jasco) was used to detect the polymer. Molecular weights were calculated using a calibration performed with PMMA standards (Polymer Standards Services GmbH). GPC in water was performed with buffered aqueous solution (50 mM sodium phosphate, 150 mM sodium chloride, pH 7). The following parts were used: Jasco pump (pU-2086 Plus series), a Jasco UV/vis detector (UV-2077 Plus), a Jasco RI-detector (Jasco RI 2031 Plus series). The flow rate was set to 0.4 mL $min^{-1}$. A Superose6 10/300 GL column was used. Calibration was done using protein standards. In all cases the elution diagram was evaluated with PSS WinGPC from Polymer Standard Service Mainz.

CD spectroscopy was performed on a Jasco J-815 spectrometer in a cell with a path length of 1 mm. The temperature was kept constant at 20 °C. Spectra were recorded at concentrations of 0.5, 0.3, and 0.2 mg $mL^{-1}$ for PLys(Z) (in HFIP, 3 g $L^{-1}$ $K^+TFA^-$), PLys(TFA) (in HFIP, 3 g $L^{-1}$ $K^+TFA^-$), and deprotected Lys (in Milli-Q water), respectively. Each spectrum represents the average of two scans. The mean residual ellipticity $\theta_{MR}$ was calculated from the observed ellipticity $\theta$ (see Supporting Information).

For the spectra of the fractions Z1_S and Z1_B, which were collected from the GPC outlet, the concentration was unknown, and 30 scans were recorded and averaged.

The mass measurements were carried out with a REFLEX MALDI-TOF mass spectrometer (Bruker, Bremen, Germany), equipped with a 337 nm nitrogen laser. 1,8,9-Trihydroxyanthracene (dithranol) was used as a MALDI matrix. For sample preparation 50 μL of the matrix dissolved in THF (5 mg/mL) was mixed with 10 μL of the fractionated polymer solution and crystallized on the stainless steel target immediately before measurement. The mass spectrometer was calibrated externally with a $C_{60}/C_{70}$ fullerene mixture. The instrument was operated in linear and reflection mode, both at an acceleration voltage of 20 kV. In linear mode the Bruker HIMAS detector was used, providing high sensitivity and low resolution and in reflection mode dual channel plates were used, providing high resolution but lower sensitivity. The mass spectra were smoothed and baseline corrected with the XMASS data processing program (Bruker). mMass was used to evaluate the spectra.[30−32]

**Initiator Synthesis.** To 5 mL (5.59 g, 36.74 mmol) of $HBF_4{\cdot}Et_2O$, 4.31 mL (3.20 g, 36.74 mmol) of neopentylamine was slowly added. The mixture was cooled using an ice bath. The addition resulted in the precipitation of a slightly brown solid. The ether was removed *in vacuo*, and the solid was recrystallized two times from ethyl acetate and washed with cyclohexane. The product was dried *in vacuo* yielding 2.72 g (15.59 mmol, 42%) of neopentylammonium tetrafluoroborate as a colorless solid. $^1H$ NMR (300 MHz, DMSO-$d_6$) $\delta$ [ppm] = 7.58 (s, 3H, $NH_3^+$), 2.63 (s, 2H, $CH_2$), 0.93 (s, 9H, $C(CH_3)_3$). $^{13}C$ NMR (75 MHz, DMSO-$d_6$) $\delta$ [ppm] = 49.94 ($CH_2NH_3^+$), 30.21 ($C(CH_3)_3$), 26.78 (3H, $C(CH_3)_3$). FT-IR: (neat) $\nu$ [$cm^{-1}$] = 3248 br w ($NH_3^+$), 2963 w (CH), 1617 w, 1508 m, 1482 w, 990 br s ($BF_4^-$), 849 w, 520 m. Anal. Calcd for $C_5H_{14}NBF_4$: C, 34.43; H, 8.06; N, 8.00. Found: C, 34.35; H, 7.99; N, 8.07.

*Monomer Synthesis. N-ε-Benzyloxycarbonyl-L-lysine-N-carboxyanhydride.* In a three-necked flask equipped with a reflux condenser, dropping funnel, and septum, 14.00 g (50 mmol) of Z-protected lysine was suspended in 100 mL of THF, and the suspension was heated to 70 °C. 6.6 mL (55 mmol) of diphosgene was added over 30 min. The solution was heated until all solid disappeared (30−60 min). Dry nitrogen was then bubbled through the solution for 2−3 h to remove excess HCl and phosgene. The solution was concentrated *in vacuo*, and dry cyclohexane was added to precipitate the NCA. The suspension was stored in the fridge for 1 h. The solid was collected by filtration in an inert atmosphere and washed with cyclohexane. It was then dissolved in the smallest possible volume of THF, and again cyclohexane was added to precipitate the product. The suspension was left in the fridge overnight. The solid was then collected by filtration in an inert atmosphere and dried in a stream of dry nitrogen. The recrystallization was repeated and $AgNO_3$ (0.1 M in water) was added to the filtrate, to confirm the absence of chloride ions (AgCl would immediately precipitate as a colorless solid). The Lys(Z)NCA, a colorless solid (13.23 g, 43.2 mmol, 86%), was transferred to a Schlenk tube and stored at −80 °C; mp 99.1 °C. $^1H$ NMR (400 MHz, DMSO-$d_6$) $\delta$ [ppm] = 9.09 (s, 1H, CO*NH*−$C_\alpha$), 7.40−7.25 (m, 6H, *Ar H*, *NH*(Z)), 5.01 (s, 2H, NH$CH_2$Ph), 4.43 (t, 1H, $C_\alpha H$), 2.99 (q, 2H, $CH_2$NH), 1.74−1.63 (m, 2H, CH−$CH_2$), 1.44−1.31 (m, 4H, $CH_2$−$CH_2$). $^{13}C$ NMR (75 MHz, DMSO-$d_6$) $\delta$ [ppm] = 171.61 ($C_\alpha$COOC=NH), 156.06 (COOBn), 151.94 ($C_\alpha$NHCOO), 137.21 (*Ar*), 128.29 (*Ar*), 127.67 (*Ar*), 65.09 (PhCH$_2$), 56.97 ($C_\alpha$), 39.88 ($C_\varepsilon$), 30.58 ($C_\beta$), 28.72 ($C_\delta$), 21.55 ($C_\gamma$).


dx.doi.org/10.1021/ma5000392 | *Macromolecules* 2014, 47, 928−936

*N-ε-Trifluoroacetyl-L-lysine-N-caboxyanhydride.* 10.15 g (41.9 mmol) of Lys(TFA) was dissolved in 150 mL of ethyl acetate, and the suspension was heated to 70 °C. 6 mL (50.3 mmol) of diphosgene was added via a syringe over 3 h. After an additional 1 h, in which the solution remained slightly cloudy, a nitrogen stream was passed through the solution for 1 h, and solution was concentrated *in vacuo*. 35 mL of AcOEt was added, and the cloudy solution was filtered. The product was crystallized by slowly adding 80 mL of hexane to the filtrate over 2.5 h. The product was collected by filtration and recrystallized two times. The product was dried *in vacuo* for 1 h, yielding 7.12 g (26.6 mmol; 63%) of colorless crystals. The product was stored at −80 °C; mp 101 °C. $^1$H NMR (400 MHz, DMSO-$d_6$, $\delta$) 9.40 (t, 1H; CO*NH*−$CF_3$), 9.09 (s, 1H; CO*NH*−CH), 4.43 (t, 1H; *CH*−$CH_2$), 3.17 (q, 2H; *CH*$_2$−NH), 1.86−1.58 (m, 2H; CH−*CH*$_2$), 1.50 (p, 2H; $CH_2$−*CH*$_2$−$CH_2$−NH), 1.44−1.21 (m, 2H; *CH*$_2$−$CH_2$−$CH_2$−NH).

*Polymer Synthesis.* Polymerizations were carried out in Schlenk tubes in DMF. The NCAs were dissolved in in DMF at a concentration of 100 mg cm$^{-3}$. The initiator (neopentylammonium tetrafluoroborate or neopentylamine) was added to initiate the polymerization. The solution was stirred at 20 °C, and the reaction flask was opened to the Schlenk line to allow the escape of $CO_2$. The progress of the polymerization was monitored by IR. Once all NCA bands disappeared the mixture was precipitated in diethyl ether and washed with ether twice. The polymer was then freeze-dried from dioxane yielding a colorless fluffy powder (typical yields 80%).

*Poly(N-ε-benzyloxycarbonyl-L-lysine) by Neopentylammonium Tetrafluoroborate.* 694 mg of Lys(Z)NCA (2.27 mmol) was transferred under nitrogen counter flow into a predried Schlenk-tube equipped with a stir bar and again dried in high vacuum for 1 h. Then the NCA was dissolved in 6.6 mL of dry DMF. To prepare a stock solution of the initiator (always used on the day of preparation) 104 mg of neopentylammonium tetrafluoroborate was weighed in and dissolved in 10 mL of dry DMF ($c$ = 0.59 mol L$^{-1}$). 38 mL (0.023 mmol) of the initiator solution was added to the monomer solution via syringe. The solution was stirred at 40 °C and kept at a constant pressure of 1.25 bar of dry nitrogen via the Schlenk line to prevent impurities from entering the reaction vessel while allowing $CO_2$ to escape. Completion of the reaction was confirmed by IR spectroscopy (disappearance of the NCA peaks (1853 and 1786 cm$^{-1}$)). Directly after completion of the reaction the polymer was precipitated in cold ether and centrifuged (4500 rpm at 4 °C for 15 min). After discarding the liquid fraction, new ether was added and the polymer was resuspended using sonication. The suspension was centrifuged again and the procedure was repeated. The polymer was then dissolved in dioxane and lyophilized to obtain a fluffy powder (595 mg, yield 83%). For different chain lengths yields ranged from 72 to 91%. $^1$H NMR (400 MHz, DMSO-$d_6$): $\delta$ [ppm] = 8.80−7.80 (64H (1n), br, −*NH*−CO−CH−), 7.54−6.98 (358H (5n), br, −$C_6H_5$), 5.30−4.65 (115H (2n), br, −O−*CH*$_2$−$C_6H_5$), 4.35−3.60 (43H (1n), br, −*CH*−$CH_2$), 3.12−2.68 (115H (2n), br, CH−*CH*$_2$), 2.70−1.70 (332H (4n), br, $CH_2$−*CH*$_2$−*CH*$_2$−NH), 0.82 (9H (initiator), br, C(*CH*$_3$)$_3$).

The synthesis of poly(*N*-ε-benzyloxycarbonyl-L-lysine) initiated by neopentylamine was carried out in the same way at 20 °C. Since neopentylamine is a liquid, it was added using an Eppendorf pipet.

The synthesis of PLys(TFA) was analogous to the synthesis of PLys(Z). Exact details can be found in the Supporting Information.

**Deprotection of Poly(*N*-ε-benzyloxycarbonyl-L-lysine).** 143 mg of PLys(Z) was dissolved in 4.3 mL of acetic acid ($c$ = 0.033 g/mL), and 0.6 mL of HBr in acetic acid (33% w/w; 4-fold excess in regard to (Z)-group) was added. After the addition of the HBr a precipitate formed. The mixture was stirred vigorously for 90 min at room temperature. Then 1 mL of water was added which dissolved part of the precipitate. After an additional 30 min another mL of water was added which dissolved the rest of the precipitate. The solution was stirred for 1 h and extracted two times with ether to remove acetic acid. The aqueous phase was freeze-dried, yielding 86 mg of PLys. $^1$H NMR (300 MHz, $D_2O$) $\delta$ [ppm] = 4.30 (t, 37H (1n), $C_\alpha H$), 3.00 (t, 79H (2n), CH−*CH*$_2$), 1.90−1.60 (m, 146H (4n), *CH*$_2$−*CH*$_2$), 1.59−1.27 (m, 82H (2n), *CH*$_2$−$NH_2$), 0.85 (s, 9H, (*CH*$_3$)$_3$).

**GPC at Different Time Points.** GPC data at different time points of the polymerization were obtained by taking samples after the appropriate times and precipitation in ethyl ether, followed by washing and lyophilization as mentioned above.

**Fractionation of the Polymer via GPC.** 20 mg of Z1 (see Table 1) was dissolved in 0.5 mL of HFIP containing 3 g L$^{-1}$ K$^+$TFA$^-$. 100 μL of this solution was injected using the autosampler. Fractions of 0.4 mL were collected and combined to Z1_S (elution volume: ca. 19.5−21 mL) and Z1_B (elution volume: ca. 16.5−19 mL) for CD measurement. Afterward, the solvent was removed under reduced pressure for MALDI-TOF analysis.

**Table 1. PLys(Z) Polymers and Their Analytical Data from NMR, HFIP GPC, and DMF GPC**

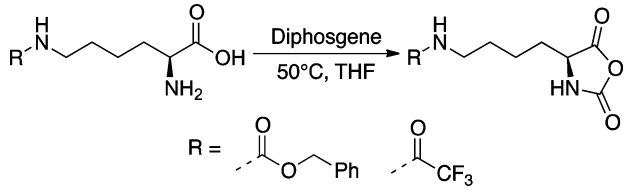
| polymer | $X_n$ (NMR) | $M_n$ (NMR) [kg/mol] | $M_n$ (HFIP GPC)[a] [kg/mol] | $Đ$ (HFIP GPC) | $M_n$ (DMF GPC)[a] [kg/mol] | $Đ$ (DMF GPC) |
|---|---|---|---|---|---|---|
| Z1 | 24 | 6.2 | 6.7 | 1.38 | 4.0 | 1.70 |
| Z2 | 57 | 15.0 | 11.5 | 1.29 | 7.5 | 1.52 |
| Z3 | 84 | 22.1 | 14.9 | 1.19 | 13.4 | 1.20 |
| Z4 | 196 | 51.3 | 20.2 | 1.15 | 16.6 | 1.15 |

[a]Relative to PMMA standards.

## RESULTS AND DISCUSSION

In this work, we have synthesized various PLys(Z) and PLys(TFA) homopolymers with different degrees of polymerization (between 20 and 200) and investigated their behavior in a variety of standard analytical methods like NMR, GPC, MALDI-TOF, and CD spectroscopy. The polymers were synthesized by ring-opening polymerization of the corresponding NCAs. All NCAs were synthesized using the Fuchs−Farthing method (Scheme 1).[33] Lys(Z) and Lys(TFA) NCAs were purified by recrystallization. Yields and melting points (Lys(Z)-NCA: mp = 99.1 °C; Lys(TFA)-NCA: mp = 101 °C) were in agreement with published literature.[2] The silver nitrate test was negative for all purified NCAs. All synthesized NCAs have been stored at −80 °C over months without any detectable degradation or oligomerization.

**Scheme 1. Synthesis of Lys(Z) and Lys(TFA) NCAs**




The ring-opening polymerization of NCAs was performed in absolute dimethylformamide (DMF) using neopentylammonium tetrafluoroborate[10] and neopentylamine as initiators (Scheme 2). The polymerizations were allowed to proceed under the exclusion of light until full conversion of the NCAs was observed by FT-IR (i.e., when the NCA associated carbonyl peaks at 1853 and 1786 cm$^{-1}$ are no longer detectable). Afterward, the polymers were purified by precipitation in ethyl ether, washed with diethyl ether, and freeze-dried from dioxane.

Tables 1 and 2 display the analytical data of the synthesized polymers. Precise $^1$H NMR analytics were enabled—even at high degrees of polymerization—by using neopentylamine as initiator, allowing integration of the 9 initiator protons (see


dx.doi.org/10.1021/ma5000392 | Macromolecules 2014, 47, 928−936

**Scheme 2. Polymerization of Lys(Z) and Lys(TFA) Using Neopentylamine**

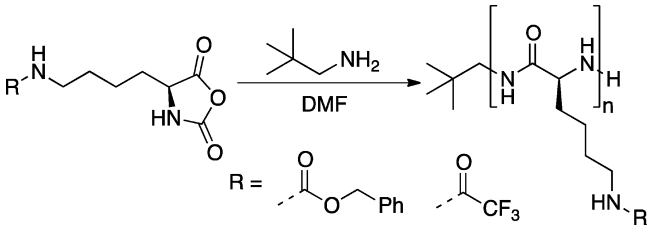



**Table 2. PLys(TFA) Polymers and Their Analytical Data from NMR (DMSO-$d_6$), HFIP GPC, and DMF GPC**

| polymer | $X_n$ (NMR) | $M_n$ (NMR) [kg/mol] | $M_n$ (HFIP GPC)[a] [kg/mol] | Đ (HFIP GPC) | $M_n$ (DMF GPC)[a] [kg/mol] | Đ (DMF GPC) |
|---|---|---|---|---|---|---|
| T1 | 20 | 4.5 | 3.4 | 1.59 | 6.2 | 1.95 |
| T2 | 65 | 14.6 | 9.5 | 1.10 | 21.9 | 1.21 |
| T3 | 90 | 20.2 | 11.9 | 1.11 | 27.9 | 1.21 |
| T4 | 143 | 32.1 | 19.0 | 1.13 | 44.9 | 1.16 |

[a]Relative to PMMA standards.

Figure 4). It was ensured that the initiator itself is well soluble at the applied precipitation conditions for the polymer. Thus, initiator peaks are polymer associated, which was additionally confirmed by DOSY (diffusion ordered spectroscopy) NMR experiments (Figure S9).

As can be seen by comparing the molecular weights from NMR and GPC in Tables 1 and 2, molecular weights are underestimated when a correction regarding weight per bond is not taken into account. After correcting the molecular weight by a factor of 1.75 (Table S1)—correcting for the differences in molecular weight per bond in the repeating unit of the polymer—the molecular weight is overestimated by GPC. This effect gets less pronounced for large polymers as already reported by Flory and co-workers[34,35] and even drops to an underestimation of a factor of 0.69 in Z4. We would like to emphasize the difference in molecular weights obtained by HFIP and DMF GPC. We are well aware that mentioning these findings seems trivial but underlines the problems of molecular weight termination by GPC of polylysine or polypeptides in general.

It can be seen in Tables 1 and 2 that the shorter polymers have a very broad distribution that—as can be seen in Figures 1 and 2—stems from a bimodal distribution. These bimodal

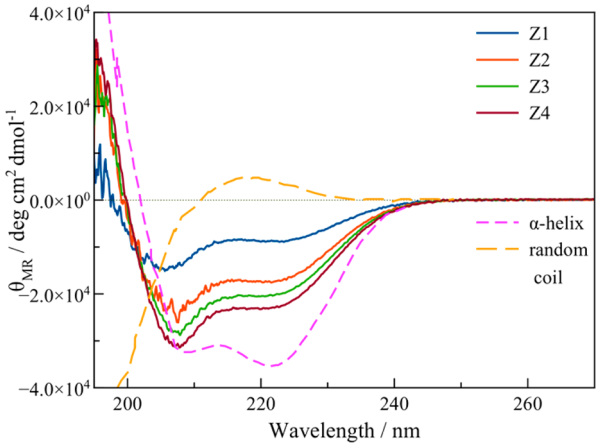



**Figure 2.** CD spectrum of PLys(Z) in HFIP, showing different secondary structures at different length.

distributions can be seen in HFIP GPC (Figure 1), DMF GPC (Figure 2), and DOSY NMR (Figure S9). The effect is strongest for Z1 (DP = 24), less pronounced in the polymer Z2 (DP = 57), and almost absent in polymer Z3 (DP = 84). The PLys(TFA) polymers show the same trend, as can be seen in Figure S1.

These bimodal distributions can emerge due to two different reasons, namely termination reactions or faster propagation than initiation of polymer chains as well as effects caused by secondary structure formation in solution. Thus, chemical as well as secondary structure associated effects have to be analyzed to attribute the observed effects to one of them. Chemically, the termination of a portion of the chains might stop their growth, while other chains continue to grow, leading to a bimodal distribution. It appears, however, that in the case of PLys(Z) and PLys(TFA) the presence of a bimodal distribution vanishes for larger polymers, which is counter-intuitive since side reactions such as activated monomer mechanism as well as urea formation should be more pronounced in higher molecular weight polymers due to elongated reaction times. In fact, when samples were taken and analyzed during different time points of the polymerization (Figure S2), the distribution undergoes a transition from monomodal to bimodal and back to monomodal. This effect was present in neopentylammonium tetrafluoroborate as well as neopentylamine initiated polymerizations. Thus, side reactions are very unlikely responsible for the bimodal distributions.

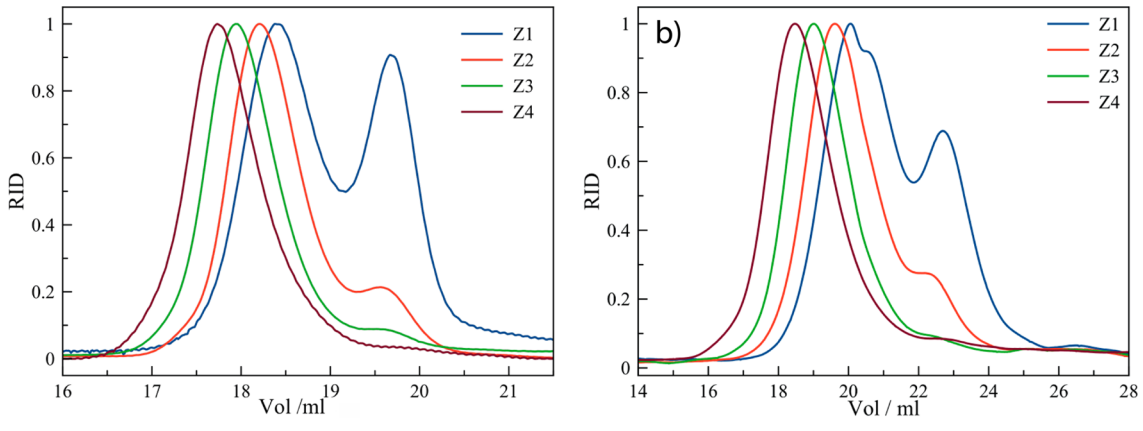



**Figure 1.** GPC of PLys(Z) at different length in HFIP (a) and DMF (b). Short polymers show bimodal distributions.


dx.doi.org/10.1021/ma5000392 | Macromolecules 2014, 47, 928−936

Interestingly, this effect is much less pronounced in the polymerization of Glu(OBn) NCA.[10]

However, it has been argued that PLys synthesized by NCA polymerization has dead chain ends (e.g., 62% at an M/I ratio of 20, due to reaction with DMF)[13] so MALDI-TOF spectrometry was used to determine the integrity of the polypeptides. In this case, termination of propagating species due to side reactions can be analyzed. Because of the mass discrimination,[36,37] it was, however, not possible to obtain MALDI-TOF spectra for the complete bimodal molecular weight distribution of PLys(Z) polymers. Thus, our smallest PLys(Z) (Z1) was separated into 12 fractions by HFIP GPC. The individual fractions could be analyzed individually (Figures S3 and S4). The fractions did not display side products, and the peaks nicely correspond to the individual PLys(Z) chains with neopentylamine initiator and amine end group (Tables S2 and S3). Within the detection limits of MALDI-TOF we could not verify dead chain ends even at higher molecular weights; however, MALDI-TOF analysis is limited in the case of PLys(Z) to a molecular weight of around 13 000 g/mol, which corresponds to a degree of polymerization of 50−60. Signal intensity for higher molecular weight polymers vanishes and thus limits the use of this methodology for PLys(Z) polymers. In addition, it limits the comparison of intensities between different fractions. In summary, MALDI-TOF analysis of the low molecular weight fraction displayed the absence of termination products within the accuracy of the analytical method. All species in the spectra are PLys(Z) polymers initiated by neopentylamine. These findings exclude termination, initiation by dimethylamine, or the presence of activated monomer mechanism as reasons for the bimodal distribution of PLys(Z). These findings were independent from the polymerization methodology used.

Physically, different secondary structures can strongly influence the hydrodynamic radius, leading in turn to a difference in elution volume, even if the polymer chains have similar length. CD spectroscopy is one of the most convenient ways to estimate the secondary structures of proteins.[19,20] We used this technique to gain insight into the variation in secondary structure with polymer length. In contrast to DMF the use of HFIP as solvent enables the correlation of GPC and CD analysis. As in GPC analysis, we observed a trend from smaller to larger polymers. As can be seen in Figure 2, the circular dichroism is stronger in larger polymers, further supporting the assumption of different secondary structures at different length. This trend is again also present in PLys(TFA) (Figure S5).

The CD spectra were measured at 20 °C to prevent solvent loss during measurement. However, it was made sure that the temperature has no influence on the secondary structure (Figure S10).

However, it has to be noted that, although it looks very similar, the CD spectrum does not fully match the $\alpha$-helix spectrum. This might be due to the protecting groups which both contain amide bonds. This itself might lead to a different CD spectrum for an $\alpha$-helix but might also induce a different kind of secondary structure that does not occur in natural peptides. Additionally, it is most likely that the polypeptide does not form a single helix but an assembly of helical units with random coil like defects as reported by Flory and co-workers.[34]

$\beta$-Sheet formation, as reported by Kricheldorf (by $^{13}$C NMR in the solid state),[25] is not very likely to occur in our case, since polymers remained soluble at all times during the polymerization.

The finding, that secondary structures play a strong role in PLys analytics, was further supported by NMR spectra in DMSO-$d_6$ and HFIP-$d_2$, which displayed two peaks for the $\alpha$-proton as well as the initiator (Figure 3, Figures S6 and S7). In

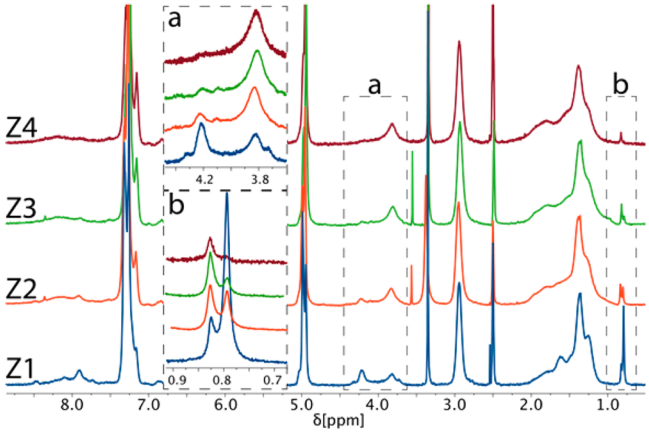



**Figure 3.** $^1$H NMR spectra of pLys(Z) with different chain length in DMSO-$d_6$. Inset a shows the $\alpha$-protons and inset b the initiator.

Figure 3, two $\alpha$-proton peaks of PLys(Z) can be seen at about 4.2 and 3.8 ppm (insert a). These chemical shifts are similar to the shifts of a lysine $\alpha$-proton in a random coil or $\alpha$-helical structure, respectively. The shifts cannot be attributed to $\beta$-sheet formation since these structures would cause a signal of the $\alpha$-proton at around 4.7 ppm.[18] Furthermore, a similar trend can also be observed for the initiator protons (inset b, Figure 3) whose chemical shift also depends on the secondary structure. As in the GPC traces, comparison to PLys(TFA) shows a similar trend with two distinct $\alpha$ and initiator protons for the short polypeptides and only one signal for the longer ones (Figure S6).

Using size exclusion chromatography, we were able to separate the two secondary structures of the short polymer (Z1 in Table 1) based on their different hydrodynamic volumes. Selected fractions are shown in Figure 4. The green fraction is especially interesting since it falls between the two GPC peaks. It can be seen that this spectrum is basically a mixture of the two adjacent fractions and is rather broad compared to the other fractions, fitting well into the hypothesis that the difference in chain length in this fraction is much larger than the difference in hydrodynamic volume. However, it has to be kept in mind that mass discrimination might distort the MALDI-TOF spectra.

All fractions belonging to one peak were combined (Z1_S, ca. 19.5−21 mL elution volume, smaller polymer; Z1_B, ca. 16.5−19 mL elution volume, bigger polymer). The MALDI-TOF analysis of the two different fractions showed that Z1_S was composed of polymers of an $X_n$ between 5 and 15, while the polymers in Z1_B had up to 60 repeating units (Figure S8).

The CD spectra (Figure 5) of both peaks clearly show that only the peak of the fraction containing the longer polymers (Z1_B) shows strong CD activity. The Z1_S fraction shows close to no CD peaks since the CD of a random coil is weaker than that of an $\alpha$-helix, and further the chain is very short, generating only a weak signal. Also, the concentration in Z1_S is slightly lower than in Z1_B.


dx.doi.org/10.1021/ma5000392 | Macromolecules 2014, 47, 928−936

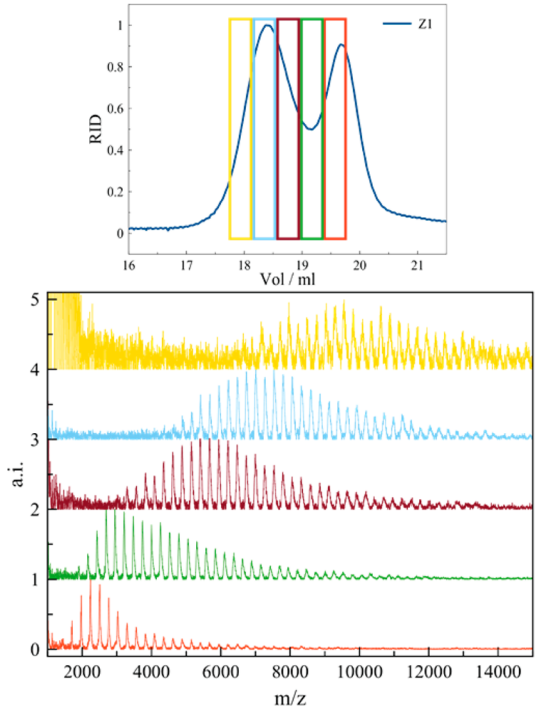



**Figure 4.** MALDI-TOF spectra of different fractions of Z1.

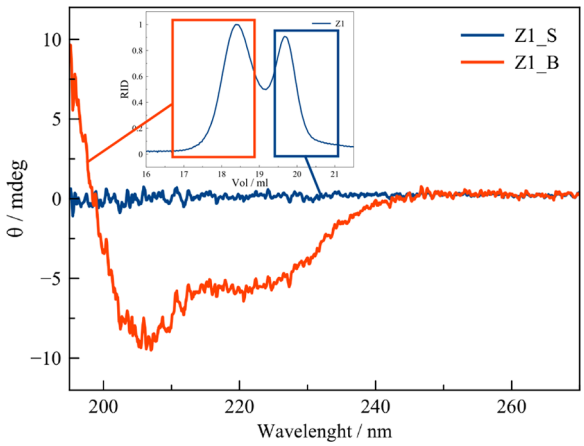


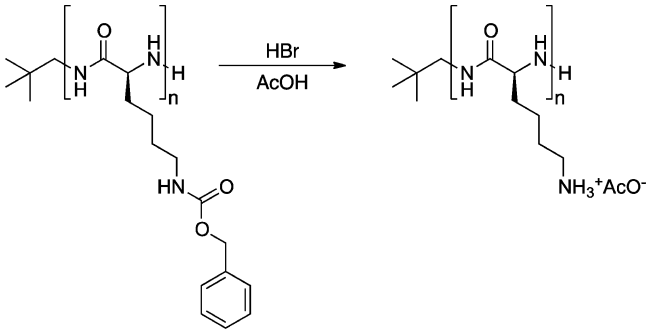

**Figure 5.** CD spectra of the two GPC peaks of Z1 indicating that only the peak corresponding to the larger polymer fraction shows significant CD activity.

Since it is well-known that unprotected PLys shows a random coil conformation at neutral pH,[38] we deprotected Z1—the polymer showing the strongest bimodal distribution—using HBr in acetic acid (see Scheme 3).[39] We expected that the deprotection should remove any effect coming from different secondary structures, thus leading to a monomodal distribution.

As shown by NMR in Figure 6, the deprotection was quantitative (in the experimental limits of NMR) and the signals from the α-helix disappeared.

Further, GPC measurements in buffer showed only a monomodal peak in the GPC elugram corresponding to a monomodal molecular weight distribution (Figure 7a) for the deprotected polylysine. CD spectroscopy finally revealed that the only secondary structure present was a random coil (Figure 7b), conclusively proving that different secondary structures are

**Scheme 3. Deprotection of PLys(Z) with HBr**




indeed responsible for bimodal distributions in short, protected polylysines.

In summary, we have demonstrated that the observed bimodal molecular weight distributions observed by DMF as well as HFIP GPC of PLys(Z) or PLys(TFA) polypeptides are not due to side reactions (termination, initiation by impurities, or activated monomer mechanism). It is much more based on differences in the hydrodynamic radii of both secondary structures. In line with our argumentation the bimodal distribution of a sample vanishes after its deprotection, when all polymers are in a random coil conformation. We would also like to mention that the above observed effects are less prominent in poly(glutamic acid) (PGA)[10] and can also be hidden by a second block, e.g. poly(ethylene glycol) or polysarcosine, in block copolymers. When the nonpeptidic block has a degree of polymerization above 150, a bimodal distribution is hardly detectable.

**Estimates of the Size Difference of α-Helix and Random Coil.** In the last part of this work we would like to discuss a possible explanation for the bimodal character of the GPC elugrams in detail. The theoretical model presented here is based on assumptions, which are most likely but cannot be proven by the authors. Thus, the presented model can only serve for a qualitative discussion of the size change that is difficult or even impossible to determine for polydispers sample experimentally. However, we feel that this model provides a deeper understanding of the reported effects.

We have already demonstrated that PLys(Z) and PLys(TFA) exist in different superstructures (random coil and α-helix) depending on the degree of polymerization. For P(Lys) in a random coil conformation the worm-like chain (WLC) model is a suitable model. To estimate the hydrodynamic radius of a random coil of PLys(Z) with a DP of 15, the model by Yamakawa and Fujii[40] as corrected by Dorfman et al. was selected.[41]

The length of the chain is estimated to be $L$ = 5.6 nm (15 × 374 pm) with a Kuhn length of $l_k$ = 2 nm and a diameter of $d$ = 2.94 nm. This leads to an approximate radius, which is calculated (see Supporting Information) to be

$$R_h^{WLC} = 1.7 \text{ nm}$$

If we assume an α-helical conformation, with a pitch of 5.4 Å incorporating 3.6 residues[42] and a DP of 15, we have a helix with a length of 2.25 nm spanning 4.2 turns.

The approximate length of the extended side chain can be calculated by assuming a median bond length of $l$ = 150 pm and an angle between bonds of Θ = 110°. The effective length of the side chain ($n$ = 12 bonds) is thus


dx.doi.org/10.1021/ma5000392 | Macromolecules 2014, 47, 928−936

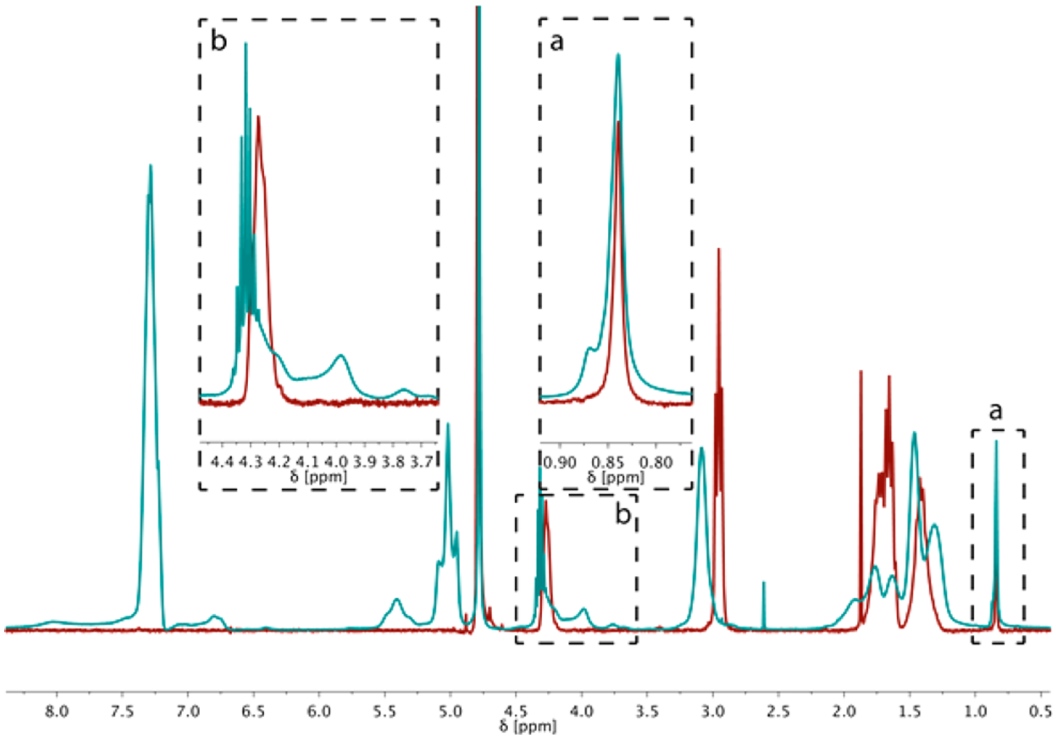



**Figure 6.** NMR spectrum of PLys before (Z1, green, HFIP-$d_2$) and after (L1, red, $D_2O$) deprotection, showing that only one secondary structure is present after deprotection.

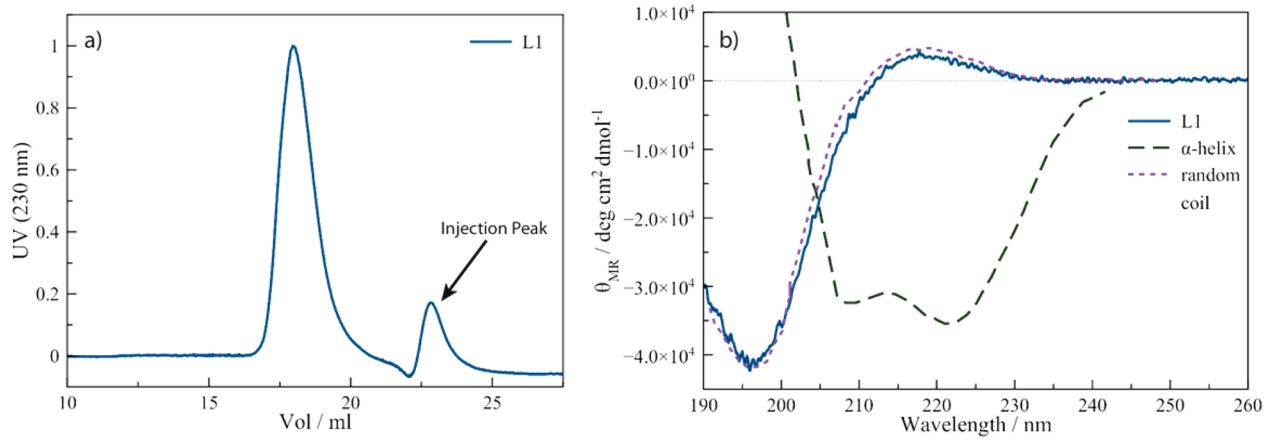



**Figure 7.** (a) GPC trace of L1 ($PLys_{25}$) after deprotection, showing monomodal distribution. (b) CD spectrum of PLys after deprotection in water, showing random coil conformation.

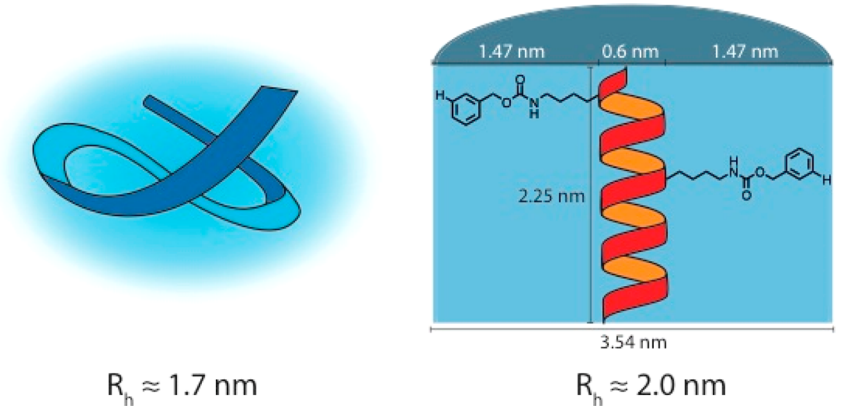



**Figure 8.** Hydrodynamic radii of Lys(Z) with DP = 15 in random coil and $\alpha$-helical conformation.


dx.doi.org/10.1021/ma5000392 | *Macromolecules* 2014, 47, 928−936

$$l_{\text{eff}} = nl \sin(\theta/2) = 1.47 \text{ nm}$$

Since the diameter of an $\alpha$-helix without side chains is approximately 0.6 nm, the diameter of the whole helix can be approximated to be 3.54 nm (see Figure 8).

This leads to a cylinder with a length of $2a$ = 2.25 nm and a diameter of $2b$ = 3.54 nm (see Figure 8). The equivalent radius (radius of a sphere with the same volume) of a cylinder is given by

$$R_{\text{eq}}^{\text{cy}} = \sqrt[3]{\frac{3}{4}\left(\frac{a}{b}\right)^2}\, a = 3.48 \text{ nm}$$

leading to a hydrodynamic radius of

$$R_{\text{h}}^{\text{cy}} = R_{\text{eq}}^{\text{cy}} \frac{f}{f_0} = 2.0 \text{ nm}$$

where $f$ is the translational friction coefficient of the cylinder and $f_0$ is the translational friction coefficient of a sphere having the same volume. An equation for $f/f_0$ was modeled by Ortega et al. (see Supporting Information).[43] Since the volume is proportional to $R_h^3$, the volume ratios would be 1:1.6 for worm-like chain and cylinder, respectively.

To obtain a random coil with a hydrodynamic radius of 2.0 nm (the $R_h$ of an $\alpha$-helix with DP = 15), the PLys(Z) chain would need to have a degree of polymerization of 30. These rough estimates show that the transition from coil to helix appears in the GPC as if the molecular weight of the polymer would suddenly double. This model can qualitativly explain the bimodal GPC elugram by the presence of two superstructures that differ significantly in the individual hydrodynamic volumes.

## CONCLUSIONS

Surprisingly, a change in secondary structure—although well-known—is seldom discussed in context with *standard analytics* of polypeptides made by NCA polymerization. The most common method for determining the molecular weight of polymers is size exclusion chromatography, which is certainly simple to perform and gives good estimates of polymer definition and molecular weights. It has to be kept in mind that molecular weight determination by GPC depends on the correlation of hydrodynamic volumes between sample and standard, which is greatly influenced by secondary structure. Because of their ability to form secondary structures via hydrogen bonds, polypeptides are often not in a random coil conformation or even one distinct conformation.

In this work we have demonstrated that secondary structure formation needs to be carefully considered when performing analysis on polypeptides. It has been shown that for PLys(Z) the helix−coil transition occurs around a degree of polymerization of 15 but does not seem to be sharp. More importantly, standard GPC analysis tends to overestimate polymer dispersity tremendously whenever the sample consists of polymers differing in conformation, like random coil or helical (rod-like) structure.

It should also be noted that the difference in secondary structures is in no way problematic for the application of polypeptides. When going to larger polypeptides, an all-$\alpha$-helical structure is observed, and even for smaller polypeptides the difference in secondary structures is lost after deprotection.

## ASSOCIATED CONTENT

### Supporting Information

GPC, CD, and NMR and experimental data of PLys(TFA), additional MALDI-TOF spectra, an NMR spectrum in HFIP, DOSY spectrum, GPC traces at different time points of the polymerization, and corrected molecular weights of PLys(Z) as well as equations for the $R_h$ estimations. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: Barz@uni-mainz.de (M.B.).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors are indebted to Prof. M. Schmidt for very fruitful discussions about the dimensions of different superstructures in solution. M.B. acknowledges financial support by the SFB 1066 and the NMFZ Mainz. D.H. acknowledges the support by the "Verband der Chemischen Industrie" (VCI) and the "Max Planck Graduate Center with the Johannes Gutenberg-Universität Mainz" (MPGC).

## ABBREVIATIONS

AMM, activated monomer mechanism; CD, circular dichroism; DMF, dimethylformamide; DMSO, dimethyl sulfoxide; DOSY, diffusion ordered spectroscopy; DP, degree of polymerization; GPC, gel permeation chromatography; HFIP, hexafluoroisopropanol; NCA, $\alpha$-amino acid-*N*-carboxyanhydride; PEG, poly(ethylene glycol); PMMA, poly(methyl methacrylate); PLys, poly(L-lysine); PLys(TFA), poly(*N*-$\varepsilon$-trifluoracetamide-L-lysine); PLys(Z), poly(*N*-$\varepsilon$-benzyloxycarbonyl-L-lysine); ROP, ring-opening polymerization; THF, tetrahydrofuran; WLC, worm-like chain.

## REFERENCES

(1) Leuchs, H. *Ber. Dtsch. Chem. Ges.* **1906**, *39*, 857−861.
(2) Kricheldorf, H. R. *α-Amino Acid-N-Carboxyanhydrides and Related Heterocycles: Syntheses, Properties, Peptide Synthesis, Polymerization*; Springer: Berlin, 1987.
(3) Kricheldorf, H. R. *Angew. Chem., Int. Ed.* **2006**, *45*, 5752−5784.
(4) Deming, T. J. *Nature* **1997**, *390*, 386−389.
(5) Aliferis, T.; Iatrou, H.; Hadjichristidis, N. *Biomacromolecules* **2004**, *5*, 1653−1656.
(6) Dimitrov, I.; Schlaad, H. *Chem. Commun.* **2003**, 2944−2945.
(7) Vayaboury, W.; Giani, O.; Cottet, H.; Deratani, A.; Schué, F. *Macromol. Rapid Commun.* **2004**, *25*, 1221−1224.
(8) Lu, H.; Cheng, J. *J. Am. Chem. Soc.* **2007**, *129*, 14114−14115.
(9) Peng, H.; Ling, J.; Shen, Z. *J. Polym. Sci., Part A: Polym. Chem.* **2012**, *50*, 1076−1085.
(10) Conejos-Sánchez, I.; Duro-Castano, A.; Birke, A.; Barz, M.; Vicent, M. J. *Polym. Chem.* **2013**, *4*, 3182−3186.
(11) Zou, J.; Fan, J.; He, X.; Zhang, S.; Wang, H.; Wooley, K. L. *Macromolecules* **2013**, *46*, 4223−4226.
(12) Hadjichristidis, N.; Iatrou, H.; Pitsikalis, M.; Sakellariou, G. *Chem. Rev.* **2009**, *109*, 5528−5578.
(13) Vayaboury, W.; Giani, O.; Cottet, H.; Bonaric, S.; Schué, F. *Macromol. Chem. Phys.* **2008**, *209*, 1628−1637.
(14) Habraken, G. J. M.; Wilsens, K. H. R. M.; Koning, C. E.; Heise, A. *Polym. Chem.* **2011**, *2*, 1322−1330.
(15) Vayaboury, W.; Giani, O.; Cottet, H.; Deratani, A.; Schué, F. *Macromol. Rapid Commun.* **2004**, *25*, 1221−1224.
(16) Doyle, D. A. *Science* **1998**, *280*, 69−77.

(17) Wüthrich, K. *Nat. Struct. Mol. Biol.* **2001**, *8*, 923−925.
(18) Wishart, D. S. *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *58*, 62−87.
(19) Pelton, J. T.; McLean, L. R. *Anal. Biochem.* **2000**, *277*, 167−176.
(20) Whitmore, L.; Wallace, B. A. *Biopolymers* **2008**, *89*, 392−400.
(21) Surewicz, W. K.; Mantsch, H. H.; Chapman, D. *Biochemistry* **1993**, *32*, 389−394.
(22) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Mol.* **1995**, *30*, 95−120.
(23) Goodman, M.; Verdini, A. S.; Toniolo, C.; Phillips, W. D.; Bovey, F. A. *Proc. Natl. Acad. Sci. U. S. A.* **1969**, *64*, 444−450.
(24) Müller, D.; Kricheldorf, H. *Polym. Bull.* **1981**, *6*, 101−108.
(25) Kricheldorf, H.; Müller, D. *Macromolecules* **1983**, *16*, 615−623.
(26) Vicent Docon, M. J.; Barz, M.; Canal, F.; Conejos Sanchez, I.; Duro Castano, A. Sintesis Controlada De Poliglutamatos Con Baja Polidispersidad Y Arquitecturas Versátiles, P201131713, 2011.
(27) Mazia, D. *J. Cell Biol.* **1975**, *66*, 198−200.
(28) Christie, R. J.; Nishiyama, N.; Kataoka, K. *Endocrinology* **2010**, *151*, 466−473.
(29) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512−7515.
(30) Strohalm, M.; Hassman, M.; Kosata, B.; Kodícek, M. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 905−908.
(31) Strohalm, M.; Kavan, D.; Novák, P.; Volný, M.; Havlícek, V. *Anal. Chem.* **2010**, *82*, 4648−4651.
(32) Niedermeyer, T. H. J.; Strohalm, M. *PLoS One* **2012**, *7*, e44913.
(33) Oya, M.; Katakai, R.; Nakai, H.; Iwakura, Y. *Chem. Lett.* **1973**, 1143−1144.
(34) Miller, W. G.; Flory, P. J. *J. Mol. Biol.* **1966**, *15*, 298−314.
(35) Miller, W. G.; Brant, D. A.; Flory, P. J. *J. Mol. Biol.* **1967**, *23*, 67−80.
(36) Schriemer, D. C.; Li, L. *Anal. Chem.* **1997**, *69*, 4176−4183.
(37) Schriemer, D. C.; Li, L. *Anal. Chem.* **1997**, *69*, 4169−4175.
(38) Tiffany, M. L.; Krimm, S. *Biopolymers* **1969**, *8*, 347−359.
(39) Ben-Ishai, D.; Berger, A. *J. Org. Chem.* **1968**, *17*, 1564−1570.
(40) Yamakawa, H.; Fujii, M. *Macromolecules* **1973**, *6*, 407−415.
(41) Tree, D. R.; Muralidhar, A.; Doyle, P. S.; Dorfman, K. D. *Macromolecules* **2013**, *46*, 8369−8382.
(42) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 205−211.
(43) Ortega, A.; García de la Torre, J. *J. Chem. Phys.* **2003**, *119*, 9914−9919.