

Imperial College London

Department of Chemical Engineering

**Synthesis and functionalisation of defined molecular weight poly(ethylene glycol) and its application as a new class of PEGylation agents**

Submitted by

Maria Johanna Burggraef

Supervised by Professor Andrew G. Livingston

December 2023

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in  
Chemical Engineering from Imperial College London

## Abstract

The polymer poly(ethylene glycol) (PEG) is a highly hydrophilic polymer which is widely used in the pharmaceutical industry to mask therapeutics and that way increase their blood-residence time and decrease their immunogenicity.<sup>1</sup> The technique of covalently binding one or more PEG units to a therapeutic is referred to as PEGylation and as of 2023, there were 40 FDA or EMA-approved PEGylated therapeutics on the market, most of which were proteins. The mass of the used PEG ranges from <1 kDa to 60 kDa and it is conventionally synthesised through the anionic polymerisation of ethylene oxide <sup>2</sup>, which results in a Gaussian distribution of PEG of different chain lengths. This inherent dispersity creates problems with the characterisation of the PEGylation agents, as well as the PEGylated products, resulting in unidentified impurities or batch-to-batch variations. The disperse nature of the PEG makes it difficult to locate the exact PEGylation sites but identifying them is valuable for predicting potential interactions of the PEG with the active site of a therapeutic.

To overcome these problems, a new class of defined molecular weight ( $M_w$ ) PEGylation agents has been developed in this work. A previously established liquid-phase approach based on Nanostar Sieving was further developed and used to synthesise defined, linear PEGs up to 5 kDa. These were then chemically modified to obtain a library of defined  $M_w$ , active PEGylation agents. A methoxy-PEG active ester was used to perform first PEGylation reactions on the protein bovine serum albumin (BSA). A comparative study regarding the purification and characterisation of mono-PEG BSA was performed between the synthesised defined PEG derivative and a commercially available disperse comparator. It was demonstrated that defined PEGylation agents not only allow for easy monitoring of chemical modification reactions, but also facilitate the straight-forward identification of the PEGylated sites within the protein via a simple Peptide mapping approach using UPLC-MS.

## Declaration of Originality

I hereby declare that the work presented in this thesis is my own. Information derived from the published and unpublished work of others has been acknowledged in the text and the relevant references are included in this thesis.

Maria Burggraef

December 2023

## Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non-Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Maria Burggraef  
December 2023

## Acknowledgements

First, I would like to thank my PhD supervisor Professor Andrew Livingston for his patient supervision. He helped me a lot with his structured, analytical, and rational approach, but also offered support in difficult times ("A PhD is a marathon, not a sprint"). I'm especially grateful that Andrew always, without a single miss, selected the next friendly and kind member for the research group!

I also would like to thank Dr. Piers Gaffney for training me in the lab at the early stages of my PhD. Next I have to thank the people who supported me in the lab, helped me getting through my experiments, showed me patiently how to use a spanner or read through this thesis- thank you everyone!

Special thanks to the friends I met on the way, in and outside of work- the Livingston group, starting with day one at Imperial. It never got boring with you guys, and it made lab-life a lot easier. It was beautiful to share these chapters of our lives with each other and witness you guys getting drunk, getting married, buying houses, finishing PhDs and so on.

Special thanks to all the people that made my life in London so fun, exciting, and interesting, made me get over the occasional home sickness and eventually made me enjoy the city. And thank you, London, for being home to all these interesting, strange, and kind people. And for the music and the food and the movie-like setting.

I also want to thank my friends back in Germany, who loyally kept coming to the Christmas drinks at home every year and helped me staying connected to the "Heimat".

Most importantly, I want to thank my family- Mama, Ricarda, Papa, Levent, Leni, Toni, Eli und Rapha- ohne euch haette ich das niemals geschafft. Danke für die konstante Unterstützung (sei es über Telefon oder live), für die zig Besuche (vor allem Toni und Leni- die vielen gemeinsamen Tage in London werde ich nie vergessen) und die Geduld, die ihr auch in anstrengenden Zeiten mit mir hattet.

To my siblings (the “cult”), I’m incredibly grateful, that we manage to stay so connected, despite living quite far from each other (for our standards). I’m sure that we will all go different, exciting ways in life, and that we will always find ways to share it and be there for each other.

This thesis is dedicated to all of you, I love you!

Maria

# Contents

|                                                                              |    |
|------------------------------------------------------------------------------|----|
| Copyright Declaration.....                                                   | 4  |
| Acknowledgements .....                                                       | 5  |
| Publications and Conferences .....                                           | 10 |
| List of Tables.....                                                          | 11 |
| List of Figures.....                                                         | 12 |
| Acronyms and Abbreviations .....                                             | 22 |
| Chapter 1 Introduction.....                                                  | 24 |
| Chapter 2 Literature Review.....                                             | 27 |
| 2.1 Poly(ethylene glycol) and PEGylation.....                                | 27 |
| 2.1.1 Introduction .....                                                     | 27 |
| 2.1.2 Historical Development and PEGylation Chemistry.....                   | 29 |
| 2.1.3 Pharmacological Behaviour of PEGylated Proteins .....                  | 33 |
| 2.1.3.1 Branched vs. linear PEGylation agents .....                          | 35 |
| 2.1.3.2 Toxicity and Immunogenicity .....                                    | 37 |
| 2.1.3.3 Dispersity of PEGylation agents.....                                 | 38 |
| 2.1.4 Purification and Characterisation of PEGylated Proteins .....          | 40 |
| 2.1.4.1 Introduction .....                                                   | 40 |
| 2.1.4.2 Separation and analysis of PEGamers .....                            | 42 |
| 2.1.4.3 Separation and identification of positional isomers .....            | 46 |
| 2.1.4.4 An example- the FDA report of the Neulasta biosimilar MYL-1401H..... | 50 |
| 2.2 Synthesis of defined PEG and PEGylation agents .....                     | 54 |
| 2.2.1 Introduction .....                                                     | 54 |
| 2.2.2 Dispersity .....                                                       | 55 |
| 2.2.3 Synthesis of defined polymers.....                                     | 60 |
| 2.2.4 Early approaches for the synthesis of defined $M_w$ PEG .....          | 61 |
| 2.2.5 Novel approaches for synthesising defined $M_w$ PEGs .....             | 64 |
| 2.2.6 Organic Solvent Nanofiltration.....                                    | 67 |
| 2.2.7 Nanostar Sieving.....                                                  | 74 |
| Chapter 3 Research Objectives .....                                          | 77 |
| Chapter 4 Synthesis of defined $M_w$ PEG using Nanostar Sieving .....        | 80 |

|                                                                                                             |     |
|-------------------------------------------------------------------------------------------------------------|-----|
| 4.1 Introduction .....                                                                                      | 81  |
| 4.2 Materials and Methods .....                                                                             | 83  |
| 4.2.1 Materials .....                                                                                       | 83  |
| 4.2.2 Methods .....                                                                                         | 84  |
| 4.3 Results and Discussion .....                                                                            | 96  |
| 4.3.1 Synthesis of defined $M_w$ 5 kDa PEG based on a $Eg_{12}$ Building Block .....                        | 96  |
| 4.3.2 Moving from a $Eg_{12}$ to a $Eg_{28}$ Building Block .....                                           | 101 |
| 4.3.2.1 Analytical Considerations and Optimisation .....                                                    | 101 |
| 4.3.2.2 Chain Extension Reaction Kinetics .....                                                             | 106 |
| 4.3.2.3 Critical Investigation of Impurity Formation .....                                                  | 108 |
| 4.3.3 Defined $M_w$ PEG synthesis <i>via</i> Chromatography vs. Nanostar Sieving- a comparative study ..... | 115 |
| 4.3.3.1 Membrane selection and synthesis .....                                                              | 115 |
| 4.3.3.2 Comparison to a chromatography-based synthesis .....                                                | 117 |
| 4.4 Conclusions .....                                                                                       | 119 |
| Chapter 5 Functionalisation of defined $M_w$ PEGs as a third generation PEGylation agents .....             | 122 |
| 5.1 Introduction .....                                                                                      | 123 |
| 5.2 Materials and Methods .....                                                                             | 124 |
| 5.2.1 Materials .....                                                                                       | 124 |
| 5.2.2 Methods .....                                                                                         | 124 |
| 5.2.2.1 Monofunctional PEGylation agents .....                                                              | 125 |
| 5.2.2.2 Heterobifunctional PEGylation agents .....                                                          | 138 |
| 5.3 Results and Discussion .....                                                                            | 142 |
| 5.3.1 Monofunctional PEGylation agents .....                                                                | 142 |
| 5.3.1.1 Carboxylic Acids, Active Esters and Aldehydes .....                                                 | 143 |
| 5.3.1.2 Other functionalities .....                                                                         | 148 |
| 5.3.1 Heterobifunctional PEGylation agents .....                                                            | 149 |
| 5.4 Conclusions .....                                                                                       | 150 |
| Chapter 6 The PEGylation of Bovine Serum Albumin – a comparative study .....                                | 152 |
| 6.1 Introduction .....                                                                                      | 153 |
| 6.2 Materials and Methods .....                                                                             | 156 |
| 6.2.1 Materials .....                                                                                       | 156 |
| 6.2.2 Methods .....                                                                                         | 156 |

|                                                                                  |     |
|----------------------------------------------------------------------------------|-----|
| 6.3 Results and Discussion .....                                                 | 157 |
| 6.3.1 PEGylation and Peptide Mapping of BSA using mPEG <sub>5163</sub> -SS ..... | 157 |
| 6.3.1 PEGylation and Peptide Mapping of BSA using mPEG <sub>5135</sub> -SP ..... | 161 |
| 6.4 Conclusions.....                                                             | 163 |
| Chapter 7 Conclusions and Future Work.....                                       | 165 |
| 7.1 Conclusions.....                                                             | 165 |
| 7.2 Future work.....                                                             | 167 |
| 7.2.1 Reaction Development.....                                                  | 167 |
| 7.2.2 Process Optimisation.....                                                  | 167 |
| 7.2.3 Higher M <sub>w</sub> PEGylation agents.....                               | 168 |
| 7.2.4 Defined M <sub>w</sub> therapeutics .....                                  | 169 |
| Chapter 8 General Methods.....                                                   | 170 |
| References .....                                                                 | 172 |
| Appendix .....                                                                   | 189 |

## Publications and Conferences

ChemEng Day 2019, Edinburgh; “Iterative Chromatography-free Synthesis of Unimolecular Poly(ethylene glycol)” (Co-presentation with Adam Oxley).

2. ADC Conference 2020, London, Poster “Nanostar Sieving Produces Precise, Unimolecular PEGs for PEGylation” (Ruijiao Dong, Maria Burggraef, Adam Oxley, Naveed A. Zaidi, Piers R. J. Gaffney, Anna Li Volsi, Andrew G. Livingston\*).

Maria Burggraef, Adam Oxley, Naveed Zaidi, Pedro Cutillas, Piers R. J. Gaffney and Andrew G. Livingston, “Exactly defined molecular weight poly(ethylene glycol) allows for facile identification of PEGylation sites on proteins”, *Manuscript prepared, currently hold back due to patent filing process.*

## List of Tables

|                                                                                                                                                      |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 1. List of FDA, or EMA-approved PEGylated therapeutics (as of 20/11/2023).....                                                                 | 28  |
| Table 2. Overview of recently published approaches for synthesising defined $M_w$ PEG (since 2014). .....                                            | 65  |
| Table 3. Amounts of nanostar-PEG-alcohol (5a-d) used and the corresponding yields for membrane-based synthesis. ....                                 | 92  |
| Table 4. Amounts of nanostar-PEG-alcohol (9a-d) used and corresponding yields for chromatography-based synthesis. ....                               | 92  |
| Table 5. Overview of reaction times for chain extensions in reactors comprising different materials.....                                             | 100 |
| Table 6. Overview of results on the reaction completion of chain extension with different amounts of BB used. ....                                   | 108 |
| Table 7. Summary of the three detected species after treatment of DmtrO-Eg <sub>28</sub> -OH or Hub(Eg <sub>28</sub> -OH) <sub>3</sub> with NaH..... | 113 |

## List of Figures

|                                                                                                                                                                                                                                                                                                                                                                                                                             |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1. Activation of mPEG-OH with cyanuric chloride and coupling to amino group of a protein. Method developed by ABUCHOWSKI <i>et al.</i> in 1977. Reproduced with permission from [20]. .....                                                                                                                                                                                                                          | 29 |
| Figure 2. Reactions of mPEG-NHS-derivatives with primary amines. Top: Reaction of NHS-carbonate with amine to form a carbamate/urethane linkage. Bottom: Reaction of NHS-ester with amine to form an amide linkage; X= (CH <sub>2</sub> ) <sub>n</sub> . Reproduced with permission from [24]. .....                                                                                                                        | 30 |
| Figure 3. Different first-generation PEGylation reagents. a, based on cyanuric chloride. b, based on succinimidyl succinate. c, NHS ester (left) and NHS carbonate (right).....                                                                                                                                                                                                                                             | 31 |
| Figure 4. Structure of the branched PEG-NHS PEGylation agent based on Lysine. Developed by MONFARDINI <i>et al.</i> [36]. .....                                                                                                                                                                                                                                                                                             | 32 |
| Figure 5. Schematic PEGylated drug derived from acylation (top) and alkylation (bottom) with a methoxy-PEG agent. Reproduced from [24]. .....                                                                                                                                                                                                                                                                               | 32 |
| Figure 6. <i>In vivo</i> activity and <i>in vitro</i> activity of a PEGylated therapeutic dependent on the PEG molecular weight. The cell culture assay <sup>3</sup> H-thymidine uptake is a measure for the <i>in vitro</i> activity and decreases with rising PEG molecular weight, whereas the <i>in vivo</i> cell proliferation increases with increasing PEG molecular weight. Copied with permission from [26]. ..... | 34 |
| Figure 7. Schematic depiction of a therapeutic PEGylated with linear or branched PEG. The “umbrella” like structure of the branched version is hindering other molecules from approaching the drug. Reproduced with permission from [44]. .....                                                                                                                                                                             | 36 |
| Figure 8. Correlation of the molecular weight of linear, disperse PEG and the corresponding half life in blood after intravenous administration into mice. Copied with permission from [44]. .....                                                                                                                                                                                                                          | 40 |
| Figure 9. Overview of the steps involved in the purification and characterisation of PEGylated proteins. This involves 1: the separation of the crude PEGylation mixture to separate the PEGamers (identification through MALDI-TOF), 2: the separation of the different PEGamers through for example, size-exclusion or ion exchange chromatography, 3: the separation of each                                             |    |

PEGamer into the different positional isomers (usually through IEC or RPC) and the identification of each positional isomer through one of three different approaches described in chapter 2.1.4.3.

..... 41

Figure 10. Overview of the different approaches of identifying the positional isomers of a PEGylated protein. .... 47

Figure 11. Analytical data from product review report for *Neulasta* biosimilar MYL 1401H. a, exemplary MALDI-TOF spectrum of MYL-1401H. b, Scatter plot of MALDI-TOF results of different samples of MYL1401H, U.S.-licensed *Neulasta* and EU-approved *Neulasta*. Copied with permission from [55]. .... 51

Figure 12. Results of the Glu-C digestion of a sample of MYL-1401H and the U.S.-approved *Neulasta*. a, UV trace of the UPLC chromatogram (bottom: *Neulasta* US, top MYL-1401H). The peaks at 32.3 min and 32.4 min were identified as PEG species. b, Underlying ESI traces for the same chromatograms. The ESI (right side) shows the characteristic PEG pattern with a difference of 44 kDa between each peak. Copied with permission from [55]. .... 53

Figure 13. Mechanism of anionic polymerisation of ethylene oxide to synthesise poly(ethylene glycol) (PEG). Ethylene oxide undergoes nucleophilic attack by an alkoxide, which results in ring opening and bond formation between the alkoxide oxygen and the ethylene oxide -CH<sub>2</sub>- group. .... 54

Figure 14. Simulated molecular weight distributions for a theoretical polymer with a median molecular weight of 27 kDa. a, A theoretical normal distribution of molecular weights as it is usually derived from conventional polymerisation reactions (such as the anionic polymerisation of ethylene oxide) with a dispersity index of 1.01; b, The theoretical molecular weight distribution for an ideal single mass polymer with an exact mass of 27 kDa and a dispersity index of 1.0; c, Two theoretical, symmetrical distributions for polymers derived from iterative processes, with the same width. The distribution with the lower average M<sub>w</sub> (blue, D = 1.00036) has a higher dispersity index than the distribution with the higher average M<sub>w</sub> (red, D = 1.00034). .... 58

|                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 15. a, schematic depiction of a defined sequence oligomer or polymer, consisting of different repeating units. b, schematic depiction of a defined weight oligomer or polymer, consisting solely of the repeating unit “A”. c, example for a synthetic defined weight oligomer (octa(ethylene glycol)). d, the peptide hormone Oxytocin, an example of a naturally occurring peptide with a defined amino acid sequence.....             | 60 |
| Figure 16. Different strategies to achieve sequence-defined polymers. Each circle stands for one building block unit. a, chain doubling; b, bidirectional iterative growth; c, unidirectional iterative growth; not depicted: chain tripling; d, an example for the chain doubling approach; e, an example for the bidirectional growth method; f, an example for the unidirectional growth method. Reproduced with permission from [13]. ..... | 62 |
| Figure 17. Overview of membrane filtration processes and their classification by operating pressure $P$ , the molecular weight of the rejected species and the pore size $r_p$ . Adapted from [131]. .....                                                                                                                                                                                                                                      | 68 |
| Figure 18. Simulated rejections of an uncharged solute by NF membranes with different pore sizes, dependent on the $M_w$ or radius of the solute. The ideal rejection curve is represented by the blue line. Copied with permission from [136].....                                                                                                                                                                                             | 70 |
| Figure 19. The reaction steps involved towards the fabrication of DBX-crosslinked PBI, modified with PPG/PEG polymers for increased anti-fouling behaviour. Adapted with permission from [135].....                                                                                                                                                                                                                                             | 73 |
| Figure 20. a, overview of a 2-stage diafiltration system (R1: retentate first stage, R2: retentate second stage, C1-C5: membrane filtration cells 1-5, P1-P5: permeate lines from cells 1-5, V1+2: pressure valves); b, modelled yields dependent on the rejections for a single stage and two stage DF system. Copied with permission from [147].....                                                                                          | 74 |
| Figure 21. General overview of Nanostar Sieving process to synthesise DSPs and DMPs. Each cycle consists of chain extension (i), deprotection (ii) and diafiltration (iii). The cycle is repeated to achieve DSPs or DMPs with the desired sequence/ chain length.....                                                                                                                                                                          | 75 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 22. Reaction scheme of the reaction of the nanostar-aldehyde ( $\text{Hub}(\text{CHO})_3$ ) to the nanostar-alcohol ( $\text{Hub}(\text{C-OH})_3$ ) and the nanostar-tribromide $\text{HubBr}_3$ (1).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 84 |
| Figure 23. Reaction scheme of the synthesis of the building blocks $\text{DmtrO-Eg}_{12}\text{-OTs}$ (2) and $\text{DmtrO-Eg}_{28}\text{-OTs}$ (3).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 85 |
| Figure 24. Reaction scheme of the synthesis of $\text{Hub}(\text{Eg}_{28}\text{-OH})_3$ (5a).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 88 |
| Figure 25. Reaction scheme of the chain extension and deprotection reactions using the $\text{Eg}_{12}$ building block (2). .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 89 |
| Figure 26. Reaction scheme of the chain extension and deprotection reactions using the $\text{Eg}_{28}$ building block (3). .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 90 |
| Figure 27. Schematic representation of the 2-stage diafiltration rig used for this work. C1-5: Membrane cells, each equipped with a flat sheet of DBX-crosslinked PBI membrane. R1/R2: Retentate lines. P1-5: Permeate lines. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 93 |
| Figure 28. UPLC-ELSD signal calibration curve for $\text{Eg}_{28}$ Diol.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 95 |
| Figure 29. Nanostar Sieving based synthesis of defined $\text{mPEG-OH}_{4966}$ . a, chain elongation cycle consisting of the coupling of the $\text{DmtrO-Eg}_{12}\text{-OTs}$ to the free nanostar-PEG alcohol (chain extension, i), removal of the protecting group (deprotection, ii) and diafiltration (iii). Repeating this cycle seven times results in the nanostar- $\text{Eg}_{112}\text{-OH}$ species. b, the terminal hydroxyls are then methylated (vi) and subsequently the PEG chains are cleaved from the nanostar (vii) to obtain defined $\text{mPEG-OH}$ with a molecular weight of 4966 Da.....                                                                         | 97 |
| Figure 30. Example of chain extension monitoring for the first chain extension from $\text{Eg}_{28}$ to $\text{Eg}_{40}$ . a, from bottom to top: UPLC-UV traces (full range) of the chain extension at 0 min, 30 min, 60 min, 20 hrs; the peaks can be identified through the underlying $\text{ESI}^+$ trace as the BB (blue), 2-arm species (yellow), the product $\text{Hub}(\text{Eg}_{40}\text{ODmtr})_3$ (green) and BB dimer (red); the top trace is of the deprotected product ( $\text{Hub}(\text{Eg}_{40}\text{OH})_3$ ) after membrane-filtration. b, the $\text{ESI}^+$ traces corresponding to the BB (blue), the dimer (red) and the chain extension product (green). ..... | 98 |

Figure 31. Overview of nanostar sieving-based synthesis of defined  $M_w$  PEG 5 kDa ( $Eg_{112}$ ) using an  $Eg_{12}$  building block. a, the reaction scheme for the chain extension and deprotection cycle; b, summary of the three  $Eg_{12}$  based syntheses performed within this work, including the flask material, final product amounts and overall yield for the different batches; c, photograph of the final  $Hub(Eg_{112}-OH)_3$  material (derived from batch 3); d, MALDI-TOF spectrum of  $MeO-Eg_{112}-OH$  ( $mPEG_{4966}-OH$ ) derived from batch 3; e, MALDI-TOF spectrum of commercially available disperse 5 kDa  $mPEG-OH$ ..... 99

Figure 32. Comparison of high- and low-quality analytical data for chain extension monitoring. A, UPLC-UV trace (290 nm, 6 min gradient, 70%-95% B) of  $Eg_{52}$  to  $Eg_{64}$  chain extension ( $t = 0, 30$  min and 1 h); blue: starting material, yellow: 1-arm reacted, red: 2-arm reacted, green: product. Example for low-quality monitoring, the peaks are broad and not well resolved; b, UPLC-UV trace (290 nm, 6 min gradient, 70%-95% B) of  $Eg_{52}$  to  $Eg_{64}$  chain extension. Example for high-quality monitoring, the peaks are sharp and well resolved. c:  $ESI^+$  spectrum of  $Hub-Eg_{112}-OH$  species before PEG cleaning procedure; d,  $ESI^+$  spectrum of  $Hub-Eg_{112}-OH$  species after PEG cleaning procedure with 10% ammonia solution/ iso-propanol. .... 103

Figure 33. Results of the MALDI-TOF analysis (reflective mode, 1000 shots) of the same sample of  $MeO-Eg_{112}-OH$ . a, MALDI-TOF spectra of the same  $MeO-Eg_{112}-OH$  sample obtained at different laser powers; b, corresponding calculated relative chain length impurities calculated from the different spectra (the spectrum of run 2 is not depicted). .... 105

Figure 34. Overview of the reaction progress for a  $Eg_{112}$  to  $Eg_{140}$  chain extension at different reaction concentrations. The weight percentages (10 wt%, 20 wt% and 35 wt%) are the weight percentages of combined starting materials ( $Hub(Eg_{112}-OH)_3$  and  $Eg_{28}$  building block) in the solvent (diglyme). The red dotted line indicates when the reactions reached 100 % conversion. .... 107

Figure 35. Overview of a side-reaction involving the reaction with an ethyl tosylate contamination. a, final UPLC-UV chromatogram of an  $Eg_{28}$  to  $Eg_{40}$  chain extension, exhibiting one

unknown peak besides the signals for the building block (BB), dimer and the product; The underlying ESI-MS of the impurity peak reveals an m/z series of 1522, 1222., 1021 and 878, matching with a mass of 6015 Da; the proposed impurity structure is the  $\text{Hub}(\text{Eg}_{40}\text{ODmtr})_2(\text{Eg}_{28}\text{OEt})$ . b,  $^1\text{H}$  spectrum of ethyl tosylate isolated from the building block. c,  $^1\text{H}$  spectra of the building block before purification (bottom) and after purification (top)..... 110

Figure 36. Overview of the Hub-related purity of different Hub-PEG species derived from Nanostar Sieving based and chromatography-based synthesis. a, UPLC-UV traces of final  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$  material derived from a Nanostar Sieving based synthesis (bottom) and a chromatography-based synthesis (top); b, photographs of the final  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$  material derived from the Nanostar Sieving based synthesis (left) and a chromatography-based synthesis (right); c, relative Hub purities for the intermediated  $\text{Hub}(\text{Eg}_{64}\text{-OH})_3$  to  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$  for three Nanostar Sieving based batches (Rig 1-3) and one chromatography-based batch (Biotage)..... 112

Figure 37. a, UPLC-UV and TIC traces of a solution of  $\text{Dmtr-Eg}_{28}\text{-OH}$  in THF before (0 h) and 20 h after adding sodium hydride (20 h); b, the proposed pathway for the oxidation of  $\text{DmtrO-Eg}_{28}\text{-OH}$ , resulting in the formation of peroxide dimer 9 and diol 11. .... 114

Figure 38. Rejection and yield data for the membrane-purification based synthesis of defined  $\text{mPEG-OH}_{4966}$ . a, Screening results for the 16 wt% and 14 wt% membranes. Rejections of  $\text{Eg}_{28}$  Diol (orange),  $\text{Eg}_{56}$  Diol (green),  $\text{Hub}(\text{Eg}_{28}\text{-OH})_3$  (purple) and  $\text{Hub}(\text{Eg}_{56}\text{-OH})_3$ . b, set-up and results for the four steps of the membrane-based synthesis of  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$ ..... 116

Figure 39. Yields and purity data for the synthesis of defined  $\text{mPEG-OH}_{4966}$ . a, Functionalisation reaction to obtain UV-active derivative 5. b, d, UV-chromatograms (254 nm, 15 cm  $\text{C}_{18}$  column, 30 min 20-95 % B, A:  $\text{H}_2\text{O} + 5 \text{ mM } \text{NH}_4\text{OAc}$ , B  $\text{MeCN: MeOH } 4:1$ ) of 5 kDa  $\text{MeO-Eg}_{112}\text{-p-methoxybenzoate}$  derived from membrane-based synthesis (black trace) and chromatography-based synthesis (red trace); Peak P:  $\text{MeO-Eg}_{112}\text{-p-methoxybenzoate}$  (5), Peak I1:  $-\text{Eg}_{28}$  impurity ( $\triangleq \text{MeO-Eg}_{84}\text{-p-methoxybenzoate}$ ), Peak I2:  $+\text{Eg}_{28}$  impurity ( $\triangleq \text{MeO-Eg}_{140}\text{-p-methoxybenzoate}$ ), Peaks I3+I4: unidentified impurities; c, MALDI-TOF-MS of  $\text{mPEG-OH}_{4966}$  synthesised via

|                                                                                                                                                                                                                            |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| membrane-based approach in this work (bottom) and commercially available disperse mPEG-OH (top). d, summary of yields and purities for chromatography-based and membrane-based synthesis of MeO-Eg <sub>112</sub> -OH..... | 119 |
| Figure 40. Reaction scheme of the methylation of Hub(Eg <sub>112</sub> -OH) <sub>3</sub> (5d) to reach Hub(Eg <sub>112</sub> -OMe) <sub>3</sub> (13).....                                                                  | 125 |
| Figure 41. Reaction scheme of the cleavage reaction to yield linear MeO-Eg <sub>112</sub> -OH (mPEG <sub>4967</sub> alcohol, 7). .....                                                                                     | 126 |
| Figure 42. Reaction scheme of the synthesis of MeO-Eg <sub>112</sub> carboxylic acid (mPEG <sub>5024</sub> -CA, 14)..                                                                                                      | 127 |
| Figure 43. Reaction scheme of the synthesis of MeO-Eg <sub>112</sub> -succinimidyl carboxymethyl ester (mPEG <sub>5121</sub> -SCM, 15).....                                                                                | 128 |
| Figure 44. Reaction scheme for the synthesis of DmtrO-(CH <sub>2</sub> )-OH (26).....                                                                                                                                      | 129 |
| Figure 45. Reaction scheme of the synthesis of DmtrO-(CH <sub>2</sub> ) <sub>3</sub> -OTs (27).....                                                                                                                        | 129 |
| Figure 46. Reaction scheme of synthesis of MeO-Eg <sub>112</sub> -(CH <sub>2</sub> ) <sub>3</sub> -OH (20).....                                                                                                            | 131 |
| Figure 47. Reaction scheme of the reaction of MeO-Eg <sub>112</sub> -(CH <sub>2</sub> ) <sub>2</sub> -OH (20) to aldehyde (21).....                                                                                        | 132 |
| Figure 48. Reaction scheme of the reaction of MeO-Eg <sub>112</sub> -OH (7) to MeO-Eg <sub>112</sub> -succinimidyl succinate (mPEG <sub>5163</sub> -SS, 17).....                                                           | 133 |
| Figure 49. Reaction scheme of the reaction of MeO-Eg <sub>112</sub> -OH (7) to MeO-Eg <sub>112</sub> -succinimidyl propionate (mPEG <sub>5135</sub> -SP, 19).....                                                          | 134 |
| Figure 50. Reaction scheme for the reaction of MeO-Eg <sub>112</sub> -OH (7) to MeO-Eg <sub>112</sub> -OTs (mPEG <sub>5120</sub> -t, 22).....                                                                              | 135 |
| Figure 51. Reaction scheme for the reaction of MeO-Eg <sub>112</sub> -OTs (22) to MeO-Eg <sub>112</sub> -N <sub>3</sub> (mPEG <sub>4991</sub> -az, 23).....                                                                | 136 |
| Figure 52. Reaction scheme for the reaction of MeO-Eg <sub>112</sub> -N <sub>3</sub> (23) to MeO-Eg <sub>112</sub> -NH <sub>2</sub> (mPEG <sub>4065</sub> -amine, 24).....                                                 | 137 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 53. Reaction scheme for the tosylation of Hub(Eg <sub>112</sub> -OH) <sub>3</sub> (5h) to yield Hub(Eg <sub>112</sub> -OTs) <sub>3</sub> (29).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 138 |
| Figure 54. Reaction scheme for the conversion of Hub(Eg <sub>112</sub> -OTs) <sub>3</sub> (29) to Hub(Eg <sub>112</sub> -N <sub>3</sub> ) <sub>3</sub> (30).<br>.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 139 |
| Figure 55. Reaction scheme for the cleavage of HO-Eg <sub>112</sub> -N <sub>3</sub> (31) from Hub(Eg <sub>112</sub> -N <sub>3</sub> ) <sub>3</sub> (30).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 140 |
| Figure 56. Reaction scheme of the reaction of N <sub>3</sub> -Eg <sub>112</sub> -OH (31) with tert-butyl bromoacetate to reach product (32).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 141 |
| Figure 57. Overview of the linear, monofunctional defined M <sub>w</sub> 5 kDa PEGylation agents synthesised within this work. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 143 |
| Figure 58. General overview of the synthesis of a mPEG active ester using an ester of bromoacetic acid as an alkylating agent. Hydrolysis of the formed ester results in the carboxylic acid which is the activated with <i>N</i> -hydroxy succinimide/ DCC to reach the NHS-ester. ....                                                                                                                                                                                                                                                                                                                                               | 144 |
| Figure 59. Overview of the successful synthesis approaches for the three defined M <sub>w</sub> active ester derivatives mPEG <sub>5121</sub> -SCM, mPEG <sub>5135</sub> -SP and mPEG <sub>5163</sub> . The grey structures were intermediates and not isolated.....                                                                                                                                                                                                                                                                                                                                                                   | 145 |
| Figure 60. DMP-route to synthesise mPEG propionaldehyde (21). ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 146 |
| Figure 61. Overview of the synthetic pathways to reach the defined M <sub>w</sub> mPEG <sub>5120</sub> -tosylate, mPEG <sub>4991</sub> -azide and mPEG <sub>4965</sub> -amine.....                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 148 |
| Figure 62. Reaction pathway to achieve heterobifunctional defined M <sub>w</sub> PEG (32). ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 149 |
| Figure 63. Comparison of the UPLC-MS data of the defined M <sub>w</sub> mPEG SP synthesised in this work and a commercially available disperse derivative. a, UPLC/ELSD traces of the synthesised defined mPEG-SPA (top) and the commercially available disperse derivative (bottom) (15 min 29-95 % B gradient with A: H <sub>2</sub> O + 5 mM NH <sub>4</sub> OAc, B: MeCN/ MeOH 4:1). The defined sampled shows 4 peaks (I1: traces of starting material mPEG-OH <sub>4966</sub> , I2: hydrolysed product/ carboxylic acid, P: product (mPEG-SPA), I3: NHS-breakdown product). For the disperse sample peak I1 is broad and peak I2 |     |

and P are not separated. b, For the defined sample the underlying ESI<sup>+</sup> traces lead to the corresponding mass. For example (shown) the m/z series 1301.5, 1044.8, 873.7, 751.6, 659.8 corresponds to a M<sub>w</sub> of 5135.11, which is the mPEG-SPA ([M+4 NH<sub>4</sub>]<sup>4+</sup>, M+5 NH<sub>4</sub>]<sup>5+</sup>, M+6 NH<sub>4</sub>]<sup>6+</sup>). For the disperse sample the determination of a precise molecular weight is not possible. .... 151

Figure 64. Comparative study of the random PEGylation and characterisation of BSA using defined and disperse PEG 5 kDa. a, Conventional PEGylation: uses disperse PEGylation agent. Random PEGylation at the primary amine groups of BSA, using defined or disperse 5 kDa linear mPEG-SP (1). Size-exclusion chromatography to isolate the mono-PEG BSA isomers (2). Digestion (3): enzymatic digestion of the mono-PEG BSA results in a mixture of PEGylated and un-PEGylated peptide fragments. The different peptide/ PEG-peptide species are then separated via RP-UPLC (4). In the case of the disperse PEG the resulting data cannot be interpreted. b, In this work a defined, 5 kDa PEGylation agent is used for the first time to randomly PEGylate BSA. After steps 1-4 in-line ESI-MS allows for extraction of the masses for identifying individual PEG-peptide fragments. The masses of the PEGylated peptide fragments can be calculated and (equivalent to a standard peptide mapping procedure) these can then be compared to an online digestion database to identify the PEGylated fragments. .... 155

Figure 65. Overview of the characterisation and separation of PEGylated BSA through MALDI-TOF-MS and size-exclusion chromatography. a, MALDI-TOF-MS of the crude PEGylation mixture derived from PEGylating BSA with mPEG<sub>5163</sub>-SS after 1 hour. The spectrum shows peaks for the native BSA (66.5 kDa), the mono-PEG BSA (71.5 kDa), the di-PEG BSA (76.5 kDa), and the tri-PEG BSA (81.5 kDa). b, Size-exclusion chromatogram of the PEGylation mixture after 1 h. c, MALDI-TOF-MS of three of the fractions collected from the SEC; the spectra confirm the presence of native BSA (66.5 kDa, green peak), the mono-PEG BSA (71.5 kDa, blue peak) and di-PEG BSA (76.5 kDa, grey peak). .... 159

Figure 66. Overview of the first peptide mapping results performed with BSA PEGylated with the mPEG<sub>5163</sub>-SS reagent. .... 160

Figure 67. Overview of the results of the identification of the PEGylated fragments in mono-PEGylated BSA through a peptide mapping approach. a, size-exclusion chromatogram of PEGylated BSA (top) and the MALDI-TOF spectra of the fraction of remaining native BSA (yellow) and the mono-PEG BSA (blue). b, TIC trace of UPLC chromatogram of Trypsin-digestion samples of mono-PEG BSA, PEGylated with a disperse (bottom, red) and defined (top, green) 5 kDa mPEG-Succinimidyl Propionate. Examples are shown for the ESI<sup>+</sup> traces extracted from the chromatograms. For the defined species (RT 20.59 min), the m/z series (1045.8, 899.0, 788.9, 703.2, 634.7) corresponds to a total mass of 6166.68 Da. Subtracting the PEG mass (5021.03 Da) identified the PEGylated fragment (1145.65 Da) as Pos. 236-245 (AS sequence -AWSVARLSQK-). The disperse derivative does not allow for extraction of exact masses. c, summary of the seven identified PEGylated fragments in mono-PEGylated BSA..... 162

## Acronyms and Abbreviations

|                |                                             |
|----------------|---------------------------------------------|
| AA             | Amino acid                                  |
| CID            | Collision induced dissociation              |
| DCA            | Dichloroacetic acid                         |
| DCM            | Dichloromethane                             |
| DMF            | Dimethylformamide                           |
| Eg             | Ethylene glycol                             |
| ELSD           | Evaporative light scattering detector       |
| ESI-           | Electro spray ionisation (negative mode)    |
| ESI+           | Electro spray ionisation (positive mode)    |
| g              | Grams                                       |
| HPLC           | High-performance liquid chromatography      |
| IEC            | Ion exchange chromatography                 |
| LG             | Leaving group                               |
| m/z            | Mass-to-charge ratio                        |
| MALDI          | Matrix assisted laser desorption ionisation |
| MeCN           | Acetonitrile                                |
| MeOH           | Methanol                                    |
| mg             | Milligram                                   |
| ml             | Millilitre                                  |
| MS             | Mass spectrometry                           |
| M <sub>w</sub> | Molecular weight                            |
| NMR            | Nuclear Magnetic Resonance                  |
| OEG            | Oligo(ethylene glycol)                      |

|      |                                         |
|------|-----------------------------------------|
| PDI  | Polydispersity index                    |
| PEG  | Poly(ethylene glycol)                   |
| PG   | Protecting group                        |
| RP   | Reversed-phase                          |
| RPC  | Reversed-phase chromatography           |
| SEC  | Size exclusion chromatography           |
| TOF  | Time of flight                          |
| UC   | Ultra-centrifugation                    |
| UF   | Ultrafiltration                         |
| UPLC | Ultra-performance liquid chromatography |
| UV   | Ultraviolet                             |
| VIS  | Visible                                 |

## Chapter 1 Introduction

PEGylation (the covalent attachment of one or more poly(ethylene glycol) (PEG) units to therapeutics) is a technique widely used in pharmaceutical industry for increasing a drug's blood-residence-time and that way increasing its bioactivity.<sup>3</sup> This is partially a consequence of the highly hydrophilic character of PEG, resulting in a higher solubility in aqueous environments and a increased hydrodynamic radius, leading to less renal clearance.<sup>4</sup> Furthermore, the “stealth effect” provides a protective barrier for polymer-modified entities by making them less accessible to blood components such as antibodies or enzymes, further increasing the blood-residence time.<sup>5,6</sup> There are currently 40 PEGylated pharmaceuticals approved by the FDA (as of 18/12/2023), 31 of which are based on proteins.

The currently used PEGylation agents all show some degree of dispersity in their molecular weights, despite often being labelled with the term “monodisperse”. This is because PEG is still conventionally synthesised through the anionic ring opening polymerisation of ethylene oxide, which results in a Gaussian distribution of PEG of different chain length.<sup>7</sup> As a result of compromised chromatographic behaviour and unclear mass spectrometric analysis, the disperse nature of these PEGylation agents results in batch-to-batch variations and difficulties in purification and characterisation of the PEGylated therapeutics.<sup>8-11</sup> These problems become more pronounced both with the increasing molecular weight of the PEG and with an increasing number of PEGamers (species with varying numbers of pendant PEG molecules) or positional isomers incorporated in the sample. Despite the availability of advanced analytical methods, the exact binding site of the PEG in a PEGylated protein often remains unclear, or PEG-related impurities remain unidentified.

Using defined molecular weight ( $M_w$ ) PEG instead of the currently used disperse derivatives would overcome these analytical problems and could have a substantial impact on the overall purity and quality of PEGylated therapeutics. Furthermore, therapeutics PEGylated with defined

PEG could potentially exhibit an improved biodistribution profile compared to the disperse derivatives.

There has been increased interest in producing defined  $M_w$  PEGs and several synthesis approaches were developed, mostly based on extensive chromatographic purification or solid-phase techniques.<sup>12-18</sup> Those techniques suffer from limited scalability or impeded reaction kinetics, making them economically unfeasible, especially for higher molecular weight PEG (> ~3 kDa).

LIVINGSTON *et al.* have previously developed a technique for synthesising sequence-defined polymers using an iterative approach based on Nanostar Sieving.<sup>19</sup> This approach is solely relying on organic solvent nanofiltration (OSN) membranes for purification and offers the advantages of scalability and liquid-phase reactions.

The work in hand had three main aims: the optimisation of using the Nano Star Sieving approach for synthesising defined  $M_w$  PEG, the synthesis of active defined  $M_w$  PEGylation agents and investigating the application of these in PEGylated therapeutics.

To give a deeper understanding on the fields of PEGylation and the synthesis of defined  $M_w$  PEG, a thorough literature review is provided in **Chapter 2** of this thesis.

In **Chapter 3** the three main objectives of this work were defined. The first objective was focused on the synthesis of defined  $M_w$  PEG and included the optimisation of the nanostar sieving-based synthesis approach to routinely synthesise 5kDa defined  $M_w$  PEG on a multi-gram scale. The nanostar sieving-based synthesis was compared to an equivalent chromatography-based one regarding the purities, yields and economic viability. Additionally, the synthesis of defined  $M_w$  PEG derivatives up to a  $M_w$  of 20 kDa was investigated. The according results are discussed in **Chapter 4**.

In the next step, the previously synthesised 5 kDa defined  $M_w$  PEGs were chemically modified to obtain active PEGylation reagents (**Chapter 5**). First studies on the analytical comparison of defined  $M_w$  and disperse PEGylation agents were performed.

The final aim of this work was to investigate the application of defined  $M_w$  PEG as PEGylation agents. Within the scope of **Chapter 6** a comparative study on the PEGylation of the 66 kDa protein bovine serum albumin (BSA) was performed. A previously synthesised defined 5 kDa methoxy-PEG active ester and a commercially available disperse derivative were used to randomly PEGylate and subsequently characterise the PEGylated BSA. It was demonstrated that the defined  $M_w$  PEG both allows for facile monitoring of chemical modification reactions during the synthesis of the PEGylation agents, and facilitates straightforward identification of the PEGylated fragments within a PEGylated protein through a simple peptide mapping approach using UPLC-MS.

**Chapter 7** summarises the results and potential impact of this thesis and gives an outlook into future work that must be pursued to translate the obtained results on defined  $M_w$  PEGylation agents into a practical and industrially applicable process.

## Chapter 2 Literature Review

### 2.1 Poly(ethylene glycol) and PEGylation

#### 2.1.1 Introduction

Poly(ethylene oxide) (PEG) is a highly hydrophilic polymer consisting of ethylene oxide (-CH<sub>2</sub>-CH<sub>2</sub>-O-) repeating units, with each repeating unit having a molecular weight of 44.05 Da. It is conventionally derived from the anionic polymerisation of ethylene oxide.<sup>7</sup>

PEG has a wide range of applications, for example, in the cosmetic industry (in skin products), in the food industry (for example, as a binding agent) or in various chemical applications (for example, as a preservation agent for historical artifacts).

Since its discovery by ABUCHOWSKI *et al.* in 1977<sup>20</sup>, the masking of small drugs, proteins or peptides with linear or branched PEG has become a well-established method for increasing the blood-residence time of therapeutics and reducing their immunogenicity.

Numerous reviews on the different aspects of PEGylation have been published in the last decades. The following literature review is based on the publications by Harris *et al.*<sup>1,2,21</sup>, Knop *et al.*<sup>5</sup>, D'souza *et al.*<sup>6</sup>, Veronese<sup>22</sup>, Swierczweska *et al.*<sup>23</sup> and Turecek *et al.*<sup>24</sup>.

PEG is a hydrophilic, flexible, and inert molecule and in combination with its low immunogenicity it is the ideal candidate for shielding therapeutics. Due to its highly hydrophilic character PEG molecules tend to form a protective barrier around PEGylated species. This results in decreased interactions with the blood components (such as enzymes) and thus in a reduced clearance of PEGylated molecules and a prolonged blood circulation time. This phenomenon is referred to as the "stealth effect".<sup>5</sup> Due to the coordination of water molecules (6-7 per Eg-unit), PEG molecules also exhibit a much larger hydrodynamic radius compared to, for instance, proteins of similar molecular weight.<sup>25</sup> PEGylated drugs thus present the advantage of slower renal excretion (additionally contributing to a prolonged blood circulation time). Despite the observation of

decreasing *in vitro* activity of a PEGylated drug with increasing molecular weight of the PEG, the *in vivo* activity shows the opposite trend.<sup>26,27</sup>

Of the 40 FDA- or EMA approved PEGylated therapeutics (as of 18/12/2023, **Table 1**), 31 are proteins or enzymes.

**Table 1.** List of FDA, or EMA-approved PEGylated therapeutics (as of 20/11/2023).

|    | Name       | Year of approval | Drug size [kDa] | PEG size [kDa] | PEG architecture | Number PEG molecules | Drug form      |
|----|------------|------------------|-----------------|----------------|------------------|----------------------|----------------|
| 1  | Adagen     | 1990             | 96- 126         | 5              | linear           | 11 to 17             | Enzyme         |
| 2  | Oncaspar   | 1994             | 483- 548        | 5              | linear           | 69 to 82             | Enzyme         |
| 3  | Doxil      | 1995             | n.a.            | 2              | linear           | <i>n.a.</i>          | Liposome       |
| 4  | PegIntron  | 2001             | 31              | 12             | linear           | 1                    | Protein        |
| 5  | Pegasys    | 2002             | 60              | 40             | branched         | 1                    | Protein        |
| 6  | Neulasta   | 2002             | 39              | 20             | linear           | 1                    | Protein        |
| 7  | Somavert   | 2003             | 42- 52          | 5              | linear           | 4 to 6               | Protein        |
| 8  | Macugen    | 2004             | 50              | 40             | branched         | 1                    | Aptamer        |
| 9  | Mircera    | 2007             | 60              | 30             | linear           | 1                    | Protein        |
| 10 | Cimzia     | 2008             | 91              | 40             | branched         | 1                    | FAB fragment   |
| 11 | Krystexxa  | 2010             | 500             | 10             | linear           | 4 x 9                | Enzyme         |
| 12 | Asclera    | 2010             | 0.6             | 0.4            | linear           | 1                    | Small Molecule |
| 13 | Sylatron   | 2011             | 31              | 12             | linear           | 1                    | Protein        |
| 14 | Omontys*   | 2012             | 45              | 40             | branched         | 1                    | Peptide        |
| 15 | Lonquex    | 2013             | 39              | 20             | linear           | 2                    | Protein        |
| 16 | Plegridy   | 2014             | 44              | 20             | linear           | 1                    | Protein        |
| 17 | Movantik   | 2014             | 0.32            | 0.34           | linear           | 1                    | Small Molecule |
| 18 | Rebinyn    | 2017             | n.a.            | 40             | branched         | 2                    | Protein        |
| 19 | Adynovate  | 2015             | 265             | 20             | linear           | 2                    | Protein        |
| 20 | Onivyde    | 2015             | n.a.            | 2              | linear           | <i>n.a.</i>          | Liposome       |
| 21 | Asparlas   | 2018             | n.a.            | 5              | <i>n.a.</i>      | 31-39                | Enzyme         |
| 22 | Jivi       | 2018             | 234             | 60             | branched         | 1                    | Protein        |
| 23 | Fulphila** | 2018             | 39              | 20             | linear           | 1                    | Protein        |
| 24 | Revcovi    | 2018             | 40              | 5              | linear           | 12 - 14              | Enzyme         |
| 25 | Palynzig   | 2018             | 820             | 20             | linear           | 9                    | Enzyme         |
| 26 | Udenyca**  | 2018             | 39              | 20             | linear           | 1                    | Protein        |
| 27 | Ziextenzo  | 2019             | 39              | 20             | linear           | 1                    | Protein        |
| 28 | Esperoct   | 2019             | n.a.            | 40             | <i>n.a.</i>      | 1                    | Protein        |
| 29 | Nyvepria   | 2020             | 39              | 20             | linear           | 1                    | Protein        |
| 30 | Empaveli   | 2021             | 2 x 1.7         | 40             | linear           | 1                    | Peptide        |

|    |             |      |             |             |             |             |                    |
|----|-------------|------|-------------|-------------|-------------|-------------|--------------------|
| 31 | Skytrofa    | 2021 | 63          | 10          | linear      | 4           | Protein            |
| 32 | BESREMi     | 2021 | <i>n.a.</i> | 40          | <i>n.a.</i> |             | Protein            |
| 33 | Fylnetra**  | 2022 | 39          | 20          | linear      | 1           | Protein            |
| 34 | Stimufend** | 2022 | 39          | 20          | linear      | 1           | Protein            |
| 35 | Rolvedon    | 2022 | 72          | 3.4         | linear      | 1           | Protein            |
| 36 | Elfabrio    | 2023 | 116         | 2.3         | linear      | 9           | Enzyme             |
| 37 | Izervay     | 2023 | 56          | 43          | branched    | 1           | RNA aptamer        |
| 38 | Onpattro    | 2018 | <i>n.a.</i> | <i>n.a.</i> | linear      | <i>n.a.</i> | Lipid Nanoparticle |
| 39 | Corminaty   | 2020 | <i>n.a.</i> | <i>n.a.</i> | linear      | <i>n.a.</i> | Lipid Nanoparticle |
| 40 | Spikevax    | 2020 | <i>n.a.</i> | <i>n.a.</i> | linear      | <i>n.a.</i> | Lipid Nanoparticle |

\* The approval for Omontys was withdrawn in 2013/2014 due to hypersensitivity issues.<sup>28</sup>

\*\*Biosimilars of the drug Neulasta.

*n.a.*: data not available

### 2.1.2 Historical Development and PEGylation Chemistry

The first successful PEGylation was published by ABUCHOWSKI *et al.* in 1977, who laid the foundation for the following “first generation” PEGylation processes.<sup>20</sup> The researchers synthesised conjugates of bovine serum albumin and methoxy(polyethylene glycol) (mPEG) of 1.9 and 5 kDa by activating the mPEG with cyanuric chloride and then coupling it to free amino groups of the protein (**Figure 1**).

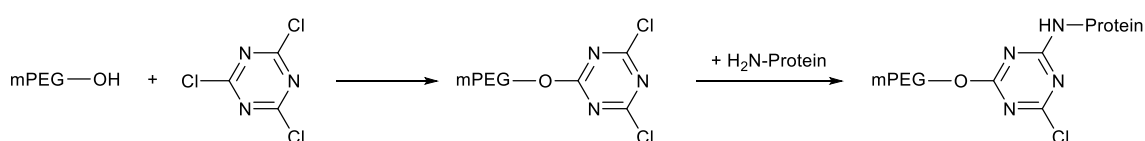


**Figure 1. Activation of mPEG-OH with cyanuric chloride and coupling to amino group of a protein.** Method developed by ABUCHOWSKI *et al.* in 1977. Reproduced with permission from [20].

PEGylation through reacting amino groups with an activated PEG derivative is still the most common method, although many different and improved variations exist. In the early days of PEGylation, the activated PEG was usually randomly attached to the abundant ε-amino groups of lysine, which make up around 10 % of most proteins.<sup>29</sup> Random PEGylation includes the formation of different isomers (positional and weight-related), which causes problems regarding

batch-to-batch variations and the characterisation of the final product. Additionally, only low molecular weight PEGs (< 10 kDa) were used in first generation PEGylation. Since then it has been found, that mono-PEGylation or site-directed PEGylation with fewer PEGs of higher  $M_w$  delivers better results than PEGylation with multiple PEGs of lower  $M_w$ , as the probability of interaction of the PEG and the proteins active site is lowered.<sup>1</sup> Furthermore, early PEGylation agents were contaminated with a large amount of PEG diol (up to 15 %), resulting in cross-linking and aggregation of the proteins.<sup>30</sup> This can be avoided by specific synthesis techniques (discussed in **Section 2.2**) or extensive purification.

The most common PEGylation agents are based on a *N*-hydroxysuccinimide (NHS) activated ester and form either a carbamate/urethane linkage or an amide linkage with available primary amine groups (from for example lysine, histidine or serine), thus reducing the charge of the protein (**Figure 2**).

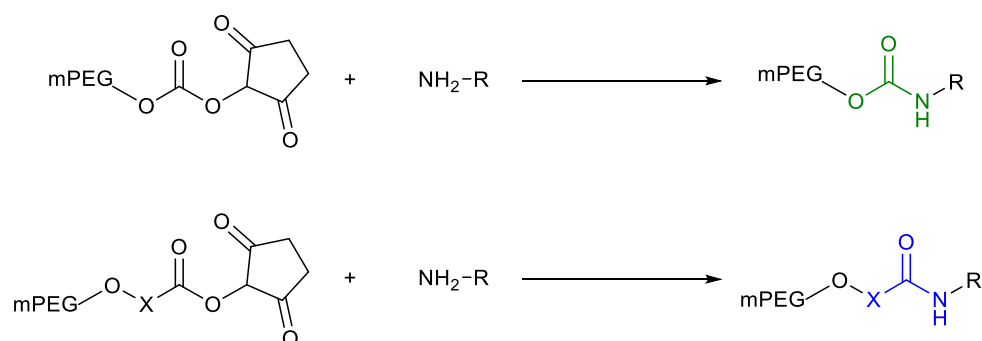


**Figure 2. Reactions of mPEG-NHS-derivatives with primary amines.** Top: Reaction of NHS-carbonate with amine to form a carbamate/urethane linkage. Bottom: Reaction of NHS-ester with amine to form an amide linkage; X= (CH<sub>2</sub>)<sub>n</sub>. Reproduced with permission from [24].

**Figure 3** shows a selection of different first-generation PEGylation agents, including the early cyanuric chloride-based derivatives (**Figure 3a**). The di-substituted cyanuric chloride was the first attempt at creating a branched PEGylating agent.<sup>31</sup>

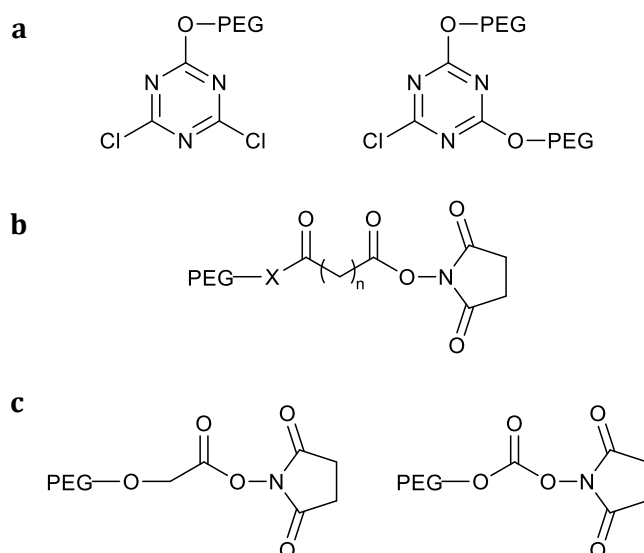


**Figure 3. Different first-generation PEGylation reagents. a**, based on cyanuric chloride. **b**, based on succinimidyl succinate. **c**, NHS ester (left) and NHS carbonate (right).

Although more advanced PEGylation methods are available nowadays, the two first FDA-approved therapeutics *Adagen* and *Oncaspar* (**Table 1**, 1 and 2) are still on the market. Both are enzymes randomly PEGylated with linear PEG chains with an average weight of 5 kDa, achieved by reaction of 5 kDa PEG succinimidyl glutarate (an NHS ester) with lysine moieties of the protein.<sup>32,33</sup> About ten years after the approval of *Adagen*, the first PEGylated therapeutic Interferon (Interferon  $\alpha 2$ -b) was marketed as *PegIntron* (**Table 1**, 4). This was the first example for a mono-PEGylation, which nevertheless still delivered a product mixture of 14 different isomers.<sup>34</sup> Mono-PEGylation using a 12 kDa PEG succinimidyl carbonate was accomplished by conducting the reaction in a mildly acidic environment (pH 6.5). This led to more than fifty percent of the resulting compound being mono-PEGylated at Histidine 34, while the remainder consisted of positional isomers resulting from the mono-PEGylation of other amino acid (AA) residues. In the same year BAILON *et al.* mono-PEGylated an Interferon  $\alpha 2$ -a with a much larger, branched PEG structure, developed previously by MONFARDINI *et al.*<sup>35,36</sup> This branched PEGylation

agent is reached by adding two linear PEGs (each 20 kDa) to both amino residues of a lysine molecule, which is then converted to an NHS ester (**Figure 4**).<sup>36</sup> This can be selectively attached to lysine residues of a protein, which, in case of the drug *Pegasys*, results in 6 different positional isomers (**Table 1, 5**). First studies on the effect of using branched instead of linear PEG structures have been published and are discussed in **Chapter 2.1.3.1**.

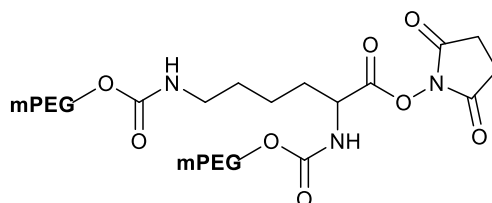


**Figure 4. Structure of the branched PEG-NHS PEGylation agent based on Lysine.** Developed by MONFARDINI *et al.* [36].

Besides the usage of NHS-PEGs described above, which can result in a relatively selective PEGylation of different primary amine residues, more sophisticated PEGylation methods have been developed, also known as “second-generation” PEGylation agents. As the acylation with an NHS ester results in a reduction of the protein charge, in some cases an alkylation is favoured (**Figure 5**).

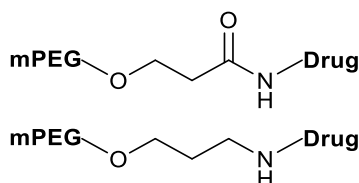


**Figure 5. Schematic PEGylated drug derived from acylation (top) and alkylation (bottom) with a methoxy-PEG agent.** Reproduced from [24].

This can be achieved by using PEG aldehydes as PEGylating agents, which leads to a selective PEGylation of the *N*-terminus in lower pH regions through reductive amination. This is due to the fact that the *N*-termini have lower  $pK_A$  values than other amine groups in a protein.<sup>37</sup> This

PEGylation technique has been used for FDA approved protein therapeutics *Neulasta* and *Plegridy* (**Table 1**, 6 and 16).

Another important class of second-generation PEGylation agents are thiol selective. If the drug is equipped with a cysteine group prior to PEGylation, and other potentially present cysteine groups are not accessible for the PEGylating agent, mono-PEGylation can be achieved. This technique was used in the marketed drugs *Cimzia* and *Jivi*.<sup>38</sup> In both cases the PEG is a branched structure based on Lysine and functionalised with maleimide which selectively reacts with thiol groups.

In general, when choosing the PEGylation agent to create a PEGylated therapeutic, the potential accessible linking points (AAs in case of protein/peptide therapeutics) must be identified. For instance, in many proteins the *N*-terminus is buried in the folded protein and is hence not suitable for PEGylation. Furthermore, potential hindrance of the active site of the therapeutic must be avoided. The number of PEG molecules attached, and the total molecular weight of the PEG can affect the PEGylated proteins' pharmacological behaviour, which will be discussed in more detail below.

### 2.1.3 Pharmacological Behaviour of PEGylated Proteins

The bioactivity of PEGylated proteins represents a dynamic interplay between their pharmacodynamics and pharmacokinetics. The pharmacokinetics of a therapeutic describe its pathway through the body, which includes the stages of absorption, distribution, metabolism, excretion and toxicity, commonly referred to with the acronym "ADMET".<sup>39</sup> A drug's pharmacodynamics on the other hand describe the physiological effects it has, based on its chemical interactions within the body.<sup>40</sup>

In case of PEGylated proteins, it is a well-known phenomenon that their *in vivo* efficacy cannot be directly inferred from their *in vitro* activity.<sup>12</sup> In most cases cell culture assays result in a decreased *in vitro* activity with increasing PEG molecular weight, whilst the *in vivo* activity shows the inverse trend (**Figure 6**).<sup>12</sup> To give a better understanding of these trends, the

pharmacokinetic and pharmacodynamic characteristics of PEGylated proteins are explained in the following.

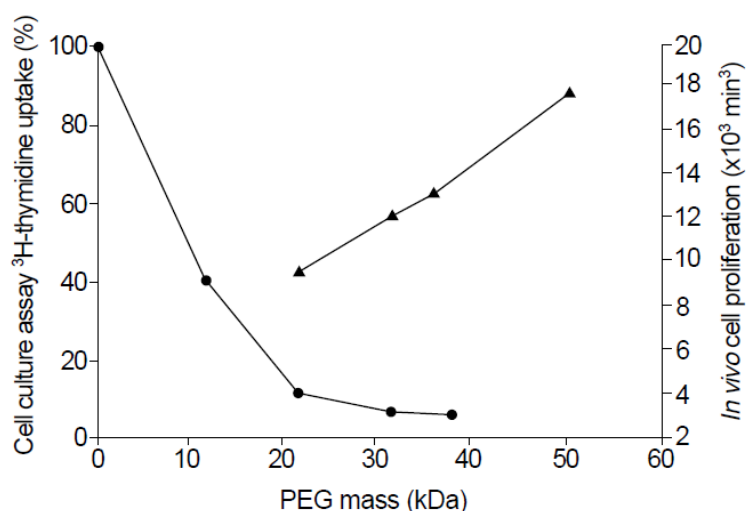


**Figure 6. *In vivo* activity and *in vitro* activity of a PEGylated therapeutic dependent on the PEG molecular weight.** The cell culture assay <sup>3</sup>H-thymidine uptake is a measure for the *in vitro* activity and decreases with rising PEG molecular weight, whereas the *in vivo* cell proliferation increases with increasing PEG molecular weight. Copied with permission from [26].

The PEGylation of proteins and other therapeutics carries the risk of compromised pharmacodynamics, as the attached PEG molecule potentially interacts with the drug's active region and thus might decrease its binding affinity. However, this potential drawback is typically counterbalanced by a series of improved pharmacokinetic parameters.<sup>41</sup>

The previously described "stealth effect" is resulting in a higher stability towards enzymatic degradation and a decreased formation of antibodies, which equals a lower immunogenicity.<sup>5</sup> Moreover, the highly hydrophilic nature of PEGylated proteins translates into a larger hydrodynamic radius, which subsequently leads to decreased renal clearance. The threshold for renal clearance is commonly reported in  $M_w$ , even though it is rather dependent on the hydrodynamic radius. Nevertheless, both are proportional to each other for one molecule type and structure. For linear PEG molecules the renal clearance threshold is around 30 kDa, whilst molecules with a higher  $M_w$  are cleared through liver uptake or the immune system.<sup>21,42</sup>

### 2.1.3.1 Branched vs. linear PEGylation agents

From **Table 1** it becomes apparent that in FDA approved PEGylated therapeutics, PEGylation agents with a molecular weight of 40 kDa or more are consistently branched structures. This raises the question as to why these branched structures are used, rather than the simple linear counterparts with equivalent molecular weight. There is limited research on the properties of branched PEGylation agents compared to the previously used linear ones. In the context of introducing the most popular branched PEGylation agent based on lysine, MONFARDINI *et al.* conducted studies on different properties of enzymes PEGylated with conventional linear PEG and the new branched structure.<sup>36</sup> The pH- and temperature-stability of the derived structures, the kinetics of the PEGylated enzymes and their stability towards proteolytic digestions were investigated. It is important to mention that the linear PEG structure used within this study had a  $M_w$  of 5 kDa, whilst the branched structure (consisting of two 5 kDa linear PEGs attached to one central lysine) had a  $M_w$  of 10 kDa. The researchers did not find any difference in stability towards pH or temperature changes of branched and linear PEG. In contrast, it was found that the enzymes PEGylated with the branched version had a significantly higher stability towards proteolytic digestion. This was explained by a greater steric hindrance and thus the prevention of macromolecules (for example, proteolytic enzymes, antibodies) approaching the enzyme. Furthermore, the researchers suggested advantages of the branched structure regarding immunogenicity and blood residence time but did not provide data to support this claim. A later study by VERONESE *et al.* included the comparison of a 10 kDa branched structure and a linear PEG of the same molecular weight.<sup>43</sup> Again it was found, that enzymes (uricase) PEGylated with the branched structure exhibit a significantly higher stability against proteolytic digestion. According to the researchers, the branched PEG adopts an “umbrella-shaped” conformation, hindering other macromolecules from approaching the enzyme (**Figure 7**). This theory provides an explanation for the longer blood residence time observed, as the enzymes are more stable.<sup>44</sup> Another observation within this publication was the higher activity of uricase PEGylated with the branched PEG derivative, compared to PEGylation with the linear one at the same degree of

PEGylation. It is important to note that a random PEGylation was performed with PEGylation degrees around 30 % (i.e., 30 % of the amino groups were PEGylated). It could be possible, that the linear PEGylation reagent has a higher reactivity towards amine groups close to the active site (which is found in a narrow cleft within the protein structure), compared to the branched reagent. The high PEGylation degrees could thus result in a lower enzymatic activity due to masking of the active site.

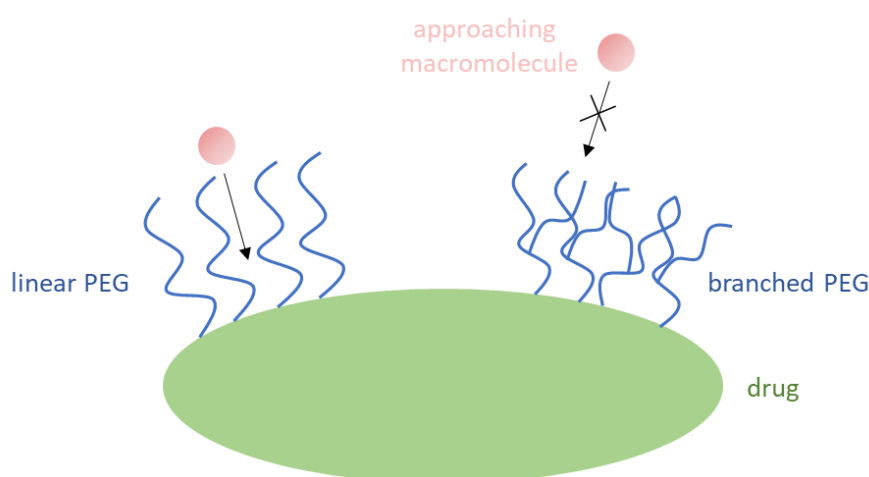


**Figure 7. Schematic depiction of a therapeutic PEGylated with linear or branched PEG.** The “umbrella” like structure of the branched version is hindering other molecules from approaching the drug. Reproduced with permission from [44].

So far, there is no evidence for improved bioactivity by mono-PEGylation with a branched PEG compared to mono-PEGylation with the linear analogue.

VERONESE’s assumption of the branched PEG having a higher hydrodynamic radius than the linear PEG, was disputed by FEE in 2007, proving the viscosity radii of a linear and a branched PEG of the same molecular weight to be the same.<sup>45</sup>

In summary, to achieve a balance of pharmacodynamic and pharmacokinetic behaviour within a PEGylated therapeutic, the following points must be considered:

1. The molecular weight of the PEGylation agent needs to be sufficiently high to avoid renal clearance.

2. Mono-PEGylation is favoured over multi-PEGylation, as the chance of interaction with active sites is minimised.
3. A mixture of different positional isomers should be avoided. The most active positional isomer is preferred and might be reached by choosing the right PEGylation chemistry.
4. Therapeutics PEGylated with branched versions of PEG may exhibit a higher *in vivo* stability than their analogues PEGylated with linear derivatives.

#### **2.1.3.2 Toxicity and Immunogenicity**

PEG has been extensively utilised in pharmaceuticals and biotechnology due to its ability to improve drug solubility, extend circulation time, and reduce immunogenicity. However, it is essential to recognise that PEG might create an immune response itself, and its immunogenicity has been a topic of growing interest in recent years.<sup>46-49</sup> PEG has been initially thought to be a biologically inert polymer. Earlier studies were primarily focused on low molecular weight PEG (for example PEG 400) and have not revealed significant immunogenic responses.<sup>50</sup> However, as the use of high molecular weight PEG in PEGylated therapeutics had become more prevalent, concerns about PEG immunogenicity have emerged. In general, the molecular weight cut-off for renal clearance of PEG is lower than that of proteins, with about 30 kDa.<sup>3</sup> This, in combination with a lack in biodegradability leads to an increased risk for accumulation of PEG within the body, potentially resulting in macromolecular syndrome.<sup>51</sup> Hence, there is an increased interest in developing degradable alternatives, but so far, all FDA-approved PEGylated drugs contain native PEG and there is a lack of research on the pharmacological behaviour of potential alternatives. One recent study addressed the problem of biodegradability by developing a new masking polymer based on a degradable PEG-amide backbone.<sup>52</sup> The study revealed that conjugates of PEG-amide structures and proteins are smaller than equivalent structures based on a pure PEG backbone.<sup>52</sup> This suggests that the amide groups interact with the protein (in the form of hydrogen bonding) and the polymer wraps around it more tightly. As a result, the desired increase in blood-circulation time might be impeded.

Since PEG is nowadays widely used in every-day applications such as cosmetics, potential patients are likely to have been exposed to PEG before being treated with a PEGylated therapeutic. Hence, the risk of patients carrying PEG antibodies and creating an immune response upon treatment with a PEGylated drug has increased.<sup>46</sup> According to literature the prevalence of anti-PEG antibodies in patients who have not been treated with a PEGylated therapeutic before, rose from 0.2 % in 1984 to 70 % in 2016.<sup>46,47</sup> This number might now be higher, especially since the wide application of the PEGylated SARS-CoV-2 RNA vaccines Spikevax (Moderna) and BNT162b2 (Pfizer-BioNTech).<sup>49</sup>

Anti-PEG antibodies do not always trigger a severe immune response, but can also simply result in the drug being cleared from the body more quickly (the so-called accelerated blood clearance response/ ABC), which reduces its efficiency.<sup>46,53,54</sup> In a detailed critical review on the immunogenicity of PEG, CHEN and ROFFLER pointed out that the ABC response is usually only significant for PEGylated drugs administered at low doses, whereas hypersensitivity gets more problematic at higher doses.<sup>49</sup> They furthermore stress that the immune response of a PEGylated drug is not only determined by the PEG itself, but largely depends on the immunogenicity of the active ingredient. For example, totally human proteins normally do not trigger a human response, whereas non-human proteins do and thus their PEGylated derivatives are more likely to result in the formation of PEG-antibodies.<sup>49</sup> The authors make an appeal to introduce routine testing for PEG-antibody levels to mitigate immune-response related problems.

In conclusion, the immunogenicity of PEG is a complex interplay of different factors that have to be monitored, but in general the benefits of PEGylation still outweigh its risks and PEGylation remains the gold-standard in the masking of therapeutics.

#### **2.1.3.3 Dispersity of PEGylation agents**

The conventional synthesis of PEG through anionic ring-opening polymerisation results in relatively narrow distributions of PEG of different molecular weights, with the index for the molar mass dispersity ( $D_M$ )  $\leq 1.01$ , which can be further improved using chromatography.<sup>9</sup> For instance, the PEGylation agent used in the FDA-approved drug *Neulasta* exhibits a dispersity index of

1.002.<sup>55</sup> Nevertheless, each sample of the therapeutic still contains species with varying molecular weights, which has been shown to result in problems during the chemical characterisation, batch-to-batch variations and could also potentially affect the biodistribution behaviour of the drug.<sup>55</sup>

The adverse effects of the dispersity of the PEGylation agents on the characterisability of PEGylation agents and PEGylated therapeutics is frequently mentioned in literature.<sup>8-10,56,57</sup> Nevertheless, due to the availability of defined  $M_w$  PEGylation agents being limited to a  $M_w$  up to about 1.5 kDa, so far there is no data available on the biodistribution behaviour of therapeutics PEGylated with defined  $M_w$  PEG compared to the disperse derivatives.

A recent study investigated a new defined  $M_w$  masking polymer, based on a PEG-amide backbone, in which oligo(ethylene glycol) (OEG) units were successively attached to each other through an amide bond.<sup>52</sup> This resulted in defined  $M_w$  PEG-like polymers with molecular weights of 5.8 kDa and 11.4 kDa.<sup>52</sup> These were covalently attached to different albumins and the sizes and immunogenicity of the resulting polymer-protein species was compared to derivatives where a conventional 5 kDa disperse PEGylation agent was used. It was shown that the new defined  $M_w$  OEG polyamide-protein species exhibit a lower hydrodynamic radius and a similar immunogenicity compared to the disperse PEG-protein species. No investigations of the biodistribution behaviour or bioactivity were included in this study.

So far there is no research data available on the potential differences in the pharmacological properties of therapeutics PEGylated with disperse vs. defined  $M_w$  PEGylation agents. One assumption can be made based on data showing that the blood-residence time of a PEG increases with increasing  $M_w$  (**Figure 8**). In the region between about 10 kDa and 50 kDa the curve is particularly steep and a small change in the  $M_w$  potentially translates into a significant change in the blood-residence time. Even though the curve might be shifted to higher molecular weights for PEGylated therapeutics (PEGylated proteins in particular), the same trend can be expected. It can be speculated, that the differences in blood-residence times of the different  $M_w$  species present in

a disperse sample would result in a larger volume of distribution compared to a sample of a defined  $M_w$ . As mentioned above, due to a lack in the availability of defined  $M_w$  PEG of higher molecular weight, there is no research data available to support this theory, which underlines the importance of producing defined  $M_w$  PEGylation agents to lay the foundation for future pharmacological experiments on defined  $M_w$  PEGylated therapeutics.

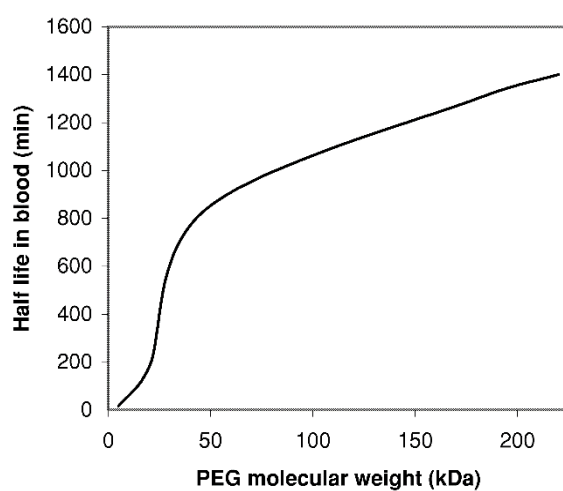


**Figure 8. Correlation of the molecular weight of linear, disperse PEG and the corresponding half life in blood after intravenous administration into mice.** Copied with permission from [44].

## **2.1.4 Purification and Characterisation of PEGylated Proteins**

### **2.1.4.1 Introduction**

The purification and full characterisation of PEGylated protein therapeutics is complex, but important to ensure the drug's efficiency and safety for the patient. Purification involves the separation of unreacted PEGylation agent, unreacted protein, and, in case of random PEGylation, the separation of proteins PEGylated with varying numbers of PEG molecules (so called PEGamers) or different positional isomers. Especially the latter is often difficult, as the different positional isomers have identical sizes and conventional size-based separation approaches

cannot be applied. An overview of the steps involved and techniques that can be used to separate and identify PEGamers and positional isomers of a PEGylated protein is shown in **Figure 9**.

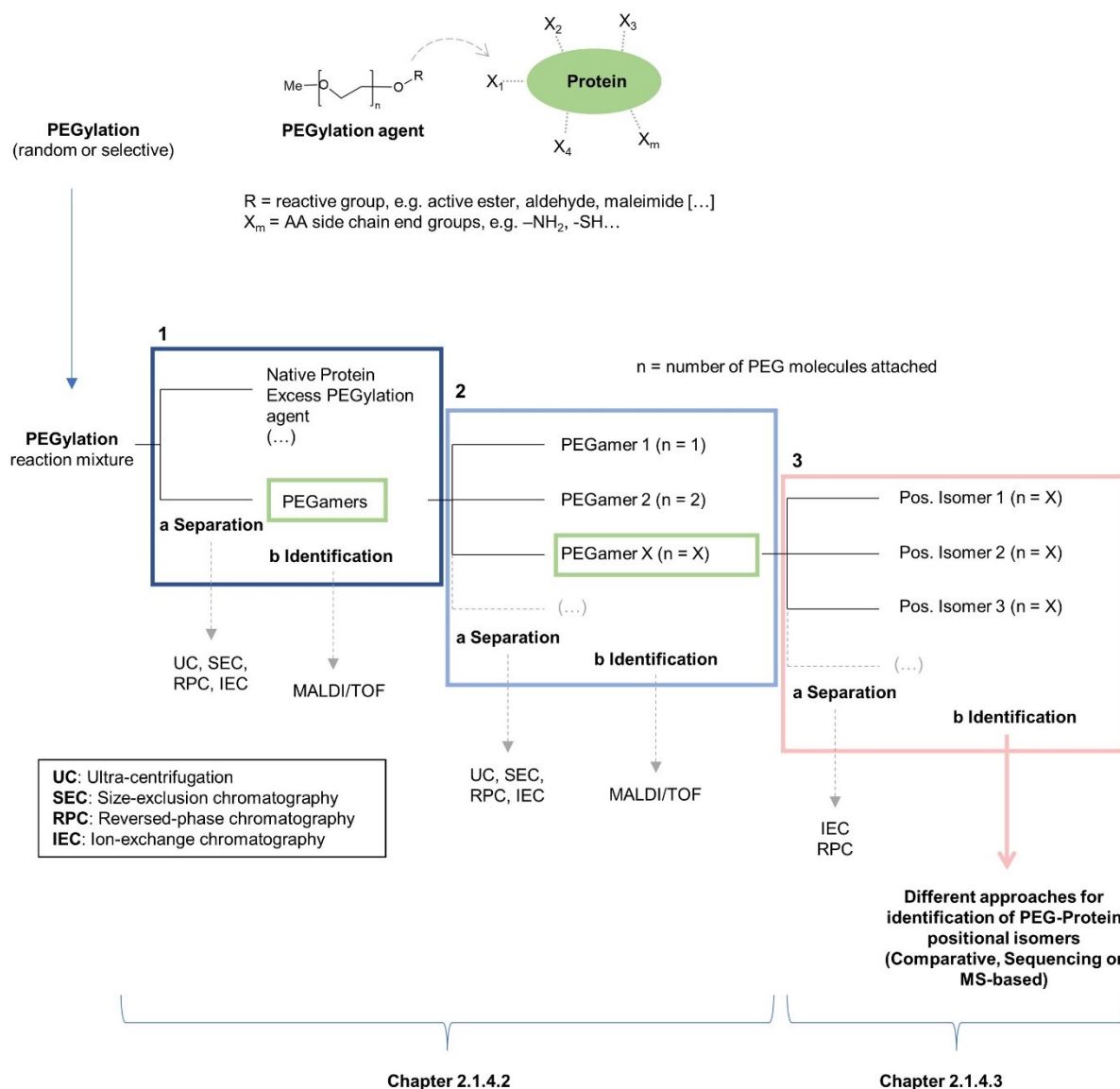


**Figure 9. Overview of the steps involved in the purification and characterisation of PEGylated proteins.** This involves **1**: the separation of the crude PEGylation mixture to separate the PEGamers (identification through MALDI-TOF), **2**: the separation of the different PEGamers through for example, size-exclusion or ion exchange chromatography, **3**: the separation of each PEGamer into the different positional isomers (usually through IEC or RPC) and the identification of each positional isomer through one of three different approaches described in chapter 2.1.4.3.

#### 2.1.4.2 Separation and analysis of PEGamers

PEGamers have been separated using a variety of chromatographic techniques, most importantly ion exchange (IE), size-exclusion (SE) and reversed-phase (RP) chromatography.<sup>29,58-61</sup>

One of the first techniques to be used on an analytical and preparative scale was size-exclusion chromatography (SEC), which separates different species based on their viscosity radius.<sup>29</sup> The viscosity radii of PEG in aqueous solutions are significantly higher than those of proteins of the same molecular weight. Hence, the viscosity radii of PEGylated proteins of a certain molecular weight differ from the viscosity radii of PEG or protein of the same molecular weight.

HAGEL found the following equation to describe the viscosity radius of a globular protein  $R_{h,prot}$ <sup>62</sup>:

$$R_{h,prot} \approx (0.82 \pm 0.02)M_r^{1/3} \quad (1)$$

FEE and van ALSTINE derived equation (2) to describe the viscosity radius of PEG, based on experimental data for the viscosity radii of spherical PEG of different molecular weights (0.2 kDa – 1200 kDa) in water published by KUGA.<sup>58,63</sup>

$$R_{h,PEG} \approx 0.1912M_{r,PEG}^{0.559} \quad (2)$$

They used equation (1) and (2) to prepare calibration curves for the size-exclusion of proteins and PEG of different sizes, showing that they follow the same trend.<sup>58</sup> The derived calibration curves were then used to calculate the viscosity radii for different PEGylated protein species. It was shown that the following equation (3) can be used to calculate the viscosity radius  $R_{h,PEGprot}$  of a PEGylated protein.<sup>61</sup>

$$R_{h,PEGprot} = \frac{A}{6} + \frac{2}{3A} R_{h,PEG}^2 + \frac{1}{3} R_{h,PEG} \quad (3)$$

where

$$A = [108 R_{h,prot}^3 + 8 R_{h,PEG}^3 + 12(81 R_{h,prot}^6 + 12 R_{h,prot}^3 R_{h,PEG}^3)^{\frac{1}{2}}]^{\frac{1}{3}} \quad (4)$$

Within these studies it was shown that the viscosity radius of a PEGylated protein does not depend on the PEGylation extent, but on the molecular weight of the protein and the total molecular weight of the PEG attached. I.e., a protein PEGylated with two units of 5 kDa PEG has the same viscosity radius as the same protein PEGylated with one single 10 kDa PEG molecule. FEE and VAN ALSTINE conclude that the PEG forms a dynamic layer around the protein, rather than extending into solution, and that way maintains a constant surface-to-volume ratio.

In general, the separability of PEGamers through SEC depends on the size of the protein and the PEG attached. For a specific combination of protein and PEG size, the separability is getting more difficult with increasing degree of PEGylation.<sup>64</sup> For example for their study on the viscosity radii, FEE and VAN ALSTINE randomly PEGylated  $\alpha$ -lactalbumin with a 5 kDa mPEG active ester.<sup>65</sup> SEC of the PEGamer mixture resulted in baseline-separation of the native protein and the mono-PEGylated species, whereas the higher molecular weight species were not fully separable. Overall, SEC can often be used efficiently to purify PEGylated proteins with a low degree of PEGylation (max.  $n = 3$ ) and low molecular weight PEG.<sup>29</sup>

In ion-exchange chromatography (IEC), species are separated based on their net surface charge and it is the most used technique to purify PEGylated proteins on an industrial scale.<sup>64</sup> As modifying an amine group with a PEG molecule changes the net surface charge of a protein, this technique is often used to separate PEGamers derived from amine-PEGylation.<sup>64</sup> But also for PEGylation techniques that do not directly affect the net surface charge, the PEG shields the surface charges on a protein in a way that it still has an effect on ion-exchange behaviour.<sup>29</sup> This

was first observed by AZARKAN *et al.* when they PEGylated the enzyme chymopapain at a thiol residue and discovered the different retention behaviour of the native protein and the PEGylated protein on an ion-exchange column, even though the isoelectric point remained unchanged.<sup>66</sup> SEELY and RICHEY made similar observations when PEGylating a protein at its cysteine residue using a PEG-vinyl sulfone. The mono-PEG species could easily be separated from the native protein and a protein-PEG-protein species, using cation-, or anion-exchange chromatography.<sup>67</sup> As the net surface charge was the same for all three species it was concluded that the separation was based on a charge masking effect of the PEG. The use of IEC on preparative scale is not always economically viable, as the systems have to be kept dilute to maintain sufficient separation and the columns tend to have a low lifespan.<sup>68</sup>

Hydrophobic-interaction chromatography has been mentioned in literature, but is not known to be implemented into analytical or preparative separation of PEGylated proteins.<sup>69</sup> Reversed-phase chromatography (RPC) is mostly used for analytically separating mixtures of PEGylated proteins, and is especially important for separating peptide mixtures within a peptide mapping process.<sup>61</sup>

Due to the drawbacks the described chromatographic techniques have in common (limited separation efficiency and low yields, lack in scalability), research is currently focusing on finding non-chromatographic approaches.<sup>33</sup> Those approaches include membrane separations, aqueous two-phase systems (ATPS), electrophoresis and capillary electrophoresis.<sup>60,64</sup>

Membrane filtration (ultrafiltration) is a promising, scalable technique to separate a PEGylated protein from native protein and excess PEGylation reagent based on their different sizes. For example MOLEK and ZYDNEY investigated the separation of a mixture of  $\alpha$ -lactalbumin PEGylated with a 20 kDa PEG-NHS ester, the native protein (14.2 kDa) and unreacted PEG using an Ultracel 30 kDa membrane and a negatively charged 100 kDa Ultracel membrane.<sup>70</sup> The first, uncharged membrane was shown to be suitable for removing the native  $\alpha$ -lactalbumin from the mixture. For the separation of the unreacted 20 kDa PEG from the PEGylated protein, the negatively charged

membrane was tested. It was shown that the rejection for the PEGylated protein was significantly increased at lower ionic strength, due to the stronger electrostatic repulsion of the negatively charged protein from the negatively charged membrane. The researchers managed to purify the mono-PEG protein with a two-stage system consisting of the two membranes in a yield of 75 %. Interestingly another study revealed that the correlation between the total molecular weight of the PEG attached to the protein and the resulting equilibration partition coefficient in SEC (discovered by FEE and VAN ALSTINE<sup>61</sup>), is not transferable to membrane filtration processes.<sup>59,61</sup> During membrane filtration the rejections seem to be dependent on the degree of PEGylation, for example a protein PEGylated with four linear 5 kDa PEG units would have higher rejection, than a protein PEGylated with two 10 kDa. Furthermore, another study has shown that the rejection of a PEGylated protein decreases with increasing flux and this effect is more pronounced with longer PEGs.<sup>71</sup> This phenomenon has been explained by an elongation of the flexible PEG chains along the fluid stream into the membrane pores. Ultrafiltration seems to be an efficient and scalable alternative to chromatography-based processes, but the rejection behaviours of the different species have to be taken into account and the process parameters adjusted accordingly. Furthermore, it is questionable if membrane filtration is useful for separating different PEGamers, as the hydrodynamic volumes might be too close to each other for matching with the currently reachable molecular weight cut-offs.

Aqueous two-phase systems (ATPS) form a class of liquid-liquid fractionation, in which two aqueous phases are formed, by adding different phase-forming compounds to them.<sup>64</sup> Usually these compounds are a combination of PEG and either another polymer (for example dextran) or a salt (for example phosphates or citrate) and the phase-formation is mostly based on hydrophobic repulsion.<sup>72</sup> Depending on their partition behaviour the different components of a PEGylation mixture tend to partition into one of the two phases. DELGADO *et al.* have shown a linear relationship between  $\log K$  ( $K$ : partition coefficient) and the degree of PEGylation.<sup>73</sup> ATPS can be a cheap and efficient tool to separate PEGylated proteins from native protein, but it is not suitable to separate different PEGamers.

Electrophoresis using polyacrylamide gels (polyacrylamide gel electrophoresis, PAGE) and capillary electrophoresis can both be used to separate PEGylated proteins for analytical purposes.<sup>60</sup>

Nowadays mostly site selective PEGylation is used to avoid problems related to the presence of different PEGamers, but not every protein allows for PEGylation at a specific site. For example, the selective PEGylation of the *N*-terminus using a PEG aldehyde derivative requires the accessibility of the *N*-terminal end of the peptide chain, which is sometimes buried. Furthermore, modification with only one PEG unit requires the usage of higher molecular weight PEG to reach the same viscosity radius that is reached with several shorter PEGs. The higher molecular weight PEG could potentially result in problems related to decreased clearance from the body.<sup>29</sup>

#### **2.1.4.3 Separation and identification of positional isomers**

Even more challenging than separating different PEGamers is the separation of different positional isomers of a PEGylated protein on an analytical or preparative scale. Even when a selective PEGylation technique is used, the PEGylation reagent can potentially react with other groups within the protein. A full purity analysis includes the identification of all positional isomers present and is crucial for a potential FDA approval.

Identifying positional isomers of a PEGylated protein is generally a challenging and laborious task, with only a limited number of techniques available. The inherent dispersity of PEGylation agents used leads to variations between batches and adversely affects the chromatographic behaviour of PEGylated proteins. Analysing them through simple UPLC-MS is not feasible due to the dispersity introduced through the PEG moieties.

The different approaches for identifying positional isomers within a sample of a PEGamer of a PEGylated protein are summarised in **Table A 1** and the included literature is discussed below. As described previously (**Section 2.1.4.2**) the first step of purification (on a preparative or analytical scale) after a PEGylation reaction is the separation of the different PEGamers. This is usually a straightforward process based on UC, SEC, IEC or RPC. The next step towards identifying

the positional isomers of a PEGylated protein is their separation. Due to their identical sizes, UC or SEC cannot be used for this purpose. IEC is the most widely used method to separate positional isomers, and there are some examples for successful application of RPC.<sup>68</sup> The subsequent identification of the positional isomers can be divided into three different approaches (**Figure 10 A, B and C**).

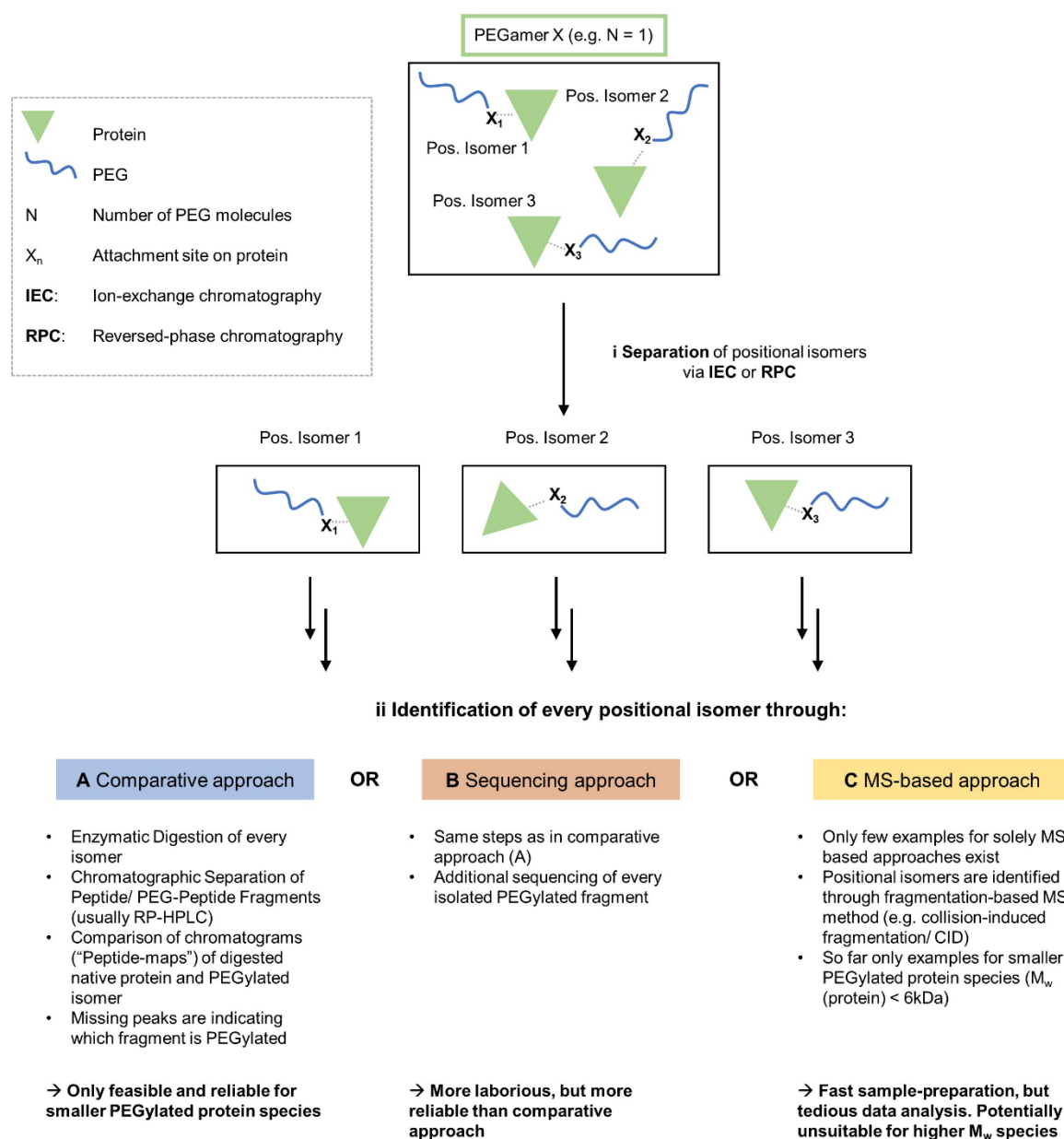


**Figure 10. Overview of the different approaches of identifying the positional isomers of a PEGylated protein.**

In 1996 CLARK *et al.* presented an early attempt at identifying PEGylation sites in a protein.<sup>74</sup> The researchers randomly PEGylated the 22 kDa human growth hormone (hGH) with a 5 kDa active ester and separated PEGamers with  $n = 2-7$  ( $n$  being the number of PEG molecules attached) for analytical purposes, using an ion-exchange sulfopropyl resin (SP-Sepharose). Fractions were collected and the PEGamers present in each sample were identified through MALDI-TOF. Pure samples of each PEGamer with  $n = 2-7$  were collected and subject to tryptic digestion. The resulting peptide mixtures were analysed through RP-HPLC-MS. When comparing the resulting chromatograms to the chromatograms of digested native hGH, the authors observed a decrease in certain peaks. They concluded these peaks to represent the PEGylated fragments. However, the authors did not provide any supporting mass spectrometric data, nor did they comment on the emergence of new peaks exhibiting the characteristic underlying mass spectrum for disperse PEG. The reliability of this comparative/ exclusion approach raises doubts, as PEGylation at one site of the protein may affect the accessibility of enzymes in other regions due to the tendency of PEG to envelop the protein and form a dense layer. This approach can be summarised as the comparative approach and is summarised in **Figure 10 A**.

KINSTLER *et al.* used a similar, yet more advanced strategy to prove the *N*-terminal PEGylation of the 18.8 kDa protein rhG-CSF (recombinant human G-CSF).<sup>75</sup> They initially PEGylated the protein using an acylation method and a 6 kDa PEGylation agent. Subsequently, they utilised IEC to separate the mono-PEG species from any native protein or higher molecular weight PEGamers. The researchers then enzymatically digested the native protein and the mono-PEG protein with EndoLysC and compared the resulting RP-HPLC chromatograms. The PEG-protein chromatogram displayed the absence of one peak that was present in the digest of the native protein, while simultaneously exhibiting a new peak. The fractions of both peaks were collected, and the sequence of the corresponding peptides were identified using amino-acid analysis an *N*-terminal sequencing. In another instance, they employed an alkylation-based method for PEGylation. They performed the same analysis as described above, but the resulting PEG-peptide fragment could not be analysed using the previously used sequencing methods. Hence, they treated the PEG-

peptide fragment with CNBr to cleave off the PEG residue with the terminal methionine Met(0). The identity of the remaining peptide fragment was then confirmed through MALDI-MS.

In conclusion the approach presented by KINSTER *et al.* is more laborious, but also more reliable, than the method previously used by CLARK *et al.* It is the first example for full characterisation, including sequencing of the PEGylated fragments and can be summarised as the sequencing approach (**Figure 10 B**). Nevertheless, they just applied it to a relatively small, mono-PEGylated protein and higher  $M_w$  PEGamers were separated out before further analysis.

MONKARSH *et al.* used a similar approach to analyse a more complex randomly PEGylated protein, resulting in more positional isomers.<sup>11</sup> They generated a mono-PEGylated derivative of the ~ 19 kDa protein Interferon  $\alpha$ -2a (IFN) based on a 5 kDa PEG and a urea linkage. The researchers then developed a method they termed an “ampholyte-free chromatofocusing-like method”. This technique involved employing a cation-exchange resin with a pH-gradient. Remarkably, they achieved successful separation of 11 distinct positional isomers of the mono-PEGylated IFN using this approach. Subsequently samples of the 11 collected fractions were subjected to enzymatic digestion (either Lys-C or Glu-C) and the resulting mixtures of peptides and PEG-peptides were separated using RP-HPLC. The peaks corresponding to PEGylated peptide fragments were identified through the characteristic underlying mass spectrum of disperse PEG, where the peaks have a mass difference of  $n \times 44$  Da. Fractions of the peaks identified as PEG-containing were then collected and subjected to Edman Sequencing. This way the exact position of every PEGylation site was determined.

There are some more recent attempts of identifying the positional isomers within a PEGylated protein sample solely based on mass spectrometry (**Figure 10 B**).<sup>56,76,77</sup> For example, Yoo *et al.* used a method involving MALDI-TOF MS in combination with reflector in-source decay (reISD) analysis to identify the positional isomers of a 3.5 kDa peptide PEGylated with a 20 kDa linear PEG.<sup>76</sup> The comparative interpretation of the MS data is complex and based on the published literature, these approaches also seem to be limited to small peptide or protein species (< 6 kDa).

In summary the identification of positional isomers of PEGylated proteins is a complex and often work-intensive process, especially if multiple PEGylation sites are present. As the full characterisation of a PEGylated therapeutic is crucial for assuring its purity and safety for the patient, there is a high interest in finding faster and easier techniques for this purpose.

#### **2.1.4.4 An example- the FDA report of the Neulasta biosimilar MYL-1401H**

There is a scarcity of accessible information regarding the analytical methods used industrially to analyse currently FDA-approved PEGylated drugs, primarily due to patent protection preventing the public release of most analytical reports. However, in the case of the blockbuster drug *Neulasta* (pegfilgrastim), its patent expired in the United States in October 2015. By 2023, the FDA had approved five biosimilars for *Neulasta*. For one of them (Fulphila, MYL-1401H) the results of an extensive analytical comparison with the previously marketed *Neulasta* is publicly available.<sup>55</sup> The report gives an example for how PEGylated therapeutics are analysed within an FDA approval process and will be described more detailed in the following.

*Neulasta* and its biosimilars consist of the 18.8 kDa protein granulocyte colony stimulating factor (G-CSF) which is selectively PEGylated at its *N*-terminus using a methoxy-PEG-aldehyde derivative with an average weight of 20 kDa. Within the FDA product review of MYL-1401H, its analytical similarity to the U.S.-approved and the EU-approved versions of *Neulasta* (US *Neulasta* and EU *Neulasta* in the following) is investigated using different analytical techniques.

Firstly, MALDI-TOF is applied to determine the intact mass of different batches of the three therapeutics (**Figure 11a**). Besides the peak observed at about 40 kDa, there are two more peaks at ~ 60 kDa and ~ 80 kDa. The report states that these are the di-PEGylated version (~60 kDa) and the dimer of two mono-PEGylated proteins (~ 80 kDa). It does not become clear from the available data, how the ~80 kDa species is identified as the aggregate and distinguished from a potentially present tri-PEGylated protein. It is important to note, that the mass determined from MALDI-TOF is only the maximum of each peak, which does not account for the dispersity introduced through the PEGylation agent. A scatter plot of the intact masses of samples of

different batches of the three derivatives (**Figure 11b**) shows the variation between the batches. The average mass for EU Neulasta deviates between  $\sim 40\,000$  Da and  $\sim 40\,800$  Da for the analysed batches.

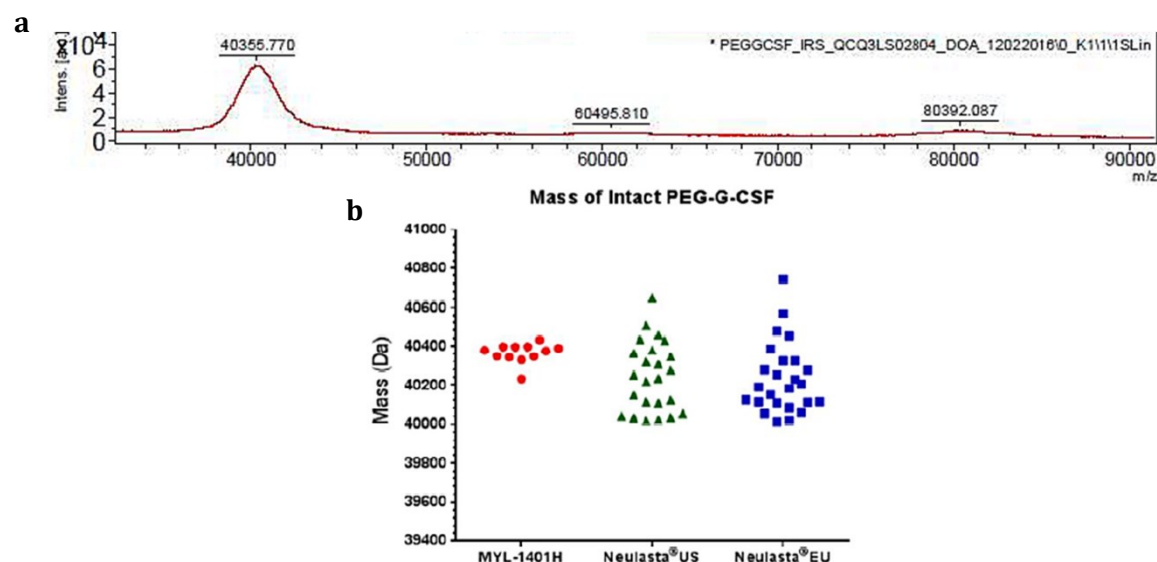


**Figure 11. Analytical data from product review report for *Neulasta* biosimilar MYL 1401H. a,** exemplary MALDI-TOF spectrum of MYL-1401H. **b,** Scatter plot of MALDI-TOF results of different samples of MYL1401H, U.S.-licensed *Neulasta* and EU-approved *Neulasta*. Copied with permission from [55].

To confirm the *N*-terminal PEGylation, a laborious combination of three different peptide mapping approaches, MALDI-TOF and MS-MS was utilised. Firstly, a sample of MYL-1401H, as well as the EU- and the U.S.-approved *Neulasta*, were digested using the enzyme Glu-C. This enzyme cleaves at the *N*-terminal side of aspartic acid and glutamic acid. The resulting peptide mixture was then subjected to UPLC-MS analysis and one peak was suspected to be the PEGylated *N*-terminus (**Figure 12a**). This was based on the observation of a characteristic difference of 44 kDa between each peak in the underlying ESI+ trace (**Figure 12b**). It is worth noting that there was a discrepancy in retention time for this peak between the two shown samples (32.3 min for MYL-1401H vs. 32.4 min for US *Neulasta*). The sponsor attributed this difference to variations in the molecular weight of the PEG, which emphasises the inconsistency in the distribution of PEG molecular weights among different batches.

To verify the molecular weights corresponding to the peak suspected to be the *N*-terminal PEG-peptide fragment, the respective fraction was collected and analysed off-line using MALDI-TOF. The same procedure was also applied to samples digested with the enzyme trypsin, which cleaves at arginine and lysine residues of the protein. The calculated theoretical masses for the PEGylated peptide fragments were 22 133 Da (Glu-C digestion) and 21 788 Da (trypsin digestion). The MALDI-TOF analysis of the PEG-peptide fragment of the different samples revealed a mass range of 24 057 Da – 24 195 Da for the Glu-C digested samples and 23 793 Da – 23 909 Da for the trypsin digested samples. The reviewers considered these masses to be sufficiently close to the theoretical values to provide support for the presence of *N*-terminal PEGylation. However, it is crucial to note that there is a possibility of another positional isomer with the same retention time and a PEG-peptide fragment mass close to the suspected *N*-terminal PEG-peptide fragment. This alternative isomer would not be detectable using the methods described above. Hence the sponsor provided data on a third, double digest method. In this, the peptide-PEG fragment is treated with cyanogen bromide (CNBr) and trypsin. CNBr cleaves at the C-terminus of methionine (Met), which constitutes the *N*-terminus of filgrastim. This cleavage results in a PEG-Met fragment and the remaining peptide fragment. The latter is subsequently analysed through LC-MS-MS, which analyses the composition of the peptide fragment through collision induced dissociation (CID). According to the reviewers, the data obtained confirms that the only lysine residue in this fragment (K<sub>17</sub>) is not PEGylated. However, the basis for this conclusion is not entirely clear, as a mixture of *N*-terminal PEGylated and K<sub>17</sub>-PEGylated filgrastim could yield similar results as those presented in the report.

In summary, the comprehensive evaluation of the product quality report for MYL-1401H highlights the challenges associated with identifying positional isomers and demonstrating the purity of a PEGylated protein. This difficulty persists even in the case of a relatively small version like MYL-1401H, where a selective PEGylation method was employed and can be attributed to the PEGs dispersity.

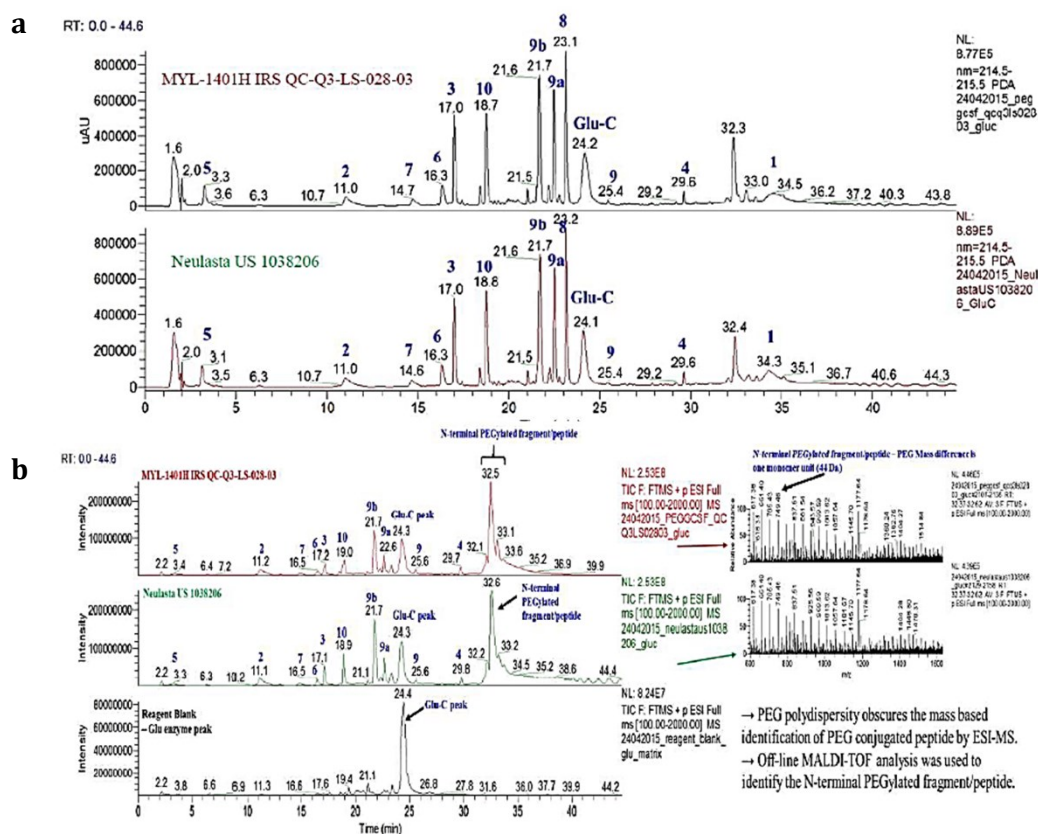


**Figure 12. Results of the Glu-C digestion of a sample of MYL-1401H and the U.S.-approved *Neulasta*.**  
**a**, UV trace of the UPLC chromatogram (bottom: *Neulasta* US, top MYL-1401H). The peaks at 32.3 min and 32.4 min were identified as PEG species. **b**, Underlying ESI traces for the same chromatograms. The ESI (right side) shows the characteristic PEG pattern with a difference of 44 kDa between each peak. Copied with permission from [55].

## 2.2 Synthesis of defined PEG and PEGylation agents

### 2.2.1 Introduction

The conventional industrial process to produce large amounts of PEG, used for example for detergents or in cosmetics, is the anionic polymerisation of ethylene oxide (**Figure 13**).<sup>7</sup>

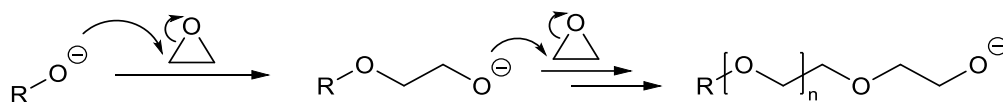


**Figure 13. Mechanism of anionic polymerisation of ethylene oxide to synthesise poly(ethylene glycol) (PEG).** Ethylene oxide undergoes nucleophilic attack by an alkoxide, which results in ring opening and bond formation between the alkoxide oxygen and the ethylene oxide -CH<sub>2</sub>- group.

This polymerisation technique, requiring a nucleophilic initiator and a basic catalyst, results in a Gaussian distribution of different chain length and can produce PEG up to a M<sub>w</sub> of 50 kDa.<sup>7</sup> Heterobifunctional PEGs for more sophisticated applications, for example the PEGylation of pharmaceuticals, can be achieved by either using an initiator bearing the desired function, a functionalised terminating agent or post-synthesis modification of the synthesised PEG diol.<sup>78-84</sup> To achieve PEG derivatives with a low dispersity, the products obtained by anionic polymerisation need extensive chromatographic purification. For example, the company Polypure AS is using low-pressure sample displacement chromatography to increase the homogeneity of low-molecular weight, disperse PEG and that way reach PEG of high chain length purity.<sup>85-87</sup> Sample displacement chromatography is based on the relative affinities of different sample components with towards the stationary phase of a column and can be performed in different modes, like for instance a reversed-phase or ion-exchange set-up.<sup>87-89</sup> To increase the homogeneity of the final product, the sample is usually subjected to several fractionation steps. For example, Polypure AS isolated an Eg<sub>8</sub> oligomer from a disperse PEG400 sample and reached a oligomer purity of 99.2 % after three fractionation steps.<sup>90</sup>

The longest defined PEG derivative available from Polypure AS is a diol with 28 Eg repeating units (Eg<sub>28</sub>).<sup>91</sup>

Low-dispersity (“monodisperse”) PEGylation agents are only commercially available from a few suppliers, limited to a molecular weight of up to about 1.5 kDa and are expensive (for example, \$430 for 0.5 g of MeO-Eg<sub>36</sub>-NH<sub>2</sub>, as of 26/12/2023).<sup>92</sup>

Due to the inefficiency of the available chromatography-based techniques and their limitations regarding higher molecular weight PEG, current research is focused on finding iterative synthesis solutions and efficient and scalable purification methods. The following sections of the literature review will critically review the definition of the term “dispersity” and its corresponding parameters and give an overview of the different methods to synthesise polymers (specifically poly(ethylene glycol)) of defined mass and chain length.

### 2.2.2 Dispersity

The term “dispersity” is most used in the field of polymer sciences to describe the extent of distribution of the molar mass or degree of polymerisation of the macromolecules within a polymeric sample. Over a long time, the term “polydisperse” has been used for samples with a broad distribution of different chain lengths/masses and the term “monodisperse” has been used for describing low dispersity. IUPAC has recommended to substitute these contradictory terms with the use of high and low dispersity.<sup>93</sup> The index for the molar mass dispersity ( $\mathcal{D}_M$ ) can be mathematically derived from the following equation.

$$\mathcal{D}_M = \frac{\bar{M}_w}{\bar{M}_n} \quad (5)$$

$\bar{M}_w$ : mass-average molar mass;  $\bar{M}_n$ : number-average molar mass

$\bar{M}_w = \frac{\sum_j N_j M_j^2}{\sum_j N_j M_j}$  and  $\bar{M}_n = \frac{\sum_j N_j M_j}{\sum_j N_j}$  with  $N_j$  being the number of molecules of the molecular mass  $M_j$ .

In this work the term “defined  $M_w$ ” will be used for polymers consisting only of one defined chain length.

It has been common to define the “(poly)dispersity index” (previously PDI, now  $D_M$ ) as a measure for the width of the distribution of different molecular weights within a polymeric sample. It is important to note that the dispersity index does not describe the absolute width of the distribution, but should only be used as a value relative to the number-average molar mass.<sup>94,95</sup> This can be explained by the following mathematical correlations. The absolute width of a probability function can be described as the variance of distribution, which in case of molecular weights can be mathematically described as:

$$\sigma_d^2 = \bar{M}_n^2 \left( \frac{\bar{M}_w}{\bar{M}_n} - 1 \right) \quad (6)$$

The second term of equation (6) (in brackets) describes the deviation of the ratio of weight-average molecular weight and number-average from the value 1, so the larger the dispersity index, the broader the distribution. The first part ( $\bar{M}_n^2$ ) shows the additional contribution of the number-average molecular weight, which underlines the fact that the width of two distributions with identical dispersity indices do not have to be the same.

In summary only the dispersity indices of polymer samples of similar number-average molecular weights should be compared, or the experimental data must be normalised before interpretation. Furthermore, the dispersity data for polymers derived from different synthetic methods (following different probability functions) cannot be compared with each other.

Another factor affecting the reliability of the dispersity data is the underlying analytical method. Typically, gel permeation chromatography (GPC) or size exclusion chromatography (SEC) are employed to assess the dispersity of a polymer sample in combination with a light-scattering detector such as a refractive index detector (RID) or evaporative light scattering detector (ELSD).<sup>96,97</sup> The reliability of the data depends on different factors, for example the purity of the

used calibration standard employed, the solvent purity or the column.<sup>96,97</sup> GPC/SEC are relative methods, so they must be calibrated using a standard of known molecular weight. The calibration standard must exhibit the same chemical structure as the polymer to be analysed, as the hydrodynamic radius in different solvents is different for each type of polymer. Furthermore, it is crucial to choose the right column with a suitable packing material (to avoid interactions of the sample with the column) and pore size distribution. Conclusively, the dispersity data for different polymer samples can only be compared with each other if the polymer samples are of the same type (same underlying synthetic method) and, if the data was derived using the same analytical conditions and data processing.

Taking into account the considerations described above, the dispersity index might still be a useful parameter to compare the dispersity of polymer samples of the same nature.

Defined  $M_w$  polymers are obtained through step-wise/ iterative synthesis approaches (described in **Section 2.2.3**) and the resulting molecular weight distributions do not follow a certain continuous probability function, as each chain elongation is composed of different reaction- and purification-steps that contribute independently to the final chain length purity. **Figure 14** shows

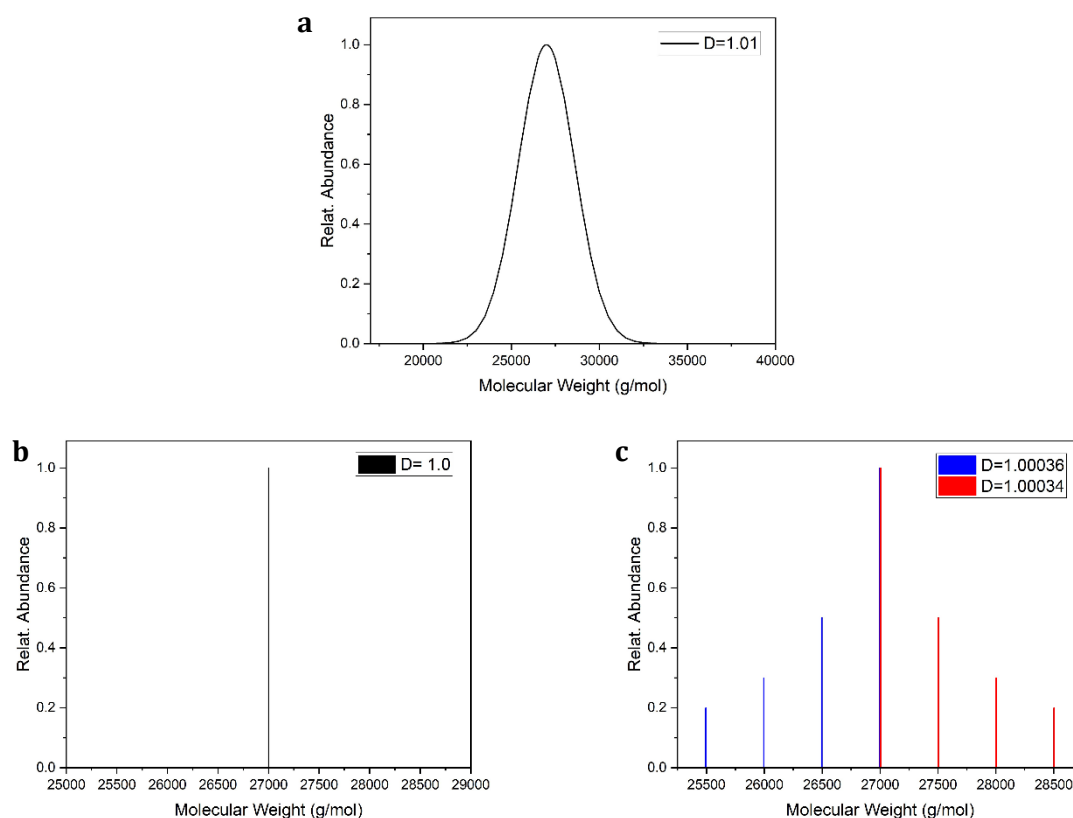


**Figure 14. Simulated molecular weight distributions for a theoretical polymer with a median molecular weight of 27 kDa.** **a**, A theoretical normal distribution of molecular weights as it is usually derived from conventional polymerisation reactions (such as the anionic polymerisation of ethylene oxide) with a dispersity index of 1.01; **b**, The theoretical molecular weight distribution for an ideal single mass polymer with an exact mass of 27 kDa and a dispersity index of 1.0; **c**, Two theoretical, symmetrical distributions for polymers derived from iterative processes, with the same width. The distribution with the lower average  $M_w$  (blue,  $D = 1.00036$ ) has a higher dispersity index than the distribution with the higher average  $M_w$  (red,  $D = 1.00034$ ).

simulated molecular weight distributions for a theoretical polymer of a median molecular weight of 27 kDa. **Figure 14a** shows a typical normal distribution, as it is usually derived from controlled polymerisation techniques, like the anionic ring-opening polymerisation of ethylene oxide to reach PEG. Even though the dispersity index for this distribution is low (1.01) and would be

described as “monodisperse”, the distribution covers a wide range of molecular weights (~22.5 kDa – 32.5 kDa). **Figure 14b** shows the molecular weight data for an ideal single-mass polymer with a mass of 27 kDa and a dispersity index of 1.. **Figure 14c** shows the molecular weight distributions of two theoretical polymers derived from an iterative synthesis approach. The distributions are mirrored symmetrically, but still the lower molecular weight distribution (blue) has a higher dispersity index (1.00036), than the higher molecular weight distribution (red, D = 1.00034). This represents the previously discussed relative dependency of the dispersity index on the absolute molecular weights present in the polymer sample. Furthermore, the actual chain length purity of the presented samples (given that the desired molecular weight is 27 kDa) is only 50 %, which is not represented by the low dispersity indices. Besides that, it is questionable if GPC/SEC would be a suitable technique to accurately measure the dispersity index of a defined  $M_w$  sample, as the results might be more a representation of the pore-size of the used columns, than of the polymers molecular weight distribution.

In general, for defined  $M_w$  polymers it is desirable to find analytical methods to separate and quantify the different molecular weight species and to then calculate the chain length purity  $P$  as:

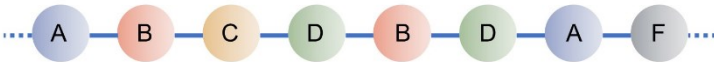
$$P = \frac{N_x}{\sum(N_x + N_i)} \times 100 \% \quad (7)$$

with  $N_x$  being the molar amount of the desired species  $X$  in the sample and  $N_i$  being the molar amount of the undesired chain length impurities in the sample.

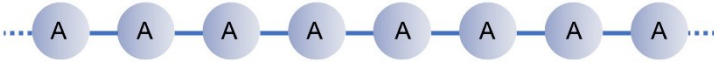
The difficulties related to a lack in analytical techniques for identifying these parameters (especially for higher molecular weight polymers), will be discussed later (**Sections 4.3.2.1 and 4.3.3.2**).

In general, a defined Mw oligomer or polymer (DMO/DMP) is equivalent to a defined sequence oligomer or polymer (DSO/DSP), where the sequence is composed of a single kind of repeating unit. The synthetic approaches developed for DSPs can also be used for producing DMPs (not always the other way around). The definition of “oligomers” and “polymers” changes depending on the chemical nature of molecule that is referred to. For example, in case of peptides and proteins, peptides are commonly defined as amino acid (AA) sequences of 2-50 AAs, and AA sequences longer than this are referred to as proteins.<sup>98</sup> For the sake of simplicity and because the work in hand is focusing on the synthesis and application of defined M<sub>w</sub> poly(ethylene glycol), the terms will be limited to “polymer” (DMP/ DSP) in the following.

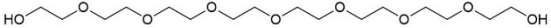
**a** Defined **Sequence** Oligomer/ Polymer



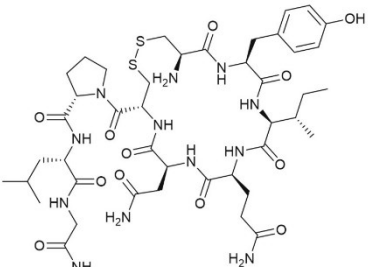
**b** Defined **Weight** Oligomer/ Polymer



**c**



**d**



60

amino acids, for example the peptide hormone Oxytocin, **Figure 15d**) or DNA and RNA (defined sequences of nucleotides).<sup>99</sup>

There is growing research interest in synthesising DSPs efficiently, especially in the fields of information-containing polymers<sup>100-104</sup>, and peptide- or oligonucleotide-therapeutics<sup>105-108</sup>. A relatively new field is the self-assembly of macromolecules into sophisticated 3-dimensional structures.<sup>109</sup> A prominent example for this is the controlled self-assembly of DNA strands to form complex engineered 3D-structures ("DNA origami").<sup>110,111</sup>

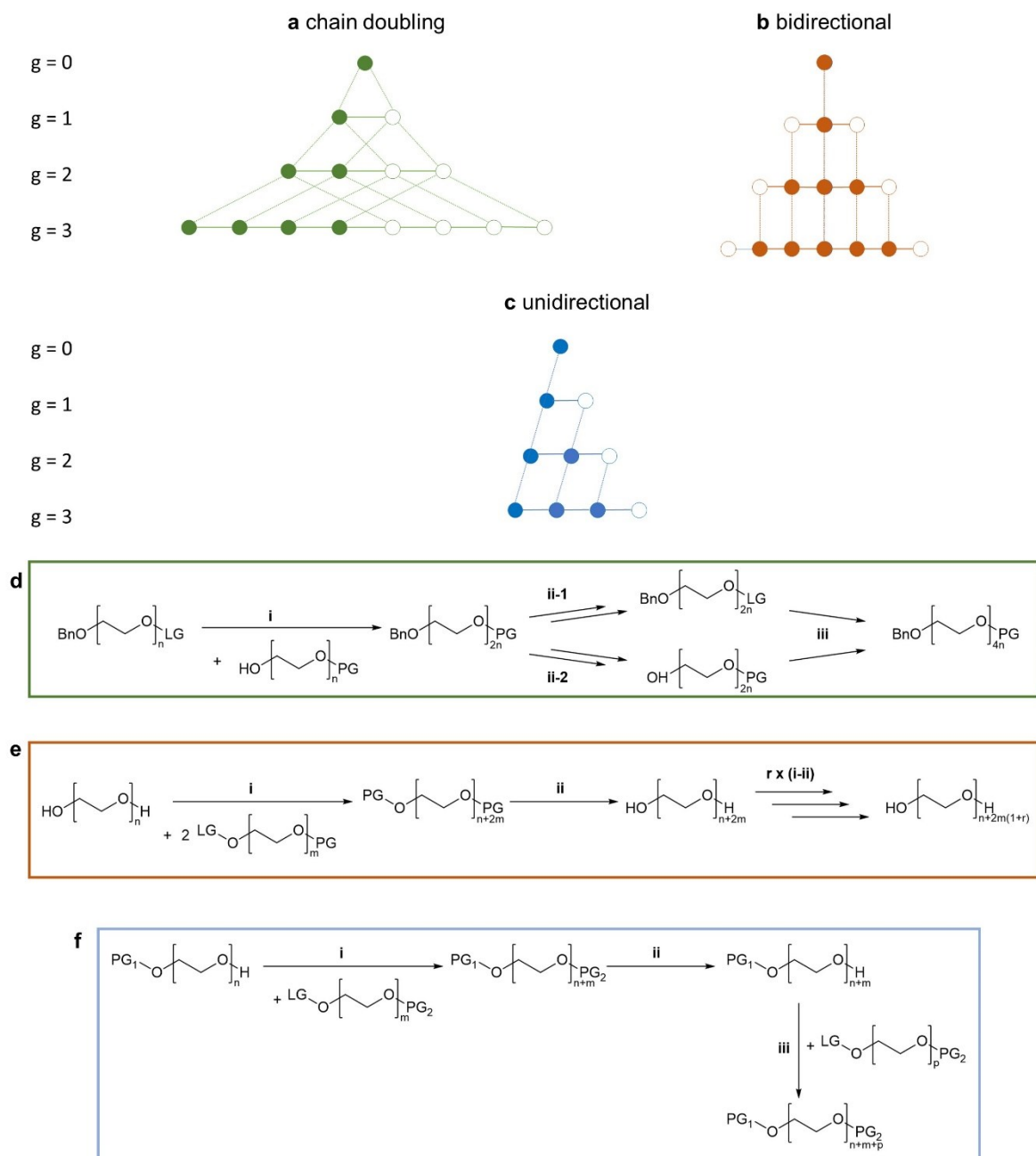
DMPs are polymers with a homogenous backbone and an exactly defined chain length/ molecular weight (**Figure 15b** and **c**). DMPs recently got more attention for replacing their disperse derivatives in drug-delivery system (for example to mask lipid-carriers for RNA-based therapeutics).<sup>112</sup>

The synthesis of defined polymers is challenging and there is intense research conducted in this field. In addition to achieving precise control over the polymer's sequence and molecular weight, challenges also arise in reaching a high chemical purity of the end product and ensuring scalability of the processes. In general, to reach a defined  $M_w$  or defined sequence polymer, either orthogonal growth approaches or those based on protecting group chemistry have to be employed to build up the polymer in an iterative manner.

#### **2.2.4 Early approaches for the synthesis of defined $M_w$ PEG**

The following section of the literature review deals with the growth approaches based on protecting group chemistry that were developed in the past with the aim of synthesising defined  $M_w$  PEG.

These methods can be divided into the groups of convergent chain doubling or tripling (iterative exponential growth), bidirectional iterative growth, and unidirectional iterative approaches (**Figure 16a-c**).<sup>113</sup>



Bn: benzyl, PG: protecting group, LG: leaving group

**Figure 16. Different strategies to achieve sequence-defined polymers.** Each circle stands for one building block unit. **a**, chain doubling; **b**, bidirectional iterative growth; **c**, unidirectional iterative growth; not depicted: chain tripling; **d**, an example for the chain doubling approach; **e**, an example for the bidirectional growth method; **f**, an example for the unidirectional growth method. Reproduced with permission from [13].

In case of the chain doubling approach, two orthogonally monoprotected, PEG based building blocks (BBs) are synthesised and coupled to each other (**Figure 16 a and d**). Through orthogonal deprotection, two bigger BB species for the next coupling are reached, which then are again coupled and deprotected in two different batches to reach the next BB species.

This chain extension process results in an exponential growth of the chain (rather than linear) and is limited to the synthesis of a polymer with an alternating sequence. The first example of using a chain doubling approach for the synthesis of defined  $M_w$  PEG was published by LOISEAU *et al.* in 2004, who synthesised  $Eg_{24}$  by chain doubling of a previously synthesised dodeca(ethylene glycol) BB.<sup>12</sup> In 2009, FRENCH *et al.* achieved a highly pure  $Eg_{48}$  by consecutive chain doubling starting from  $Eg_4$  Diol.<sup>13</sup> An exemplary scheme for the basic chain doubling approach used by FRENCH *et al.* is given in **Figure 16 d**. Chain doubling or tripling requires various desymmetrisation steps within a synthesis to assure the formation of one specific chain length. They still result in PEG diols, which need further desymmetrisation to be converted into valuable heterobifunctional PEG structures that can be used for PEGylating therapeutics.

Another way of synthesising discrete PEGs is the bidirectional growth method, which results in a symmetric extension of the chain at both ends (**Figure 16 b and e**). This starts from a bifunctional starting material and the chain is extended successively in both directions by adding protected BB, followed by deprotection and the next chain extension. By using this method and starting from a di-tosylated core molecule, KEEGSTRA *et al.* synthesised  $Eg_{12}$  Diol starting from di- and octa(ethylene glycol).<sup>114</sup> Starting from  $Eg_8$  and attaching an  $Eg_4$  to both ends in a stepwise manner, AHMED *et al.* reached  $Eg_{44}$  Diol in an overall yield of 17 %.<sup>115</sup> Equivalent to the chain doubling and tripling approaches, bidirectional growth methods result in a homobifunctional PEG. Desymmetrisation techniques, especially on longer PEG chains result in significant yield losses, as the different species have to be purified using column chromatography.

The only approaches that can be used to synthesise DSPs and result in heterobifunctional polymers are based on unidirectional iterative growth (**Figure 16 c and f**).<sup>12,13,15-18,116-121</sup> A first

example for the synthesis of a defined  $M_w$  heterobifunctional PEG was published by HILL and LOISEAU in 2004.<sup>12</sup> They synthesised unsymmetric PEG BBs based on tri- and hexa(ethylene glycol) and a mono-protected PEG alcohol. By successively reacting the mono-protected PEG-alcohol with the building blocks (**Figure 16f**) they reached the first heterobifunctional  $Eg_{24}$ .

It is important to note, that all of the early attempts for synthesising defined  $M_w$  PEG were exclusively based on bidirectional growth methods, resulting in homofunctional products up to  $Eg_{48}$ .<sup>14,114,115,122-126</sup> Furthermore, these processes, as well as the above described unidirectional growth methods, were relying on laborious chromatographic techniques between each chain extension step to purify the growing chain from the remaining reaction debris. Because most PEGylation applications require heterobifunctional PEGs of higher molecular weight (5 kDa and higher), an increased focus has been placed on the development of desymmetrisation techniques and unidirectional growth methods as well as replacing chromatography with cheaper and scalable means of purification.<sup>12,13,15-18,116-121</sup>

### 2.2.5 Novel approaches for synthesising defined $M_w$ PEGs

When synthesising a DMP/DSP, the critical factors for the purity of the final product and the economic efficiency of the process are the reaction kinetics and the purification technique that is used to separate the growing polymer from remaining reaction debris after each step.

The most common method to industrially synthesise precise polymers is a solid-phase approach, initially developed by Bruce Merrifield in 1963 in order to synthesise sequence-defined peptides.<sup>127,128</sup> In this method a peptide is bound to solid resin and reacted with a protected amino acid. After deprotection the remaining reaction debris is removed through filtration and the chain is iteratively grown by repeating the cycle of chain extension, deprotection and filtration. Solid-phase processes are well-established in industry to synthesise peptides and oligonucleotides, due to the efficiency of the purification method and a high degree of automation. Recently a solid-phase approach for the synthesis of defined  $M_w$  PEG was presented by KHANAL and FANG, who first used a Wang resin to synthesise heterobifunctional PEG up to  $Eg_{12}$ .<sup>17</sup> Fang *et al.* later used the

same approach implemented into a modified peptide synthesiser to synthesise defined  $M_w$  PEG up to  $Eg_{20}$  in a fully automated manner.<sup>18</sup> Despite the clear advantage of a high degree of automation, solid-phase approaches suffer from poor reaction kinetics, which are impeded by the mass-transfer from bulk to the resin pores.<sup>113,129-131</sup> This might become problematic when attempting the solid-phase synthesis of higher  $M_w$  defined PEGs. Furthermore solid-phase approaches are restricted to a limited choice of solvents and coupling chemistry<sup>132</sup>, and cannot easily be monitored. Finally, the scale-up of processes based on solid-phase is often difficult, due to swelling of the resin and the resulting instability of the packed bed.<sup>113,129-131</sup>

In contrast, liquid-phase processes exhibit a high homogeneity of the reaction mixture, wider possibilities for real-time monitoring and the (depending on the used method for purification) feasibility of large-scale processes. The more recent liquid-phase approaches for synthesising defined  $M_w$  PEG offer solutions for the problem of desymmetrisation, by using iterative unidirectional growth methods. Nevertheless, they mostly still use chromatographic processes for purification after each synthesis step, which limits the scalability, has a high consumption of solvents and often results in significant yield losses.<sup>133</sup>

**Table 2.** Overview of recently published approaches for synthesising defined  $M_w$  PEG (since 2014).

| Authors (Year) <sup>Citation</sup>            | Growth Method                           | Separation Method        | Type/ length of PEG                       |
|-----------------------------------------------|-----------------------------------------|--------------------------|-------------------------------------------|
| Livingston <i>et al.</i> (2014) <sup>15</sup> | Iterative unidirectional (liquid phase) | Column Chromatography    | Heterobifunctional <b>Eg<sub>56</sub></b> |
| Jiang <i>et al.</i> (2014) <sup>117</sup>     | Iterative unidirectional (liquid phase) | Solid-phase extraction   | Heterobifunctional <b>Eg<sub>20</sub></b> |
| Jiang <i>et al.</i> (2015) <sup>118</sup>     | Iterative unidirectional (liquid phase) | Column Chromatography    | Heterobifunctional <b>Eg<sub>64</sub></b> |
| Kinbara <i>et al.</i> (2016) <sup>121</sup>   | Iterative unidirectional (liquid phase) | Liquid-liquid extraction | Heterobifunctional <b>Eg<sub>16</sub></b> |
| Fang & Khanal (2017) <sup>17</sup>            | Iterative unidirectional (solid phase)  | Filtration               | Heterobifunctional <b>Eg<sub>12</sub></b> |
| Fang <i>et al.</i> (2022) <sup>18</sup>       | Iterative unidirectional (solid phase)  | Filtration               | Heterobifunctional <b>Eg<sub>20</sub></b> |

In 2014 SZEKELY *et al.* presented a new unidirectional synthesis approach, in which PEG chains were iteratively grown from a 3-armed, benzylic nanostar (Hub) using an  $Eg_8$  building block. At

the time, this was the longest defined, heterobifunctional PEG derivative (Eg<sub>56</sub>), which was synthesised in a 12 % overall yield on a gram scale.<sup>15,16</sup> The researchers used column chromatography as means of purification, but mentioned that the size of the PEG-nanostar species potentially would allow for replacing column chromatography with membrane-separation. This is further discussed later in this review. Alternatives to the conventional orthogonal protecting group chemistries have been developed, for example an approach based on macrocyclic sulphates, resulting in the up to this work longest defined, heterobifunctional PEG with 64 ethylene glycol repeating units.<sup>119</sup> This and most of the other approaches are still relying on tedious chromatographic techniques for separating the growing PEG molecule from the reaction debris after each chain elongation step. Problems include low yields, increasing separation difficulties with increasing chain length, and poor scalability.

To synthesise defined, heterobifunctional PEG building blocks up to Eg<sub>16</sub> in high overall yield (62% for HO-Eg<sub>16</sub>-OTs), KINBARA *et al.* recently presented a method solely based on extraction, which has the potential to be scaled up to an industrial scale.<sup>121</sup> Nevertheless this method is only applicable up to a chain length of 16 Eg units and therefore only allows for the large scale synthesis of shorter building blocks.

In summary, the efficiency and economic viability of liquid-phase approaches are largely dependent on the purification method used to separate the growing polymer from the remaining reaction debris after each extension step. Hence the focus of research has recently been on finding more efficient and economically viable purification techniques to replace conventional column chromatography in liquid phase approaches. A highly promising approach is the application of organic solvent nanofiltration (OSN). This describes using a solvent-stable, selective membrane, that has a pore size which allows for permeation of smaller molecular species, whilst rejecting bigger molecular species (like a growing polymer). The theory of organic solvent nanofiltration is described in more detail in **Section 2.2.6**.

In general, a combination of liquid-phase chemistry and membrane-separation combines the advantages of liquid-phase reaction kinetics and the possibility for real-time monitoring, whilst also allowing for automation and easy scale-up of the process. Hence, it is of high significance to optimise this process and investigate its applicability to synthesise defined higher  $M_w$  polymers, such as defined  $M_w$  PEG. In the following section of the literature review, the foundations of organic solvent nanofiltration and nanostar sieving will be discussed in more detail.

### 2.2.6 Organic Solvent Nanofiltration

Organic Solvent Nanofiltration (OSN) is a separation process that has gained significant interest in recent years due to its potential applications in various industries, including the pharmaceutical industry, the petrochemical industry and the production of fine chemicals.<sup>134</sup> OSN involves the separation of solutes and/or solvents from a mixture using a membrane with selective permeability towards one or more solutes and/or the solvent.<sup>131</sup> OSN is a filtration technique that operates in the nanometre range and uses solvent-stable membranes, making it ideal for the separation of small solute molecules from organic solvents. Filtration methods can be categorised into several distinct processes based on the size of the retained solute, the molecular weight cutoff (MWCO) of the membrane, the operating pressure, and the underlying transport mechanism through the membrane (**Figure 17**). For example, microfiltration employs porous membranes that typically have MWCO values above 100,000 g mol<sup>-1</sup> and operate at pressures below 1 bar. This method is useful for separating relatively large molecular species, such as bacteria or pigments. In ultra filtration (UF) processes, operating pressures typically range between 1 and 5 bar. The membranes have MWCO values that allow the retention of molecules with sizes ranging from 2 to 100 nm. Applications include the separation of viruses and proteins. OSN falls within the category of nanofiltration category, where operating pressures range from 5 to 40 bar.<sup>131</sup> NF can be used to separate molecular species at a nanometre scale,

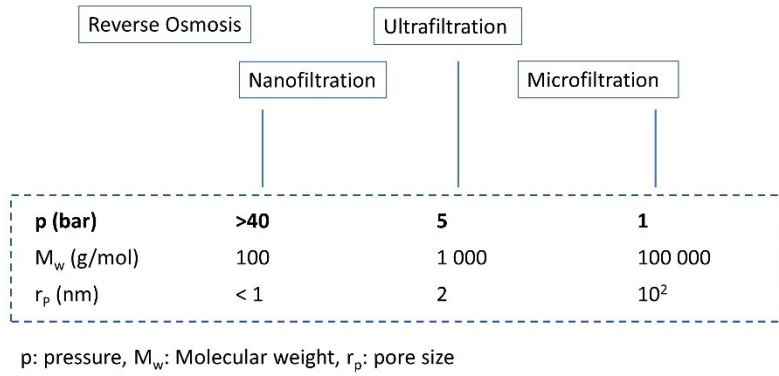


**Figure 17. Overview of membrane filtration processes and their classification by operating pressure P, the molecular weight of the rejected species and the pore size  $r_p$ .** Adapted from [131].

typically smaller than 2 nm, and is suitable for the separation of small molecules and ions. Reverse osmosis (RO) is well-known for desalination processes, as it can reject solutes on the smallest molecular level, with MWCO values below 1 nm. This requires operating pressures above 40 bar.

There are different parameters that can be used to characterise the separation performance of membranes with the most widely reported ones being the flux (or permeance) and rejection (or selectivity).<sup>131, 135</sup>

The flux is defined through the solvent volume that is passing a certain membrane area in a certain time and the permeance is the flux normalised to the applied trans-membrane pressure (**Equation (8)**).<sup>131, 135</sup>

$$B = \frac{V}{A \cdot t \cdot \Delta p} \quad (8)$$

with V = collected volume (L), A = membrane area (m<sup>2</sup>), t = sample collection time (h), p = trans-membrane pressure (bar).

In general, high permeances are desirable in membrane-based separated processes, to achieve a fast and economically viable diafiltration.

The rejection of a solute is a measure for the ratio of the concentration of this solute in the permeate and the retentate.

The rejection of a solute  $i$  can be derived from the following equation:

$$R_i = \left(1 - \frac{c_{P,i}}{c_{R,i}}\right) \quad (9)$$

with  $c_{P,i}$  being the concentration of the solute  $i$  in the permeate and  $c_{R,i}$  the concentration of  $i$  in the retentate.

The selectivity of a membrane for the separation of two components  $i$  and  $j$  can be described by the separation factor  $\beta_{i/j}$  with:

$$\beta_{i/j} = \left(1 - \frac{R_j}{R_i}\right) \quad (10)$$

with  $R_j$  being the rejection of component  $j$  and  $R_i$  being the rejection of component  $i$ .<sup>136</sup>

The MWCO of a membrane is a critical parameter for defining its selectivity for solutes during OSN.<sup>131,137,138</sup> It is defined as the  $M_w$  of a reference solute that is retained with 90% rejection.<sup>137</sup>

For an ideal membrane, the rejection curve (the rejection of a certain species plotted against its radius or molecular weight) would show one sharp, step-like increase at one specific value (blue, solid line in **Figure 18**). The species with a  $M_w$  below the MWCO would show 0 % rejection, whereas the species with a  $M_w$  higher than the MWCO would be 100% rejected. This type of membrane does not exist, and the rejection curves always increase over a range of solute  $M_w$ /radii, depending on the pore size distribution of the membrane.

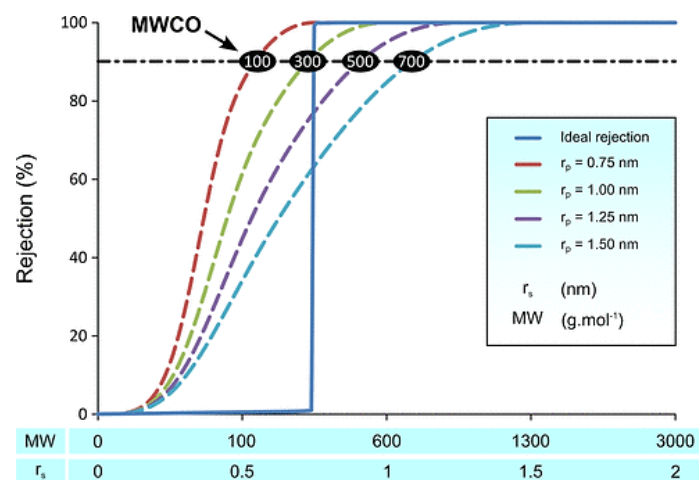


**Figure 18. Simulated rejections of an uncharged solute by NF membranes with different pore sizes, dependent on the  $M_w$  or radius of the solute.** The ideal rejection curve is represented by the blue line. Copied with permission from [136].

It is important to note that the MWCO can be a complex and sometimes ambiguous parameter due to its dependence on various factors. Different molecular species with the same molecular weight may exhibit significantly different hydrodynamic radii, which affects their behaviour in OSN. For example, proteins and polyethylene glycol (PEG) can have similar molecular weights but different sizes and interactions with membranes. The choice of solvent also plays a role in determining the molecular radii and membrane behaviour.

In general, filtration processes can be run in different set-ups and separation modes, which are concentration, solvent exchange, or purification. In the process of concentration, at least one solute is rejected by the membrane, whilst only the solvent is permeating through. In a solvent-exchange set-up, one or more solutes are retained by the membrane, whilst the initial solvent is permeating and the retentate is constantly fed with a second solvent, which results in the replacement of the first solvent with the second. In the last process mode, the purification, one or more solutes are rejected by the membrane, whilst other solutes are permeating through together with the solvent. The retentate is fed with fresh solvent and this results in the separation of the different solute species.

OSN is a relatively new field in the area of membrane separation processes, but a powerful tool for the separation of solutes and solvents on the nanometre scale. OSN-based processes might be capable of replacing highly energy-intensive thermal techniques such as distillation, or only limited scalable and highly solvent-consumptive processes like chromatography.

In order to develop highly selective and efficient OSN processes, intense research is conducted in producing membranes with a high stability and a well-defined MWCO. The membranes currently available for OSN are polymeric membranes/ mixed matrix membranes and ceramic membranes.<sup>131</sup> Polymeric or mixed membranes can be distinguished into integrally skinned asymmetric (ISA) and thin film composite (TFC) membranes.<sup>131</sup>

Ceramic membranes are made from an inorganic material, like oxides of aluminium, zirconium or titanium.<sup>139</sup> They usually show high thermal, chemical and mechanical stability and are less prone to fouling compared to polymeric membranes.<sup>131</sup> Disadvantages of ceramic membranes include their high hydrophilicity (problematic for the usage of non-polar solvents), as well as their high weight and production costs.<sup>131,140</sup> Furthermore they are more brittle and less easy to scale-up than polymeric membranes.<sup>106</sup>

Polymeric membranes are generally more cost-effective to manufacture than ceramic membranes and can be engineered with a wide range of pore sizes/ MWCOs and surface chemistries. This versatility allows them to be tailored to specific separation requirements, making them suitable for a broad spectrum of applications. Polymeric membranes are typically easier to scale up for large-scale processes due to their flexibility and the ability to produce membranes in various module configurations. This is important in industrial applications where a high throughput is necessary. They also bear the advantage of a low weight, which can simplify handling and installation.

The first type of polymeric membranes are thin film composite (TFC) membranes, in which an ultrathin film of a polymer is carried by a porous support of a different chemical composition.<sup>136</sup> These types of membranes are prepared by interfacial polymerisation, where the reactive

monomers are dissolved in two immiscible phases and polymerise at their interface. TFC membranes offer a variety of different polymers that can be used and techniques to fine-tune the pore size.<sup>106</sup> For example JIANG *et al.* implemented functionalised macrocycles into a TFC membrane which resulted in the alignment of the macrocycles to form defined pores of < 1 nm size.<sup>141</sup> These membranes exhibited a significantly higher methanol permeance and selectivity compared to the thicker, unorganised structures.

Another type of polymeric membranes are integrally skinned asymmetric (ISA) membranes, which consist of only one material, with different physical structures across the layers.<sup>136</sup> The top skin layer is tighter than the more porous sublayer and is forming the selective part of the membrane. ISA membranes are formed by phase-inversion immersion-precipitation processes, in which a solution is applied to a nonwoven backing material and then immersed in a bath of a solvent the polymer is not soluble in. That way the polymer layer precipitates from top to bottom to form the membrane. Similarly to TFC membranes, ISA membranes offer a wide selection of polymers to choose from, for example polyacrylonitrile (PAN), poly(ether ether ketone) (PEEK) or polybenzimidazole (PBI).<sup>106</sup> As PBI membranes are the type of membrane used in this work, its properties are going to be discussed in more detail in the following.

PBI membranes consist of a polymer with a benzimidazole repeating unit (**Figure 19**) and they recently gained interest due to their high thermal and mechanical stability.<sup>142</sup> Additional crosslinking with agents like  $\alpha,\alpha'$ -dibromo-*p*-xylene (DBX) makes this type of membrane highly stable towards a wide range of solvents.<sup>143</sup> The selectivity and permeability of PBI (as other types of ISA membranes) can be tuned by altering the polymer concentration in the casting solution.<sup>144,145</sup> In general, the higher the polymer concentration is, the more dense the skin layer will be, which results in a higher selectivity and in turn lower permeance.<sup>146</sup>

Another way to control the selectivity and performance of an ISA membrane is post-crosslinking modification with grafting agents. OXLEY *et al.* investigated the effect of post-crosslinking modification of DBX-crosslinked PBI membranes with a 1000 g mol<sup>-1</sup> copolymer based on PEG

and poly(propylene glycol) (PPG) and a small molecule modifier (2-methoxyethylamine, 75 g mol<sup>-1</sup>).<sup>135</sup> They managed to alter the membranes MWCO between 2000 g mol<sup>-1</sup> and 20 000 g mol<sup>-1</sup> by altering the ratio of the two grafting agents.

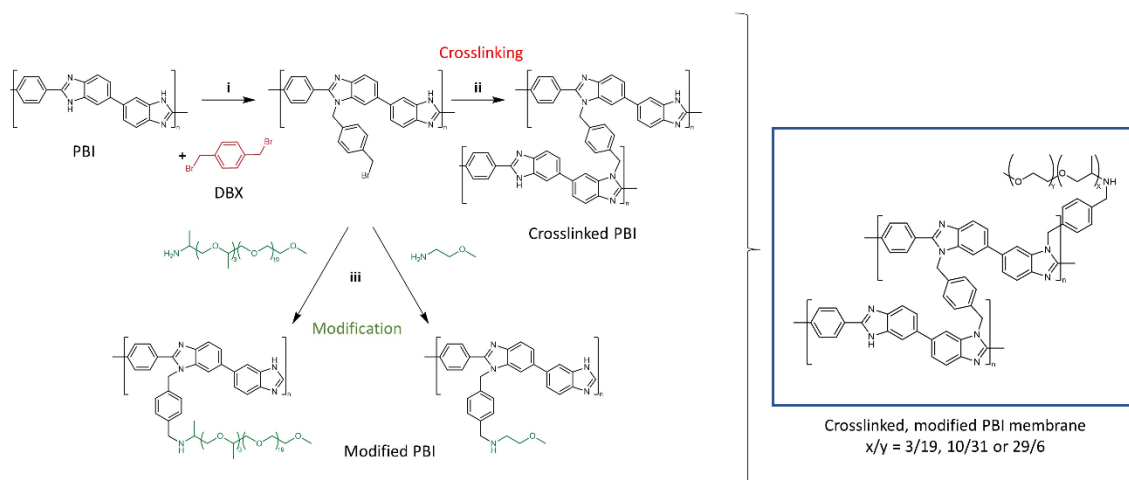


**Figure 19.** The reaction steps involved towards the fabrication of DBX-crosslinked PBI, modified with PPG/PEG polymers for increased anti-fouling behaviour. Adapted with permission from [135].

In a later study OXLEY *et al.* investigated the PEG/PPG modifications effect on the fouling behaviour of the membranes.<sup>145</sup> They found that depending on the chosen polymer (hydrophilic or hydrophobic), the grafting modification significantly improves the anti-fouling behaviour in different solvents, depending on the interactions of the polymer with the solvent. In water, PEG-based modifiers were found to show the best anti-fouling properties, whereas in polar organic solvents the PPG-based polymers delivered better results.

Other than tuning membrane properties in order to increase the efficiency of a membrane-based separation process, the set-up of the process itself can also be adapted to increase yields and the overall efficiency of the separation. For example, KIM *et al.* managed to significantly increase the yield of a membrane-based separation, by extending the conventional single-stage diafiltration to a cascade-like two-stage process.<sup>147</sup> In a model system the researchers purified an active pharmaceutical ingredient from two impurities using a crosslinked PBI membrane. The yield of

the separation process increased from 58 % in the single-stage system to 95 % in the two-stage system.

In summary, OSN is a valuable technique for the separation of solutes and solvents on the nano

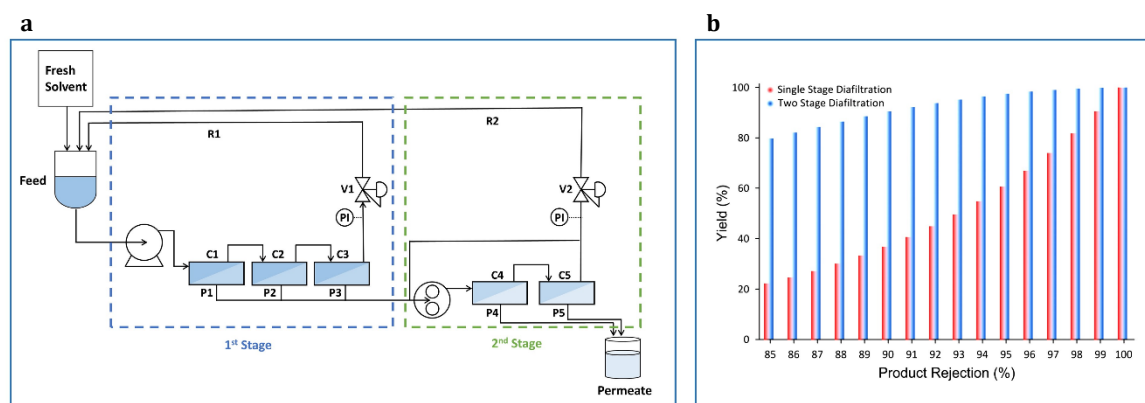


**Figure 20. a**, overview of a 2-stage diafiltration system (R1: retentate first stage, R2: retentate second stage, C1-C5: membrane filtration cells 1-5, P1-P5: permeate lines from cells 1-5, V1+2: pressure valves); **b**, modelled yields dependent on the rejections for a single stage and two stage DF system. Copied with permission from [147].

meter scale, particularly in industrial processes where high selectivity and efficiency are required.

## 2.2.7 Nanostar Sieving

As previously discussed, the challenges for synthesising DSPs/ DMPs involve separating the growing polymer from remaining reaction debris, such as excess building block (BB) or other byproducts formed during chain extension or deprotection. This purification step is crucial in an iterative synthesis, to avoid implementation of leftover BB in following reaction steps, or cross-reactions with other impurities. Most conventionally used separation techniques are inefficient, not easily scalable or involve reactions in the solid phase, which creates new problems related to impeded reaction kinetics.

To use OSN as an efficient separation technique, the size difference between the two species to be separated has to be sufficiently high. With the introduction of the 3-armed nanostar molecule

(**Figure 21**) , from which the polymer is grown successively, SZEKELY *et al.* have offered a solution to this problem and opened the door to membrane-based DSP/ DMP manufacturing. <sup>15,16</sup>

Dong *et al.* were the first ones to use the nanostar technology in combination with OSN to synthesise low molecular weight, multi-functional, sequence-defined polymers.<sup>19</sup> The polymers were based on a PEG backbone and were extended using a simple Williamson etherification reaction in the liquid phase. The growing chain was separated from excess building block and other reaction debris using selective poly(benzimidazole) (PBI) membranes.<sup>19</sup>

YEO *et al.* further used a similar approach to develop a one-pot, liquid-phase synthesis of highly pure peptides and thereby overcoming the limitation of the conventional solid-phase processes.<sup>148</sup>

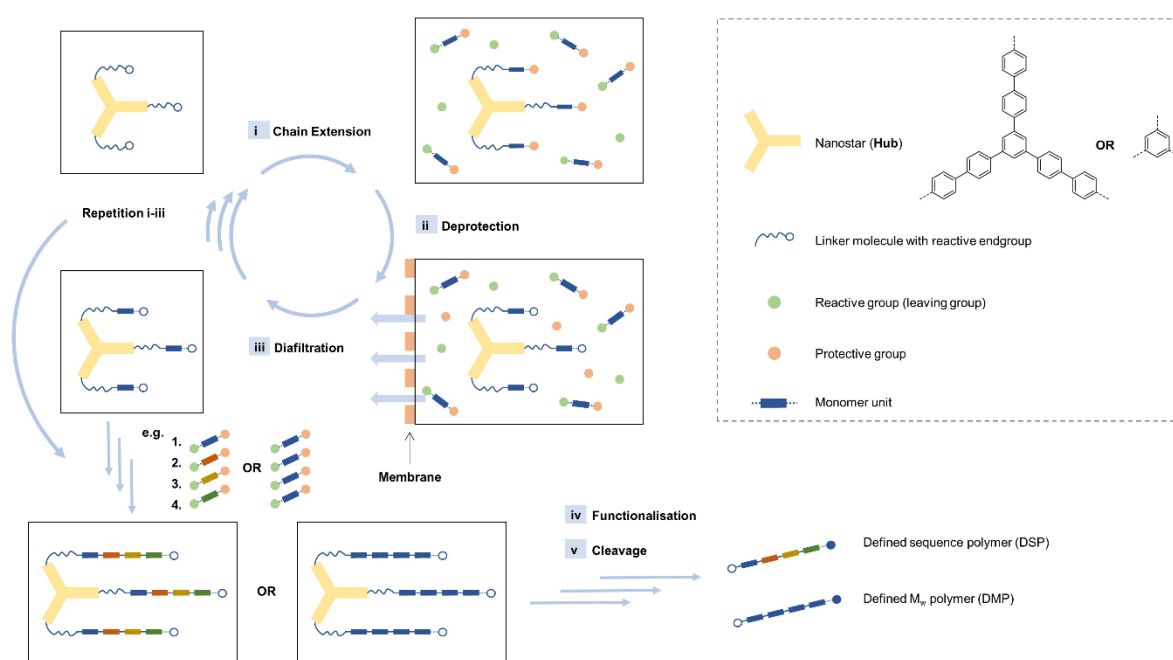


**Figure 21. General overview of Nanostar Sieving process to synthesise DSPs and DMPs.** Each cycle consists of chain extension (i), deprotection (ii) and diafiltration (iii). The cycle is repeated to achieve DSPs or DMPs with the desired sequence/ chain length.

The general process of a nanostar sieving based synthesis of a DSP or DMP is shown in **Figure 21**. The benzylic nanostar molecule is pre-loaded with a linker molecule and a reactive end group. In a first step the chains are extended by reaction with a BB that is equipped with a reactive group

on one side, and a protecting group on the other (**Figure 21, i**). Subsequently the protecting groups are removed (**Figure 21, ii**) and the resulting mixture is then subjected to diafiltration using a selective membrane, that rejects the growing nanostar-polymer species, whilst other reaction components can permeate through (**Figure 21, iii**). This cycle of chain extension, deprotection and diafiltration can be repeated until the desired sequence or chain length is achieved. To achieve a DMP with a homogenous composition, the same type of BB is used in every chain extension. In case of DSPs different BBs are used in every chain extension. For example, in the work of DONG *et al.* the BBs were carrying different side-groups, like benzyl or azide.

Besides allowing for liquid-phase synthesis and fast purification by OSN, the benzylic hub molecule shows a distinctive UV absorption maximum at 310 nm. This facilitates easy reaction monitoring by UPLC coupled with a UV detector.

Using the nanostar sieving approach for synthesising defined  $M_w$  PEGylation agents is the subject of the work in hand. A detailed overview of the research objectives is given in the following chapter (**Chapter 3**).

## Chapter 3 Research Objectives

From the literature review it becomes apparent that using defined  $M_w$  PEGylation agents could offer significant advantages over the currently used disperse agent. So far, there has been no process reported for the synthesis of defined  $M_w$  PEG beyond  $\sim 3$  kDa and there are no studies on the potential advantages of defined  $M_w$  PEGylation agents. Nanostar Sieving offers a promising approach for producing defined  $M_w$  PEG in an efficient and scalable manner. The work in hand deals with the synthesis of defined  $M_w$  PEGylation agents using Nanostar Sieving, as well as their application and potential benefits.

### **Objective 1 Develop and optimise the Nanostar Sieving-based synthesis of defined $M_w$ PEG**

In order to investigate if the Nanostar Sieving approach could be a competitive method for synthesising defined  $M_w$  PEG, the first objective within this work was to investigate and optimise the process in terms of reaction kinetics, step-reduction and final purity of the product. As the lowest weight PEGylation agent which is used in multiple FDA-approved PEGylated therapeutics is 5 kDa (5 out of 40, **Table 1**), the first target was to develop and optimise the multi-gram synthesis of defined  $M_w$  PEG with 112 Eg repeating units ( $Eg_{112}$ ) which is equivalent to a  $M_w$  of 5 kDa. Analytical monitoring was optimised and the chain extension and deprotection reactions were investigated and optimised regarding their kinetics. The usage of a longer building block ( $Eg_{28}$  instead of  $Eg_{12}$ ) was investigated and the diafiltration process was adapted accordingly. First investigations for the synthesis of higher  $M_w$  defined PEG were pursued.

The different research steps to achieve Objective 1 are the following and the results are discussed in **Chapter 4**:

- Investigate and optimise the Nanostar Sieving-based  $Eg_{112}$  synthesis in terms of choice of reactor material, solvent and reaction concentration
- Investigate and optimise the Nanostar Sieving-based  $Eg_{112}$  synthesis in terms of the formation of impurities and the final purity of the product

- Optimise the reaction monitoring and end-product characterisation to assure the highest standard for following experiments
- Increase the process efficiency by using a longer building block and in this way reduce the number of synthesis steps needed
- Validate the final results through comparison with a conventional chromatography-based synthesis approach
- Investigate the potential production higher  $M_w$  defined PEG

### **Objective 2 Synthesise active defined $M_w$ PEGylation agents**

To use PEG as an active PEGylation agent it must be capped with a chemically inert group at one end and activated at the other end by introducing a reactive functional group. As this work focuses on the PEGylation of protein and peptide therapeutics, the second objective was to synthesise defined  $M_w$  PEGylation agents reactive towards functional groups commonly found in proteins and peptides.

To achieve Objective 2 the following research strategy was followed, and the results are presented in **Chapter 5**.

- Synthesise defined  $M_w$  PEGylation agents with a  $M_w$  of 5 kDa for random PEGylation of proteins/peptides
- Synthesise defined  $M_w$  PEGylation agents with a  $M_w$  of 5 kDa for the site-selective PEGylation of proteins/peptides
- Perform first comparative studies on defined  $M_w$  and disperse PEGylation reagents regarding synthesis and purity

### **Objective 3 Investigate the application of defined $M_w$ PEGs in the PEGylation of proteins**

The final aim of this work was to investigate the application of defined  $M_w$  PEG as PEGylation agents in the PEGylation of protein and peptide therapeutics. The literature review revealed the problems related to the dispersity of currently used PEGylation agents regarding the analysis and characterisation of PEGylated therapeutics. A potential application of defined  $M_w$  PEGylation for

the facile identification of PEGylation sites was identified and the results are presented in **Chapter 5**. The following research strategy was followed:

- Random PEGylation of bovine serum albumin as an inexpensive model, using the previously synthesised defined  $M_w$  5 kDa PEGylation agents
- Purification and characterisation of the PEGylated protein
- Perform a comparative study on a peptide mapping procedure on BSA PEGylated with defined  $M_w$  PEG and with a commercially available disperse derivative

## Chapter 4 Synthesis of defined $M_w$ PEG using Nanostar Sieving

### Abstract

Synthesising defined  $M_w$  PEG of higher molecular weight in an efficient and scalable manner is of high interest, especially due to the potential application as defined  $M_w$  PEGylation agents. Current approaches do not exceed a  $M_w$  of 2.8 kDa (Eg<sub>64</sub>) and the majority rely on extensive chromatographic purification or solid-phase reactions. This chapter deals with the Nanostar Sieving based synthesis of defined  $M_w$  PEG up to a molecular weight of 5 kDa (Eg<sub>112</sub>), using DBX-crosslinked PBI membranes for purification. A conventional liquid-phase approach based on the Williamson etherification was used for iteratively growing the PEG, combined with an acid-labile protecting group to allow for the synthesis of defined chain length. Aiming for the introduction of longer building blocks, the process was optimised regarding the monitoring of reactions and reaction kinetics, as well as the formation and removal of impurities. Within these studies, the validity of methods for analysing the final chain length purity of the product were investigated. Finally, an Eg<sub>28</sub> building block was introduced to cut the numbers of steps involved and increase the overall process efficiency. Defined  $M_w$  PEG 5 kDa (Eg<sub>112</sub>) was prepared on a multigram scale, and the process efficiency and final product purity was evaluated in a comparative study including an equivalent synthesis based on chromatographic purification.

### Keywords

Defined  $M_w$  poly(ethylene glycol) (PEG), monodisperse, Nanostar Sieving, Organic solvent nanofiltration,

### Highlights

- Defined PEG 5 kDa (Eg<sub>112</sub>) can be synthesised efficiently on a multi-gram scale using Nano Star sieving and an Eg<sub>28</sub> building block, which reduces the number of cycles required from seven to three compared to a synthesis based on an Eg<sub>12</sub> building block

- MALDI-TOF is not a suitable technique to quantitatively determine the chain length purity of PEG
- An approach involving the functionalisation with a UV-active group and analysis through UPLC-MS can be used to determine the chain length purity based on growth defects occurring during the iterative synthesis of defined  $M_w$  PEG up to 5 kDa
- The synthesis of defined  $M_w$  PEG 5 kDa based on the Nanostar Sieving approach results in double the yield and similar product purity, compared to an equivalent process using chromatography for purification

## 4.1 Introduction

To establish the Nanostar Sieving based synthesis of defined  $M_w$  PEG as a competitive and economically feasible technique, different parameters must be investigated and adapted. Each cycle of the process involves multiple chemical reactions (hub loading, chain extension and deprotection), alongside the critical membrane filtration (diafiltration)(**Figure 29**). To obtain a product of maximum purity (chemical and chain length purity), the chemical reaction steps must be optimised to achieve 100 % conversion, whilst avoiding side-products. Another way is to optimise the purification, so that it removes 100 % of impurities that are present in the reaction mixture at the time of purification.

Optimisation and reliable analytical techniques are the basis of a reliable process in both cases. Additionally, economic viability is based on optimum speed, including reaction kinetics, practical handling, and separation time. Furthermore, the process must be optimised in terms of material costs, so an optimum atom-efficiency is desired.

The foundation of the Nanostar Sieving based synthesis of defined MW PEG is the base-induced Williamson etherification, in which the growing PEG chain is coupled with a reactive, PEG-based building block. This  $S_N2$ -type reaction is a well-established method that has first been used for the synthesis of short, defined  $M_w$  PEG by PERRY and HIBBERT in 1936.<sup>122</sup> The early approaches involved reacting PEG alkoxides with a PEG chloride and relied on harsh reaction conditions with

temperatures of up to 175 °C.<sup>122,123</sup> These reactions conditions were found to result in side reactions, including the base-induced depolymerisation of the PEG chain or elimination reactions at the chain end, resulting in  $\text{Eg}_{n-x}$  impurities.<sup>149</sup> Later, species bearing a tosylate group as a more reactive group were introduced to allow for reactions at lower temperatures and with shorter reaction times.<sup>114,124</sup> It has to be noted, that the reported reactions involved relatively short BBs (up to  $\text{Eg}_8$ ) and reactions up to a length of  $\text{Eg}_{56}$  and no data on the kinetic behaviour of higher molecular weight chain extensions has been published.<sup>15</sup>

The protecting groups previously used with the Williamson-etherification based synthesis of defined  $M_w$  PEG were either benzyl- or trityl-species that can be cleaved off using either reduction with  $\text{H}_2$  an Pd/C or with acids like dichloroacetic acid (DCA) or trifluoroacetic acid (TFA).

12,13,15,16,114,115,124,150

To produce the PEG alkoxide almost solely sodium hydride has been used as the base, except a few examples in which potassium tert-butoxide (KOTBu) or sodium metal were used.

12,13,15,16,114,115,124,150

The solvent for the Williamson etherification must be aprotic and capable of dissolving the reactants. In almost all cases tetrahydrofuran (THF) was chosen, besides some examples using *N,N'*-Dimethylformamide (DMF). It has to be noted that DMF (as well as acetonitrile) have been shown to result in side products in sodium hydride-mediated Williamson etherification reactions, as sodium hydride can react as both, a base and a reducing agent.<sup>151</sup>

The foundation for a membrane-based process to synthesise defined  $M_w$  PEG is to find a membrane with superior stability and performance. It has previously been shown that DBX-crosslinked PBI membranes exhibit a high thermal, mechanical, and chemical stability and their selectivity can be adapted by using different concentrations of PBI in the initial polymer solution or by post-modification with a linear copolymer consisting of PEG and PPG repeating units.<sup>135,142-</sup>

<sup>145</sup> Modification with a hydrophobic copolymer with a PPG/PEG ratio of 29/5 results in superior anti-fouling properties in MeOH, which is also an ideal solvent for PEG.

Within this chapter, defined  $M_w$  PEG 5 kDa was first synthesised on a multi-gram scale using a Nanostar Sieving based approach and an  $Eg_{12}$  building block. To reach the 5 kDa species ( $Eg_{112}$ ), seven cycles of chain extension, deprotection and purification were needed. The process was investigated and optimised, firstly by investigating the reaction kinetics of the chain extensions regarding reactor material and reaction concentration. Furthermore, the quality and reliability of the analytical monitoring and characterisation procedures were improved.

The formation and separation of impurities was studied and the reaction conditions were adapted to increase the final purity of the product. Previously used techniques to determine chain length purities of defined  $M_w$  PEG were critically reviewed and a solution to quantitatively determine the growth-related chain length impurities was developed. Finally, a synthesis based on using an  $Eg_{28}$  BB was developed and the Nanostar Sieving based synthesis of defined  $M_w$  PEG 5 kDa ( $Eg_{112}$ ) was evaluated in a comparative study, including an equivalent synthesis using chromatography as means of purification.

## 4.2 Materials and Methods

### 4.2.1 Materials

#### Chemicals

Dichloroacetic Acid (>98%) (TCI), sinapinic acid (matrix substance for MALDI, ≥99 %, TCI), pyrrole (>99.0%, TCI), sodium hydride (60 % in mineral oil, Sigma-Aldrich), *trans*-2-[3-(4-*tert*-butylphenyl)-2-methyl-2-propenylidene]malononitrile (DCTB, matrix substance for MALDI, ≥99 %, TCI), triethylamine (for synthesis, Sigma-Aldrich), 4,4'-dimethoxytriphenylmethyl chloride (>97.0 %, TCI),  $NaHCO_3$  (ACS reagent, > 99.7 %, Sigma-Aldrich), *N*-methylimidazole (for synthesis, Sigma-Aldrich), *p*-toluenesulfonyl chloride (TsCl) (ReagentPlus®, ≥99%, Sigma-Aldrich),  $Eg_{12}$  Diol (Dodecathylene glycol, Polypure AS) and  $Eg_{28}$  Diol (Octacosathylene glycol, Polypure AS), 5''-(4'-formyl-[1,1'-biphenyl]-4-yl)-[1,1':4',1'':3'',1''':4''',1''''-quinquephenyl]-4,4''''-dicarbaldehyde (CAS: 805246-78-0, Chemextension ET Co. Ltd.) All solvents were purchased from VWR.

All PBI membranes used within this work were prepared by Adam Oxley, as described in [135] and [145].

The Hub-Br<sub>3</sub> (**1**) and the building blocks (**2**) and (**3**) were prepared with help from Naveed Zaidi.

#### 4.2.2 Methods

All reactions were performed under argon, unless otherwise stated. Solvents labelled as “dry” were dried over activated 3Å molecular sieve for 24 hours and tested for their water content using a Karl-Fischer coulometric titrator (<50 ppm).

The Nano Star species is referred to as “Hub” in chemical formule.

##### *Synthesis of Hub-Tribromide (Hub-Br<sub>3</sub>, (**1**))*

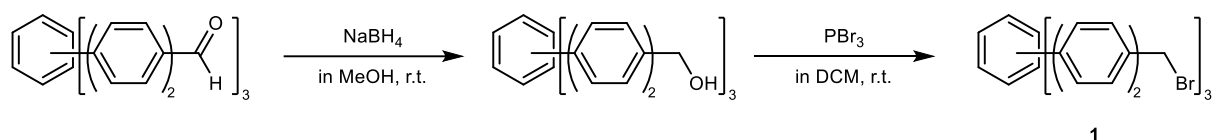


**Figure 22. Reaction scheme of the reaction of the nanostar-aldehyde (Hub(CHO)<sub>3</sub> to the nanostar-alcohol (Hub(C-OH)<sub>3</sub> and the nanostar-tribromide HubBr<sub>3</sub> (**1**).**

To a suspension of Hub(CHO)<sub>3</sub> (618.7 g mol<sup>-1</sup>, 1.0 g, 1.6 mmol) in methanol (50 mL) was added NaBH<sub>4</sub> (0.2 g, 4.7 mmol, 2.9 eq.) stepwise and the mixture was left to stir at room temperature for 30 minutes. Subsequently the reaction was diluted with water to destroy excess NaBH<sub>4</sub> and the mixture was filtered through a sinter funnel and the solid was washed with water (2 x 20 mL) and once with methanol, dried in vacuo. The product (Hub(OH)<sub>3</sub>) was obtained as a white solid (624.8 g mol<sup>-1</sup>, 1.00 g, 1.6 mmol, y = 99 %).

To a solution of Hub(OH)<sub>3</sub> (624.8 g mol<sup>-1</sup>, 0.70 g, 1.1 mmol) in dichloromethane was added excess of PBr<sub>3</sub> (3.0 g, 11.0 mmol) at room temperature and the mixture was left to stir overnight. The next day the brown solution was poured onto ice in a beaker, and it was left to stir for 2 h until the formation of HCl gas ceased. The solid was filtered through a sinter funnel and washed with

cold water and recrystallised using ethyl acetate/ pentane to obtain the product (**1**) as a colourless solid (813.5 g mol<sup>-1</sup>, 0.9 g, 1.1 mmol, y = 95 %).

**<sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>)** δ = 8.01 (m, 9H, Hub **CH**), 7.84 (d, *J* = 8.2 Hz, 6H, Hub **CH**), 7.76 (d, *J* = 7.9 Hz, 6H, Hub **CH**), 7.58 (d, *J* = 7.9 Hz, 6H, Hub **CH**), 4.79 (s, 6H, Ar**CH**<sub>2</sub>Br).

**<sup>13</sup>C NMR (101 MHz, DMSO-d<sub>6</sub>)** δ = 141.13 (3C, 3 x Hub **C**), 139.55 (3C, 3 x Hub **C**), 139.31 (3C, 3 x Hub **C**), 138.80 (3C, 3 x Hub **C**), 137.35 (3C, 3 x Hub **C**), 129.98 (3C, 3 x Hub **CH**), 127.78 (6C, 6 x Hub **CH**), 127.56 (6C, 6 x Hub **CH**), 127.21 (6C, 6 x Hub **CH**), 126.91 (6C, 6 x Hub **CH**), 34.34 (3 C, 3 x **CH**<sub>2</sub>Br).

### Synthesis of Building Blocks (DmtrO-Eg<sub>12</sub>-OTs (**2**), DmtrO-Eg<sub>28</sub>-OTs (**3**))

#### DmtrO-Eg<sub>12</sub>-OTs (**2**)

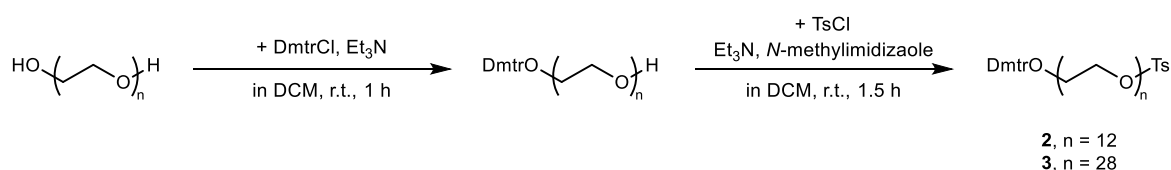


**Figure 23. Reaction scheme of the synthesis of the building blocks DmtrO-Eg<sub>12</sub>-OTs (**2**) and DmtrO-Eg<sub>28</sub>-OTs (**3**).**

Eg<sub>12</sub> diol (546.7 g mol<sup>-1</sup>, 10.0 g, 18.3 mmol) was evaporated from anhydrous acetonitrile twice. The residue was dissolved in anhydrous DCM (70 mL) and 4,4'-dimethoxytriphenylmethyl chloride (Dmtr-Cl, 338.8 g mol<sup>-1</sup>, 6.2 g, 18.3 mmol, 1.0 eq.) was added. The solution was cooled in an ice-bath and triethylamine (101.2 g mol<sup>-1</sup>, 5.1 ml, 36.6 mmol, 2.0 eq.) was added. The cooling bath was removed after 5 minutes, and the reaction stirred for 1 hour.

Upon completion, the reaction was diluted with more DCM and more triethylamine was added to avoid detritylation. The crude product was loaded onto silanised silica and purified using flash chromatography before further reaction (849.0 g mol<sup>-1</sup>, 7.0 g, 8.2 mmol, y = 45 %).

In the next step DmtrO-Eg<sub>12</sub>-OH (849.0 g mol<sup>-1</sup>, 10.0 g, 11.8 mmol) was evaporated from anhydrous acetonitrile twice. The residue was dissolved in anhydrous DCM to which was added

triethylamine (4.9 ml, 35.4 mmol, 3.0 eq.) and *N*-methylimidazole (0.9 ml, 11.8 mmol, 1.0 eq.). The reaction was cooled in a water bath and *p*-toluenesulfonyl chloride (TsCl, 3.4 g, 17.7 mmol, 1.5 eq.) was added. After 5 minutes the water bath was removed, and the reaction was stirred at room temperature for 1.5 h. THF (50ml) was added to the crude mixture and the DCM was evaporated off (at room temperature) and to the residue was added more THF (3x40ml). The salt (Et<sub>3</sub>N·HCl) was removed by filtration, washing the solids with further THF. NaHCO<sub>3</sub> was added (9.9 g, 118 mmol, 10 eq.) was added to the crude mixture, then water (100 ml) and the solution was stirred for 3 hours to hydrolyse the remaining TsCl.

The crude product was loaded onto silanised silica and purified using flash chromatography. The product **2** was obtained as a yellow, viscous liquid (1003.2 g mol<sup>-1</sup>, 9.2 g, 9.2 mmol, *y* = 78 %).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)**  $\delta$  = 7.79 (d, *J* = 8.3 Hz, 2H, CH<sub>2</sub> (Ar, Ts)), 7.45 (d, *J* = 7.1 Hz, 2H, CH<sub>2</sub> (Ar, Ts)), 7.33 (d, *J* = 8.9 Hz, 6H, CH<sub>2</sub> (Ar, Dmtr)), 6.80 (d, *J* = 8.9 Hz, 4H, CH<sub>2</sub> (Ar, Dmtr)), 4.18 – 4.11 (m, 2H, -CH<sub>2</sub>-), 3.77 (s, 6H, Dmtr-O-CH<sub>3</sub>), 3.71 – 3.59 (m, 40H, -(CH<sub>2</sub>-CH<sub>2</sub>-O)-), 3.57 (s, 4H, -CH<sub>2</sub>-), 3.21 (dd, *J* = 5.6, 4.9 Hz, 2H, -CH<sub>2</sub>-), 2.43 (s, 3H, Ts-O-CH<sub>3</sub>).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)**  $\delta$  = 158.50 (2 x C (Ar, Dmtr)), 145.21 (1 x C (Ar, Ts)), 144.88 (1 x C (Ar, Ts)), 136.47 (2 x CH<sub>2</sub> (Ar, Ts)), 130.18 (4 x CH<sub>2</sub> (Ar, Dmtr)), 129.94 (2 x C (Ar, Dmtr)), 128.33 (2 x CH<sub>2</sub> (Ar, Dmtr)), 128.10 (2 x CH<sub>2</sub> (Ar, Dmtr)), 127.84 (2f x C (Ar, Dmtr)), 126.75 (2 x CH<sub>2</sub> (Ar, Ts)), 113.15 (4 x CH<sub>2</sub> (Ar, Dmtr)), 86.04 (1 x C (Dmtr)), 70.69 (21 x CH<sub>2</sub> (PEG)), 69.35 (1 x CH<sub>2</sub> (PEG)), 68.80 (1 x CH<sub>2</sub> (PEG)), 63.27 (1 x CH<sub>2</sub> (PEG)), 55.32 (2 x CH<sub>3</sub> (Dmtr)), 21.76 (1 x CH<sub>3</sub> (Ts)).

### DmtrO-Eg<sub>28</sub>-OTs (**3**)

The Eg<sub>28</sub> BB was synthesised in the same way as the Eg<sub>12</sub> BB but using Eg<sub>28</sub> diol as the starting material and using automated flash chromatography for purification. The amounts of the reagents, as well as the yield can be found below:

### Tritylation

Eg<sub>28</sub> Diol (1251.5 g mol<sup>-1</sup>, 20.0 g, 16.0 mmol), DCM (70 mL), 4,4'-dimethoxytriphenylmethyl chloride (Dmtr-Cl, 338.8 g mol<sup>-1</sup>, 5.41 g, 16.0 mmol, 1.0 eq.), triethylamine (101.2 g mol<sup>-1</sup>, 4.50 mL, 32.3 mmol, 2.0 eq.) .

Upon completion, the reaction was diluted with more DCM and more triethylamine was added to avoid detritylation. The crude product was loaded onto silanised silica and purified using automated flash chromatography (Column: Biotage Sfaer C18 D, Duo 100 Å 30 µm, 400 g; flowrate 100 mL/min; gradient: 25-95 %B over 12 column volumes; A: H<sub>2</sub>O + 1 % Et<sub>3</sub>N, B: Acetone).

### Tosylation

DmtrO-Eg<sub>28</sub>-OH (1553.9 g mol<sup>-1</sup>, 24.0 g, 15.4 mmol), triethylamine (7.4 mL, 54.0 mmol, 3.0 eq.) and *N*-methylimidazole (1.5 mL, 18.0 mmol, 1.0 eq.), toluene sulfonyl chloride (Tos-Cl, 4.5 g, 23.6 mmol, 1.5 eq.), THF (50 mL), NaHCO<sub>3</sub> (15 g, 180 mmol, 10 eq.), water (100 mL)

The crude product was loaded onto silanised silica and purified using automated flash chromatography (Column: Biotage Sfaer C18 D, Duo 100 Å 30 µm, 400 g; flowrate 100 mL/min; gradient: 10-95 %B over 16 column volumes; A: H<sub>2</sub>O + 1 % Et<sub>3</sub>N, B: Acetone). The fractions containing the pure product (confirmed via UPLC-MS) were combined and evaporated off to yield substance **3** (1708.05 g mol<sup>-1</sup>, 22.2 g, 13.0 mmol, *y* = 84 %).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)** δ = 7.78 (d, *J* = 8.0 Hz, 2H, CH<sub>2</sub>(Ar, Ts)), 7.43 (d, *J* = 7.7 Hz, 2H, CH<sub>2</sub>(Ar, Ts)), 7.32 (d, *J* = 8.4 Hz, 6H, CH<sub>2</sub>(Ar, Dmtr)), 6.79 (d, *J* = 8.5 Hz, 4H, CH<sub>2</sub>(Ar, Dmtr)), 4.14 (m, 2H, -CH<sub>2</sub>-), 3.76 (s, 6H, Dmtr-O-CH<sub>3</sub>), 3.62 (m, 104H, (-CH<sub>2</sub>-CH<sub>2</sub>-O-)), 3.56 (s, 4H, -CH<sub>2</sub>-), 3.20 (t, *J* = 5.3 Hz, 2H, -CH<sub>2</sub>-), 2.42 (s, 3H, Ts-O-CH<sub>3</sub>).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)** δ = 158.38 (2 x C (Ar, Dmtr)), 145.09 (1 x C (Ar, Ts)), 144.76 (1 x C (Ar, Ts)), 136.37 (2 x CH<sub>2</sub> (Ar, Ts)), 130.05 (4 x CH<sub>2</sub> (Ar, Dmtr)), 129.81 (2 x C (Ar, Dmtr)), 128.21 (2 x CH<sub>2</sub> (Ar, Dmtr)), 127.98 (2 x CH<sub>2</sub> (Ar, Dmtr)), 127.72 (2 x C (Ar, Dmtr)), 126.62 (2 x CH<sub>2</sub> (Ar,

Ts)), 113.03(4 x CH<sub>2</sub> (Ar, Dmtr)), 85.92 (1 x C (Dmtr)), 70.59 (53 x CH<sub>2</sub> (PEG)), 69.23 (1 x CH<sub>2</sub> (PEG)), 68.68 (1 x CH<sub>2</sub> (PEG)), 63.14 (1 x CH<sub>2</sub> (PEG)), 55.19 (2 x CH<sub>3</sub> (Dmtr)), 21.63 (1 x CH<sub>3</sub> (Ts)).

Loading of nanostar with Eg<sub>28</sub> to reach Hub(Eg<sub>28</sub>(OH)<sub>3</sub>) (5a)

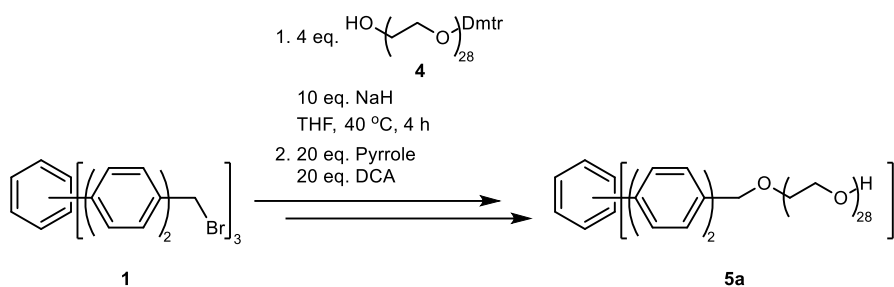


**Figure 24. Reaction scheme of the synthesis of Hub(Eg<sub>28</sub>-OH)<sub>3</sub> (5a).**

For the first coupling, DmtrO-Eg<sub>28</sub>-OH (**4**, 5.35 g, 3.44 mmol, 4 eq.) was azeotroped from dry acetonitrile and then dissolved in dry THF (20 ml) to which sodium hydride (60 % dispersion in mineral oil, 0.34 g, 8.60 mmol, 10 eq.) was added. The solution was stirred for 5 minutes and then the hub-tribromide (**1**, 0.70 g, 0.86 mmol) was added and the reaction was stirred at 40 °C for 4 hours. The reaction was allowed to cool down to room temperature and then quenched with satd. ammonium chloride solution. In the case of a batch based on chromatographic purification, the crude material was purified through chromatography once before and once after detritylation. Before loading onto silanised silica, the crude product was diluted with THF and filtered through a paper filter. For deprotection, the crude product was redissolved in DCM, and pyrrole (1.19 ml, 17.2 mmol, 20 eq.) and DCA (1.42 mL, 17.2 mmol, 20 eq.) were added. The solution was stirred until the reaction was completed (confirmed by UPLC-MS) and was then neutralised with triethylamine until pH ≈ 7. The DCM was removed under reduced pressure and the residue was dissolved in MeOH, any solids were by paper filtration and the clear solution was either transferred to the diafiltration rig for membrane purification, or the material was loaded onto silanised silica for chromatographic purification. After complete purification, the product (**5a**) was obtained as a yellow, waxy solid (membrane purification: 3.24 g, 87 %, chromatographic purification: 2.46 g, 86%/78%).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)**  $\delta$  = 7.86 (s, 3H, Hub **CH**), 7.78 (d, J = 8.2 Hz, 6H, Hub **CH**), 7.70 (d, J = 8.0 Hz, 6H, Hub **CH**), 7.62 (d, J = 8.0 Hz, 6H, Hub **CH**), 7.43 (d, J = 7.9 Hz, 6H, Hub **CH**), 4.60 (s, 6H, 3 x OCH<sub>2</sub>Ar), 3.72 – 3.53 (m, 336H, 168 x CH<sub>2</sub>O).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)**  $\delta$  = 142.01 (3C, 3 x Hub **C**), 140.23 (3C, 3 x Hub **C**), 140.02 (3C, 3 x Hub **C**), 139.97 (3C, 3 x Hub **C**), 137.61 (3C, 3 x Hub **C**), 131.06 (3C, 3 x Hub **CH**), 128.34 (6C, 6 x Hub **CH**), 127.77 (6C, 6 x Hub **CH**), 127.58 (6C, 6 x Hub **CH**), 127.09 (6C, 6 x Hub **CH**), 73.01+72.62 (3 x OCH<sub>2</sub>Ar), 70.60- 70.58 (m, 159 C, 159 x CH<sub>2</sub>O), 70.32 (3 C, 3 x CH<sub>2</sub>O), 69.59 (3 C, 3 x CH<sub>2</sub>O), 61.72 (3C, 3 x CH<sub>2</sub>OH).

The UPLC-MS data corresponding to structure **5a** can be found in the appendix (**Figure A 1**).

#### General protocol for chain extension and deprotection

##### Synthesis using Eg<sub>12</sub> based building block (**2**)

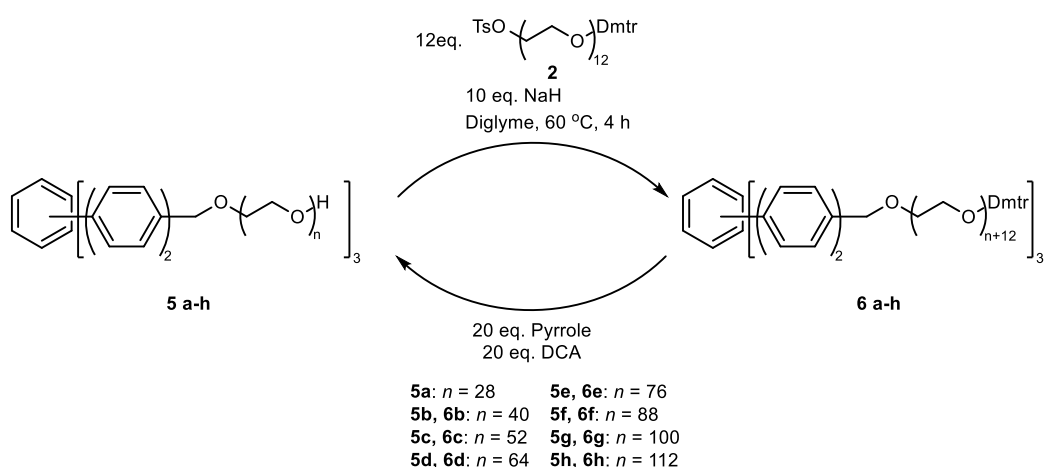


**Figure 25.** Reaction scheme of the chain extension and deprotection reactions using the Eg<sub>12</sub> building block (**2**).

For every chain extension, the combined nanostar-PEG-alcohol (**5a-h**, Hub(Eg<sub>n</sub>-OH)<sub>3</sub>) and the BB (**2**, DmtrO-Eg<sub>12</sub>-OTs) were azeotroped from dry acetonitrile, then dissolved in dry diglyme. Sodium hydride (60 wt% dispersion in mineral oil, 10 eq.) was added under argon and the

reaction was stirred in an oil bath set to 60 °C until the reaction was fully complete (confirmation through UPLC-MS). The reaction was quenched with sat. ammonium chloride and the solvent was removed under reduced pressure. For detritylation, the crude product was redissolved in DCM, and pyrrole (20 eq.) then DCA (20 eq.) were added. The solution was stirred until the reaction was complete (confirmation through UPLC-MS) and was then neutralised with triethylamine until pH  $\approx$  7. The DCM was removed under reduced pressure and the crude product was dissolved in MeOH, any solids were removed by paper filtration and the clear solution was either transferred to the diafiltration rig for membrane purification, or the material loaded onto silanised silica for chromatographic purification.

Syntheses using  $Eg_{28}$  based building block (3)

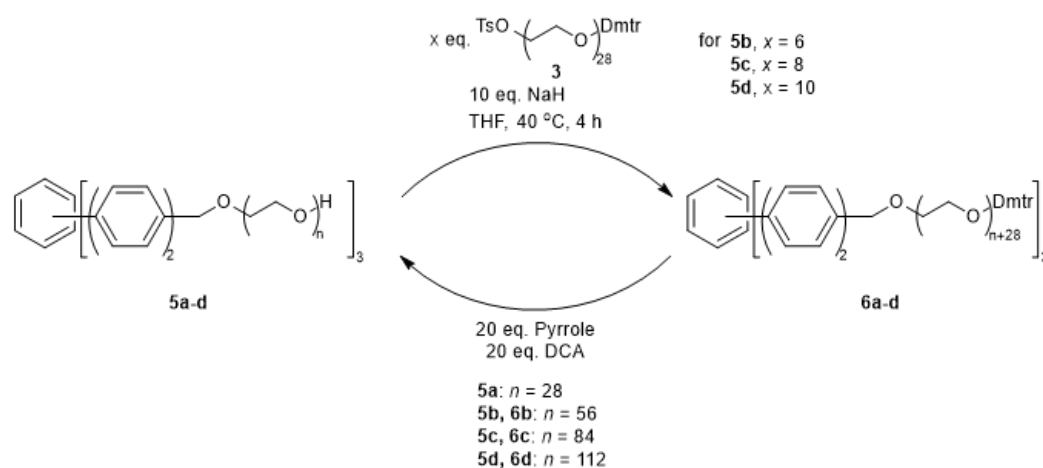


**Figure 26.** Reaction scheme of the chain extension and deprotection reactions using the  $Eg_{28}$  building block (3).

The chain extension using the  $Eg_{28}$  BB **3** were conducted in an equivalent manner to the chain extensions with  $Eg_{12}$  BB **2**, with the following alterations:

- Different equivalents of BB **3** were used for every step. The specific amounts are given in **Table 3** and **Table 4**
- THF was used as the solvent instead of diglyme and the reaction was performed at 40 °C

### Protocol for membrane purification

The diafiltration rig was a 2-stage, crossflow system (see supplementary information for scheme), the first stage consisting of three circular cells, each equipped with a flat sheet membrane of active area 51 cm<sup>2</sup>, and the second stage consisting of two cells with membranes of the same area. A diaphragm pump (Wanner, Hydra-Cell G20) both raised the pressure and circulation flow (1.76 ml/min) in the first stage, and a gear pump provided circulation flow in the second stage; rapid cross-flow over the membranes is required to minimise concentration polarisation. Depending on the rejection of the Hub-PEG-alcohol species, the rig was fitted with different types of membranes. The rig was configured in either the two-stage set-up to maximise selectivity at lower molecular weight nanostar, or the second stage was disconnected at high molecular weight nanostar to accelerate diafiltration. Each set of membranes was compacted in methanol at 10 bar for at least 8 hours before using it for purification.

After each chain extension/deprotection cycle, the deprotected Hub-PEG-alcohol species was transferred to the diafiltration rig dissolved in MeOH and the sample flask was diluted to reach a total system volume of 500 ml. The diafiltration rig was run at a trans membrane pressure of 10 bar and with an average permeance of 1.7 L m<sup>-2</sup> h<sup>-1</sup>bar<sup>-1</sup>. The diafiltration progress was monitored by taking samples from the first stage, the second stage, and the permeate which were analysed through UPLC-MS. The diol species (Eg<sub>28</sub> and Eg<sub>56</sub>) were detected using an ELSD detector (a calibration curve for the signal-to-concentration relation was prepared beforehand), whereas the 290 nm UV trace was used for the nanostar species. The diafiltration was stopped once the diol reached <1% of the initial concentration. The purified product was collected from first and second stage and the methanol was evaporated to obtain the Hub-PEG-OH as a waxy solid, which was analysed by UPLC-MS and NMR.

*The UPLC-MS data corresponding to products **5b-d** can be found in the appendix (Figure A 2-4).*

**Table 3. Amounts of nanostar-PEG-alcohol (5a-d) used and the corresponding yields for membrane-based synthesis.**

| Reaction       | Mass SM [g] | n Hub [mmol] | Equivalents BB | n BB [mmol] | Mass BB [g] | THF [ml] | Mass Product [g] | Yield |
|----------------|-------------|--------------|----------------|-------------|-------------|----------|------------------|-------|
| <b>5a → 5b</b> | 3.24        | 0.75         | 6.0            | 4.50        | 7.67        | 55       | 5.29             | 88 %  |
| <b>5b → 5c</b> | 5.25        | 0.65         | 8.0            | 5.23        | 8.93        | 70       | 6.60             | 86 %  |
| <b>5c → 5d</b> | 6.55        | 0.56         | 10.0           | 5.6         | 9.62        | 80       | 8.20             | 95 %  |

*Protocol for chromatographic purification*

Chromatographic purification was performed using a Biotage Isolera flash chromatography instrument equipped with a UV detector.

The crude material was loaded onto silanised silica and packed in a Biotage dry loading vessel. The separation was conducted on a Biotage Sfaer C18 D column (120 g, 100 Å, 30 µm) using H<sub>2</sub>O (+ 0.5% Formic Acid) and MeCN/H<sub>2</sub>O (4:1) as the mobile phase. The gradient was 10-95 % over 12 column volumes. Fractions were detected and collected at 290 nm and subsequently analysed through UPLC-MS. Pure fractions were combined, and the organic solvent stripped off on the rotary evaporator. The aqueous solution was saturated with NaCl and extracted with chloroform (5 times). The combined organic layers were dried over sodium sulfate, filtered off and the solvent was removed under reduced pressure. The product **9b -d** were obtained as waxy solids with the yields stated in **Table 4**.

**Table 4. Amounts of nanostar-PEG-alcohol (9a-d) used and corresponding yields for chromatography-based synthesis.**

| Reaction       | Mass SM [g] | n Hub [mmol] | Equivalents BB | n BB [mmol] | Mass BB [g] | THF [ml] | Mass Product [g] | Yield |
|----------------|-------------|--------------|----------------|-------------|-------------|----------|------------------|-------|
| <b>9a → 9b</b> | 2.40        | 0.56         | 6.0            | 3.36        | 5.73        | 30       | 3.30             | 74%   |
| <b>9b → 9c</b> | 3.25        | 0.41         | 8.0            | 3.28        | 5.59        | 45       | 3.54             | 75%   |
| <b>9c → 9d</b> | 3.50        | 0.30         | 10.0           | 3.00        | 5.11        | 45       | 4.12             | 89%   |

## Diafiltration

### Set-up 2-stage rig

The diafiltration rig was set up as a 2-stage crossflow system (**Figure 27**), with the first stage consisting of three circular cells each equipped with a flat sheet membrane with an active area of 51 cm<sup>2</sup> and the second stage consisting of two cells with membranes of the same area. A diaphragm pump (Wanner, Hydra-Cell G20) was implemented in the first stage and a gear pump in the second stage to minimise concentration polarisation.

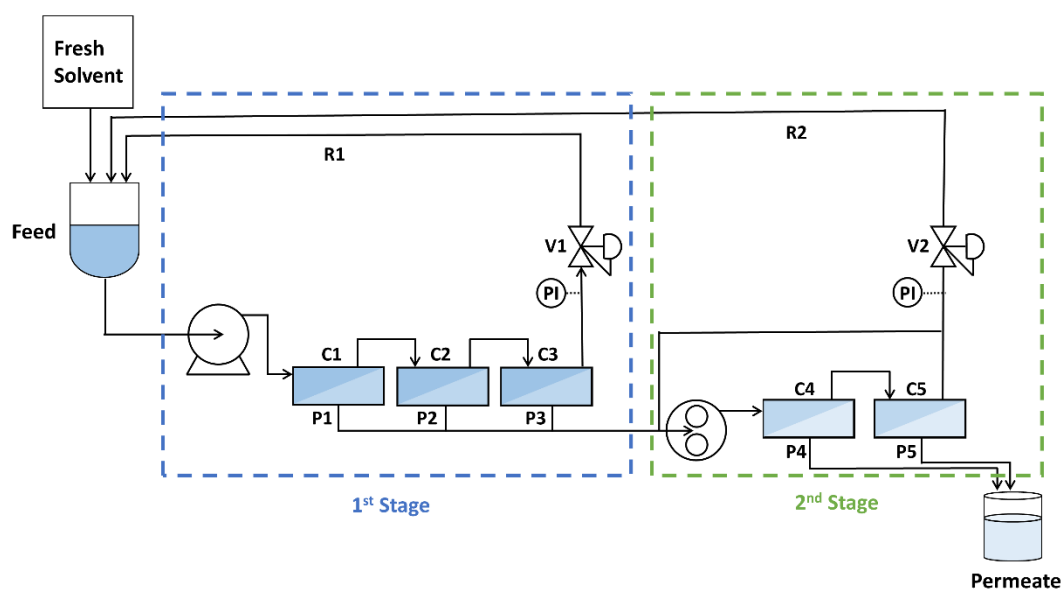


**Figure 27.** Schematic representation of the 2-stage diafiltration rig used for this work. **C1-5:** Membrane cells, each equipped with a flat sheet of DBX-crosslinked PBI membrane. **R1/R2:** Retentate lines. **P1-5:** Permeate lines.

### Membrane Screening

In each screening, five membranes were screened at the same time. The different PEG-preserved membranes were cut into circular pieces with a diameter of 8.1 cm and soaked in methanol for 5 minutes. Subsequently they were implemented into the circular cells together with a stainless-steel sinter disk of the same size. The cells were filled with methanol to avoid drying out of the membranes and then tightened and implemented into the 2-stage rig. The system was washed

with 500 mL of methanol and then ran at 10 bar for at least 8 hours to fully compact the membranes.

A screening solution consisting of Hub(Eg<sub>28</sub>-OH)<sub>3</sub> (1.4 g), Hub(Eg<sub>56</sub>-OH)<sub>3</sub> (0.9 g), Eg<sub>28</sub> Diol (0.4 g) and Eg<sub>56</sub> Diol (0.5 g) in 100 ml methanol was prepared, filtered through a paper filter and transferred to the rig. The feed flask was topped up to reach a total system volume of 500 ml. After distributing the screening solution evenly through the system, the rig was run at 10 bar whilst feeding the permeate back into the feed flask. Samples of 0.50 mL volume were taken from each cells permeate line (C1- C5) and the second-stage retentate (R2) and a sample of 0.10 mL was taken from the first-stage retentate (R1) and diluted in a ratio of 1:2.

#### Calculation of Rejections

The samples were analysed on an Agilent 1290 Infinity II UPLC equipped with a DAD Detector, a 1290 Infinity II ELSD detector, and a 6130 Quadrupole mass spectrometer (operated in positive electrospray ionisation mode). The column was a Acquity UPLC BEH C18 1.7um, 2.1 x 50 mm and the mobile phases used were H<sub>2</sub>O (with 50 mM NH<sub>4</sub>OAc) and MeCN/MeOH (4:1) at a flow rate of 0.3 ml/min and a column temperature of 60 °C. The samples were run with an 8 minute long gradient from 20 % B to 95 % B.

For calculating the Hub species rejections for a certain cell/membrane, the respective peaks in the UV chromatograms at 290 nm were integrated. For the first-stage cells the rejections were calculated as follows:

$$R_i = 1 - \frac{A_i (P)}{3 \times A_i (R1)} \quad (11)$$

With  $R_i$  being the rejection for the species  $i$  (Hub(Eg<sub>28</sub>-OH)<sub>3</sub> or Hub(Eg<sub>56</sub>-OH)<sub>3</sub>),  $A_i (P)$  being the peak area corresponding to species  $i$  in the 290 nm chromatogram derived from the permeate line sample and  $A_i (R1)$  the peak area corresponding to species  $i$  in the 290 nm chromatogram derived from the first-stage retentate.

The rejections for the second-stage cells were calculated in a similar manner, but without using the correction factor of 3, as the R2 samples were not diluted.

For the rejections of the PEG Diols (Eg<sub>28</sub> Diol and Eg<sub>56</sub> Diol) a calibration curve was prepared for the UPLC-ELSD signal, as (in contrast to the UV signal) the ELSD signal is not proportional to the concentration of the analysed substance. The calibration curves were prepared using samples of the respective diol with a concentration ranging from 0.00 mg to 1.00 mg (**Figure 28**). For calculating the PEG Diol rejections, the corrected ELSD peaks areas were used.

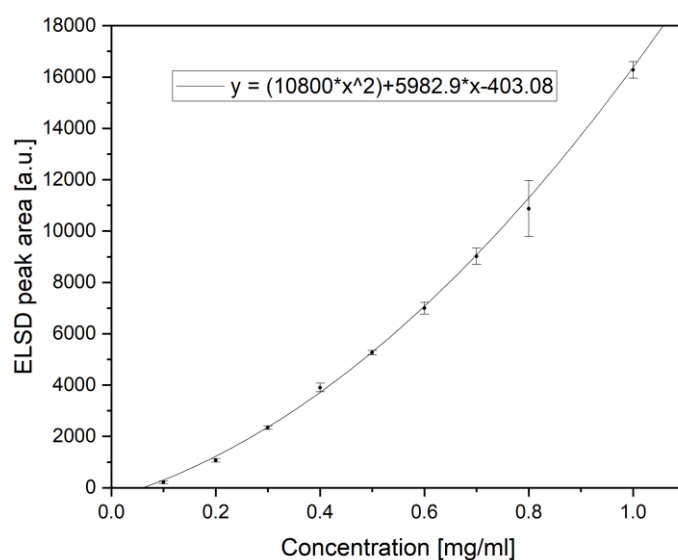


**Figure 28.** UPLC-ELSD signal calibration curve for Eg<sub>28</sub> Diol.

## 4.3 Results and Discussion

### 4.3.1 Synthesis of defined $M_w$ 5 kDa PEG based on a $Eg_{12}$ Building Block

The early Nanostar Sieving based syntheses of defined  $M_w$  5 kDa PEG (112 Eg units  $\triangleq$   $Eg_{112}$ ) within this work were based on using a BB with 12 Eg repeating units ( $Eg_{12}$ ), which was equipped with a 4,4'-dimethoxytrityl (Dmtr) group as the protecting group and a *p*-toluenesulfonyl (tosylate, Tos, Ts) as the leaving group. The 3-armed, benzylic nanostar was coupled to  $Eg_{28}$  diol in a first step, and the resulting  $Hub(Eg_{28}-OH)_3$  was initially purified using reversed-phase chromatography. Starting from  $Hub(Eg_{28}-OH)_3$  a total of seven cycles of chain extension, deprotection and membrane-purification were necessary to reach the desired  $Hub(Eg_{112}-OH)_3$  (**Figure 29**). Three different batches of this type were initially prepared, aiming for increasing the batch size to > 10 g.

Diglyme was used for the  $S_N2$  type chain extension reaction (Williamson etherification), which requires an aprotic and polar solvent that offers high solubility for the PEG starting materials. Furthermore, it has a high boiling point (162 °C) and thus allows for reactions at increased temperatures. Sodium hydride is a widely used base for Williamson etherification reactions of PEG and was also used in this work, as a 60wt% suspension in mineral oil, which simplifies the handling and increases the safety. During the chain extension reaction, the BB reacts with the terminal hydroxyls of the nanostar-PEG (**Figure 29, i**) resulting in extended, Dmtr-capped species. The hydrophobicity of the Dmtr-group contributes to a significant change of retention time of the different species on a RP-UPLC column. The retention time increases with every BB unit attached, with the starting material having the lowest and the fully reacted (3-arm) species having the highest retention time. This, in combination with the strong UV-absorbance of the benzylic nanostar and the defined MS spectra of the involved species, allows for facile monitoring of the chain extension through UPLC-UV-MS (**Figure 31**). During the chain extension the BB tends to undergo a side reaction, in which a BB dimer (DmtrO- $Eg_{56}$ -ODmtr) is formed (peak marked with red box in **Figure 31a**).

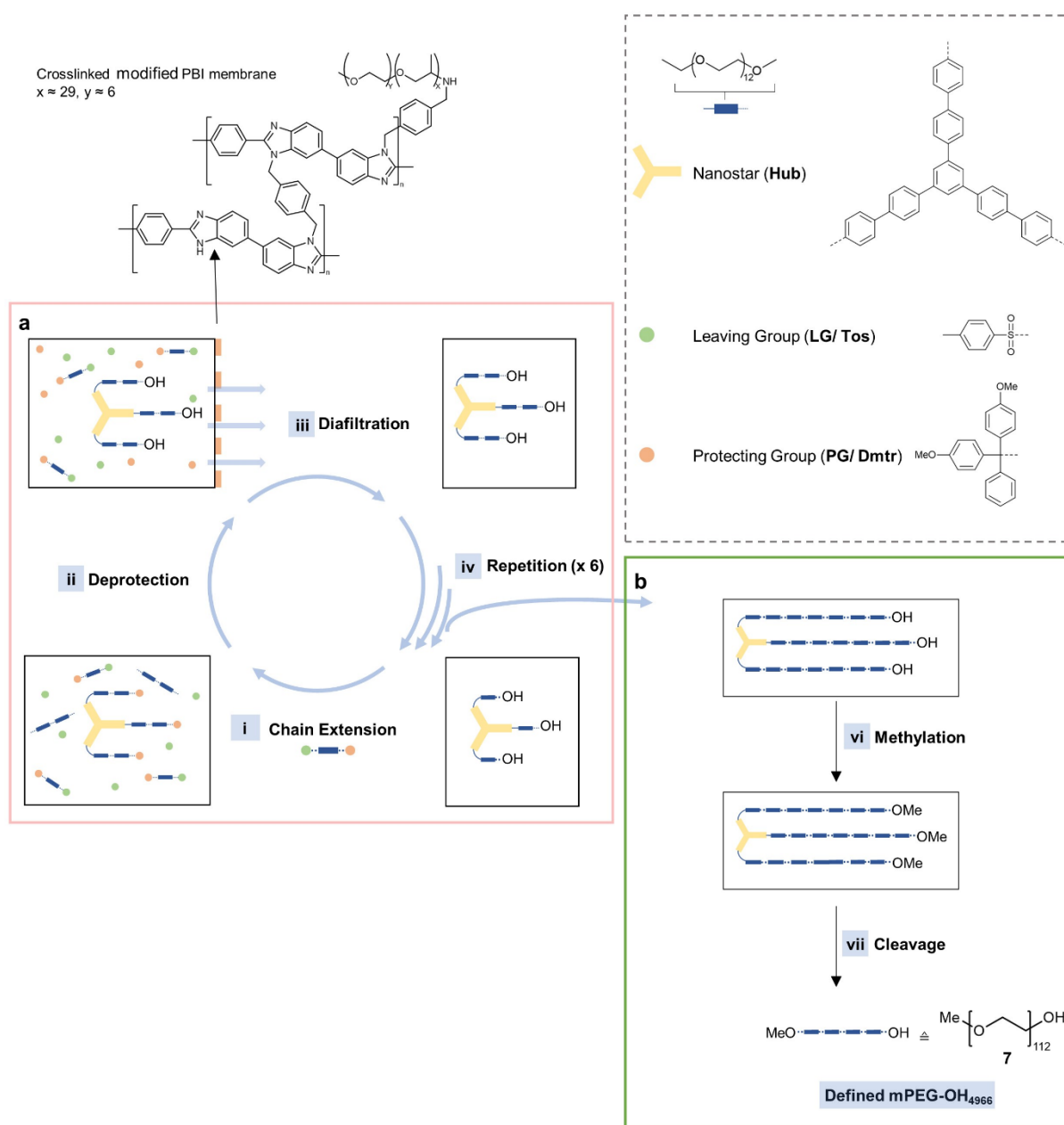


**Figure 29. Nanostar Sieving based synthesis of defined mPEG-OH<sub>4966</sub>.** **a**, chain elongation cycle consisting of the coupling of the DmtrO-Eg<sub>12</sub>-OTs to the free nanostar-PEG alcohol (chain extension, i), removal of the protecting group (deprotection, ii) and diafiltration (iii). Repeating this cycle seven times results in the nanostar-Eg<sub>112</sub>-OH species. **b**, the terminal hydroxyls are then methylated (vi) and subsequently the PEG chains are cleaved from the nanostar (vii) to obtain defined mPEG-OH with a molecular weight of 4966 Da.

Due to its increased size the dimer is the diafiltration limiting species and the membranes must have a MWCO that allows for permeation of the BB and the dimer, whilst the nanostar-PEG species is being rejected,

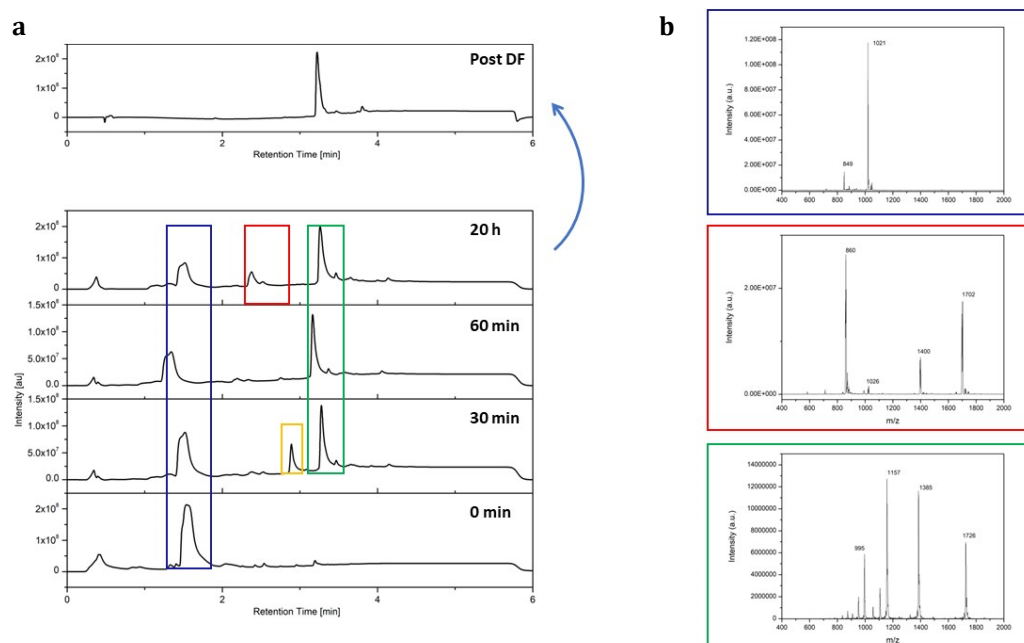


**Figure 30. Example of chain extension monitoring for the first chain extension from Eg<sub>28</sub> to Eg<sub>40</sub>.** **a**, from bottom to top: UPLC-UV traces (full range) of the chain extension at 0 min, 30 min, 60 min, 20 hrs; the peaks can be identified through the underlying ESI<sup>+</sup> trace as the BB (blue), 2-arm species (yellow), the product Hub(Eg<sub>40</sub>ODmtr)<sub>3</sub> (green) and BB dimer (red); the top trace is of the deprotected product (Hub(Eg<sub>40</sub>OH)<sub>3</sub>) after membrane-filtration. **b**, the ESI<sup>+</sup> traces corresponding to the BB (blue), the dimer (red) and the chain extension product (green).

Before the diafiltration the acid-labile Dmtr-protecting group was removed using dichloroacetic acid and pyrrole as a cation scavenger.

The membrane-filtration process used for the Eg<sub>12</sub> based syntheses were previously established in the group. A 2-stage, crossflow system was used, the first stage consisting of three circular cells, each equipped with a flat sheet membrane with an active area of 51 cm<sup>2</sup>, and the second stage consisting of two cells with membranes of the same area.

The membranes used for the Eg<sub>12</sub> based synthesis were DBX-crosslinked PBI membranes, cast with a 18wt% dope solution and modified with 2 kDa PPG polymer. These membranes exhibited

suitable rejection profiles, with a rejection of > 90 % for the  $\text{Hub}(\text{Eg}_{40}\text{-OH})_3$  and > 98 % for the  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$ . The diafiltrations were performed in methanol and were monitored through UPLC-MS. As the BB dimer is converted to  $\text{Eg}_{56}$  diol through the deprotection, it loses its UV-activity and the removal of the dimer was thus followed using the ELSD detector of the UPLC.

An overview and the results of three of the early batches of  $\text{Eg}_{112}$  synthesised using an  $\text{Eg}_{12}$  building is given in **Figure 31**. The chain extensions and deprotections were performed according to the reaction scheme depicted in **Figure 31a**. The chain extension reactions were initially performed in standard glass laboratory equipment (round bottom flasks) and it was suspected that the free silanol groups on the glass surface could potentially react with the sodium hydride and this way slow down the reactions. Hence, three different reactor materials were tested within this work: standard laboratory glass, silanised glass and PTFE. For the silanised glass, a standard

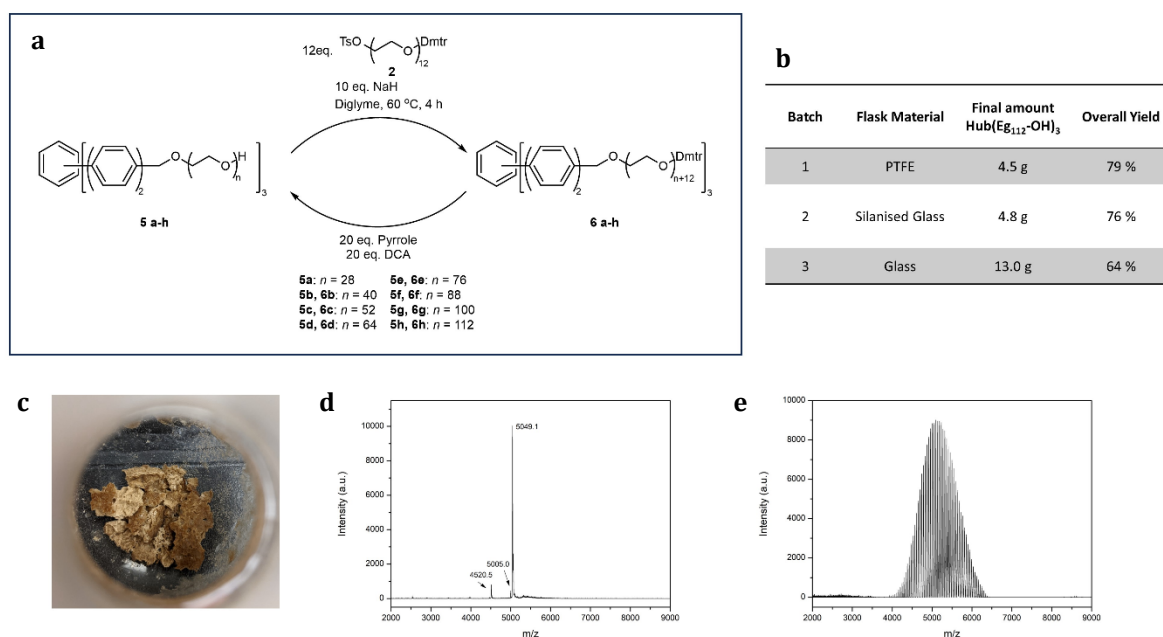


**Figure 31. Overview of nanostar sieving-based synthesis of defined  $M_w$  PEG 5 kDa ( $\text{Eg}_{112}$ ) using an  $\text{Eg}_{12}$  building block. a**, the reaction scheme for the chain extension and deprotection cycle; **b**, summary of the three  $\text{Eg}_{12}$  based syntheses performed within this work, including the flask material, final product amounts and overall yield for the different batches; **c**, photograph of the final  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$  material (derived from batch 3); **d**, MALDI-TOF spectrum of  $\text{MeO-Eg}_{112}\text{-OH}$  ( $\text{mPEG}_{4966}\text{-OH}$ ) derived from batch 3; **e**, MALDI-TOF spectrum of commercially available disperse 5 kDa  $\text{mPEG-OH}$ .

laboratory glass round bottom flask was passivated with trimethylchlorosilane. During the chain extension reactions samples were taken in 30 to 60 minute intervals to monitor the reaction progress (the results are shown in **Table 5**). There were no significant differences observed in the reaction time needed until 100 % conversion was reached.

**Table 5. Overview of reaction times for chain extensions in reactors comprising different materials.**

| Reactor material | Reaction time interval in which 100 % was reached [min] |                                     |                                     |                                     |                                     |                                      |                                       |
|------------------|---------------------------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|--------------------------------------|---------------------------------------|
|                  | Eg <sub>28</sub> → Eg <sub>40</sub>                     | Eg <sub>40</sub> → Eg <sub>52</sub> | Eg <sub>52</sub> → Eg <sub>64</sub> | Eg <sub>64</sub> → Eg <sub>76</sub> | Eg <sub>76</sub> → Eg <sub>88</sub> | Eg <sub>88</sub> → Eg <sub>100</sub> | Eg <sub>100</sub> → Eg <sub>112</sub> |
| PTFE             | 30 - 60                                                 | 60 - 90                             | 120 - 180                           | 120 - 180                           | 120 - 180                           | 120 - 180                            | 180 - 240                             |
| Silanised Glass  | 30 - 60                                                 | 60 - 90                             | 90 - 120                            | 120 - 180                           | 120 - 180                           | 180 - 180                            | 180 - 240                             |
| Glass            | 30 - 60                                                 | 60 - 90                             | 90 - 120                            | 120 - 180                           | 120 - 180                           | 120 - 180                            | 180 - 240                             |

In general, it was observed, that the reaction times increase with increasing chain length of the PEG on the nanostar, which might be due to the increased viscosity of the reaction mixture and a lower overall concentration. With reaction times of below 4 hours, the chain extension reactions up to Eg<sub>112</sub> were still reasonably fast. In the course of investigating the usage of a longer BB and going to higher molecular weights, a study was conducted on the ideal concentration for the chain extensions at higher molecular weight (**Section 4.3.2.2**).

The first two batches (1 and 2 in **Figure 31b**) were performed on a similar scale (4.5 g and 4.8 g of final Hub(Eg<sub>112</sub>-OH)<sub>3</sub>) and obtained similar overall yields (79 % and 76 %). When the batch-size was increased to an intended final amount of around 15 g, the overall yield decreased to 64 %. As the same 2-stage diafiltration system was used for the membrane-purifications, as for the previous batches, the concentration of material in the system was significantly higher. If the concentration of the rejected solute in a membrane-filtration process is too high, it accumulates in a boundary layer between bulk and membrane.<sup>152,153</sup> The increased concentration (and higher gradient between one side of the membrane and the other), results in a higher osmotic pressure,

which in turn results in a decreased rejection of the solute. This concentration polarisation effect explains the observed yield losses when increasing the batch size of the Eg<sub>112</sub> synthesis.

The final material (Hub(Eg<sub>112</sub>-OH)<sub>3</sub>) was methylated with iodomethane, cleaved from the nanostar using BCl<sub>3</sub> and subsequently analysed through MALDI-TOF MS. The resulting spectrum (**Figure 31d**) exhibited one main peak for the MeO-Eg<sub>112</sub>-OH species, and two minor peaks for MeO-Eg<sub>111</sub>-OH and MeO-Eg<sub>100</sub>-OH impurities. The MeO-Eg<sub>100</sub>-OH is a synthesis-related impurity, resulting from either incomplete chain extension reactions or incomplete deprotections, whilst the MeO-Eg<sub>100</sub>-OH is a contamination derived from the Eg<sub>28</sub> starting material (contaminated with Eg<sub>27</sub>). Whilst the MALDI-TOF MS of the such derived Eg<sub>112</sub> allows for straightforward interpretation, the MADLI-TOF MS of a commercially available disperse mPEG 5 kDa derivative (**Figure 31e**) showed the typical Gaussian distribution of PEG of different chain lengths/ molecular weights reaching from ~4 kDa to ~6.5 kDa.

#### 4.3.2 Moving from a Eg<sub>12</sub> to a Eg<sub>28</sub> Building Block

##### 4.3.2.1 Analytical Considerations and Optimisation

The foundation of a chemical process resulting in a product of the highest possible purity are the analytical methods used to monitor the chemical reactions and separation steps. The final goal of the synthesis of defined M<sub>w</sub> PEG is a chain length purity of 100 %, i.e., the sample consists of molecules with only one specific chain length. A first source for chain length impurities is an incomplete chain extension reaction, leaving one or two arms of the nanostar species not extended. This results in Eg<sub>a-nb</sub> impurities, with **a** being the desired number of repeating units in the product, **b** being the number of repeating units in the used BB and **n** being the number of missed chain extensions. Due to the unique absorption maximum of the benzylic nanostar species at 290 nm, the chain extension reaction can be followed using UPLC-MS coupled with a UV-VIS detector.

The relative quantity of the starting material, each intermediate and the product of the chain extension can be derived through integration of the specific peak in the UV chromatogram.

Nevertheless, the reliability of this method strongly depends on the chosen method and the cleanliness of the different parts of the UPLC stack. The different species must be fully separated and signal-suppression through contaminations must be ruled out. Otherwise, the obtained data might not represent the real stage of the reaction, resulting in accumulation of chain length impurities at every step and a disperse final product. Hence, one part of this work focused on optimising the methods used for monitoring the chain extensions and increasing the analytical reliability.

The separation of the different species on a RP-UPLC column is affected by the chosen method, including time, the solvent gradient, the pH and the temperature, as well as the state of the column and the degree of contamination of the machine. **Figure 32a** gives an example for low-quality chromatograms of a chain extension reaction from Eg<sub>52</sub> to Eg<sub>64</sub>, in which the peaks are broad and not well resolved.

**Figure 32b** shows the same samples run on a new column. The peaks are sharp and well resolved, and the 1 h trace (top) shows a peak for the 2-arm species (red), which was not visible on the previously recorded chromatograms using an older column.

As PEG tends to stick to the tubing of the UPLC-MS, it accumulates in the machine over time and the quality of the chromatograms and MS spectra decreases. It was found that regularly washing the machine with a mixture of ammonia and iso-propanol has a drastic effect on the obtained signal-to-noise ratio in the according mass spectra (**Figure 32c and d**).

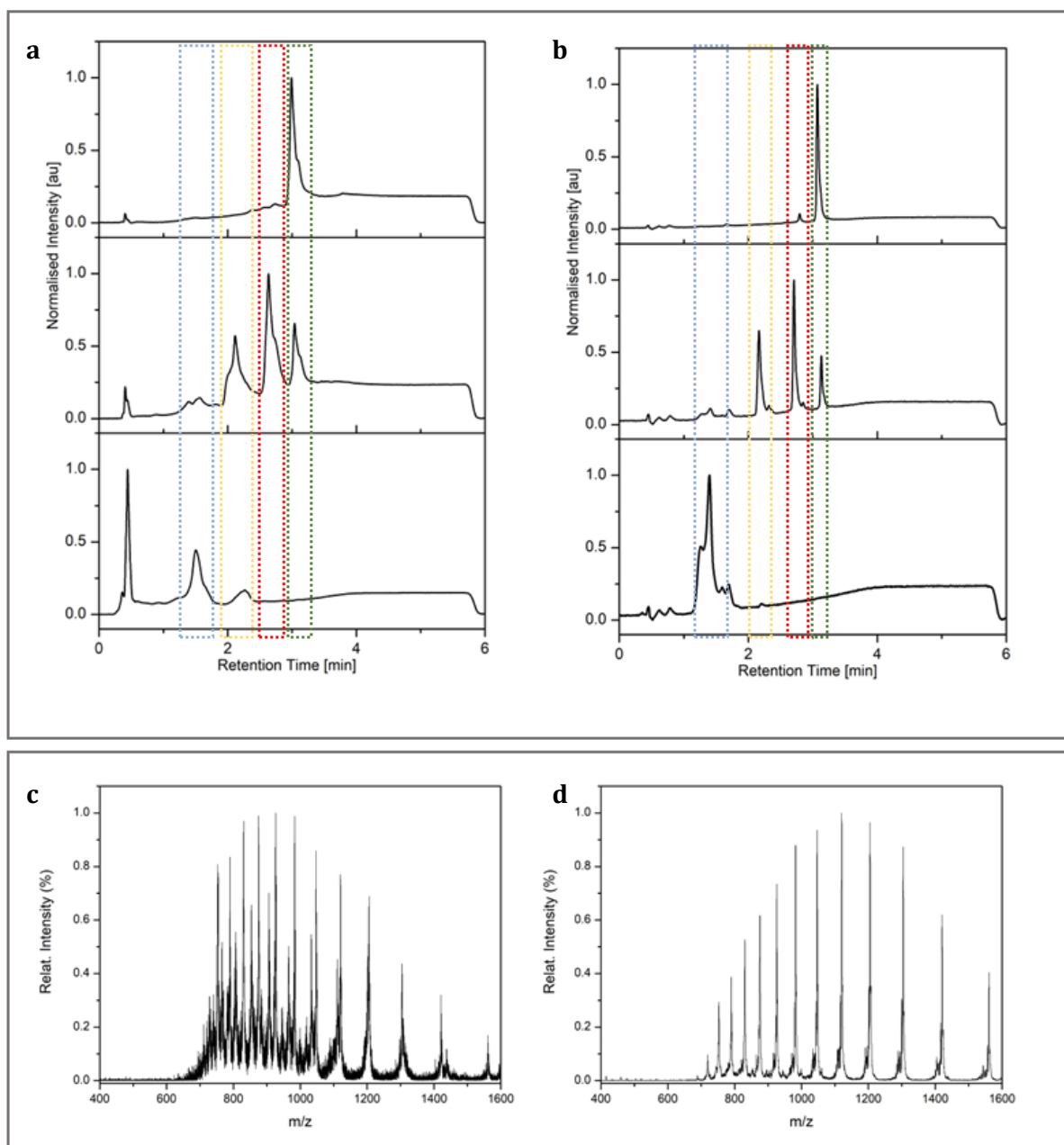


**Figure 32. Comparison of high- and low-quality analytical data for chain extension monitoring.** **A**, UPLC-UV trace (290 nm, 6 min gradient, 70%-95% B) of Eg<sub>52</sub> to Eg<sub>64</sub> chain extension ( $t = 0, 30 \text{ min}$  and  $1 \text{ h}$ ); blue: starting material, yellow: 1-arm reacted, red: 2-arm reacted, green: product. Example for low-quality monitoring, the peaks are broad and not well resolved; **b**, UPLC-UV trace (290 nm, 6 min gradient, 70%-95% B) of Eg<sub>52</sub> to Eg<sub>64</sub> chain extension. Example for high-quality monitoring, the peaks are sharp and well resolved. **c**: ESI<sup>+</sup> spectrum of Hub-Eg<sub>112</sub>-OH species before PEG cleaning procedure; **d**, ESI<sup>+</sup> spectrum of Hub-Eg<sub>112</sub>-OH species after PEG cleaning procedure with 10% ammonia solution/ iso-propanol.

A second source for chain length impurities is incomplete deprotection, which leaves one or more arms of the nanostar species unreactive and as a result unextended in the subsequent chain extension step, which then results again in  $Eg_{a-nb}$  impurities. Conveniently the progress and completion of the deprotection can be followed through UPLC-MS in a way equivalent to the chain extension. With the deprotection of every arm the species is losing a hydrophobic trityl residue, which results in significantly lower retention time on the UPLC. Again, the cleanliness and sensitivity of the machine is crucial to assure the reliability of the data.

### **Purification Monitoring**

A third source for chain length impurities in the final product is the incomplete removal of BB-based impurities. The chain extension reactions are performed with the BB in excess, to assure 100 % conversion. As the BB is bearing a UV-active tosyl group on one end, the removal of BB during diafiltration can be monitored through the UPLC-UV chromatogram. The BB dimer is present as the diol after deprotection and does not exhibit any UV-activity anymore. An evaporative-light scattering detector can be used for monitoring the removal of BB dimer during a diafiltration, but as these detectors are prone to contamination-related sensitivity problems, it is crucial to regularly record calibration curves and to determine the limit of detection. Within this work the limit of detection for  $Eg_{56}$  diol was determined to be below 0.1 mg/ml and hence low enough to monitor the removal of BB dimer.

### **Determination of chain length purity**

As described previously in the literature review (**Section 2.2.2**), the widely used parameter to describe a polymers dispersity is the dispersity index, which is commonly derived through GPC. This method is sufficient to compare the dispersity of polymer samples derived from the same type of conventional polymerisation technique. In case of defined  $M_w$  polymers, a technique for quantifying the actual relative amounts of every chain length present is desirable. MALDI-TOF is a widely used technique to qualitatively visualise a polymers mass distribution and has also been

used by researchers to quantitatively measure the chain length impurities present in defined  $M_w$  PEGs.

Within this work, the final  $Eg_{112}$  samples were analysed using MALDI-TOF MS in linear and reflective mode and at different laser powers. A selection of the results is presented in **Figure 33**. It became apparent, that the relative intensity of the signals is largely dependent on the laser

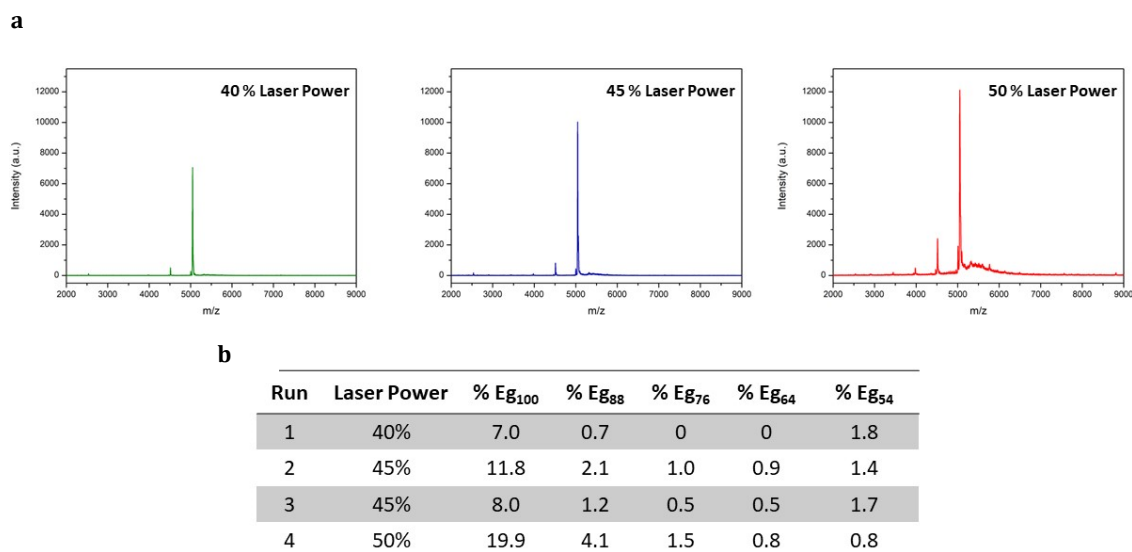


**Figure 33. Results of the MALDI-TOF analysis (reflective mode, 1000 shots) of the same sample of MeO- $Eg_{112}$ -OH.** **a**, MALDI-TOF spectra of the same MeO- $Eg_{112}$ -OH sample obtained at different laser powers; **b**, corresponding calculated relative chain length impurities calculated from the different spectra (the spectrum of run 2 is not depicted).

power employed. At low laser power (40 %), a low overall intensity is reached, and the relative intensity of the chain length impurities is low. With increasing laser power, the total intensity increases, and the relative intensity of the chain length impurities increases as well. Even when the same laser power is used for measuring the same sample, significantly different results are obtained for different runs. For example, when the same sample of MeO- $Eg_{112}$ -OH was measured twice at a laser power of 45 %, from one run (**Figure 33b, 2**) a contamination with  $Eg_{100}$  of 11.8 % was calculated, whilst the second run (**Figure 33b, 3**) resulted in a contamination with  $Eg_{100}$  of 8.0 %. These results underline that MALDI-TOF MS can be used as a qualitative analytical method, but is not a suitable technique to measure the absolute chain length purity of a polymer.

Within this work one approach was developed to reliably determine the synthesis-related chain length purity of a defined  $M_w$  PEG 5kDa sample, which will be discussed below (**Section 4.3.3.2**).

#### **4.3.2.2 Chain Extension Reaction Kinetics**

##### **Reaction Concentration**

As the chain extension reactions were observed to get slower with increasing chain length, a study was performed to identify the ideal concentration of reactants when moving to a larger BB and higher molecular weight PEG. Small scale chain extension reactions starting from  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$  and using an  $\text{Eg}_{28}$  BB were performed at three different concentrations, samples were taken from the reaction every 10 minutes to 1 hour and progress was monitored based on the UV chromatogram at 290 nm. Diglyme was used as the solvent at 60 °C, with 12 equivalents of BB and 10 equivalents of sodium hydride in each reaction. Three different weight percentages were tested (10 wt%, 20 wt% and 35 wt%). 35 wt% was chosen as the upper limit, as above that the reaction mixture was too viscous to stir with a magnetic stirrer. The dilute reaction (10 wt%) was slow, with the reaction reaching full conversion after between 2 and 3 hours. It was found that 20 wt% delivered the fastest reaction time, with 100 % conversion reached after 50 minutes, whereas the 35 wt% reactions were slightly slower (60 minutes), which is presumably due to the very high viscosity of the reaction mixture. In general, the reaction times were surprisingly low compared to previously performed chain extensions. This might be due to a scale effect, as the previous chain extension reactions were all performed on a multi gram scale, whereas the chain extensions for this experiment were performed on a small scale with only 150 mg of  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$ .

Based on these results it was decided to perform all chain extension involving the Eg<sub>28</sub> BB with a concentration of 20 wt%.

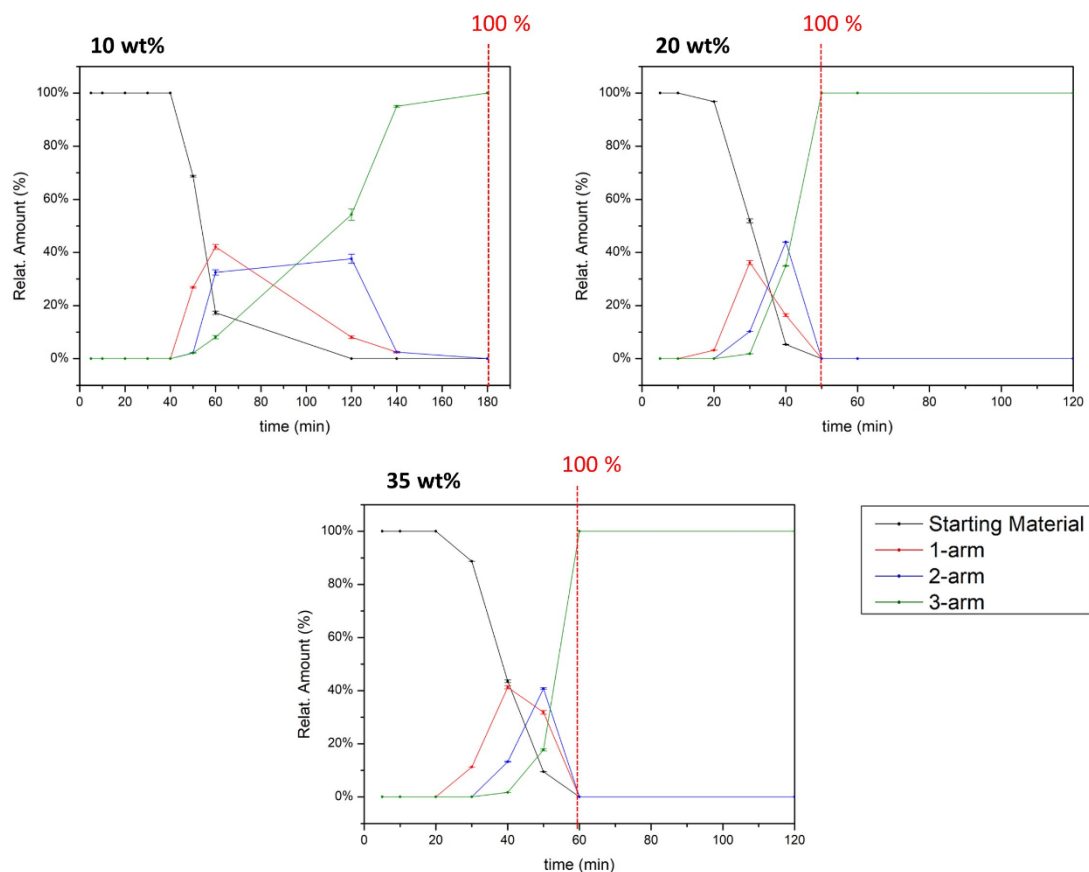


**Figure 34. Overview of the reaction progress for a Eg<sub>112</sub> to Eg<sub>140</sub> chain extension at different reaction concentrations.** The weight percentages (10 wt%, 20 wt% and 35 wt%) are the weight percentages of combined starting materials (Hub(Eg<sub>112</sub>-OH)<sub>3</sub> and Eg<sub>28</sub> building block) in the solvent (diglyme). The red dotted line indicates when the reactions reached 100 % conversion.

### Stoichiometry

It was noted that the chain extension reactions become slower the longer the PEG chain gets. At the same time the chain extension is competing with the formation of dimer, which is consuming the valuable BB. In early syntheses within this work, 12 equivalents of BB (4 equivalents per arm) were used in each chain extension, to assure 100 % conversion of the reaction. When moving from using the Eg<sub>12</sub> based BB, to using the Eg<sub>28</sub> BB, an efficient stoichiometry of the reaction becomes even more important due to the higher cost of the Eg<sub>28</sub> diol (£50/g for Eg<sub>28</sub> diol

compared to £25/g for Eg<sub>12</sub> diol, as of December 2022, from Polypure AS). Hence, the three chain extension steps involved to reach the Eg<sub>112</sub> derivative, starting from Hub(Eg<sub>28</sub>-OH)<sub>3</sub> were tested on a small scale with different amounts of BB, starting with a minimum of 6 equivalents. The reactions were left stirring for 24 hours and the progression of the reaction was determined through UPLC-MS. It was found that the first chain extension (Eg<sub>28</sub> to Eg<sub>56</sub>) only requires 6 equivalents of BB to reach 100 % conversion (**Table 6**). For the next step (Eg<sub>56</sub> to Eg<sub>84</sub>) the reaction did not go to completion with 6 equivalents of BB and after 24 hours all BB reacted to the dimer species. When the same reaction was performed with 8 equivalents of BB, the reaction reached 100 % conversion in 24 hours. For the last chain extension (Eg<sub>84</sub> to Eg<sub>112</sub>) 10 equivalents of BB were necessary to achieve 100 % conversion.

**Table 6.** Overview of results on the reaction completion of chain extension with different amounts of BB used.

| Chain Extension Step    | Eg <sub>28</sub> →Eg <sub>56</sub> | Eg <sub>56</sub> →Eg <sub>84</sub> | Eg <sub>56</sub> →Eg <sub>84</sub> | Eg <sub>84</sub> →Eg <sub>112</sub> | Eg <sub>84</sub> →Eg <sub>112</sub> | Eg <sub>84</sub> →Eg <sub>112</sub> |
|-------------------------|------------------------------------|------------------------------------|------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Equivalents of BB used  | 6.0                                | 6.0                                | 8.0                                | 6.0                                 | 8.0                                 | 10.0                                |
| Conversion after 24 hrs | 100 %                              | 96 %                               | 100 %                              | 93 %                                | 95 %                                | 100 %                               |
| BB left? (yes/no)       | yes                                | no                                 | yes                                | no                                  | no                                  | yes                                 |

Based on these results, all following chain extensions based on the Eg<sub>28</sub> BB were performed with 6 equivalents of BB for the first (Eg<sub>28</sub> to Eg<sub>56</sub>), 8 equivalents for the second (Eg<sub>56</sub> to Eg<sub>84</sub>) and 10 equivalents for the third (Eg<sub>84</sub> to Eg<sub>112</sub>) chain extension to reach full conversion and high efficiency.

#### 4.3.2.3 Critical Investigation of Impurity Formation

Besides chain length-related impurities, the formation of chemical impurities (i.e. impurities bearing a different end-function than -OH) has to be avoided. During the optimisation of the Nanostar Sieving based synthesis of defined M<sub>w</sub> PEG different side reactions were observed. A first source for impurity formation are contaminations in the starting materials (the Hub(Eg<sub>28</sub>-

OH)<sub>3</sub> or the BB). A main advantage of working with defined M<sub>w</sub> PEGs is, that they are exhibiting superior separation behaviour on RP-UPLC and the different species exhibit clear mass spectra from which the exact masses can be calculated. This allows for facile mechanistic investigations and troubleshooting in case impurities are formed.

For example, during one of the early Eg<sub>12</sub> based syntheses, a formation of a significant amount (about 20 % relative to the main product species) of a side-product occurred during the first chain extension from Eg<sub>28</sub> to Eg<sub>40</sub>, which was visible as a peak with a slightly lower retention time than the main product peak (**Figure 35a**). The underlying mass spectrum showed a m/z series matching with a mass of 6015 Da. The lower retention time (in the range of the 2-arm species) indicated that the species was only bearing two trityl groups and was a not fully extended version of the nanostar-PEG. The mass matched with 2-arm extended species, where the third arm is capped with an ethyl unit. Further investigations revealed that the BB used contained an impurity, which was visible from the UPLC chromatogram, but did not have an underlying mass spectrum. The contamination was chromatographically isolated, and the <sup>1</sup>H-NMR proved this contamination to be ethyl tosylate (**Figure 35b**), which has a mass so low (200.3 Da) that it was not detected by the ESI-MS. The <sup>1</sup>H NMR spectrum of the BB shows very low intensity peaks for the ethyl tosylate species, which disappeared after further purification (**Figure 35b**). It is possible that the ethyl tosylate has a much higher reactivity towards the PEG-hydroxyls than the BB, which would explain why even a small contamination would result in a significant amount of ethyl-capped species. It is not known where the ethyl tosylate contamination was derived from, but one potential source could be traces of ethanol which were present during the tosylation reaction to derive the BB.

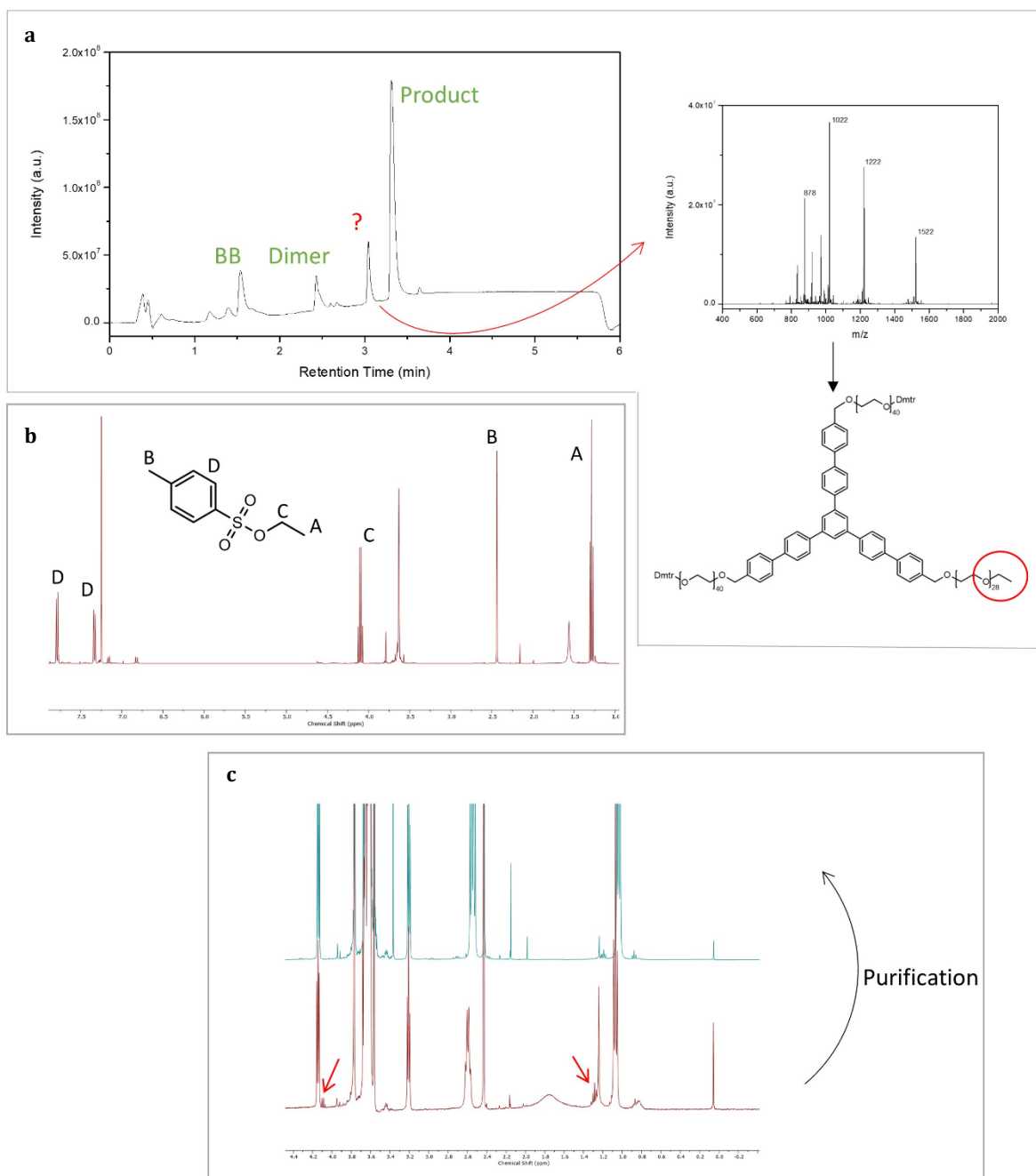


**Figure 35. Overview of a side-reaction involving the reaction with an ethyl tosylate contamination.** **a**, final UPLC-UV chromatogram of an Eg<sub>28</sub> to Eg<sub>40</sub> chain extension, exhibiting one unknown peak besides the signals for the building block (BB), dimer and the product; The underlying ESI-MS of the impurity peak reveals an m/z series of 1522, 1222., 1021 and 878, matching with a mass of 6015 Da; the proposed impurity structure is the Hub(Eg<sub>40</sub>ODmtr)<sub>2</sub>(Eg<sub>28</sub>OEt). **b**, <sup>1</sup>H spectrum of ethyl tosylate isolated from the building block. **c**, <sup>1</sup>H spectra of the building block before purification (bottom) and after purification (top).

This investigation presents an important illustration of how using defined  $M_w$  PEGs allows for reliable monitoring of reactions and contributes to the final purity of the product.

In the early Nanostar Sieving based syntheses based on the  $Eg_{12}$  BB, the UPLC-UV chromatograms exhibited peaks besides the main product peak, that could be assigned to species attached to the nanostar, due to their maximum absorbance at 290 nm. The underlying mass spectra did not show discrete masses, which could result from either different chemical species overlapping due to similar retention times, or from depolymerisation of the PEG chain. When an equivalent synthesis was performed using chromatography as means of purification (instead of membrane purification), these impurities decreased (**Figure 36a**). Furthermore, the nanostar intermediates derived after each chain extension step had a brown colour (that developed during the chain extension), which increased in intensity with each cycle, with the final  $Hub(Eg_{112}-OH)_3$  being obtained as a dark brown solid. Compared to that, the material derived from the chromatography-based synthesis was a colourless solid (**Figure 36b**). To investigate the hub-related impurities, a first comparative study was performed, in which the nanostar-PEG species from different steps of different  $Eg_{112}$  batches were chromatographically separated on the UPLC. The 290 nm UV-chromatograms were analysed, and each peak was integrated, and the relative (hub-related) purity of the main nanostar-PEG species was calculated. Three Nanostar Sieving based batches and one chromatography-based batch were compared to each other, and the results are shown in **Figure 36c**, which shows the relative Hub purity in dependency of the chain length for each step between a chain length of  $Eg_{64}$  and  $Eg_{112}$ . It became apparent that the Hub-related purity decreased with each step for the Nanostar Sieving -based batches (black, blue and red graphs). In the Nanostar Sieving-based batch number 3 (red graph) the final  $Hub(Eg_{112}-OH)_3$  was purified through column chromatography and the Hub-related purity increased from about 70 % to about 90 %, indicating that the impurities were removed. These results suggested that there were impurities forming during the chain extension/ deprotection reactions that were related to the growing PEG chains attached to the nanostar. These impurities seem to exhibit different retention times on a reversed-phase column, compared to the desired nanostar-PEG alcohol species, which

explains why they would be removed through reversed-phase chromatography, but not through Nanostar Sieving, as they would exhibit similar sizes to the desired Hub-PEG alcohol species.

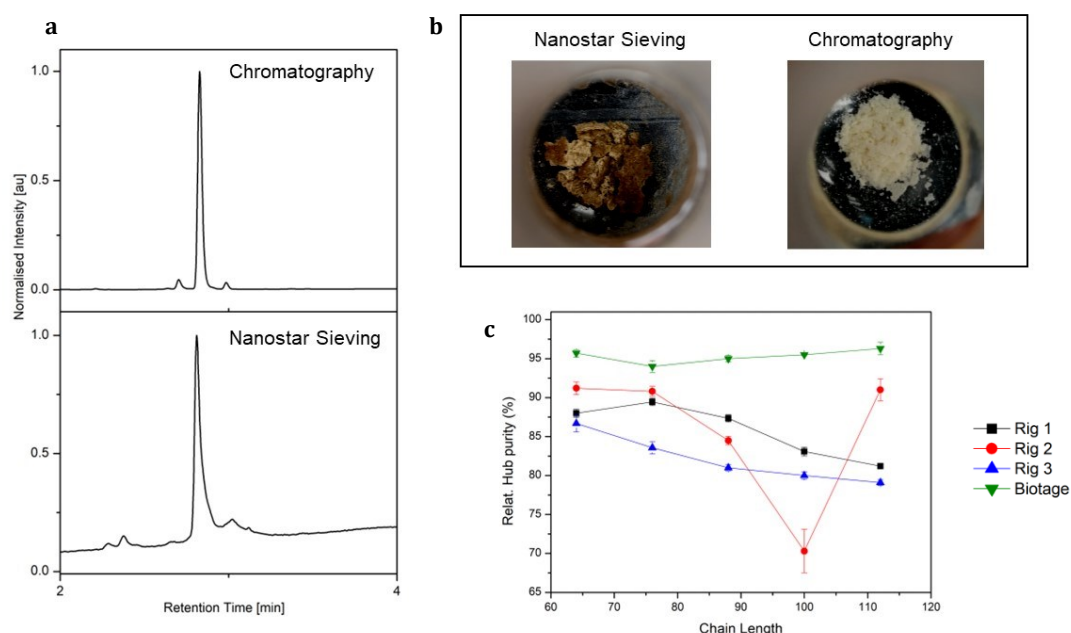


**Figure 36. Overview of the Hub-related purity of different Hub-PEG species derived from Nanostar Sieving based and chromatography-based synthesis.** **a**, UPLC-UV traces of final Hub(Eg<sub>112</sub>-OH)<sub>3</sub> material derived from a Nanostar Sieving based synthesis (bottom) and a chromatography-based synthesis (top); **b**, photographs of the final Hub(Eg<sub>112</sub>-OH)<sub>3</sub> material derived from the Nanostar Sieving based synthesis (left) and a chromatography-based synthesis (right); **c**, relative Hub purities for the intermediated Hub(Eg<sub>64</sub>-OH)<sub>3</sub> to Hub(Eg<sub>112</sub>-OH)<sub>3</sub> for three Nanostar Sieving based batches (Rig 1-3) and one chromatography-based batch (Biotage).

A further study was performed, in which different PEG species were dissolved in either THF or diglyme and then treated with sodium hydride for 20 hours. Samples were taken and analysed on the UPLC. For DmtrO-Eg<sub>28</sub>-ODmtr, no change was observed in either the colour of the reaction or the UPLC chromatogram. For DmtrO-Eg<sub>28</sub>-OTs, the formation of dimer (DmtrO-Eg<sub>56</sub>-ODmtr) was observed, but besides that no other side-products were detected and there was no change in colour. When the same experiment was performed with PEG species bearing a free hydroxyl group (like DmtrO-Eg<sub>28</sub>-OH and Hub(Eg<sub>28</sub>-OH)<sub>3</sub>, the colour of the solution turned dark brown overnight. The UPLC chromatograms showed the appearance of new peaks (**Figure 37a**), with the underlying MS traces exhibiting a degradation profile with a mass difference of -44 Da

between each species. It was possible to extract the  $m/z$  series of the (heaviest) main species of each peak and calculate the according mass (**Table 7**) For example, for the DmtrO-Eg<sub>28</sub>-OH, one peak was detected at lower retention time than the starting material, with an underlying mass spectrum matching with a mass of about 1570 Da, which is 16 Da heavier than the starting material (1554 Da). This indicated the addition of one oxygen atom. Another species was detected at a higher retention time with a mass of about 3122 Da, which is roughly double the mass of the starting material (+ 14 Da) and indicated the formation of a dimer.

**Table 7. Summary of the three detected species after treatment of DmtrO-Eg<sub>28</sub>-OH or Hub(Eg<sub>28</sub>-OH)<sub>3</sub> with NaH.**

|   | Species                                | T<br>[°C] | Solvent | Peaks detected after 20h |             |                      | Identified Species<br>(M <sub>w</sub> [Da])                             | % of<br>SM<br>Peak<br>(UV) |
|---|----------------------------------------|-----------|---------|--------------------------|-------------|----------------------|-------------------------------------------------------------------------|----------------------------|
|   |                                        |           |         |                          | RT<br>[min] | Approx. Mass<br>[Da] |                                                                         |                            |
| A | DmtrO-Eg <sub>28</sub> -OH             | 50        | THF     | 1.                       | 2.33        | 1570                 | DmtrO-Eg <sub>28</sub> -OH(OH) (1569.9)                                 | 4.2                        |
|   |                                        |           |         | 2.                       | 3.37        | 1554                 | DmtrO-Eg <sub>28</sub> -OH (1553.9)                                     | -                          |
|   |                                        |           |         | 3.                       | 4.01        | 3122                 | Peroxide-Dimer (3121.7)                                                 | 11.1                       |
| B | DmtrO-Eg <sub>28</sub> -OH             | 50        | Diglyme | 1.                       | 2.33        | 1570                 | DmtrO-Eg <sub>28</sub> -OH(OH) (1569.9)                                 | 0.5                        |
|   |                                        |           |         | 2.                       | 3.37        | 1554                 | DmtrO-Eg <sub>28</sub> -OH (1553.9)                                     | -                          |
|   |                                        |           |         | 3.                       | 4.01        | 3122                 | Peroxide-Dimer (3121.7)                                                 | 2.2                        |
| C | Hub(Eg <sub>28</sub> -OH) <sub>3</sub> | 50        | THF     | 1.                       | 2.49        | 4341                 | Hub(Eg <sub>28</sub> -OH) <sub>2</sub> (Eg <sub>28</sub> -OOH) (4341.2) | 11.3                       |
|   |                                        |           |         | 2.                       | 3.06        | 4323                 | Hub(Eg <sub>28</sub> -OH) <sub>3</sub> (4523.2)                         | -                          |
|   |                                        |           |         | 3.                       | 3.39        | 8664                 | Peroxide-Dimer (8664.4)                                                 | 1.8                        |
| D | Hub(Eg <sub>28</sub> -OH) <sub>3</sub> | 50        | Diglyme | 1.                       | 2.49        | 4341                 | Hub(Eg <sub>28</sub> -OH) <sub>2</sub> (Eg <sub>28</sub> -OOH) (4341.2) | 5.0                        |
|   |                                        |           |         | 2.                       | 3.06        | 4323                 | Hub(Eg <sub>28</sub> -OH) <sub>3</sub> (4325.2)                         | -                          |
|   |                                        |           |         | 3.                       | 3.39        | 8664                 | Peroxide-Dimer (8664.4)                                                 | 2.1                        |

These results, in combination with the dark brown colour of the solution, suggested the occurrence of oxidation or peroxidation processes. The oxidative degradation and autooxidation of PEG have been reported in literature before.<sup>154-159</sup> The formation of PEG based radicals can either be thermal- or light-induced or through chemical initiators or the presence of transition

metals.<sup>160</sup> In case of the work in hand a likely source of radical species is the solvent, THF or diglyme, which are both peroxide forming solvents.

Based on the detected masses, the dimeric peroxide structure **9** and the diol **10** shown in **Figure 37b** are the species that are formed through the oxidation/ peroxidation of DmtrO-Eg<sub>28</sub>-OH and the masses detected within the experiment with Hub(Eg<sub>28</sub>-OH)<sub>3</sub> match with the according dimer and diol.

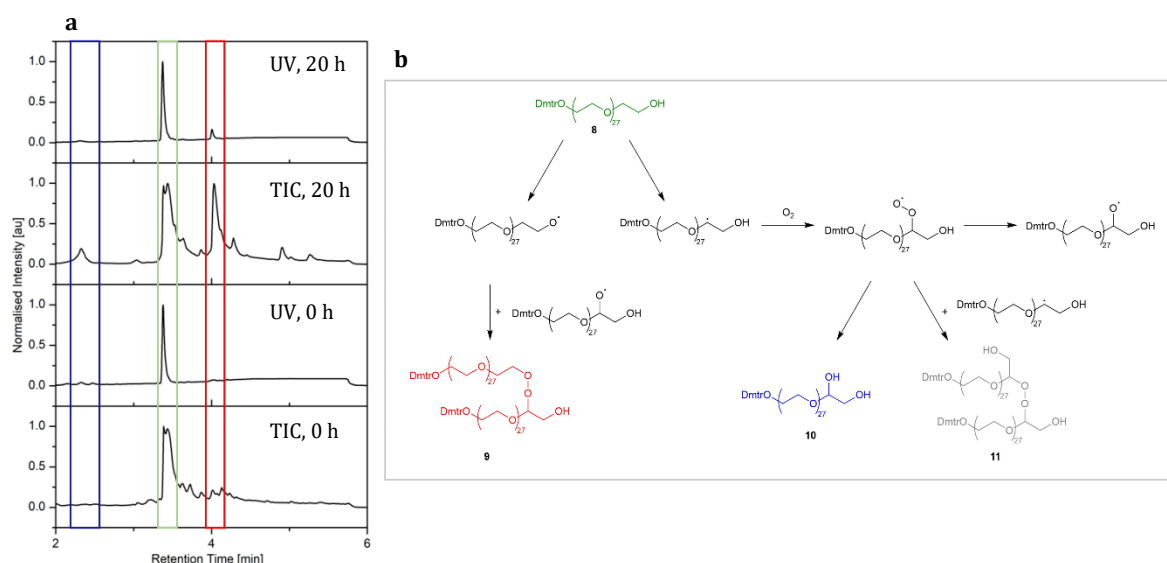


**Figure 37.** **a**, UPLC-UV and TIC traces of a solution of Dmtr-Eg<sub>28</sub>-OH in THF before (0 h) and 20 h after adding sodium hydride (20 h); **b**, the proposed pathway for the oxidation of DmtrO-Eg<sub>28</sub>-OH, resulting in the formation of peroxide dimer **9** and diol **11**.

As these species cannot be removed through membrane-purification and result in decrease the overall chemical purity of the final PEG material, it is of great interest to find a solution for this problem. The peroxide content in the used solvents should be investigated. A source for oxygen would be not sufficiently pure inert gas (Argon) used during the synthesis, which could be overcome by adding a filter unit.

### **4.3.3 Defined $M_w$ PEG synthesis *via* Chromatography vs. Nanostar Sieving- a comparative study**

The final aim within this chapter was to synthesise defined  $M_w$  PEG 5 kDa based on Nanostar Sieving and using an Eg<sub>28</sub> BB. The previously performed studies showed that the chain extension with Eg<sub>28</sub> was kinetically feasible and optimised reaction and monitoring conditions allowed for a low BB consumption and a high final purity. Within this chapter a comparative study was conducted, in which a defined  $M_w$  linear 5 kDa mPEG-OH was synthesised, with one synthesis solely based on membrane-purification, whereas in the second one automated chromatography was used for the purification between each step.

#### **4.3.3.1 Membrane selection and synthesis**

The remaining challenge was to find membranes that would allow for the removal of the bigger Eg<sub>56</sub> dimer, whilst still rejecting the nanostar-PEG species. Due to their high solvent permeance, high stability and high estimated separation factors, different DBX-crosslinked PBI membranes were screened to investigate their suitability for the defined PEG synthesis. As previous work showed, the pore size of PBI membranes can be changed by altering the concentration of PBI in the dope solution<sup>145</sup>. To select the right membranes for the different separation steps, PBI membranes synthesised with different concentrations of PBI in the dope solution were screened. Additionally, crosslinked PBI membranes modified with 2 kDa poly(propylene glycol) (PPG) chains are known to show improved anti-fouling behaviour in methanol, a relatively cheap, low-boiling solvent, which can easily dissolve all PEG-based species present in the reaction mixtures.<sup>145</sup> Hence, a combination of PPG-modified, crosslinked PBI membranes and methanol as the solvent was chosen for the diafiltrations in this work.

During the loading of DmtrO-Eg<sub>28</sub>-OH to the Hub-tribromide species, no dimer formation occurs, and the separation-limiting species (i.e., the largest species to remove) is the Eg<sub>28</sub> Diol. In the next following three cycles DmtrO-Eg<sub>28</sub>-OTs was coupled to the nanostar-PEG-alcohol species. One

side-product is the building block dimer, that is present as the Eg<sub>56</sub> Diol after deprotection and represents the separation-limiting species for these steps.

Membrane screening was performed in the same system as the subsequent synthesis, a two-stage crossflow set-up, consisting of five membrane cells with separate sample outlets. For the screening, a solution of the biggest species to remove (Eg<sub>28</sub> Diol, Eg<sub>56</sub> Diol) and the smallest species to retain (Hub(Eg<sub>28</sub>-OH)<sub>3</sub> and Hub(Eg<sub>56</sub>-OH)<sub>3</sub>) in methanol was used. Samples were taken from the retentate and the permeate line of each cell and the rejections were calculated from the UPLC-ELSD chromatogram (diols) and the UPLC-UV (290 nm) chromatogram (hub species) respectively. A 16 wt% membrane (16 wt% of PBI in the dope solution) was identified as suitable for the separation of the first step (coupling of DmtrO-Eg<sub>28</sub>-OH to the Hub-alcohol) with a rejection of 85 % for the Hub(Eg<sub>28</sub>-OH)<sub>3</sub> and 38 % for the Eg<sub>28</sub> Diol (**Figure 38a**). For the next step (the first chain extension) a 14 wt% membrane was chosen, which had a rejection of 88 % for the Hub(Eg<sub>56</sub>-OH)<sub>3</sub> and a rejection of 30 % for the Eg<sub>56</sub> Diol (**Figure 38a**).

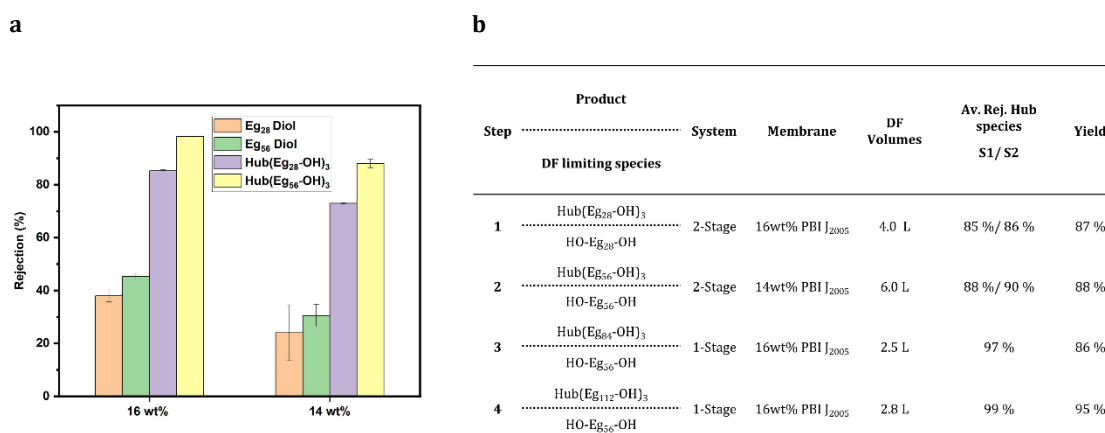


**Figure 38.** Rejection and yield data for the membrane-purification based synthesis of defined mPEG-OH<sub>4966</sub>. **a**, Screening results for the 16 wt% and 14 wt% membranes. Rejections of Eg<sub>28</sub> Diol (orange), Eg<sub>56</sub> Diol (green), Hub(Eg<sub>28</sub>-OH)<sub>3</sub> (purple) and Hub(Eg<sub>56</sub>-OH)<sub>3</sub>. **b**, set-up and results for the four steps of the membrane-based synthesis of Hub(Eg<sub>112</sub>-OH)<sub>3</sub>.

For the final membrane-based synthesis two different set-ups were used: in the first two steps the rejection for the nanostar-PEG species was close to 90 % and a two stage system was used, as this has been shown to increase the yield significantly.<sup>147</sup> Due to their increased size the third

(Hub(Eg<sub>84</sub>-OH)<sub>3</sub>) and fourth (Hub(Eg<sub>112</sub>-OH)<sub>3</sub>) intermediate allowed for purification using the 16 wt% membrane in a single stage system, which decreased the needed diavolumes for complete purification significantly. The synthesis was done on a multi-gram scale and resulted in an overall yield of 63 % with 8.3 g of final Hub(Eg<sub>112</sub>-OH)<sub>3</sub> after four iterations of the cyclic process starting from the Hub-tribromide (the yields for each step are given in **Figure 38b**).

#### 4.3.3.2 Comparison to a chromatography-based synthesis

An equivalent synthesis (starting from the same amount of starting material) was performed using automated reversed-phase chromatography for the purification in each cycle, whilst keeping all chain extension and deprotection conditions the same. The final MeO-Eg<sub>112</sub>-OH (mPEG-OH<sub>4966</sub>) was then comparatively analysed in terms of its purity. Despite MALDI-TOF being widely used for determining a final polymers chain length purity, it is not a quantitative method, as previously shown in section 4.3.2.1. Hence, a quantitative method relying on UPLC-MS was developed instead, using the MS trace for identifying the UV chromatogram for relatively quantifying each species. The two MeO-Eg<sub>112</sub>-OH samples (derived from the membrane-based synthesis vs. from the chromatography-based synthesis) were functionalised with anisoyl chloride to make them detectable through UV-Vis (**Figure 39a**). A UPLC method was developed, involving using a 15 cm long C<sub>18</sub> UPLC column (instead of the 5 cm long ones normally used within this work), which resulted in successful separation and relative quantification of the MeO-Eg<sub>112</sub>-anisoyl derivative **5** and the respective impurities, including +/- Eg<sub>28</sub> species. -Eg<sub>x</sub> impurities resulting from Eg<sub>27</sub> in the Eg<sub>28</sub> starting material, or potential unzipping of the PEG chain cannot be separated and hence not quantified. It should be noted that each peak showed a clear underlying MS trace (ES<sup>+</sup>) due to the uniformity of the PEG. This allowed for identification of the exact nature of all respective species.

The M-Eg<sub>28</sub> species, derived from either incomplete chain extension or deprotection, was detected for both samples (**Figure 39b**, Peak “1” corresponds to MeO-Eg<sub>84</sub>-X, shorter derivatives were not detected). Also, the M+Eg<sub>28</sub> (**Figure 39b**, Peak “3” corresponds to MeO-Eg<sub>140</sub>-X, longer

derivatives were not detected) species was detected for both samples, which can be attributed to small amounts of TsO-Eg<sub>28</sub>-OH present in the building block. From this the chain length purities (in terms of +/- Eg<sub>28</sub> species) were calculated as 96.8 % for the chromatography-based synthesis and 95.8 % for the membrane-based synthesis. The samples derived from the membrane-based synthesis showed two more peaks at higher retention times, that could not be identified (**Figure 39b** Peak "4" and "5"). It is suspected that they are derived from side reactions on the PEG chain end, that result in impurities with a reverse-phase retention behaviour significantly different to the Hub-PEG-alcohol species. This would result in removal during chromatographic purification, whereas they cannot be removed by membrane-purification due to the same size of the different Hub-PEG-species. Due to this the overall purity for the membrane-based process is slightly lower (95.5 %) than for the chromatography-based process (98.2 %) (**Figure 39d**). With the membrane-based process an overall yield of 63 % was achieved, which is a significant improvement compared to the synthesis relying on chromatography (33 %).

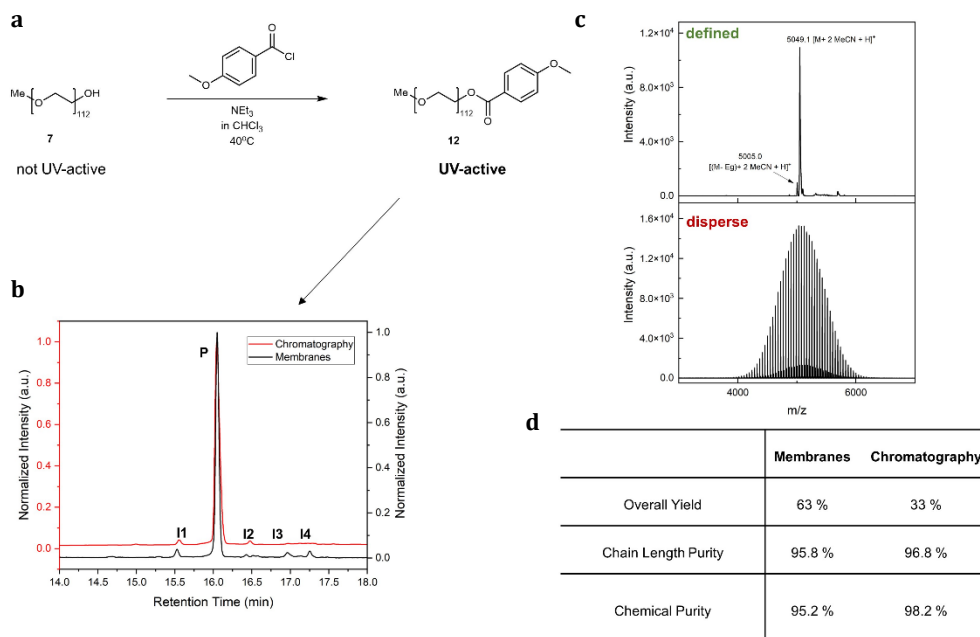


**Figure 39. Yields and purity data for the synthesis of defined mPEG-OH<sub>4966</sub>.** **a**, Functionalisation reaction to obtain UV-active derivative **5**. **b**, **d**, UV-chromatograms (254 nm, 15 cm C<sub>18</sub> column, 30 min 20-95 % B, A: H<sub>2</sub>O + 5 mM NH<sub>4</sub>OAc, B MeCN: MeOH 4:1) of 5 kDa MeO-Eg<sub>112</sub>-p-methoxybenzoate derived from membrane-based synthesis (black trace) and chromatography-based synthesis (red trace); Peak **P**: MeO-Eg<sub>112</sub>-p-methoxybenzoate (**5**), Peak **I1**: -Eg<sub>28</sub> impurity ( $\triangleq$  MeO-Eg<sub>84</sub>-p-methoxybenzoate), Peak **I2**: +Eg<sub>28</sub> impurity ( $\triangleq$  MeO-Eg<sub>140</sub>-p-methoxybenzoate), Peaks **I3+I4**: unidentified impurities; **c**, MALDI-TOF-MS of mPEG-OH<sub>4966</sub> synthesised via membrane-based approach in this work (bottom) and commercially available disperse mPEG-OH (top). **d**, summary of yields and purities for chromatography-based and membrane-based synthesis of MeO-Eg<sub>112</sub>-OH.

## 4.4 Conclusions

Within this chapter, different investigations, and optimisations on the Nanostar Sieving based synthesis of defined  $M_w$  Peg 5 kDa were conducted, starting from the preparation of multi-gram batches using an Eg<sub>12</sub> building block. This process needed a total of seven cycles of chain extension, deprotection and membrane-purification to go from the Hub(Eg<sub>28</sub>-OH)<sub>3</sub> (which was chromatographically purified beforehand) to the final Hub(Eg<sub>112</sub>-OH)<sub>3</sub>. The largest batch yielded 13 g of Hub(Eg<sub>112</sub>-OH)<sub>3</sub> and the MALDI-TOF spectrum of the cleaved MeO-Eg<sub>112</sub>-OH exhibited the characteristic discrete nature attributed to the iterative synthetic pathway.

With the final aim of using a longer  $\text{Eg}_{28}$  based BB and synthesising PEG of higher molecular weight, different investigations were conducted to optimise the product purity and process efficiency.

The reliability of the used UPLC-MS based techniques to monitor the reactions and purifications was improved through implementing cleaning and calibration procedures. It was shown that MALDI-TOF MS is not suitable to determine the chain length purity of defined  $M_w$  PEG and a solution was offered to determine the synthesis-related chain length purity of defined  $M_w$  PEG 5 kDa, by introducing a UV-active marker and analysing the different species through UPLC-MS.

An investigation on the reaction concentration identified a concentration of 20 wt% to deliver the fastest chain extension reactions when a  $\text{Eg}_{28}$  BB is used. Furthermore, it was found that at lower chain lengths lower amounts of BBs can be used without impeding the conversion progress of the reaction. This way, the overall efficiency of the synthesis was increased.

A critical investigation of side-reactions revealed that impurities contained in the starting materials (such as ethyl tosylate) can result in cross-reactions and undesired side-products. It was shown that defined  $M_w$  PEG allows for facile investigation of such side-reactions through UPLC-MS, which contributes to a high purity of the final material.

One unresolved issue during the Williamson etherification-based synthesis of PEG is the oxidation/peroxidation that was shown to occur upon treatment of free PEG-alcohol species with sodium hydride, resulting in linear and higher  $M_w$  branched oxidised PEG species. The successful detection and identification of those impurities lays the foundation for finding a solution to this problem.

In the final part of this chapter, a defined  $M_w$  PEG 5 kDa was synthesised using an  $\text{Eg}_{28}$  BB, which decreased the number of needed chain extension cycles from seven to three and that way significantly increased the efficiency of the synthesis process. A comparative study was conducted by performing one synthesis based on membrane-purification and another one using automated chromatography for the purification between each step. A total of 8.3 g of  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$  was

synthesised through the membrane-based synthesis with an overall yield of 63 %. While the derived chain length purities were similar for both approaches, the chromatography-based synthesis results in a overall yield of only 33 %.

These results offer the foundation for an efficient and scalable synthesis of defined  $M_w$  PEG. The molecular weight of 5 kDa makes these an interesting candidate for the application as defined  $M_w$  PEGylation agents. The functionalisation of defined  $M_w$  PEG 5 kDa to reach active PEGylation agents will be discussed in the next chapter (**Chapter 5**).

## Chapter 5 Functionalisation of defined $M_w$ PEGs as a third generation

### PEGylation agents

#### Abstract

Despite often being labelled as “monodisperse”, the PEGylation agents currently used in FDA-approved PEGylated therapeutics all exhibit a degree of dispersity in their molecular weight/chain length. In this chapter, a selection of linear, defined  $M_w$  5 kDa PEGylation agents was synthesised, offering a first-of-its kind solution for random and site-selective protein or peptide PEGylation to produce defined  $M_w$  protein or peptide therapeutics. It was demonstrated how the previously discussed Nanostar Sieving technique allows for the facile synthesis of monofunctional and heterobifunctional defined  $M_w$  PEGylation agents. Due to the defined mass of the PEG species, all chemical modification reactions could easily be followed using UPLC-MS. It was shown that the defined PEG species exhibit an improved chromatographic separability on reverse-phase UPLC, compared to the disperse derivatives. A range of defined  $M_w$  5 kDa mPEG active esters was synthesised as candidates for the random PEGylation at primary amine groups. A mPEG aldehyde was produced to offer a solution for the site-selective *N*-terminal PEGylation of proteins.

#### Keywords

PEGylation agents, defined, disperse, monodisperse, selective PEGylation, random PEGylation

#### Highlights

- seven different monofunctional defined  $M_w$  PEGylation agents were synthesised
- three defined  $M_w$  mPEG-active esters were synthesised, offering different reactivities towards primary amine groups of proteins
- a first defined  $M_w$  heterobifunctional PEGylation agent (5 kDa) was synthesised
- working with defined  $M_w$  PEGylation agents allows for facile purity analysis through UPLC-MS

## 5.1 Introduction

One of the potential applications for defined  $M_w$  PEG is as PEGylation agents to mask protein- or peptide-therapeutics, small molecules, or nano carriers for therapeutics. To achieve an active PEGylation agent, one end of the molecule must be modified with a functional group that is active towards the desired functional group within the therapeutic. The other end (or two ends in case of a two-armed, branched structure) must be capped with a chemically inert group to avoid cross reactions that could change the PEGylated therapeutic after modification. In the following, the PEG species bearing only one reactive functional group will be called mono-reactive. In all cases of FDA-approved PEGylated therapeutics, the inactive ends of the PEG are capped with a methoxy group (MeO-) and the resulting PEG structures are called methoxy-PEGs or mPEGs. The choice of the reactive functionality depends on the chemical nature of the species to be PEGylated and the desired outcome of the PEGylation reaction. One of the most popular structures for PEGylating proteins are amine-reactive active esters (or NHS esters), which are carboxylates that have been activated through carbodiimide-mediated reaction with *N*-hydroxysuccinimide. These species are highly reactive towards primary amine groups, which are usually the most abundant functional group on the surface of proteins and exhibit a nucleophilic character under physiological pH. Depending on the number and location of primary amine groups in the substrate, PEGylation reactions with NHS esters are usually random. Depending on their exact structure different mPEG-NHS esters exhibit different reactivities and hydrolysis half lives.

Another type of PEGylation agent frequently used in FDA-approved PEGylated drugs, is the mPEG aldehyde. This type of reagent can be used to selectively PEGylate the *N*-terminus of a protein.

Heterobifunctional PEG reagents are promising candidates for linker molecules in antibody-drug conjugates.

Within the following chapter, defined  $M_w$  5 kDa derivatives of the different groups of PEGylation agents were synthesised.

## 5.2 Materials and Methods

### 5.2.1 Materials

Boron trichloride solution (1.0 M in methylene chloride, Sigma-Aldrich), dichloroacetic acid (>98%, TCI), ethyl 3-bromopropionate (Alfa Aesar), iodomethane (copper as stabilizer, 99%, Sigma-Aldrich), *N*-hydroxysuccinimide (98%, Sigma-Aldrich), *N,N'*-dicyclohexylcarbodiimide (DCC, 99%, Sigma-Aldrich), sodium hydride (60 % in mineral oil, Sigma-Aldrich), triethylamine (for synthesis, Sigma-Aldrich), triethylammonium chloride (for synthesis, Sigma-Aldrich), sodium bicarbonate (technical, >98 %, VWR), potassium *tert*-butoxide (1 M in THF, TCI), ethyl bromoacetate (reagent grade, 98 %, Sigma-Aldrich), sodium hydroxide (NaOH, 98 %, Thermo Scientific), sodium sulfate (anhydrous, Fisher Chemical), 4,4'-dimethoxytrityl chloride (>97.0 %, TCI), *N*-methylimidazole (for synthesis, Sigma-Aldrich), *p*-toluene sulfonyl chloride (ReagentPlus®, ≥99%, Sigma-Aldrich), Dess-Martin periodinane (DMP, 97 %, Sigma-Aldrich), succinic anhydride (for synthesis, Sigma-Aldrich), magnesium sulfate (Fisher Chemical), sodium azide (ReagentPlus®, >99.5 %, Sigma-Aldrich), citric acid (VWR), triphenylphosphine (>95.0 %, TCI).

### 5.2.2 Methods

All reactions were performed under argon, unless otherwise stated. Solvents labelled as “dry” were dried over activated 3Å molecular sieve for 24 hours and tested for their water content using a Karl-Fischer coulometric titrator (<50 ppm).

The nanostar species is referred to as “Hub” in chemical formulas.

### 5.2.2.1 Monofunctional PEGylation agents

Methylation of Hub(Eg<sub>112</sub>-OH)<sub>3</sub> (**13**).

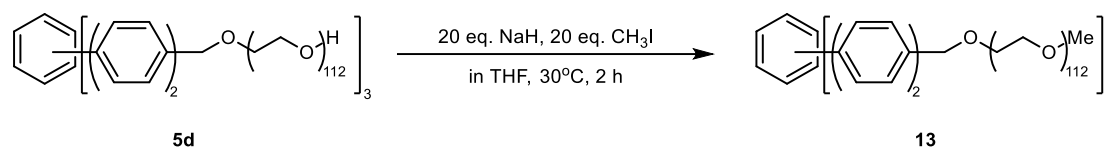


**Figure 40.** Reaction scheme of the methylation of Hub(Eg<sub>112</sub>-OH)<sub>3</sub> (**5d**) to reach Hub(Eg<sub>112</sub>-OMe)<sub>3</sub> (**13**).

The purified Hub(Eg<sub>112</sub>-OH)<sub>3</sub> (**5d**, 1.00 g, 0.065 mmol) was azeotroped from dry acetonitrile and then dissolved in dry THF (5 ml). Sodium hydride (60% dispersion in mineral oil, 0.05 g, 1.30 mmol, 20 eq.) was added and then immediately Iodomethane (0.08 ml, 1.30 mmol, 20.0 eq.). The flask was placed in an oil bath set to 30 °C and the reaction mixture was stirred for 2 hours. After confirming completion of the reaction through the UPLC-UV chromatogram), it was quenched with a solution of triethylammonium chloride (Et<sub>3</sub>NCl, 0.18 g, 1.30 mmol, 20 eq.), filtered, dissolved in MeOH and then transferred to the diafiltration rig for purification. After diafiltration purification, 0.96 g of pure Hub<sup>3</sup>(Eg<sub>112</sub>-OMe)<sub>3</sub> (**13**) was obtained (0.062 mmol, *y* = 95 %).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>).**  $\delta$  = 7.83 (s, 3H, Hub *CH*), 7.75 (d, *J* = 8.0 Hz, 6H, Hub *CH*), 7.67 (d, *J* = 8.1 Hz, 6H, Hub *CH*), 7.59 (d, *J* = 7.8 Hz, 6H, Hub *CH*), 7.40 (d, *J* = 7.8 Hz, 6H, Hub *CH*), 4.57 (s, 6H, 3 x OCH<sub>2</sub>Ar), 3.84 – 3.28 (m, 1344H, 672 x CH<sub>2</sub>O), 3.25 (s, 3H, OCH<sub>3</sub>).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>).**  $\delta$  = 141.92 (3C, 3 x Hub *C*), 140.13 (3C, 3 x Hub *C*), 139.93 (3C, 3 x Hub *C*), 137.62 (3Cf, 3 x Hub *C*), 137.54 (3C, 3 x Hub *C*), 131.03 (3C, 3 x Hub *C*), 128.25 (6C, 6 x Hub *CH*), 127.69 (6C, 6 x Hub *CH*), 127.50 (6C, 6 x Hub *CH*), 127.00 (6C, 6 x Hub *CH*), 72.92+72.5zz8 (3 x OCH<sub>2</sub>Ar), 71.00- 69.99 (m, 663 x CH<sub>2</sub>O), 70.25 (3 C, 3 x CH<sub>2</sub>O), 69.52(3 C, 3 x CH<sub>2</sub>O), 61.61 (3C,3 x CH<sub>2</sub>OH), 59.09 (C, 1 x CH<sub>2</sub>OCH<sub>3</sub>).

***m/z* (ESI<sup>+</sup>)** = 1952 [M+8 NH<sub>4</sub><sup>+</sup>]<sup>8+</sup>, 1736 [M+9 NH<sub>4</sub><sup>+</sup>]<sup>9+</sup>, 1565 [M+10 NH<sub>4</sub><sup>+</sup>]<sup>10+</sup>, 1425 [M+11 NH<sub>4</sub><sup>+</sup>]<sup>11+</sup>, 1307 [M+12 NH<sub>4</sub><sup>+</sup>]<sup>12+</sup>, 1208 [M+13 NH<sub>4</sub><sup>+</sup>]<sup>13+</sup>; *calc.* (C<sub>720</sub>H<sub>1386</sub>O<sub>339</sub>) = 1951 [M+8 NH<sub>4</sub><sup>+</sup>]<sup>8+</sup>, 1736

$[M+9 \text{ NH}_4^+]^{9+}$ , 1564  $[M+10 \text{ NH}_4^+]^{10+}$ , 1424  $[M+11 \text{ NH}_4^+]^{11+}$ , 1307  $[M+12 \text{ NH}_4^+]^{12+}$ , 1208  $[M+13 \text{ NH}_4^+]^{13+}$

*Cleavage to yield MeO-Eg<sub>112</sub>-OH (mPEG<sub>4967</sub> alcohol, 7).*



**Figure 41. Reaction scheme of the cleavage reaction to yield linear MeO-Eg<sub>112</sub>-OH (mPEG<sub>4967</sub> alcohol, 7).**

For cleaving MeO-Eg<sub>112</sub>-OH from the nanostar, Hub(Eg<sub>112</sub>-OMe)<sub>3</sub> (**13**, 0.50 g, 0.03 mmol) was dissolved in DCM. The flask was placed in a cooling bath consisting of acetone and dry ice to reach a temperature of -70 °C. Once cooled down, boron trichloride (1 M solution in DCM) was added to the stirred reaction solution and it was allowed to warm up to -30 °C. The reaction was maintained at this temperature until it reached completion (confirmation through UPLC-MS) and then cooled back to -70 °C to then slowly add MeOH and sodium bicarbonate to quench the reaction. The solution was concentrated under reduced pressure and then precipitated in ice cold ether. The precipitate was filtered off, washed with ice-cold ether and then dried under high vacuum, Pure MeO-Eg<sub>112</sub>-OH (**7**, 0.38 g, y = 85%) was obtained as a colourless, crystalline solid.

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>).** δ = 3.83- 3.42 (m, 448, 224 x CH<sub>2</sub>O).), 3.37 (s, 3H, OCH<sub>3</sub>).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>).** δ = 72.66 (C, 1 x CH<sub>2</sub>OCH<sub>3</sub>), 71.96 (C, 1 x CH<sub>2</sub>CH<sub>2</sub>OH), 70.59 (220 C, 220 x CH<sub>2</sub>O), 61.71 (C, 1 x CH<sub>2</sub>OH), 59.09 (C, 1 x CH<sub>2</sub>OCH<sub>3</sub>).

***m/z* (ESI+)** = 1673 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1260 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1011 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 846 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.*calc.*

(C<sub>225</sub>H<sub>452</sub>O<sub>113</sub>) = 1673 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1259 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1011 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 846 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

***m/z* (MALDI-ToF+).** 5049.1 ([M + 2 MeCN + H]<sup>+</sup>).

Purity (UPLC-ELSD)= 99 %

Synthesis of MeO-Eg<sub>112</sub> carboxylic acid (mPEG<sub>5024</sub>-CA, **14**).

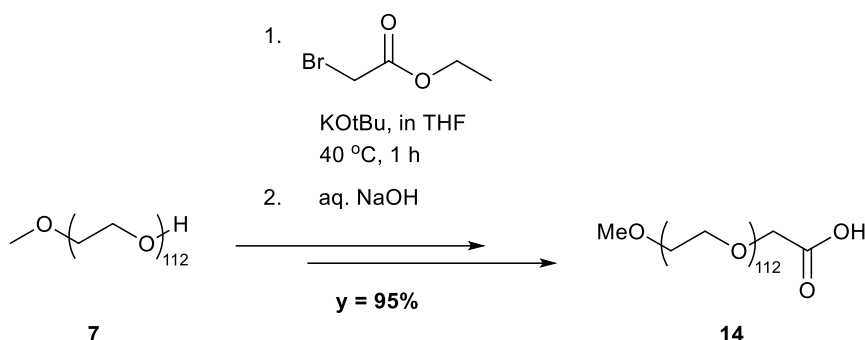


**Figure 42.** Reaction scheme of the synthesis of MeO-Eg<sub>112</sub> carboxylic acid (mPEG<sub>5024</sub>-CA, **14**).

To reach the mPEG carboxylic acid **14**, first the ethyl ester was synthesised. For this the mPEG-OH (**7**, 200 mg, 0.04 mmol) was azeotroped from dry acetonitrile (3 x), then from dry THF (1 x) and then dissolved in dry THF (4.00 ml). To the stirred solution potassium tert-butoxide solution was added (0.25 ml, 1M in THF, 0.25 mmol, 6.25 eq.) and then immediately ethyl bromoacetate was added (0.03 ml, 0.25 mmol, 6.25 eq.). The reaction was stirred at 40 °C for 1 hour and completion of the reaction was confirmed through UPLC-MS (ELSD chromatogram, 20-95 % MeCN:MeOH 4:1, 10 min gradient). The solution was allowed to cool down to room temperature and water (2 ml) and NaOH (10 mg, 0.25 mmol, 6.25 eq.) were added to hydrolyse the ester to the according carboxylic acid (**14**). After completion the THF was removed under reduced pressure and the aqueous solution was extracted with DCM twice. The pH of the aqueous phase was adjusted to 4 using 0.1 M HCL and then it was extracted with DCM three times. This organic layer was dried over sodium sulfate, filtered and then the solvent was removed under reduced pressure to yield MeO-Eg<sub>112</sub>-COOH (**14**, 190 mg,  $y = 95\%$ ).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>).**  $\delta = 3.77 - 3.73$  (m, 2H, 2 x CH<sub>2</sub>O), 3.60-3.56 (m, 448H, 224 x CH<sub>2</sub>O)), 3.50 – 3.47 (, 2H, 2 x CH<sub>2</sub>O), 3.43 – 3.38 (m, 2H, 2 x CH<sub>2</sub>O), 3.32 (s, 3H, OCH<sub>3</sub>).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>).**  $\delta = 72.65$  (C, 1 x CH<sub>2</sub>OCH<sub>3</sub>), 71.93 (C, 1 x CH<sub>2</sub>CH<sub>2</sub>OH), 70.56 (220 C, 220 x CH<sub>2</sub>O), 61.70 (C, 1 x CH<sub>2</sub>OH), 59.07 (C, 1 x CH<sub>2</sub>OCH<sub>3</sub>).

$M_w$  (Product) = 5024.01 Da

$m/z$  (ESI+) = 1693 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1774 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1023 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 866 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.*calc.*

(C<sub>227</sub>H<sub>454</sub>O<sub>115</sub>) = 1693 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1774 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1023 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 855 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

Purity (UPLC-ELSD) = 99 %

*Synthesis of MeO-Eg<sub>112</sub>-succinimidyl carboxymethyl ester (mPEG<sub>5121</sub>-SCM, **15**).*

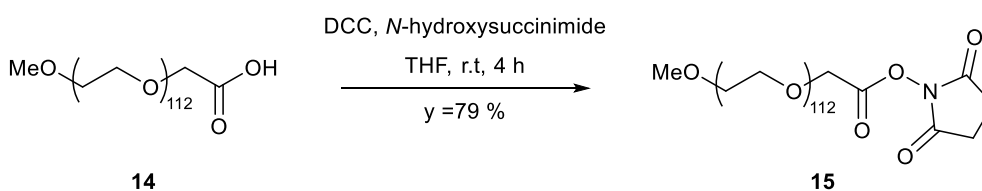


**Figure 43.** Reaction scheme of the synthesis of MeO-Eg<sub>112</sub>-succinimidyl carboxymethyl ester (mPEG<sub>5121</sub>-SCM, **15**).

A stock solution of *N,N'*-dicyclohexylcarbodiimide (DCC) in dry THF was prepared (1 M). The dry, isolated mPEG carboxylic acid (**14**, 190 mg, 0.038 mmol) was dissolved in dry THF and 14.8 mg of *N*-hydroxysuccinimide were added (0.13 mmol, 4 eq.). To the stirred solution 0.13 ml (0.13 mmol, 4 eq.) of the DCC solution were added and the reaction was stirred at room temperature for 4 hours. After completion, any solids were filtered off and the filtrate was precipitated in ice cold diethylether to yield the active ester **15** (154 mg, 0.030 mmol, *y* = 79 %). The highly reactive mPEG-NHS ester was stored under argon at -4 °C.

$M_w$  (Product) = 5121.09 Da

$m/z$  (ESI+) = 1725 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1298 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1042 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 871 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.*calc.*

(C<sub>231</sub>H<sub>457</sub>O<sub>117</sub>) = 1725 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1298 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1042 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 871 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

Purity (UPLC-ELSD) = 90 % (8 % carboxylic acid, **14**)

*Synthesis of DmtrO-(CH<sub>2</sub>)<sub>3</sub>-OH (26).*

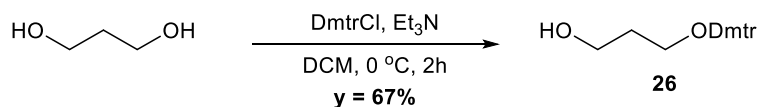


**Figure 44.** Reaction scheme for the synthesis of DmtrO-(CH<sub>2</sub>)<sub>3</sub>-OH (26).

1,3-propanediol (2.00 g, 26.3 mmol, 10.0 eq.) was azeotroped from dry MeCN (3 x) and then dissolved in dry DCM (50 ml). 4,4'-dimethoxytrityl chloride (0.9 g, 2.6 mmol) was added and then triethylamine (0.7 ml, 5.3 mmol, 2.0 eq.) was added dropwise. The reaction was cooled in an ice bath and stirred for 2 hours. The crude reaction mixture was purified through flash chromatography (normal phase, Ø 6.5 cm, 9 cm height, mobile phase: CHCl<sub>3</sub> + 0.1 % Et<sub>3</sub>N) and the product was obtained as a colourless solid (**26**, 0.7 g, y= 67 %).

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>).** δ = 7.46 – 7.37 (m, 2H, 2 x CH<sub>Ar</sub>), 7.35 – 7.15 (m, 8H, 8 x CH<sub>Ar</sub>), 6.87 – 6.77 (m, 4H, 4 x CH<sub>Ar</sub>), 3.78 (dd, *J* = 4.7, 2.5 Hz, 8H, 2 x OCH<sub>3</sub> + 1x -CH<sub>2</sub>-), 3.26 (tt, *J* = 5.9, 1.9 Hz, 2H, 1x -CH<sub>2</sub>-), 1.84 (tdd, *J* = 10.3, 4.7, 2.8 Hz, 2H, 1x -CH<sub>2</sub>-).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>).** δ = 158.55 (2C, 2 x C<sub>Ar</sub>), 136.24 (2C, 2 x C<sub>Ar</sub>), 130.06 (4C, 4 x C<sub>Ar</sub>), 128.17 (3C, 3 x C<sub>Ar</sub>), 127.94 (3C, 3 x C<sub>Ar</sub>), 126.85 (2C, 2 x C<sub>Ar</sub>), 113.22 (4C, 4 x C<sub>Ar</sub>), 86.56 (1C, 1 x C), 62.51 (1C, 1 x CH<sub>2</sub>), 62.05 (1C, 1 x CH<sub>2</sub>), 55.29 (2C, 2 x CH<sub>3</sub>), 32.52 (1C, 1 x CH<sub>2</sub>).

*Synthesis of DmtrO-(CH<sub>2</sub>)<sub>3</sub>-OTs (27).*

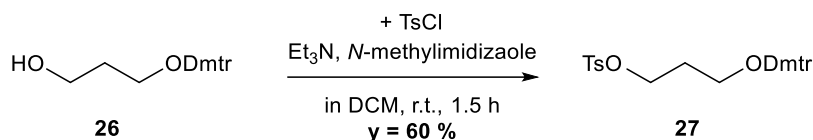


**Figure 45.** Reaction scheme of the synthesis of DmtrO-(CH<sub>2</sub>)<sub>3</sub>-OTs (27).

DmtrO-(CH<sub>2</sub>)<sub>3</sub>-OH (0.4 g, 1.3 mmol) was evaporated from anhydrous acetonitrile twice. The residue was dissolved in anhydrous DCM to which was added triethylamine (0.54 ml, 3.9 mmol, 3.0 eq.) and *N*-methylimidazole (NMI, 0.1 ml, 1.3 mmol, 1.0 eq.). The reaction was cooled in a

water bath and p-toluene sulfonyl chloride (Tos-Cl, 4.5 g, 23.6 mmol, 1.5 eq.) was added. After 5 minutes the water bath was removed, and the reaction was stirred at room temperature for 1.5 h. To this crude mixture was added THF (15 ml) and DCM was evaporated off (at room temperature) and to the residue was more THF (3x5 ml). The salt ( $\text{Et}_3\text{N}\cdot\text{HCl}$ ) was removed by filtration, washing the solids with further THF. To the filtrate was added  $\text{NaHCO}_3$ , then water (7 ml) and the solution was stirred for 3 hours to hydrolyse the remaining Tos-Cl.

The crude reaction mixture was purified through flash chromatography (normal phase,  $\varnothing$  6.5 cm, 9 cm height, mobile phase:  $\text{CHCl}_3$  + 0.1 %  $\text{Et}_3\text{N}$ ) and the product **27** was obtained as a colourless solid (0.38 g,  $y=60\%$ ).

**$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ).**  $\delta = 7.73$  (d,  $J = 8.4$  Hz, 2H, 2 x  $\text{CH}_{\text{Ar}}$ ), 7.35 – 7.26 (m, 4H, 4 x  $\text{CH}_{\text{Ar}}$ ), 7.22 – 7.14 (m, 6H, 6 x  $\text{CH}_{\text{Ar}}$ ), 6.82 – 6.74 (m, 4H, 4 x  $\text{CH}_{\text{Ar}}$ ), 4.18 (t,  $J = 6.4$  Hz, 2H, 1 x  $\text{CH}_2$ ), 3.77 (s, 6H, 2 x  $\text{CH}_3$ ), 3.09 (t,  $J = 6.0$  Hz, 2H, 1 x  $\text{CH}_2$ ), 2.41 (s, 3H, 1 x  $\text{CH}_3$ ), 1.86 (t,  $J = 6.2$  Hz, 2H, 1 x  $\text{CH}_2$ ).

*Coupling of DmtrO-(CH<sub>2</sub>)<sub>3</sub>-OTs to MeO-Eg<sub>112</sub>-OH and Deprotection.*

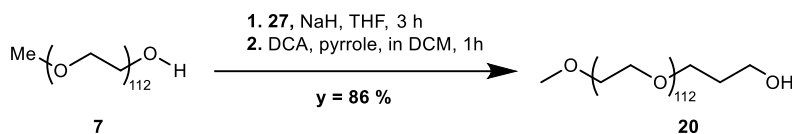


**Figure 46. Reaction scheme of synthesis of MeO-Eg<sub>112</sub>-(CH<sub>2</sub>)<sub>3</sub>-OH (**20**).**

MeO-Eg<sub>112</sub>-OH (0.45 g, 0.09 mmol) and DmtrO-(CH<sub>2</sub>)<sub>3</sub>-OTs (0.14 g, 0.27 mmol, 3.0 eq.) were azeotroped from dry MeCN (3 x 30 ml) and then dissolved in dry THF (5 ml). Sodium hydride (60 % disp., 0.036 g, 0.9 mmol, 10.0 eq.) were added and the reaction was stirred at room temperature for 3 hours. Upon completion (monitored through UPLC-MS) the reaction was quenched with a mixture of H<sub>2</sub>O/ MeCN and Et<sub>3</sub>NCl (triethylammonium chloride, 0.12 g, 0.9 mmol, 10.0 eq.). The THF was removed under reduced pressure and the residue redissolved in DCM (5 ml). Pyrrole was added (0.06 ml, 0.9 mmol, 10.0 eq.) and then dichloroacetic acid (DCA, 0.07 ml, 0.9 mmol, 10.0 eq.). The reaction was stirred at room temperature for 1 hour (completion was confirmed through UPLC-MS) and then neutralised with triethylamine until pH ≈ 7. Any solids were filtered off and the residue was concentrated under vacuum. The concentrate was precipitated in ice-cold ether and the precipitate washed with more ice-cold ether. The precipitate was then washed into a flask with MeCN and subsequently dried under vacuum. The product (**20**) was obtained as a colourless waxy solid (0.39 g, 86 %).

M<sub>w</sub> (Product) = 5024.06 Da

**m/z (ESI<sup>+</sup>)** = 1692 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1274 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1023 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 855 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>. *calc.*

(C<sub>228</sub>H<sub>458</sub>O<sub>114</sub>) = 1692 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1274 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1023 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 855 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

Purity (UPLC-ELSD) = 96 %

Reaction of MeO-EG<sub>112</sub>-(CH<sub>2</sub>)<sub>3</sub>-OH to MeO-Eg<sub>112</sub>-propionaldehyde (mPEG<sub>5022</sub>-PA, **21**).

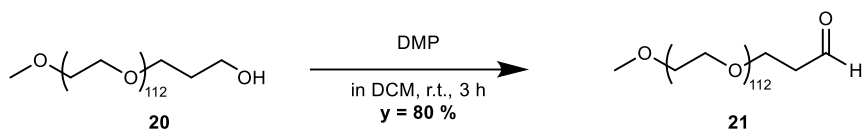


**Figure 47. Reaction scheme of the reaction of MeO-Eg<sub>112</sub>-(CH<sub>2</sub>)<sub>2</sub>-OH (**20**) to aldehyde (**21**).**

The starting material **20** (0.19 g, 0.04 mmol) was azeotroped from dry MeCN and then redissolved in dry DCM (6 ml). Dess-Martin periodinane (0.07 g, 0.17 mmol, 4.5 eq.) were added to the solution and the reaction was stirred at room temperature for 3 hours. Upon completion of the reaction (confirmation through UPLC-MS) first sat. NaHCO<sub>3</sub> (1 ml) and sat. Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (10 ml) were added and then H<sub>2</sub>O (20 ml) and DCM (20 ml). The organic layer was separated and the aqueous was washed with further DCM (3 x). The combined organic layers were concentrated under reduced pressure and the concentrated solution was precipitated in ice-cold ether and washed with more ether. The filter cake was redissolved in MeCN, transferred into a flask and the solvent was striped off under reduced pressure. After drying under high vacuum, the product **21** was obtained as a colourless, waxy solid (0.14 g, 73 %)

M<sub>w</sub>(Product) = 5022.04 Da

**m/z (ESI+)** = 1692 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1273 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1023 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 855 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>. *calc.*

(C<sub>228</sub>H<sub>456</sub>O<sub>114</sub>) = 1692 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1274 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1022 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 855 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

Purity (UPLC-ELSD)= 97 %

*Synthesis of MeO-Eg<sub>112</sub>-succinimidyl succinate (mPEG<sub>5163</sub>-SS, **17**).*

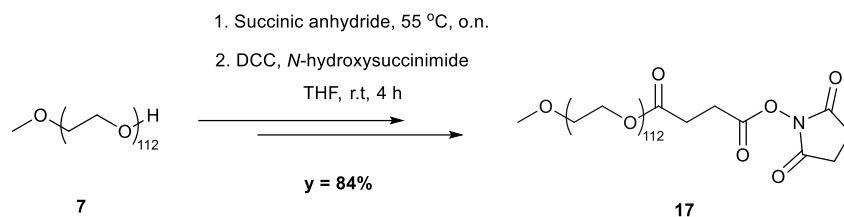


**Figure 48.** Reaction scheme of the reaction of MeO-Eg<sub>112</sub>-OH (**7**) to MeO-Eg<sub>112</sub>-succinimidyl succinate (mPEG<sub>5163</sub>-SS, **17**).

The starting material (**7**, MeO-Eg<sub>112</sub>-OH, 0.50 g, 0.1 mmol) was azeotroped from dry MeCN and subsequently dissolved in chloroform (15 ml). Succinic anhydride (0.05 g, 0.5 mmol, 5.0 eq.) and triethylamine (0.08 ml, 0.5 mmol, 5.0 eq.) were added and the reaction was stirred at room temperature overnight (completion of the reaction was confirmed through UPLC-MS). The chloroform was stripped off and the residue was redissolved in 10wt% NaHCO<sub>3</sub>. The solution was filtered, subsequently cooled down in an ice bath and acidified with 0.1 M HCl. The aqueous phase was extracted with chloroform (4 x). The combined organic layers were dried over MgSO<sub>4</sub>, filtered and the solution was concentrated under reduced pressure. The concentrated product was precipitated in ice-cold ether, filtered, and washed with more ether. The filter cake was redissolved in acetonitrile, transferred into a flask and the solvent was removed under reduced pressure. The resulting succinic acid was dissolved in dry DCM (6 ml) and *N*-hydroxysuccinimide (0.045 g, 0.4 mmol, 4.0 eq.) was added under stirring. DCC (0.083 g, 0.4 mmol, 4.0 eq.) was dissolved in dry THF and added to the solution dropwise and the solution was stirred under argon for four hours. After the reaction was complete the solvent was stripped off on the rotary evaporator. The crude product was re-dissolved in acetonitrile and filtered through a paper filter, the filtrate was collected, and the solvent removed on the rotary evaporator. This process was repeated twice. Subsequently the remaining solid was dissolved in a small amount of DCM and precipitated in ice-cold diethyl ether. The precipitate was filtered off, washed with more ice-cold diethyl ether and then dried under reduced pressure. Compound **17** was obtained as a colourless, waxy solid (0.43 g, *y* = 84 %).

$M_w = 5163.12 \text{ Da}$

$m/z$  (ESI+) = 1739 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1309 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1051 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 879 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.*calc.*

(C<sub>233</sub>H<sub>459</sub>O<sub>118</sub>) = 1739 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1309 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1051 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 879 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

Purity (UPLC-ELSD)= 96 %

*Synthesis of MeO-Eg<sub>112</sub>-succinimidyl propionate (mPEG<sub>5135</sub>-SP, **19**).*

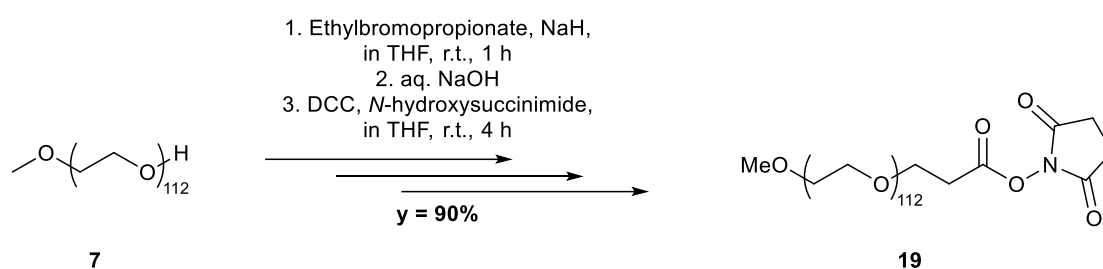


**Figure 49.** Reaction scheme of the reaction of MeO-Eg<sub>112</sub>-OH (**7**) to MeO-Eg<sub>112</sub>-succinimidyl propionate (mPEG<sub>5135</sub>-SP, **19**).

MeO-Eg<sub>112</sub>-OH (**7**, 0.30 g, 0.060 mmol) was azeotroped from dry acetonitrile and then dissolved in dry THF (6 ml). The solution was cooled down to 0 °C and first sodium hydride (0.005 g, 0.125 mmol, 2.1 eq.) was added and then ethyl 3-bromopropionate (75 µL, 0.60 mmol, 10.0 eq.). The reaction was allowed to warm up from 0 °C to room temperature under stirring and then heated to 40 °C. The reaction was complete after 1 h (monitored through UPLC-MS) and quenched sat. NH<sub>4</sub>Cl. Aqueous NaOH was added to hydrolyse the ester. The THF was evaporated off and the pH of the remaining aqueous adjusted to pH ≈ 4. The aqueous was saturated with NaCl and extracted with dichloromethane (3 x). The combined organic layers were dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and the solvent was stripped off under reduced pressure. Subsequently the propionic acid (**18**) (0.20 g, 0.04 mmol) was dissolved in dry DCM (3 ml) and N-hydroxysuccinimide (0.018 g, 0.16 mmol, 4.0 eq.) was added under stirring. DCC (0.033 mg, 0.16 mmol, 4.0 mmol) was dissolved in dry THF and added to the solution dropwise and the solution was stirred under argon for four hours. After the reaction was complete the solvent was

stripped off on the rotary evaporator. The crude product was re-dissolved in acetonitrile and filtered through a paper filter, the filtrate was collected, and the solvent removed on the rotary evaporator. This process was repeated twice. Subsequently the remaining solid was dissolved in a small amount of DCM and precipitated in ice-cold diethyl ether. The precipitate was filtered off, washed with more ice-cold diethyl ether and then dried under reduced pressure. Compound **19** was obtained as a colourless, waxy solid (0.16 g, 0.03 mmol,  $y = 75\%$ ).

$M_w = 5135.11$  Da

$m/z$  (ESI+) = 1302 ( $[M+4\text{ NH}_4]^{4+}$ ), 1045 ( $[M+5\text{ NH}_4]^{5+}$ ), 874 ( $[M+6\text{ NH}_4]^{6+}$ ), 752 ( $[M+7\text{ NH}_4]^{7+}$ ), 660 ( $[M+8\text{ NH}_4]^{8+}$ ); *calc.* ( $\text{C}_{225}\text{H}_{452}\text{O}_{113}$ ): 1302 ( $[M+4\text{ NH}_4]^{4+}$ ), 1045 ( $[M+5\text{ NH}_4]^{5+}$ ), 874 ( $[M+6\text{ NH}_4]^{6+}$ ), 752 ( $[M+7\text{ NH}_4]^{7+}$ ), 660 ( $[M+8\text{ NH}_4]^{8+}$ ).

Purity (UPLC-ELSD) = 92% (5 % hydrolysed derivative)

*Synthesis of MeO-Eg<sub>112</sub>-OTs (mPEG<sub>5120</sub>-t, **22**)*

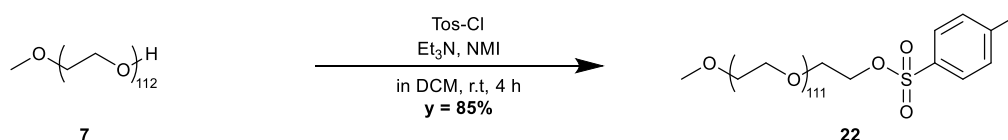


**Figure 50. Reaction scheme for the reaction of MeO-Eg<sub>112</sub>-OH (**7**) to MeO-Eg<sub>112</sub>-OTs (mPEG<sub>5120</sub>-t, **22**).**

MeO-Eg<sub>112</sub>-OH (**7**, 0.50 g, 0.1 mmol) was azeotroped from dry acetonitrile (3 x) and then dissolved in dry DCM (3 ml). Triethylamine (0.08 ml, 0.6 mmol, 6 eq.) and *N*-methylimidazole (8  $\mu\text{l}$ , 0.1 mmol, 1 eq) were added to the solution and then *p*-toluenesulfonyl chloride (0.11 g, 0.6 mmol, 6 eq.) were added. The reaction was stirred at room temperature for 4 hours and completion of the reaction was confirmed through UPLC-MS (ELSD chromatogram, 20-95 % MeCN:MeOH 4:1/ $\text{H}_2\text{O}$ , 6 min gradient). The DCM was removed under reduced pressure and the residue was taken up in THF (5 ml) and solids ( $\text{Et}_3\text{N}\cdot\text{HCl}$ ) were filtered off.  $\text{NaHCO}_3$  (0.10 g, 1.2 mmol, 12 eq.) and  $\text{H}_2\text{O}$  (3 ml) was added and the solution was stirred for 2 hours to hydrolyse all remaining Tosyle

$$M_w = 5120.16 \text{ Da}$$

Purity (UPLC-UV) = 98 %

The synthesised MeO-Eg<sub>112</sub>-OTs (**22**, 0.45 g, 0.08 mmol) was dissolved in DMF (2 ml) and sodium azide (0.08 g, 1.2 mmol, 15 eq.) was added. The reaction was stirred at 120 °C over night (under Argon atmosphere). Water (5 ml) and NaHCO<sub>2</sub> were added and the solution was saturated with NaCl and then extracted with chloroform (4 x). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and then stripped of under reduced pressure. The product (**23**) was obtained as a pale yellow solid (326 mg, y = 82 %).

$$M_w = 4974.99 \text{ Da}$$

Purity (UPLC-ELSD) = 98 %

Reduction to yield MeO-Eg<sub>112</sub>-NH<sub>2</sub> (mPEG<sub>4965</sub> amine, **24**)

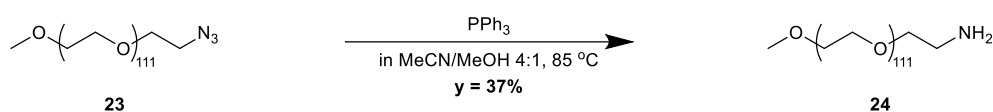


**Figure 52. Reaction scheme for the reaction of MeO-Eg<sub>112</sub>-N<sub>3</sub> (**23**) to MeO-Eg<sub>112</sub>-NH<sub>2</sub> (mPEG<sub>4065</sub>-amine, **24**).**

MeO-Eg<sub>112</sub>-N<sub>3</sub> (1.05 g, 0.21 mmol) was dissolved in MeCN/H<sub>2</sub>O (9:1, 180 mL) and triphenylphosphine (0.55 g, 2.1 mmol, 10 eq.) was added. The reaction was heated to 85 °C and stirred under reflux overnight. After completion the MeCN was removed under reduced pressure. The remaining aqueous solution was diluted with water to 50 ml and then the pH was adjusted to 4 using citric acid. A small amount of brine was added, and the aqueous layer was washed with chloroform (3 x). The chloroform was back-extracted with citric acid buffer solution. The pH of the combined aqueous layers was adjusted to >7 with K<sub>2</sub>CO<sub>3</sub>. The solution was saturated with NaCl and then extracted with chloroform (4 x). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and evaporated off. After final chromatographic purification the product (**24**) was obtained as a colourless solid (383 mg, y = 37 %).

M<sub>w</sub> = 4964.99 Da

**m/z (ESI<sup>+</sup>)** = 1673 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1259 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1011 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 845 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>. *calc.* (C<sub>233</sub>H<sub>459</sub>O<sub>118</sub>) = 1673 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1259 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1011 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 845 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

**<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>).** δ = 3.84-3.41 (m, 444 H, 222 x CH<sub>2</sub>), 3.36 (s, 3H, 2 x CH<sub>3</sub>), 3.01 (t, J = 6.2 Hz, 2 H, 1 x CH<sub>2</sub>), 2.21 (t, J = 6.2 Hz, 2 H, 1 x CH<sub>2</sub>).

**<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>).** δ = 71.90 (222C, 222 x CH<sub>2</sub>), 51.15 (3C, 1 x CH<sub>3</sub>), 29.66 (1C, 1 x CH<sub>2</sub>), 22.78 (1C, 1 x CH<sub>2</sub>).

### 5.2.2.2 Heterobifunctional PEGylation agents

#### Tosylation of Hub(Eg<sub>112</sub>-OH)<sub>3</sub> (**5h**).

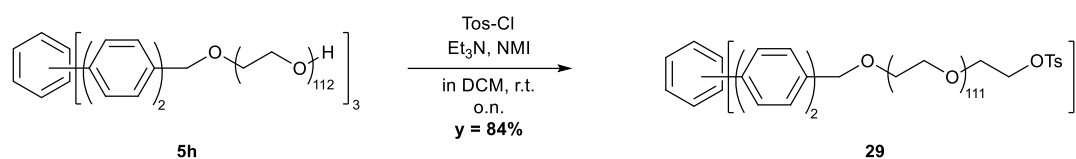


**Figure 53.** Reaction scheme for the tosylation of Hub(Eg<sub>112</sub>-OH)<sub>3</sub> (**5h**) to yield Hub(Eg<sub>112</sub>-OTs)<sub>3</sub> (**29**).

The starting material (Hub(Eg<sub>112</sub>-OH)<sub>3</sub>, 0.5 g, 0.03 mmol) was azeotroped from dry acetonitrile (3 x) and then dissolved in dry DCM (35 ml). Triethylamine (0.08 ml, 0.6 mmol, 20.0 eq.) and *N*-methylimidazole (0.02 ml, 0.3 mmol, 10 eq.) and then *p*-toluenesulfonyl chloride (Tos-Cl, 0.05 g, 0.3 mmol, 10 eq.) were added and the reaction was stirred at room temperature overnight. The next day, the DCM was removed under reduced pressure (at room temperature) and the residue taken up in THF (5 ml). The solution was filtered to remove any Et<sub>3</sub>N·HCl salts and to the filtrate NaHCO<sub>3</sub> (0.05 g, 0.6 mmol, 20 eq.) and H<sub>2</sub>O (5 ml) was added and the solution was stirred for 2 hours to hydrolyse any remaining Tos-Cl. The THF was removed under reduced pressure, the remaining aqueous saturated with NaCl and then extracted with chloroform (4 x). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and the solvent was removed under high vacuum. The product **29** was obtained as a pale yellow solid (0.43 g, y = 84%).

M<sub>w</sub> = 15905.18

***m/z* (ESI+)** = 1785 [M+8 NH<sub>4</sub><sup>+</sup>]<sup>8+</sup>, 1608 [M+9 NH<sub>4</sub><sup>+</sup>]<sup>9+</sup>, 1463 [M+10 NH<sub>4</sub><sup>+</sup>]<sup>10+</sup>, 1343 [M+11 NH<sub>4</sub><sup>+</sup>]<sup>11+</sup>, 1241 [M+12 NH<sub>4</sub><sup>+</sup>]<sup>12+</sup>, 1154 [M+13 NH<sub>4</sub><sup>+</sup>]<sup>13+</sup>; *calc.* (C<sub>720</sub>H<sub>1386</sub>O<sub>339</sub>) = 1785 [M+8 NH<sub>4</sub><sup>+</sup>]<sup>8+</sup>, 1608 [M+9 NH<sub>4</sub><sup>+</sup>]<sup>9+</sup>, 1463 [M+10 NH<sub>4</sub><sup>+</sup>]<sup>10+</sup>, 1343 [M+11 NH<sub>4</sub><sup>+</sup>]<sup>11+</sup>, 1241 [M+12 NH<sub>4</sub><sup>+</sup>]<sup>12+</sup>, 1154 [M+13 NH<sub>4</sub><sup>+</sup>]<sup>13+</sup>.

Purity (UPLC-UV) = 96 %

Conversion of  $\text{Hub}(\text{Eg}_{112}\text{-OTs})_3$  to  $\text{Hub}(\text{Eg}_{112}\text{-N}_3)_3$ .

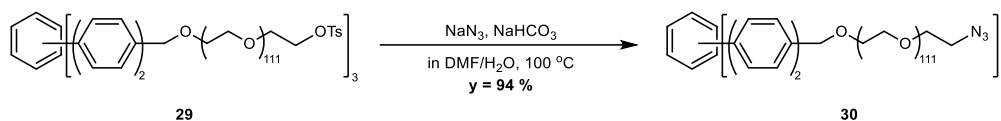


**Figure 54.** Reaction scheme for the conversion of  $\text{Hub}(\text{Eg}_{112}\text{-OTs})_3$  (**29**) to  $\text{Hub}(\text{Eg}_{112}\text{-N}_3)_3$  (**30**).

The starting material (**29**, 0.38 g, 0.024 mmol) was dissolved in DMF (2.1 ml) and water (0.4 ml) and  $\text{NaHCO}_3$  (0.049 g, 0.58 mmol, 24 eq.) was added to the solution. Sodium azide (0.070 g, 1.08 mmol, 45 eq.) was added and the reaction was stirred at  $70\text{ }^\circ\text{C}$  for 24 hours. The completion of the reaction was confirmed through UPLC-MS and the reaction was worked up by first removing the solvent under high vacuum (at room temperature) and then redissolving the residue in chloroform. The organic layer was washed with saturated sodium bicarbonate (3 x) and then dried over  $\text{Na}_2\text{SO}_4$ , filtered off and the solvent was removed under reduced pressure. The product **30** was obtained as a yellow solid (0.35 g,  $y = 94\%$ ).

$M_w = 15517.67$

$m/z$  (ESI+) = 1742  $[\text{M}+8\text{ NH}_4^+]^{8+}$ , 1570  $[\text{M}+9\text{ NH}_4^+]^{9+}$ , 1429  $[\text{M}+10\text{ NH}_4^+]^{10+}$ , 1311  $[\text{M}+11\text{ NH}_4^+]^{11+}$ , 1212  $[\text{M}+12\text{ NH}_4^+]^{12+}$ , 1127  $[\text{M}+13\text{ NH}_4^+]^{13+}$ ; *calc.* ( $\text{C}_{720}\text{H}_{1386}\text{O}_{339}$ ) = 1742  $[\text{M}+8\text{ NH}_4^+]^{8+}$ , 1570  $[\text{M}+9\text{ NH}_4^+]^{9+}$ , 1429  $[\text{M}+10\text{ NH}_4^+]^{10+}$ , 1311  $[\text{M}+11\text{ NH}_4^+]^{11+}$ , 1212  $[\text{M}+12\text{ NH}_4^+]^{12+}$ , 1127  $[\text{M}+13\text{ NH}_4^+]^{13+}$ .

Purity (UPLC-UV) = 97 %

Cleavage of HO-Eg<sub>112</sub>-N<sub>3</sub> (HO-PEG<sub>4977</sub>-N<sub>3</sub>) from Hub.

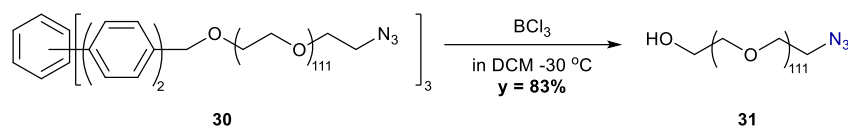


**Figure 55. Reaction scheme for the cleavage of HO-Eg<sub>112</sub>-N<sub>3</sub> (**31**) from Hub(Eg<sub>112</sub>-N<sub>3</sub>)<sub>3</sub> (**30**).**

For cleaving HO-Eg<sub>112</sub>-N<sub>3</sub> from the nanostar, Hub(Eg<sub>112</sub>-N<sub>3</sub>)<sub>3</sub> (**30**, 0.33 g, 0.021 mmol) was dissolved in DCM (15 ml). The flask was placed in a cooling bath consisting of acetone and dry ice to reach a temperature of -70 °C. Once cooled down, boron trichloride (1 M solution in DCM, 0.75 ml, 0.08 mmol, 4 eq.) was added to the stirred reaction solution and it was allowed to warm up to -30 °C. The reaction was maintained at this temperature until it reached completion (confirmation through UPLC-MS) and then cooled back to -70 °C to then slowly add MeOH. Sodium bicarbonate (0.04 g, 0.4 mmol, 20 eq.) was added to quench the reaction. The solution was concentrated under reduced pressure and then precipitated in ice cold ether. The precipitate was filtered off, washed with ice-cold ether and then dried under high vacuum. Pure MeO-Eg<sub>112</sub>-N<sub>3</sub> (**31**, 0.087 g, y = 83%) was obtained as a colourless, crystalline solid.

M<sub>w</sub> = 4976.97

**m/z (ESI+)** = 1677 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1262 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1013 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 847 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>. *calc.*

(C<sub>233</sub>H<sub>459</sub>O<sub>118</sub>) = 1677 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1262 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1013 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 847 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

Purity (UPLC-UV)= 98 %

Reaction of HO-Eg<sub>12</sub>-N<sub>3</sub> with *tert*-butyl bromoacetate.

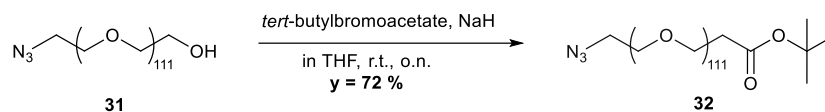


**Figure 56. Reaction scheme of the reaction of N<sub>3</sub>-Eg<sub>12</sub>-OH (31) with *tert*-butyl bromoacetate to reach product (32).**

The starting material **31** (0.01 g, 0.02 mmol) was dried by azeotropeing from dry acetonitrile (3 x) and then dissolved in dry THF (2.0 ml). Sodium hydride was added (10 eq., 0.2 mmol, 8 mg) and the reaction mixture was stirred at room temperature for 20 minutes. *Tert*-butyl bromoacetate was added and the reaction was stirred at room temperature over night. The next day the reaction was quenched with HNEt<sub>3</sub>Cl dissolved in MeCN/ H<sub>2</sub>O. The solvent was removed under vacuum and the residue redissolved in MeCN, filtered and the filtrate was dried under reduced pressure. The product (**32**) was obtained as a pale yellow, waxy solid (0.72 mg,  $y = 72\%$ ).

$M_w = 5091.11$

$m/z$  (ESI+) = 1715 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1290 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1036 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 866 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>. *calc.*

(C<sub>233</sub>H<sub>459</sub>O<sub>118</sub>) = 1715 [M+3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1290 [M+4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 1036 [M+5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 866 [M+6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>.

Purity (UPLC-ELSD)= 99 %

## 5.3 Results and Discussion

### 5.3.1 Monofunctional PEGylation agents

All PEGylation agents used in the currently FDA-approved PEGylated therapeutics (**Table 1**) are monofunctional species, where one end of the PEG chain is equipped with a reactive functional group, whilst one or (in case of the branched versions) two ends are capped with a chemically inert methoxy group (MeO- or H<sub>3</sub>CO-) in order to avoid post-PEGylation reactions. Hence, this work focused on the synthesis of linear, monofunctional mPEG-X species, with X being a reactive functionality that can be used to PEGylate different moieties present in proteins. The reactions were performed on Eg<sub>112</sub> species to reach defined M<sub>w</sub> PEGylation agents equivalent to a weight of about 5 kDa. In general, the Nanostar Sieving approach has the advantages of not only allowing for purification through membrane filtration, but also being a desymmetrisation technique in itself: as the growing PEG chains are attached to the nanostar on one end, when the desired PEG chain length is reached, the other end can be modified. After cleavage from the nanostar, an unsymmetric, linear PEG species is obtained. The prerequisite is that the added functional group withstand the cleavage with BCl<sub>3</sub>. Within the work for this chapter, the chain ends of previously synthesised Hub(Eg<sub>112</sub>-OH)<sub>3</sub> (**Chapter 4**) were methylated using iodomethane (MeI) to reach a methoxy-capped PEG chain, which is not affected by the cleavage and results in a monofunctional MeO-Eg<sub>112</sub>-OH. This species was used as the starting material for several modification reactions to reach the active PEGylation agents shown in **Figure 57**. Structure **15 - 19** are all active esters obtained through the DCC-initiated activation of a carboxylic acid with *N*-hydroxysuccinimide. These structures react with primary amine groups in a protein (from for example lysine, histidine, serine) which forms an amide bond under the liberation of *N*-hydroxysuccinimide. The PEGylation with active esters is random, so if there is more than one primary amine group present, PEGylation with an active ester will result in the formation of different PEGamers and positional isomers.

To achieve a candidate for the selective PEGylation of proteins, a mPEG propionaldehyde (mPEG<sub>5022</sub> PA) was synthesised (**Figure 57**, blue box). If the *N*-terminus of a protein is not buried, it can be selectively PEGylated through reaction with a mPEG aldehyde under a specific pH, as the  $pK_a$  of the *N*-terminus differs from the  $pK_a$  of other amine groups within the protein.

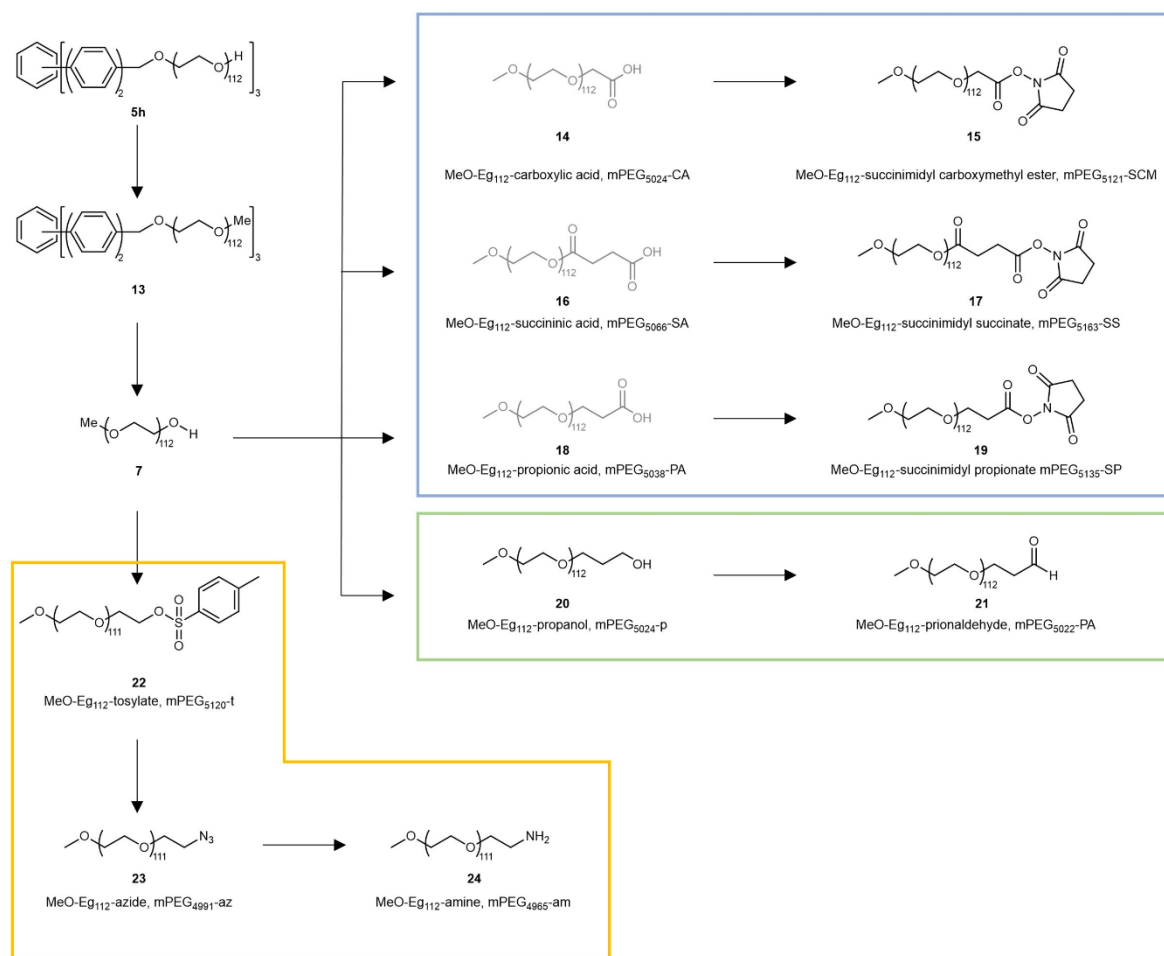


**Figure 57. Overview of the linear, monofunctional defined  $M_w$  5 kDa PEGylation agents synthesised within this work.**

### 5.3.1.1 Carboxylic Acids, Active Esters and Aldehydes

The final aim of the work in hand was to perform a first random PEGylation of a protein using a defined  $M_w$  5 kDa PEG reagent. Due to high abundance of primary amine groups within most proteins, an mPEG active ester was chosen as the first PEGylation reagent to be synthesised.

Direct oxidation methods, using for example  $\text{CrO}_3$  (Jones oxidation) or  $\text{KMnO}_4$  are not recommended to synthesise PEG carboxylic acids or aldehydes, as PEG is prone to oxidative degradation.<sup>161 154,155,157,159</sup> The increase in dispersity would not easily be detectable for an initially disperse material, but in case of a defined  $M_w$  PEG starting material the initially precise mass would be turned into a disperse mixture of PEG of different chain length, which would result in impeded analysability through mass spectrometry. Hence, other ways of synthesising PEG carboxylic acids and aldehydes were explored within this thesis.

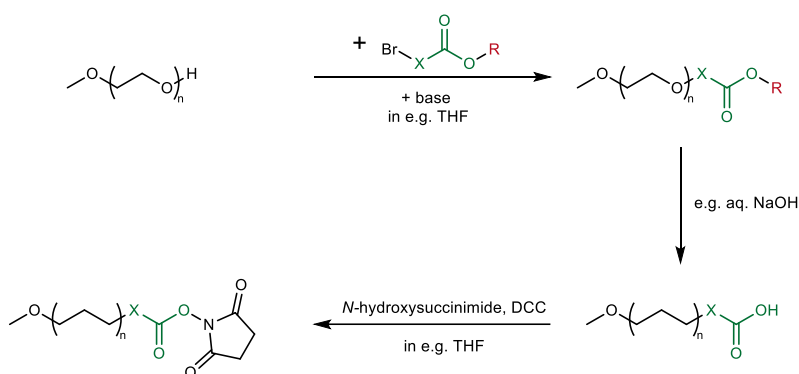


**Figure 58. General overview of the synthesis of a mPEG active ester using an ester of bromoacetic acid as an alkylating agent.** Hydrolysis of the formed ester results in the carboxylic acid which is the activated with *N*-hydroxy succinimide/ DCC to reach the NHS-ester.

A first alternative way to synthesise PEG active esters is to react the PEG alcohol with an ester of bromoacetic acid, which is a strong alkylating agent (**Figure 58**).<sup>15,82,84,162,163</sup> The formed ester can be hydrolysed using for example sodium hydroxide, to reach the according carboxylic acid (**Figure 58**). Subsequent reaction with *N*-hydroxysuccinimide (NHS) using *N,N'*-dicyclohexylcarbodiimide (DCC) as a coupling reagent, yields the active ester. Depending on the structure of the linking unit between the PEG chain and the ester group, the active esters exhibit different hydrolysis lifetimes and reactivity towards aminolysis.<sup>164</sup>

Within this work the first defined  $M_w$  mPEG active ester derivative to be synthesised was a MeO-Eg<sub>112</sub>-succinimidyl carboxymethyl ester with a molecular weight of 5121 Da (mPEG<sub>5121</sub>-SCM, **Figure 59**). In a first step the MeO-Eg<sub>112</sub>-OH was reacted with ethyl bromoacetate, using sodium hydride as the base. Due to the precise  $M_w$  of the starting material the reaction could be followed

through UPLC-MS and was determined to be finished within 1 hour. The resulting ester was not isolated but hydrolysed straightaway using aq. NaOH to reach the carboxylic acid, which was then converted to the corresponding active ester. The synthesis involved only extraction and ether precipitation as means of purification and had an overall yield (starting from the MeO-Eg<sub>112</sub>-OH) of 77 %.

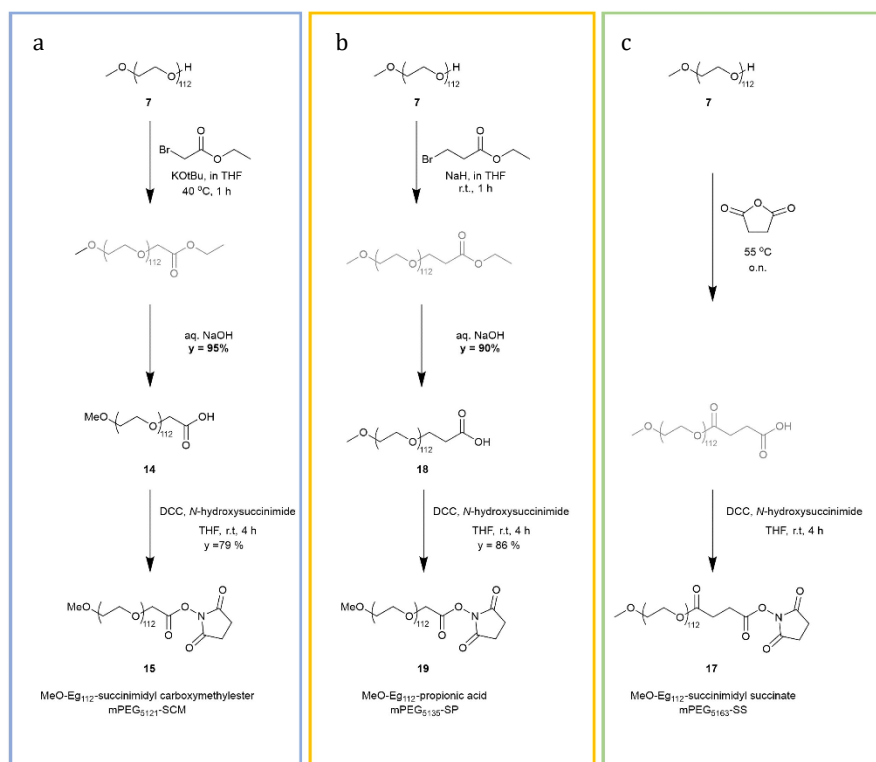


**Figure 59. Overview of the successful synthesis approaches for the three defined  $M_w$  active ester derivatives mPEG<sub>5121</sub>-SCM, mPEG<sub>5135</sub>-SP and mPEG<sub>5163</sub>.** The grey structures were intermediates and not isolated.

Subsequent work revealed that the mPEG-SCM reagent was too unstable for use as a PEGylation agent and would hydrolyse rapidly when in contact with the aqueous buffer solution used for protein PEGylation. Depending on the structure of the linking unit between the PEG chain and the ester group, the active esters exhibit different hydrolysis lifetimes ( $t_{1/2}$ ) and reactivity towards aminolysis.<sup>164</sup> The synthesised mPEG-SCM has been reported to hydrolyse within 0.75 minutes at pH = 8.<sup>164</sup> Hence, it was decided to synthesise two more stable alternatives: the mPEG

succinimidyl propionate ( $t_{1/2}$  = 16.5 min at pH = 8) and the mPEG succinimidyl succinate ( $t_{1/2}$  = 9.8 min at pH = 8).

The first attempted route to reach the more stable MeO-Eg<sub>112</sub>-succinimidyl propionate (mPEG<sub>5135</sub>-SP) was based on the work by GUSTAVINO *et al.*<sup>165</sup> In this work the researchers equipped an mPEG alcohol starting material with an aliphatic linker by first converting it to a mPEG mesylate and then reacting it with selected alkoxides. For the work in hand, it was decided to leave the mPEG alcohol unfunctionalised, but instead to synthesise a reactive “building block”, based on 1,3-propanediol (**Figure 60**). As shown previously within the synthesis of defined M<sub>w</sub> PEG, two tosylate-functionalised chain ends can react with each other to form a dimer species. In case of the short propyl-based spacer, this dimer is easily removable as it is soluble in diethyl ether, whilst the long PEG species are not. A dimer formed of two 5 kDa PEG chains would have been difficult to remove and would have resulted in yield-loss of the valuable Eg<sub>112</sub> species.

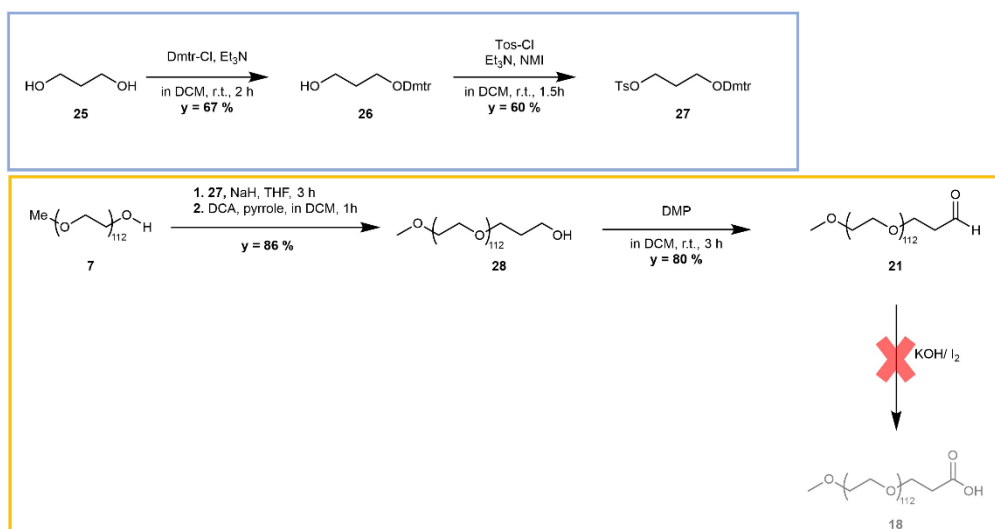


**Figure 60. DMP-route to synthesise mPEG propionaldehyde (21).**

The propyl-based linker unit was obtained in a straightforward 2-step synthesis (**Figure 60**) and based on low-cost materials. The linker was then attached to the mPEG alcohol using sodium hydride as the base, and after deprotection and ether precipitation the pure mPEG propanol was obtained in 86 % yield. Based on the work by GUSTAVINO *et al.* the reagent Dess-Martin

periodinane (**Figure 60**) was used as a mild oxidising agent and the MeO-Eg<sub>112</sub> propionaldehyde (mPEG<sub>5022</sub>-PA) was reached in a yield of 80 %. <sup>165</sup> This presents a facile route to achieve valuable defined M<sub>w</sub> PEG propionaldehydes, which can be used to selectively PEGylate the *N*-terminus of a protein when the reaction is performed at a specific pH). For example, the block buster drug Neulasta and its five biosimilar are composed of the protein filgrastim, which is selectively PEGylated at its *N*-terminus using a 20 kDa mPEG propionaldehyde.

In a next step it was attempted to further oxidise the obtained mPEG aldehyde to the according carboxylic acid. Reproducing the proposed route by using I<sub>2</sub> in the presence of KOH remained unsuccessful. Similarly, the attempt of oxidising the mPEG propionaldehyde with pyridinium dichromate resulted in no reaction.

In order to successfully synthesise the desired mPEG<sub>5135</sub>-SP, a different route was chosen, equivalent to the previously used method to synthesise the mPEG<sub>5121</sub>-SCM through reaction of the mPEG alcohol with ethyl bromoacetate (**Figure 59b**). In this case the mPEG alcohol was reacted with ethyl bromopropionate and through subsequent hydrolysis the propionic acid was reached in 90 % yield.<sup>15,162</sup> Activation with *N*-hydroxysuccinimide and DCC delivered the mPEG<sub>5135</sub>-SP with a yield of 86 %.

Converting a mPEG alcohol into the succinimidyl succinate derivative is a straightforward route, involving only two steps (**Figure 59c**).<sup>166</sup> Within this work, the MeO-Eg<sub>112</sub>-OH was reacted with succinic anhydride to reach the succinic acid. Subsequently this was converted to the NHS-ester using the previously used method based on NHS/DCC.

### 5.3.1.2 Other functionalities

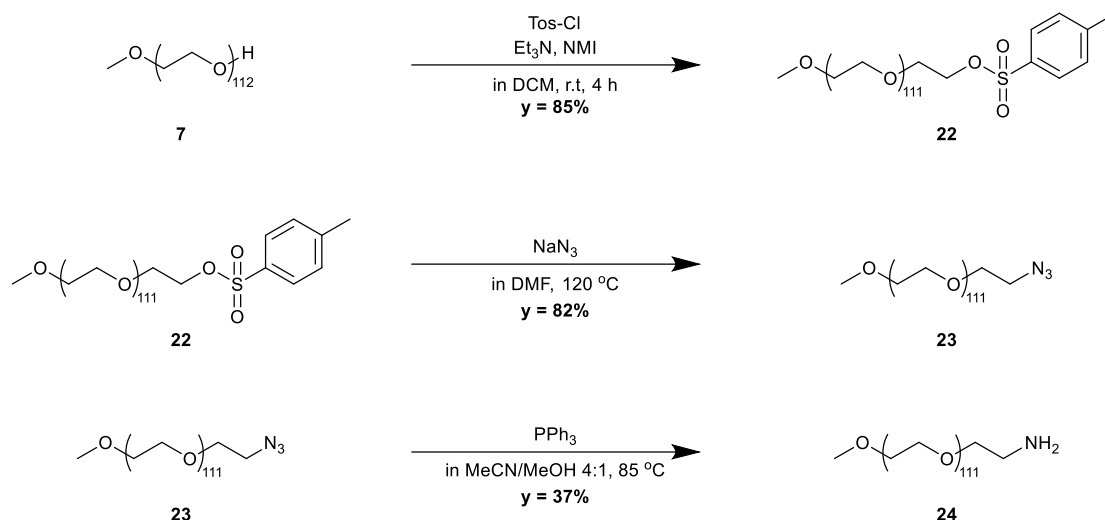


**Figure 61. Overview of the synthetic pathways to reach the defined  $M_w$  mPEG<sub>5120</sub>-tosylate, mPEG<sub>4991</sub>-azide and mPEG<sub>4965</sub>-amine.**

Other defined, monofunctional linear PEG 5kDa derivatives synthesised within this work, included the MeO-Eg<sub>112</sub> tosylate, MeOEg<sub>112</sub> azide and MeO-Eg<sub>112</sub> amine. The tosylate function is a useful leaving group, that can either be further converted (as in this synthesis) or used to attach the PEGylation agent to a hydroxyl function. mPEG-azides are valuable materials for click-chemistry reactions with alkynes. The amine functionality can for example be used to attach the PEG to an active ester moiety. The syntheses of the MeO-Eg<sub>112</sub> tosylate and MeOEg<sub>112</sub> azide were straightforward, resulting in yields of 85 % and 82 % respectively (**Figure 61**).

The reduction of the amine using triphenylphosphine resulted in a yield of only 37 %. This might be due to an increased stability of the phosphazene intermediate and low hydrolysis to the final amine. An alternative synthesis approach would be the reductive hydrogenation with Pd/C under increased pressure.

### 5.3.1 Heterobifunctional PEGylation agents

A first attempt of synthesising heterobifunctional defined  $M_w$  PEGylation agents for the potential usage as linker units, for example in antibody-drug conjugates was pursued in this work as a proof-of-concept. To achieve a heterobifunctional defined  $M_w$  PEG, the first function (or a protected version of it) has to be attached before the PEG is cleaved from the nanostar, in order to keep the unsymmetric nature. This functionality must withstand the cleaving conditions with  $\text{BCl}_3$ . The heterobifunctional structure chosen for this work was the linear 5 kDa PEG equipped with a tert-butyl ester group on one side and an azide group on the other (**32**). Further reduction of the azide would result in an amine, whereas the ester could be hydrolysed to the acid and further be converted to an active ester.

For synthesising structure **32**, the  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$  was first tosylated and subsequently the tosylate was displaced with an azide function. Both steps result in acceptable yield of 84 % and 94 % (**Figure 62**). Subsequent cleavage resulted in the  $\text{HO-Eg}_{112}\text{-N}_3$ , which was then further reacted with tert-butyl bromoacetate to achieve structure **32**. This is the first example of a defined  $M_w$  heterobifunctional 5 kDa PEG.

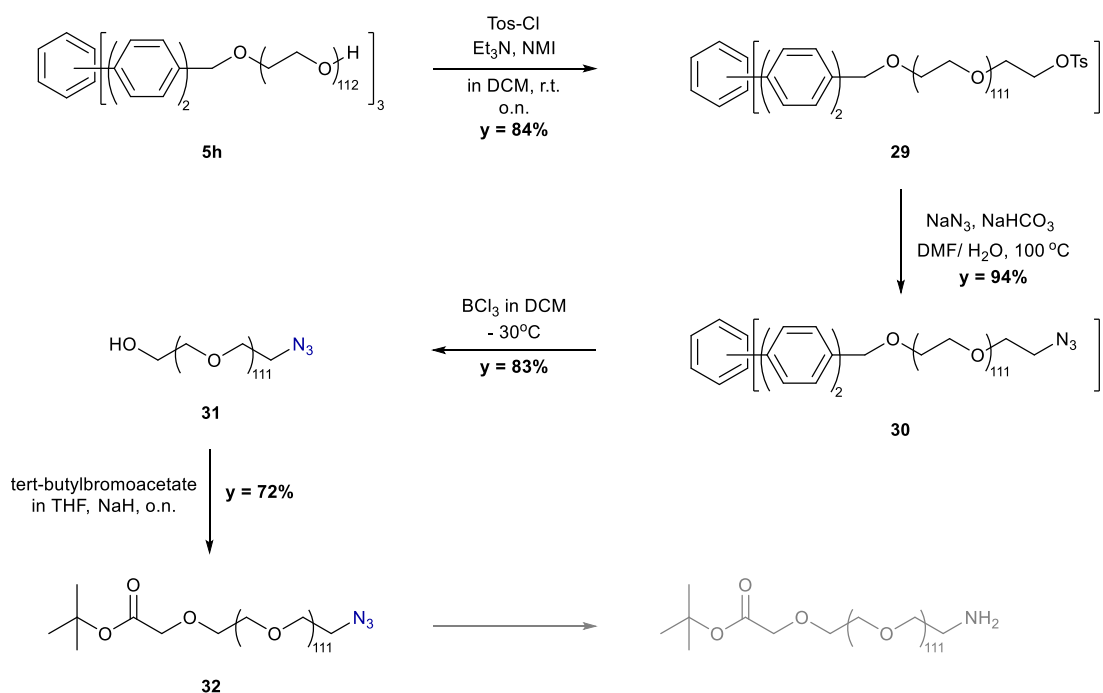


**Figure 62.** Reaction pathway to achieve heterobifunctional defined  $M_w$  PEG (**32**).

## 5.4 Conclusions

Within this chapter a library of different monofunctional defined  $M_w$  PEGylation agents with a molecular weight of close to 5 kDa has been synthesised successfully, which is laying the foundation for the synthesis of defined  $M_w$  PEGylated therapeutics. Different active esters were synthesised, which can be used for the random PEGylation of peptides and proteins, as well as a PEG propionaldehyde which can be used for the site-selective N-terminal PEGylation of proteins.

Furthermore, a first example of a heterobifunctional defined  $M_w$  PEG 5 kDa, bearing an ester group on one end and an azide function on the other, has been produced successfully.

A significant advantage of the defined  $M_w$  PEGylation agents is that they exhibit a clear mass spectrum with exactly defined masses. This allows for monitoring of most chemical reaction through UPLC-MS, if the species (starting material and product) have different retention times. Within this work it has been shown, that defined  $M_w$  PEGs exhibit superior separation behaviour on a reversed-phase UPLC column, again allowing for easier reaction monitoring through UPLC-MS.

In case of disperse PEG, the completion of a reaction is usually either followed through TLC, which is not an accurate method. Detailed reaction monitoring or the investigation of reaction kinetics requires more sophisticated assays or post-reaction analysis of samples through NMR.

For example, during the synthesis of the mPEG<sub>5135</sub>-SP the defined nature of the PEG allowed for clear monitoring of the different reaction steps through simple UPLC (equipped with an ultraviolet diode array detector, an evaporative light scattering detector and an electrospray mass spectrometer) and for straight-forward identification and quantification of any intermediates or side-products. For example, in the final UPLC-ELSD chromatograms (**Figure 63a**) and the corresponding MS traces of the defined mPEG-SP (**Figure 63b**) the product and the two impurities were clearly distinguishable (Peak 1: MeO-Eg<sub>112</sub>-OH and Peak 2: the hydrolysed NHS-Ester/ carboxylic acid). Baseline separation of the peaks allowed for quantification through

integration of the ELSD chromatogram. In case of the disperse derivative, the mass-distribution resulted in poor separation, as well as an unanalysable ESI<sup>+</sup> trace.

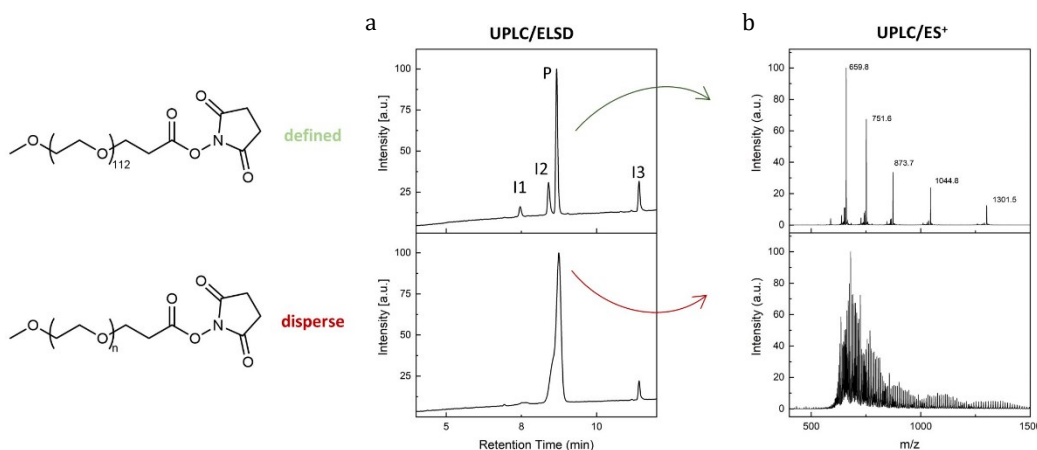


**Figure 63. Comparison of the UPLC-MS data of the defined  $M_w$  mPEG SP synthesised in this work and a commercially available disperse derivative.** **a**, UPLC/ELSD traces of the synthesised defined mPEG-SPA (top) and the commercially available disperse derivative (bottom) (15 min 29-95 % B gradient with A: H<sub>2</sub>O + 5 mM NH<sub>4</sub>OAc, B: MeCN/ MeOH 4:1). The defined sample shows 4 peaks (I1: traces of starting material mPEG-OH<sub>4966</sub>, I2: hydrolysed product/ carboxylic acid, P: product (mPEG-SPA), I3: NHS-breakdown product). For the disperse sample peak I1 is broad and peak I2 and P are not separated. **b**, For the defined sample the underlying ESI<sup>+</sup> traces lead to the corresponding mass. For example (shown) the m/z series 1301.5, 1044.8, 873.7, 751.6, 659.8 corresponds to a  $M_w$  of 5135.11, which is the mPEG-SPA ([M+4 NH<sub>4</sub>]<sup>4+</sup>, M+5 NH<sub>4</sub>]<sup>5+</sup>, M+6 NH<sub>4</sub>]<sup>6+</sup>). For the disperse sample the determination of a precise molecular weight is not possible.

## Chapter 6 The PEGylation of Bovine Serum Albumin – a comparative study

### Abstract

A first comparative study on the random PEGylation and subsequent characterisation of the protein bovine serum albumin was performed, using both of the previously synthesised defined  $M_w$  mPEG SP active esters (mPEG<sub>5163</sub>-SS and mPEG<sub>5135</sub>-SP) and a commercially available disperse 5 kDa mPEG SP. First, the mono-PEG BSA was isolated using size-exclusion chromatography. In a next step, the mono-PEG BSA sample underwent a peptide mapping procedure based on tryptic digestion. In case of PEGylation with the mPEG<sub>5135</sub>-SP, a mixture of peptide and PEG-peptide fragments was obtained that was analysed through UPLC-MS. Two positional isomers of mono-PEG BSA were identified. It was demonstrated that the defined  $M_w$  PEGylation agents facilitate the straightforward identification of the PEGylated fragments within a PEGylated protein through a simple peptide mapping approach using UPLC-MS.

### Keywords

BSA, protein, PEGylation, peptide mapping, positional isomers, PEGylation sites

### Highlights

- BSA was PEGylated with the first 5 kDa defined PEGylation agent and an equivalent disperse derivative
- Samples of mono-PEG BSA were isolated through size-exclusion chromatography (SEC)
- The sample were subjected to a peptide mapping approach based on tryptic digestion
- When the mPEG<sub>5135</sub>-SP was used, the PEG-peptide fragments were stable enough to separate and identify them through UPLC-MS, which was not possible for the disperse derivative
- Two positional isomers of mono-PEG BSA were identified that way

## 6.1 Introduction

Of the currently 40 FDA-approved PEGylated therapeutics, 31 are PEGylated therapeutic proteins or peptides.

The disperse nature of the used PEGylation agents results in batch-to-batch variations and difficulties in purification and characterisation of the PEGylated therapeutics.<sup>8-11</sup> This is a result of compromised chromatographic behaviour and unclear analysis in terms of mass spectrometry. These problems become more pronounced with increasing molecular weight of the PEG and with an increasing number of PEGamers (species with varying numbers of PEG molecules) or positional isomers incorporated in the sample. Despite the availability of advanced analytical methods, the exact binding site of the PEG in a PEGylated protein often remains unclear, or PEG-related impurities remain unidentified.

Even if a protein is selectively mono-PEGylated at its *N*-terminus, the analytical confirmation is often complicated. An example of these analytical difficulties can be found in the FDA product quality review of the therapeutic Fulphila, a biosimilar of the blockbuster drug Neulasta (pegfilgrastim) approved by the FDA in 2018.<sup>55</sup> This therapeutic is widely used to treat neutropenia and is composed of the 18 kDa protein filgrastim that is selectively mono-PEGylated with a 20 kDa linear PEG. In the product quality review, the analytical similarities between the original U.S.-licensed drug Neulasta and the new biosimilar Fulphila (MYL-1401H) are compared. It is shown that the average PEG molecular weights of the two samples are not identical, despite 20 kDa PEG being used for both products. Due to the inherent dispersity, the *N*-terminal PEGylated fragment cannot be identified straightaway using the HPLC-MS chromatogram of the Glu-C or Tryptic digest. Rather, the identification requires a full peptide map of the protein, collection of the PEGylated fragment and analysis of this fraction through MADLI-TOF. This reveals that the heterogeneity of the PEG leads to significant deviations between the expected masses and the masses obtained through MALDI-TOF, potentially resulting in undetected impurities.

Using defined PEG instead of the currently used disperse derivatives would overcome these analytical problems and could have a substantial impact on the overall purity and quality of PEGylated therapeutics. Furthermore, therapeutics PEGylated with defined PEG could potentially exhibit an improved biodistribution profile compared to the disperse derivatives. The correlation curve of PEG half-life in blood and the PEG molecular weight is particularly steep between 0 and 40 kDa<sup>44</sup>, which is in the range of most reported PEGylation agents. This suggests that using defined PEGylation agents could result in a decreased volume of distribution, which might have a significant impact on the dosing controllability, and the efficiency of the PEGylated therapeutic.

Two of the active defined  $M_w$  PEGylation agents synthesised within the scope of **Chapter 4** (mPEG<sub>5163</sub>-SS and mPEG<sub>5135</sub>-SP ) were used to randomly PEGylate native bovine serum albumin (BSA, **Figure 64**). The same reaction was performed with a commercially available, disperse derivative. The mono-PEGylated derivative (PEGamer) was then isolated from both reaction mixtures using size exclusion chromatography and its purity was confirmed through MALDI-TOF. Both mono-PEG BSA samples were enzymatically digested with trypsin, resulting in a mixture of peptide and PEG-peptide fragments, which were separated and analysed using UPLC-MS. The chain length homogeneity of the defined PEG allowed for separation and exact mass determination of the different PEG-peptide fragments. The exact sequences of the PEGylated fragments were identified by calculating the differences between the molecular weights of PEG-peptide fragments and the attached PEG residue and comparing those masses with a peptide mapping online database (ProteinProspector v 6.4.5). These results show that defined PEGs are a powerful tool to easily identify the PEGylated regions within a protein using relatively straightforward techniques. Furthermore, the general characterisation of the PEGylation agent,

as well as the PEGylated therapeutic is more straight-forward than for disperse agents. Together these can be used to increase the overall quality of PEGylated therapeutics.

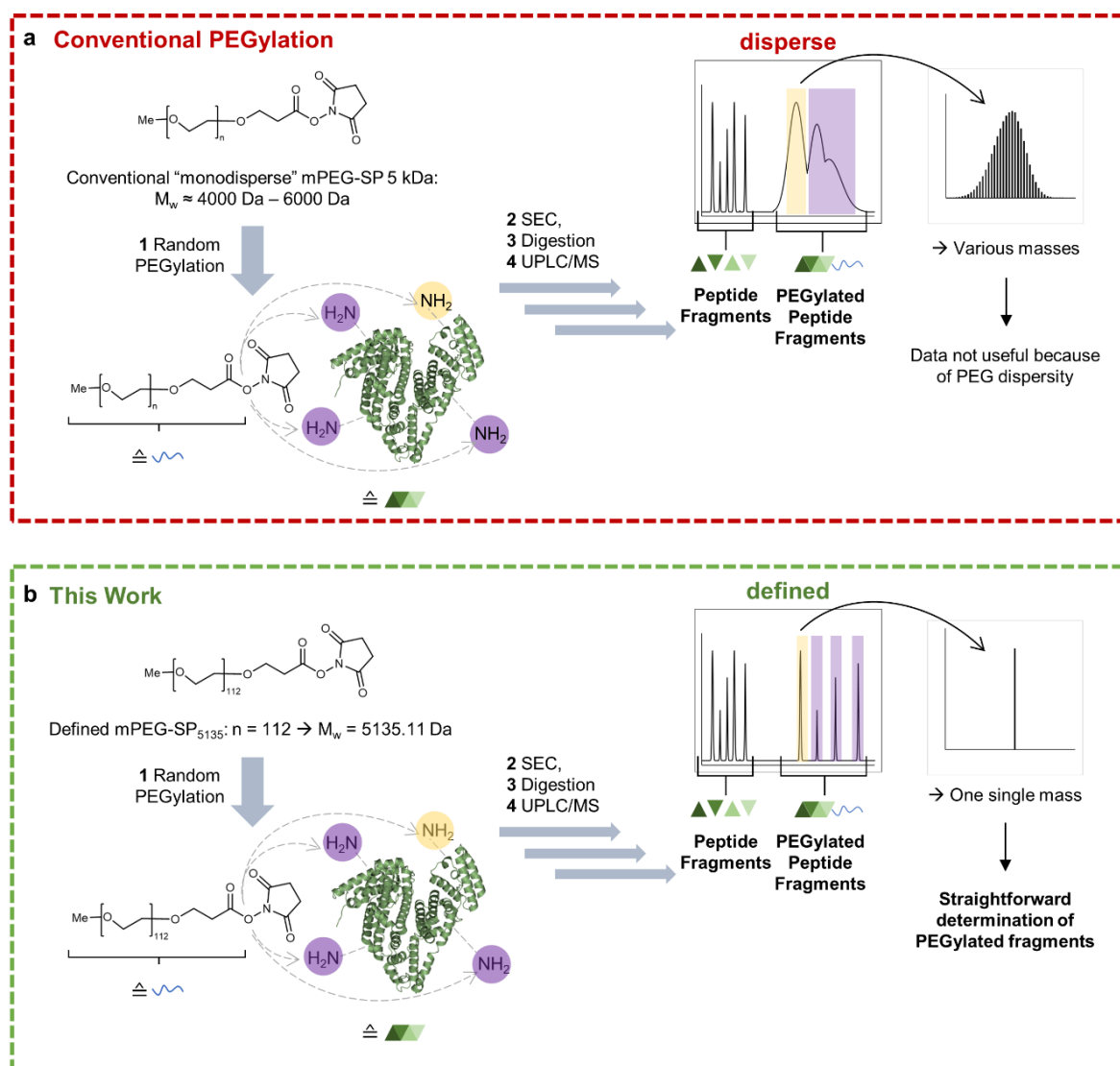


**Figure 64. Comparative study of the random PEGylation and characterisation of BSA using defined and disperse PEG 5 kDa.** **a**, Conventional PEGylation: uses disperse PEGylation agent. Random PEGylation at the primary amine groups of BSA, using defined or disperse 5 kDa linear mPEG-SP (1). Size-exclusion chromatography to isolate the mono-PEG BSA isomers (2). Digestion (3): enzymatic digestion of the mono-PEG BSA results in a mixture of PEGylated and un-PEGylated peptide fragments. The different peptide/ PEG-peptide species are then separated via RP-UPLC (4). In the case of the disperse PEG the resulting data cannot be interpreted. **b**, In this work a defined, 5 kDa PEGylation agent is used for the first time to randomly PEGylate BSA. After steps 1-4 in-line ESI-MS allows for extraction of the masses for identifying individual PEG-peptide fragments. The masses of the PEGylated peptide fragments can be calculated and (equivalent to a standard peptide mapping procedure) these can then be compared to an online digestion database to identify the PEGylated fragments.

## 6.2 Materials and Methods

### 6.2.1 Materials

Bovine Serum Albumin (heat shock fraction, protease free, fatty acid free, essentially globulin free, pH 7, ≥98%) (Sigma-Aldrich), mPEG-Succinimidyl Propionate (5 kDa, Creative PEGworks), Phosphate Buffered Saline (PBS, pH 7.4, liquid, sterile-filtered) (Sigma-Aldrich), Sinapinic acid (matrix substance for MALDI, ≥99 %), *trans*-2-[3-(4-*tert*-Butylphenyl)-2-methyl-2-propenylidene]malononitrile (DCTB, matrix substance for MALDI, ≥99 %) (TCI), Trypsin (from bovine pancreas, Type I, ~10 000 BAEE units/mg protein) (Sigma-Aldrich), 1,4-Dithiothreitol (DTT) (Merck) were used as received without further purification. All solvents used for the syntheses, purifications and analytics were purchased from VWR.

### 6.2.2 Methods

#### *PEGylation of BSA.*

The procedure for the small scale batch PEGylation of BSA with 5 kDa mPEG-SP and mPEG-SS was based on a paper published by van Alstine and Fee in 2004.<sup>65</sup>

50 mg of bovine serum albumin (BSA, 66 kDa) was dissolved in 5 mL of PBS buffer (pH = 7.4). The solution was stirred in an open vial at room temperature and 50 mg of PEGylation agent was added (either disperse or defined 5 kDa mPEG-SP or 5 kDa mPEG-SS). The reaction was finished after 1 hour (monitored through SEC) and quenched with 0.1 M HCl.

#### *Purification of PEGylated BSA through SEC.*

A sample of 0.5 ml was pushed through a 0.25 micron syringe filter and 100 µl were injected onto a Cytiva Superdex 200 Increase 10/300 GL SEC column, which was implemented into an Agilent 1100 HPLC stack with UV detector. 0.1 M NH<sub>4</sub>OAc was used as the mobile phase. Flowrate: 0.5 ml/min, length: 60 min, injection volume: 50 µl, T = 30 °C.

0.5 ml (1 min) fractions were collected starting from 18 min, up to 40 minutes. Each sample was analysed through MALDI-TOF, as described above.

The 23-24 min sample (disperse PEG) and 24- 25 min sample (defined MW PEG) contained only mono-PEGylated BSA (according to MALDI spectrum). The concentration was estimated by calculating the relative peak area of the mono-PEG peak, compared to the other protein-related peaks. The samples were subsequently used for enzymatic digestion experiments.

#### *Enzymatic Digestion of mono-PEGylated BSA.*

Three of the previously collected SEC samples containing mono-PEG BSA were concentrated into one by using an Amicon Ultra 0.5 centrifugal filter and redissolved in 0.5 ml 5 mM DTT/ 50 mM NaHCO<sub>3</sub> in ultrapure water. The solution was incubated at 37 °C for one hour to denature the protein. Subsequently 25 µl Trypsin solution (1 mg/ml in 50 mM acetic acid) was added to the sample to reach a ratio Enzyme:Protein of 1:20 (w:w). The sample was incubated at 37 °C overnight and then analysed through UPLC-MS (Mobile Phase: MeCN/MeOH 4:1 and 0.1% Formic Acid in Water, 15 cm C18 Column, 30 min 1-50 % gradient). The underlying ESI<sup>+</sup> trace of each peak was matched with the corresponding peptide fragment using (exact procedure see supplementary information).

## **6.3 Results and Discussion**

### **6.3.1 PEGylation and Peptide Mapping of BSA using mPEG<sub>5163</sub>-SS**

The 66.5 kDa protein bovine serum albumin (BSA) was chosen as an inexpensive and readily available test candidate for random PEGylation on its primary amine residues. Random PEGylation generally results in a heterogenous mixture of species with different numbers of PEG-molecules bound to the protein (PEGamers), with every PEGamer usually having different positional isomers. In order to allow for a simple first proof-of-concept study it was decided, to isolate the mono-PEGylated BSA for further experiments.

The first PEGylation experiments were performed using the previously synthesised active ester reagent mPEG<sub>5163</sub>-SS, due to the more straightforward synthesis. BSA was randomly PEGylated by adding the PEGylation agent to a stirred solution of the protein in PBS buffer (pH = 7.4). To

monitor the reaction progress, samples were taken from the solution initially every 30 minutes (up to 4 hours) and analysed on a HPLC equipped with a size-exclusion column and a UV-detector. Fractions were collected in the time frame the products eluted, and the fractions were analysed using MALDI-TOF-MS. In this way, it was possible to assign every peak to a certain species (native BSA, mono-PEGylated BSA, di-PEGylated BSA) and to collect fractions containing only one of the different PEGamers. It was found, that the PEGylation reaction was completed after 1 hour, with the mono-PEG BSA being the most abundant PEGamer (**Figure 65**).

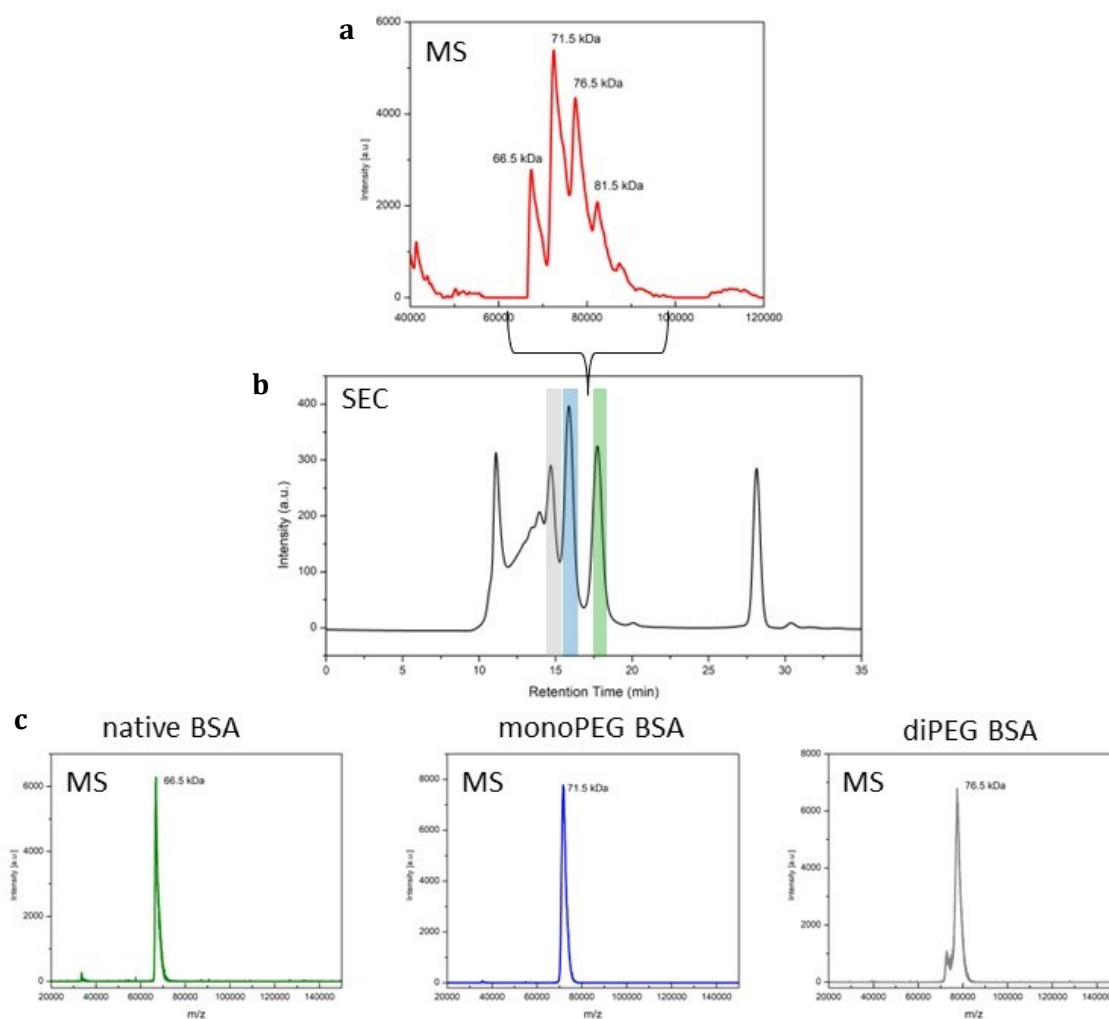


**Figure 65. Overview of the characterisation and separation of PEGylated BSA through MALDI-TOF-MS and size-exclusion chromatography.** **a**, MALDI-TOF-MS of the crude PEGylation mixture derived from PEGylating BSA with mPEG<sub>5163</sub>-SS after 1 hour. The spectrum shows peaks for the native BSA (66.5 kDa), the mono-PEG BSA (71.5 kDa), the di-PEG BSA (76.5 kDa), and the tri-PEG BSA (81.5 kDa). **b**, Size-exclusion chromatogram of the PEGylation mixture after 1 h. **c**, MALDI-TOF-MS of three of the fractions collected from the SEC; the spectra confirm the presence of native BSA (66.5 kDa, green peak), the mono-PEG BSA (71.5 kDa, blue peak) and di-PEG BSA (76.5 kDa, grey peak).

Collected samples of the mono-PEG BSA were concentrated using an ultracentrifuge filter and subsequently subjected to a peptide mapping procedure based on digestion with trypsin. Trypsin is known to cleave proteins at their lysine and arginine residues and was chosen for the enzymatic digestion of the mono-PEG BSA. Because of potential inhibition of the digestion due to PEGylation, the protein was denatured prior to digestion using dithiothreitol (DTT, reduction of disulfide bonds), to make the amino acid sequence more accessible. The digestion reaction was quenched

after 16 hours, and the sample of the resulting peptide/PEG-peptide mixture was purified through ultracentrifugation and then run on a UPLC equipped with a 5 cm  $C_{18}$  column and a mass spectrometer with an electrospray/single quadrupole set-up. The analysis was performed in positive mode and using  $NH_4OAc$  as the buffer, so that only  $NH_4^+$  adducts were formed. The chromatograms exhibited a region at lower retention times which (based on the underlying mass spectra) contained only peptide fragments of BSA (peptide region/ BSA fingerprint region) (**Figure 66**). From a retention time of 17 min and higher, the 1 h sample showed signs of species with a molecular weight of 5 kDa and higher, indicating the presence of PEG species. Nevertheless, the signals were too weak to allow for a reliable analysis. Suspecting incomplete digestion to be the cause for the low intensity signals, another sample was taken and analysed after 16 hours. This time, the  $\sim 5$  kDa signals present in the 1 h sample were even lower in intensity, but a strong signal had appeared at a higher retention time. The underlying mass spectrum showed the clear  $m/z$  series for a  $M_w$  of 4965 Da. This indicated the hydrolysis of the ester bond present in the PEG-unit during the digestion reaction, reducing the concentration of detected PEG-peptide species and making an identification of these impossible. It was hence decided, to perform an equivalent

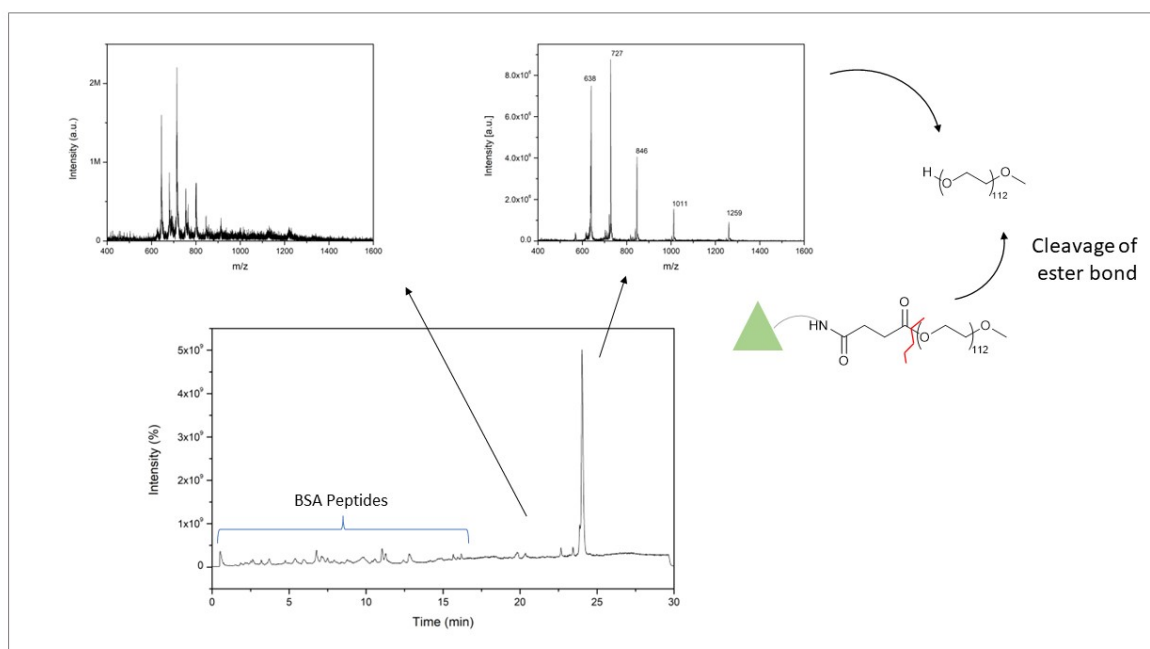


**Figure 66.** Overview of the first peptide mapping results performed with BSA PEGylated with the mPEG<sub>5163</sub>-SS reagent.

experiment using mPEG<sub>5135</sub>-SP as the PEGylation agent, which does not exhibit the labile ester bond.

### 6.3.1 PEGylation and Peptide Mapping of BSA using mPEG<sub>5135</sub>-SP

For this the next attempt, BSA was randomly PEGylated with the previously synthesised defined 5 kDa mPEG-SP and a commercially available disperse derivative. The mono-PEG BSA for both samples was isolated through size exclusion chromatography and subsequently subjected to the previously used peptide mapping procedure.

Comparison of the total ion count traces (**Figure 67a**) of both samples (derived from PEGylation with both the defined and the disperse mPEG-SPA, referred to as “defined PEG” and “disperse PEG”) revealed significant differences. Although both chromatograms show the same characteristics in the protein fingerprint region (with the same underlying MS traces), at retention times over 10 minutes the defined PEG sample shows a clear separation of species with discrete m/z series, whilst the disperse PEG sample is composed of broad and overlapping peaks. The underlying m/z series show the characteristic dispersity pattern for PEG, with a mass difference of +/- 44 for every ion species, which does not allow for further analysis. In contrast, the TIC of the defined PEG sample showed several discrete m/z series (different PEG-Peptide fragments), for which the according molecular weights were calculated. After subtracting the molecular weight of the attached PEG for every found mass (5021.03 g mol<sup>-1</sup>), the resulting masses were matched with digestion data obtained from an online database search and analytical tool (ProteinProspector v 6.4.5) to identify two main PEGylated peptide fragments in mono-PEG BSA (**Figure 67b**). For example, at retention time 20.8 min the underlying ESI<sup>+</sup> trace shows an m/z series of 968, 849, 757 which corresponds to a molecular weight of about 6650 Da (968: [M + 7 NH<sub>4</sub><sup>+</sup>]<sup>7+</sup>, 849 [M + 8 NH<sub>4</sub><sup>+</sup>]<sup>8+</sup>, 757 [M + 9 NH<sub>4</sub><sup>+</sup>]<sup>9+</sup>). This means the corresponding peptide

fragment has a molecular weight of around 1629 g mol<sup>-1</sup> and the only matching digestion fragment is the sequence position 286 – 299, with a molecular weight of 1628.85 g mol<sup>-1</sup>.

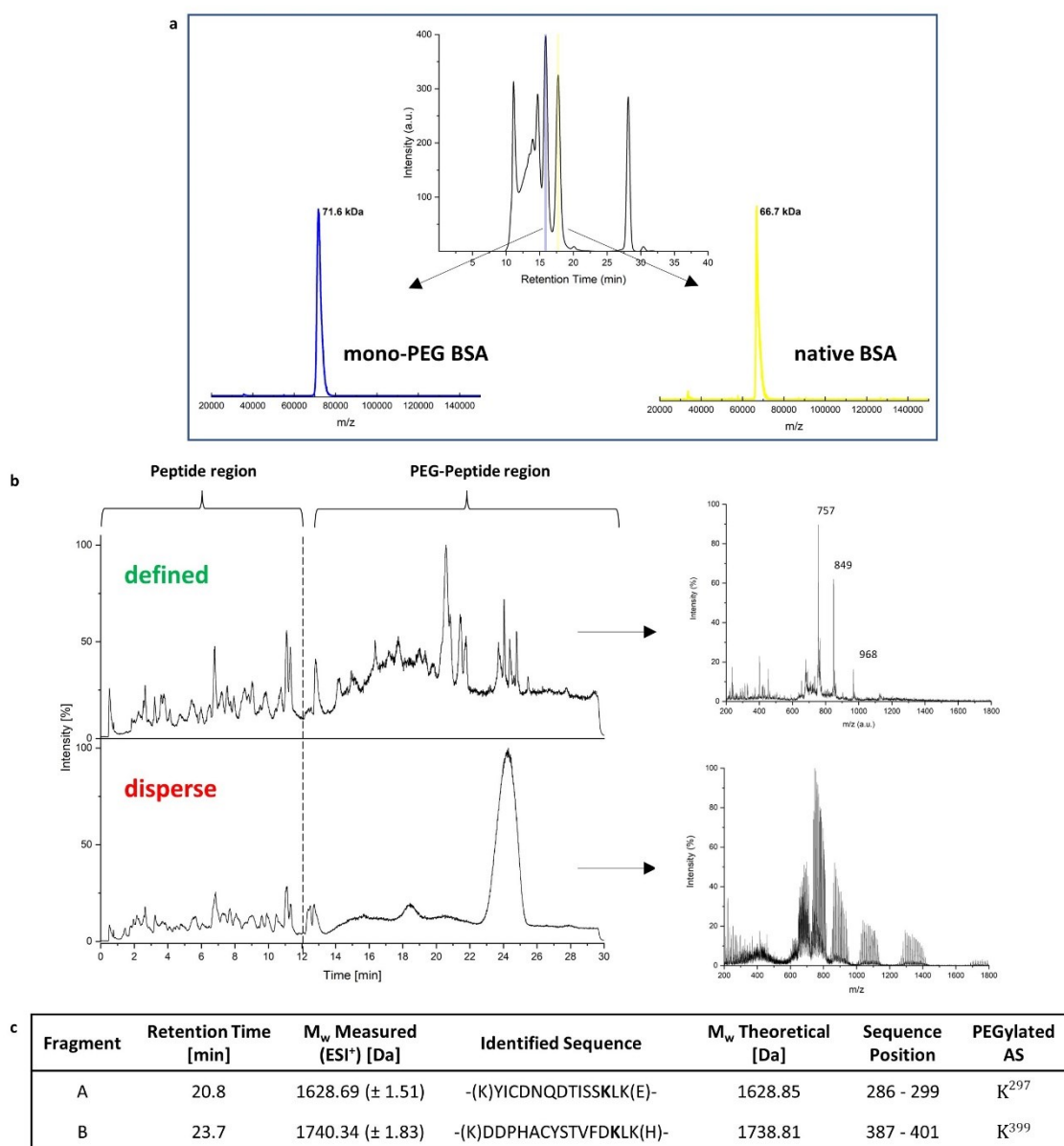


**Figure 67. Overview of the results of the identification of the PEGylated fragments in mono-PEGylated BSA through a peptide mapping approach.** **a**, size-exclusion chromatogram of PEGylated BSA (top) and the MALDI-TOF spectra of the fraction of remaining native BSA (yellow) and the mono-PEG BSA (blue). **b**, TIC trace of UPLC chromatogram of Trypsin-digestion samples of mono-PEG BSA, PEGylated with a disperse (bottom, red) and defined (top, green) 5 kDa mPEG-Succinimidyl Propionate. Examples are shown for the ESI<sup>+</sup> traces extracted from the chromatograms. For the defined species (RT 20.8 min), the m/z series (968, 849, 757) corresponds to a total mass of 6650 Da. Subtracting the PEG mass (5021.03 Da) identified the PEGylated fragment (1629 Da) as Pos. 286 - 299 (AS sequence -YICDNQDTISSK(LK-). The disperse derivative does not allow for extraction of exact masses. **c**, data of the two identified PEGylated fragments in mono-PEGylated BSA.

This way two positional isomers contained in the previously isolated mono-PEG BSA sample were identified down to the PEGylated sequence (**Figure 67c**). The presence of more positional isomers is possible, but could not be confirmed within these experiments, due to the relatively weak signals in the resulting mass spectra. Further digestion experiments, preferably with enzymes cleaving at different amino acids would be necessary to confirm the obtained data.

In general, the presented strategy provides the basis for a fast and efficient technique for identifying the PEGylation sites within a PEGylated protein. It can be applied to PEG-protein species of varying sizes and to mixtures consisting of different positional isomers. Knowing the PEGylation sites within a therapeutic is important for understanding and predicting the drug's efficacy, as well as for quality control and safety assessments. Using defined PEGylation agents would make the synthesis and purification processes more reliable and easier control, and at the same time the defined PEGylated therapeutics might show improved biodistribution behaviour. Overall, this could offer significant improvements in terms of purity and required dosages, thereby decreasing potential risks and minimising adverse effects for patients.

## 6.4 Conclusions

Within the last chapter of this thesis an initial comparative investigation was conducted to assess the random PEGylation and subsequent characterisation of bovine serum albumin (BSA) using two previously synthesised defined  $M_w$  mPEG SP active esters (mPEG<sub>5163</sub>-SS and mPEG<sub>5135</sub>-SP) alongside a commercially available disperse 5 kDa mPEG SP. Initially, mono-PEG BSA was separated through size-exclusion chromatography. Subsequently, the mono-PEG BSA sample underwent a peptide mapping procedure based on tryptic digestion. Using the mPEG<sub>5135</sub>-SP resulted in stabile PEG-peptide fragments that were separated and analysed through UPLC-MS. This analysis resulted the identification of seven distinct positional isomers of mono-PEG BSA. The study demonstrated that the use of defined  $M_w$  PEGylation agents facilitates the direct identification of PEGylated fragments within a PEGylated protein using a simple peptide mapping approach with UPLC-MS. As described in **Section 2.1.4.3** conventional approaches of identifying

the PEGylation sites in PEGylated proteins involve a series of different techniques (chromatographic separation of PEGamers and positional isomers, digestion, and sequencing of the PEGylated fragments) and the results often remain inconclusive. The herein presented approach was tested on a mono-PEGylated protein but can potentially also be applied to the native PEGylation mixture (consisting of several different PEGamers), without the need for prior separation of the PEGamers. The use of defined  $M_w$  PEG would thus allow for the straightforward identification of PEGylation sites involving only two steps: a digestion step, followed by UPLC-MS analysis of the resulting crude digestion mixture. This approach furthermore offers the potential to be applicable to various combinations of PEG and proteins, including more complex mixtures and higher molecular weight species. Limitations might arise from the involved digestion procedure, which requires a bond between protein and PEG that is stable towards the enzymatic digestion. Furthermore, the herein presented work only involved the use of a linear, defined  $M_w$  PEG of a molecular weight of 5 kDa, but most FDA-approved PEGylated protein therapeutics are made from PEG with a  $M_w$  of 20 kDa or higher and often involve branched PEG structures. PEGylation with higher molecular weight and branched PEG might inhibit the enzymatic digestion, which could result in obstacles with the herein presented approach. Hence, future research on the efficient and scalable synthesis of defined  $M_w$  PEG beyond 5 kDa is necessary in order to investigate the transferability of the presented results to higher molecular weight PEGylation agents.

## Chapter 7 Conclusions and Future Work

### 7.1 Conclusions

One of the main high-value applications of poly(ethylene glycol) (PEG) is in the form of PEGylation agents in the pharmaceutical industry. All currently marketed PEGylation agents above 1.5 kDa are disperse polymers, which is due to a lack in synthesis techniques for defined versions. The PEGylation agents used in FDA-approved PEGylated therapeutics almost exclusively exhibit a mass of 5 kDa or higher.

Within **Chapter 4** of this work, defined  $M_w$  linear 5 kDa PEG has been synthesised using an iterative approach based on a Williamson etherification reaction and the efficient and scalable Nanostar Sieving technique. The synthesis was previously based on using an  $Eg_{12}$  based BB and within this work an  $Eg_{28}$  based BB was introduced. This reduced the number of chain extension cycles needed to get from the initial  $Eg_{28}$  species to the 5 kDa equivalent ( $Eg_{112}$ ) from seven to three, which significantly increases the efficiency of the process. The atom efficiency of the synthesis was improved by determining the minimum equivalents of BB needed for each chain extension to reach 100 % conversion.

Investigations regarding the formation of impurities revealed the formation of oxidised PEG side products under the Williamson etherification conditions. It was shown how defined  $M_w$  PEG allows for facile reaction monitoring through UPLC-MS, which facilitates mechanistic investigations and contributes to the final purity of the material.

At the end of **Chapter 4**, a comparative study was conducted which included the preparation of two batches of  $Hub(Eg_{112}-OH)_3$ , one based on Nanostar Sieving, and another one based on chromatography. The final material was produced on a multigram scale, and its chain length purity was confirmed through a newly developed UV-labelling approach. It was shown that the Nanostar Sieving-based approach exhibits significant advantages over the synthesis based on chromatography, not only due to its scalability. The overall yield of the synthesis using Nanostar

Sieving was almost double in terms of the overall yield compared to the chromatography-based approach (63 % vs. 33 %).

Within **Chapter 5** the previously synthesised linear, defined  $M_w$  5 kDa PEGs were chemically modified to obtain the first active, defined  $M_w$  PEGylation agents of this size. Seven different monofunctional derivatives were synthesised starting from the mPEG<sub>4966</sub>-alcohol. Three of them were active ester derivatives (mPEG<sub>5135</sub>-SP, mPEG<sub>5163</sub>-SS, mPEG<sub>5121</sub>-SCM), which exhibit different stabilities and reactivities towards primary amine groups. One mPEG aldehyde (mPEG<sub>5022</sub>-PA) was successfully synthesised, which is a candidate for the site-selective (N-terminal) PEGylation of proteins. Three further monofunctional species (mPEG<sub>5120</sub>-tosylate, mPEG<sub>4991</sub>-azide, mPEG<sub>4965</sub>-amine) were produced, which covers a range of other potential linker chemistries. Also in **Chapter 5** a first example of a defined  $M_w$  heterobifunctional 5 kDa PEG has been synthesised. Heterobifunctional PEGs have recently gained more interest due to their application as linker-molecules in antibody-drug conjugates.

The final chapter (**Chapter 6**) had the aim of investigating the application of the previously synthesised defined  $M_w$  PEGylation agents in the PEGylation of proteins. The protein bovine serum albumin (BSA) was chosen as a low-cost test base and was randomly PEGylated with two of the defined  $M_w$  active esters (mPEG<sub>5135</sub>-SP and mPEG<sub>5163</sub>-SS). A comparative study was conducted, including the isolated monoPEG-BSA derived from PEGylation with mPEG<sub>5135</sub>-SP and a derivative based on a commercially available, disperse mPEG-SP. Both monoPEG-BSA derivatives were subjected to tryptic digestion, and it was shown that, in case of the defined  $M_w$  protein, the PEGylated fragments can be identified using a simple peptide-mapping approach.

In conclusion, the first defined, highly pure 5 kDa PEGylation agents were successfully synthesised within this work, using a scalable liquid-phase approach solely relying on membrane-filtration for purification. It was shown that the PEGylation of proteins with defined PEGs results in significant advantages in terms of characterisation and allows for straight-forward identification of the PEGylated region within the protein. Using this new class of defined

PEGylation agents has the potential for better controllability of PEGylated therapeutics, increasing the overall purity, reducing batch-to-batch variations, and could result in enhanced biodistribution behaviour.

## **7.2 Future work**

### **7.2.1 Reaction Development**

One of the main constraints during the Nanostar Sieving based synthesis of defined  $M_w$  PEG is the removal of the relatively large dimer species, that is formed through reaction of two BB units with each other. To improve the overall speed and efficiency of the synthesis, it would be desirable to avoid the dimer formation completely. Up to now it is not exactly known why the dimer forms, but potentially the BB is losing its tosylate group on one point during the synthesis, which results in a DmtrO-Eg<sub>n</sub>-OH species, that then reacts with the remaining BB. There are different potential reasons for this and a structured study would be necessary to identify the cause and optimise the synthesis accordingly.

Another problem that needs further investigation is the formation of oxidation-related side products, that have a similar size to the main nanostar-PEG species, are thus not removable through diafiltration and decrease the final products overall chemical purity. There are several potential sources for initiation of the oxidation (for example peroxides in the solvents, oxygen, autooxidation of the PEG due to increased temperature during azeotroping) and a structured study would be necessary to identify the cause.

### **7.2.2 Process Optimisation**

The Nanostar Sieving based process was shown to be an efficient way to synthesise defined  $M_w$  PEG. Nevertheless, one problem that repeatedly occurred during the different syntheses was the accumulation of particles on the membranes, which increased the rejection drastically. Hence, the membranes had to be changed frequently. One possible source for the fouling with a dark particulate is the gear pump that was used for creating the cross-flow in the second stage of the

system. Therefore, a first step in future work should be to replace the gear pump with a diaphragm pump.

The scale-up of the synthesis from ~ 10 g to > 100 g batches was not part of the work in hand, but should be investigated in the future for making the process industrially relevant. The reactions can easily be scaled up through using larger batch reactors, but would require in-depth risk assessments, as they would involve the handling of large amounts of sodium hydride and the development of hydrogen gas. The membrane purification allows for straightforward scale-up, as the membranes can be implemented into spiral-wound modules, which offer high surface areas.

### **7.2.3 Higher $M_w$ PEGylation agents**

11 of the 40 FDA-approved PEGylated therapeutics are PEGylated with PEG of 5 kDa (or lower), the rest uses PEG with a  $M_w$  of 10-60 kDa. Up to 30 kDa the PEGylation agents are linear structures. For higher molecular weights (40 kDa and 60 kDa), two linear PEG molecules are attached to each other through a Lysine entity, resulting in a 2-armed branched structure.

Within this work the first defined  $M_w$  5 kDa PEGylation agents have been synthesised. To produce comparable analytical and pharmacological data and to meet the market needs, one focus of future work should be to produce defined  $M_w$  PEG with molecular weights of 10 kDa and higher. This could be achieved by continuing the synthesis from  $\text{Hub}(\text{Eg}_{112}\text{-OH})_3$  and iteratively extending the chains with an  $\text{Eg}_{28}$  BB, but 12 more extension cycles would be needed to reach the  $\text{Eg}_{448}$  (which is the equivalent to a 20 kDa PEG). Hence, it might be necessary to investigate the usage of longer BBs, like for example based on an  $\text{Eg}_{56}$  backbone.

Moving forward, the primary objective should be to synthesise 20 kDa defined  $M_w$  PEG. Once this is reached, it can easily be converted to an active ester, which then allows for linking two 20 kDa PEG molecules through a lysine entity, which would produce the branched, defined  $M_w$  40 kDa derivative.

#### 7.2.4 Defined $M_w$ therapeutics

Within this work the first defined  $M_w$  PEGylated protein was produced by PEGylating BSA with defined  $M_w$  PEG. The subsequent comparative study on the characterisation of the PEGylated protein and the identification of the PEGylated regions, offered first insights into the advantage of defined  $M_w$  PEGylation agents. Future work should include the synthesis of first defined  $M_w$  active PEGylated therapeutics and to obtain data on their biodistribution and pharmacological behaviour, compared to the marketed disperse derivatives. The blockbuster drug Neulasta (pegfilgrastim) is a potentially ideal candidate for this purpose, which is a protein PEGylated with a linear 20 kDa PEG. As the patent of this therapeutic has expired and the analytical data for its biosimilar MYL1401H is publicly available, selection of Neulasta as a model drug would allow for a comparative analysis, providing meaningful insights into the potential advantages of our defined  $M_w$  PEGylated therapeutic. The mPEG propionaldehyde used to PEGylate this therapeutic has been synthesised within this work as a 5 kDa derivative.

While data on the enhanced biodistribution behaviour of defined  $M_w$  PEGylated therapeutics is currently unavailable, existing data on the blood residence time of PEG of different  $M_w$  supports this theory. The blood-residence time of PEG increases with increasing  $M_w$  (Fig. 1 in [44]). The curve is especially steep below PEG weights of 50 kDa, which leads to the assumption that a small change in the commonly for PEGylation used  $M_w$  can result in a significantly different blood residence time. This implies that a defined PEGylated therapeutic may have a different volume of distribution compared to a disperse version, offering potential advantages in terms of enhanced biodistribution and, subsequently, therapeutic efficacy.

## Chapter 8 General Methods

**NMR.**  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra were recorded on a Bruker AVANCE III HD-400 spectrometer, with working frequencies of 400 ( $^1\text{H}$ ) and 101 ( $^{13}\text{C}$ ) MHz using deuterated chloroform ( $\text{CDCl}_3$ ) as a solvent at 293 K.

**UPLC-MS.** UPLC-MS analysis was performed on a Waters Acquity UPLC stack equipped with a PDA e $\lambda$  detector and a micromass ZQ mass spectrometer (operated in positive electrospray ionisation mode) or on an Agilent 1290 Infinity II stack equipped with a DAD Detector, a 1290 Infinity II ELSD detector, and a 6130 Quadrupole mass spectrometer (operated in positive electrospray ionisation mode). The columns used were either a Acquity UPLC BEH C18 1.7 $\mu\text{m}$ , 2.1 x 50 mm or a Acquity UPLC BEH C18 1.7 $\mu\text{m}$ , 2.1 x 150 mm. The mobile phases used were  $\text{H}_2\text{O}$  (with 50 mM  $\text{NH}_4\text{OAc}$ , **A**) and  $\text{MeCN}/\text{MeOH}$  (4:1, **B**) at a flow rate of 0.3 ml/min and a column temperature of 60 °C. The column was equilibrated with the initial solvent composition prior to injection (the injection volume was 4  $\mu\text{L}$ ).

The samples were run at different gradients, with an increasing fraction of the organic solvent **B** in the aqueous solvent **A**. The gradients are specified for the runs as for example “20 -95 % B, 8 min”, meaning that the solvent composition changed from 20 % B in A to 95 % and the total gradient time was 8 minutes.

**MALDI-TOF-MS.** MALDI-TOF-MS was performed on a Bruker autoflex MALDI-TOF spectrometer at 70 % laser power, using either Sinapinic acid (for samples containing protein species) or DCTB (for samples containing pure PEG) as the matrices. The typical sample preparation protocols are described in the following.

Protein-containing samples: 10  $\mu\text{L}$  of sample (derived from PEGylation mixture or after SEC) were purified using a C<sub>4</sub> ZipTip and eluted into 10  $\mu\text{L}$  of saturated Sinapinic Acid (in  $\text{MeCN}:\text{H}_2\text{O}$  30%:70% with 0.1% TFA). 0.8  $\mu\text{L}$  of the solution was spotted onto a MALDI plate and left to dry.

PEG samples: a 1 mg/ml solution of the sample in MeCN and a saturated solution of the matrix (DCTB) in 50 mM NaOAc in H<sub>2</sub>O/MeCN 1:1 were prepared. Both solutions were mixed in a 1:1 ratio, 0.8 µL of the mixture was spotted onto the plate and left to dry.

## References

- [1] Zalipsky, S. & Harris, J. M. in *Poly(ethylene glycol)* Vol. **680** *ACS Symposium Series* Ch. 1, 1-13 (American Chemical Society, 1997).  
<https://doi.org/10.1021/bk-1997-0680.ch001>.
- [2] Roberts, M. J., Bentley, M. D. & Harris, J. M. Chemistry for peptide and protein PEGylation. *Adv. Drug Del. Rev.* **54**, 459-476 (2002).  
[https://doi.org/10.1016/S0169-409X\(02\)00022-4](https://doi.org/10.1016/S0169-409X(02)00022-4).
- [3] Srichana, T. & Suwandecha, T. in *Biodegradable Polymers in Clinical Use and Clinical Development* Vol. **1** (eds Abraham J. Domb, Neeraj Kumar, & Aviva Ezra) Ch. 13, 451-493 (John Wiley & Sons, Inc., 2011).  
<https://doi.org/10.1002/9781118015810.ch13>.
- [4] Veronese, F. M. & Pasut, G. PEGylation, successful approach to drug delivery. *Drug Discov. Today* **10**, 1451-1458 (2005).  
[https://doi.org/10.1016/s1359-6446\(05\)03575-0](https://doi.org/10.1016/s1359-6446(05)03575-0).
- [5] Knop, K., Hoogenboom, R., Fischer, D. & Schubert, U. S. Poly(ethylene glycol) in drug delivery: pros and cons as well as potential alternatives. *Angew. Chem. Int. Ed. Engl.* **49**, 6288-6308 (2010).  
<https://doi.org/10.1002/anie.200902672>.
- [6] D'souza, A. A. & Shegokar, R. Polyethylene glycol (PEG): a versatile polymer for pharmaceutical applications. *Expert Opin. Drug Deliv.* **13**, 1257-1275 (2016).  
<https://doi.org/10.1080/17425247.2016.1182485>.
- [7] Herzberger, J., Niederer, K., Pohlitz, H., Seiwert, J., Worm, M., Wurm, F. R. & Frey, H. Polymerization of Ethylene Oxide, Propylene Oxide, and Other Alkylene Oxides: Synthesis, Novel Polymer Architectures, and Bioconjugation. *Chem. Rev.* **116**, 2170-2243 (2016).  
<https://doi.org/10.1021/acs.chemrev.5b00441>.
- [8] Veronese, F. M. Peptide and protein PEGylation: a review of problems and solutions. *Biomaterials* **22**, 405-417 (2001).  
[https://doi.org/10.1016/S0142-9612\(00\)00193-9](https://doi.org/10.1016/S0142-9612(00)00193-9).
- [9] Gupta, V., Bhavanasi, S., Quadir, M., Singh, K., Ghosh, G., Vasamreddy, K., Ghosh, A., Siahaan, T. J., Banerjee, S. & Banerjee, S. K. Protein PEGylation for cancer therapy: bench to bedside. *J. Cell Commun. Signal* **13**, 319-330 (2019).  
<https://doi.org/10.1007/s12079-018-0492-0>.
- [10] Park, E. J. & Na, D. H. Characterization of the Reversed-Phase Chromatographic Behavior of PEGylated Peptides Based on the Poly(ethylene glycol) Dispersity. *Anal. Chem.* **88**, 10848-10853 (2016).  
<https://doi.org/10.1021/acs.analchem.6b03577>.

- [11] Monkarsh, S. P., Ma, Y., Aglione, A., Bailon, P., Ciolek, D., Debarbieri, B., Graves, M. C., Hollfelder, K., Michel, H., Palleroni, A., Porter, J. E., Russoman, E., Roy, S. & Pan, Y.-C. E. Positional Isomers of Monopegylated Interferon  $\alpha$ -2a: Isolation, Characterization, and Biological Activity. *Anal. Biochem.* **247**, 434-440 (1997).  
<https://doi.org/10.1006/abio.1997.2128>.
- [12] Loiseau, F. A., Hii, K. K. & Hill, A. M. Multigram Synthesis of Well-Defined Extended Bifunctional Polyethylene Glycol (PEG) Chains. *J. Org. Chem.* **69**, 639-647 (2004).  
<https://doi.org/https://doi.org/10.1021/jo035042v>.
- [13] French, A. C., Thompson, A. L. & Davis, B. G. High-Purity Discrete PEG-Oligomer Crystals Allow Structural Insight. *Angew. Chem. Int. Ed.* **48**, 1248-1252 (2009).  
<https://doi.org/10.1002/anie.200804623>.
- [14] Maranski, K., Andreev, Y. G. & Bruce, P. G. Synthesis of Poly(ethylene oxide) Approaching Monodispersity. *Angew. Chem. Int. Ed.* **53**, 6411-6413 (2014).  
<https://doi.org/10.1002/anie.201403436>.
- [15] Szekely, G., Schaepertoens, M., Gaffney, P. R. & Livingston, A. G. Beyond PEG2000: synthesis and functionalisation of monodisperse PEGylated homostars and clickable bivalent polyethyleneglycols. *Chemistry* **20**, 10038-10051 (2014).  
<https://doi.org/10.1002/chem.201402186>.
- [16] Székely, G., Schaepertoens, M., Gaffney, P. R. J. & Livingston, A. G. Iterative synthesis of monodisperse PEG homostars and linear heterobifunctional PEG. *Polymer Chemistry* **5**, 694-697 (2014).  
<https://doi.org/10.1039/C3PY01367G>.
- [17] Khanal, A. & Fang, S. Solid Phase Stepwise Synthesis of Polyethylene Glycols. *Chemistry – A European Journal* **23**, 15133-15142 (2017).  
<https://doi.org/10.1002/chem.201703004>.
- [18] Eriyagama, D. N. A. M., Yin, Y. & Fang, S. Automated stepwise PEG synthesis using a base-labile protecting group. *Tetrahedron* **119**, 132861 (2022).  
<https://doi.org/10.1016/j.tet.2022.132861>.
- [19] Dong, R., Liu, R., Gaffney, P. R. J., Schaepertoens, M., Marchetti, P., Williams, C. M., Chen, R. & Livingston, A. G. Sequence-defined multifunctional polyethers via liquid-phase synthesis with molecular sieving. *Nat. Chem.*, (2018).  
<https://doi.org/10.1038/s41557-018-0169-6>.
- [20] Abuchowski, A., van Es, T., Palczuk, N. C. & Davis, F. F. Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol. *J. Biol. Chem.* **252**, 3578-3581 (1977).  
[https://doi.org/10.1016/S0021-9258\(17\)40291-2](https://doi.org/10.1016/S0021-9258(17)40291-2).
- [21] Harris, J. M. & Chess, R. B. Effect of pegylation on pharmaceuticals. *Nat. Rev. Drug Discov.* **2**, 214-221 (2003).  
<https://doi.org/10.1038/nrd1033>.

- [22] Veronese, F. M. Peptide and protein PEGylation: a review of problems and solutions. *Biomaterials* **22**, 405-417 (2001).  
[https://doi.org/10.1016/s0142-9612\(00\)00193-9](https://doi.org/10.1016/s0142-9612(00)00193-9).
- [23] Swierczewska, M., Lee, K. C. & Lee, S. What is the future of PEGylated therapies? *Exp. Opin. Emerg. Drugs* **20**, 531-536 (2015).  
<https://doi.org/10.1517/14728214.2015.1113254>.
- [24] L. Turecek, P., Bossard, M., Schoetens, F. & Ivens, I. PEGylation of Biopharmaceuticals: A Review of Chemistry and Nonclinical Safety Information of Approved Drugs. *J. Pharm. Sci.* **105**, 460-475 (2016).  
<https://doi.org/10.1016/j.xphs.2015.11.015>.
- [25] Harris, J. M. in *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications*, (ed J. Milton Harris) 1-14 (Springer US, 1992).  
[https://doi.org/10.1007/978-1-4899-0703-5\\_1](https://doi.org/10.1007/978-1-4899-0703-5_1).
- [26] Bailon, P. & Berthold, W. Polyethylene glycol-conjugated pharmaceutical proteins. *Pharm. Sci. Technol. today* **1**, 352-356 (1998).  
[https://doi.org/10.1016/S1461-5347\(98\)00086-8](https://doi.org/10.1016/S1461-5347(98)00086-8).
- [27] Molineux, G. The Design and Development of Pegfilgrastim (PEG-rmetHuG-CSF, Neulasta). *Curr. Pharm. Des.* **10**, 1235-1244 (2004).  
<https://doi.org/10.2174/1381612043452613>.
- [28] Schiller, L. J. in *Notice 84 FR 3795*, 3795-3796 (U.S. Food & Drug Administration, Department of Health and Human Services, 2019).  
<https://www.govinfo.gov/content/pkg/FR-2019-02-13/pdf/2019-02146.pdf>
- [29] Fee, C. J. & Van Alstine, J. M. PEG-proteins: Reaction engineering and separation issues. *Chem. Eng. Sci.* **61**, 924-939 (2006).  
<https://doi.org/10.1016/j.ces.2005.04.040>.
- [30] Dust, J. M., Fang, Z. H. & Harris, J. M. Proton NMR characterization of poly(ethylene glycols) and derivatives. *Macromolecules* **23**, 3742-3746 (1990).  
<https://doi.org/10.1021/ma00218a005>.
- [31] Sasaki, H., Ohtake, Y., Matsushima, A., Hiroto, M., Kodera, Y. & Inada, Y. Reduction of Immunoreactivity of Bovine Serum Albumin Conjugated with Comb-Shaped Polyethylene Glycol Derivatives. *Biochem. Biophys. Res. Commun.* **197**, 287-291 (1993).  
<https://doi.org/10.1006/bbrc.1993.2473>.
- [32] Hershfield, M. S., Buckley, R. H., Greenberg, M. L., Melton, A. L., Schiff, R., Hatem, C., Kurtzberg, J., Markert, M. L., Kobayashi, R. H., Kobayashi, A. L. & et al. Treatment of adenosine deaminase deficiency with polyethylene glycol-modified adenosine deaminase. *N. Engl. J. Med.* **316**, 589-596 (1987).  
<https://doi.org/10.1056/nejm198703053161005>.

- [33] Abuchowski, A., Kazo, G. M., Verhoest, C. R., Jr., Van Es, T., Kafkewitz, D., Nucci, M. L., Viau, A. T. & Davis, F. F. Cancer therapy with chemically modified enzymes. I. Antitumor properties of polyethylene glycol-asparaginase conjugates. *Cancer Biochem. Bioph.* **7**, 175-186 (1984).
- [34] Wang, Y. S., Youngster, S., Grace, M., Bausch, J., Bordens, R. & Wyss, D. F. Structural and biological characterization of pegylated recombinant interferon alpha-2b and its therapeutic implications. *Adv Drug Deliv Rev* **54**, 547-570 (2002).  
[https://doi.org/10.1016/S0169-409X\(02\)00027-3](https://doi.org/10.1016/S0169-409X(02)00027-3).
- [35] Bailon, P., Palleroni, A., Schaffer, C. A., Spence, C. L., Fung, W.-J., Porter, J. E., Ehrlich, G. K., Pan, W., Xu, Z.-X., Modi, M. W., Farid, A., Berthold, W. & Graves, M. Rational Design of a Potent, Long-Lasting Form of Interferon: A 40 kDa Branched Polyethylene Glycol-Conjugated Interferon  $\alpha$ -2a for the Treatment of Hepatitis C. *Bioconjugate Chem.* **12**, 195-202 (2001).  
<https://doi.org/10.1021/bc000082g>.
- [36] Monfardini, C., Schiavon, O., Caliceti, P., Morpurgo, M., Harris, J. M. & Veronese, F. M. A Branched Monomethoxypoly(ethylene glycol) for Protein Modification. *Bioconjugate Chem.* **6**, 62-69 (1995).  
<https://doi.org/10.1021/bc00031a006>.
- [37] Wong, S. S. & Jameson, D. M. *Chemistry of Protein and Nucleic Acid Cross-Linking and Conjugation*. 2nd edn, (CRC Press 2012).  
<https://doi.org/10.1201/b11175>.
- [38] Shah, A., Coyle, T., Lalezari, S., Fischer, K., Kohlstaedde, B., Delesen, H., Radke, S. & Michaels, L. A. BAY 94-9027, a PEGylated recombinant factor VIII, exhibits a prolonged half-life and higher area under the curve in patients with severe haemophilia A: Comprehensive pharmacokinetic assessment from clinical studies. *Haemophilia* **24**, 733-740 (2018).  
<https://doi.org/10.1111/hae.13561>.
- [39] Guan, L., Yang, H., Cai, Y., Sun, L., Di, P., Li, W., Liu, G. & Tang, Y. ADMET-score – a comprehensive scoring function for evaluation of chemical drug-likeness. *Med. Chem. Comm.* **10**, 148-157 (2019).  
<https://doi.org/10.1039/C8MD00472B>.
- [40] Negus, S. S. & Banks, M. L. Pharmacokinetic-Pharmacodynamic (PKPD) Analysis with Drug Discrimination. *Curr. Top. Behav. Neurosci.* **39**, 245-259 (2018).  
[https://doi.org/10.1007/7854\\_2016\\_36](https://doi.org/10.1007/7854_2016_36).
- [41] Fishburn, C. S. The Pharmacology of PEGylation: Balancing PD with PK to Generate Novel Therapeutics. *J. Pharm. Sci.* **97**, 4167-4183 (2008).  
<https://doi.org/10.1002/jps.21278>.
- [42] Yamaoka, T., Tabata, Y. & Ikada, Y. Distribution and Tissue Uptake of Poly(ethylene glycol) with Different Molecular Weights after Intravenous Administration to Mice. *J. Pharm. Sci.* **83**, 601-606 (1994).  
<https://doi.org/10.1002/jps.2600830432>.

- [43] Veronese, F. M., Caliceti, P. & Schiavon, O. Branched and Linear Poly(Ethylene Glycol): Influence of the Polymer Structure on Enzymological, Pharmacokinetic, and Immunological Properties of Protein Conjugates. *J. Bioact. Compatible Polym.* **12**, 196-207 (1997).  
<https://doi.org/10.1177/088391159701200303>.
- [44] Caliceti, P. & Veronese, F. M. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)–protein conjugates. *Adv. Drug Del. Rev.* **55**, 1261-1277 (2003).  
[https://doi.org/10.1016/S0169-409X\(03\)00108-X](https://doi.org/10.1016/S0169-409X(03)00108-X).
- [45] Fee, C. J. Size comparison between proteins PEGylated with branched and linear poly(ethylene glycol) molecules. *Biotechnol. Bioeng.* **98**, 725-731 (2007).  
<https://doi.org/10.1002/bit.21482>.
- [46] Hoang Thi, T. T., Pilkington, E. H., Nguyen, D. H., Lee, J. S., Park, K. D. & Truong, N. P. The Importance of Poly(ethylene glycol) Alternatives for Overcoming PEG Immunogenicity in Drug Delivery and Bioconjugation. *Polymers (Basel)* **12** (2020).  
<https://doi.org/10.3390/polym12020298>.
- [47] Kong, Y. W. & Dreaden, E. C. PEG: Will It Come Back to You? Polyethelyne Glycol Immunogenicity, COVID Vaccines, and the Case for New PEG Derivatives and Alternatives. *Front. Bioeng. Biotechnol.* **10** (2022).  
<https://doi.org/10.3389/fbioe.2022.879988>.
- [48] Yu, X., Li, H., Dong, C., Qi, S., Yang, K., Bai, B., Peng, K., Buljan, M., Lin, X., Liu, Z. & Yu, G. Poly(ethyl ethylene phosphate): Overcoming the "Polyethylene Glycol Dilemma" for Cancer Immunotherapy and mRNA Vaccination. *ACS Nano* **17**, 23814-23828 (2023).  
<https://doi.org/10.1021/acsnano.3c07932>.
- [49] Chen, B.-M., Cheng, T.-L. & Roffler, S. R. Polyethylene Glycol Immunogenicity: Theoretical, Clinical, and Practical Aspects of Anti-Polyethylene Glycol Antibodies. *ACS Nano* **15**, 14022-14048 (2021).  
<https://doi.org/10.1021/acsnano.1c05922>.
- [50] Working, P. K., Newman, M. S., Johnson, J. & Cornacoff, J. B. in *Poly(ethylene glycol)* Vol. **680 ACS Symposium Series** Ch. 4, 45-57 (American Chemical Society, 1997).  
<https://doi.org/10.1021/bk-1997-0680.ch004>.
- [51] Hoang Thi, T. T., Pilkington, E. H., Nguyen, D. H., Lee, J. S., Park, K. D. & Truong, N. P. The Importance of Poly(ethylene glycol) Alternatives for Overcoming PEG Immunogenicity in Drug Delivery and Bioconjugation. *Polymers* **12**, 298 (2020).  
<https://doi.org/10.3390/polym12020298>.
- [52] Wang, J., Deng, T., Liu, Y., Chen, K., Yang, Z. & Jiang, Z.-X. Monodisperse and Polydisperse PEGylation of Peptides and Proteins: A Comparative Study. *Biomacromolecules* **21**, 3134-3139 (2020).  
<https://doi.org/10.1021/acs.biomac.0c00517>.

- [53] Richter, A. W. & Åkerblom, E. Antibodies against Polyethylene Glycol Produced in Animals by Immunization with Monomethoxy Polyethylene Glycol Modified Proteins. *International Archives of Allergy and Applied Immunology* **70**, 124-131 (1983).  
<https://doi.org/10.1159/000233309>.
- [54] Moreno, A., Pitoc, G. A., Ganson, N. J., Layzer, J. M., Hershfield, M. S., Tarantal, A. F. & Sullenger, B. A. Anti-PEG Antibodies Inhibit the Anticoagulant Activity of PEGylated Aptamers. *Cell Chem. Biol.* **26**, 634-644.e633 (2019).  
<https://doi.org/10.1016/j.chembiol.2019.02.001>.
- [55] Nichols, S. Product Quality Review(s) 761075Orig1s000. 92-153 (U.S. Food & Drug Administration, Center for Drug Evaluation and Research, Department of Health and Human Services, 2018).  
[https://www.accessdata.fda.gov/drugsatfda\\_docs/nda/2018/761075Orig1s000ChemR.pdf](https://www.accessdata.fda.gov/drugsatfda_docs/nda/2018/761075Orig1s000ChemR.pdf)
- [56] Gerislioglu, S., Adams, S. R. & Wesdemiotis, C. Characterization of singly and multiply PEGylated insulin isomers by reversed-phase ultra-performance liquid chromatography interfaced with ion mobility mass spectrometry. *Anal. Chim. Acta* **1004**, 58-66 (2018).  
<https://doi.org/10.1016/j.aca.2017.12.009>.
- [57] Monkarsh, S. P., Ma, Y., Aglione, A., Bailon, P., Ciolek, D., DeBarbieri, B., Graves, M. C., Hollfelder, K., Michel, H., Palleroni, A., Porter, J. E., Russoman, E., Roy, S. & Pan, Y. C. Positional isomers of monopegylated interferon alpha-2a: isolation, characterization, and biological activity. *Anal. Biochem.* **247**, 434-440 (1997).  
<https://doi.org/10.1006/abio.1997.2128>.
- [58] Fee, C. J. & Van Alstine, J. M. in *Asian Pacific Confederation of Chemical Engineering congress*, 355-355 (2004).
- [59] Molek, J. R. & Zydney, A. L. Ultrafiltration characteristics of pegylated proteins. *Biotechnol. Bioeng.* **95**, 474-482 (2006).  
<https://doi.org/10.1002/bit.21020>.
- [60] Mayolo-Deloya, K., González-Valdez, J., Guajardo-Flores, D., Aguilar, O., Benavides, J., & Rito-Palomares, M. Current advances in the non-chromatographic fractionation and characterization of PEGylated proteins. *J. Chem. Technol. Biotechnol.* **86**, 18-25 (2010).  
<https://doi.org/https://doi.org/10.1002/JCTB.2498>.
- [61] Fee, C. J. & Van Alstine, J. M. in *Protein Purification*, 339-362 (2011).  
<https://doi.org/https://doi.org/10.1002/9780470939932.ch14>.
- [62] Hagel, L. in *Principles, High Resolution Methods, an Application*, (ed J.-C. Janson and L. Ryden) 63-106 (VCH Publishers, 1989).
- [63] Kuga, S. Pore size distribution analysis of gel substances by size exclusion chromatography. *J. Chromatogr. A* **206**, 449-461 (1981).  
[https://doi.org/10.1016/S0021-9673\(00\)88914-1](https://doi.org/10.1016/S0021-9673(00)88914-1).

- [64] Santos, J. H. P. M., Torres-Obreque, K. M., Meneguetti, G. P., Amaro, B. P. & Rangel-Yagui, C. O. Protein PEGylation for the design of biobetters: from reaction to purification processes. *Braz. J. Pharm. Sci.* **54** (2018).  
<https://doi.org/10.1590/s2175-97902018000001009>.
- [65] Fee, C. J. & Van Alstine, J. M. Prediction of the Viscosity Radius and the Size Exclusion Chromatography Behavior of PEGylated Proteins. *Bioconjugate Chem.* **15**, 1304-1313 (2004).  
<https://doi.org/10.1021/bc049843w>.
- [66] Azarkan, M., Maes, D., Bouckaert, J., Thi, M.-H. D., Wyns, L. & Looze, Y. Thiol pegylation facilitates purification of chymopapain leading to diffraction studies at 1.4 Å resolution. *J. Chromatogr. A* **749**, 69-72 (1996).  
[https://doi.org/10.1016/0021-9673\(96\)00360-3](https://doi.org/10.1016/0021-9673(96)00360-3).
- [67] Seely, J. E. & Richey, C. W. Use of ion-exchange chromatography and hydrophobic interaction chromatography in the preparation and recovery of polyethylene glycol-linked proteins. *J. Chromatogr. A* **908**, 235-241 (2001).  
[https://doi.org/10.1016/S0021-9673\(00\)00739-1](https://doi.org/10.1016/S0021-9673(00)00739-1).
- [68] Gaberc-Porekar, V., Zore, I., Podobnik, B. & Menart, V. Obstacles and pitfalls in the PEGylation of therapeutic proteins. *Curr. Opin. Drug Discov. Deliv.* **11**, 242-250 (2008).
- [69] Fee, C. J. in *PEGylated Protein Drugs: Basic Science and Clinical Applications*, (ed Francesco M. Veronese) 113-125 (Birkhäuser Basel, 2009).  
[https://doi.org/10.1007/978-3-7643-8679-5\\_7](https://doi.org/10.1007/978-3-7643-8679-5_7).
- [70] Molek, J. R. & Zydney, A. L. Separation of PEGylated  $\alpha$ -Lactalbumin from Unreacted Precursors and Byproducts Using Ultrafiltration. *Biotechnol. Progr.* **23**, 1417-1424 (2007).  
<https://doi.org/10.1021/bp070243w>.
- [71] Latulippe, D. R., Molek, J. R. & Zydney, A. L. Importance of Biopolymer Molecular Flexibility in Ultrafiltration Processes. *Ind. Eng. Chem. Res.* **48**, 2395-2403 (2009).  
<https://doi.org/10.1021/ie8005337>.
- [72] Asenjo, J. A. & Andrews, B. A. Aqueous two-phase systems for protein separation: A perspective. *J. Chromatogr. A* **1218**, 8826-8835 (2011).  
<https://doi.org/10.1016/j.chroma.2011.06.051>.
- [73] Delgado, C., Malik, F., Selisko, B., Fischer, D. & Francis, G. E. Quantitative analysis of polyethylene glycol (PEG) in PEG-modified proteins/cytokines by aqueous two-phase systems. *J. Biochem. Biophys. Methods* **29**, 237-250 (1994).  
[https://doi.org/10.1016/0165-022X\(94\)90035-3](https://doi.org/10.1016/0165-022X(94)90035-3).
- [74] Clark, R., Olson, K., Fuh, G., Marian, M., Mortensen, D., Teshima, G., Chang, S., Chu, H., Mukku, V., Canova-Davis, E., Somers, T., Cronin, M., Winkler, M. & Wells, J. A. Long-acting growth hormones produced by conjugation with polyethylene glycol. *J. Biol. Chem.* **271**, 21969-21977 (1996).  
<https://doi.org/10.1074/jbc.271.36.21969>.

- [75] Kinstler, O. B., Brems, D. N., Lauren, S. L., Paige, A. G., Hamburger, J. B. & Treuheit, M. J. Characterization and Stability of N-terminally PEGylated rhG-CSF. *Pharm. Res.* **13**, 996-1002 (1996).  
<https://doi.org/10.1023/A:1016042220817>.
- [76] Yoo, C., Suckau, D., Sauerland, V., Ronk, M. & Ma, M. Toward top-down determination of PEGylation site using MALDI in-source decay MS analysis. *J. Am. Soc. Mass. Spectrom.* **20**, 326-333 (2009).  
<https://doi.org/10.1016/j.jasms.2008.10.013>.
- [77] Lu, X., Gough, P. C., DeFelippis, M. R. & Huang, L. Elucidation of PEGylation site with a combined approach of in-source fragmentation and CID MS/MS. *J. Am. Soc. Mass. Spectrom.* **21**, 810-818 (2010).  
<https://doi.org/10.1016/j.jasms.2010.01.011>.
- [78] Mahou, R. & Wandrey, C. Versatile Route to Synthesize Heterobifunctional Poly(ethylene glycol) of Variable Functionality for Subsequent Pegylation. *Polymers* **4**, 561-589 (2012).  
<https://doi.org/10.3390/polym4010561>.
- [79] Harris, J. M. & Kozlowski, A. Poly(ethylene glycol) and related polymers monosubstituted with propionic or butanoic acids and functional derivatives thereof for biotechnical applications. 5,672,662 (U.S. Patent and Trademark Office, 1995).  
<https://patents.google.com/patent/US5672662A/en>
- [80] Zhang, J., Zhao, Y.-J., Su, Z.-G. & Ma, G.-H. Synthesis of monomethoxy poly(ethylene glycol) without diol poly(ethylene glycol). *J. Appl. Polym. Sci.* **105**, 3782-3786 (2007).  
<https://doi.org/10.1002/app.26457>.
- [81] Zalipsky, S., Gilon, C. & Zilkha, A. Attachment of drugs to polyethylene glycols. *Eur. Polym. J.* **19**, 1177-1183 (1983).  
[https://doi.org/10.1016/0014-3057\(83\)90016-2](https://doi.org/10.1016/0014-3057(83)90016-2).
- [82] Bückmann, A. F., Morr, M. & Johansson, G. Functionalization of poly(ethylene glycol) and monomethoxy-poly(ethylene glycol). *Makromol. Chem.* **182**, 1379-1384 (1981).  
<https://doi.org/10.1002/macp.1981.021820509>.
- [83] Vojkovsky, T., Sullivan, B. & Sill, K. N. Synthesis of heterobifunctional polyethylene glycols: Polymerization from functional initiators. *Polymer* **105**, 72-78 (2016).  
<https://doi.org/10.1016/j.polymer.2016.10.015>.
- [84] Li, J. & Kao, W. J. Synthesis of Polyethylene Glycol (PEG) Derivatives and PEGylated–Peptide Biopolymer Conjugates. *Biomacromolecules* **4**, 1055-1067 (2003).  
<https://doi.org/10.1021/bm034069l>.
- [85] Moghaddam, A. J. A. b. l. Use of polyethylene glycol polymers for bioconjugations and drug development. *Am. Biotech. Lab.* **19**, 42-44 (2001).
- [86] Agner, E. Method for Displacement Chromatography. US 6,576,134 B1,(United States Patent and Trademark Office, 1999).

- [87] Agner, E. Purification of Peptides and Oligonucleotides by Sample Displacement Chromatography process and apparatus. US 6,245,238 B1,(Unites States Patent and Trademark Office, 1999).  
<https://patents.google.com/patent/US6245238B1/en>
- [88] Srajer Gajdosik, M., Clifton, J. & Josic, D. Sample displacement chromatography as a method for purification of proteins and peptides from complex mixtures. *J. Chromatogr. A* **1239**, 1-9 (2012).  
<https://doi.org/10.1016/j.chroma.2012.03.046>.
- [89] Veeraragavan, K., Bernier, A. & Braendli, E. Sample displacement mode chromatography: purification of proteins by use of a high-performance anion-exchange column. *J. Chromatogr. A* **541**, 207-220 (1991).  
[https://doi.org/10.1016/S0021-9673\(01\)95993-X](https://doi.org/10.1016/S0021-9673(01)95993-X).
- [90] *Monodispersity & PDI* (2023).  
<https://polypure.com/sites/default/files/Monodispersity%20%26%20PDI.pdf> (accessed December 26, 2023).
- [91] *PEG-28* (2023). <https://polypure.com/products/peg/peg-28> (accessed
- [92] Biopharma PEG. *Featured Monodispersed PEGs*  
<https://www.biochempeg.com/product/541/> (accessed December 26, 2023).
- [93] Gilbert, R., Hess, M., Jenkins, A., Jones, R., Kratochvil, P. & Stepto, R. Dispersity in polymer science. *Pure Appl. Chem.* **81**, 351-353 (2009).  
<https://doi.org/10.1351/PAC-REC-08-05-02>.
- [94] Rogošić, M., Mencer, H. J. & Gomzi, Z. Polydispersity index and molecular weight distributions of polymers. *Eur. Polym. J.* **32**, 1337-1344 (1996).  
[https://doi.org/10.1016/S0014-3057\(96\)00091-2](https://doi.org/10.1016/S0014-3057(96)00091-2).
- [95] Rane, S. S. & Choi, P. Polydispersity Index: How Accurately Does It Measure the Breadth of the Molecular Weight Distribution? *Chem. Mater.* **17**, 926-926 (2005).  
<https://doi.org/10.1021/cm048594i>.
- [96] Agilent Technologies. *An Introduction to Gel Permeation Chromatography and Size Exclusion Chromatography* <https://www.agilent.com/cs/library/primers/Public/5990-6969EN%20GPC%20SEC%20Chrom%20Guide.pdf> (accessed December 26, 2023).
- [97] Bouvier, E. S. P. & Koza, S. M. Advances in size-exclusion separations of proteins and polymers by UHPLC. *Trends in Analytical Chemistry* **63**, 85-94 (2014).  
<https://doi.org/10.1016/j.trac.2014.08.002>.
- [98] Davenport, A. P., Scully, C. C. G., de Graaf, C., Brown, A. J. H. & Maguire, J. J. Advances in therapeutic peptides targeting G protein-coupled receptors. *Nat. Rev. Drug Discov.* **19**, 389-413 (2020).

<https://doi.org/10.1038/s41573-020-0062-z>.

- [99] Lutz, J.-F. Defining the Field of Sequence-Controlled Polymers. *Macromol. Rapid Commun.* **38**, 1700582 (2017).  
<https://doi.org/10.1002/marc.20170052>.
- [100] Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E. M., Sipos, B. & Birney, E. Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **494**, 77-80 (2013).  
<https://doi.org/10.1038/nature11875>.
- [101] Colquhoun, H. & Lutz, J.-F. Information-containing macromolecules. *Nat. Chem.* **6**, 455 (2014).  
<https://doi.org/10.1038/nchem.1958>.
- [102] Lutz, J.-F. Coding Macromolecules: Inputting Information in Polymers Using Monomer-Based Alphabets. *Macromolecules* **48**, 4759-4767 (2015).  
<https://doi.org/10.1021/acs.macromol.5b00890>.
- [103] Martens, S., Landuyt, A., Espeel, P., Devreese, B., Dawyndt, P. & Du Prez, F. Multifunctional sequence-defined macromolecules for chemical data storage. *Nature Comm.* **9**, 4451 (2018).  
<https://doi.org/10.1038/s41467-018-06926-3>.
- [104] Aksakal, R., Mertens, C., Soete, M., Badi, N. & Du Prez, F. Applications of Discrete Synthetic Macromolecules in Life and Materials Science: Recent and Future Trends. *Advanced Science* **8**, 2004038 (2021).  
<https://doi.org/10.1002/advs.202004038>.
- [105] Alm, E. W., Oerther, D. B., Larsen, N., Stahl, D. A. & Raskin, L. The oligonucleotide probe database. *Appl. Environ. Microbiol.* **62**, 3557-3559 (1996).  
<https://doi.org/10.1128/aem.62.10.3557-3559.1996>.
- [106] Dias, N. & Stein, C. A. Antisense Oligonucleotides: Basic Concepts and Mechanisms. *Mol. Cancer Ther.* **1**, 347-355 (2002).
- [107] Muttenthaler, M., King, G. F., Adams, D. J. & Alewood, P. F. Trends in peptide drug discovery. *Nat. Rev. Drug Discov.* **20**, 309-325 (2021).  
<https://doi.org/10.1038/s41573-020-00135-8>.
- [108] Lau, J. L. & Dunn, M. K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Biorg. Med. Chem.* **26**, 2700-2707 (2018).  
<https://doi.org/10.1016/j.bmc.2017.06.052>.
- [109] Lutz, J.-F., Lehn, J.-M., Meijer, E. W. & Matyjaszewski, K. From precision polymers to complex materials and systems. *Nature Reviews Materials* **1**, 16024 (2016).  
<https://doi.org/10.1038/natrevmats.2016.24>.

- [110] Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302 (2006).  
<https://doi.org/10.1038/nature04586>.
- [111] Dey, S., Fan, C., Gothelf, K. V., Li, J., Lin, C., Liu, L., Liu, N., Nijenhuis, M. A. D., Saccà, B., Simmel, F. C., Yan, H. & Zhan, P. DNA origami. *Nat. Rev. Methods Primers* **1**, 13 (2021).  
<https://doi.org/10.1038/s43586-020-00009-8>.
- [112] Hartmann, L. & Börner, H. G. Precision Polymers: Monodisperse, Monomer-Sequence-Defined Segments to Target Future Demands of Polymers in Medicine. *Adv. Mater.* **21**, 3425-3431 (2009).  
<https://doi.org/doi:10.1002/adma.200801884>.
- [113] Solleder, S. C., Schneider, R. V., Wetzel, K. S., Boukis, A. C. & Meier, M. A. R. Recent Progress in the Design of Monodisperse, Sequence-Defined Macromolecules. *Macromol. Rapid Commun.* **38**, 1600711 (2017).  
<https://doi.org/doi:10.1002/marc.201600711>.
- [114] Keegstra, E. M. D., Zwikker, J. W., Roest, M. R. & Jenneskens, L. W. A highly selective synthesis of monodisperse oligo(ethylene glycols). *J. Org. Chem.* **57**, 6678-6680 (1992).  
<https://doi.org/10.1021/jo00050a065>.
- [115] Ahmed, S. A. & Tanaka, M. Synthesis of Oligo(ethylene glycol) toward 44-mer. *J. Org. Chem.* **71**, 9884-9886 (2006).  
<https://doi.org/10.1021/jo0617464>.
- [116] Niculescu-Duvaz, D., Getaz, J. & Springer, C. J. Long Functionalized Poly(ethylene glycol)s of Defined Molecular Weight: Synthesis and Application in Solid-Phase Synthesis of Conjugates. *Bioconjugate Chem.* **19**, 973-981 (2008).  
<https://doi.org/10.1021/bc060242+>.
- [117] Li, Y., Guo, Q., Li, X., Zhang, H., Yu, F., Yu, W., Xia, G., Fu, M., Yang, Z. & Jiang, Z.-X. Fluorous synthesis of mono-dispersed poly(ethylene glycols). *Tetrahedron Lett.* **55**, 2110-2113 (2014).  
<https://doi.org/10.1016/j.tetlet.2014.02.047>.
- [118] Li, Y., Qiu, X. & Jiang, Z.-X. Macrocyclic Sulfates as Versatile Building Blocks in the Synthesis of Monodisperse Poly(ethylene glycol)s and Monofunctionalized Derivatives. *Org. Process Res. Dev.* **19**, 800-805 (2015).  
<https://doi.org/10.1021/acs.oprd.5b00142>.
- [119] Zhang, H., Li, X., Shi, Q., Li, Y., Xia, G., Chen, L., Yang, Z. & Jiang, Z. X. Highly Efficient Synthesis of Monodisperse Poly(ethylene glycols) and Derivatives through Macrocyclization of Oligo(ethylene glycols). *Angew. Chem. Int. Ed.* **54**, 3763-3767 (2015).  
<https://doi.org/10.1002/anie.201410309>.
- [120] Zhang, Q., Ren, H. & Baker, G. L. A practical and scalable process to selectively monofunctionalize water-soluble  $\alpha,\omega$ -diols. *Tetrahedron Lett.* **55**, 3384-3386 (2014).  
<https://doi.org/10.1016/j.tetlet.2014.04.066>.

- [121] Wawro, A., Muraoka, T. & Kinbara, K. Chromatography-free synthesis of monodisperse oligo(ethylene glycol) mono-p-toluenesulfonates and quantitative analysis of oligomer purity. *Polym. Chem.* **7** (2016).  
<https://doi.org/10.1039/C6PY00127K>.
- [122] Perry, S. Z. & Hibbert, H. Studies on reactions relating to carbohydrates and polysaccharides: XLVIII. Ethylene Oxide and related compounds: Synthesis of the polyethylene glycols. *Can. J. Res.* **14b**, 77-83 (1936).  
<https://doi.org/10.1139/cjr36b-011>.
- [123] Fordyce, R., Lovell, E. L. & Hibbert, H. Studies on Reactions Relating to Carbohydrates and Polysaccharides. LVI. The Synthesis of the Higher Polyoxyethylene Glycols. *J. Am. Chem. Soc.* **61**, 1905-1910 (1939).  
<https://doi.org/10.1021/ja01876a072>.
- [124] Harada, A., Li, J. & Kamachi, M. Preparation and Characterization of a Polyrotaxane Consisting of Monodisperse Poly(ethylene glycol) and .alpha.-Cyclodextrins. *J. Am. Chem. Soc.* **116**, 3192-3196 (1994).  
<https://doi.org/10.1021/ja00087a004>.
- [125] French, A. C., Thompson, A. L. & Davis, B. G. High-purity discrete PEG-oligomer crystals allow structural insight. *Angew Chem Int Ed Engl* **48**, 1248-1252 (2009).  
<https://doi.org/10.1002/anie.200804623>.
- [126] Bohn, P. & Meier, M. A. R. Uniform poly(ethylene glycol): a comparative study. *Polym. J.* **52**, 165-178 (2020).  
<https://doi.org/10.1038/s41428-019-0277-1>.
- [127] Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154 (1963).  
<https://doi.org/10.1021/ja00897a025>.
- [128] Merrifield, B. Solid phase synthesis. *Science* **232**, 341 (1986).  
<https://doi.org/10.1126/science.3961484>.
- [129] Gravert, D. J. & Janda, K. D. Organic Synthesis on Soluble Polymer Supports: Liquid-Phase Methodologies. *Chem. Rev.* **97**, 489-510 (1997).  
<https://doi.org/10.1021/cr960064l>.
- [130] Pon, R. T. in *Protocols for Oligonucleotides and Analogs: Synthesis and Properties*, (ed Sudhir Agrawal) 465-496 (Humana Press, 1993).  
<https://doi.org/10.1385/0-89603-281-7:465>.
- [131] Marchetti, P., Jimenez Solomon, M. F., Szekely, G. & Livingston, A. G. Molecular Separation with Organic Solvent Nanofiltration: A Critical Review. *Chem. Rev.* **114**, 10735-10806 (2014).  
<https://doi.org/10.1021/cr500006j>.
- [132] Bayer, E. & Mutter, M. Liquid Phase Synthesis of Peptides. *Nature* **237**, 512-513 (1972).

<https://doi.org/10.1038/237512a0>.

- [133] Lyddiatt, A. Process chromatography: current constraints and future options for the adsorptive recovery of bioproducts. *Curr. Opin. Biotechnol.* **13**, 95-103 (2002).  
[https://doi.org/10.1016/S0958-1669\(02\)00293-8](https://doi.org/10.1016/S0958-1669(02)00293-8).
- [134] Priske, M., Lazar, M., Schnitzer, C. & Baumgarten, G. Recent Applications of Organic Solvent Nanofiltration. *Chem. Ing. Tech.* **88**, 39-49 (2016).  
<https://doi.org/https://doi.org/10.1002/cite.201500084>.
- [135] Oxley, A., Gaffney, P. R. J., Kim, D., Marchetti, P. & Livingston, A. G. Graft modification of polybenzimidazole membranes for organic solvent ultrafiltration with scale up to spiral wound modules. *J. Membr. Sci.* **647**, 120199 (2022).  
<https://doi.org/10.1016/j.memsci.2021.120199>.
- [136] Marchetti, P., Peeva, L. & Livingston, A. The Selectivity Challenge in Organic Solvent Nanofiltration: Membrane and Process Solutions. *Annu. Rev. Chem Biomol.* **8**, 473-497 (2017).  
<https://doi.org/10.1146/annurev-chembioeng-060816-101325>.
- [137] Vandezande, P., Gevers, L. E. M. & Vankelecom, I. F. J. Solvent resistant nanofiltration: separating on a molecular level. *Chem. Soc. Rev.* **37**, 365-405 (2008).  
<https://doi.org/10.1039/B610848M>.
- [138] Kim, J. F., Freitas da Silva, A. M., Valtcheva, I. B. & Livingston, A. G. When the membrane is not enough: A simplified membrane cascade using Organic Solvent Nanofiltration (OSN). *Sep. Purif. Technol.* **116**, 277-286 (2013).  
<https://doi.org/10.1016/j.seppur.2013.05.050>.
- [139] Fujioka, T., Khan, S. J., McDonald, J. A. & Nghiem, L. D. Nanofiltration of trace organic chemicals: A comparison between ceramic and polymeric membranes. *Sep. Purif. Technol.* **136**, 258-264 (2014).  
<https://doi.org/10.1016/j.seppur.2014.08.039>.
- [140] Luque, S., Gómez, D. & Álvarez, J. R. in *Membrane Science and Technology* Vol. **13** 177-216 (Elsevier, 2008).  
[https://doi.org/10.1016/S0927-5193\(07\)13006-0](https://doi.org/10.1016/S0927-5193(07)13006-0).
- [141] Jiang, Z., Dong, R., Evans, A. M., Biere, N., Ebrahim, M. A., Li, S., Anselmetti, D., Dichtel, W. R. & Livingston, A. G. Aligned macrocycle pores in ultrathin films for accurate molecular sieving. *Nature* **609**, 58-64 (2022).  
<https://doi.org/10.1038/s41586-022-05032-1>.
- [142] Asadi Tashvigh, A. & Chung, T.-S. Robust polybenzimidazole (PBI) hollow fiber membranes for organic solvent nanofiltration. *J. Membr. Sci.* **572**, 580-587 (2019).  
<https://doi.org/10.1016/j.memsci.2018.11.048>.

- [143] Valtcheva, I. B., Kumbharkar, S. C., Kim, J. F., Bhole, Y. & Livingston, A. G. Beyond polyimide: Crosslinked polybenzimidazole membranes for organic solvent nanofiltration (OSN) in harsh environments. *J. Membr. Sci.* **457**, 62-72 (2014).  
<https://doi.org/10.1016/j.memsci.2013.12.069>.
- [144] See-Toh, Y. H., Ferreira, F. C. & Livingston, A. G. The influence of membrane formation parameters on the functional performance of organic solvent nanofiltration membranes. *J. Membr. Sci.* **299**, 236-250 (2007).  
<https://doi.org/10.1016/j.memsci.2007.04.047>.
- [145] Oxley, A. & Livingston, A. G. Anti-fouling membranes for organic solvent nanofiltration (OSN) and organic solvent ultrafiltration (OSU): graft modified polybenzimidazole (PBI). *J. Membr. Sci.* **662**, 120977 (2022).  
<https://doi.org/10.1016/j.memsci.2022.120977>.
- [146] Hołda, A. K. & Vankelecom, I. F. J. Understanding and guiding the phase inversion process for synthesis of solvent resistant nanofiltration membranes. *J. Appl. Polym. Sci.* **132** (2015).  
<https://doi.org/10.1002/app.42130>.
- [147] Kim, J. F., Székely, G., Valtcheva, I. B. & Livingston, A. G. Increasing the sustainability of membrane processes through cascade approach and solvent recovery—pharmaceutical purification case study. *Green Chemistry* **16**, 133-145 (2014).  
<https://doi.org/10.1039/C3GC41402G>.
- [148] Yeo, J., Peeva, L., Chung, S., Gaffney, P., Kim, D., Luciani, C., Tsukanov, S., Seibert, K., Kopach, M., Albericio, F. & Livingston, A. Liquid Phase Peptide Synthesis via One-Pot Nanostar Sieving (PEPSTAR). *Angew. Chem. Int. Ed.* **60**, 7786-7795 (2021).  
<https://doi.org/10.1002/anie.202014445>.
- [149] Boden, N., Bushby, R. J., Clarkson, S., Evans, S. D., Knowles, P. F. & Marsh, A. The design and synthesis of simple molecular tethers for binding biomembranes to a gold surface. *Tetrahedron* **53**, 10939-10952 (1997).  
[https://doi.org/10.1016/S0040-4020\(97\)00698-4](https://doi.org/10.1016/S0040-4020(97)00698-4).
- [150] Kinbara, K. Monodisperse engineered PEGs for bio-related applications. *Polym. J.* **50**, 689-697 (2018).  
<https://doi.org/10.1038/s41428-018-0074-2>.
- [151] Hesek, D., Lee, M., Noll, B. C., Fisher, J. F. & Mobashery, S. Complications from Dual Roles of Sodium Hydride as a Base and as a Reducing Agent. *J. Org. Chem.* **74**, 2567-2570 (2009).  
<https://doi.org/10.1021/jo802706d>.
- [152] Luis, P. in *Fundamental Modelling of Membrane Systems*, (ed Patricia Luis) 1-23 (Elsevier, 2018).  
<https://doi.org/10.1016/B978-0-12-813483-2.00001-0>.

- [153] McDonogh, R. M., Bauser, H., Stroh, N. & Chmiel, H. Concentration polarisation and adsorption effects in cross-flow ultrafiltration of proteins. *Desalination* **79**, 217-231 (1990).  
[https://doi.org/10.1016/0011-9164\(90\)85007-W](https://doi.org/10.1016/0011-9164(90)85007-W).
- [154] Robnik, B., Naumoska, K. & Časar, Z. A Novel Testing Approach for Oxidative Degradation Dependent Incompatibility of Amine Moiety Containing Drugs with PEGs in Solid-State. *Pharmaceutics* **12**, 37 (2020).
- [155] Han, S., Kim, C. & Kwon, D. Thermal/oxidative degradation and stabilization of polyethylene glycol. *Polymer* **38**, 317-323 (1997).  
[https://doi.org/10.1016/S0032-3861\(97\)88175-X](https://doi.org/10.1016/S0032-3861(97)88175-X).
- [156] Babity, S. & Brambilla, D. Data on the removal of peroxides from functionalized polyethylene glycol (PEG) and effects on the stability and sensitivity of resulting PEGylated conjugates. *Data Brief* **32**, 106258 (2020).  
<https://doi.org/10.1016/j.dib.2020.106258>.
- [157] Hemenway, J. N., Carvalho, T. C., Rao, V. M., Wu, Y., Levons, J. K., Narang, A. S., Paruchuri, S. R., Stamato, H. J. & Varia, S. A. Formation of reactive impurities in aqueous and neat polyethylene glycol 400 and effects of antioxidants and oxidation inducers. *J. Pharm. Sci.* **101**, 3305-3318 (2012).  
<https://doi.org/10.1002/jps.23198>.
- [158] Hecht, E. S., Yeh, G. K. & Zhang, K. Evidence of free radical generation from the interaction of polyethylene glycol with PVC medical tubing. *J. Pharm. Biomed. Anal.* **197**, 113955 (2021).  
<https://doi.org/10.1016/j.jpba.2021.113955>.
- [159] Rahman, Z., Kohli, K., Khar, R. K., Ali, M., Charoo, N. A. & Shamsheer, A. A. Characterization of 5-fluorouracil microspheres for colonic delivery. *AAPS PharmSciTech* **7**, E47 (2006).  
<https://doi.org/10.1208/pt070247>.
- [160] Zhang, K., Pellett, J. D., Narang, A. S., Wang, Y. J. & Zhang, Y. T. Reactive impurities in large and small molecule pharmaceutical excipients – A review. *TrAC, Trends Anal. Chem.* **101**, 34-42 (2018).  
<https://doi.org/10.1016/j.trac.2017.11.003>.
- [161] Ciufolini, M. A. & Swaminathan, S. Synthesis of a model depsipeptide segment of Luzopeptins (BBM 928), potent antitumor and antiretroviral antibiotics. *Tetrahedron Lett.* **30**, 3027-3028 (1989).  
[https://doi.org/10.1016/S0040-4039\(00\)99393-6](https://doi.org/10.1016/S0040-4039(00)99393-6).
- [162] Roberts, C. Z., Yi, Beaumier, Francis; Lepissier, Luce. Compounds for the modulation of MYC activity. US patent WO2016196910A1, (World Intellectual Property Organization, 2016).
- [163] Topchiyeva, I. N. Synthesis of biologically active polyethylene glycol derivatives. A review. *Polymer Science U.S.S.R.* **32**, 833-851 (1990).

[https://doi.org/10.1016/0032-3950\(90\)90214-Q](https://doi.org/10.1016/0032-3950(90)90214-Q).

- [164] Hermanson, G. T. in *Bioconjugate Techniques* Vol. **2** Ch. 18, 711-742 (Elsevier, 2008).
- [165] Guastavino, J. F., Vaillard, V. A., Cristaldi, M. D., Rossini, L. & Vaillard, S. E. Simple Synthesis of Aldehyde and Carboxylic Acid Terminated Methoxypoly(ethylene glycol). *Macromol. Chem. Phys.* **217**, 1745-1751 (2016).  
<https://doi.org/10.1002/macp.201600094>.
- [166] Ferruti, P., Tanzi, M. C., Rusconi, L. & Cecchi, R. Succinic half-esters of poly(ethylene glycol)s and their benzotriazole and imidazole derivatives as oligomeric drug-binding matrices. *Makromol. Chem.* **182**, 2183-2192 (1981).  
<https://doi.org/10.1002/macp.1981.021820807>.
- [167] Lee, K. C., Moon, S. C., Park, M. O., Lee, J. T., Na, D. H., Yoo, S. D., Lee, H. S. & DeLuca, P. P. Isolation, characterization, and stability of positional isomers of mono-PEGylated salmon calcitonins. *Pharm. Res.* **16**, 813-818 (1999).  
<https://doi.org/10.1023/a:1018861616465>.
- [168] Wang, Y.-S., Youngster, S., Bausch, J., Zhang, R., McNemar, C. & Wyss, D. F. Identification of the Major Positional Isomer of Pegylated Interferon Alpha-2b. *Biochemistry* **39**, 10634-10640 (2000).  
<https://doi.org/10.1021/bi000617t>.
- [169] Wang, Y.-S., Youngster, S., Grace, M., Bausch, J., Bordens, R. & Wyss, D. F. Structural and biological characterization of pegylated recombinant interferon alpha-2b and its therapeutic implications. *Adv. Drug Del. Rev.* **54**, 547-570 (2002).  
[https://doi.org/10.1016/S0169-409X\(02\)00027-3](https://doi.org/10.1016/S0169-409X(02)00027-3).
- [170] Lee, H. & Park, T. G. Preparation and Characterization of Mono-PEGylated Epidermal Growth Factor: Evaluation of in Vitro Biologic Activity. *Pharm. Res.* **19**, 845-851 (2002).  
<https://doi.org/10.1023/A:1016113117851>.
- [171] Lee, H., Jang, I. H., Ryu, S. H. & Park, T. G. N-terminal site-specific mono-PEGylation of epidermal growth factor. *Pharm. Res.* **20**, 818-825 (2003).  
<https://doi.org/10.1023/a:1023402123119>.
- [172] Foser, S., Schacher, A., Weyer, K. A., Brugger, D., Dietel, E., Marti, S. & Schreitmüller, T. Isolation, structural characterization, and antiviral activity of positional isomers of monopegylated interferon alpha-2a (PEGASYS). *Protein Expr Purif* **30**, 78-87 (2003).  
[https://doi.org/10.1016/s1046-5928\(03\)00055-x](https://doi.org/10.1016/s1046-5928(03)00055-x).
- [173] Youn, Y. S., Na, D. H., Yoo, S. D., Song, S.-C. & Lee, K. C. Chromatographic separation and mass spectrometric identification of positional isomers of polyethylene glycol-modified growth hormone-releasing factor (1-29). *J. Chromatogr. A* **1061**, 45-49 (2004).  
<https://doi.org/10.1016/j.chroma.2004.10.062>.

- [174] Cai, Y. & Yue, P. Separation of exenatide analogue mono-PEGylated with 40 kDA polyethylene glycol by cation exchange chromatography. *J. Chromatogr. A* **1218**, 6953-6960 (2011).  
<https://doi.org/10.1016/j.chroma.2011.07.107>.
- [175] Qian, X., Dong, H., Tian, H., Tong, Y., Guo, L., Hu, X., Gao, X. & Yao, W. Characterization of a site-specific PEGylated analog of exendin-4 and determination of the PEGylation site. *Int. J. Pharm.* **454**, 553-558 (2013).  
<https://doi.org/10.1016/j.ijpharm.2013.06.059>.

## Appendix

### ***Additional Analytical Data***

#### ***UPLC-MS data for Hub(Eg<sub>28</sub>-OH)<sub>3</sub> (5a)***

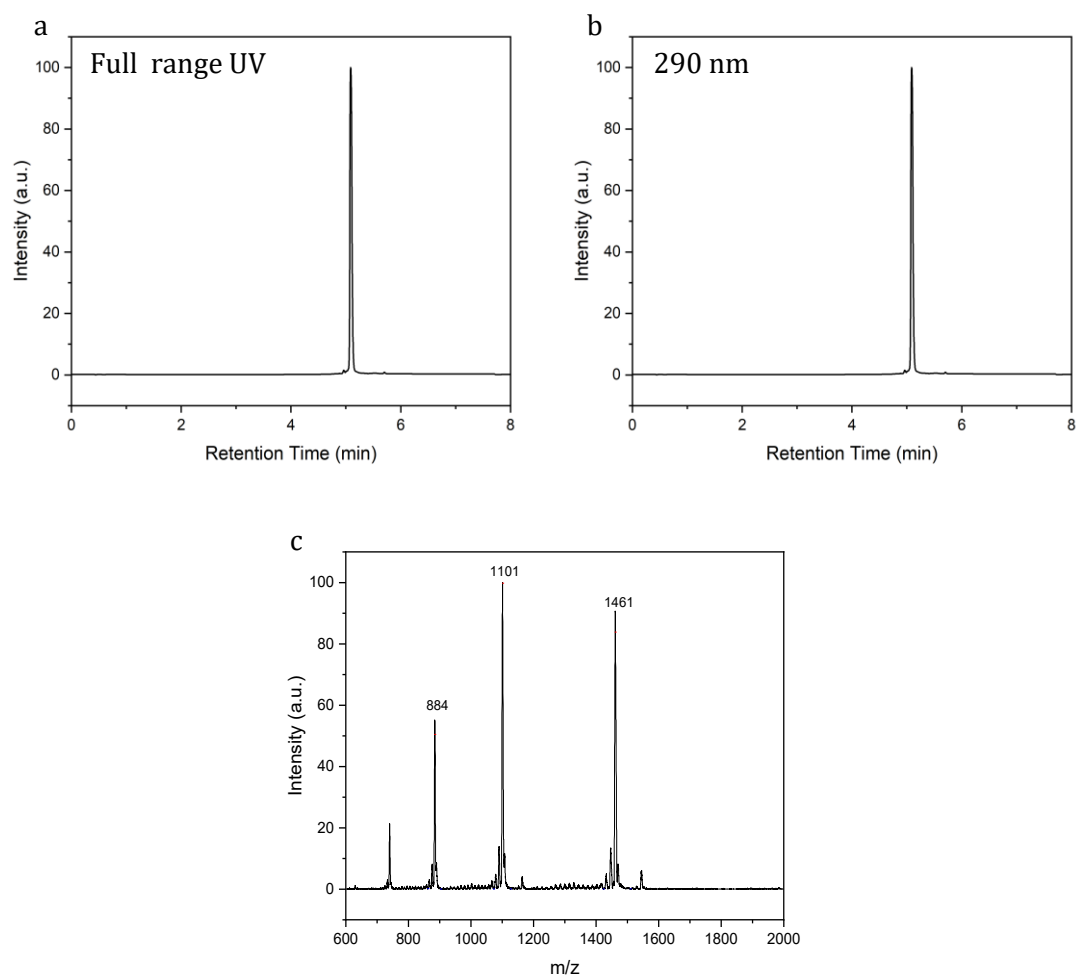


**Figure A 1. UPLC-MS data of synthesised Hub(Eg<sub>28</sub>OH)<sub>3</sub> (5a) derived from membrane-based synthesis.** **a**, Full range UV chromatogram (5 cm C<sub>18</sub> column, 8 min 20-95 % B in A, A: H<sub>2</sub>O + 5 mM NH<sub>4</sub>OAc, B MeCN: MeOH 4:1); the products retention time is 5.10; **b**, 290 nm UV chromatogram of same run; **c**, underlying ESI<sup>+</sup> spectrum of peak at 5.10 min. The *m/z* series 1461, 1101, 884 confirms the identity of Hub(Eg<sub>28</sub>-OH)<sub>3</sub> (*M<sub>w</sub>*= 4325.23 Da) with 1461 = [M+ 3 NH<sub>4</sub><sup>+</sup>]<sup>3+</sup>, 1101 = [M+ 4 NH<sub>4</sub><sup>+</sup>]<sup>4+</sup>, 884 = [M+ 5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>.

UPLC-MS data for *Hub(Eg<sub>56</sub>-OH)<sub>3</sub>* (**5b**)

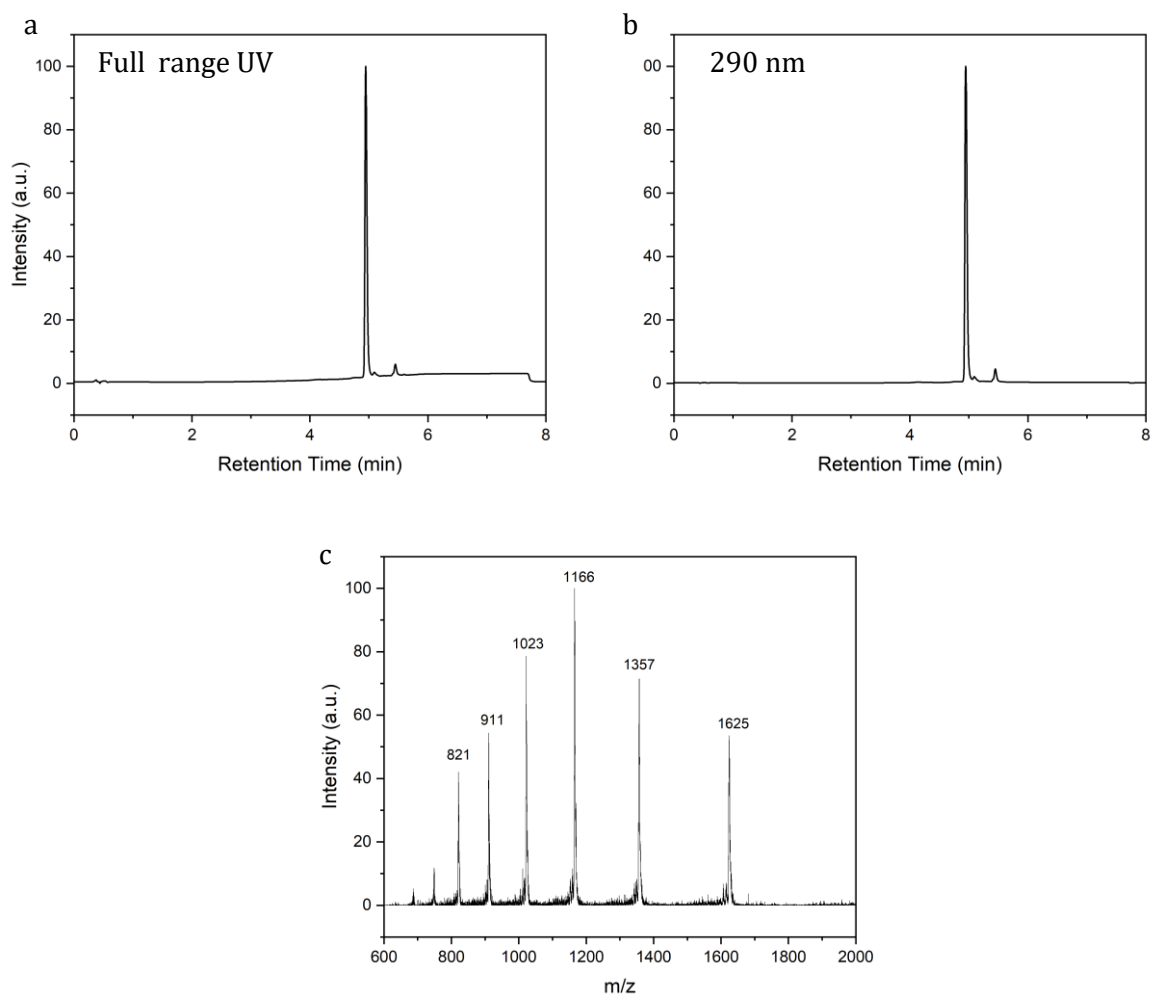


**Figure A 2. UPLC-MS data of synthesised *Hub(Eg<sub>56</sub>OH)<sub>3</sub>* (**5b**) derived from membrane-based synthesis. a**, Full-range UV trace chromatogram (5 cm C<sub>18</sub> column, 8 min 20-95 % B in A, A: H<sub>2</sub>O + 5 mM NH<sub>4</sub>OAc, B MeCN: MeOH 4:1); the products retention time is 4.90 min. **b**, 290 nm UV chromatogram of the same run; **c**, underlying ESI<sup>+</sup> spectrum of peak at 4.90 min. The *m/z* series 1625, 1357, 1166, 1023, 911, 821 confirms the identity of *Hub(Eg<sub>56</sub>-OH)<sub>3</sub>* (*M<sub>w</sub>*= 8025.68 Da) with 1625 = [M+ 5 NH<sub>4</sub><sup>+</sup>]<sup>5+</sup>, 1357 = [M+ 6 NH<sub>4</sub><sup>+</sup>]<sup>6+</sup>, 1166 = [M+ 7 NH<sub>4</sub><sup>+</sup>]<sup>7+</sup>, 1023 = [M+ 7 NH<sub>4</sub><sup>+</sup>]<sup>7+</sup>, 911 = [M+ 7 NH<sub>4</sub><sup>+</sup>]<sup>7+</sup>, 822 = [M+ 7 NH<sub>4</sub><sup>+</sup>]<sup>7+</sup>.

*UPLC-MS data for Hub(Eg<sub>84</sub>-OH)<sub>3</sub> (5c)*

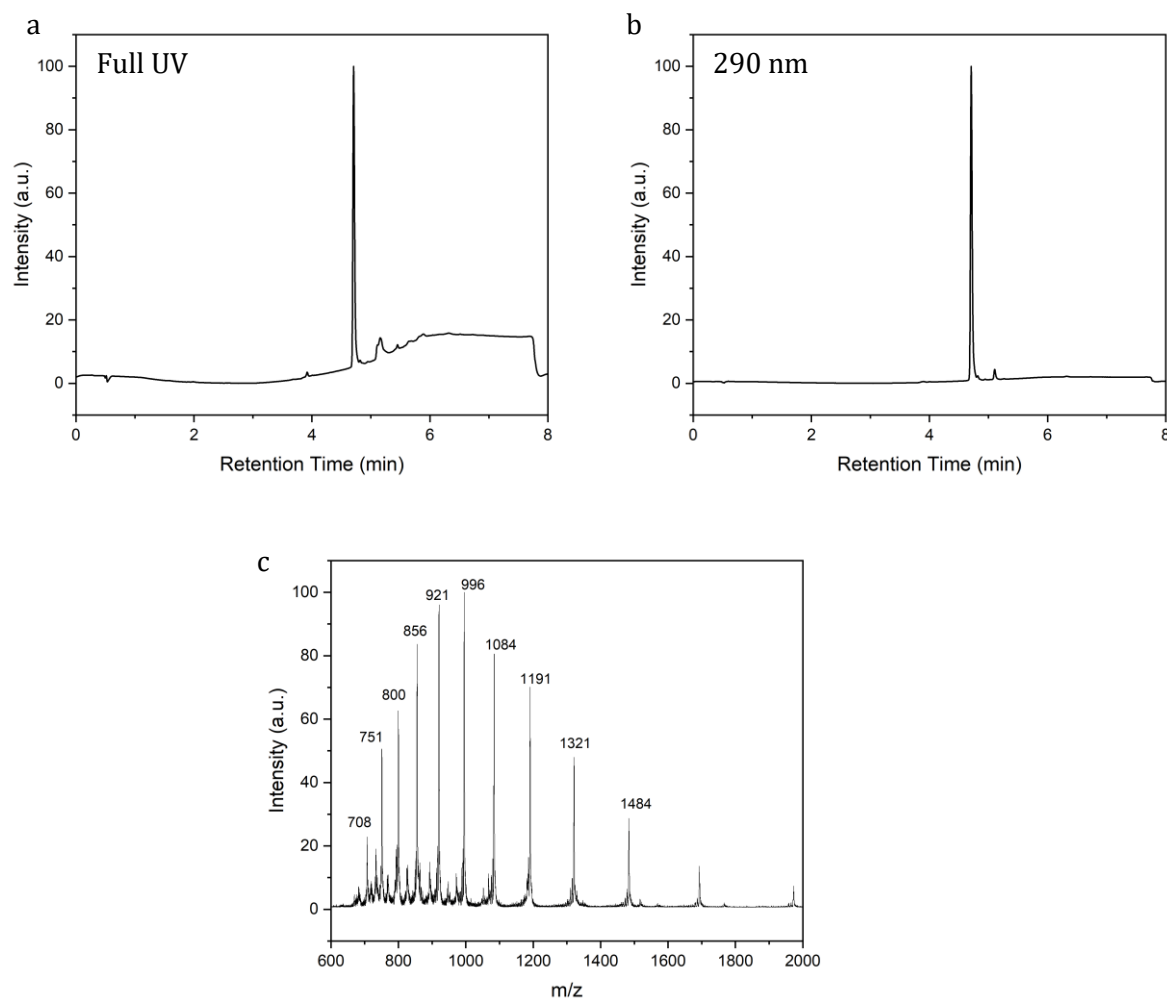


**Figure A 3. UPLC-MS data of synthesised Hub(Eg<sub>84</sub>OH)<sub>3</sub> derived from membrane-based synthesis (5c).** **a**, Full-range UV chromatogram (5 cm C<sub>18</sub> column, 8 min 20-95 % B in A, A: H<sub>2</sub>O + 5 mM NH<sub>4</sub>OAc, B MeCN: MeOH 4:1); the products retention time is 4.70 min. **b**, 290 nm UV chromatogram; **c**, underlying ESI<sup>+</sup> spectrum of peak at 4.10 min. The *m/z* series confirms the identity of Hub(Eg<sub>84</sub>-OH)<sub>3</sub> (*M<sub>w</sub>*= 11726.14 Da ) with 1484 = [M+ 8 NH<sub>4</sub><sup>+</sup>]<sup>8+</sup>, 1321 = [M+ 9 NH<sub>4</sub><sup>+</sup>]<sup>9+</sup>, 1191 = [M+ 10 NH<sub>4</sub><sup>+</sup>]<sup>10+</sup>, 1084 = [M+ 11 NH<sub>4</sub><sup>+</sup>]<sup>11+</sup>, 996 = [M+ 12 NH<sub>4</sub><sup>+</sup>]<sup>12+</sup>, 921 = [M+ 13 NH<sub>4</sub><sup>+</sup>]<sup>13+</sup>, 856 = [M+ 14 NH<sub>4</sub><sup>+</sup>]<sup>14+</sup>, 800 = [M+ 15 NH<sub>4</sub><sup>+</sup>]<sup>15+</sup>, 751 = [M+ 16 NH<sub>4</sub><sup>+</sup>]<sup>16+</sup>, 708 = [M+ 17 NH<sub>4</sub><sup>+</sup>]<sup>17+</sup>.

*UPLC-MS data for Hub(Eg<sub>112</sub>-OH)<sub>3</sub> (5d)*

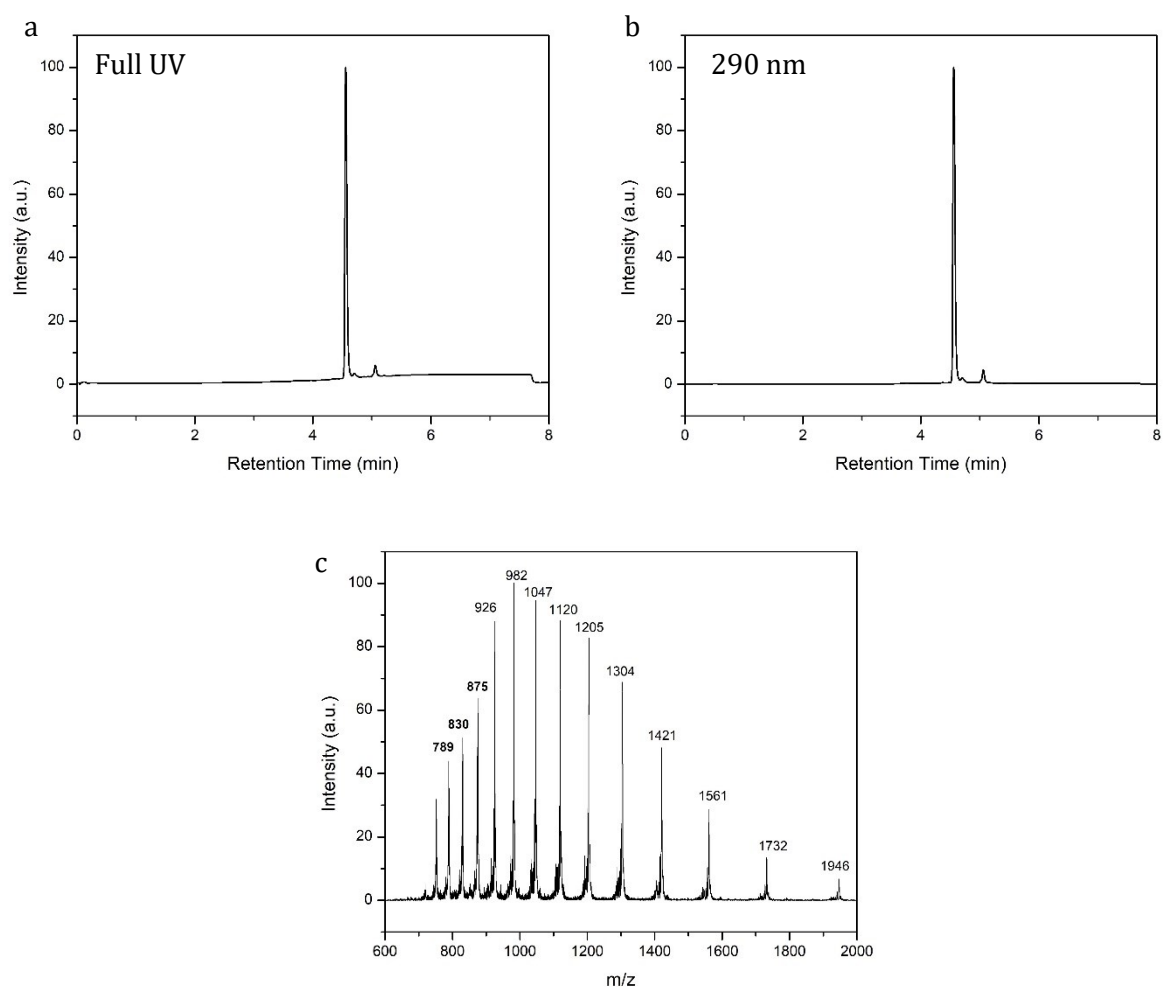


**Figure A 4. UPLC-MS data of synthesised Hub(Eg<sub>112</sub>OH)<sub>3</sub> derived from membrane-based synthesis (5d).** **a**, Full-range UV trace 5 cm C<sub>18</sub> column, 8 min 20-95 % B in A, A: H<sub>2</sub>O + 5 mM NH<sub>4</sub>OAc, B MeCN: MeOH 4:1); the products retention time is 4.30 min. **b**, 290 nm UV trace of the same run; **c**, underlying ESI<sup>+</sup> spectrum of peak at 4.10 min. The *m/z* series confirms the identity of Hub(Eg<sub>112</sub>-OH)<sub>3</sub> (*M<sub>w</sub>*= 15426.59 Da) with 1946 = [M+ 8 NH<sub>4</sub><sup>+</sup>]<sup>8+</sup>, 1732 = [M+ 9 NH<sub>4</sub><sup>+</sup>]<sup>9+</sup>, 1561 = [M+ 10 NH<sub>4</sub><sup>+</sup>]<sup>10+</sup>, 1421 = [M+ 11 NH<sub>4</sub><sup>+</sup>]<sup>11+</sup>, 1304 = [M+ 12 NH<sub>4</sub><sup>+</sup>]<sup>12+</sup>, 1205 = [M+ 13 NH<sub>4</sub><sup>+</sup>]<sup>13+</sup>, 1120 = [M+ 14 NH<sub>4</sub><sup>+</sup>]<sup>14+</sup>, 1047 = [M+ 15 NH<sub>4</sub><sup>+</sup>]<sup>15+</sup>, 982 = [M+ 16 NH<sub>4</sub><sup>+</sup>]<sup>16+</sup>, 926 = [M+ 17 NH<sub>4</sub><sup>+</sup>]<sup>17+</sup>, 875 = [M+ 18 NH<sub>4</sub><sup>+</sup>]<sup>18+</sup>, 830 = [M+ 19 NH<sub>4</sub><sup>+</sup>]<sup>19+</sup>, 789= [M+ 20 NH<sub>4</sub><sup>+</sup>]<sup>20+</sup>.

**Table A 1.** Overview of the literature on the separation and identification of positional isomers of PEGylated proteins.

| Publication                                 | Approach    | Protein type/size         | PEGylation technique (PEG size) | Number identified isomers (number PEG units) | PEGamer/ Pos. Isomer Separation | Identification PEGylation Site |
|---------------------------------------------|-------------|---------------------------|---------------------------------|----------------------------------------------|---------------------------------|--------------------------------|
| Clark <i>et al.</i> (1996) <sup>74</sup>    | Comparative | hGH (22 kDa)              | random (5 kDa)                  | 8 (n = 2-7)                                  | IEX                             | MALDI-TOF                      |
|                                             |             |                           |                                 |                                              |                                 | Tryptic Digestion              |
|                                             |             |                           |                                 |                                              |                                 | RP-HPLC-MS                     |
|                                             |             |                           |                                 |                                              |                                 | Endoproteinase peptide mapping |
| Kinstler <i>et al.</i> (1996) <sup>75</sup> | Sequencing  | rhG-CSF (18.8 kDa)        | selective (6 kDa)               | 1 (n = 1)                                    | IEX                             | RP-HPLC                        |
|                                             |             |                           |                                 |                                              |                                 | Edman sequencing               |
|                                             |             |                           |                                 |                                              |                                 | MALDI-MS                       |
|                                             |             |                           |                                 |                                              |                                 | Cyanogen bromide               |
| Monkarsh <i>et al.</i> (1997) <sup>11</sup> | Sequencing  | IFN $\alpha$ -2a (19 kDa) | random (5 kDa)                  | 11 (n = 1)                                   | IEX-type                        | Lys-C/ Glu-C digestion         |
|                                             |             |                           |                                 |                                              |                                 | RP-HPLC                        |
|                                             |             |                           |                                 |                                              |                                 | MALDI-TOF                      |
|                                             |             |                           |                                 |                                              |                                 | Edman sequencing               |
| Lee <i>et al.</i> (1999) <sup>167</sup>     | Comparative | sCT (3.4 kDa)             | random (5 kDa)                  | 3 (n = 1)                                    | SEC                             | MALDI-TOF                      |
|                                             |             |                           |                                 |                                              | RP-HPLC                         | Tryptic Digestion              |
|                                             |             |                           |                                 |                                              |                                 | RP-HPLC                        |

|                                   |                     |                           |                     |             |                |                       |
|-----------------------------------|---------------------|---------------------------|---------------------|-------------|----------------|-----------------------|
| Wang et al. (2000) <sup>168</sup> | NMR                 | IFN $\alpha$ -2b (19 kDa) | random (12 kDa)     | 1 (n = 1)   | IEX            | Subtilisin Digestion  |
|                                   |                     |                           |                     |             |                | SEC                   |
|                                   |                     |                           |                     |             |                | MALDI-MS              |
|                                   |                     |                           |                     |             |                | NMR                   |
| Wang et al. (2002) <sup>169</sup> | Sequencing +<br>NMR | IFN $\alpha$ -2b (19 kDa) | random              | 14 (n = 1)  | IEX            | MALDI-MS              |
|                                   |                     |                           |                     |             |                | Enzym. Digestion      |
|                                   |                     |                           |                     |             |                | N-terminal sequencing |
|                                   |                     |                           |                     |             |                | NMR                   |
| Lee et al. (2002) <sup>170</sup>  | Comparative         | rhEGF (6 kDa)             | random (3.4 kDa)    | 3 (n = 1)   | GPC<br>RP-HPLC | MALDI-TOF             |
|                                   |                     |                           |                     |             |                | Tryptic Digestion     |
|                                   |                     |                           |                     |             |                | RP-HPLC               |
|                                   |                     |                           |                     |             |                | SEC-HPLC              |
| Lee et al (2003) <sup>171</sup>   | Comparative         | rhEGF (6 kDa)             | selective (2/5 kDa) | 1 (n = 1)   | n.a.           | MALDI-TOF             |
|                                   |                     |                           |                     |             |                | Tryptic Digestion     |
|                                   |                     |                           |                     |             |                | RP-HPLC               |
|                                   |                     |                           |                     |             |                |                       |
| Foser et al (2003) <sup>172</sup> | Comparative         | IFN $\alpha$ -2a (19 kDa) | random (40 kDa)     | 9 (n = 1))  | IEX            | Lys-C Digestion       |
|                                   |                     |                           |                     |             |                | RP-HPLC               |
|                                   |                     |                           |                     |             |                | MALDI-TOF             |
|                                   |                     |                           |                     |             |                |                       |
| Youn et al (2004) <sup>173</sup>  | Comparative         | GRF (3.4 kDa)             | random (1 kDa)      | 7 (n = 1-3) | RP-HPLC        | Lys-C Digestion       |
|                                   |                     |                           |                     |             |                | MALDI-TOF             |

|                                         |             |                             |                    |           |                |                                                                  |
|-----------------------------------------|-------------|-----------------------------|--------------------|-----------|----------------|------------------------------------------------------------------|
| Yoo et al. (2008) <sup>76</sup>         | MS          | Synth. Peptide<br>(3.5 kDa) | selective (20 kDa) | 1 (n =1)  | RP-HPLC        | Reflector in-source decay (reISD)<br>MALDI-TOF                   |
| Lu et al. (2010) <sup>77</sup>          | MS          | Glucagon (3.5 kDa)<br>IgG4  | random (5/20 kDa)  | n.a.      | n.a.           | In-source fragmentation (ISF)<br>CID-MS-MS                       |
| Cai and Yue (2011) <sup>174</sup>       | Comparative | Hybr-003 (4.3 kDa)          | random (40 kDa)    | 4 (n = 1) | RP-HPLC<br>IEX | MALDI-TOF<br>Tryptic Digestion<br>RP-HPLC                        |
| Qian et al. (2013) <sup>175</sup>       | Sequencing  | Ex4C (4.3 kDa)              | selective (20 kDa) | n.a.      | RP-HPLC        | MALDI-TOF<br>Tryptic Digestion<br>RP-HPLC<br>Amino acid analysis |
| Gerislioglu et al. (2018) <sup>56</sup> | MS          | Insulin (5.7 kDa)           | random (5 kDa)     | n.a.      | n.a.           |                                                                  |

### Permission summary table for third party copyright works

| Figure Number | Source of Work                                                                                                                                                                                                                                                                                                                                  | Copyright Holder | Date of Permission Requested | Approved |
|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|------------------------------|----------|
| Figure 1      | Abuchowski, A., van Es, T., Palczuk, N. C. & Davis, F. F. Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol. <i>J. Biol. Chem.</i> <b>252</b> , 3578-3581 (1977).<br><br><a href="https://doi.org/10.1016/S0021-9258(17)40291-2">https://doi.org/10.1016/S0021-9258(17)40291-2</a> . | Elsevier         | 27/12/2023                   | Yes      |
| Figure 2      | L. Turecek, P., Bossard, M., Schoetens, F. & Ivens, I. PEGylation of Biopharmaceuticals: A Review of Chemistry and Nonclinical Safety Information of Approved Drugs. <i>J. Pharm. Sci.</i> <b>105</b> , 460-475 (2016).<br><br><a href="https://doi.org/10.1016/j.xphs.2015.11.015">https://doi.org/10.1016/j.xphs.2015.11.015</a> .            | Elsevier         | 27/12/2023                   | Yes      |
| Figure 4      | Monfardini, C., Schiavon, O., Caliceti, P., Morpurgo, M., Harris, J. M. & Veronese, F. M. A Branched Monomethoxypoly(ethylene glycol) for Protein Modification. <i>Bioconjugate Chem.</i> <b>6</b> , 62-69 (1995).<br><br><a href="https://doi.org/10.1021/bc00031a006">https://doi.org/10.1021/bc00031a006</a> .                               | ACS              | 27/12/2023                   | Yes      |
| Figure 5      | Monfardini, C., Schiavon, O., Caliceti, P., Morpurgo, M., Harris, J. M. & Veronese, F. M. A Branched Monomethoxypoly(ethylene glycol) for Protein Modification. <i>Bioconjugate Chem.</i> <b>6</b> , 62-69 (1995).<br><br><a href="https://doi.org/10.1021/bc00031a006">https://doi.org/10.1021/bc00031a006</a> .                               | ACS              | 27/12/2023                   | Yes      |
| Figure 6      | Bailon, P. & Berthold, W. Polyethylene glycol-conjugated pharmaceutical proteins. <i>Pharm. Sci. Technol. today</i> <b>1</b> , 352-356 (1998).<br><br><a href="https://doi.org/10.1016/S1461-5347(98)00086-8">https://doi.org/10.1016/S1461-5347(98)00086-8</a> .                                                                               | Elsevier         | 27/12/2023                   | Yes      |
| Figure 7      | Caliceti, P. & Veronese, F. M. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. <i>Adv. Drug Del. Rev.</i> <b>55</b> , 1261-1277 (2003).                                                                                                                                                             | Elsevier         | 27/12/2023                   | Yes      |

|           |                                                                                                                                                                                                                                                                                                                                                                      |                |            |     |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|------------|-----|
|           | <a href="https://doi.org/10.1016/S0169-409X(03)00108-X">https://doi.org/10.1016/S0169-409X(03)00108-X</a> .                                                                                                                                                                                                                                                          |                |            |     |
| Figure 8  | Caliceti, P. & Veronese, F. M. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)–protein conjugates. <i>Adv. Drug Del. Rev.</i> <b>55</b> , 1261-1277 (2003).<br><a href="https://doi.org/10.1016/S0169-409X(03)00108-X">https://doi.org/10.1016/S0169-409X(03)00108-X</a> .                                                                   | Elsevier       | 27/12/2023 | Yes |
| Figure 16 | French, A. C., Thompson, A. L. & Davis, B. G. High-Purity Discrete PEG-Oligomer Crystals Allow Structural Insight. <i>Angew. Chem. Int. Ed.</i> <b>48</b> , 1248-1252 (2009).<br><a href="https://doi.org/10.1002/anie.200804623">https://doi.org/10.1002/anie.200804623</a> .                                                                                       | Wiley          | 27/12/2023 | Yes |
| Figure 17 | Marchetti, P., Jimenez Solomon, M. F., Szekely, G. & Livingston, A. G. Molecular Separation with Organic Solvent Nanofiltration: A Critical Review. <i>Chem. Rev.</i> <b>114</b> , 10735-10806 (2014).<br><a href="https://doi.org/10.1021/cr500006j">https://doi.org/10.1021/cr500006j</a> .                                                                        | ACS            | 27/12/2023 | Yes |
| Figure 18 | Marchetti, P., Peeva, L. & Livingston, A. The Selectivity Challenge in Organic Solvent Nanofiltration: Membrane and Process Solutions. <i>Annu. Rev. Chem Biomol.</i> <b>8</b> , 473-497 (2017).<br><a href="https://doi.org/10.1146/annurev-chembioeng-060816-101325">https://doi.org/10.1146/annurev-chembioeng-060816-101325</a> .                                | Annual Reviews | 27/12/2023 | Yes |
| Figure 19 | Oxley, A., Gaffney, P. R. J., Kim, D., Marchetti, P. & Livingston, A. G. Graft modification of polybenzimidazole membranes for organic solvent ultrafiltration with scale up to spiral wound modules. <i>J. Membr. Sci.</i> <b>647</b> , 120199 (2022).<br><a href="https://doi.org/10.1016/j.memsci.2021.120199">https://doi.org/10.1016/j.memsci.2021.120199</a> . | Elsevier       | 27/12/2023 | Yes |