Communications

High-Purity Discrete PEG-Oligomer Crystals Allow Structural Insight**

Alister C. French, Amber L. Thompson, and Benjamin G. Davis*

Poly(ethylene glycols) (PEGs) are a class of molecules that show wide-ranging application in biological contexts.[1] For example, the attachment of chains of PEG to therapeutic proteins, peptides, and even small molecule therapeutics (PEG-ylation)[2] is a well-established method used to influence stability, immunogenicity,[3] pharmacokinetics,[1,2,4,5] and mode of action.[4,5] This water-soluble, nontoxic polymer has been used to improve solubility,[6] dramatically extend serum half life,[7] and enhance the oral bioavailability[2] of substrate molecules. PEGs have also played key roles in lipid self-assembly, including liposome development[8] and even the delineation of the structural mechanisms of lipid-bilayer formation and fusion.[9]

In the development of PEG-ylated biopharmaceuticals a first potential source of heterogeneity is low site-selectivity and low conversion during attachment to, for example, the protein or substrate.[6,10,11] The second potential source of heterogeneity is that of the attached ligand itself; in polymeric structures, such as PEG, this arises from a distribution of different-length oligomers (polydispersity). Well-controlled anionic polymerization can produce monofunctional PEG with polydispersity indices (PDIs) approaching 1.04 (see Supporting Information).[12,13] Polydisperse oligomers give rise to mixtures of PEG-ylated conjugates and regulatory bodies are inclined to demand further investigations before approving products of inexactly known composition.[1,14]

To surmount these difficulties and with the goal of greater functional precision we aim to create purer PEG-ylated proteins. Notably, enhancement of in vivo serum half-life is most significant above 20 KDa.[15] Whilst discrete PEG_{512}[16] (MW 22.6 kDa) is beyond the reach of current synthesis, a strategy can be envisaged through which the same functional effects can be achieved additively by attaching multiple, shorter chains. Statistical modeling[17] of such combinations revealed that current PEG sources[18–20] would be insufficiently pure (Figure 1).
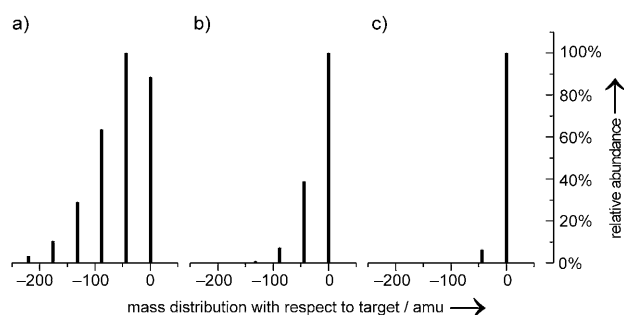
[*] A. C. French, Dr. A. L. Thompson, Prof. B. G. Davis
Department of Chemistry, University of Oxford, Chemistry Research Laboratory
12 Mansfield Road, Oxford (UK)
Fax: (+44) 1865-275-674
E-mail: ben.davis@chem.ox.ac.uk
Homepage: users.ox.ac.uk/ndplb0149

Figure 1. Statistical modeling[17] of multi-chain PEG brush: Simulated mass spectra showing the statistical distribution of 32 copies of PEG oligomers, PEG_m, at specific purities attached to a single substrate. a) 96% oligomer purity[20] (taken from analysis of Quanta BioDesign amino-$dPeg_{24}$ *tert*-butyl ester). b) 98.8% oligomer purity (target 16-mer purity chosen as benchmark). c) 99.8% oligomer purity (taken from actual purity of amino-PEG_{16} methyl ether **19** synthesized herein.

Our syntheses of 16-, 32-, and 48-mers of ethylene glycol, focused on minimizing contamination by other chain lengths. Despite advances in analytical, protecting, and coupling strategies,[21,22] Hibbert and co-workers' approaches to the first discrete EG oligomer[23] in the 1930s created PEG_{42},[24] which still remains one[25] of the longest reported (semi)-discrete oligomers. A common feature is strategic desymmetrization (e.g. protection) that prevents polymerization during coupling. There are then four conceptual ways to elongate the EG: iterative coupling of a mono-protected building block to A) one end, to B) both ends, C) chain doubling, or D) chain tripling (Figure 2). Mode D chain tripling allows rapid expansion but is limited[23,24,26] by difficulties in desymmetrizing PEGs of > 12 units.[27] Mode A/B strategies have produced symmetrical chains of up to 44[25] and 29[28] EG units by bidirectional extension of PEG diols with monobenzylated PEG monotosylates.[25,28] However, although such mode A/B iterative strategies provide longer oligomers for the first two filial coupling generations, the length of products from a mode C[29,27] chain doubling strategy follows the relationship $L = 2^g n$. Thus, once $g > 2$, the exponential relationship of a mode C doubling strategy greatly outstrips the linear growth of mode A/B iterative addition, $L = n(1+2g)$.[30]

We therefore adopted mode C multi-generation, chain doubling; tetraethylene glycol (EG_4, **1**) was chosen as a cheap, readily available starting material (100× cheaper and available in higher purity than EG_6[27] or EG_5.[31]) Highly pure PEG was accessed through an ultimately efficient and scaleable monoprotection and electrophilic activation (sulfonation) strategy on a 35–75 gram scale. Two key features allowed the preparation of highly pure oligomers of PEG_n. Firstly,
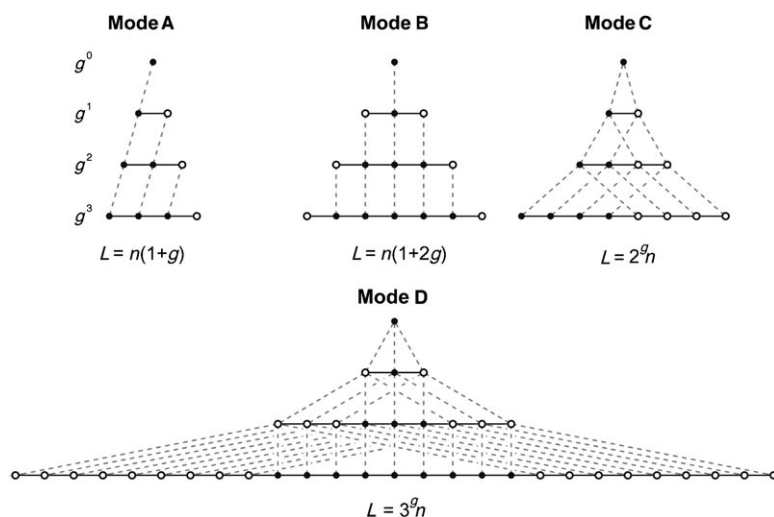
Figure 2. (Semi)discrete PEG oligomer synthesis strategies: Each circle represents an EG unit. *L* is the oligomer length after *g* generations of coupling given a starting material of length *n*. Mode A: unidirectional iterative coupling; Mode B: bidirectional iterative coupling; Mode C: chain doubling; Mode D: chain tripling.

ether protection,[37] in place of *tert*-butyl furnished substrates, such as **4**, with superior chromatographic separations. Secondly, with an empirical connection between presence of alkoxide and depolymerization established, subsequent coupling methods were designed to minimize reacting concentrations of alkoxide through slow addition of a solution of potassium *tert*-butoxide to the coupling partners, such as alcohols **4** or **7** and tosylates **5** or **12** in DMF. Thus, after coupling to form the octaethylene glycol benzyl trityl ether **7**, contaminating heptaethylene glycol benzyl trityl ether **9** was readily separated by normal-phase flash chromatography. Chain doubling, and subsequent reverse-phase flash chromatography to remove 15-mer gave hexadecaethylene glycol benzyl trityl ether **14** at 99.0% oligomer purity (PDI = 1.00009). Moreover, the strategy of minimizing alkoxide concentration through the slow addition of base[38] allowed the preparation of **14** in an even better 99.8% oligomer purity (PDI = 1.0000023) without the need for reverse-phase chromatography.

Next, we incorporated functionality useful for protein modification/conjugation. Thus, after acidolytic removal of Trt and tosylation, PEG$_{16}$ benzyl tosylate **16** was treated with sodium methoxide in DMF to give the benzyl methyl ether from which hydrogenolytic deprotection readily furnished PEG$_{16}$ monomethyl ether **17**. Further tosylation and reaction with sodium azide then allowed access to PEG$_{16}$ azido methyl ether **18** (suitable for attachment to proteins by, e.g., site-selective[39] copper-catalyzed Huisgen–Dimroth cycloaddition[40,41]), and by hydrogenation the PEG$_{16}$ amino compound **19**.

These PEG$_{16}$ derivatives are the highest purity PEGs of comparable length characterized to date, displaying PDI values between 1.000015 and 1.000002 and percentage purities between 99.6 and 99.9% single oligomer.[42] Remarkably, these exquisite purities allowed the first crystallizations of PEGs, which in turn allowed the formation of diffracting single crystals. Single-crystal X-ray diffraction experiments on these gave the first indication of the 3D structure and also a unique insight into an extended helical secondary structure of PEGs (Figure 3; see footnote and Supporting Information for experimental details). As a result of extended end-to-end packing of the PEG$_{16}$ **17** and a period of the PEG molecule that is not commensurate with the period of the helices or the lattice a continuous solution emerges. Although, given the quality of the data it is not possible to determine the structure fully, these results clearly indicate the presence of packed antiparallel helical strands with opposing handedness within which the oxygen atoms reside in the core of the helix with surrounding hydrophobic methylene groups. This behavior is strikingly similar to that proposed[43] in the formation of inverse micelles where again hydrophilic groups are sequestered in the micelle core. These helices pack together forming interleaved opposing-strands, such that each PEG is surrounded by four of the opposite handedness (Figure 3). The

avoidance of PEG-chain degrading depolymerization and, secondly, improved chromatographic contrast. This advance was accomplished through the survey of three contrasting hydroxy protecting groups: benzyl, *tert*-butyl and trityl.

EG$_4$ mono-*tert*-butyl ether **2**[32] and mono-benzyl ether **3**[33] were prepared from **1** using acid resin/isobutene and alkali/BnCl, respectively; importantly **2** and **3** purified well by normal-phase flash chromatography to yield mono-protected hydroxy nucleophilic coupling moieties for chain doubling. The electrophilic coupling components were accessed through activation as sulfonate esters; despite use in earlier examples of PEG synthesis,[24,34] attempted halogenation methods confirmed prior observations of chain degradation.[27] Thus, reaction of **3** with alkali/TsCl[25] afforded tosylate **5** on a 15 g scale (Scheme 1).

Coupling of **5** with **2** highlighted critical processes that threatened purity in PEG synthesis. Ether formation was initially performed by first treating alcohol **2** with sodium hydride in dry THF, then adding the tosylate **5** dropwise. After workup this gave the desired PEG$_8$ benzyl *tert*-butyl ether **6**. However the product was contaminated by 3% PEG$_7$ benzyl *tert*-butyl ether **8**, which resulted from partial depolymerization of the alcohol **2**;[35,36] the level of depolymerized contaminant increased with the prolonged existence of the intermediate alkoxide.
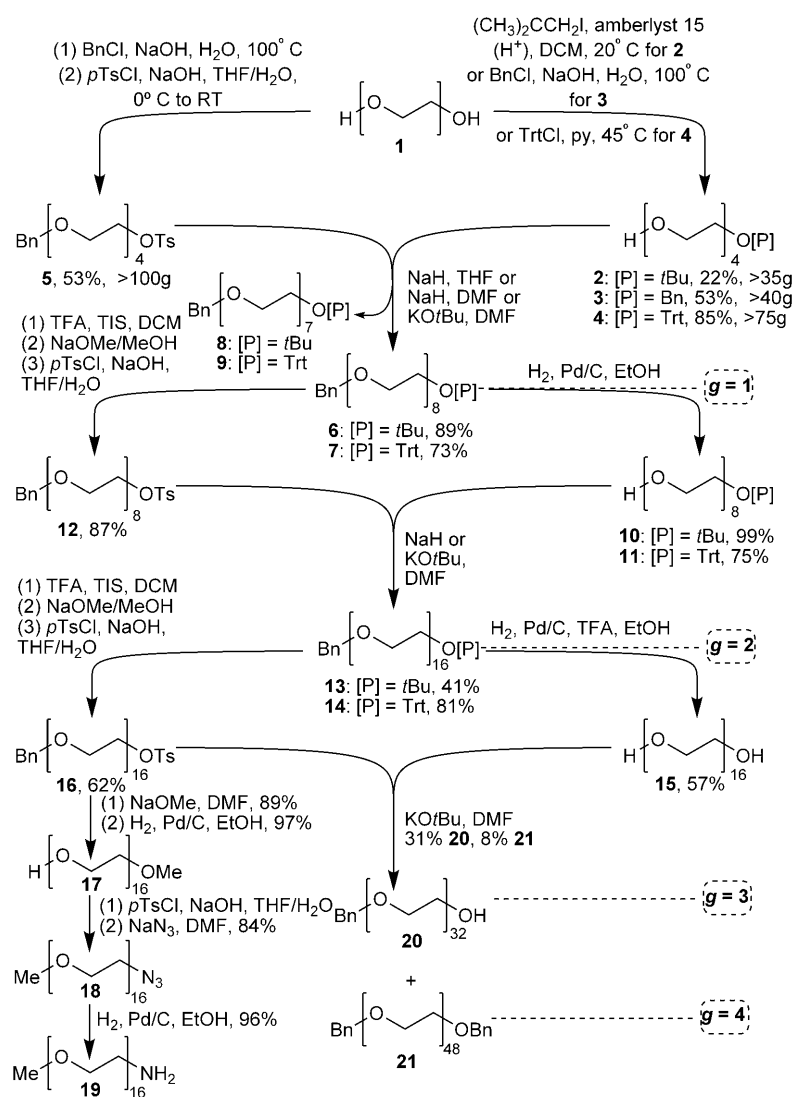
Despite considerable effort, lack of chromatographic resolution of these *t*Bu-protected PEGs **6** and **8** under a number conditions meant that complete separation or removal of truncated impurities, such as **8**, proved difficult and carriage of truncated PEG$_{n-1}$ impurities into subsequent chain-doubling generations to give 16-mer **13** and the corresponding 32-mer led only to compounded issues of contamination and a rapid collapse in purity.

Two adaptations allowed us to greatly improve the outcome of our synthetic strategy. Firstly, use of monotrityl

Scheme 1. PEG Syntheses: Chain doubling mode C synthesis of highly pure PEG oligomers. Trt = trityl, pTs = $para$-toluenesulfonyl, TIS = triisopropylsilane, DCM = dichloromethane, TFA = trifluoroacetic acid, [P] = protecting group.

distance between the centre of two similar handed coils is 6.4 Å), while opposing helices are closer (4.5 Å).

The need for exquisite purity in sample preparation was made clear by the melting points of these samples (29 °C for **17**, 31 °C for **18**, 23 °C for BnO-PEG$_{16}$-OMe) which are strikingly close to room temperature—the slightest depression in melting point would have rendered crystals unobtainable under many standard growth conditions. Although limited examples[45–47] of shorter (< 9) EG derivatives have been crystallized these have been achieved only through heavy-metal derivatization,[45,46] templating,[45–47] and/or unusual end-group modifications;[46] the resulting structures are not extended, give little insight into secondary structure, and are reminiscent only of crown ethers.

From these unique structures it is clear that PEG is capable of adopting an intriguing 3_{10}-helix (using Bragg nomenclature[48]) structure that is strikingly similar in shape and pitch to the rare 3_{10}-alpha-helical motifs found in some natural polypeptides.[49] We consider this first clear structural insight into the extended shape of PEG to be an important step in understanding the structural biology of the molecule class that is most commonly used for therapeutic protein modification and that is playing an important role in the understanding of bilayer systems. Moreover, the snapshot of lateral helix-to-helix packing observed in this case may also offer detailed insight into crystallinity effects seen in PEG-terminated self-assembled monolayers.[50]

Finally, higher generation discrete PEGs, PEG$_{32}$ (**20**) and PEG$_{48}$ (**21**), were synthesized. When selective hydrogenolytic deprotection of the benzyl trityl ether **14** was attempted to allow further chain doubling, fully deprotected hexadecaethylene glycol diol **15** was instead obtained as the major product. To explore couplings to give access to longer PEG oligomers, **15** was used in a trial desymmetrizing chain doubling with PEG$_{16}$ benzyl tosylate **16** (Scheme 1). Excitingly, this allowed not only formation of 3rd generation but also 4th generation PEGs, giving after chromatography, not only PEG$_{32}$ derivative **20** but also PEG$_{48}$ **21** in 98.9 % and 98.0 % purity, respectively.

In summary, we have produced alcohol, azide, and amine derivatives of hexadecaethylene glycol monomethyl ether in gram quantities at purities of > 99.8 % single oligomer. This is sufficiently pure for a 32-chain protein conjugate to be > 90 % single mass; these syntheses have also yielded PEG$_{32}$ and PEG$_{48}$. Moreover, these highly pure PEGs have given rise to the first X-ray crystal structures that have allowed a unique insight into PEG morphology. Work is currently ongoing to allow a more complete understanding of the complex factors that promote depolymerization and elimination in PEG synthesis in addition to the development of methods that will incorporate these highly pure PEGs into biomolecules.

Experimental Section

Characterization and methods for all compounds can be found in the Supporting Information.

17 (outline of synthesis): Na metal (0.5 g) was washed in petroleum ether and dissolved in dry MeOH (10 mL) overnight. **16** (7.0 g, 7.3 mmol) was taken up in dry DMF (90 mL) and treated with the resulting NaOMe solution. After 4 h, the reaction mixture was concentrated, worked up (NH$_4$Cl (aq)), extracted (CHCl$_3$), dried (Na$_2$SO$_4$) evaporated, and exposed to a 2nd reaction cycle. After 3 h, ESI-MS indicated > 99 % conversion. Work-up gave 6.9 g of crude product which was purified by gradient flash chromatography [Biotage SP4 system 40 mL min^{-1}; SNAP 100 g KP Silica cartridge; Solvent A ≡ DCM, solvent B ≡ 30 % methanol in DCM; column equilibrated at 0 % B; crude product diluted with 13 mL DCM and injected; column run 1 CV at 0 % B, 3 CV to 2 % B, 3 CV to 6 % B,
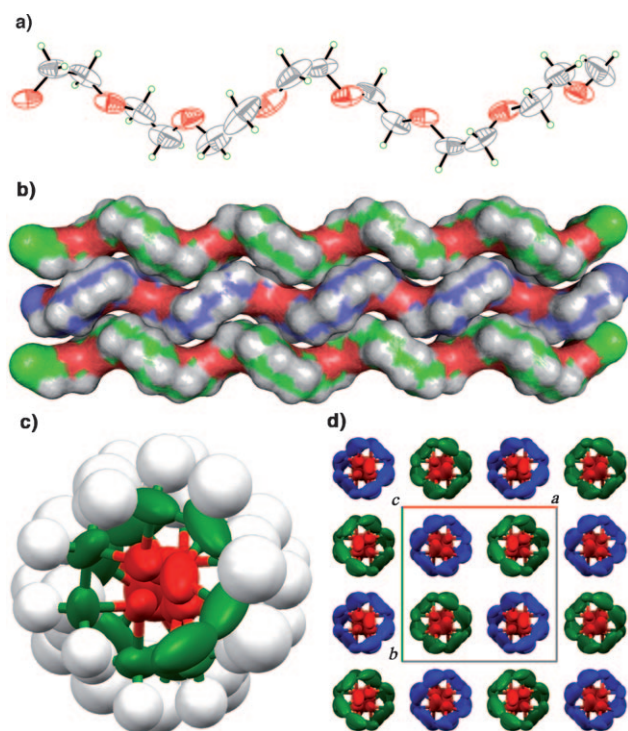
Figure 3. X-ray structure of PEG$_{16}$ (**17**). a) the close-packed 3$_{10}$-helix dominated by a *gauche* arrangement of individual EG units; red O, gray C, green H; b) accessible-surface diagram showing three antiparallel PEGs in blue and green; c) Thermal ellipsoid plot of **17** showing the oxygen core surrounded by the hydrophobic methylene groups; d) Packing in **17**, showing the two different helix hands in green and blue.[44]

3 CV to 14% B, and 3 CV to 30% B]. to give 5.11 g (6.18 mmol, 89% yield) of **17**, a colorless oil which crystallized upon being left to stand at room temperature.

Single-crystal X-ray diffraction data were collected using an Enraf–Nonius KappaCCD diffractometer (Mo$_{K\alpha}$ radiation; $\lambda = 0.71073$ Å) equipped with a Cryostream N2 open-flow cooling device,[51] and the data were collected at 150(2) K. Data were processed using the DENZO-SMN package[52] and the structure was solved by direct methods.[53] A number of models were examined using PLATON (ADSYMM)[54,55] and the final model was refined on F^2 using the CRYSTALS suite.[56] CCDC 707050 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

Single Crystal X-ray Diffraction Data: "C$_{20}$H$_{80}$O$_{20}$", $M_r = 881.06$, $0.20 \times 0.35 \times 0.40$ mm, Tetragonal ($I\bar{4}c2$), $a = 9.0621(11)$, $c = 28.680(5)$ Å, $V = 2355.2(6)$ Å^3, $Z = 2$, $\mu = 0.098$ mm^{-1}, $\rho_{calcd} = 1.242$ Mg m^{-3}, $T = 150(2)$ K, 6852 reflections collected, 583 independent [$R(int) = 0.129$], $R1 = 0.1403$, $wR2 = 0.2205$ [$I > 2\sigma(I)$], $\rho_{min,max} = -0.30$, 0.42 e Å^3.

[1] R. Haag, F. Kratz, *Angew. Chem.* **2006**, *118*, 1218; *Angew. Chem. Int. Ed.* **2006**, *45*, 1198.

[2] J. Whelan, C. Pan (Bayer Pharmaceuticals Corporation, USA), PCT Application WO 2005-US20469, **2005**.

[3] A. Abuchowski, T. van Es, N. C. Palczuk, F. F. Davis, *J. Biol. Chem.* **1977**, *252*, 3578.

[4] S. Zalipsky, *Adv. Drug Delivery Rev.* **1995**, *16*, 157.

[5] R. B. Greenwald, Y. H. Choe, J. McGuire, C. D. Conover, *Adv. Drug Delivery Rev.* **2003**, *55*, 217.

[6] S. Brocchini, A. Godwin, S. Balan, J.-w. Choi, M. Zloh, S. Shaunak, *Adv. Drug Delivery Rev.* **2008**, *60*, 3.

[7] P. Caliceti, F. M. Veronese, *Adv. Drug Delivery Rev.* **2003**, *55*, 1261.

[8] A. Gabizon, *Adv. Drug Delivery Rev.* **1997**, *24*, 337.

[9] B. R. Lentz, J. Lee, *Mol. Membr. Biol.* **1999**, *16*, 279.

[10] G. G. Kochendoerfer, *Curr. Opin. Chem. Biol.* **2005**, *9*, 555.

[11] Refs. [6,10] give examples of approaches that move beyond the statistical mixtures of products obtained by partial reaction of exposed protein side chains during typical methods using for example, PEG-acylating reagents, PEG-alkylating reagents, or reductive amination.

[12] N. N. Reed, K. D. Janda, *J. Org. Chem.* **2000**, *65*, 5843.

[13] J. Zhang, Y.-J. Zhao, Z.-G. Su, G.-H. Ma, *J. Appl. Polym. Sci.* **2007**, *105*, 3782.

[14] M. J. Vicent, L. Dieudonne, R. J. Carbajo, A. Pineda-Lucena, *Expert Opin. Drug Delivery* **2008**, *5*, 593.

[15] D. Filpula, H. Zhao, *Adv. Drug Delivery Rev.* **2008**, *60*, 29.

[16] We describe a discrete oligomer of *n* ether-linked ethylene units as PEG$_n$.

[17] See Supporting Information for modelling of the effects of even small impurities in multivalent conjugations.

[18] Polypure advertise material of > 95% single oligomer, up to PEG$_{28}$ prepared by sample displacement chromatography. However, the strategy is inefficient if a specific length is required and scales poorly for longer PEGs, limited by the diminishing relative differences between successive oligomers.

[19] Quanta BioDesign currently advertise discrete oligomers up to PEG$_{24}$; ESI-MS analysis of "Amino-dPEG$_{24}$ *tert*-Butyl Ester" (lot 314521) indicates that 96% of this material is the named oligomer.

[20] Whilst PDI is useful for discussing mixed ensembles of oligomers, it becomes cumbersome in the situation where one principal species is contaminated with a few others in minor quantities. Herein, we therefore discuss oligomer purity primarily on a percentage basis.

[21] M. B. Andrus, T. M. Turner, E. P. Updegraff, Z. E. Sauna, S. V. Ambudkar, *Tetrahedron Lett.* **2001**, *42*, 3819.

[22] A. Harada, J. Li, M. Kamachi, *J. Am. Chem. Soc.* **1994**, *116*, 3192.

[23] H. Perry, *Can. J. Res. Sect. B* **1936**, *14*, 77.

[24] R. Fordyce, E. L. Lovell, H. Hibbert, *J. Am. Chem. Soc.* **1939**, *61*, 1905.

[25] S. A. Ahmed, M. Tanaka, *J. Org. Chem.* **2006**, *71*, 9884.

[26] Hibbert's original synthesis was a chain tripling (where $L = 3\,gn$) that reached PEG$_{42}$ by the third filial generation ($g = 3$).

[27] F. A. Loiseau, K. K. Hii, A. M. Hill, *J. Org. Chem.* **2004**, *69*, 639.

[28] D. Niculescu-Duvaz, J. Getaz, C. J. Springer, *Bioconjugate Chem.* **2008**, *19*, 973.

[29] Two-generation chain doubling synthesis has allowed the synthesis of 24 unit EGs.

[30] It has been suggested (Ref. [28]) that bidirectional iterative growth is quicker and more efficient than chain doubling, yet our statistical analysis reveals that this is only true beyond $L = 16\,n$. It should also be noted that all the steps are not of equal strategic importance, since it is only during ether coupling that the most difficult to remove by-products are formed.

[31] Sigma–Aldrich prices 2008.

[32] R. D. Lousenberg, M. S. Shoichet, *J. Org. Chem.* **1997**, *62*, 7844.

[33] T. Gibson, *J. Org. Chem.* **1980**, *45*, 1095.

[34] G. Coudert, M. Mpassi, G. Guillaumet, C. Selve, *Synth. Commun.* **1986**, *16*, 19.

[35] N. Boden, R. J. Bushby, S. Clarkson, S. D. Evans, P. F. Knowles, A. Marsh, *Tetrahedron* **1997**, *53*, 10939.

[36] Base-catalyzed EG depolymerization during Williamson ether synthesis is seldom noted in the literature. See Boden et al. for a brief but excellent discussion (Ref. [35]); > 15 % of EG unit loss was observed in some examples.

[37] V. E. M. Kaats-Richters, J. W. Zwikker, E. M. D. Keegstra, L. W. Jenneskens, *Synth. Commun.* **1994**, *24*, 2399.

[38] Interestingly, the result was much reduced levels of depolymerization with higher levels of elimination of tosylate; whilst this is wasteful of tosylate, the resulting vinyl ether is much easier to remove than a truncated but chemically identical oligomer.

[39] S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105.

[40] C. W. Tornoe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057.

[41] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596.

[42] We have analyzed the highest purity material that is commercially available (Ref. [19]); it should be noted that a lack of mass spectra or spectral expansions in impurity regions in previous syntheses make purities difficult to assess.

[43] V. Luzzati, R. Vargas, A. Gulik, P. Mariani, J. M. Seddon, E. Rivas, *Biochemistry* **1992**, *31*, 279.

[44] Figure prepared using Pymol [W. L. DeLano. The PyMOL Molecular Graphics System (2008). DeLano Scientific, Palo Alto, California, USA. http://www.pymol.org] and Mercury [C. F. Macrae, P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor, M. Towler, J. van de Streek, *J. Appl. Crystallogr.* **2006**, *39*, 453–457].

[45] H. Ohmoto, Y. Kai, N. Yasuoka, N. Kasai, S. Yanagida, M. Okahara, *Bull. Chem. Soc. Jpn.* **1979**, *52*, 1209.

[46] G. Weber, W. Saenger, *Acta Crystallogr. Sect. B* **1979**, *35*, 3093.

[47] D. Britton, M. K. Chantooni, Jr., *J. Chem. Crystallogr.* **2001**, *31*, 5.

[48] L. Bragg, J. C. Kendrew, M. F. Perutz, *Proc. R. Soc. London Ser. A* **1950**, *203*, 321.

[49] C. Tonlolo, E. Benedetti, *Trends Biochem. Sci.* **1991**, *16*, 350.

[50] P. Harder, M. Grunze, R. Dahint, G. M. Whitesides, P. E. Laibinis, *J. Phys. Chem. B* **1998**, *102*, 426.

[51] J. Cosier, A. M. Glazer, *J. Appl. Crystallogr.* **1986**, *19*, 105.

[52] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, 276.

[53] A. Altomare, G. Cascarano, C. Giacovazzo, A. Guagliardi, M. C. Burla, G. Polidori, M. Camalli, *J. Appl. Crystallogr.* **1994**, *27*, 435.

[54] A. L. Speck, Utrecht University, **2001**.

[55] Y. Le Page, *J. Appl. Crystallogr.* **1988**, *21*, 983.

[56] P. W. Betteridge, J. R. Carruthers, R. I. Cooper, K. Prout, D. J. Watkin, *J. Appl. Crystallogr.* **2003**, *36*, 1487.