J|A|C|S
COMMUNICATIONS

Published on Web 09/18/2003

Synthesis of Monodisperse Polymers from Proteins

Jerry Yang,[†] Irina Gitlin,[†] Vijay M. Krishnamurthy,[†] Jenny A. Vazquez,[‡] Catherine E. Costello,[‡] and George M. Whitesides*,[†]

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, and Mass Spectrometry Resource, Departments of Biochemistry and Biophysics, Boston University School of Medicine, 715 Albany Street R-806, Boston, Massachusetts 02118

Received May 6, 2003; E-mail: gwhitesides@gmwgroup.harvard.edu

This paper describes the syntheses of representative members of a new class of monodisperse, semisynthetic polymers − polymers obtained by perfunctionalization of amine groups on a protein. Monodisperse polymers are centrally important in understanding the relationships between molecular and physical properties of macromolecules.[1] They may be useful where well-defined macromolecular structures are necessary (e.g., polyvalent biological agents[2]), or where physical properties must be tightly controlled (e.g., in molecular weight standards, or as subjects of biological or spectroscopic study). Truly monodisperse polymers are, however, difficult to prepare using the methods of conventional polymer synthesis. We hypothesized that by starting with polymeric backbones that were already monodisperse − proteins − and by modifying their side chains chemically, we might be able to synthesize monodisperse polymers with precisely tailored sequence and side chain functionality.

Proteins have a limited number of functional groups (e.g., Lys-NH_2, Asp-CO_2H, and Glu-CO_2H, Cys-CH_2SH, Ser/Thr-OH, Tyr-C_6H_4OH, Arg-$NHC(NH_2)NH$),[3] but the reactivities of these groups differ greatly depending on their position in the folded protein; as a result of these differences, their reaction in intact proteins is seldom complete. We have explored reactions of proteins, both in native and in denatured form, in the presence of sodium dodecyl sulfate (SDS). We demonstrate complete modification (>99% conversion of each amino group; 73−97% overall purity of the fully modified protein, after dialysis) of the NH_2 groups (Lys-NH_2 and N-terminal) of three proteins when they are denatured in SDS micelles.[4] This procedure provides a practical route, in principle, to a range of monodisperse polyamide-based polymers, having controlled sequence and functional, nonbiological side chains. These polymers may, of course, lose the native structure and function of the protein from which they were derived. In this work, however, we consider proteins only as readily available, truly monodisperse polyamides useful as the backbones for other, functional polymers.

We investigated the nucleophilic reaction of lysine side chains of three different proteins with activated carboxylic acid groups. We performed the reactions under pH-controlled conditions[5] using 0.1 M HEPES buffer (pH = 8.2) in the presence of 10 mM SDS at 25 °C (the SDS was present to render all lysine residues accessible to derivatization[4]). We arbitrarily set the concentration of protein at 100 μM in all experiments. The Supporting Information summarizes the complete experimental details.

In initial studies, we used acetic anhydride to acetylate the amino groups of the proteins. To demonstrate the generality of the procedure, we examined three proteins available commercially in high purity and having different values of molecular weight and number of modifiable side chains: ubiquitin (8.6 kDa, 7 lysines, 1 N-terminal amine), lysozyme (EC 3.2.1.17, 14.3 kDa, 6 lysines, 1 N-terminal amine), and bovine carbonic anhydrase II (BCA) (EC 4.2.1.1, 29 kDa, 18 lysines, post-translationally acetylated N-terminal amine). We have studied the partial acetylation of the lysine ϵ-NH_2 residues of these proteins with acetic anhydride previously.[6]

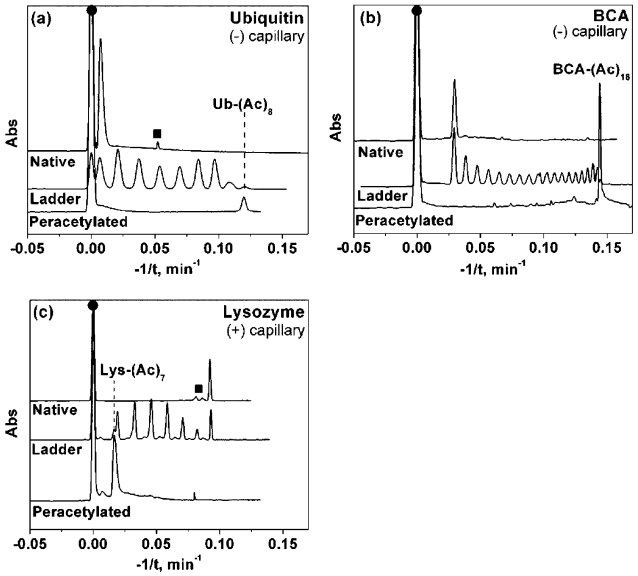

Figure 1. CE electropherograms (monitored at $\lambda = 214$ nm; $1/t$ is directly proportional to electrophoretic mobility) of native, partially acetylated (ladder), and peracetylated (a) ubiquitin, (b) bovine carbonic anhydrase (BCA), and (c) lysozyme: (●) indicates mesityl oxide added to measure electroosmotic flow; (■) indicates impurities in the starting protein. For analysis of lysozyme (a positively charged protein), we used a capillary with positively charged walls: using this capillary, the most negatively charged species eluted first (i.e., at the smallest value of electrophoretic mobility). The peracetylated derivatives of each protein correspond to the most negatively charged peak of the charge ladder. In (c), the last two peaks of the charge ladder are more closely spaced than the other peaks due to the difference in pK_a, and thus in the effective charge, between the Lys-NH_2 and N-terminal NH_2 groups.[6] The *y*-axis in (a−c) has arbitrary units and is not meant to reflect absolute absorbance.

We allowed solutions of ubiquitin, lysozyme, or BCA (in the presence of SDS) to react with acetic anhydride (20 mol equiv per amine group of the protein) and analyzed the resulting mixture of proteins by capillary electrophoresis (CE). We determined the extent of modification (i.e., the number of functionalized amine groups) by comparison of the electropherogram of the perfunctionalized species with a charge ladder[6,7] derived from that protein. We observed (Figure 1) peracetylated derivatives of these proteins as the only products (in ~93%, ~90%, and ~90% purity for ubiquitin, lysozyme, and BCA, respectively[8]). Mass spectrometric (MS) analysis of peracetylated ubiquitin indicated that this compound was indeed perfunctionalized on the amines, but that a species

[†] Harvard University.
[‡] Boston University School of Medicine.

10.1021/ja035978l CCC: $25.00

Table 1. Electrophiles Used To Perfunctionalize Ubiquitin

compound ID	modifying reagent	relative purities (%)[b,c]
1	$(CH_3CO)_2O$	93
2	$CH_3O(CH_2CH_2O)_3CH_2CH_2CO-(NHS)$[a]	73
3	C_6H_5COCl	97
4	$CO(CH_2)_3COO$[d]	89
5	$CF_3CH_2CO-(NHS)$[a]	84
6	$ICH_2CO-(NHS)$[a,d]	84

[a] NHS = *N*-hydroxysuccinimide. [b] As determined by CE. [c] See ref 8. [d] See Supporting Information.

containing one additional acetyl group was also present. Ubiquitin has a single tyrosine residue, and we hypothesized that the phenolic hydroxyl group of this amino acid might be acetylated. To remove this ester group selectively, we incubated perfunctionalized ubiquitin (100 μM) in 0.1 M LiOH for 1 h at 4 °C. Dialysis of the hydrolyzed protein solution against water and reanalysis by mass spectrometry gave a single peak, consistent in mass with one acylation per amine in ubiquitin. We subjected all other derivatized proteins to the same de-esterification procedure prior to CE and MS analysis. With ubiquitin and BCA, we achieved >93% recovery of soluble protein after analysis of solutions by UV (λ = 280 nm) before and after reaction with acetic anhydride (see Supporting Information). To demonstrate that this procedure for modification can yield useful quantities of derivatized protein, we converted 1 g of lysozyme to the peracetylated species (90% purity).

To establish conditions that consistently perfunctionalize a protein regardless of the size of or number of reactive residues, we explored the role of protein denaturation on the reaction of the lysine and N-terminal amine groups with electrophiles. We subjected ubiquitin and BCA, both in their native form, and denatured in SDS, to acetic anhydride (20 mol equiv per amine group in the protein) in 0.1 M HEPES buffer (pH = 8.2) and analyzed the products by CE, after de-esterification and dialysis. There was little difference in the reactivity of ubiquitin in native and denatured forms with acetic anhydride. With BCA (Figure 1b), however, we found essentially a single product (>90% purity) corresponding to 18 acetylated lysines when we carried out the reaction in SDS-containing solution, while we found a mixture of partially reacted products when we allowed BCA to react in its native form, in the absence of SDS.[6] This result suggests that for larger proteins (>30 kDa), denaturation is important to achieve complete derivatization of the amino groups. These reactions proceeded with indistinguishable results at temperatures ranging from 0 to 25 °C. Active pH control in the absence of buffer − that is, simultaneous titration of the electrophile (dissolved in dioxane) and aqueous 1 N NaOH − resulted in lower conversions to acylated products than did reactions performed in buffered solutions. Although smaller excesses of electrophile may be used in some cases, we used a 20-fold molar excess (per amine on the protein) of electrophile in all of the experiments we describe.

To investigate the generality of the conditions we used to derivatize proteins, we tested several active esters of carboxylic acids for their ability to acylate ubiquitin (Table 1). The results indicate that we can introduce a range of potentially useful chemical properties using this method, that is, functionality with hydrophilic (**2**), hydrophobic (**3**), charged (**4**), fluorinated (**5**), or chemically reactive (**6**) properties.

We have, thus, demonstrated that acylation of denatured proteins in SDS-containing buffer provides a method for synthesizing macromolecular linear polyamides with well-defined chain lengths and presenting nonbiological functional groups attached at known points on their backbones. This approach to the use of proteins as *polymers* complements methods based on genetic engineering.[9,10] For larger (here, BCA) proteins, this method requires the denaturation of the protein backbones with a denaturant such as SDS before reaction of the protein with acylating agents; SDS can be removed by dialysis after the modification step. For smaller proteins, SDS may not be necessary.

An efficient route to truly monodisperse polymers with tailored chemical groups may provide access to materials with defined structure (e.g., rigid rodlike molecules that display liquid crystalline behavior[10] or supramolecular organization[11]), with useful mechanical properties (e.g., by analogy with silk[12]), or with applications as adhesives (e.g., by analogy with mussel proteins[13]). Monodisperse polymeric materials may also be useful in fundamental studies in biochemistry,[14] biophysics, and polymer science, or as scaffolds for the polyvalent display of ligands and receptors.[2]

Acknowledgment. This work was supported by the NIH (grants GM51559, GM30367, and RR10888). I.G. acknowledges support from an NSF predoctoral fellowship. V.M.K. acknowledges support from NDSEG for a predoctoral research fellowship.

Supporting Information Available: Additional experimental details and mass spectra data (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Craver, C.; Carraher, C. *Applied Polymer Science 21st Century*; Elsevier: Amsterdam, 2000.
(2) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2755−2794.
(3) Walker, J. *The Protein Protocols Handbook*; Humana Press: Totowa, 2002.
(4) Magdassi, S. *Surface Activity of Proteins*; Marcel Dekker: New York, 1996. Jones, M. N. *Chem. Soc. Rev.* **1992**, 127−136. Turro, N. J.; Lei, X. G.; Ananthapadmanabhan, K. P.; Aronson, M. *Langmuir* **1995**, *11*, 2525−2533.
(5) Cline, G. W.; Hanna, S. B. *J. Am. Chem. Soc.* **1987**, *109*, 3087−3091. Cline, G. W.; Hanna, S. B. *J. Org. Chem.* **1988**, *53*, 3583−3586. King, J. F.; Rathore, R.; Lam, J. Y. L.; Guo, Z. R.; Klassen, D. F. *J. Am. Chem. Soc.* **1992**, *114*, 3028−3033.
(6) Gitlin, I.; Mayer, M.; Whitesides, G. M. *J. Phys. Chem. B* **2003**, *107*, 1466−1472. Carbeck, J. D.; Colton, I. J.; Gao, J.; Whitesides, G. M. *Acc. Chem. Res.* **1998**, *31*, 343−350 and references therein.
(7) A charge ladder is a collection of protein derivatives differing in the number of modified lysine (and N-terminal) amino groups and, thus, in charge. CE separates these derivatives into distinct groups of peaks. See ref 6.
(8) We estimated purity by integration of the major (peracetylated) peak relative to the integrated area of all observable deviations from baseline, assuming equal molar absorptivity coefficients for all species detected; these deviations may arise from partially functionalized products, decomposition, other side reactions, aggregated polymeric species, or interaction of the polymers with the walls of the capillary.
(9) Wang, L.; Magliery, T. J.; Liu, D. R.; Schultz, P. G. *J. Am. Chem. Soc.* **2000**, *122*, 5010−5011. Kim, D. M.; Swartz, J. R. *Biotechnol. Prog.* **2000**, *16*, 385−390. Ulmer, K. M. *Science* **1983**, *219*, 666−671.
(10) Yu, S. M.; Conticello, V. P.; Zhang, G.; Kayser, C.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *Nature* **1997**, *389*, 167−170. van Hest, J. C. M.; Tirrell, D. A. *Chem. Commun.* **2001**, *19*, 1897−1904.
(11) Lecommandoux, S.; Achard, M. F.; Langenwalter, J. F.; Klok, H. A. *Macromolecules* **2001**, *34*, 9100−9111.
(12) McGrath, K.; Kaplan, D. *Protein-Based Materials*; Birkhauser: Boston, 1997. Winkler, S.; Szela, S.; Avtges, P.; Valluzzi, R.; Kirschner, D. A.; Kaplan, D. *Int. J. Biol. Macromol.* **1999**, *24*, 265.
(13) Strausberg, R. L.; Link, R. P. *Trends Biotechnol.* **1990**, *8*, 53−57.
(14) Hoare, D. G.; Koshland, D. E., Jr. *J. Biol. Chem.* **1967**, *242*, 2447−2453. Oppenheimer, H. L.; Labouesse, B.; Hess, G. P. *J. Biol. Chem.* **1966**, *241*, 2720−2730. Labouesse, J.; Gervais, M. *Eur. J. Biochem.* **1967**, *2*, 215−223.

JA035978L