Journal of Macromolecular Science, Part A
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597274

Biosynthetic Incorporation and Chemical Modification of Alkene Functionality in Genetically Engineered Polymers

Timothy J. Deming[ab]; Maurille J. Fournier[c]; Thomas L. Mason[c]; David A. Tirrell[a]
[a] Departments of Polymer Science and Engineering, University of Massachusetts at Amherst, Amherst, Massachusetts [b] Department of Materials, University of California Santa Barbara, Santa Barbara, CA [c] Biochemistry and Molecular Biology, University of Massachusetts at Amherst, Amherst, Massachusetts

To cite this Article Deming, Timothy J. , Fournier, Maurille J. , Mason, Thomas L. and Tirrell, David A.(1997) 'Biosynthetic Incorporation and Chemical Modification of Alkene Functionality in Genetically Engineered Polymers', Journal of Macromolecular Science, Part A, 34: 10, 2143 — 2150
To link to this Article: DOI: 10.1080/10601329708010331
URL: http://dx.doi.org/10.1080/10601329708010331

J.M.S.—PURE APPL. CHEM., A34(10), pp. 2143–2150 (1997)

BIOSYNTHETIC INCORPORATION AND CHEMICAL MODIFICATION OF ALKENE FUNCTIONALITY IN GENETICALLY ENGINEERED POLYMERS

Timothy J. Deming,†§ Maurille J. Fournier, ‡ Thomas L. Mason,‡ and David A. Tirrell*†
†Departments of Polymer Science and Engineering and
‡Biochemistry and Molecular Biology
University of Massachusetts at Amherst
Amherst, Massachusetts 01003

ABSTRACT

Repetitive polypeptides of sequence $[(AlaGly)_3ProGluGly]_{16}$, **3a**, have been prepared in *Escherichia coli* as overexpressed recombinant proteins. Replacement of more than 90% of the naturally occurring proline (Pro) residues with 3,4-dehydroproline (Dhp) in sequence **3a** was achieved by *in vivo* expression of the target protein in medium containing Dhp and lacking Pro. The resulting material (**3b**) was treated with H_2O_2 or Br_2 to yield polymers containing 3,4-dihydroxyproline (Dhy, **3c**) and 3,4-dibromoproline (Dbr, **3d**), respectively, in place of the Dhp residue. These results represent the first demonstration of the incorporation and modification of alkene functionality in recombinant proteins.

INTRODUCTION

The incorporation of unnatural amino acid analogs into proteins allows the introduction of unusual reactive sites into these biological materials. Biosynthetic

*To whom correspondence should be addressed.
§ = Current address: Department of Materials, University of California Santa Barbara, Santa Barbara, CA 93106

incorporation of analogs has the advantage of placing these reactive sites at precisely defined locations in the polypeptide sequence [1]. We have previously reported the incorporation of some reactive amino acids (e.g., selenomethionine [2] and 3-thienylalanine [3]) into repetitive polypeptides for just these reasons; namely, that the reactive group can be displayed in a regular pattern along the protein chain or perhaps on surfaces or in crystalline or liquid crystalline arrays. We now report the introduction of an alkene functional group into a repetitive polypeptide and subsequent chemical modification of this residue. This is the first demonstration of alkene group incorporation into a recombinant protein.

The source of the alkene group is the proline (Pro, **1**) analog 3,4-dehydroproline (Dhp, **2**). Dhp has been shown to be incorporated *in vivo* and *in vitro* into the cellular proteins of *Escherichia coli* [4] as well as in chick embryos for studies on collagen biosynthesis [5]. Since Dhp is very similar in size and conformation to Pro, it readily replaces Pro in protein biosynthesis, and analog incorporation levels approaching 100% have been achieved [4b, 6]. We sought to incorporate Dhp into the polypeptide of sequence **3a**, since this polymer has the potential to adopt a lamellar morphology which would display the Pro (or Dhp) residues at the surface of crystalline aggregates, making the alkene groups accessible for subsequent reactions [7].

$[(AlaGly)_3XxxGluGly]_{16}$

3a: Xxx = Pro
3b: Xxx = Dhp
3c: Xxx = Dhy
d: Xxx = Dbr

EXPERIMENTAL

Analysis of Dhp Incorporation

A single colony of BL21(DE3)pLysS pET3-8 Pro⁻ was used to inoculate M9AA minimal medium (30 mL of M9 medium [8] containing the 20 natural amino acids at 20 μg/mL and supplemented with 0.1 mM $CaCl_2$, 1.0 mM $MgSO_4$, 25μg/mL chloramphenicol, and 200μg/mL ampicillin). The culture was shaken at 37°C until the O.D. (600 nm) = 0.75, at which point IPTG (final concentration = 0.4 mM) was added and shaking was continued for 10 minutes. Cells were isolated by centrifugation of 2/3 of the mixture (20 mL) and the pellet was washed

with M9AA minimal medium without proline. The cells were resuspended in M9AA minimal medium without proline (20 mL). To the remaining 1/3 of original culture (10 mL) was added a solution of ^{35}S-Met (20 μL, 300 μCi) and to the sample of washed culture (20 mL) was also added ^{35}S-Met (40 μL, 600 μCi). The 20 mL culture was divided into 2 equal portions. To one portion was added Dhp (400 μg of the racemate as a 1 mg/mL solution) and to the other portion nothing was added. All 3 samples were shaken at 37°C for 3.5 hours and aliquots (1.5 mL) were removed at time intervals of 15, 60, 120, and 210 minutes after addition of the analog. Cells were isolated from each aliquot by centrifugation and the supernatents were discarded. Cells were suspended in M9AA minimal medium without proline (1.0 mL), centrifuged again, and the pellets were then lysed and fractionated on a 12% polyacrylamide gel. The samples were electrophoresed at 12 mA constant current overnight and the proteins stained with Coomassie Brilliant Blue G-250. After destaining with 45% MeOH/10% HOAc the gels were dried and then exposed to X-ray film for 2 d at -80°C.

Isolation of the Polymer

Polymer **3a** was prepared by growth of the host strain in rich medium (2x YT) [8] followed by expression of the target sequence by addition of isopropyl β-D-thiogalactopyranoside (IPTG). The polymer was expressed as a fusion protein of overall sequence, MASMTGGQQMGRDPMFKYSRDPMG $[AGAGAGPEG]_{16}$ ARMHIRPGRYQLDPAANKARKEAELAAATAEQ; the flanking segments of which were encoded by pET3b and p937.51 vectors used for cloning and expression [7]. For the synthesis of **3b**, cells were grown in rich medium and then shifted to medium containing Dhp followed by induction with IPTG. The recombinant proteins partitioned into the soluble portion of the cell lysate and were isolated by lowering the pH of the lysate to 4.0 and then precipitating the target protein with increasing concentrations of $(NH_4)_2SO_4$. Proteins **3a** and **3b** were found to precipitate in the range of 40-60% $(NH_4)_2SO_4$ saturation at ambient temperature. Removal of the fusion segments with CNBr gave the target polypeptides which were purified by dissolution in 96% formic acid (1.0 mL per 100 mg polypeptide) and then precipitated by addition of 3 volumes of methanol. The precipitates were washed with 5 x 4 volumes of methanol and then dried *in vacuo* to give the polypeptides in overall yields of 30 mg/L (**3a**) and 20 mg/L (**3b**) which represent approximately 20% and 13% of total cellular protein, respectively, at the cell densities used in these experiments [3].

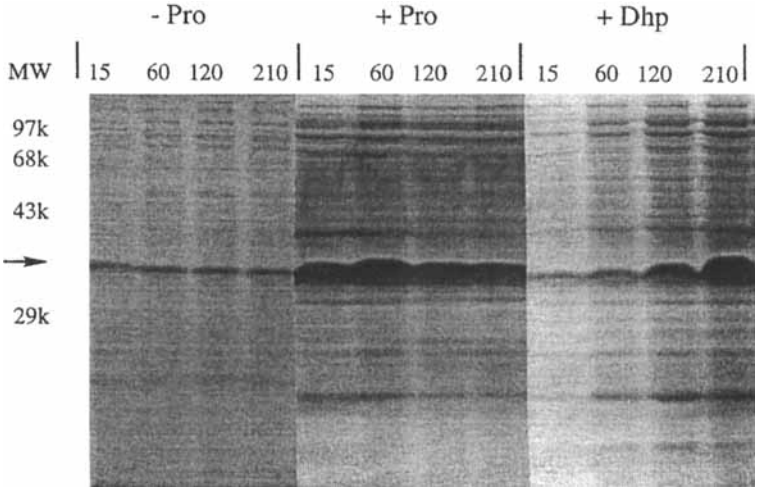


Figure 1. SDS-PAGE analysis of whole cell lysates visualized by ^{35}S-Met labeling. Lanes 1-4 contain target protein from cells shifted to medium lacking Pro; lanes 5-8 contain target protein from cells incubated in medium containing Pro; and lanes 9-12 contain target protein from cells shifted to medium containing Dhp. The target protein is indicated by an arrow. Molecular weight standards are listed to the left of the autoradiogram: the apparent molecular weights of **3a** and **3b** were estimated from these size standards. The target protein migrates at a molecular weight of about 38,000, which is significantly higher than the predicted value (16,900). This anomalous migration behavior is consistent with previous observations and is not a result of genetic instability [7].

RESULTS AND DISCUSSION

Construction of an artificial gene encoding **3a** has been described previously [7]. Proteins **3a** and **3b** were synthesized by using a T7 RNA polymerase expression system in *E. coli;* the host strain BL21(DE3)pLysS pET3-8 Pro$^-$ was grown in medium containing Pro or Dhp respectively (Figure 1). From the figure it can be seen that protein accumulates to significantly higher levels in Dhp medium than in proline-free controls, suggesting that Dhp is incorporated into the polypeptide. The relative amounts of Pro and Dhp in the polypeptides were first determined using amino acid compositional analyses (Table 1). The data show that the samples are at least 90% pure and that the extent of substitution with Dhp

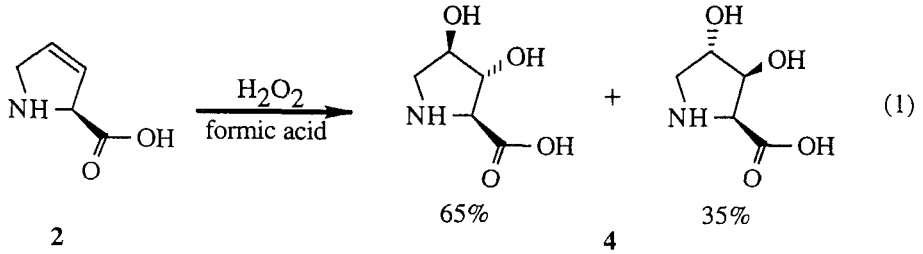
TABLE 1. Amino Acid Compositions of Proteins **3a**, **3b**, **3c**, and **3d**.

		mol % (obsd)			
amino acid	mol % (theor)	**3a**	**3b**	**3c**[d]	**3d**[d]
glycine	29.0	27.9	27.4	31.0	31.2
alanine	34.5	35.3	34.5	34.8	37.9
glutamic acid	9.5	10.6	10.5	10.8	11.4
proline	10.0[a]	10.1	1.4	1.6	1.5
dehydroprolin e	10.0[b]	ND[c]	9.0	ND	ND

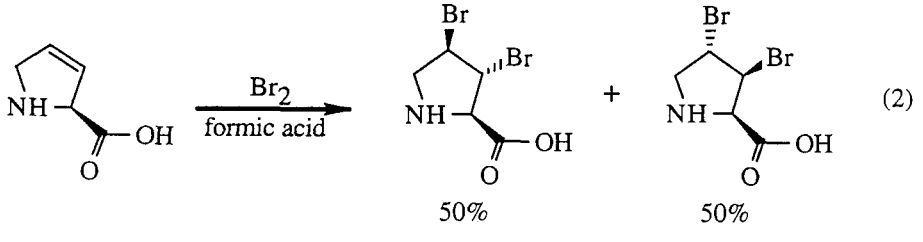
[a] Expected for **3a**.
[b] Expected for **3b**.
[c] None detected.
[d] Dhy and Dbr were found to degrade under the conditions (6N HCl or CH_3SO_3H) used to hydrolyze and analyze the peptides and so could not be measured directly by this technique.

is about 90% for protein **3b**. The extent of analog substitution can also be estimated from ^1H NMR spectra of the purified polymers. Integration of Dhp resonances at 6.12 and 5.88 ppm versus those of Ala (1.36 ppm) and Gly (4.13 ppm) placed the amount of analog incorporation at >95%. Electrospray ionization mass spectrometry gave the mass of polymer **3b** as 11,090 after CNBr cleavage, compared with an expected mass of 11,094, assuming 100% Dhp substitution and pairing of 3 Na^+ ions with the glutamate side chains [9].

Both **3a** and **3b** were soluble in water and were found to be reluctant to crystallize. However, protein **3b** has the advantage of possessing chemically reactive sites at the putative turn positions (Pro residues) [7] in the sequence. To explore the reactivity of Dhp, we converted it as the free amino acid into 3,4-dihydroxyproline (Dhy, **4**) [10] or 3,4-dibromoproline (Dbr, **5**) via the model reactions shown in Equations 1 and 2. Both of these reactions proceeded essentially quantitatively [11] to yield the diastereomers indicated. When applied to the polymer these reactions should convert the small, hydrophobic Dhp residue to the hydrophilic Dhy and to the large, heavy atom-labeled Dbr residues, respectively.



2 4



2 5

Samples of protein **3b** were treated with either H_2O_2 or Br_2 in formic acid solutions at room temperature for 3 d to prepare samples of **3c** and **3d**. ^1H NMR analysis of the resulting materials revealed that the alkene proton resonances (6.12 and 5.88 ppm) had disappeared, indicating complete reaction of the Dhp units. The other resonances in the spectra remained unaffected by the treatment, attesting to the specificity of the reagents for the alkene functionality. Amino acid analysis of **3c** and **3d** also indicated complete reactivity of the Dhp units (Table 1).

CONCLUSION

These results demonstrate the insertion of multiple reactive alkene functionalities into genetically engineered polypeptides and the subsequent site-specific modification of these groups. This strategy should be useful for introducing a variety of novel side groups into recombinant proteins. The physical properties of modified polymers **3c** and **3d** are currently under investigation.

ACKNOWLEDGMENTS

This work was supported by Contract No. DAAHO4-93-G-0217 from the U. S. Army Research Office and by a grant (DMR-8914359) from the Polymers and

Genetics Programs of the National Science Foundation. T. J. D. acknowledges the National Institutes of Health for a National Research Service Award postdoctoral training fellowship. NMR spectra were recorded in the University of Massachusetts NMR Facility, which is supported in part by the NSF Materials Research Science and Engineering Center at the University.

REFERENCES

[1] (a) M. J. Wilson and D. L. Hatfield, *Biochim. Biophys. Acta* , *781*, 205-215 (1984); (b) M. H. Richmond, *Bacteriol. Rev.*, *26*, 398-420 (1962); (c) G. Hortin and I. Boime, *Meth. Enzymol.*, *96*, 777-784 (1983).

[2] M. J. Dougherty, S. Kothakota, M. J. Fournier, T. L. Mason, and D. A. Tirrell, *Macromolecules*, *26*, 1779-1781 (1993).

[3] S. Kothakota, T. L. Mason, D. A. Tirrell, M. J. Fournier, *J. Am. Chem. Soc.*, *117*, 536-537 (1995).

[4] (a) H. Tristram and C. F. Thurston, *Nature*, *212*, 74-75 (1966); (b) L. Fowden, S. Neale, and H. Tristram, *Nature*, *199*, 35-38 (1963); (c) L. C. Smith, J. M. Ravel, C. G. Skinner, and W. Shive, *Arch. Biochem. Biophys.*, *99*, 60-64 (1962); (d) A. B. Mauger and B. Witkop, *Chem. Rev.*, *66*, 47-86 (1966).

[5] (a) J. Rosenbloom and D. J. Prockop, *J. Biol. Chem.*, *245*, 3361-3368 (1970); (b) S. Jimenez, and J. Rosenbloom, *Arch. Biochem. Biophys.*, *163*, 459-465 (1974).

[6] In order to obtain high levels of analog incorporation it is necessary to utilize a host strain which cannot produce Pro (an auxotroph) so that the cells and medium can be depleted of the natural amino acid. Since Dhp does not support the growth of *E. coli* it is necessary to grow cells first in Pro-containing medium and then induce expression of the target protein in medium containing Dhp.

[7] K. P. McGrath, M. J. Fournier, T. L. Mason, and D. A. Tirrell, *J. Am. Chem. Soc.*, *114*, 727-733 (1992).

[8] J. H. Miller, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, New York, 1972, p. 419.

[9] The observed mass of the Pro form of the polymer was 11,065, in excellent agreement with the expected mass of 11,065, given binding of one Na^+ counterion.

[10] (a) T. Nakajima and B. E. Volcani, *Science*, *164*, 1400-1401 (1969); (b) I. L. Karle, J. W. Daly, and B. Witkop, *Science*, *164*, 1401-1402 (1969);. (c) A. Buku, H. Faulstich, T. Wieland, and T. Dabrowski, *J. Proc. Natl. Acad. Sci., USA*, *77*, 2370-2371 (1980).

[11] Yields estimated by ^1H NMR spectroscopy. 3,4-Dihydroxyproline: Anal. calcd. for $C_5H_9O_4N$ with 6.3% H_2O: C, 38.5%; H, 5.8%; N, 9.0%. Found: C, 38.1%; H, 5.5%; N, 9.4%. 3,4-Dibromoproline hydrobromide: Anal. calcd. for $C_5H_8O_2NBr_3$: C, 17.0%; H, 2.3%; N, 4.0%. Found: C, 17.0%; H, 2.2%; N, 3.9%.