ORGANIC LETTERS

2003
Vol. 5, No. 2
121−124

Synthesis of Long-Chain Fatty Acid Enol Esters Isolated from an Environmental DNA Clone

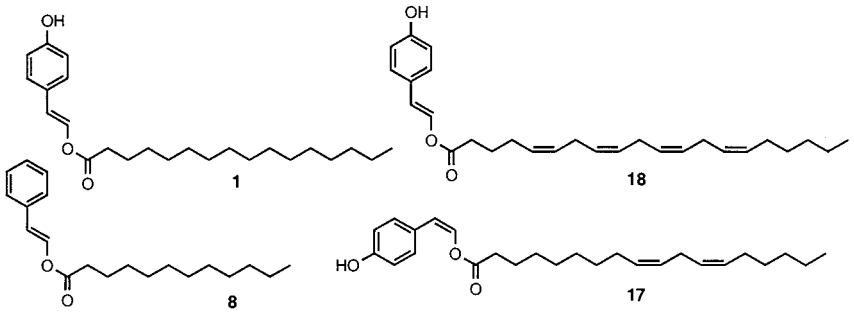
Sean F. Brady and Jon Clardy*

Department of Chemistry and Chemical Biology, Cornell University, Baker Laboratory, Ithaca, New York 14853-1301

jcc12@cornell.edu

Received August 22, 2002

ABSTRACT



Long-chain fatty acid enol ester 1 is the major metabolite of a new family of small molecules isolated from the heterologous expression of environmentally derived DNA. A versatile synthesis of 1, in which an aromatic acetaldehyde is O-acylated with a long-chain acyl chloride allowed for the rapid construction of both the isolated product (1) and a number of structural analogues (including 8, 17, and 18).

For the past 50 years natural products from cultured soil microbes have played a central role in drug discovery. Nucleic acid based studies now suggest that these cultured bacteria represent only a small fraction of the bacterial diversity present in most environmental samples.[1] The uncultured majority is a potentially rich source of structurally unique natural products. Heterologous expression of DNA extracted directly from environmental samples (environmental DNA, eDNA) in easily cultured hosts should provide access to natural products produced by many previously uncultured bacteria.[2] The first new natural product biosynthetic gene cluster characterized using this method, the *fee* (*f*atty acid *e*nol *e*sters) gene cluster, produces long-chain fatty acid enol esters that have not been previously characterized from the extensive screening of cultured bacteria.[3] A synthesis of the major metabolite in this family was undertaken to confirm the proposed structure **1** and provide access to natural-product-like derivatives not found in the culture broth of this heterologous expression system.

The *fee* gene cluster contains 13 opening reading frames (*fee*A−*fee*M) and was cloned directly from soil collected on the Cornell University campus. The *N*-acyl amino acid synthase (NAS, FeeM) of the *fee* gene cluster together with an acyl carrier protein (ACP) are thought to produce long-

(1) Torsvik, V.; Sorheim, R.; Goksoyr, J. *J. Ind. Microbiol.* **1996**, *17*, 170−178. Torsvik, V.; Goksoyr, J.; Daae, F. L. *Appl. Environ. Microbiol.* **1990**, *56*, 782−787. Hugenholtz, P.; Goebel, B. M.; Pace, N. R. *J. Bacteriol.* **1998**, *180*, 4765−4774. Stackebrandt, E.; Liesack, W.; Goebel, B. M. *FASEB J.* **1993**, *7*, 232−236. Ward, D. M.; Weller, R.; Bateson, M. M. *Nature* **1990**, *345*, 63−65.

(2) Brady, S. F.; Clardy, J. *J. Am. Chem. Soc.* **2000**, *122*, 12903−12904. Brady, S. F.; Chao, C. J.; Handelsman, J.; Clardy, J. *Org. Lett.* **2001**, *3*, 1981−1984. Gillespie, D. E.; Brady, S. F.; Betterman, A. D.; Cianciotto, N. P.; Liles, M. R.; Rondon, M. R.; Clardy, J.; Goodman, R. M.; Handelsman, J. *Appl. Environ. Microbiol.* **2002**, *68*, 4301−4306. Wang, G.; Graziani, E.; Waters, B.; Pan, W.; Li, X.; McDermott, J.; Meurer, G.; Saxena, G.; Anderson, R.; Davies, J. *Org. Lett.* **2000**, *2*, 2401−2404.

(3) Brady, S. F.; Chao, C. J.; Clardy, J. *J. Am. Chem. Soc.* **2002**, *124*, 9968−9969.

10.1021/ol0267681 CCC: $25.00
Published on Web 12/24/2002

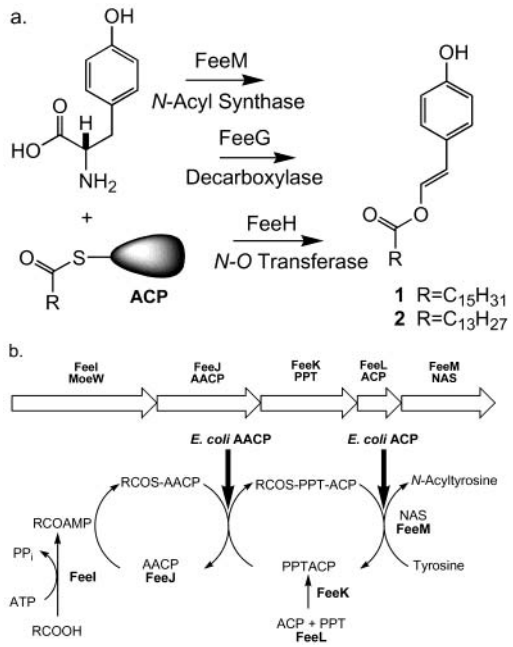


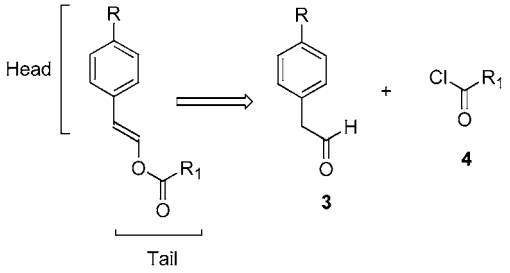
Figure 1. (a) Long-chain enol esters **1** and **2** are derived from tyrosine in successive steps catalyzed by an *N*-acyl amino acid synthase (NAS),[4] a decarboxylase, and an N−O transferase. (b) When the *fee* gene cluster is expressed in *E. coli* the fatty acids found in these natural products are likely specified by the *E. coli* fatty acid loading machinery and not the fatty acid loading system specific to the *fee* gene cluster.

chain *N*-acyltyrosines[4] that are oxidatively decarboxylated (FeeG) and converted from the eneamide to the enol ester (FeeH) to give **1** and **2** (Figure 1). Transposon mutagenesis studies suggest that the ACP could either be the *fee*-specific ACP (FeeL) or the endogenous *E. coli* ACP. In each case it appears that the native *E. coli* fatty acid loading machinery and not the pathway-specific fatty acid loading system is likely used to charge the ACP (Figure 1).[3] The biological properties of many natural products that contain long-chain fatty acids (i.e. homoserine lactones, Nod factors, insect pheromones, and lipid A) are highly dependent on the specific long-chain fatty acids found within these molecules. Because the *E. coli* fatty acid loading machinery is being used in the biosynthesis of these molecules, the specific fatty acid(s) encoded by the fatty acid loading system of the *fee* gene cluster may not be incorporated into **1** and **2**. Thus the isolated metabolites **1** and **2** may not be the compounds produced by the *fee* gene cluster in the organism from which the eDNA was derived. The synthesis of compounds related to **1** was undertaken to gain access to a pool of molecules resembling metabolites that might be produced by the *fee* gene cluster in the wild.

(4) The proposed biosynthesis of long-chain *N*-acyl amino acids is inferred from published studies on the biosynthesis of long-chain homoserine lactones. More, M. I.; Finger, L. D.; Stryker, J. L.; Fuqua, C.; Eberhard, A.; Winans, S. C. *Science* **1996**, *272*, 1655−1658. Watson, W. T.; Minogue, T. D.; Val, D. L.; Bec von Bodman, S.; Churchill, M. E. A. *Mol. Cell* **2002**, *9*, 685−694.

We sought a flexible synthesis of **1** that could be used to create a library of natural-product-like compounds for future biological screening. It seemed likely that the most versatile route to **1** would exploit the O-acylation of an aldehyde (**3**) with an acid chloride (**4**) (Scheme 1). Because

Scheme 1



numerous long-chain acid chlorides and aromatic aldehydes are commercially available the head, tail and configuration of the double bond could be easily varied with this approach.

Phenylacetaldehyde (**5**) was used to construct a family of long-chain fatty acid enol esters, **6**−**8**, as a model for the formation of long-chain enol esters from an aromatic acetaldehyde and a long-chain acyl chloride (Scheme 1, Figure 2). These compounds are of particular interest because

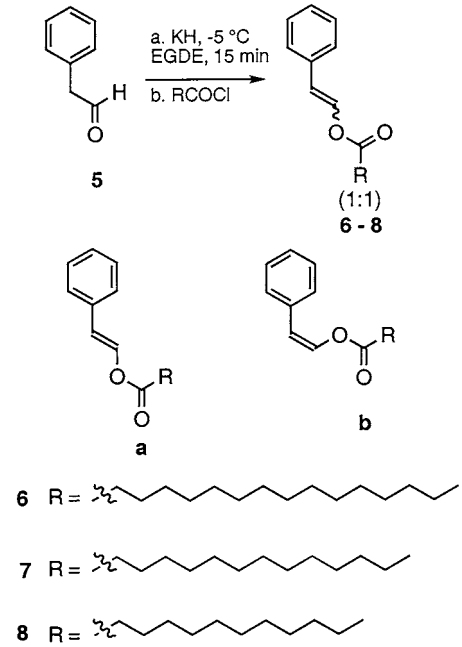


Figure 2. Long-chain enol esters prepared from phenylacetaldehyde as examples of synthetic phenylalanine derivatives of the natural tyrosine based enol esters **1** and **2**.

they represent the products of a *fee*-like biosynthetic pathwaythat uses phenylalanine in place of tyrosine as the amino acid precursor.

The potassium enolate of phenylacetaldehyde (**5**) was prepared using KH in ethylene glycol diethyl ether (EGDE) at −5 °C.[5] Upon the addition of a long-chain acid chloride (2 equiv) to the potassium enolate, long-chain enol esters **6**−**8** were obtained in 64−84% yield. In general we have observed that our yields decrease as the length of the fatty acid increases. After purification by normal phase flash chromatography (hexanes/dichloromethane) the cis and trans isomers of each long-chain enol ester were obtained as approximately 1:1 mixtures. The initial six members of the phenylalanine-based library are cis and trans isomers of long-chain fatty acid enol esters containing palmitoyl (**6**),[6] myristoyl (**7**), and lauryl (**8**) groups (Figure 2).

The acetaldehyde needed for the synthesis of long-chain enol esters based on tyrosine was obtained from the oxidation of *tert*-butyldimethylsilyl (TBS)-protected 4-hydroxyphenethyl alcohol (**9**). 4-Hydroxyphenethyl alcohol (**9**) was protected as the bis TBS silyl ether[7] (**10**) followed by selective deprotection of the secondary silyl ether with I_2 in methanol[8] to afford **11** in 87% yield over two steps. A Swern oxidation of the TBS protected 4-hydroxyphenethyl alcohol (**11**) provided aldehyde **12** in 67% yield.[9]

The synthesis of long-chain enol esters from **12** using KH gave only poor yields in our hands. Even when acetyl chloride was used as a model reactant in place of the less soluble longer-chain acid chlorides, a large amount of the starting aldehyde (**12**) was recovered from the reaction. Under these conditions it is appears that KH is not able to efficiently form the K enolate of the para-substituted acetaldehyde **12**. We found that replacing KH with potassium bis(trimethylsilane)amide (KHMDS) under essentially the same reaction conditions greatly increased the yield of long-chain enol esters.[10] Using KHMDS to trap the potassium enolate of **12**, the TBS-protected long-chain enol esters of **1** and **17**−**19** were obtained in 65−85% yield. Each of the TBDMS-protected enol esters were then deprotected using 1.1 equiv of TBAF in THF (56−85% yield).[11] With the appropriate mobile phase (hexanes/dichloromethane) the cis and trans isomers of the deprotected products could be separated by normal phase flash chromatography.

The large coupling constant (J = 13 Hz) observed in the natural sample of **1** suggested that the enolate contained a trans double bond. The cis and trans isomers of **1** were separated cleanly by silica gel flash chromatography (90:10, hexanes/dichloromethane) and the synthetic trans isomer **1a** (**1a** trans, J = 13 Hz, **1b** cis, J = 7.5 Hz) was found to be spectroscopically identical to the isolated compound. In addition to **1**, cis and trans isomers of three analogues that differ in the fatty acid attached to the tyrosine derived headgroup were constructed. These derivatives include analogues containing linoleic (**17**), arachidonic (**18**), and capric (**19**) fatty acids (Figure 3).

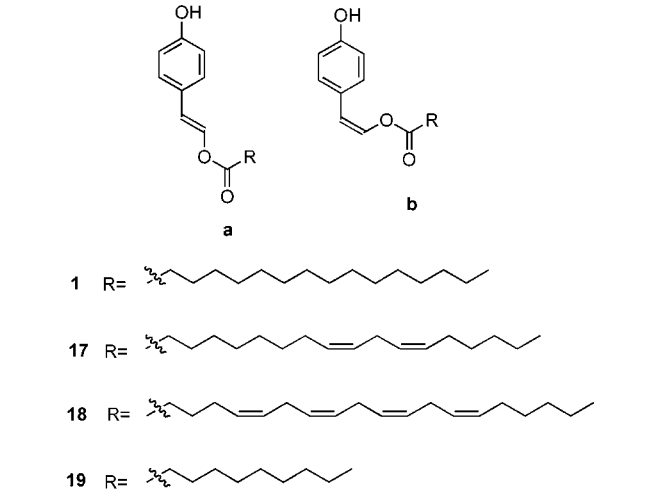


Figure 3. The isolated long-chain enol ester **1a** and analogues constructed as initial members of a library of compounds based on **1** and **2**.

The general strategy presented here (Scheme 2) is applicable to the rapid synthesis of a modest library of

(5) The procedure for the formation of the enol ester from a potassium enolate was modified slightly from: Ladjama, D.; Riehl, J. J. *Synthesis* **1979**, *7*, 504−507. KH (1.1 equiv) was cleaned of mineral oil using three pentane washes and then resuspended in EGDE at 1 mmol/500 μL; 0.5 mmol of the aldehyde (**5**) in 200 μL of EGDE was added to 1.1 equiv of KH stirring in EGDE at −5 °C. Two equivalents of acid chloride (chilled in 400 μL of EGDE) was added to the K enolate after 10 min and then stirred at room temperature for 15 min. After the addition of 1 mL of water the reaction was extracted five times with pentane. The dried organic extracts were then chromatographed using normal phase flash chromatography with hexanes/diethyl ether and hexanes/dichloromethane.

(6) NMR data for representative long-chain enol esters with phenylalanine-derived headgroups **6a**,**b** (cis:trans): ^{1}H NMR (500 MHz, CD_2Cl_2) δ 7.89 (d, J = 13, 1H), 7.62 (m, 2H), 7.30−7.38 (m, 7H), 7.22−7.28 (m, 2H), 6.41 (d, J = 13, 1H), 5.72 (d, J = 7, 1H), 2.54 (t, J = 7.5, 2H), 2.45 (t, J = 7.5, 2H), 1.66−1.75 (m, 4H), 1.24−1.42 (m, 48H), 0.91 (t, J = 7.5, 6H); ^{13}C NMR (100 MHz, CD_2Cl_2) δ 171.3, 170.8, 137.0, 134.9, 134.8, 134.6, 129.7, 129.3, 129.0, 127.9, 127.8, 126.7, 115.4, 112.0, 34.7, 34.5, 32.6, 30.3, 30.3, 30.2, 30.1, 30.0, 29.9, 29.7, 25.3, 25.3, 23.3, 14.5.

(7) Corey, E. J.; Venkateswarlu, A. *J. Am. Chem. Soc.* **1972**, *94*, 6190−6191.

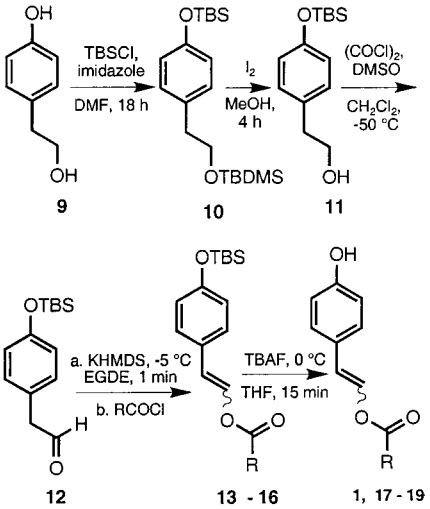
(8) Lipshutz, B. H.; Keith, J. *Tetrahedron Lett.* **1998**, *39*, 2495−2498.

(9) Mancuso, A. J.; Huang, S.-L.; Swern, D. *J. Org. Chem.* **1978**, *43*, 2480−2482.

(10) KHMDS (1.1 equiv, as a 0.5 M solution in THF) was added to 0.5 mmol of **12** in 100 μL of EGDE. After 1 min at −5 °C a chilled solution of acid chloride (2 equiv) dissolved in EGDE (160 μL) was added to the K enolate and allowed to stir at room temperature for 15 min. After the addition of 1 mL of water the reactions were extracted five times with pentane. The dried organic extracts were then chromatographed using normal phase flash chromatography with hexanes/ethyl acetate.

(11) NMR data for representative long-chain enol esters with tyrosine-derived headgroups **1a** (trans): ^{1}H NMR (500 MHz, CD_2Cl_2) δ 7.74 (d, J = 13, 1H), 7.22, (d, J = 9, 2H), 6.78 (d, J = 9, 2H), 6.33 (d, J = 13, 1H), 2.42 (t, J = 7.5, 2H), 1.66 (m, 2H), 1.2−1.4 (m, 24H), 0.88 (t, J = 6.5, 3H); ^{13}C NMR (100 MHz, CD_2Cl_2) δ 171.5, 155.8, 135.6, 128.1, 127.4, 116.1, 115.0, 34.5, 32.5, 30.3, 30.2, 30.2, 30.0, 29.9, 29.8, 29.6, 25.3, 23.3, 14.5. **1b** (cis): ^{1}H NMR (500 MHz, CD_2Cl_2) δ 7.49 (d, J = 9, 2H), 7.20 (d, J = 7.5, 1H), 6.81 (d, J = 9, 2H), 5.63 (d, J = 7.5, 1H), 2.52 (t, J = 7, 2H), 1.71 (m, 2H), 1.2−1.4 (m, 24H), 0.88 (t, J = 6.5, 3H); ^{13}CNMR (100 MHz, CD_2Cl_2) δ 170.9, 155.4, 133.1, 131.1, 127.6, 115.8, 111.5, 34.7, 32.5, 30.3, 30.2, 30.2, 30.0, 29.9, 29.8, 29.6, 25.3, 23.3, 14.4. **19a**,**b** (cis and trans): ^{1}H NMR (500 MHz, CD_2Cl_2) δ 7.73 (d, J = 13), 7.49 (d, J = 8.5), 7.20−7.23 (m), 6.81 (d, J = 8.5), 6.78 (d, J = 8.5), 6.33 (d, J = 13), 5.63 (d, J = 7), 2.52 (t, J = 8), 1.6−1.8 (m), 1.2−1.4 (m), 0.88 (t, J = 6.5); ^{13}CNMR (100 MHz, CD_2Cl_2) δ 171.5, 171.0, 155.8, 155.5, 135.5, 133.1, 131.1, 128.1, 127.6, 127.4, 116.1, 115.8, 115.1, 111.6, 34.7, 34.5, 32.4, 30.0, 29.8, 29.6, 25.3, 25.3, 23.2, 14.4.

Scheme 2



compounds similar to **1**. Using this approach we both confirmed the structure of the isolated product (**1**) and produced 12 analogues to examine the overall applicability of this strategy. A larger library is currently being made and will be screened in a number of bioassays in an attempt to define the role of this new family of long-chain enol esters. Using organic synthesis to understand the roles that biosynthetic small molecules play among uncultured soil bacteria is a unique contribution of chemistry to illuminating the unknown world of uncultured microorganisms.

Acknowledgment. We thank Professor David B. Collum for his prophetic and helpful suggestions. This work was supported by a grant from the Ellison Medical Foundation and CA24487.

Supporting Information Available: ^{1}H and ^{13}C NMR spectra for representative long-chain enol esters based on phenylalanine (**6a**,**b**) and tyrosine (**1a**,**b** and **19a**,**b**). This material is available free of charge via the Internet at http://pubs.acs.org.

OL0267681