Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters 18 (2008) 3117–3121
Bioorganic & Medicinal Chemistry Letters

Synthetic libraries of tyrosine-derived bacterial metabolites

Savvas N. Georgiades and Jon Clardy*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA

Received 19 September 2007; accepted 16 October 2007
Available online 22 October 2007

Abstract—The preparation of a collection of 131 small molecules, reminiscent of families of long chain *N*-acyl tyrosines, enamides and enol esters that have been isolated from heterologous expression of environmental DNA (eDNA) in *Escherichia coli*, is reported. The synthetic libraries of *N*-acyl tyrosines and their 3-keto counterparts were prepared via solid-phase routes, whereas the enamides and enol esters were synthesized in solution-phase.

Microbes, especially soil-dwelling bacteria, have made enormous contributions to our stock of biologically active small molecules.[1] Discovering these molecules usually requires isolating the producing organism, culturing it in the laboratory and assaying the culture extracts for biological activity. The realization that only a tiny and unrepresentative minority of soil and other microorganisms can be cultured by currently described techniques[2] led many laboratories to develop alternative strategies for accessing the small molecules produced by the uncultured majority.[3,4] Such processes typically involve obtaining DNA—not the producing organism—directly from the environment (thus termed environmental DNA or eDNA) and incorporating it into alternative hosts to discover gene-host combinations with the capacity to produce biologically active compounds. In our laboratory we have employed this approach to express eDNA-encoded pathways in *Escherichia coli*, and used an antibiotic assay to identify colonies producing small molecule antibiotics. The most frequently identified metabolites have been long chain *N*-acyl amino acids,[3b,5] especially *N*-acyl tyrosines (NATs), and the widespread occurrence of these compounds, which had not been previously reported as microbial natural products, raised questions about their biosynthesis and biological function(s).

One pathway uncovered using this approach produced NATs that were first converted to *N*-acyl (*E*)-enamides through an oxidative decarboxylation, and finally to *N*-acyl (*E*)-enol esters through an unusual *N*,*O*-exchange[5a] (Fig. 1a). Formation of the amide is catalyzed by an *N*-acyl synthase that couples free tyrosine to a long chain acid, delivered by an acyl carrier protein (ACP).[6] The structure, mechanism and sequence alignment of

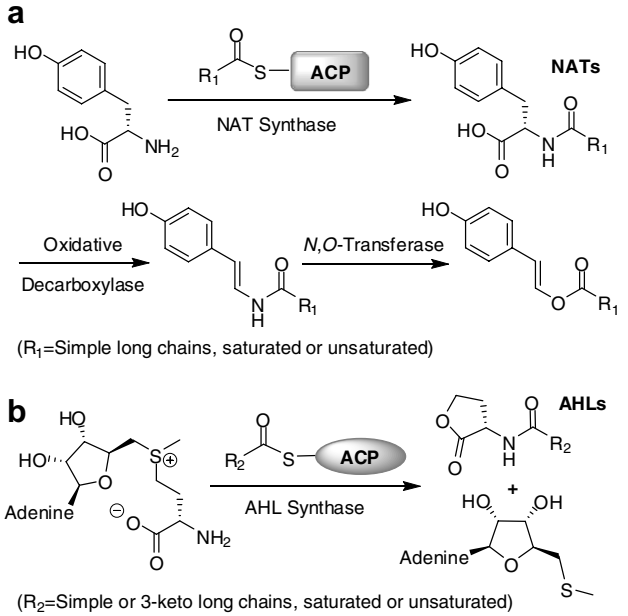


Figure 1. (a) Proposed biosynthetic pathway for the production of long chain *N*-acyl tyrosines and their conversion to (*E*)-enamides and (*E*)-enol esters. (b) Formation of acyl homoserine lactones from biosynthetic precursor *S*-adenosyl methionine (SAM).

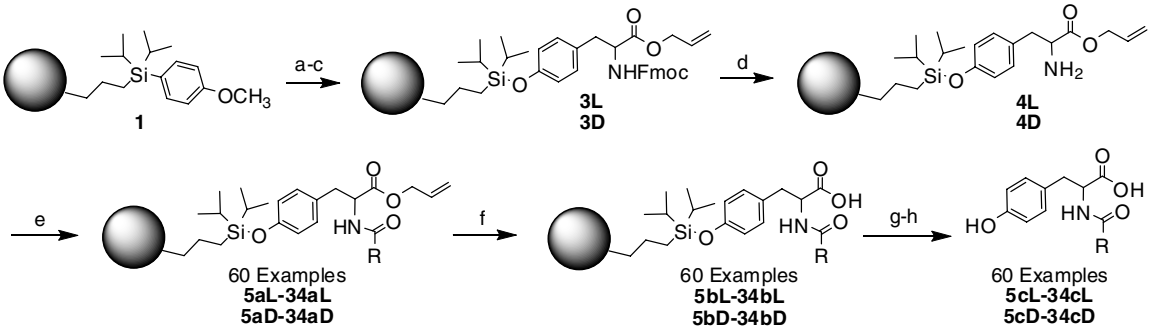
Keywords: Libraries; Bacterial metabolites; Antibiotics; Environmental DNA; Solid-phase synthesis; Microwave-assisted reactions; Acyl Meldrum's acids; Tyrosine; Long chain acids; *N*-Acyl amino acids; Enamides; Enol esters.
* Corresponding author. Tel.: +1 617 432 2845; fax: +1 617 738 3702; e-mail: jon_clardy@hms.harvard.edu

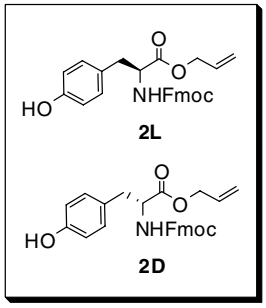
0960-894X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2007.10.058

this *N*-acyl synthase[5b,6] suggest a relationship to the synthases that produce the acyl homoserine lactones (AHLs, Fig. 1b), quorum sensing mediators in many Gram-negative bacteria.[7] In both the AHLs and NATs, a common head group—either homoserine lactone or tyrosine—is attached to a variety of long chain acids, and biological activity requires the correct acid fragment. In the *E. coli* heterologous expression system we use, ACP-bound long chain acids from *E. coli* pathways can replace those of the original producer, and consequently the identities of the biologically relevant products of NAT pathways are not known with certainty.[5a] In order to explore the relationship between acid structure and biological activity in the amide, enamide and enol ester pathway, we developed an efficient chemical synthesis of plausible library members, which is described herein.

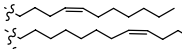
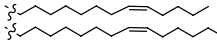
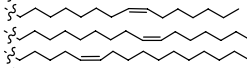
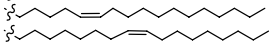
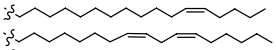
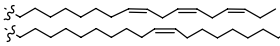
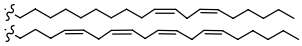
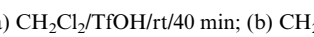
Long chain NATs were prepared via a solid-phase route (Scheme 1). Polystyrene macrobeads of 500–600 μm diameter, functionalized with a trialkylaryl silicon linker (**1**), were chosen as the solid support.[8] First, the resin was activated with excess trifluoromethanesulfonic acid in CH_2Cl_2, followed by quenching with 2,6-lutidine. The activated beads were then treated for 24 h with a saturated solution of a bis-protected tyrosine substrate, the Fmoc carbamate/allyl ester of either L- or D-tyrosine (**2L** and **2D**, respectively),[9] to afford the corresponding resin-bound species (**3L** and **3D**).[10] Complete removal of the Fmoc group with 20% (v/v) piperidine in DMF led to the resin-bound allyl esters **4L** and **4D**. After screening various amide coupling conditions, we elected to carry out this reaction in THF using diethyl phosphocyanidate (DEPC) as the activating agent in the presence of TEA, for 24 h at room temperature. Excess of acid, DEPC and TEA was required. Under these conditions, 30 simple acids with various degrees of unsaturation (**5**–**34**) were combined with the resin-bound nucleophiles in parallel to furnish 60 coupling products (**5aL**–**34aL** and **5aD**–**34aD**). All couplings were quantitative and free of impurities, as indicated by LC–MS. Parallel removal of the allyl protecting group from these intermediates was achieved with 5 mol % $Pd(PPh_3)_4$ in THF and excess morpholine in 6–8 h. The mild deprotection conditions were compatible with the presence of unsaturated long chains. This sensitive reaction required absence of light and initial degassing of the reaction container, to avoid catalyst degradation. Cleavage of the NATs from the beads was carried out using a 70/30 (v/v) HF/pyridine mixture in THF, buffered with additional pyridine. The reaction was typically quenched with TMSOMe, to provide >95% pure products (**5cL**–**34cL** and **5cD**–**34cD**) after the supernatant was separated from the beads and the volatile byproducts were evaporated.

To augment the library with long chains reminiscent of AHLs, solid-phase synthesis of NATs bearing a 3-keto functionality was carried out using a modification of the above procedure (Scheme 2). Long chain 5-acyl-2,2-dimethyl-1,3-dioxane-4,6-diones (5-acyl Meldrum's acids),[11] a masked form of the labile 3-oxo carboxylic acids, were used as building blocks in the amide coupling step. An alternative to conventional heating in

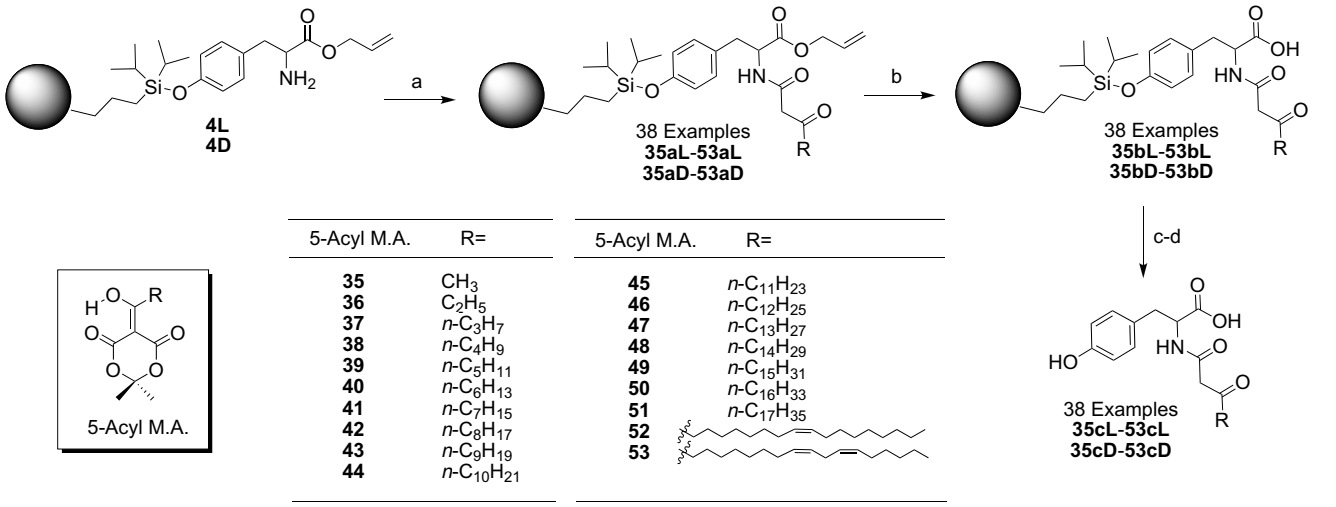




RCO_2H	R=
5	n-C_3H_7
6	n-C_4H_9
7	n-C_5H_{11}
8	n-C_6H_{13}
9	n-C_7H_{15}
10	n-C_8H_{17}
11	n-C_9H_{19}
12	n-$C_{10}H_{21}$
13	n-$C_{11}H_{23}$
14	n-$C_{12}H_{25}$
15	n-$C_{13}H_{27}$
16	n-$C_{14}H_{29}$
17	n-$C_{15}H_{31}$
18	n-$C_{16}H_{33}$
19	n-$C_{17}H_{35}$

RCO_2H	R=
20	n-$C_{18}H_{37}$
21	n-$C_{19}H_{39}$
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	

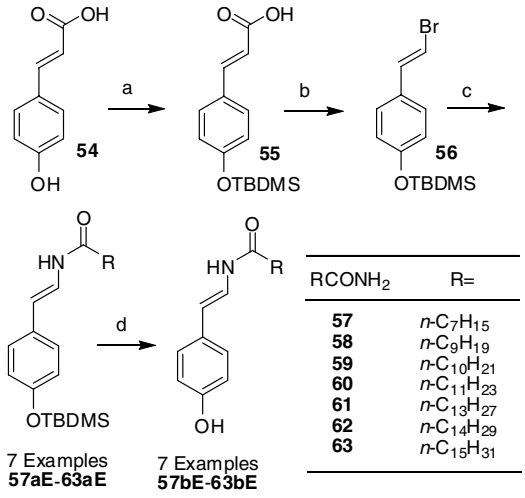
Scheme 1. Solid-phase synthesis of long chain *N*-acyl tyrosines. Reagents and conditions: (a) CH_2Cl_2/TfOH/rt/40 min; (b) CH_2Cl_2/2,6-lutidine/rt/5–10 min; (c) CH_2Cl_2/**2L** or **2D**/rt/24 h; (d) DMF/piperidine/rt/4 h; (e) RCO_2H (**5**–**34**)/THF/DEPC/TEA/rt/24 h; (f) THF/$Pd(PPh_3)_4$/morpholine/rt/8–12 h; (g) THF/HF/pyridine/rt/2.5 h; (h) TMSOMe/rt/15 min.



Scheme 2. Solid-phase synthesis of long chain 3-oxo-*N*-acyl tyrosines. Reagents and conditions: (a) NMP/5-acyl Meldrum's acid (**35**–**53**)/microwave/200 °C/10 min; (b) THF/$Pd(PPh_3)_4$/morpholine/rt/8–12 h; (c) THF/HF/pyridine/rt/2.5 h; (d) TMSOMe/rt/15 min.

the presence of an external base, which is the usual protocol of choice for a solution-phase reaction involving nucleophilic attack on a 5-acyl Meldrum's acid, was necessary in order to avoid bead degradation. Thus a microwave-assisted method was developed. Optimal reaction conditions required the use of *N*-methyl pyrrolidone (NMP) as the solvent, which has both excellent absorbance characteristics and ideal interaction with this solid-phase, excess of the 5-acyl Meldrum's acid building block and no external base, at 200 °C with controlled microwave irradiation in a commercial reactor for 10 min. Parallel coupling of 19 different saturated and unsaturated building blocks (**35**–**53**) to resin-bound substrates **4L** and **4D** led to full consumption of the substrate in all cases, and formation of 70–80% of the desired products (**35aL**–**53aL** and **35aD**–**53aD**). The 38 intermediates were successfully carried through the allyl deprotection and cleavage steps as described above, to yield 38 final 3-oxo-*N*-acyl tyrosines (**35cL**–**53cL** and **35cD**–**53cD**).

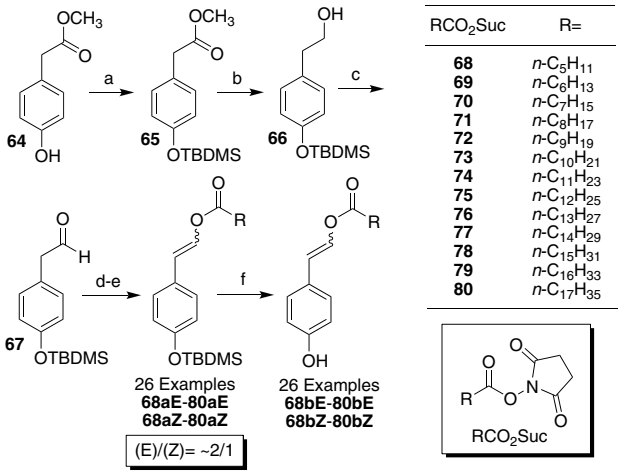
Since our devised synthetic routes leading to the enamide and enol ester families involved steps that were either heterogeneous or incompatible with the solid-phase, both were realized in solution-phase. The (*E*)-enamides were prepared as shown in Scheme 3, starting with protection of 4-hydroxycinnamic acid (**54**) as a silyl ether (**55**, 85% yield) with TBDMSCl in DMF, in the presence of imidazole. Intermediate **55** was stereoselectively converted to the corresponding (*E*)-vinyl bromide (**56**, 83% yield) via a decarboxylative bromination with *N*-bromosuccinimide.[12] This LiOAc-catalyzed Hunsdiecker transformation was carried out in CH_3CN/H_2O with mild heating. A modification of Buchwald's conditions for copper-catalyzed amidation[13] was employed for coupling of the vinyl bromide to a series of saturated long chain carboxamides (**57**–**63**),[14] generating the long chain (*E*)-enamides **57aE**–**63aE** at 65–70% yield with retention of the trans geometry. The amidation was carried out in toluene, using CuI as the copper source, *N*,*N*′-dimethylethylenediamine as the bidentate ligand



Scheme 3. Solution-phase synthesis of long chain *N*-acyl enamides. Reagents and conditions: (a) TBDMSCl/imidazole/DMF/rt/12 h/85%; (b) NBS/LiOAc/CH_3CN/H_2O/60 °C/1 h/83%; (c) $RCONH_2$ (**57**–**63**)/$(CH_3NHCH_2)_2$/CuI/Rb_2CO_3/toluene/80 °C/24 h/65–70%; (d) HF/pyridine/THF/rt/6 h/85–90%.

and Rb_2CO_3 as the base,[13] and required strictly anhydrous conditions and heating at 70 °C for 24 h. Deprotection of the intermediates was achieved with 70/30 (v/v) HF/pyridine in THF/pyridine and afforded 85–90% of the final (*E*)-enamides (**57bE**–**63bE**).

The last class of eDNA-associated compounds, the enol esters, were derived from methyl 4-hydroxyphenyl acetate (**64**, Scheme 4), which was quantitatively protected under conditions similar to the protection of **54** above to afford TBDMS-silyl ether **65**. Intermediate **65** was converted to an aldehyde in two steps: reduction to an alcohol (**66**) with $LiAlH_4$ in THF at −78 °C (94% yield), followed by alcohol oxidation to the aldehyde (**67**) with 2-iodoxybenzoic acid (IBX)[15] in DMSO (93% yield). By means of KHMDS in THF/toluene, **67** was converted to a mixture of trans and cis potassium enolates, in accord



Scheme 4. Solution-phase synthesis of long chain *N*-acyl enol esters. Reagents and conditions: (a) TBDMSCl/imidazole/DMF/rt/12 h/100%; (b) $LiAlH_4$/THF/−78 °C/1 h/94%; (c) IBX/DMSO/rt/4 h/93%; (d) KHMDS/toluene/THF/0 °C/5 min; (e) THF/RCO_2Suc (**68**–**80**)/rt/30 min/30–40%; (f) HF/pyridine/rt/6 h/85–90%.

with a previous report from our laboratory for a similar system.[16] Under the specific reaction conditions employed here, moderate selectivity for the trans enolate (precursor to the only isomeric product present in the natural extract from the eDNA clone) was observed. The enolates were trapped with saturated long chain *N*-hydroxysuccinimide esters (**68**–**80**).[17] LC–MS analysis of an aliquot from each reaction revealed in all cases a mixture of products with the same mass, that included (*E*)- and (*Z*)-enol esters as well as what appeared to be carbon acylation products. Partial purification of the crude products afforded inseparable mixtures of (*E*)- and (*Z*)-enol esters (**68aE**–**80aE** and **68aZ**–**80aZ**), for which ^1H NMR indicated a ratio of approximately 2:1. The ratio was independent of acid chain length and reflects the thermodynamic preference for the trans enolate. The (*E*)/(*Z*) mixtures were resolved after HF removal of the TBDMS protecting group in pyridine, which proceeded at 85–90% yield. Thirteen deprotected (*E*)-enol esters (**68bE**–**80bE**) and 13 (*Z*)-enol esters (**68bZ**–**80bZ**) were obtained in pure form.

In summary, solid- and solution-phase methods are described for the preparation of synthetic libraries of tyrosine-derived bacterial metabolites (131 compounds have been delivered), resembling small molecules isolated from heterologous expression of eDNA in *E. coli*. Preliminary biological studies on library members have shown antibiotic activity against *Bacillus subtilis* and moderate inhibitory potential of *Pseudomonas aeruginosa* biofilm formation. Additional assays are planned and will be reported elsewhere in due course.

Acknowledgments

We gratefully acknowledge the NIH for funding of this project (CA24487 to J.C.). We thank Dr. Sean Brady (formerly HMS) and Dr. Lauren Junker (HMS) for helpful discussions and preliminary biological studies on the library, Dr. Li Lai (formerly HMS) for assistance with solid-phase handling, and Dr. Ralph Mazitschek (Broad Institute of MIT and Harvard) and Dr. Jared Shaw (formerly Broad Institute) for useful suggestions.

Supplementary data

Detailed experimental procedures and spectroscopic data associated with this article can be found in the online version, at doi:10.1016/j.bmcl.2007.10.058.

References and notes

1. (a) Clardy, J.; Fischbach, M. A.; Walsh, C. T. *Nat. Biotechnol.* **2006**, *24*, 1541; (b) Clardy, J.; Walsh, C. *Nature* **2004**, *432*, 829; (c) Koehn, F. E.; Carter, G. T. *Nat. Rev. Drug Discov.* **2005**, *4*, 206; (d) Butler, M. S. J. *Nat. Prod.* **2004**, *67*, 2141; (e) Leeds, J. A.; Schmitt, E. K.; Krastel, P. *Expert Opin. Investig. Drugs* **2006**, *15*, 211; (f) Pelaez, F. *Biochem. Pharmacol.* **2006**, *71*, 981.
2. (a) Torsvik, V.; Salte, K.; Sørheim, R.; Goksøyr, J. *Appl. Environ. Microbiol.* **1990**, *56*, 776; (b) Torsvik, V.; Goksøyr, J.; Daae, F. L. *Appl. Environ. Microbiol.* **1990**, *56*, 782; (c) Ward, D. M.; Weller, R.; Bateson, M. M. *Nature* **1990**, *345*, 63; (d) Stackebrandt, E.; Liesack, W.; Goebel, B. M. *FASEB J.* **1993**, *7*, 232; (e) Amann, R. I.; Ludwig, W.; Schleifer, K. H. *Microbiol. Rev.* **1995**, *59*, 143; (f) Hugenholtz, P.; Goebel, B. M.; Pace, N. R. *J. Bacteriol.* **1998**, *180*, 4765.
3. (a) Wang, G.-Y.-S.; Graziani, E.; Waters, B.; Pan, W.; Li, X.; McDermott, J.; Meurer, G.; Saxena, G.; Andersen, R. J.; Davies, J. *Org. Lett.* **2000**, *2*, 2401; (b) Brady, S. F.; Clardy, J. *J. Am. Chem. Soc.* **2000**, *122*, 12903; (c) MacNeil, I. A.; Tiong, C. L.; Minor, C.; August, P. R.; Grossman, T. H.; Loiacono, K. A.; Lynch, B. A.; Phillips, T.; Narula, S.; Sundaramoorthi, R.; Tyler, A.; Aldredge, T.; Long, H.; Gilman, M.; Holt, D.; Osburne, M. S. J. *Mol. Microbiol. Biotechnol.* **2001**, *3*, 301.
4. A similar approach has been used for enzyme catalyst discovery: DeSantis, G.; Zhu, Z.; Greenberg, W. A.; Wong, K.; Chaplin, J.; Hanson, S. R.; Farwell, B.; Nicholson, L. W.; Rand, C. L.; Weiner, D. P.; Robertson, D. E.; Burk, M. J. *J. Am. Chem. Soc.* **2002**, *124*, 9024.
5. (a) Brady, S. F.; Chao, C. J.; Clardy, J. *J. Am. Chem. Soc.* **2002**, *124*, 9968; (b) Brady, S. F.; Chao, C. J.; Clardy, J. *J. Appl. Environ. Microbiol.* **2004**, *70*, 6865; (c) Brady, S. F.; Clardy, J. *J. Nat. Prod.* **2004**, *67*, 1283; (d) Brady, S. F.; Clardy, J. *Org. Lett.* **2005**, *7*, 3613.
6. (a) Van Wagoner, R. M.; Clardy, J. *Structure* **2006**, *14*, 1425; Commentary by (b) Churchill, M. E. A. *Structure* **2006**, *14*, 1342.
7. For recent reviews on AHLs and bacterial signalling, see: (a) Waters, C. M.; Bassler, B. L. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 319; (b) Bassler, B. L.; Losick, R. *Cell* **2006**, *125*, 237.
8. Preparation of this resin is described in: Tallarico, J. A.; Depew, K. M.; Pelish, H. E.; Westwood, N. J.; Lindsley, C. W.; Shair, M. D.; Schreiber, S. L.; Foley, M. A. *J. Comb. Chem.* **2001**, *3*, 312.
9. Substrates 2L and 2D were prepared from commercial Fmoc-tyrosine following the method of: Jensen, K. J.; Meldal, M.; Bock, K. *J. Chem. Soc. Perkin Trans. 1* **1993**, *17*, 2119.
10. The loading in this step was typically 1.3–1.4 mmol of substrate per gram of resin, approximately equal to the

empirical maximum for this type and batch of resin. It was determined by UV analysis as explained in the supplementary data.
11. The long chain 5-acyl Meldrum's acids were prepared from carboxylic acids and Meldrum's acid, as detailed in the supplementary data, using a modification of a published method: Raillard, S. P.; Chen, W.; Sullivan, E.; Bajjalieh, W.; Bhandari, A.; Baer, T. A. *J. Comb. Chem.* **2002**, *4*, 470.
12. (a) Kuang, C.; Yang, Q.; Senboku, H.; Tokuda, M. *Synthesis* **2005**, *8*, 1319; (b) Kuang, C.; Senboku, H.; Tokuda, M. *Synlett* **2000**, *10*, 1439.
13. Jiang, L.; Job, G. E.; Klapars, A.; Buchwald, S. L. *Org. Lett.* **2003**, *5*, 3667.
14. The carboxamide building blocks were prepared from fatty acids and NH_3 with DEPC in 1,4-dioxane/DMF, as described in the supplementary data.
15. For an efficient preparation of IBX from 2-iodobenzoic acid, see: Frigerio, M.; Santagostino, M.; Sputore, S. *J. Org. Chem.* **1999**, *64*, 4537.
16. Brady, S. F.; Clardy, J. *Org. Lett.* **2003**, *5*, 121.
17. The *N*-hydroxysuccinimide ester building blocks were prepared from fatty acids and *N*-hydroxysuccinimide with DCC in THF, as described in the supplementary data.