J|A|C|S
COMMUNICATIONS

Published on Web 10/18/2002

Skeletal Diversity via a Branched Pathway: Efficient Synthesis of 29 400 Discrete, Polycyclic Compounds and Their Arraying into Stock Solutions

Ohyun Kwon,† Seung Bum Park, and Stuart L. Schreiber*

Department of Chemistry and Chemical Biology, Howard Hughes Medical Institute, Harvard Institute of Chemistry and Cell Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received August 9, 2002

Diversity-oriented synthesis (DOS) aims to synthesize efficiently complex, small molecules broadly distributed in multidimensional descriptor space.[1] Such collections[2] are key to chemical genetics, where small molecules are used to explore biology and medicine systematically.[3] Skeletal diversity in DOS has proven to be especially challenging. Here, we report a branching DOS pathway that yields 29 400 discrete compounds comprising 10 distinct polycyclic skeletons.[4]

The six-step, stereoselective synthesis, which affords products having a central skeleton with between two and four rings and up to six stereocenters, has been achieved using an inexpensive and accessible, "one bead-one stock solution" technology platform.[5] The pathway builds on the report by Fallis and co-workers on the use of consecutive Diels−Alder reactions.[6] We have adapted their reported triene synthesis and subsequent complexity-generating reactions to phenolic aldehyde-loaded macrobeads and discovered a set of dienophiles that react only once with the Fallis-type trienes. The latter observation provides a branch point to the pathway, where diene products are formed from a single Diels−Alder cycloaddition, and monoene products are formed from consecutive Diels−Alder reactions involving either the same or different dienophiles (Figure 1). An important feature of the branched pathway is that the diastereoselection observed in the original report has been extended to reaction sequences involving different dienophiles.

To optimize the yield and purity of the library members, potential building blocks for the library were tested individually as follows. In separate reaction vessels, 64 hydroxyaldehydes were loaded onto macrobeads through silylation of their hydroxyl groups with the previously described macrobead-alkylsilyl triflate (illustrated with the silylation-loading of vanillin **1** in Figure 1).[7] Each macrobead-loaded aldehyde was separately reacted with indium dust[8] and 5-bromo-1,3-pentadiene[9] in DMF, which provided the γ-addition product (**2** in the illustrated case with vanillin).[10] Mesylation followed by elimination using DBU furnished the cross-conjugated triene (cf., **3**).[10] After cleavage with HF-py and analysis of the purity of the triene products by ^{1}H NMR, 40 of the original 64 hydroxyaldehydes (Figure 2 top) were found to yield a single identifiable compound.[10] These 40 aldehydes were used in the DOS pathway described below.

Macrobead-loaded triene **3** (Figure 1) was used to assess the reactivity and stereoselectivity of 53 disubstituted- and 44 tri- or tetrasubstituted cyclic dienophiles. In earlier pathway-development studies, we had ascertained that noncyclic dienophiles[11] afforded stereoisomeric mixtures of double cycloadducts, whereas cyclic dienophiles yielded products stereoselectively.[10] Spectroscopic analyses of single and double cycloadducts, including X-ray crystallography in five cases, verified that the selectivity reported by Fallis and co-workers was general.[12]

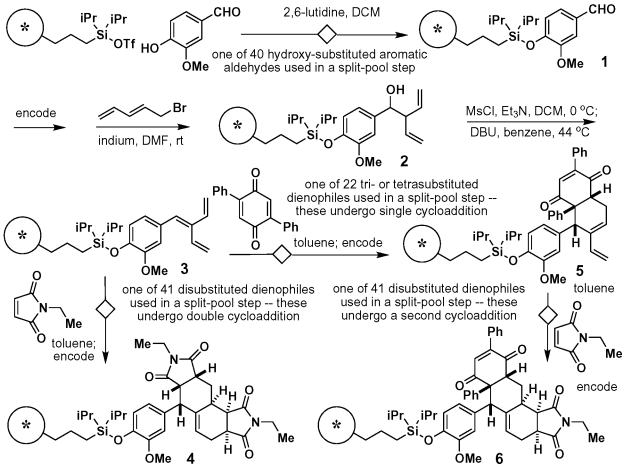

Figure 1. Hydroxyl-substituted aromatic aldehydes (most frequently, phenolic aldehydes; vanillin is illustrated) were loaded onto high capacity macrobeads (denoted by the asterisk-within-a-circle symbol), converted to trienes, and reacted with dienophiles. The degree of substitution on the dienophiles determines whether they participate in the second cycloaddition (see text for details). The diamond inserted in the arrow denotes a split-and-pool step.

An important pattern of reactivity was uncovered using **3**: disubstituted dienophiles underwent double cycloaddition (cf., **4**), whereas tri- or tetrasubstituted dienophiles underwent mono cycloaddition (cf., **5**). Using the criterion of single isomer formation (cf., **4**) in high purity from triene **3**, we selected 41 (of 53) disubstituted dienophiles (Figure 2, middle) for use in the DOS pathway.[10] Representative members of mono-cycloadduct dienes (cf., **5**) were found to undergo stereoselective Diels−Alder reactions with a second dienophile to yield tetracycles derived from two different dienophiles (cf., **6**). Using the criteria of efficient, single isomer production of both single and double cycloadducts (cf., **5** and **6**), we selected 22 (of 44) tri- or tetrasubstituted dienophiles (Figure 2, bottom) for use in the DOS pathway.[10] These dienophiles, which "interrupt" the double Diels−Alder process, provide a key skeleton-diversifying branch in the DOS pathway. Combinations of the selected skeletal building blocks are calculated to produce a maximum of 29 400 distinct compounds.[13]

Approximately 88 200 macrobeads[15] were divided into 40 equal portions and loaded with the 40 aldehydes described above in separate reaction vessels. The individual vessels of aldehyde-loaded macrobeads (cf., **1**) were tagged with diazo-based electrophoretic reporters using a binary code,[16] pooled, and converted to triene-loaded macrobeads as described above (cf., **3**). The tagged and pooled triene-containing macrobeads were divided into 23 portions.

* To whom correspondence should be addressed. E-mail: sls@slsiris.harvard.edu.
† Current address: University of California, Los Angeles.

10.1021/ja028086e CCC: $22.00

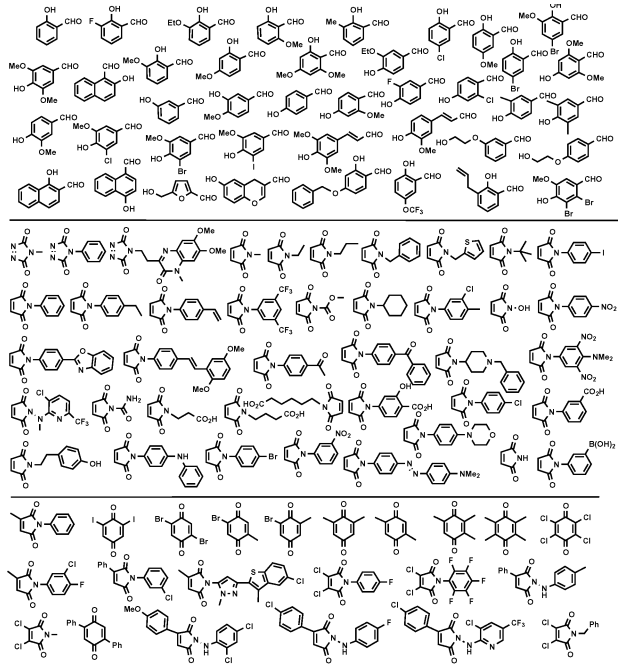

Figure 2. Forty hydroxyaldehyde- (top), 41 disubstituted dienophile- (middle), and 22 tri- or tetrasubstituted dienophile- (bottom) building blocks used in the branched DOS pathway.

One portion was recombined later with the dienes prepared below for the consecutive Diels−Alder cycloaddition (cf., **3** to **4**). Twenty-two portions were reacted individually, using optimized conditions, with 22 dienophiles (Figure 2, bottom). Each segregated collection of macrobeads was tagged using additional reporters. The 22 vessels containing single cycloadducts (cf., **5**) were pooled and subsequently divided into 4 × 41 portions (the 41 portions were grouped into four sets because we found that the subsequent Diels−Alder reactions fell into four different optimal reaction conditions depending on the reactivity of the second dienophile[10]). A 42nd portion was set aside to be combined with the collection of tetracyclic compounds, thus ensuring the presence of bicyclic dienes (cf., **5**) in the final collection of products. The 4 × 41 vessels were treated individually with the 41 dienophiles (using four different conditions) that undergo the second cycloaddition (Figure 2, middle) and tagged using additional reporters. The pooled, 88 200 encoded macrobeads serve to segregate a high percentage of the theoretical 29 400 compounds prior to automated preparation of stock solutions.

Our quality control efforts during the pathway development phase of this research identified the reaction partners expected to undergo efficient and predictable outcomes, but they also revealed reactivity patterns that further diversified the skeletons[4] of the products of this DOS pathway (Figures 3 and 4). Whereas macrobead-bound trienes (cf., **3**) reacted with tri- and tetrasubstituted dienophiles to yield the expected bicycles of structural types **S1** and **S2** (verified in **10** and **8**), they reacted with halogenated dienophiles to yield structural types **S3**, **S9**, and **S10**. These latter compounds result from cycloadditions followed by dehydrohalogenation: **S3** (verified in **11**) by dehydroiodination[10] and **S9**−**10** (verified in **13**) by dehydrobromination (dehydrohalogenation was facilitated with strontium carbonate[17] when maleimides were used as the second dienophile). Macrobead-bound dienes of **S1** react with maleimides to yield the expected tetracycle of **S4** (verified in **7′**), but they react with 4-phenyl-1,2,4-triazoline-3,5-dione (and presumably related dienophiles) to yield products having anti,anti- and syn,anti-trans-fused C−D ring junctions as in **S5** and **S10** (verified in **12** and **13**). Extending these observations to the possible combinations of

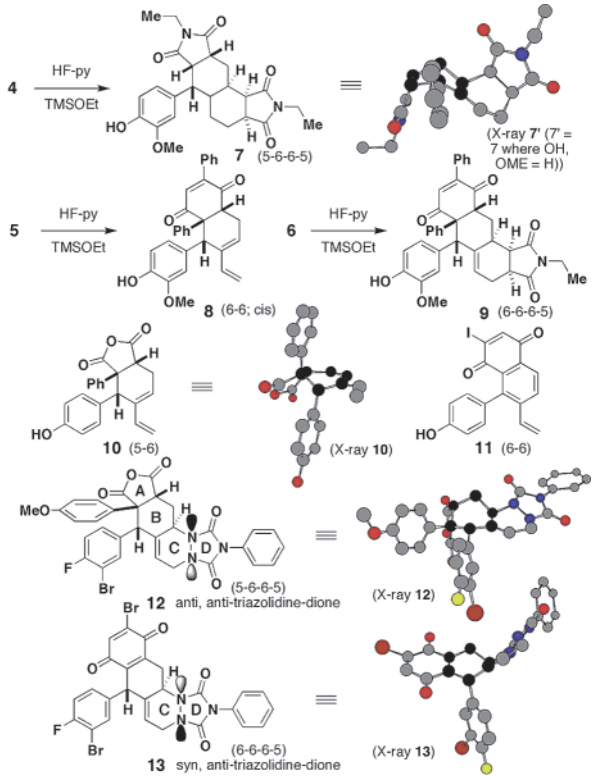

Figure 3. Products derived from the intermediates in Figure 1 and several related products characterized by X-ray crystallography. Ring B on each skeleton is highlighted in black in the Chem 3D images derived from X-ray coordinates.[14]

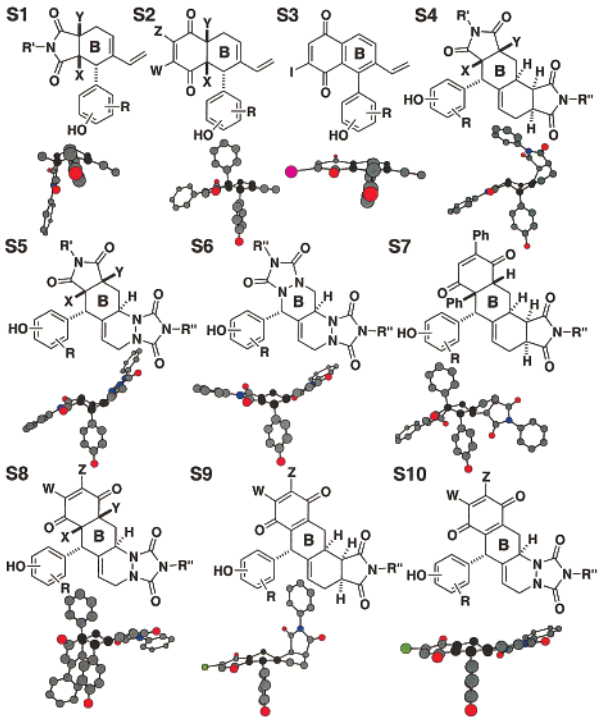

Figure 4. The branched DOS pathway leads to compounds having 10 distinct skeletons.[4]

dienophile building blocks suggests that at least 10 different skeletons[4] will be represented among the 29 400 anticipated products.

Our first step in analyzing purity and identity of these products entailed the random selection of 50 macrobeads from the final pool.

Products were eluted from the macrobeads with HF-py (and then TMSOEt), diluted to 10 mM stock solutions (DMF), and analyzed by LC/MS and stock solution decoding.[18] These data revealed acceptable levels of purity and structures consistent with expectations. Our second step in postsynthesis quality control was performed following both full arraying of all macrobeads and automated stock solution preparation.

The 88 200 individual macrobeads were first arrayed into 384-well microtiter plates using a vacuum-based bead arrayer to entrain 352 beads in an equal number of wells (two columns of wells from each plate were left empty to accommodate controls used in subsequent assays).[5a] Microtiter plates containing one bead per well were then subjected to a robotic cleavage process, in which each well was treated with 20 μL of HF-py cocktail (5% HF-py, 5% py in THF) delivered using a ceramic pump. After 300 min at room temperature, each cleavage reaction was quenched with 20 μL of TMSOEt[19] for 30 min, evaporated, and eluted from beads with three 30 μL DMF washes. DMF eluates were pooled into fresh 384-well "mother plates", each of which was mapped into five "daughter plates" by volumetric transfer using a Hydra384 syringe-array robot (50% of stock solution for cell-based assays, 20% for small molecule microarrays 2 × 10% for compound archiving, and 10% for chemical analysis).[5b]

Currently, 150 microtiter plates (52 800 single compound-containing stock solutions, approximately two theoretical copies) have been arrayed, and 61 microtiter plates (21 472 compounds, 73% of a theoretical copy) have been formatted into "daughter plates". For post-automated formatting, quality control (QC) analysis, we again used LC/MS and stock solution decoding.[18b] The structures of 88 out of 100 samples were inferred successfully by LC/MS and GC decoding. The structures of the remaining 12 were inferred by GC decoding, but could not be confirmed by LC/MS.

Preliminary analysis of the purity of resulting stock solutions and their performance in both protein-binding and phenotypic assays has revealed that the overall process is sufficient for identifying novel small molecules having specific and potent protein-binding and cellular activities. We expect that the pathway should be subject to further development and optimization, including modified pathways guided by analyses of the molecular descriptors of the small molecules in advance of their synthesis. We are optimistic that this pathway will provide many effective probes for chemical genetic studies aimed at dissecting biology.

Acknowledgment. We thank Richard Staples for performing X-ray crystallographic analysis and Cullen M. Taniguchi for assistance in the early phase of these studies. We are especially grateful to John Tallarico, Paul Clemons, and Max Narovlyansky of the Harvard ICCB for their assistance. We thank the National Institute for General Medical Sciences for support of this research, the National Cancer Institute, Merck KGaA, Merck & Co., and the Keck Foundation for support of the Harvard Institute of Chemistry and Cell Biology, and the National Cancer Institute for support of the Molecular Target Laboratory. O.K. was and S.B.P. is a Research Associate, and S.L.S. is an Investigator at the Howard Hughes Medical Institute in the Department of Chemistry and Chemical Biology, Harvard University.

Supporting Information Available: Representative experimental procedures and characterization data (PDF). An X-ray crystallographic file (CIF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Schreiber, S. L. *Science* **2000**, *287*, 1964−1969.
(2) Dolle, R. E. *J. Comb. Chem.* **2001**, *3*, 477−517.
(3) (a) Schreiber, S. L. *Bioorg. Med. Chem.* **1998**, *6*, 1127−1152. (b) Mitchison, T. J. *Chem. Biol.* **1994**, *1*, 3−6.
(4) In this paper, we define compounds with different: (1) numbers or sizes of rings, (2) ring fusion stereochemistry, or (3) degree of ring fusion saturation as having "different skeletons".
(5) (a) Blackwell, H. E.; Perez, L.; Stavenger, R. A.; Tallarico, J. A.; Eatough, E. C.; Foley, M. A.; Schreiber, S. L. *Chem. Biol.* **2001**, *8*, 1167−1182. (b) Clemons, P. A.; Koehler, A. N.; Wagner, B. K.; Sprigings, T. G.; Spring, D. R.; King, R. W.; Schreiber, S. L.; Foley, M. A. *Chem. Biol.* **2001**, *8*, 1183−1195.
(6) Woo, S.; Squires, N.; Fallis, A. G. *Org. Lett.* **1999**, *1*, 573−575.
(7) Tallarico, J. A.; Depew, K. D.; Pelish, H. E.; Westwood, N. J.; Lindsley, C. W.; Shair, M. D.; Schreiber, S. L.; Foley, M. A. *J. Comb. Chem.* **2001**, *3*, 312−318.
(8) For indium-mediated allylation of resin-bound aldehydes with sonication: Cavallaro, C. L.; Herpin, T.; MuGuinness, B. F.; Shimshock, Y. C.; Dolle, R. E. *Tetrahedron Lett.* **1999**, *40*, 2711−2714.
(9) Prevost, C.; Miginiac, P.; Miginiac-Groizeleau, L. *Bull. Soc. Chim. Fr.* **1964**, 2485−2492.
(10) See Supporting Information for details.
(11) Acyclic dienophiles tested: *trans*-β-nitrostyrene, dimethyl maleate, and dimethyl fumarate.
(12) HF-py-mediated cleavage of macrobead-loaded **4** resulting from 100 mg of 3-[diisopropyl(*p*-methoxyphenyl)silyl]propyl functionalized macrobeads yielded 32 mg (0.71 mmol/g of beads, 109 nmol/bead) of the tetracyclic product **7** (Figure 3) (single diastereomer and 95% pure by ^{1}H NMR).
(13) 800 dienes (40 aldehydes × 20 dienophiles), 2640 tetracycles from interrupted D−A with 1,2,4-triazoline-3,5-diones (40 aldehydes × 22 dienophiles × 3 disubstituted dienophiles), 24 320 tetracycles from interrupted D−A with maleimides (40 aldehydes × 16 dienophiles × 38 disubstituted dienophiles), and 1640 tetracycles from consecutive D−A (40 aldehydes × 41 disubstituted dienophiles).
(14) Compounds **7′**, **12**, and **13** were produced through solution-phase synthesis during the pathway development phase of this research.[10]
(15) Burgess, K.; Liaw, A. I.; Wang, N. *J. Med. Chem.* **1994**, *37*, 2985−2987.
(16) (a) Ohlmeyer, M. J. H.; Swanson, R. N.; Dillard, L. W.; Reader, J. C.; Asouline, G.; Kobayashi, R.; Wigler, M.; Still, W. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10922−10926. (b) Nestler, H. P.; Bartlett, P. A.; Still, W. C. *J. Org. Chem.* **1994**, *59*, 4723−4724.
(17) Pearlman, B. A.; McNamara, J. M.; Hasan, I.; Hatakeyama, S.; Sekizaki, H.; Kishi, Y. *J. Am. Chem. Soc.* **1981**, *103*, 4248−4251.
(18) (a) Stavenger, R. A.; Schreiber, S. L. *Angew. Chem., Int. Ed.* **2001**, *40*, 3417−3421. (b) Blackwell, H. E.; Perez, L.; Schreiber, S. L. *Angew. Chem., Int. Ed.* **2001**, *40*, 3421−3425.
(19) TMSOEt was used as the quenching reagent instead of previously reported TMSOMe to minimize cross-contamination in the 386-well microtiter plate.

JA028086E