Natural Product-like Combinatorial Libraries Based on Privileged Structures. 2. Construction of a 10 000-Membered Benzopyran Library by Directed Split-and-Pool Chemistry Using NanoKans and Optical Encoding

K. C. Nicolaou,* J. A. Pfefferkorn, H. J. Mitchell, A. J. Roecker, S. Barluenga, G.-Q. Cao, R. L. Affleck,† and J. E. Lillig†

Contribution from the Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, and IRORI, a Division of Discovery Partners International Company, 9640 Towne Centre Drive, San Diego, California 92121

Received June 7, 2000

Abstract: Having developed a reliable and versatile solid-phase strategy for the split-and-pool synthesis of naturally occurring and designed derivatives of the benzopyran template (see preceding paper), we now report the construction of a 10 000-membered natural product-like compound library for chemical biology studies. Concomitantly, we report an early application of the IRORI NanoKan optical encoding system for the high throughput nonchemical tagging and sorting of library members during split-and-pool synthesis. The overall synthetic strategy for library construction is discussed and the individual reaction pathways are examined in the context of specific library members, illustrating reaction conditions as well as yields and purities. The issues of building block selection and quality control of library members are also addressed and, finally, potential applications of the library to chemical biology are discussed.

Introduction

In this series of papers, we introduce a strategy for the design and synthesis of small molecule libraries based on naturally occurring privileged structural motifs.[1,2] To illustrate this strategy, we targeted the 2,2-dimethylbenzopyran moiety as an inaugural natural product-like template. Having already developed a reliable and versatile solid-phase strategy for the split-and-pool synthesis of naturally occurring and designed derivatives of this scaffold,[1] we now report the construction of a 10 000-membered natural product-like compound library for chemical biology studies. Concomitantly, we report the first published application of the IRORI NanoKan optical encoding system for the high-throughput nonchemical tagging and sorting of library members during split-and-pool synthesis.[3]

We will begin by examining the design of the overall synthetic strategy and then discuss each reaction pathway in the context of a specific library member illustrating reaction conditions as well as yields and purities of cleavage products. The selection of library building blocks will also be addressed. In addition, the integration of this chemistry with the new nonchemical encoding technology is described along with its application to the high-throughput synthesis, sorting, cleavage, and concentration of library members. Last, the issue of quality control will be addressed and the applications of this library to chemical biology will be discussed.

Results and Discussion

Library Design and Development. The selection of reaction pathways and building blocks was governed by the goal of ultimately obtaining a library of compounds that emulated the structural types observed in benzopyran-containing natural products. Prominent structural features of this class include the following: both rigid and flexible skeletons, multiple aromatic rings (including heterocycles) with various substituents (hydroxyl, alkoxy, keto, and ester), saturated and unsaturated alkyl side chains, carbohydrates, molecular weights between 200 and 600, and an average of 3 to 6 heteroatoms per structure.[1] In choosing reaction sequences to generate structures with these characteristics, the following five key criteria were considered: (a) structural resemblance of library members to natural products as described above; (b) overall diversity and complexity of the library;[4] (c) reliability of reaction sequences (i.e. yield and purity of final cleavage products); (d) availability of building blocks; and (e) potential to simplify the synthesis through split-and-pool techniques.

* Address correspondence to this author at The Scripps Research Institute.
† IRORI, a Discovery Partners International Company.

(1) Nicolaou, K. C.; Pfefferkorn, J. A.; Roecker, A. J.; Cao, G.-Q.; Barluenga, S.; Mitchell, H. J. *J. Am. Chem. Soc.* **2000**, *122*, 9939−9953.

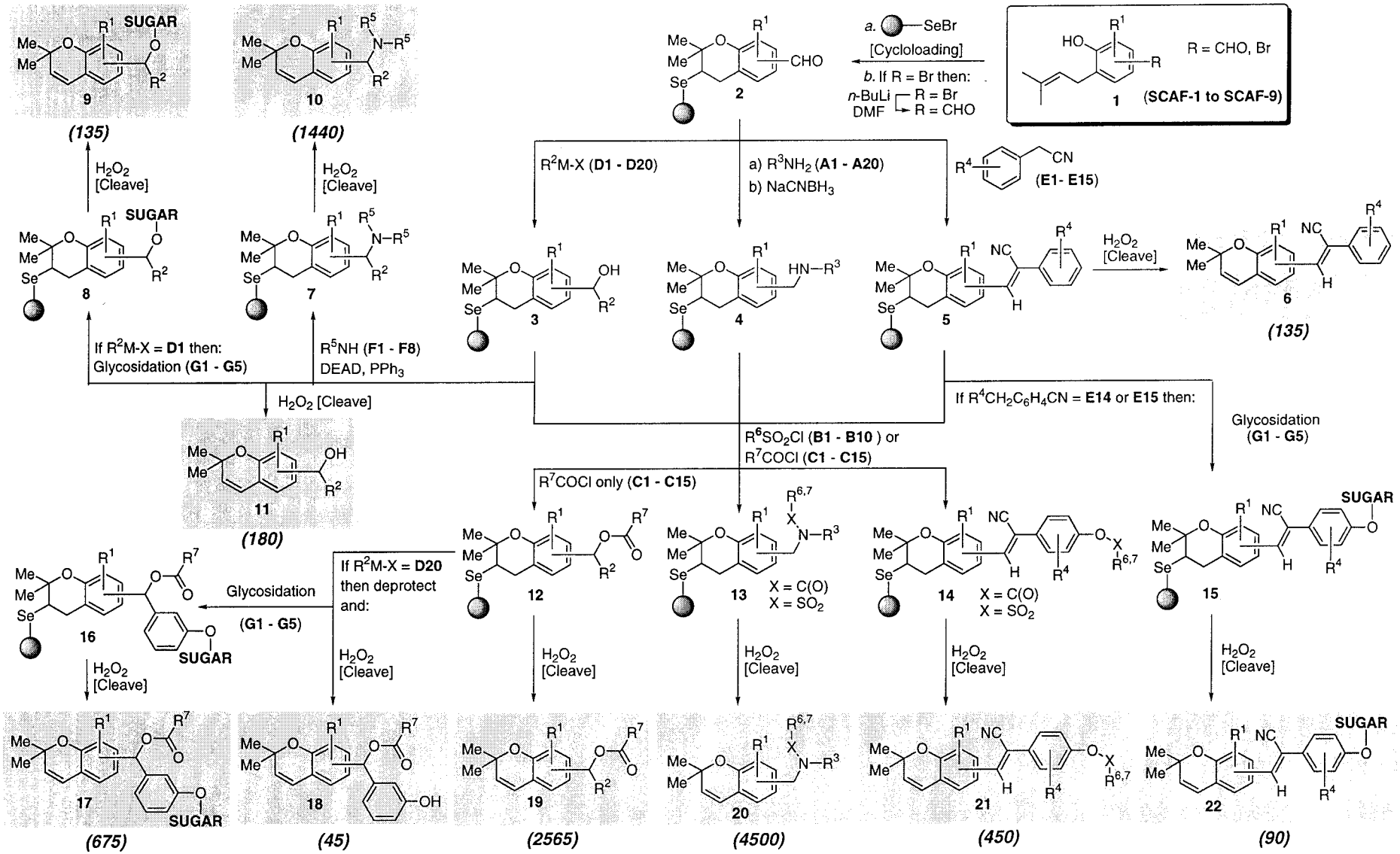
(2) Nicolaou, K. C.; Pfefferkorn, J. A.; Barluenga, S.; Mitchell, H. J.; Roecker, A. J.; Cao, G.-Q. *J. Am. Chem. Soc.* **2000**, *122*, 9968−9976.

(3) A preliminary communication of this optical encoding concept has been disclosed (see: Xiao, X.-Y.; Zhoa, C.; Potash, H.; Nova, M. P. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 780−782) wherein laser etched (6 × 6) ceramic encoding grids were embedded in polypropylene blocks. On the surface of these blocks was grafted aminomethylated polystyrene, and the blocks were employed in the synthesis of a small oligonucleotide library. The present method differs from this early system in that the NanoKan microreactors now encapsulate the polymer (as opposed to having it grafted on the surface) making the current method compatible with a wider range of commercial and specialized resins.

(4) For a review of combinatorial diversity, see: Kauvar, L. M.; Laborde, E. *Curr. Opin. Drug Discovery Dev.* **1998**, *1*, 66−70.

10.1021/ja002034c CCC: $19.00
Published on Web 09/30/2000

Scheme 1. Overall Synthetic Sequence for the Construction of a 2,2-Dimethylbenzopyran-Based Heterogeneous Natural Product-like Library[a,b]



[a] See Figure 1 for structures of building blocks. [b] Italicized numbers in parentheses below structures represent the number of compounds of that type found in the final library.

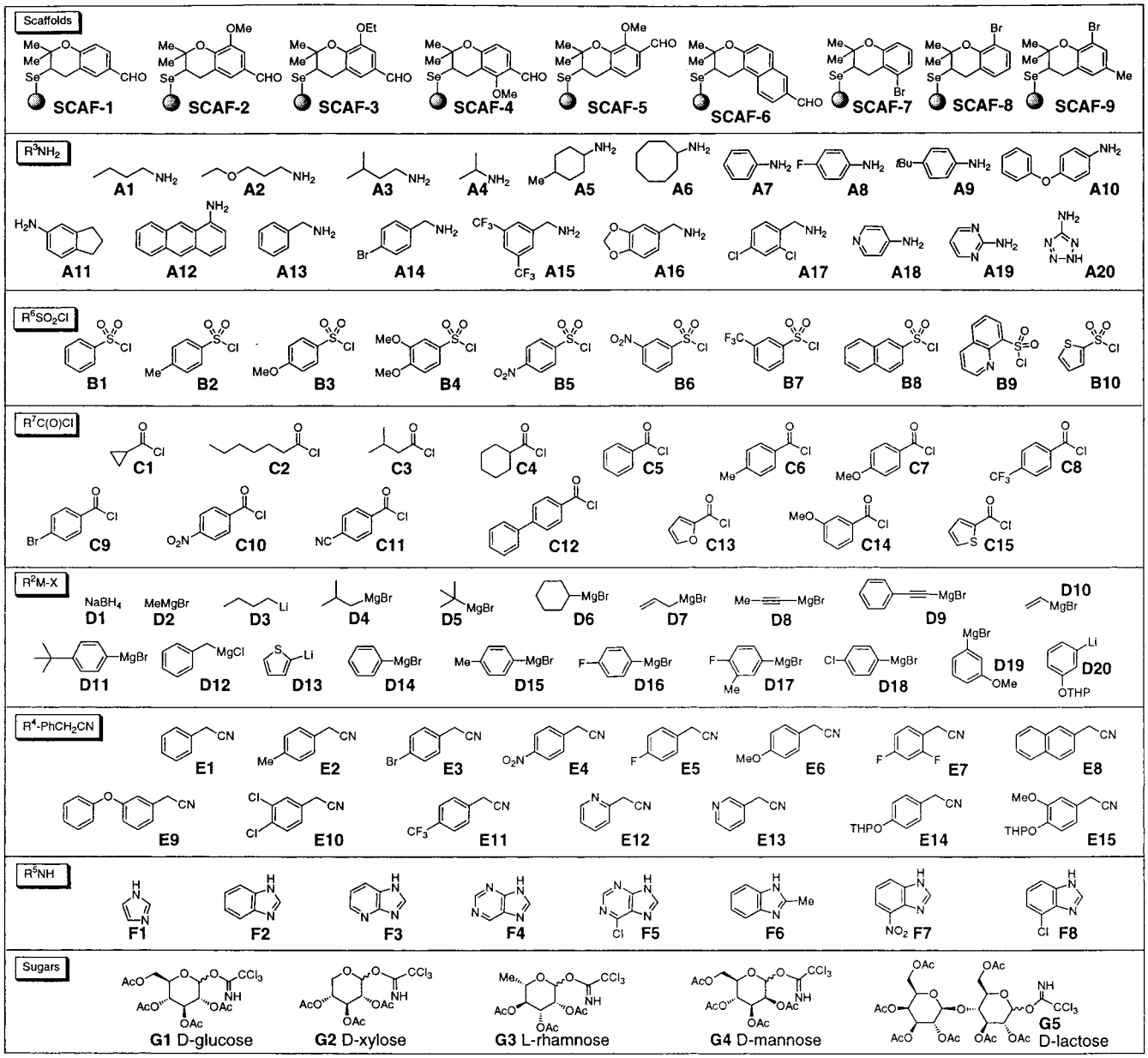


Figure 1. Building blocks employed for construction of natural product-like library based on the benzopyran template.

The complete reaction manifold employed for constructing this library is outlined in Scheme 1 with the corresponding building blocks shown in Figure 1. As illustrated, a heterogeneous synthetic strategy was selected commencing from nine aldehyde-containing benzopyran scaffolds (**2**, Scheme 1) from which all library members were derived. These scaffolds were either loaded directly from the corresponding *o*-prenylated benzaldehydes (**SCAF-1** through **SCAF-6**, Figure 1) or constructed from the corresponding aryl bromide (**SCAF-7** through **SCAF-9**, Figure 1).[1] The decision to use a common functionality (i.e. an aldehyde) for the starting point of all reactions reduced the number of overall processes required as well as the number of *o*-prenylated phenols (**1**, Scheme 1) that had to be synthesized. The aldehyde functional group was specifically selected owing to its ability to be engaged in a variety of diversity building reactions many of which introduce a second functional group which can be, in turn, further derivatized. Diversity was introduced into scaffolds themselves (**SCAF-1** through **SCAF-9**, Figure 1) by varying the position of the aldehyde on the aromatic ring as well as by adding various alkyl and alkoxy substituents. The effect that varying the aldehyde position can have on molecular architecture is illustrated in Figure 2, wherein four scaffolds (**SCAF-1**, **SCAF-5**, **SCAF-7**, and **SCAF-8**), each with an aldehyde group in a different position, were subjected to the same reaction sequence (described in Scheme 4) to provide compounds **23**−**26** (Figure 2) with quite different shapes.

The nine aldehydes represented the starting point of all synthetic operations and, as shown in Scheme 1, proceeded directly to a synthetic branch point where three reaction pathways (organometallic addition, reductive amination, and Knoevenagel condensation) were utilized to provide structures of types **3**, **4**, and **5**, respectively. The functional groups generated by these three operations were subsequently enlisted for installation of secondary elements of diversity. Hence, the alcohols (**3**), amines (**4**), and phenols (**5**, R^4 = OH) were all acylated (with acid chlorides **C1** − **C15**, Figure 1) to give esters (**12**, Scheme 1), amides [**13**, R = C(O)], and phenolic esters [**14**, X = C(O)], respectively. In addition, the secondary amines (**4**, Scheme 1) and phenols (**5**) were reacted with various sulfuryl

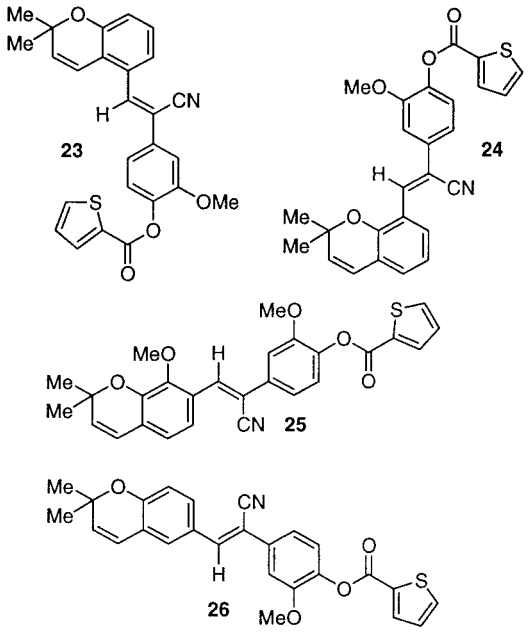


Figure 2. Representative examples of the same reaction sequence generating structurally distinct products based upon the use of four different aldehyde scaffolds (**SCAF-7** → **23**, **SCAF-8** → **24**, **SCAF-5** → **25**, **SCAF-1** → **26**). See Scheme 4 for reaction sequence.

chlorides (**B1**−**B10**, Figure 1) to provide sulfonamides [**13**, X = SO_2] and sulfonates [**14**, X = SO_2], respectively. Secondary alcohols (**3**) were also reacted with various imidazole derivatives (**F1**−**F8**, Figure 1) through a Mitsunobu reaction to produce heterocycles (**7**, Scheme 1).[5] Additionally, several structural types bearing either an alcohol or phenolic functionality were glycosidated (with trichloroacetimidates **G1**−**G5**, Figure 1)[6] to provide carbohydrates of general types **8**, **15**, and **16** (Scheme 1). Once complete, all compounds were released from the resin into separate wells of microtiter plates using an automated version of the oxidation−elimination protocol.

The evolution of this synthetic strategy was heavily influenced by our earlier experiences during the syntheses of several focused libraries as described in the previous paper. In fact, the adoption of a number of the current reaction pathways (specifically: organometallic addition, Knoevenagel condensation, Mitsunobu reaction, and glycosidation) was a direct result of the earlier investigations. Other pathways from the previous library studies were excluded because of either lack of available building blocks (i.e. pyranocoumarin library) or poor reaction reliability (i.e. chalcone library). It is noteworthy that the reductive amination sequence employed herein was not used in the construction of earlier libraries. Its inclusion in the current effort reflects the many diverse building blocks that it can accommodate as well as the very high fidelity with which the reaction proceeds. The combination of the selected pathways resulted in a collection of compounds with similar structural characteristics to the parental natural products in terms of number and types of rings, lipophilicity, types of substituents, and average molecular weight. In addition, sugars were attached to a small portion (ca. 8%) of library members in such a way as to mimic the types of benzopyran glycoside conjugates observed in nature.[1]

Specific Reaction Pathways. In light of the complexity of the overall synthetic strategy, it was deemed important to trace the fate of individual compounds through each of the pathways to better illustrate reaction conditions, yields, and estimated purities of the final products. Later in this paper, we detail how the NanoKan Optical Encoding platform, recently developed by IRORI, a Discovery Partners International Company, was used for automated tagging, sorting, cleaving, and concentrating of the library.[3] A key element of this automation platform is the microreactors (termed NanoKans) which encapsulate aliquots of resin for individual library members. Thus, to accurately simulate the actual synthetic conditions expected to be encountered during the library construction (particularly reaction kinetics), all test reactions described below were conducted in these (or related) microreactors.

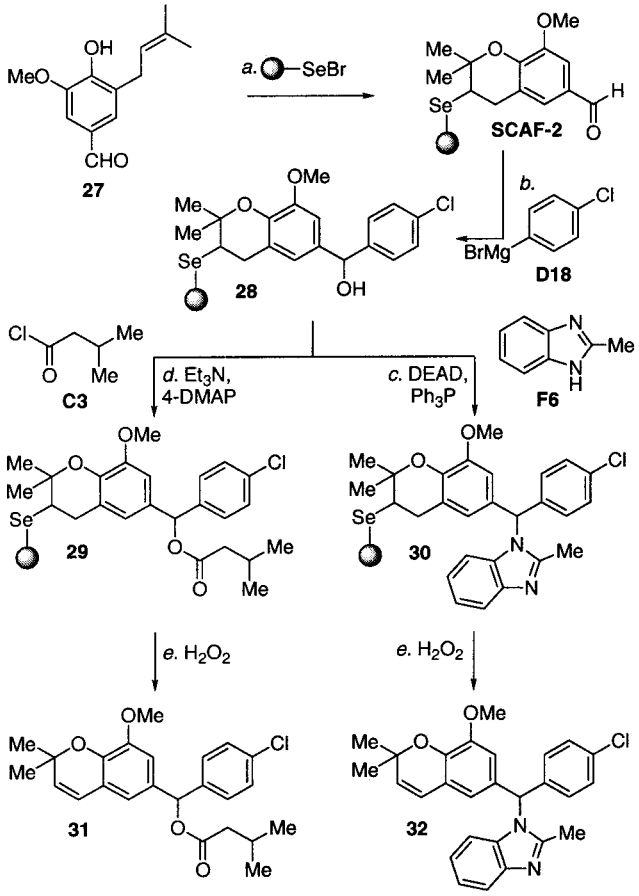
The first reaction of the branch point involved addition of alkyl and aryl organometallics to the starting aldehyde scaffolds to form secondary benzyl alcohols or benzhydrols. In the previous paper, we reported the construction of similar benzyl alcohols through a halogen−metal exchange of a resin-bound bromide and subsequent quenching with various aldehydes.[1] On the basis of comparative experimentation, however, we found that the addition of the organometallic species to a resin-bound aldehyde proceeded with higher purity and tolerated a greater variety of functionality than the reverse approach previously used. As an example, aldehyde conjugate **SCAF-2** (Scheme 2) was constructed by treatment of benzaldehyde **27** with selenenyl bromide resin. The resulting scaffold was then reacted with 4-chlorophenylmagnesium bromide (**D18**) in THF at 0 °C to afford benzhydrol **28** (98% conversion based on ^{1}H NMR analysis of cleavage product). This alcohol was then used in two separate reaction paths, the first of which required treatment of **28** with isovaleryl chloride (**C3**) in the presence of Et_3N and 4-DMAP to afford ester **29** which was cleaved giving benzopyran **31** in greater than 90% purity as suggested by ^{1}H NMR spectroscopy (illustrated in Scheme 2). In a second reaction, alcohol **28** was treated with Ph_3P, 2-methylbenzimidazole (**F6**), and DEAD to generate structure **30** which was cleaved (aqueous workup) to provide amine **32** in an estimated 80% purity (^{1}H NMR spectroscopy).[5] It is noteworthy that the ^{1}H NMR spectra shown were obtained from the cleavage of single NanoKans each encapsulating 8.0 mg of resin and, therefore, represent the actual amount of compound expected for all members of the final library. The ability to obtain such spectra attested to the fidelity of the reaction sequences and confirmed that ^{1}H NMR spectroscopy could indeed be employed to assist in the characterization of the final 10 000-membered library.

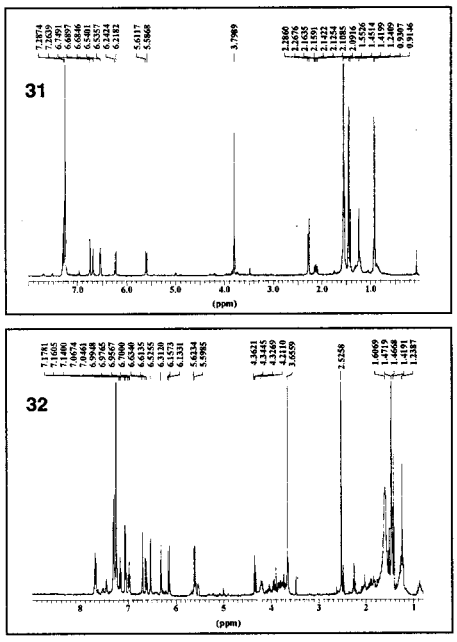
The second reaction of this branch point involved conversion of the aldehyde scaffolds to various amide and sulfonamide derivatives through a sequence involving reductive amination as exemplified in Scheme 3. Thus, resin-bound aldehyde **SCAF-5** (Scheme 3) was prepared by loading of benzaldehyde **33** onto the selenium resin and then condensing the resulting product with 4-fluoroaniline (THF, 65 °C, 5 h) to generate the corresponding imine which was reduced in situ by addition of sodium cyanoborohydride to form secondary amine **34** in 98% conversion (^{1}H NMR). This amine (**34**) was treated separately with 4-nitrobenzenesulfonyl chloride (**B5**) and cyclopropanecarbonyl chloride (**C1**) in the presence of both Et_3N and 4-DMAP to provide sulfonamide **35** and amide **36**, respectively. Cleavage of these products afforded sulfonamide **37** and amide

(5) Soll, R. M.; Dollings, P. J.; Mitchell, R. D.; Hafner, D. A. *Eur. J. Med. Chem.* **1994**, *29*, 223−232.

(6) For preparation of trichloroacetimidates, see: (a) D-Glucose: Schmidt, R. R.; Michel, J. *Angew. Chem., Int. Ed. Engl.* **1980**, *9*, 731−732. (b) D-Xylose: Mori, M.; Ito, Y.; Ogawa, T. *Carbohydr. Res.* **1990**, *195*, 199−224. (c) L-Rhamnose: Kerékgártó, J.; Agoston, K.; Batta, G.; Szurmai, Z. *Carbohydr. Res.* **1997**, *297*, 153−161. (d) D-Mannose: Toepfer, A.; Kretzschmar, K. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1311−1316. (e) D-Lactose: Hasegawa, A.; Morita, M.; Kojima, Y.; Ishida, H.; Kiso, M. *Carbohydr. Res.* **1991**, *214*, 43−53.

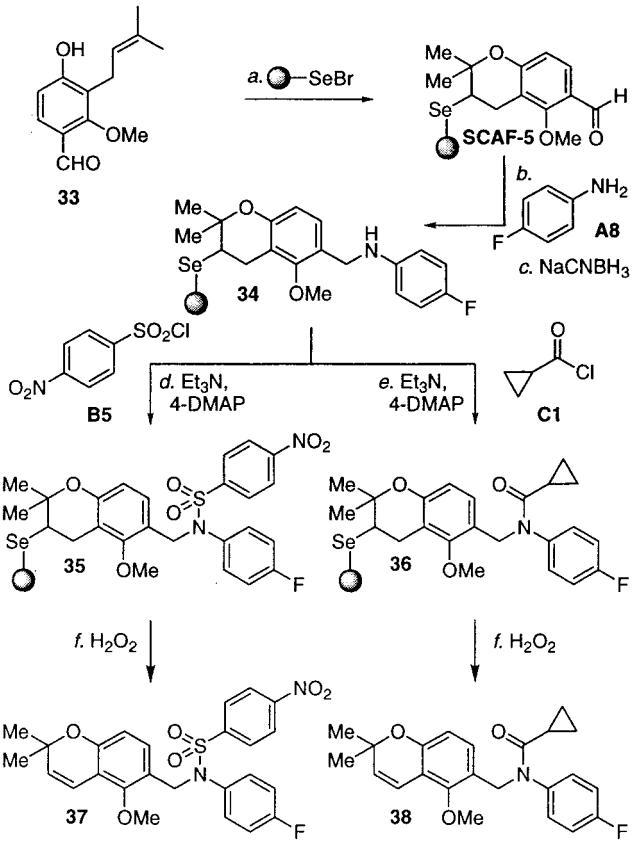
Scheme 2. Reaction Pathway Leading to Esters **31** and **32**[a]

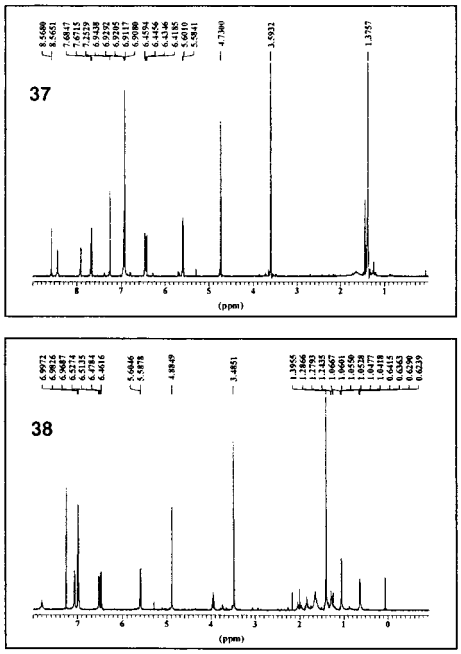




[a](a) 2.0 equiv of **27**, CH_2Cl_2, 25 °C, 1 h, 100%; (b) 15.0 equiv of 4-chlorophenylmagnesium bromide (**D18**), THF, 0 → 25 °C, 2.5 h, 98%; (c) 15.0 equiv of 2-methylbenzimidazole (**F6**), 10.0 equiv of Ph_3P, 15.0 equiv of DEAD, THF, 25 °C, 24 h, 80 %; (d) 20.0 equiv of isovaleryl chloride (**C3**), 20.0 equiv of Et_3N, 1.0 equiv of 4-DMAP, CH_2Cl_2, 25 °C, 12 h, 90%; (e) 6.0 equiv of H_2O_2, THF, 25 °C, 20 min. 4-DMAP = 4-dimethylaminopyridine, DEAD = diethyl azodicarboxylate. ^{1}H-NMR (400 MHz) were recorded in $CDCl_3$ (see Supporting Information for expansions).

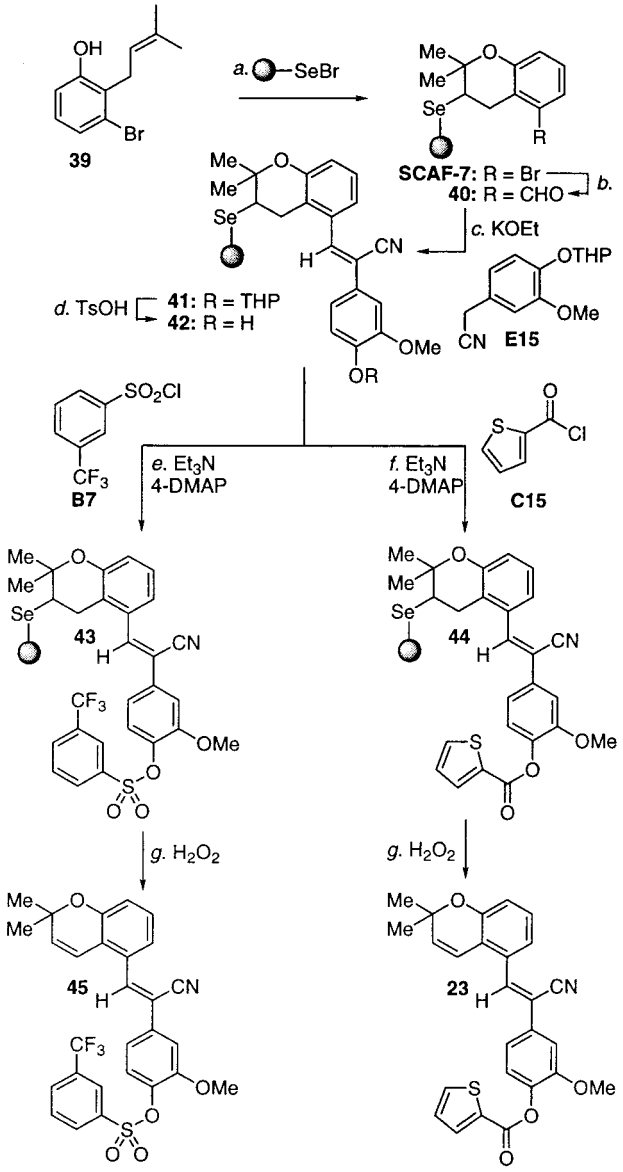
Scheme 3. Reductive Amination Pathway Leading to **37** and **38**[a]





[a](a) 2.0 equiv of **33**, CH_2Cl_2, 25 °C, 1 h, 100%; (b) 20.0 equiv of 4-fluoroaniline (**A8**), THF, 65 °C, 4 h; (c) 25.0 equiv of $NaCNBH_3$, THF:MeOH (10:1), 65 °C, 6 h, 98%; (d) 20.0 equiv of 4-nitrosulfonyl chloride (**B5**), 20.0 equiv of Et_3N, 1.0 equiv of 4-DMAP, CH_2Cl_2, 25 °C, 12 h, 90%; (e) 20.0 equiv of cyclopropanecarbonyl chloride (**C1**), 20.0 equiv of Et_3N, 1.0 equiv of 4-DMAP, CH_2Cl_2, 25 °C, 12 h, 90%; (f) 6.0 equiv of H_2O_2, THF, 25 °C, 30 min. 4-DMAP = 4-dimethylaminopyridine, ^{1}H-NMR (600 MHz) were recorded in $CDCl_3$ (see Supporting Information for expansions).

Scheme 4. Reaction Pathway Leading to Stilbenoids **23** and **45**[a]



[a](a) 2.0 equiv of **39**, CH_2Cl_2, 25 °C, 1 h, 100%; (b) 10.0 equiv of *n*-BuLi (1.6 M in hexanes), THF, -78 → 0 °C, 2 h and 20 equiv of DMF, -78 → 0 °C, 2 h then 1 M HCl (aq), 0 °C, 10 min, 99%; (c) 10.0 equiv of 3-OMe-4-OTHP-phenylacetonitrile (**E15**), 20.0 equiv of KOEt (1.0 M in EtOH), THF, 25 °C, 8 h, 92%; (d) 2.0 equiv of TsOH, THF:MeOH (9:1), 25 °C, 1 h, 100%; (e) 10.0 equiv of 3-(trifluoromethyl) benzenesulfonyl chloride (**B7**), 20.0 equiv of Et_3N, 1.0 equiv of 4-DMAP, CH_2Cl_2, 25 °C, 12 h, 75%; (f) 10.0 equiv of 2-thiophenecarbonyl chloride (**C15**), 20.0 equiv of Et_3N, 1.0 equiv of 4-DMAP, CH_2Cl_2, 25 °C, 12 h, 80%; (g) 6.0 equiv of H_2O_2, THF, 25 °C, 20 min. TsOH = *p*-toluenesulfonic acid.

38 in greater than 90% purities as suggested by ^{1}H NMR spectroscopy.

The third reaction of this branch utilized a Knoevenagel condensation of the aldehyde scaffolds to produce stilbene-type structures.[7] As illustrated in Scheme 4, scaffold **40** was prepared by loading of bromophenol **39** to give **SCAF-7** which was

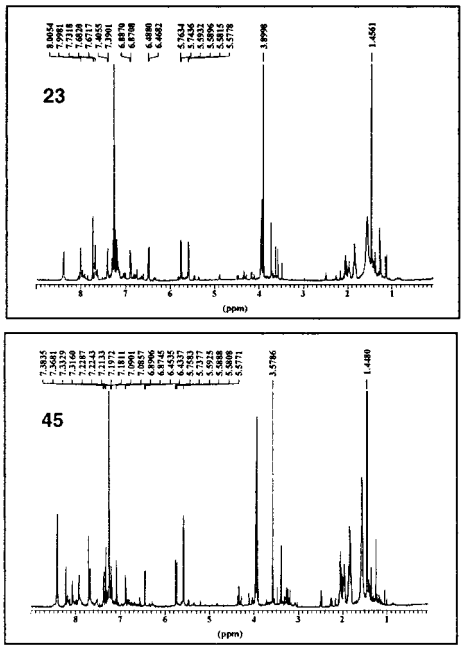


Figure 3. ^{1}H NMR spectra of stilbenoids **23** and **45** constructed as illustrated in Scheme 4. ^{1}H NMR (500 MHz) were recorded in $CDCl_3$ (see Supporting Information for expansions).

treated with *n*-BuLi (−78 → 0 °C, 2 h) and then quenched with DMF (−78 → 0 °C, 2 h).[8] Cleavage of a portion of **40** revealed that this transformation had occurred with 99% conversion. The resulting resin-bound aldehyde **40** was then condensed with 3-OMe-4-(OTHP)-phenylacetonitrile (**E15**) in the presence of KOEt to afford *E*-stilbene **41** in 92% conversion (^{1}H NMR spectroscopy). The acid labile THP protecting group was then removed by treatment with TsOH (100% conversion), and the resulting phenol was reacted with 3-(trifluromethyl)benzenesulfonyl chloride (**B7**) and 2-thiophenecarbonyl chloride (**C15**) under identical conditions used for the secondary alcohols. Cleavage of the resulting compounds afforded stilbenes **45** and **23** in 75 and 80% estimated purities, respectively, over six steps as estimated by ^{1}H NMR spectroscopy (see Figure 3 for ^{1}H NMR spectra of compounds **23** and **45**).

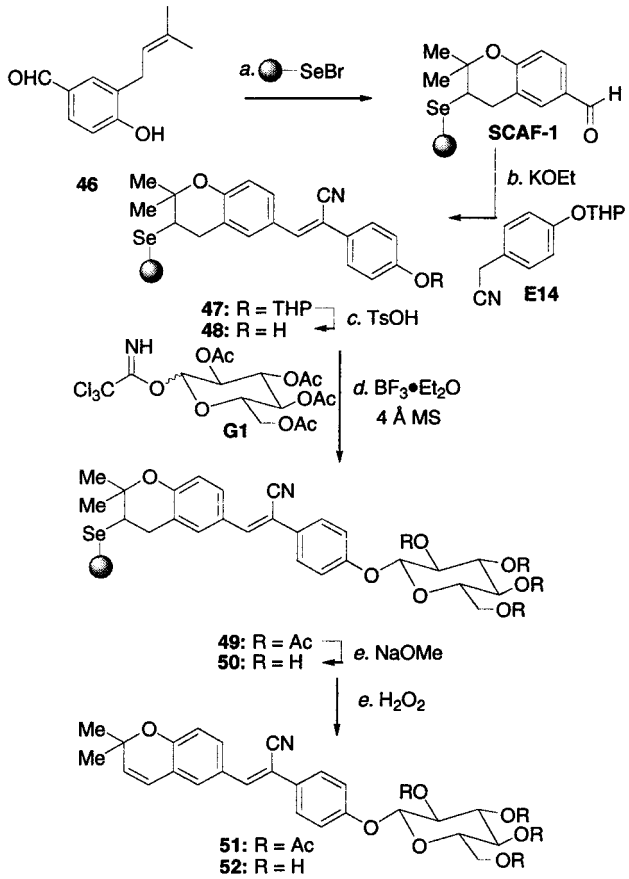
The construction of all library members proceeded through one of the above three reaction pathways. However, certain building blocks employed in these pathways contained other functionality that allowed for the installation of additional diversity elements. Of particular interest was the coupling of several structure types with assorted mono- and disaccharide units (**G1**−**G5**, Figure 1**)** as illustrated in Scheme 1. As mentioned in the preceding paper, carbohydrates are known to improve the solubility and cellular targeting of small molecules;[9] consequently, we sought to incorporate these moieties into a small portion of the library to enhance both the structural diversity and pharmacological potential of library members. An example of this strategy is outlined in Scheme 5 for the synthesis of stilbene glycosides **51** and **52**. As shown, scaffold **SCAF-5** was constructed from phenol **46** and condensed with 4-OTHP phenyl acetonitrile **E14** in the presence of KOEt to provide stilbene **47**. The THP-protected phenol, introduced during the condensation, was then liberated by treatment with TsOH, and the resulting phenol was glycosidated using conditions developed previously for the synthesis of a natural chromene glucoside isolated from *Ageratum conyzoides*. In the event,

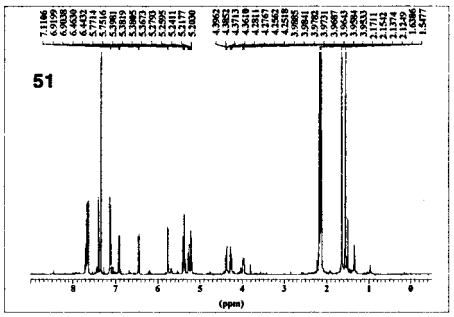
(7) Ohsumi, K.; Nakagawa, R.; Fukuda, Y.; Hatanaka, T.; Morinaga, Y.; Nihei, Y.; Ohishi, K.; Suga, Y.; Akiyama, Y.; Tsuji, T. *J. Med. Chem.* **1998**, *41*, 3022−3032.

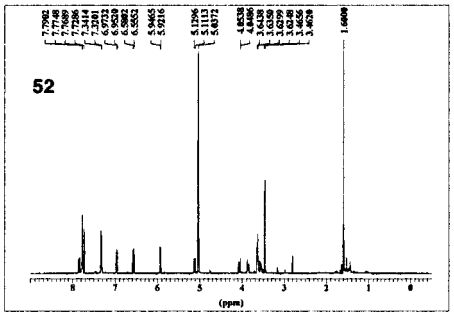
(8) Ding, C. Z. *Synth. Commun.* **1996**, *26*, 4267−4273.

(9) For discussion, see: Hecht, S. M. *Bioorganic Chemistry: Carbohydrates*; Oxford Press: New York, 1999.

Scheme 5. Reaction Pathway Leading to Glycosides **51** and **52**[a]

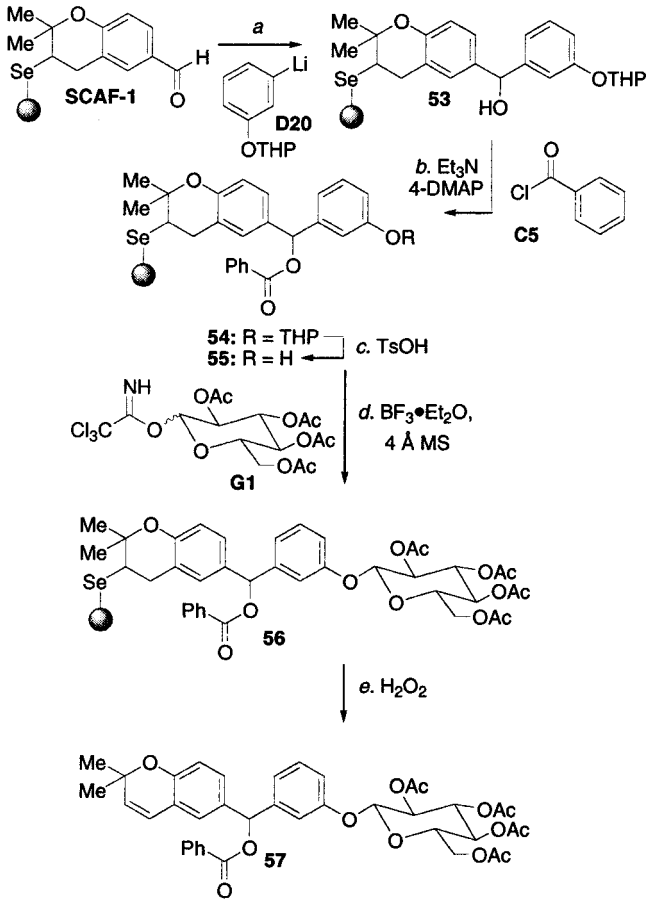


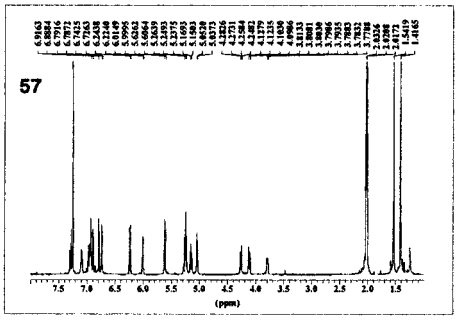




[a](a) 2.0 equiv of **46**, CH_2Cl_2, 25 °C, 1 h, 100%; (b) 10.0 equiv of 4-OTHP-phenylacetonitrile (**E14**), 20.0 equiv of KOEt (1.0 M in EtOH), THF, 25 °C, 8 h, 90%; (c) 2.0 equiv of TsOH, THF:MeOH (9:1), 25 °C, 1 h, 100%; (d) 5.0 equiv of trichloroacetimidate **G1**, 4.0 equiv of BF_3•Et_2O, CH_2Cl_2, -40 → 0 °C, 12 h, 85%; (e) 10.0 equiv of NaOMe, THF:MeOH (2:1), 25 °C,12 h, 100%, f) 6.0 equiv of H_2O_2, THF, 25 °C, 20 min. ^{1}H-NMR (400 MHz) were recorded in $CDCl_3$ (**51**) and CD_3OD (**52**) (see Supporting Information for expansions).

Scheme 6. Reaction Pathway Leading to Glycoside **57**[a]





[a](a) 10.0 equiv of 3-THPO-phenyllithium (**D20**), THF, -78 → 0 °C, 1 h, 96%; (b)10.0 equiv of benzoyl chloride (**C5**), 20.0 equiv of Et_3N, 1.0 equiv of 4-DMAP, CH_2Cl_2, 25 °C, 12 h, 100%; (c) 2.0 equiv of TsOH, THF:MeOH (9:1), 25 °C, 30 min, 100%; (d) 5.0 equiv of trichloroacetimidate **G1**, 4.0 equiv of BF_3•Et_2O, CH_2Cl_2, -40 → 0 °C, 12 h; (e) 6.0 equiv of H_2O_2, THF, 25 °C, 20 min, 90%.^{1}H-NMR (500 MHz) were recorded in $CDCl_3$ (see Supporting Information for expansions).

phenol **48** was coupled with trichloroacetimidate **G1**[6] in the presence of BF_3·Et_2O and 4 Å molecular sieves while warming from −40 to 0 °C over 12 h to afford peracetylated glycoside **49**. Oxidative cleavage of the resin in one microreactor containing **49** provided peracetylated glycoside **51** in 90% estimated purity (see ^{1}H NMR spectrum) indicating that the glycosidic coupling had occurred with high selectivity for the β-glycoside. Additional microreactors containing **49** were then treated with NaOMe in THF for 12 h at 25 °C to affect deprotection of **49** to **50** which was then released by standard oxidative cleavage affording glycoside **52** in an estimated 85% purity. A second glycoside example is illustrated in Scheme 6

wherein scaffold **SCAF-1** (from Scheme 5) was treated with 3-OTHP aryllithium **D20** to generate benzhydrol **53** in 95% yield. Esterification of the secondary alcohol of **53** was accomplished by treatment with benzoyl chloride (**C5**) to provide structure **54**, the phenol of which was then liberated by treatment with TsOH to afford **55**. Using the same conditions as above, phenol **55** was coupled with trichloroacetimidate **G1**[6] in the presence of $BF_3 \cdot Et_2O$ and 4 Å molecular sieves while warming from −40 to 0 °C over 12 h to afford peracetylated β-glycoside **56** which upon cleavage afforded glycoside **57** in 90% estimated purity (see ^{1}H NMR spectrum of **57**). Due to complications from functional group overlap, glycosidated structures such as **57** which contained ester moieties were not deprotected; instead, these compounds were cleaved directly and incorporated into the library as the peracetylated glycosides.

Selection and Validation of Building Blocks. With the various reaction pathways established, the next step was to select suitable building blocks for use in each sequence. In evaluating potential components, both reaction reliability and product diversity were considered. In most cases, a greater number of building blocks were tested than were actually used such that low yielding candidates could be excluded. In choosing components for preliminary testing, we deliberately included both alkyl and aryl moieties encompassing different electronic substituents and varying degrees of lipophilicity. Building block pools were evaluated in two types of tests to be described consecutively using the organometallic addition reaction (**2** → **3**, Scheme 1) as a representative example. Hence, a collection of 25 organometallic reagents including hydrides, alkyl- and aryllithiums, and alkyl, alkenyl, vinyl, and aryl Grignards were initially screened. These reagents were evaluated by reaction with a representative aldehyde scaffold. In the event, resin-bound scaffold **SCAF-1** (Figure 1) was loaded on resin in MicroKans and subsequently reacted individually with each organometallic candidate. The resulting benzyl alcohol products were then pooled and acetylated (Ac_2O, 4-DMAP) before being sorted and cleaved into individual vials. The cleavage products were then examined by TLC (see Figure 4), ES-MS, and ^{1}H NMR spectroscopy. In general, any building block exhibiting less than 90% conversion to a single product was excluded at this point. Subsequently, we sought to ensure that each reaction type (i.e. organometallic addition, reductive amination, etc.) would work reliably on each of the nine scaffolds. We were particularly interested here in preemptively uncovering any steric or regioselectivity problems before they could compromise the actual library synthesis. As an example, a representative organometallic reagent [i.e. phenylmagnesium bromide (**D14**)] was individually reacted with all nine aldehyde scaffolds (**SCAF-1** through **SCAF-9**), each loaded onto resin held in separate microreators. The resulting benzhydrols were then acetylated as described above and cleaved for analysis. In this case, all nine scaffolds exhibited excellent conversion (>95%) and no problems were noted. This two-step screening process was then repeated for the reductive amination (45 amines tested, from which 20 were selected) and Knoevenagel condensation (25 nitriles tested, from which 15 were selected) sequences as shown in Figure 4. In addition, the loading/cleavage of all scaffolds was evaluated and all sugars were tested in a representative glycosidation reaction. All test reactions and the resulting TLC plates for the chosen building blocks are illustrated in Figure 4. In general, most reactions and their requisite building blocks exhibited better than 90% conversion in each test, suggesting that the final library compounds would most likely be obtained in acceptable purities (i.e. >80%) for biological studies.[10]

Several interesting reactivity trends were noted during these screening studies. First, during the organometallic testing, both alkyl and aryl Grignards exhibited higher conversion and purities than their alkyl- or aryllithium counterparts. Second, testing of the reductive amination sequence revealed that highest conversions were observed for the anilines and benzylamines independent of the nature of the aryl substituents. The worst performing amines tended to be the heteroaromatics which displayed low conversion and/or low purities. This is likely attributed, in part, to the fact that most of the heteroaromatic amines tested were available only as salts which had to be neutralized during the reaction. Despite this neutralization, solubility remained a consistent problem which was exacerbated by the diffusion requirements of the microreactors thereby contributing to the low conversions observed for those reactions. Of these heteraromatic amines, only **A18**−**A20** performed sufficiently well to warrant inclusion in the library. Third, testing of the acid chlorides and sulfuryl chlorides with either amines, alcohols, or phenols revealed no obvious reactivity trends as all reagents displayed high conversion and purities. Fourth, during testing of the substituted phenyl acetonitriles in the condensation reaction it was found that the electronically neutral or electron-deficient substrates reacted faster than their electron-rich counterparts, but given sufficient reaction time all proceeded to nearly complete conversion. Interestingly, the desired trans stereochemistry was predominantly observed (>90%) in almost all of the cases tested. Finally, the imidazole derivatives and the trichloroacetimidates generally performed equally well in their respective test reactions regardless of structural properties.

Library Encoding, Synthesis, and Cleaving Strategy. A critical element of any split-and-pool library synthesis strategy is the manner in which the individual compounds are encoded such that one can rapidly identify the structures of individual library members and avoid the time-consuming task of mixture deconvolution.[11] Two general techniques have traditionally been used for this task: chemical encoding and nonchemical encoding. Chemical encoding relies on the iterative coupling of chemical tags (typically oligonucleotides, peptides, or binary small molecule systems) to orthogonally functionalized beads during the library synthesis.[12] These chemical tags can be analyzed (either on or off the bead) by appropriate analytical techniques with their constitution being a consequence of the synthetic path that the compound from that bead followed.[12] The advantages of chemical tagging include the relatively low initial and per compound costs as well as the large library sizes that can be accommodated. On the downside, this technique requires additional synthetic steps to attach the requisite tags; moreover, the sensitivity of the tags may limit the range of chemical transformations which can be used in the library construction. In some cases, the decoding of these chemical tags can also be a tedious and time-consuming process.

As an alternative to the chemical strategy, various nonchemical encoding techniques have been developed which attempt to record the synthetic history of compounds by physical means.[13]

(10) For discussion of requirements of library purity, see: *J. Comb. Chem.* **1999**, *1*, 11A−12A.

(11) For a general review, see: Czarnik, A. W. *Curr. Opin. Chem. Biol.* **1997**, *1*, 60−66.

(12) For a review of chemical encoding, see: Barnes, C.; Scott, R. H.; Balasubramanian, S. *Rec. Res. Dev. Org. Chem.* **1998**, *2*, 367−379.

(13) Xiao, X.-Y.; Nova, M. P. *Comb. Chem.* **1997**, 135−152.

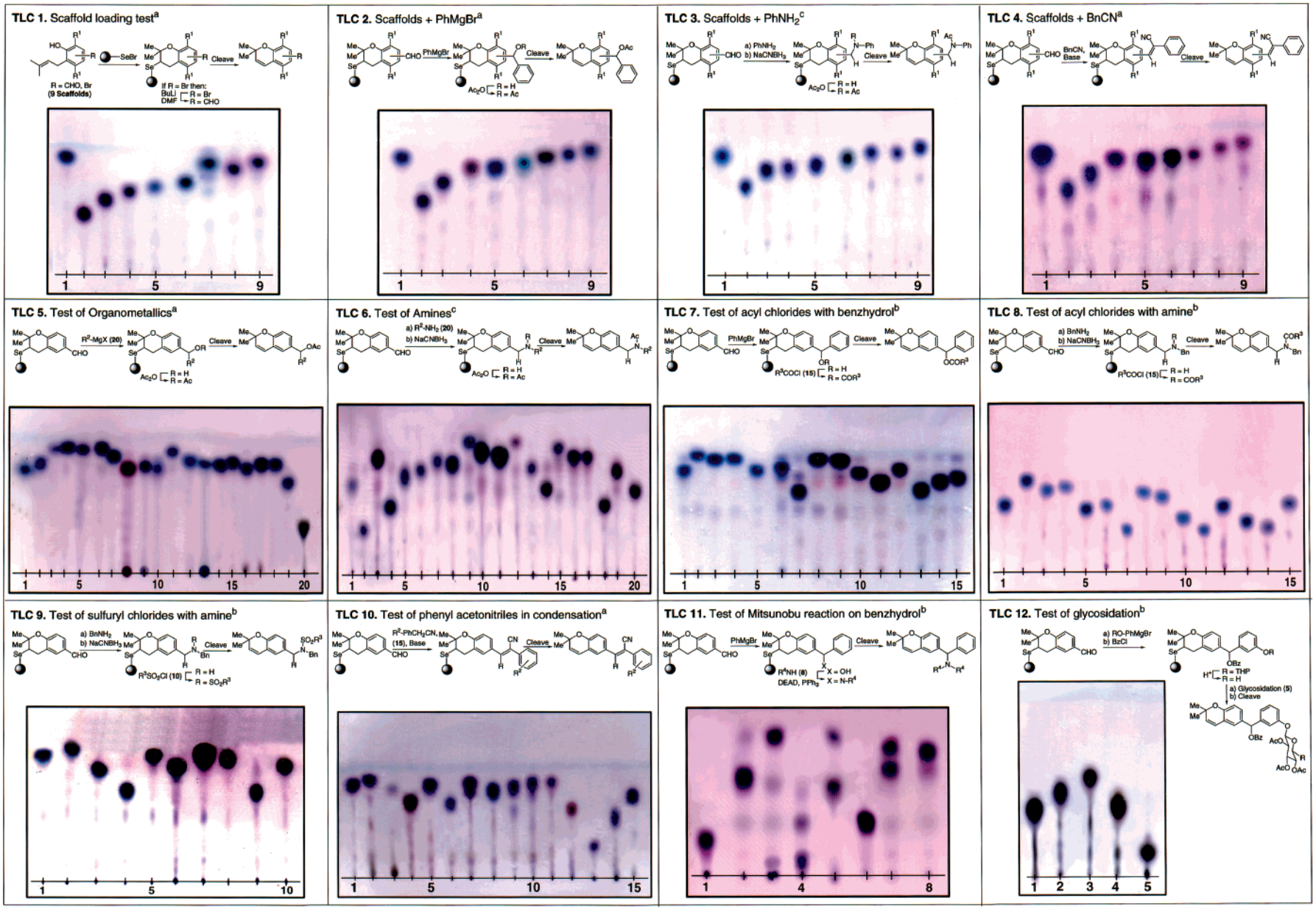


Figure 4. Thin-layer chromatographic plates of test reactions for scaffolds and building blocks. TLC solvent systems: *a* = 20% EtOAc in hexanes; *b* = 40% EtOAc in hexanes; *c* = 60% EtOAc in hexanes.

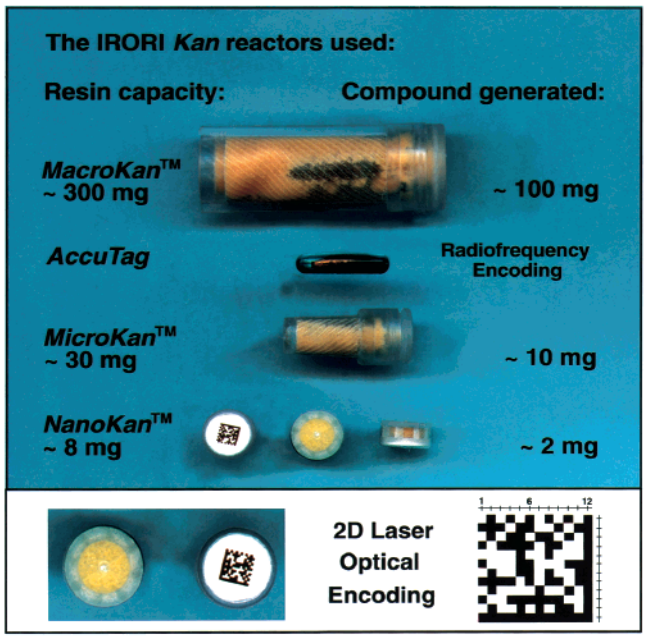


Figure 5. The IRORI family of microreactors.

Particularly useful in these laboratories has been the technology of radiofrequency (RF) encoding[14] coupled with microreactor technology developed at IRORI, now a subsidiary of Discovery Partners International Company, in 1995.[13] In this system, polypropylene microreactors with porous side panels are used to encapsulate a solid-phase resin and a radiofrequency addressable semiconductor tagging device which is capable of receiving, storing, and emitting RF signals to a suitably interfaced computer.[13] In this way, the RF tag can be used to direct its resin aliquot through a predetermined reaction pathway. In our own combinatorial studies, we have employed two types of microreactors (see Figure 5): MacroKans (capable of holding ca. 300 mg of resin) and MicroKans (capable of holding ca. 30 mg of resin). In these systems, multimilligram amounts of cleavage products can be isolated (typically 80−100 mg for Macrokans and 8−10 mg for MicroKans).[15] In practice, this radiofrequency encoding system is well suited for the rapid construction of libraries of up to several thousand members.

To accommodate larger libraries, a high-throughput variant of this technology has recently been developed by IRORI using optical encoding in place of the radiofrequency encoding.[3] This technology involves microreactors (termed NanoKans) containing a laser-etched ceramic grid (12 × 12) which can be read optically for encoding/decoding purposes (see Figure 5). The advantages of this NanoKan platform compared with the radiofrequency system, described above, are the reduced expense of the tags and the complete automation of filling, sorting, washing, and cleaving processes, thereby reducing the labor and cost involved for synthesizing large libraries. During a library synthesis, as shown in Figure 6, the above-described NanoKans are initially filled with ca. 8 mg of resin and then sealed with encoded ceramic caps. The NanoKan microreactors are sorted into vessels where the requisite reactions are performed. Subsequently, the microreactors are pooled, washed, and resorted into vessels for the next reaction. Upon completion of

(14) (a) Nicolaou, K. C.; Xiao, X.-Y.; Parandoosh, Z.; Senyei, A.; Nova, M. P. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2476−2479. (b) Moran, E. J.; Sarshar, S.; Cargill, J. F.; Shahbaz, M. J. M.; Lio, A.; Mjalli, A. M. M.; Armstrong, R. W. *J. Am. Chem. Soc.* **1995**, *117*, 10787−10788.

(15) These ranges are based upon an average resin functionalization of 1.2 mmol/g.

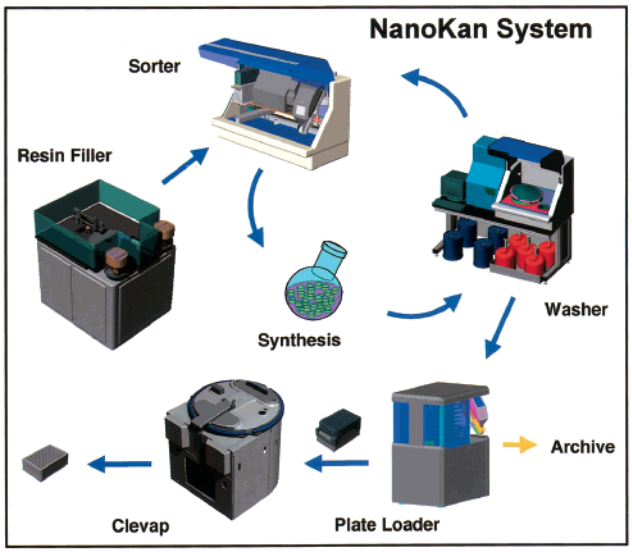


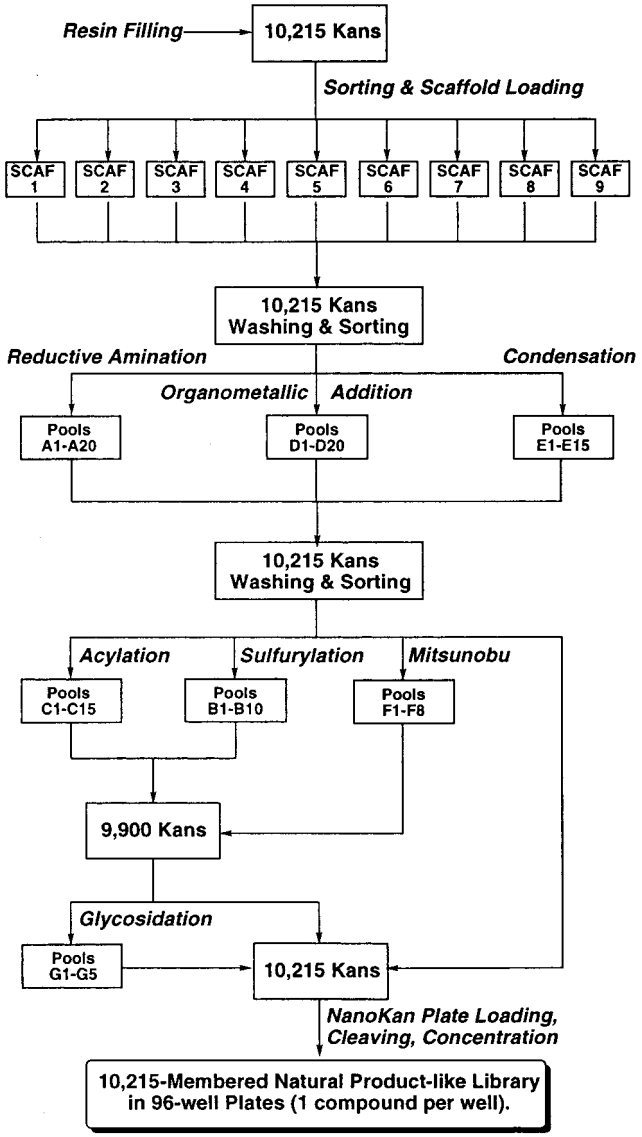
Figure 6. The IRORI NanoKan system.

the reaction sequence, the microreactors are sorted into bar coded cleavage plates (96 NanoKans per plate). Finally, using the automated Clevap system, the products in the NanoKans are released from the resin into standard 96-well microtiter plates where they are automatically concentrated. Since these NanoKans hold ca. 8 mg of resin each, cleavage typically affords ca. 1−2 mg of each library member. This quantity of material allows for analysis of representative members by standard analytical techniques (including NMR spectroscopy) and use of the library in multiple biological screens. Moreover, since the final cleavage products are released into standard 96-well microtiter plates, the entire library is formatted for immediate biological screening with standard liquid handling technologies. It was the use of this highly automated system that allowed for the complete synthesis of the present 10 000-membered natural product-like library in 8 days.

The use of either the IRORI radiofrequency or optical encoding systems offers both advantages and disadvantages as compared to the more traditional chemical tagging encoding strategies. One advantage is that these physical encoding methods require no extraneous chemical manipulations during the library synthesis nor do they limit the types of chemistries which can be performed on library members (as is sometimes the case in chemical encoding due to tag instability). Additionally, all library members are readily identifiable, and cleavage of the individual microreactors affords sufficient quantities of products such that standard analytical techniques can be employed to fully characterize library members. On the downside, this physical method of encoding has a higher initial cost as compared to chemical encoding. In addition, the effects of the microreactor environment on reaction kinetics must be addressed during the reaction validation process. Finally, even with the NanoKan system, the nature of physical encoding imposes a practical limit on library size in the range of 100 000 members as compared to the larger libraries available using chemical encoding.

Library Synthesis. With the encoding strategy selected and all reactions and building blocks validated, synthesis of the actual library commenced. As outlined in Scheme 7, 10 215 NanoKans were loaded with ca. 8 mg of selenenyl bromide resin (1.1 mmol/g functionalization) each and optically encoded using an appropriate algorithm (see Supporting Information for SynthMan algorithm). A preliminary sort of these 10 215 NanoKans provided nine individual batches to be loaded with

Scheme 7. Flowchart for Synthesis, Washing, Sorting, and Cleavage of the Benzopyran-Based Natural Product-like Library Using the IRORI NanoKan System



scaffolds **SCAF-1** through **SCAF-9**. In each event, 1135 microreactors were suspended in CH_2Cl_2 and treated with a CH_2Cl_2 solution of the *o*-prenyl phenol at 25 °C for 1 h, after which time the solvent was removed and the NanoKans thoroughly washed and dried (see Experimental Section for details). Scaffolds containing an aryl bromide (i.e. **SCAF-7** through **SCAF-9**) were then pooled and converted to the corresponding aldehydes by treatment with *n*-BuLi at −78 to 0 °C over 2 h followed by quenching with DMF at −78 to 0 °C over an additional 2 h period. At this stage, all microreactors were pooled and subjected to a global sort resulting in 55 different sets of NanoKans: 20 to be reacted with organometallics (**D1**−**D20**), 20 to be condensed with amines (**A1**−**A20**), and 15 to be condensed with phenylacetonitriles (**E1**−**E15**). As previously described, the 20 organometallic addition reactions were effected by treatment of a suspension of NanoKans in THF at 0 °C with a 1.0 M solution of reagents followed by stirring at 0 °C for 30 min and 25 °C for 4 h. The reductive aminations were effected by heating a suspension of microreactors and amines in THF at 70 °C for 4 h followed by the addition of sodium cyanoborohydride and methanol (10 vol %) and heating for an additional 6 h at 70 °C. Last, the Knoevenagel condensations were achieved by treatment of a suspension of NanoKans and phenylacetonitriles in THF with KOEt at 25 °C over a period of 5 h. Upon completion of these 55 reactions, the NanoKans were combined, washed, and sorted to afford new pools for reactions with 15 acid chlorides, 10 sulfuryl chlorides, and 8 amines (Mitsunobu inversion), along with miscellaneous microreactors destined either for cleavage or glycosidation in a later step. The sulfuryl chlorides (**B1**−**B10**) and acid chlorides (**C1**−**C15**) were reacted under identical conditions by treating a suspension of NanoKans and Et_3N in CH_2Cl_2 at 0 °C with the reagent and 4-DMAP, followed by warming to 25 °C and stirring for 12 h. The Mitsunobu reactions were carried out by addition of Ph_3P, amines (**F1**−**F8**), and DEAD to a suspension of NanoKans in CH_2Cl_2 at 25 °C followed by agitation for 48 h. At this point, the microreactors were re-pooled and sorted, resulting in 5 sets for glycosidation reactions and the remainder for cleavage. The glycosidation reactions were conducted using the previously established conditions whereby a suspension of microreactors, trichloroacetimidates (**G1**−**G5**),[6] and 4 Å molecular sieves in CH_2Cl_2 at −40 °C was treated with $BF_3{\cdot}Et_2O$ and allowed to warm to 0 °C over a 12 h period.

When the reactions were completed, all NanoKans were sorted into Clevap plates (96 microreactors per plate with one microreactor per well) which were fitted with bar coded receiving plates. The assemblies were then taken in sets of 23 and cleaved using the automated Clevap system (see Figure 6). In each event, the individual wells of the plates containing the NanoKans were filled with a solution of hydrogen peroxide in THF (6.0 equiv) and the suspension was degassed under vacuum and allowed to react for 30 min, after which time the solution was transferred (via centrifugatis) to the receiving plates. A solution of methyl sulfide in THF (15.0 equiv) was then dispensed into the individual wells and transferred to the receiving plates. This latter addition functioned both as a wash of the resin and as a quench of the excess oxidant. Finally, the NanoKans were rinsed with CH_2Cl_2 and the wells of the microtiter plates were then concentrated under vacuum centrifugatis for 2 h to provide the targeted compounds with one member in each microtiter plate well. The location/structure of individual compounds was readily traceable using the archives generated from the loading of the cleavage plates (see Figure 6).

Quality Control Assessment of the Library. With the chemistry complete, it was necessary to assess the integrity and purity of library members so as to ensure that biological screening studies would not be confounded by false positives or other complications which might arise if the composition of library members was not sufficiently established.[16] To ensure against this problem, four types of quality controls (QC) were employed during the library development and construction process. The first of these controls occurred prior to the actual library synthesis where, as described in the above section, all building blocks and scaffolds were tested and validated to ensure high conversion for individual steps. The three remaining quality control checks occurred either during or after the library synthesis and sought to address three criteria, including the following: (a) "in situ" monitoring of reactions during library synthesis to ensure complete conversion; (b) performance of automated cleavage protocol (as compared to standard cleavage protocol used previously in the earlier focused libraries[1]) in terms of both compound purity and yield; and (c) overall library purity and integrity as determined by analysis of a representative

(16) Fitch, W. L. *Annu. Rep. Comb. Chem. Mol. Diversity* **1999**, *2*, 33−39.

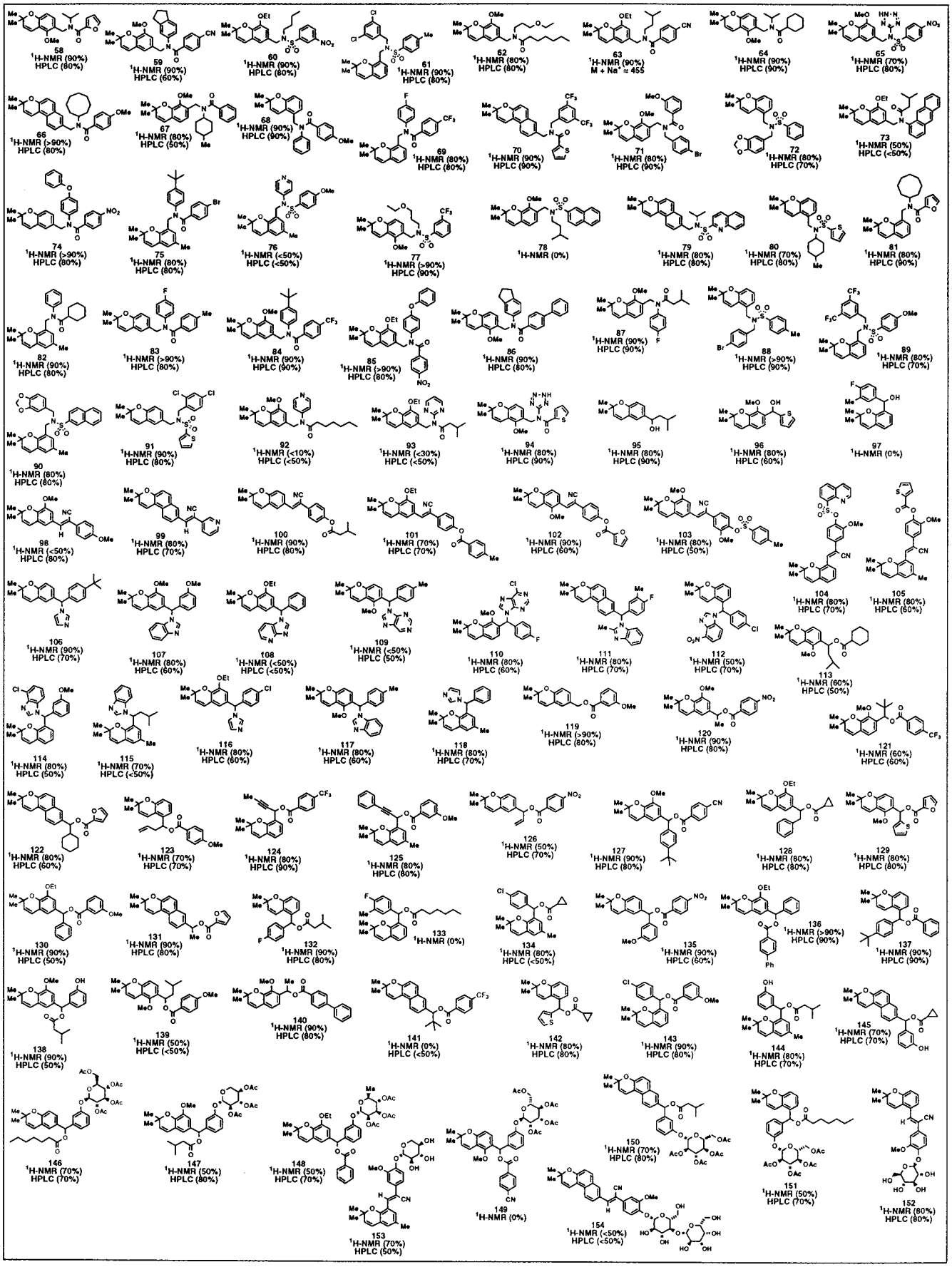


Figure 7. Structures of the systematically selected QC set (1% of library) used to monitor reaction completion and cleavage fidelity. All compounds were analyzed by ^{1}H NMR and HPLC (see Supporting Information for data) with estimated purities as shown. In addition, all structures were confirmed by MS except for compounds **78**, **97**, **141**, and **149** for which no parent or fragmentation peak was observed.

5% sampling of library members. A key component to criteria (a) and (b) was a set of 100 systematically selected library members (see Figure 7 for structures) which were chosen such that each reaction used to construct the library was represented by at least one of these compounds.[17] The selected compounds were actually encoded in quadruplicate (i.e. at the start of the library synthesis there were four separate copies with each copy in its own NanoKan). Three of these four copies were used for analytical purposes while the fourth remained in the final library so that the quality control studies would not result in a void.

To effectively monitor "in situ" reaction progress during the library synthesis, the first of the three QC sets of these compounds (Figure 7) were encapsulated in black NanoKans as compared to the standard white NanoKans which allowed them to be visually sorted out during a reaction. Thus, during (or after) each reaction, one of the designated black QC NanoKans was removed and subjected to cleavage such that the conversion of that reaction could be estimated. If analysis of the cleavage product revealed the reaction to be incomplete, all of the NanoKans in that reaction vessel were resubmitted to the reaction conditions using fresh reagents to drive the reaction to completion.

A second quality control concern regarded the fidelity of the automated cleavage protocol. Of particular interest was whether the automated system would introduce additional impurities (leached from either the apparatus itself or the cleavage plates) into the compounds. Hence, the second and third copies of these QC control compounds (Figure 7) were carried through the entire reaction sequence and then sorted out at the end. The first of these two QC sets was cleaved manually using the standard oxidative treatment and methyl sulfide quench (described in the preceding paper[1]) while the second set was cleaved using an identical *chemical protocol*, but in the Clevap apparatus. These two sets of QC compounds were then analyzed for purity and integrity by TLC, MS, HPLC, and ^{1}H NMR. In addition, the averaged yields of each QC set were determined.

The final quality control criteria required estimating the overall purity of the library through analysis of a representative 5% sample of members. These 500 compounds consisted of the 100 used above (Figure 7) along with 400 additional randomly selected compounds. All 500 compounds were analyzed by TLC, HPLC, and MS. Purities were estimated from the HPLC data based on integration and reported as 90%, 80%, 70%, 60%, 50%, <50%, or 0%. Purities were also assigned based on NMR data (when available). In making the NMR evaluations, residual water or DMSO was not considered, and examination of the signals of the aromatic and vinyl protons along with other distinguishing signals was used to estimate each compound's relative purity. Tabular HPLC, TLC, MS, and NMR data as well as HPLC traces (5% of library) and ^{1}H NMR spectra (1% of library) for crude cleavage products are provided in the Supporting Information.

Inspection of the results of these QC analyses provided encouraging evidence that the targeted library had indeed been obtained in high fidelity and good overall purity. First, comparison of the TLC, HPLC, and ^{1}H NMR data for the compounds in the two sets of QC compounds that were cleaved both manually and automatically [e.g. criteria (b) above] revealed the purities of the two sets to be strikingly similar with

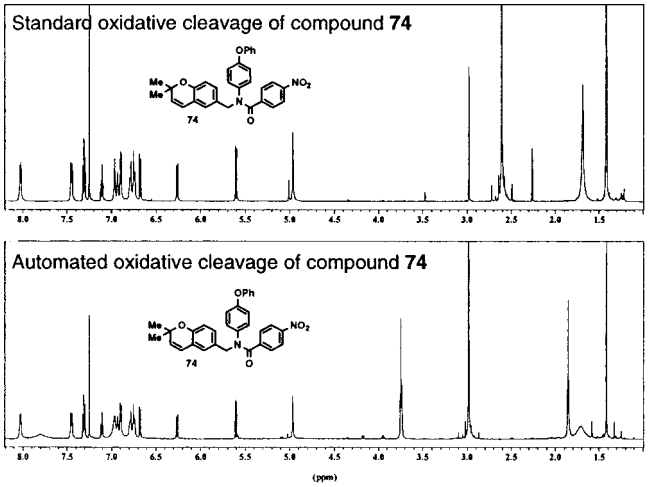


Figure 8. Comparation of standard and automated cleavage of compound **74**. ^{1}H NMR (500 MHz) were recorded in $CDCl_3$.

no notable differences attributable to accumulated impurities. A representative example is shown in Figure 8, where the ^{1}H NMR spectra of benzopyran **74** cleaved manually (top spectrum) and cleaved automatically (bottom spectrum) were identical except for small differences in the amounts of residual solvents (i.e. THF, DMSO, and H_2O). Furthermore, the average yield of each of the two QC sets was determined by carefully drying and weighing 30 compounds from each set. The average amount of compound released for the manually cleaved NanoKans was found to be 1.8 mg. Assuming an average molecular weight of 400, this recovery represented an average overall yield of 51% over 3 to 7 steps depending on the identity of the particular compound. The average amount of material released from the NanoKans cleaved automatically was determined to be 1.6 mg or about 89% of the value released during the manual cleavage studies. The similarity in compound purity and yield between these two QC sets attested to the efficiency and utility of the automated Clevap system.

Finally, inspection of the data from the 5% library sampling [e.g. criteria (c) above] provided a decisive measure of overall library quality. As previously described, these 500 compounds were analyzed by HPLC, TLC, and MS. In addition, 1% of the library was also analyzed by ^{1}H NMR. During the mass spectroscopy studies, confirmatory peaks for 458 of the 500 compounds (92%) were identified. In addition, HPLC analysis found that 347 of the 500 compounds (69%) were of 80% purity or greater while 418 of the 500 compounds (84%) were of 70% purity or greater as determined by integration.[18] In addition, evaluation of a 1% library sampling (see Figure 7) by ^{1}H NMR suggested that 75% of the compound samples were of 80% purity or greater while 82% of the samples were of 70% purity or greater.[18]

Besides confirming overall library quality, this analysis also revealed several interesting reactivity trends. First, in terms of purity, the reductive amination and Kneovanagel condensation sequences appear to have outperformed the organometallic addition sequence. In fact, the only notable failures in the reductive aminations stemmed from the heterocyclic amines **A18**−**A20** (Figure 1) likely due to limited solubility as discussed previously. As expected, the Kneovanagel condensations pro-

(17) The compounds shown in Figure 7 are a subset of the 500 library members used in the quality control studies. Examination of their structures and purities provides a concise overview of the structural characteristics of library members as well as the purities observed for different types of library members. Complete tabulated data for all 500 QC compounds can be found in the Supporting Information.

(18) The assessment of purities of glycosides derived from organometallic addition reactions by HPLC or ^{1}H NMR was partially obscured by the fact that the compounds were synthesized as a 1:1 mixture of diastereomers since the carbon of the benzylic alcohol was racemic. The reported purities for these compounds are, therefore, best estimates given this fact.

ceeded in good conversion and stereoselectivity with both electron-deficient and electron-rich phenyl acetonitriles. Somewhat unexpectedly, compounds sampled from the organometallic addition route were slightly less clean (average purities in the range of 70−80%) than was anticipated based upon the building block validation studies where all reagents tested exhibited purities of 90% or greater. Notwithstanding these minor perturbations, these sampling studies confirmed that, overall, members of this library were sufficiently pure to be used for biological screening.

Conclusion

In conclusion, we have described a new strategy for the construction of natural product-like libraries based on privileged scaffolds. Application of this strategy to the 2,2-dimethylbenzopyran structural motif required the development of a solid-phase selenium-based cycloloading methodology for the construction and elaboration of structures containing this template. Integration of this chemistry with the new IRORI NanoKan optical encoding platform has resulted in the rapid construction of a 10 000-membered natural product-like library, whose members are obtained in quantities of 1−2 mg each, to be used in chemical biology studies. In the following paper, we report a unique strategy whereby the diversity and potential pharmacological properties of these benzopyran libraries can be further enhanced through parallel solution-phase combinatorial chemistry.[2]

Currently, the discovery-oriented library described herein as well as several of the smaller focused libraries reported in the preceding paper are being employed in multiple biological studies. Specifically, these combined libraries (and subsets thereof) are being screened in both in vitro enzyme inhibition assays as well as in several cell and animal systems through collaborations with laboratories at The Scripps Research Institute, The National Institutes of Health, and other academic institutions with initial interest in anticancer, antibacterial, and antiviral targets. Preliminary results from these studies suggest that these library members are in fact cell-permeable and capable of high-affinity interactions with biological targets.[19] Additional results from these investigations will be reported in due course.

Acknowledgment. We would like to thank the staff of Discovery Partners International Company and particularly Chris Castello, Leslie Eastman, Roger Howard, and Rob Neeper for their generous assistance. Financial support for this work was provided by The Skaggs Institute for Chemical Biology and the National Institutes of Health (U.S.A.), fellowships from the Department of Defense (to J.P.), the American Chemical Society Division of Medicinal Chemistry (sponsored by Parke-Davis, Hoescht Marion Roussel, Abbott, Wyeth Ayerst, and Bristol Myers Squibb) (to J.P.), and the Ministerio de Educacion y Cultura, Spain (to S.B.), and grants from Amgen, Boehringer-Ingelheim, Glaxo, Hoffmann-La Roche, DuPont, Merck, Novartis, Pfizer, and Schering Plough.

Supporting Information Available: Experimental procedures and results (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA002034C

(19) Nicolaou, K. C.; Pfefferkorn, J. A.; Schuler, F.; Roecker, A. J.; Cao, G.-Q.; Casida, J. E. Unpublished results.