A paradigm for drug discovery employing encoded combinatorial libraries

(medical chemistry/isozyme selectivity/combinatorial chemistry)

JONATHAN J. BURBAUM*, MICHAEL H. J. OHLMEYER, JOHN C. READER, IAN HENDERSON, LAWRENCE W. DILLARD, GE LI, TROY L. RANDLE, NOLAN H. SIGAL, DANIEL CHELsky, and JOHN J. BALDWIN

Departments of Biology and Chemistry, Pharmaceopia, Inc., 101 College Road East, Princeton, NJ 08540

Communicated by Michael H. Wigler, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, February 8, 1995 (received for review November 1, 1994)

ABSTRACT Very large combinatorial libraries of small molecules on solid supports can now be synthesized and each library element can be identified after synthesis by using chemical tags. These tag-encoded libraries are potentially useful in drug discovery, and, to test this utility directly, we have targeted carbonic anhydrase (carbonate dehydratase; carbonate hydro-lyase, EC 4.2.1.1) as a model. Two libraries consisting of a total of 7870 members were synthesized, and structure–activity relationships based on the structures predicted by the tags were derived. Subsequently, an active representative of each library was resynthesized. Each of these compounds was shown to have nanomolar dissociation constants (15 and 4 nM, respectively). In addition, a focused sublibrary of 217 sulfamoylbenzamides was synthesized and revealed a clear, testable structure–activity relationship describing isozyme-selective carbonic anhydrase inhibitors.

The current national focus on health care reform has highlighted the continuing importance of developing more cost-effective methods for treating disease. Historically, the pharmaceutical industry has contributed to this effort by discovering and developing therapeutic molecules. The development of drug candidates is expensive (1, 2), and preclinical research toward the discovery and development of lead structures contributes significantly to this expense. For example, in 1991, the costs of medicinal chemistry and biological testing prior to safety assessment amounted to nearly 30% of domestic research costs (3). Apart from the economics of drug discovery, classical medicinal chemistry cannot efficiently address the plethora of new biochemical targets suggested by recent discoveries in molecular genetics, and lead structures from natural sources are often too complex for the cost-effective synthesis of analogs. Hence, targets for therapeutic intervention cannot be rapidly and efficiently exploited.

Combinatorial chemical synthesis applied to drug discovery promises to improve the productivity of medicinal chemistry both by significantly increasing the number of molecules available for testing and by providing facile routes toward synthetic analogs of active molecules. Many strategies for the generation of chemical diversity have been proposed (4–7). With a single exception (4), these approaches have been confined to the synthesis of flexible oligomeric ligands of relatively high molecular weight (e.g., peptides and oligonucleotides) which tend to be poor therapeutic agents, in part because of their lack of availability and stability in vivo.

For drug discovery, small-molecule libraries containing diverse functionality in a variety of molecular scaffolds possess the greatest utility. To construct such structurally diverse libraries while allowing the assignment of a structure to each member, we use a set of chemically stable molecular tags during solid-phase synthesis on polymeric beads (8). These tags can be used to unambiguously encode each bead with the synthetic scheme for its library member and are sufficiently robust to allow a wide range of chemical reactions during ligand construction.

A drug discovery effort based on such combinatorial libraries must pass several tests to be successful. First, the library must be diverse enough to permit the identification of a small subset of active molecules among many less active or inactive ones. Second, the most promising candidates in the library should suggest a testable structure–activity relationship. Third, since selectivity for a particular target is often important, strategies for the design of libraries containing a significant fraction of active compounds should result in optimization of selectivity as well as activity.

To explore the capabilities of encoded combinatorial libraries in a small-molecule drug discovery effort directed toward lead identification and optimization, we chose carbonic anhydrase (carbonate dehydratase; carbonate hydro-lyase, EC 4.2.1.1) as a therapeutically significant model, since inhibitors of carbonic anhydrase are known to be useful in ameliorating the symptoms of glaucoma. The inhibition of carbonic anhydrase by compounds containing primary sulfamoyl groups (-SO2NH2) has been well-characterized, both pharmacologically (9–11) and structurally (12–15). This allowed the design of test libraries that were predisposed toward showing carbonic anhydrase inhibition without explicitly including known inhibitors. Finally, since a number of different forms of the enzyme have been characterized, we were able to perform experiments to probe isoform selectivity.

MATERIALS AND METHODS

Materials. TentaGel S-NH2 resin (particle size, 130 μm) was obtained from Rapp Polymere (Tubingen, Germany). The photocleavable linker was synthesized as described (16). Carbonic anhydrase was obtained from Sigma. White 96-well polystyrene plates were obtained from Dynarec.

Library Synthesis and Analysis. Libraries were synthesized by the split synthesis protocol, and each bead thus contains one member of the library (17–19). All members were affixed via a photocleavable o-nitrobenzyl linker. To preserve the synthetic history of each bead, each library was indexed in a binary fashion using electrophoric tags. These tags were attached by using [CF3CO2]2Rh]-activated carbene insertion as described (20). After identification and isolation of beads, the tags were analyzed by gas chromatography (8).

Abbreviations: DNSA, dansylamide; hCA(I) and hCA(II), human carbonic anhydrase isozymes I and II.

*To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.
Synthesis of Dihydrobenzopyran Library. First, three dihydroxyacetophenones (2-acetyl-1,4-dihydroxybenzene, 2-acetyl-1,5-dihydroxybenzene, and 2-acetyl-6-methyl-1,5-dihydroxybenzene) were coupled using Mitsunobu conditions (triphenylphosphine/diethylazodicarboxylate in tetrahydrofuran at ambient temperature) to the o-nitrobenzyl linker, which was then coupled to the beads by carbodiimide coupling [di(isopropylcarbodiimide)/hydroxybenzotriazole in dimethylformamide]. These intermediates were cyclized on the resin with a set of seven ketones, four of which contained a protected amine functionality. Modification of the amine functionality (when present) with 31 different headgroups under standard conditions produced a library of ketones containing 381 members. The ketones were then reduced with NaBH₄, converted into the corresponding dithiolanes with ethane-1,2-dithiol, or left unaltered to produce a library of 1143 distinct compounds.

Synthesis of Acytopiperidine Library. Library elements were coupled as either a carboxylic ester or a carbonate ester of an o-nitrobenzyl alcohol. First, five piperidines (2-hydroxymethyl, 3-hydroxymethyl, 4-hydroxy-, 4-carboxyl-, and 3-carboxyl-), and two linear moieties (5-amino-3-oxopentan-1 and 6-aminohexanoic acid) were attached to this linker either by phosgene coupling in solution (for the alcohols) or directly to the resin (for the acids). Next, 31 amino acids, including 28 examples of both D and L isomers of the 20 commonly occurring α-amino acids, and 3 nonproteinogenic amino acids (piperidine-4-carboxylic acid, piperidine-3-carboxylic acid, and 6-aminohexanoic acid) were reacted with the free amine. Finally, 31 reagents were linked as carbamides, carbamates, sulfonamides, and ureas, including three primary sulfonamides (4-sulfamoylbenzamido, 4-chloro-5-sulfamoylbenzamido, and 2,4-dichloro-5-sulfamoylbenzamido).

Library Analysis for Carbonic Anhydrase Inhibition. Beads were distributed into 96-well microtiter plates. Compounds were detached from the beads by UV irradiation (350 nm) and then transferred to assay plates. An assay solution (50 μL) consisting of 0.1 M phosphate buffer (pH 7.4) containing bovine carbonic anhydrase (0.5 μM) and dansylamide (DNSA; 0.6 μM) was added to each well. Fluorescence values (λex, 274 nm; λem, 454 nm) were measured with a Perkin–Elmer model LS 50B spectrophotometer equipped with a microtiter plate reader accessory and were normalized (uninhibited, 100; empty, 0). Assay solutions from the initial screen identified as active were transferred to a second fluorescence plate and 5 μL of 2 mM DNSA in dimethyl sulfoxide was added to each well to increase [DNSA] ∼600-fold in situ. Under these conditions, chlorothiazide (Kd ∼ 75 nM) was displaced, while acetazolamide (Kd ∼ 7.5 nM) was not, suggesting that more potent inhibitors were less sensitive to displacement by elevated [DNSA].

Selectivity Analysis. Aliquots (one-third) from the same well. Concentrations of the two isozymes were matched using Kd values from the literature (21), and the [DNSA] was varied according to the literature values for Kd (DNSA) of the isozymes (22).

RESULTS AND DISCUSSION

Two model libraries for drug discovery—a 1143-member dihydrobenzopyran library and a 6727-member acytopiperidine library—were designed. The first (dihydrobenzopyran) library (Fig. 1A) was designed to include 4-sulfamoylbenzoyl amides that were expected to confer carbonic anhydrase activity. The second (acytopiperidine) library (Fig. 1B) was designed around a core aminocarboxamide derived from 31 amino acids, including nonproteinogenic amino acids and both the D and L configurations of several naturally occurring α-amino acids. The carboxamide was formed from the carboxylic acid with seven amines including four substituted piperidines. Three different sulfamoylbenzoyl groups were incorporated at the final step to confer carbonic anhydrase activity.

Activity of library members against bovine carbonic anhydrase was assessed by using a fluorescence-based ligand-displacement assay (9). Because each synthetic library contained >1000 beads per member, we characterized the activity of the library statistically to optimize our screening conditions. The dihydrobenzopyran library was determined to contain ∼1.4% actives, identified as those wells where the fluorescence decreased by >5σ. Similarly, in the acytopiperidine library, we determined that the samples contained ∼6% actives. These observations were consistent with the percentage of primary sulfonamides incorporated in the two libraries, which was predicted to be 3% and 10% for the dihydrobenzopyran and acytopiperidine libraries, respectively.

Over 2300 beads from the dihydrobenzopyran library were then assayed singly and 33 individual beads were chosen for decoding. Thus, 2.0 library equivalents were assayed, where one library equivalent is defined as the number of beads equal to the number of distinct library members. The probability that a given compound was not assayed is approximated as follows: In a library of N total beads >> M, the total number of library members, when S beads have been assayed, the probability that a given member has not been chosen is given by

\[ P_S(0) \approx \left(1 - \frac{1}{M}\right)^S, \]  

which is approximately \((e^{-1})^S/M\) for large S. The quantity \(S/M\) is \(L\), the number of library equivalents that have been assayed, so \(P_S(0) \approx e^{-L}\). For the case at hand, the fraction of the library assayed at least once is \((1 - e^{-2})\) or 86%.

The larger acytopiperidine library was pooled at the final reaction step into two smaller portions to permit a more convergent assay strategy. These pools consisted of 3472 and 3255 members, the larger portion containing all the primary sulfonamides. For the smaller portion, ∼5 library equivalents (∼17,000 beads) were assayed at 10 beads per well, and no actives were identified. Thus, in Eq. 1, \(L \approx 5\), and \(P_S(0) \approx 0.01\). Hence, >99% of the members were assayed at least once, and none were found to be active under the conditions described.

We therefore focused our continued efforts on the portion of the acytopiperidine library that contained primary sulfonamides. Here, half the eluates from a total of 4320 single beads (1.3 library equivalents) were assayed individually, and >300 actives were identified (Fig. 2A). To estimate the relative potency of the active inhibitors, the stringency of the assay was increased by raising the concentration of the competitor, DNSA. However, the associated background fluorescence made the fluorescence decrease caused by inhibitors less easily detected. Potent actives were therefore judged as differing by >3σ compared to the median active compound. Eighteen high-affinity actives were thereby identified and decoded (Fig. 2B).

As anticipated, we found that carbonic anhydrase inhibitors from either library exclusively contained the sulfamoyl group (-SO₂NH₂). From the actives, we chose two synthetic targets, one from the dihydrobenzopyran library (compound 1) and one of the more potent actives from the acytopiperidine library (compound 2). These compounds were resynthesized, their structures were confirmed spectroscopically, and their Kd values vs. carbonic anhydrase were determined experimentally (Table 1). The low nanomolar values for compounds 1 and 2 support the use of tag-derived information for structure–activity assessments in the libraries we report.

Significantly, no structures containing a 4-chloro-5-sulfamoylbenzoyl moiety were found among the tag-predicted structures for potent actives in the acytopiperidine library. Furthermore, while the other two sulfamoylbenzamides were
well-represented, they showed markedly different structure–activity relationships. For example, the 4-sulfamoylbenzamides appear to prefer a hydrophobic R2 group of the L configuration, while the 2,4-dichloro-5-sulfamoylbenzamides appear to tolerate more diversity at R2, including D-alanine. Thus, to test the validity of this observation, library-specific homologs of the acylpiperidine (compound 2) were synthesized (Table I, compounds 3 and 4). The 150-fold increase in $K_a$ seen upon substitution further validates the ability of the screen to select active compounds and supports the divergent structure–activity relationships observed with the various sulfonamide headgroups.

In developing a class of compounds for use as therapeutic agents, increased affinity is but one criterion for optimization. Isozyme selectivity is often an additional critical property required to minimize side effects. The selectivity of a large library is difficult to optimize convergently (i.e., with >1 element per assay), because such assays must measure the absence of activity against a particular target or relative levels of activity vs. multiple targets. As the percentage of actives increases, the probability that the observed affinities are the property of multiple elements increases. Thus, approaches taken to identify high-affinity peptides in a convergent fashion (e.g., see ref. 6) become impractical with selectivity, and focused sublibraries assayed singly become pertinent. To evaluate the potential for using encoded combinatorial libraries to optimize selectivity, a 217-member sublibrary (consisting of the 4-sulfamoylbenzamides from the larger acylpiperidine library) was prepared and evaluated against human carbonic anhydrase isozymes I and II [CA(I) and CA(II)].

To isolate selectivity as a variable, two factors were normalized. First, equal inhibitor concentrations were exposed to each isozyme by taking aliquots from the same bead eluate. Thus, one-third of a complete eluate was assayed vs. human CA(I), and one-third was assayed vs. CA(II). Second, the concentrations of the two isozymes were matched, but the
[DNA] was varied according to the literature values for $K_{DNA}$ of the isozymes. These requirements derive from the binding equation, given relative to isozyme $x$

$$F_x \propto \frac{[E]_{tot,x}}{[DNA]_x} = \frac{[E]_{tot,x}}{1 + \frac{K_{DNA,x}}{[DNA]_x} \frac{[I]}{K_{I,x}}}$$

[2]

Thus, when $E_{tot}$ is equal, the ratio $[DNA]/K_{DNA}$ is equal, and $[I]$ is equal, the only variable that can affect $F$ is the dissociation constant for the inhibitor, $K_i$. To identify samples for which there is a significant difference in $K_i$ for the two isozymes, the normalized fluorescence values were analyzed graphically (Fig. 3).

Table 1. Analysis of compounds and validation of tag-derived structure–activity relationships

<table>
<thead>
<tr>
<th>Compound</th>
<th>$K_d$[bCA(II)], nM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>H, SO$_2$NH$_2$, H</td>
</tr>
<tr>
<td>3</td>
<td>H, Cl, SO$_2$NH$_2$</td>
</tr>
<tr>
<td>4</td>
<td>Cl, Cl, SO$_2$NH$_2$</td>
</tr>
</tbody>
</table>

Fig. 3A shows a selectivity plot comparing the normalized fluorescence values for the two enzymes. In this figure, region A contains compounds that show potency vs. both isozymes, while region B contains compounds that are less selective for CA(II) than the median, and region C contains compounds that are more selective than the median. Sulfonamides that are not particularly active against either isozyme are found in region D. Beads from each region were decoded, and examples of the structures predicted by these codes are shown in Fig. 2B. The potent compound that was identified in screening against bovine carbonic anhydrase (Table 1, compound 2) was identified by 1H NMR spectroscopy and mass spectrometry. $K_d$ values were measured as described (7) with EXCEL software to solve the multiple equilibria numerically. b, Bovine.
We wish to thank Pharmacia’s engineering department, in particular Joseph Brzezinski and Jeffrey Davidson, for providing plates containing arrayed beads, and Peter Kieselbach and Nicole White for electron capture gas chromatography analysis of tagged beads. We also thank Laura Rokosz (Department of Biology, Pharmacia) for expert HPLC analysis. We thank Dr. Harvey Schwam, Professor W. Clark Still (Columbia University, New York), Professor Paul D. Bartlett (University of California, Berkeley), and Professor Tom Maren (University of Florida) for helpful scientific discussions and John Chabala, Gabriel Lopez, and Joseph Mollica (Pharmacia) for careful reading of the manuscript. Michael H. Wigler is a consultant to Pharmacia in the area of combinatorial chemical synthesis.