"Lead Hopping". Validation of Topomer Similarity as a Superior Predictor of Similar Biological Activities

Richard D. Cramer,*,† Robert J. Jilek,† Stefan Guessregen,‡,§ Stephanie J. Clark,‡ Bernd Wendt,‡ and Robert D. Clark†

Tripos Discovery Research, Bude-Stratton Business Park, Bude, Cornwall, EX23 8LY, U.K., and Tripos Inc., 1699 South Hanley Road, St. Louis, Missouri 63144

Received June 23, 2004

Two extensive studies quantifying the ability of topomer shape similarity to forecast a variety of biological similarities are described. In a prospective trial of "lead hopping", using topomer similarity for virtual screening and queries from the patent literature, biological assays of 308 selected compounds (representing 0.03% of those available, per assay type) yielded 11 successful "lead hops" in the 13 assays attempted. The hit rate averaged over all assays was 39% ("activity" defined as inhibition ≥20% at 10 μM), significantly greater than an unexpectedly high negative control hit rate of 15%. The average "Tanimoto 2D fingerprint similarity" between query and "lead hop" structures (0.36) was little more than the Tanimoto similarity between random drug-like structures. Topomer shape and Tanimoto 2D fingerprint similarities were also compared retrospectively, in their tendencies to concentrate together potential and actual drugs reported to belong to the same "activity class", for twenty classes. Among the most similar 3% of structures (corresponding to "≥0.85 Tanimoto" for these structures), an average of 62% of the topomer similar selection possessed a near neighbor belonging to the same activity class, roughly a one-third superiority over the "Tanimoto ≥ 0.85" selection containing 48% actives in avoiding false positives. Conversely, the least similar 75% of structures contained 0.3% actives for topomer similarity vs 1.0% actives for Tanimoto 2D fingerprint similarity, a 3-fold superiority for topomers in avoiding false negatives.

Improved prediction of the biological properties of specified chemical structures is the main goal of most computer-aided molecular design activities. Yet the validation of new prediction methodologies is usually limited to retrospective analysis of one or two literature data sets. Prospective methodology validation, though surely preferable, requires costly experiments, the testing and perhaps synthesis of many compounds. So such experiments are usually done only in connection with a larger goal of discovering promising biological activity. The resulting proprietary considerations then inhibit publication of the experimental results and may also bias the experimental design unfavorably for methodology validation.

Several general issues may be considered when evaluating any compound selection methodology validation study, retrospective or prospective.

How Diverse and Numerous Are the Validation Data Sets, Both Structurally and Biologically? The structural variety of both receptors and possible ligands is so immense that when only a few data sets are considered, there is a large risk of any apparent relationship being artifactual. A case in point is the frequency of 2D fingerprint similarity predicting biological similarity. The initial report that a compound with a 2D-fingerprint Tanimoto coefficient ≥0.85 to an active compound had an 80% chance of sharing that activity,[1] apparently confirmed by other workers,[2,3] had to be substantially lowered to a 30% chance of sharing activity when the results from testing of compounds in 115 assays were analyzed.[4]

What Sorts of Unrecognized Structural Bias May Exist within the Validation Data Sets? Published data sets are usually "polished" subsets, from which the more awkward results have been filtered. And, as has long been realized,[5] even "complete" screening sets tend to be structurally biased, containing disproportionately more of structures that apparently resemble other promising structures. Although such structural biases need not affect side-by-side comparisons of one methodology with another, they do make the absolute performance of any methodology very difficult to assess. To some degree, these performance uncertainties can be alleviated by inspecting representative structural selections.

Is Any Aspect of the Validation Prospective? As mentioned few methodology validations have been prospective. Although retrospective studies can draw on the many data sets freely available from the published literature, prospective validation experiments obviously can be far less vulnerable to bias. However, the much higher cost of prospective experiments always tends to limit their diversity and number.

Have Negative Control Experiments Been Performed? Or are the results of such experiments simply assumed? For example, a report that a methodology contributed to a particular pharmaceutical discovery,

* Author to whom correspondence should be addressed. Present address: Tripos Associates, Inc., c/o National Center for Genome Resources, 2935 Rodeo Park Drive East, Santa Fe, NM 87505. Phone: (505) 995-4425 or (314) 647-1099. Fax: (505) 995-4439. E-mail: cramer@tripos.com.

† Tripos Inc.

‡ Tripos Discovery Research.

§ Present address: Aventis Pharma Deutschland GmbH, Building G878, 65926 Frankfurt am Main, Germany.

10.1021/jm049501b CCC: $27.50
Published on Web 12/07/2004

though welcome, will surely lack a negative control: no one can say whether a discovery would have been made without the methodology. Negative control experiments are especially onerous in prospective validations when novel biological activity is also sought, as an apparent waste of resources. However, our own initial expectations in the following work, that negative control experiments were almost superfluous because the hit rates from testing random compounds would surely be close to zero, proved quite wrong.

Do the Validation Results Include Any Direct Comparison with Alternative Approaches? Because comparisons with alternatives are less vulnerable to data set biases than are absolute performances, they can be done retrospectively and provide useful confirmation to the results of prospective experiments. Also advantage can thereby be taken of the large amounts of retrospective data, to offset any concerns about generality caused by the relatively limited data ranges practical in prospective validation.

We have for many years been applying topomer similarity as a predictor of similarity in biological activity, mostly in proprietary projects with collaborators. A topomer is an invariant 3D representation of a molecular fragment, generated from its 2D topology by deterministic rules that produce absolute coordinates for each of its constituent atoms.[6] Topomers were originally devised to enable shape similarity searching of very large virtual libraries,[7,8] within ChemSpace. It was also observed in retrospective[2] studies that topomer shape similarity seemed to be at least as powerful a predictor of biological similarity as any other molecular descriptor investigated. In a prospective study, the degree of topomer similarity to the "nearest" of four A2 antagonist query structures correlated with the hit rate in the A2 assay, while correctly identifying a few unexpectedly active ligands.[9] More recently it was reported[10] that topomers provide a surprisingly robust and general solution to the vexing challenge of CoMFA alignments.

However neither of the two published topomer validation studies[2,9] was large or complete enough to address the above list of general validation issues in a more than preliminary way. So to more fully assess and illustrate the effectiveness of topomer similarity in predicting biological similarity, we undertook and now report two much more extensive studies: (1) prospective "lead hopping", by biological testing of 308 compounds that were topomerically most similar, from among approximately 80000 candidates, to 14 query structures including 13 different kinds of reported biological activity; (2) retrospective comparison of topomer similarity with the standard "Tanimoto 2D fingerprints", in their propensities for concentrating together structures reported to share a particular class of biological activity or "activity class", for 20 such activity classes as reported in the World Drug Index.

Methods

Topomer Methodologies. Complete details of current procedures for topomer shape similarity comparison and topomer generation have recently been published.[11] Topomer shape similarity differs from most other shape comparison methods in several fundamental ways:

(1) Topomer shape comparisons consider all the atoms, in contrast to pharmacophore-based "3D searching" approaches where "shape" comparison focuses on a small subset of atom-like features.[12]

(2) Topomer shape comparisons are produced by combining a set of fragment-to-fragment differences, rather than from some operation on complete structures.

(3) Topomer shape comparison usually involves only one "topomer" conformation for each fragment, rather than the indefinitely large variety of conformations that most fragments are capable of achieving.

(4) Topomers resemble one another, not to the extent that achievable geometries among assumed critical features are shared, but mostly to the extent that (a) the same spatial volume or "field" elements are occupied or avoided by the corresponding fragments (in their topomer conformations) and (b) similar features occupy similar positions within a fixed Cartesian reference frame.

(5) The topomeric difference between two molecules is the minimum value found among the many fragment set comparisons that are usually possible.

The procedure for generating the topomer of a monovalent molecular fragment may also be summarized as follows:

(1) A structurally distinctive "cap" is attached to the open valence, and a CONCORD[13] (3D) model is generated for the resulting complete structure.

(2) This model is oriented to superimpose the "cap" attachment bond ("root") onto a vector fixed in Cartesian space.

(3) Proceeding away from this "root" attachment bond, only as required to place the "most important" (typically the largest) unprocessed group farthest from the root and the next most important to the "right" of the largest (when looking along the bond away from the root), acyclic torsional angles may be adjusted, "stereocenters" inverted, and ring "puckerings" standardized.[11]

(4) Removal of the cap completes the topomer conformation.

Note that the conventional force field energy (intramolecular enthalpy) of a topomer is immaterial. For example, any internal steric clashes that may result are completely ignored.

Topomer Similarity Selection of Screening Candidates. The results reported here constitute the entirety of our first two rounds of "lead hopping": compound selection and testing. Some slight procedural differences between the first and second rounds are mentioned where appropriate.

Query structures were chosen from the weekly World Drug Alert[14] service, by filtering to exclude structures containing molecular weights outside the range 100−800, and any substructures not allowed in LeadQuest designs, notably metallic elements. Taking one typical three-week period as an example, these criteria yielded 352 usable query structures from 771 World Drug Alert (WDA) records (46%).

The source of screening candidates was the current LeadQuest inventory, which numbers roughly 80000 compounds, most composed of some 25 novel cores or scaffolds bearing two side chains or "R groups". Considerable design effort ensured that the vast majority of these candidate compounds were "drug-like" (fulfill Lipinski RO5 criteria and lack toxic substructures) and as diverse from one another as full array synthesis permitted, and that the library templates (cores/scaffolds) had some medicinal rationale. (These design criteria are emphasized as a possible partial rationalization of the unexpectedly high "random hit rate", discussed below.)

For each query structure, the topomer similar compounds within LeadQuest were identified by searching with the *dbtop* program,[15] including pharmacophoric features as well as shape in the similarity criteria. In both rounds a rather large topomer radius of 250 was used, in order to obtain hits from certain expected libraries. The second round hits were considered more satisfactory, partially because of the correction of the conceptual oversight in topomer generation (standardization of ring "puckering") that had reduced hits from those expected libraries.

Under these conditions, 141 of the 352 queries in the example yielded at least one and at most 42113 "topomer similar" structures from the ~80000 LeadQuest records (40%). These 141 queries represented over fifty different "activity

classes". It was further required that an assay judged comparable to the WDA-reported biology be commercially offered at reasonable cost, and that the WDA-reported IC_{50} be 0.5 μM or less. These criteria reduced the 141 usable example queries to 28.

The final selections of candidate queries (plus assays) reflected a combination of considerations, many more subjective and centering on the chemical series or library. How tractable would a library be for follow-up activities, with respect to numbers of compounds currently and potentially available? Would hits from a library be viewed as "lead hops" worthy of follow-up, considering other literature reports of biological activity? How certain was the functional correspondence between the original and available biological assays? A limited number of FlexS comparisons between queries and hits were also done, in particular specifically seeking query side chains lacking a counterpart within the hits. However, these comparisons were not at all decisive; for example, two of the first round of seven query structures did not yield good FlexS comparisons. Additional pragmatic criteria for deciding to test a particular compound in a particular assay were the amount of sample available (large, to maintain inventory while supplying enough sample for possible IC_{50} determination) and the number of structures selected (at least three per core, to minimize the possibility of overlooking activity).

In the first round, mainly because of an optimistic expectation of immediately obtaining substantive SAR data, the structural selections included almost every side chain represented within multiple topomer similar structures, for a total of 257 compounds, each of course to be tested in a single specified assay. In the second round, only the three to six topomerically most similar structures derived from a particular core were tested, yielding 51 compounds, again to be tested in one of seven assays. One assay, PDE4, was assigned test candidates in both rounds. Overall 308 topomer similarity selected compounds were tested in one of thirteen different assays, an average of 24 compounds tested per assay (308/13) or ~0.03% of the available compounds (24/~80000).

All biological testing was performed at PanLabs MDS,[16] with the structural identity of the individual compounds of course being unknown to that laboratory. The thirteen assays are named in the first column of Table 1, along with catalog ID numbers. The testing concentration was generally 10 μM.[16] As the minimum criterion for an "active compound" we used "≥20% response at 10 μM", a level of response which proved reproducible in all but two of fifty subsequent IC_{50} determinations. By this criterion, in the CCR5 and COX-2 assays no active compounds were found, so for those assays IC_{50} determinations were instead performed on the apparently most active compounds (18% and 17% inhibition, respectively). Results are also reported below for the more demanding activity criteria of "≥50%" and "≥90%" at 10 μM.

Negative Control Experiments. The hit rates obtained from topomer selection required suitable negative control values for full evaluation. We at first expected that the hit rate for randomly chosen compounds would be experimentally indistinguishable from zero, but this was not found. The first round of negative controls involved cross-testing fourteen of the active compounds in two of the thirteen assays, two that were functionally dissimilar to the assay that had generated the hit. A further requirement of assay assignment in cross-testing was that each of the thirteen assays had two compounds assigned, except that PDE4 had four compounds assigned because of its two query structures. Significant inhibition was seen from six of these 28 tests. So the second round assigned 28 compounds randomly chosen from LeadQuest to be tested into the same assay mix, i.e., two compounds in each assay except four in PDE4. Again six of 28 test results were active. The third round involved retesting all the apparently active compounds from the second round, as both original and repurified samples, by a different laboratory. These three comparable sets of test results agreed with one another, and so the two confirmatory retest results are excluded from the results reported here.

It became inescapable that the "average random hit rate" for these compounds in these negative control experiments was greater than zero and also likely to vary among the thirteen assays. Accordingly the next two rounds were designed to maximize the number of individual assays in which topomer selection would significantly enhance the hit rate over random selection, at minimum overall cost, by testing randomly selected compounds in particular assays. A final round of negative controls assigned additional randomly selected compounds to four assays for which the negative control hit rate was initially largest. Table 1 accumulates the results from all six rounds of negative control experiments, excluding only the retests mentioned.

Retrospectively Comparing Topomer Shape Similarity with Tanimoto Coefficients of 2D Fingerprints as Concentrators of Biologically Similar Structures. The World Drug Index[17] compiles literature and patent references to "marketed and development drugs". Its approximately 32000 entries were first filtered, by removing inorganic or polymeric structures and by requiring an explicit "MECHANISM" of biological activity, to yield 8832 structures. Among these, the twenty most frequently reported biological mechanisms, as determined by tabulation and sorting, are listed in the first column of Table 2. The structures recorded as acting by each of these mechanisms were identified by string searches. Later, by inspection of individual compounds that were structurally similar but apparently not biologically similar to some particular compound, a few mechanistic synonyms were also recognized and incorporated. The final mappings of text strings to biological mechanism are shown in a footnote of Table 2.

By treating every compound belonging to any of the twenty mechanistic classes as a query structure, two types of neighbor lists were constructed from similarity searches of all 8832 structures as candidates. First, its Tanimoto neighbors were identified and counted, by applying *dbsimilar* and standard Unity 2D fingerprints, for all Tanimoto threshold values (radii) between 0.40 and 0.95 in steps of 0.05. Second, its topomer neighbors were also identified and counted, using the *dbtop* program, for all topomer similarity radii between 90 and 240 in steps of 10. Finally, for each of the twenty types of activity and each type and radius of neighbor list, the neighbor lists matching these three criteria were merged, removing duplicate structures. Such a merged neighbor list then contains exactly those structures that are neighbors of any structure that acts by a particular biological mechanism, still labeled so that the subset of those neighbors that were reported to be active by that same mechanism can be identified. Counts from some of these merged neighbor lists are the starting points for Tables 2 and 3, in columns labeled "nbors" and "actv", respectively. Further analyses of these counts are described below.

Results

Table 1 summarizes the results of the prospective selection of screening candidates by topomer similarity. Each row of Table 1 corresponds to a particular biological assay, designated in its upper left-hand corner by a name and its catalog reference number in square brackets. The query structure associated with that assay is shown at the left, accompanied by a literature reference in square brackets and its reported potency (micromolar) in parentheses.[18−30] Next to the query structure is an example of a hit structure, one first identified by a topomer similarity search using the query structure shown and then found experimentally to have the same activity as the query, with its micromolar potency established by IC_{50} determination shown in parentheses. The structures of both query and example hit have been drawn to emphasize their apparent shape similarity. The right half of Table 1 provides the overall screening statistics obtained for this assay, in two lines, the upper line for topomer similarity selected

Table 1. Results from Prospective Validation of "Lead Hopping" Based on Topomer Similarity

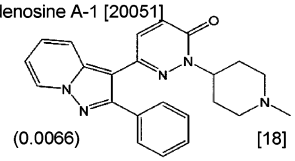
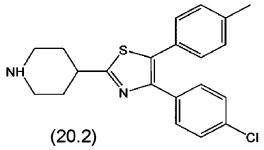
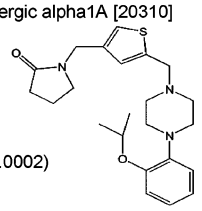
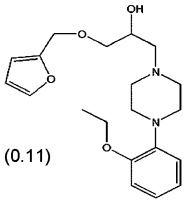
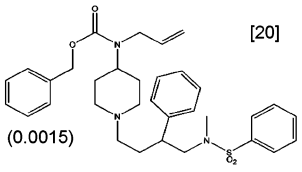
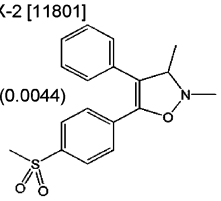
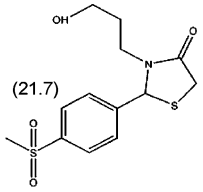
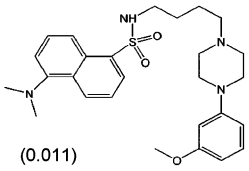
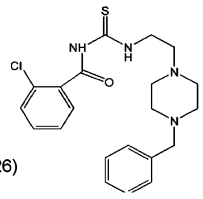
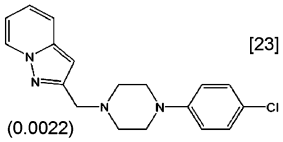
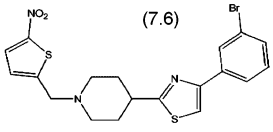
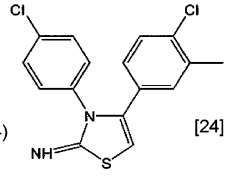
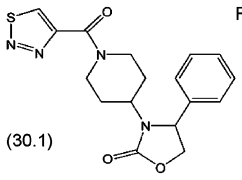
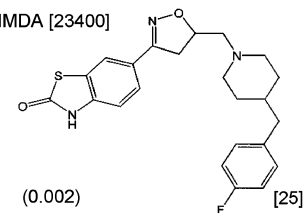
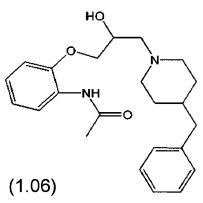
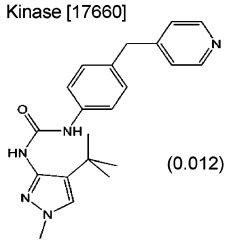
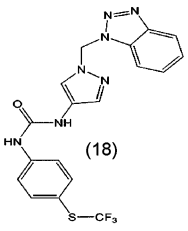
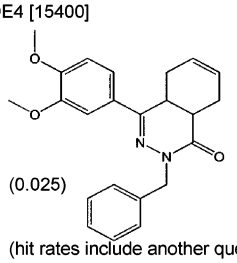
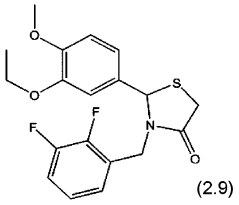
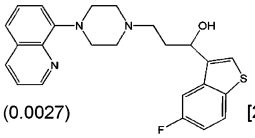
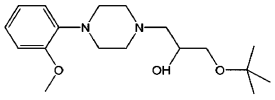
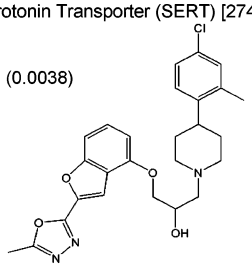
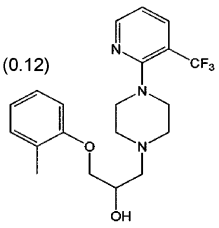
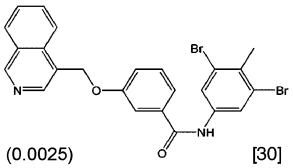
Assay [ID #] / Query (IC50, uM) [ref.]	TopSim hit (IC50, uM)	How sel'ct'd	n	Overall Hit Rates: Inhibition at 10 uM >= 20%	>= 50%	>= 90%
Adenosine A-1 [20051]						
 (0.0066) [18]	 (20.2)	TopSim	11	3 27%	0 0%	0 0%
		Random	14	2 14%	1 7%	0 0%
Adrenergic alpha1A [20310]						
 (0.0002) [19]	 (0.11)	TopSim	12	4 [c] 33%	4 [a] 33%	3 [a] 25%
		Random	17	2 12%	0 0%	0 0%
CCR5 [21770]						
 [20]	No confirmed hits (< 100 uM)	TopSim	5	0 0%	0 0%	0 0%
(0.0015)		Random	4	0 0%	0 0%	0 0%
COX-2 [11801]						
 (0.0044) [21]	 (21.7)	TopSim	14	0 0%	0 0%	0 0%
		Random	4	0 0%	0 0%	0 0%
Dopamine D2L [21960]						
 (0.011) [22]	 (26)	TopSim	26	3 12%	0 0%	0 0%
		Random	14	3 21%	3 21%	1 7%
Dopamine D4.4 [22000]						
 [23] (0.0022)	 (7.6)	TopSim	50	6 12%	2 4%	0 0%
		Random	14	5 36%	0 0%	0 0%
Dopamine transport [22032]						
 (0.024) [24]	 (30.1)	TopSim	4	1 25%	0 0%	0 0%
		Random	4	1 25%	0 0%	0 0%
NMDA [23400]						
 (0.002) [25]	 (1.06)	TopSim	6	6 [a] 100%	6 [a] 100%	4 [a] 6
		Random	4	0 0%	0 0%	0

Table 1 (Continued)

Assay [ID #] / Query (IC50, uM) [ref.]	TopSim hit (IC50, uM)	How sel'ct'd	n	Overall Hit Rates: Inhibition at 10 uM >= 20%	>= 50%	>= 90%
p38a Kinase [17660]						
 (0.012) [26]	 (18)	TopSim	7	4 [b] 57%	1 14%	0 0%
		Random	12	1 8%	1 8%	0 0%
PDE4 [15400]						
 (0.025) [27]	 (2.9)	TopSim	84	29 [a] 35%	10 [b] 12%	0 0%
		Random	23	2 9%	0 0%	0 0%
(hit rates include another query structure not shown)						
Serotonin 5HT1a [27111]						
 (0.0027) [28]	 (0.11)	TopSim	4	4 [a] 100%	4 [a] 100%	4 [a] 100%
		Random	4	0 0%	0 0%	0 0%
Serotonin Transporter (SERT) [27402]						
 (0.0038) [29]	 (0.12)	TopSim	5	5 [b] 100%	5 [a] 100%	4 [a] 80%
		Random	7	4 57%	3 43%	1 14%
Vascular Endothelial Growth Factor (VEGF) [28690]						
 (0.0025) [30]	No confirmed hits (< 100 uM)	TopSim	80	4 5%	0 0%	0 0%
		Random	14	2 14%	0 0%	0 0%
	AVERAGE TopSim			39%	28%	21%
	AVERAGE Random			15%	6%	2%
	AVERAGE Enhancement			24%	22%	19%

[a] TopSim hit rate enhancement significantly greater than Random at 95% level (Chi-square test)
[b] TopSim hit rate enhancement significantly greater than Random at 90% level (Chi-square test)
[c] TopSim hit rate enhancement significantly greater than Random at 75% level (Chi-square test)

compounds including the example hit, and the lower line for the negative control experiments, mostly compounds randomly chosen from the same (LeadQuest) screening collection as detailed above. Each of these lines contains, from left to right, the number of compounds tested (n) and the numbers and percentages of those tested compounds that were active, for various definitions of activity, responses $\geq$20%, $\geq$50%, and $\geq$90%, all at the administered concentration of 10 μM. Footnotes next to the number of active compounds mark instances where the frequency of activity for topomer similarity selected compounds was significantly greater than that for randomly selected compounds, according to a χ^2 test.[31]

At the bottom of Table 1 are given the averaged hit rates for each combination of the three activity criteria with the two methods of compound selection. Here are some important additional generalizations from Table 1.

(1) The apparent improvements of 24%, 22%, and 19% in hit rates for topomer similarity selections over random selections are statistically significant, well over the 95% confidence level for four of six comparisons

Table 2. Comparison of Topomers and Tanimoto 2D Fingerprints, in Ability to Concentrate Together Compounds Acting by the Same Mechanism According to the Word Drug Index, for Smaller Structural Differences

	Tanimoto neighbors					topomer neighbors						
			≥0.85 similarity			"neighbors" same as 0.85 Tanimoto						
biol mechanism[a]	*n*	nbrs	actv	hit rate, %	enhce	*n*	T. radius[b]	nbrs	actv	hit rate, %	enhce	actv rat.[c]
antihistamines	419	220	137	62	13	399	90	230	179	78	16	1.25
antiserotonin	218	127	49	39	16	198	100	146	77	53	21	1.37
dopamine antagonist	349	282	124	44	11	330	120	327	191	58	15	1.33
dopaminergic	200	175	92	53	23	168	130	190	96	51	22	0.96
endothelin-A	74	41	24	59	70	73	130	41	37	90	108	1.54
HIV	86	40	24	60	62	79	130	42	34	81	83	1.35
leukotriene antagonist	227	73	41	56	22	218	100	82	65	79	31	1.41
muscarinic	26	38	6	16	54	25	150	40	11	28	93	1.74
opioid	59	56	20	36	53	52	140	61	28	46	69	1.29
parasympatholytic	252	220	96	44	15	246	90	229	107	47	16	1.07
parasympathomimetic	90	60	17	28	28	82	140	65	41	63	62	2.23
prostaglandin	610	330	231	70	10	600	110	360	302	84	12	1.20
protease inhibitors	62	20	15	75	107	55	90	20	20	100	142	1.33
protein kinase inhibitors	100	61	26	43	38	81	120	66	27	41	36	0.96
purine antagonist	434	292	197	67	14	423	120	305	229	75	15	1.11
serotoninergic	120	78	20	26	19	119	120	91	32	35	26	1.37
sympatholytic alpha	285	222	85	38	12	269	110	223	111	50	15	1.30
sympatholytic	725	613	274	45	5	699	110	626	365	58	7	1.30
sympathomimetic	414	297	166	56	12	386	120	321	228	71	15	1.27
topoisomerase-I	68	51	20	39	51	60	160	54	28	52	67	1.32
average	241	165		48	32	228	119	176		62	44	1.34
std deviation				16	26		20			20	38	0.28

[a] Correspondences to the World Drug Index codes: antihistamines, "ANTIHISTAMINE"; antiserotonin, ANTISEROTONIN; dopamine antagonist, DOPAMINE-ANTAGONIST; dopaminergic, "DOPAMINERGIC"; endothelin-A, ENDOTHELIN-A and ET-A; HIV, HIV; leukotriene antagonist, LEUKOTRIENE-ANTAGONIST; muscarinic, MUSCARINIC; opioid, OPIOID; parasympatholytic, PARASYMPATHOLYTIC; parasympathomimetic, PARASYMPATHOMIMETIC; prostaglandin, PROSTAGLANDIN and THROMBOXANE; protease inhibitors, PROTEASE-INHIBITOR; protein kinase inhibitors, PROTEIN-KINASE*INHIBIT; purine antagonist, PURINE-ANTAGONIST and CALCIUM-ANTAG; serotoninergic, SEROTONINERGI; sympatholytic, SYMPATHOLYTIC; sympatholytic alpha, SYMPATHOLYTIC-ALPHA; sympathomimetic, SYMPATHOMIMETIC; topoisomerase-I, TOPOISOMERASE-I. [b] The topomer radius (in steps of 10) that enclosed a number of structures fewer than the structures having a Tanimoto similarity ≥ 0.85. See text. [c] The ratio of superiority, in either hit rate or enhancement, of topomer similarity over Tanimoto similarity.

Table 3. Comparison of Topomers and Tanimoto 2D Fingerprints, in Ability to Concentrate Together Compounds Acting by the Same Mechanism According to the Word Drug Index, for Larger Structural Differences

	Tanimoto 2D fingerprints							topomers								
					hit rates, %							hit rates, %				
	coeff > 0.65		coeff > 0.55			0.55–		radius < 200		radius < 240			200–		core/ shell	shell/ tail
mechanism	nbrs	actv	nbrs	actv	>0.65	0.65	<0.55	nbrs	actv	nbrs	actv	<200	240	>240	enrich	enrich
antihistamines	1320	247	3445	323	18.7	3.6	1.8	1216	330	3058	368	27.1	2.1	0.5	2.51	1.91
antiserotonin	1120	109	2931	156	9.7	2.6	1.1	1033	149	2869	172	14.4	1.3	0.4	3.07	1.16
dopamine antagonist	1041	242	2543	296	23.2	3.6	0.8	1303	282	3760	323	21.6	1.7	0.1	2.01	2.83
dopaminergic	734	147	1719	168	20.0	2.1	0.4	703	140	1892	165	19.9	2.1	0.0	1.01	10.26
endothelin-A	235	50	1038	55	21.3	0.6	0.2	109	51	698	56	46.8	0.8	0.2	1.61	1.59
HIV	179	46	797	54	25.7	1.3	0.4	91	54	1091	65	59.3	1.1	0.2	2.72	1.87
leukotriene antagonist	553	95	2066	132	17.2	2.4	1.4	846	152	3512	194	18.0	1.6	0.5	1.62	2.01
muscarinic	166	9	537	10	5.4	0.3	0.2	153	11	686	16	7.2	0.9	0.1	0.38	6.08
opioid	414	33	1510	43	8.0	0.9	0.2	246	36	887	46	14.6	1.6	0.1	1.07	4.95
parasympatholytic	915	170	2457	197	18.6	1.8	0.9	1056	205	3517	232	19.4	1.1	0.3	1.67	2.05
parasympathomimetic	289	33	945	39	11.4	0.9	0.6	289	55	1772	65	19.0	0.7	0.2	2.26	1.98
prostaglandin	1467	399	3411	472	27.2	3.8	2.5	1367	483	4562	561	35.3	2.4	0.9	2.00	1.81
protease inhibitors	83	32	623	38	38.6	1.1	0.3	89	40	997	44	44.9	0.4	0.1	2.94	0.83
protein kinase inhibitors	449	45	1424	51	10.0	0.6	0.7	799	46	2374	69	5.8	1.5	0.2	0.24	8.45
purine antagonist	1307	294	3211	340	22.5	2.4	1.7	1219	329	3964	382	27.0	1.9	0.8	1.50	1.59
serotoninergic	531	59	1669	86	11.1	2.4	0.5	831	80	1861	104	9.6	2.3	0.2	0.88	2.17
sympatholytic-alpha	1198	168	2616	218	14.0	3.5	1.1	1063	207	3371	247	19.5	1.7	0.4	2.83	1.32
sympatholytic	2236	505	4289	607	22.6	5.0	2.6	1913	588	4263	658	30.7	3.0	0.9	2.27	1.74
sympathomimetic	1066	282	2310	320	26.5	3.1	1.4	1052	343	2733	365	32.6	1.3	0.3	2.88	1.79
topoisomerase-I	209	31	648	37	14.8	1.4	0.4	105	34	614	40	32.4	1.2	0.2	2.53	1.34
average	776		2009		18.3	2.2	1.0	774		2424		25.3	1.5	0.3	1.90	2.89
std deviation					8.0	1.3	0.7					13.8	0.6	0.3	0.86	2.55

performed and just below the 95% confidence level for the other two comparisons. For purposes of significance testing, all the raw hit rates in Table 1 were continuity corrected.[32] Two types of significance tests were carried out:[31] paired samples (≥20%, $t = 2.16$; ≥50%, $t = 1.87$; ≥90%, $t = 1.83$) and differences of means (≥20%, $t = 1.95$; ≥50%, $t = 1.60$; ≥90%, $t = 1.70$).

(2) By the most widely accepted measure of structural diversity, Tanimoto coefficients of "2D fingerprints", the queries and example hit structures shown are not at all similar. The average Tanimoto coefficient for these paired structures is 0.36, with little variance (the maximum Tanimoto coefficient is 0.45). This level of Tanimoto similarity is little greater than the expected

mean Tanimoto similarity of two randomly chosen structures (0.333).[33]

(3) The potencies of the topomer similarity selected "lead hop" structures are much lower than those of the corresponding query structures, by factors averaging close to 1000× and ranging from 30× to 5000×.

Table 2 directly compares topomers and 2D fingerprints in their retrospective abilities to concentrate biologically similar structures together. Its baseline is the performance of a "Tanimoto 2D fingerprint coefficient of at least 0.85", probably the most widely used diversity criterion for avoiding redundancy when selecting structures for random screening.[4] Table 2 may be understood by working through an individual line of data in detail. Its first (antihistamines) row begins with the 2D fingerprint (left-hand) block, indicating that there are 419 structures (among 8832) having a recorded mechanism of ANTHISTAMINE. These 419 structures have 220 neighbors altogether, again from among 8832, when the criterion for a neighbor is "Tanimoto 2D fingerprint coefficient of at least 0.85". Of these 220 neighbor structures, 137 are included also among those 419 structures with the ANTIHISTAMINE mechanism, hence "actv", for a "hit rate" of 137/220 or 62%.

Much virtual screening research presents hit rate improvements in a different way, as an "enhancement factor" over the hit rate that would be expected for a random selection from the same set of compounds. In this case, for the ANTIHISTAMINE mechanism the random hit rate would be 419/8832 or 4.7%, so "enhce" becomes 62%/4.7% or 13. It may also be noted that there must be a maximum possible "enhancement factor" for any activity, in this case ANTIHISTAMINE. For example, if every one of the 3% nearest of structures to any ANTIHISTAMINE was also itself an ANTIHISTAMINE, the enhancement factor would be 100%/4.7% or 21. It is also likely that some of the reportedly non-ANTIHISTAMINE near neighbors of an ANTIHISTAMINE structure actually have antihistaminic properties that have never been measured or recorded. Thus the enhancement ratios reported in Table 2 may be considered as lower bounds on values that also have hard upper bounds.

The remaining right-hand block of Table 2 compares topomer results to this baseline "≥0.85 Tanimoto 2D fingerprint" performance. Within each row a topomer similarity radius was chosen to select roughly the same number of maximally similar compounds that the Tanimoto 2D fingerprint "similarity ≥ 0.85" criterion selects. Continuing along the same row of data in detail, the first column records that 20 of the 419 reported antihistamines could not be processed topomerically (usually because of the lack of any acyclic bond whose cleavage would yield two fragments each having more than three atoms), leaving 399 antihistamines. Within a topomeric radius of 90, there were 230 structures found among the 8832 that were at least that similar to one of those 399 antihistamines. (In other words, a topomer radius of 90 enclosed the fewest neighbors (230) that are still more than the 220 count found for the Tanimoto radius of 0.85.) Among these 230 topomerically similar structures, 179 were themselves recorded as antihistamines, for a hit rate of 78% and an enhancement factor of 16. The hit rate of 78% is higher than the hit rate of 62% for the compounds selected by the Tanimoto ≥ 0.85 threshold, by a ratio of 1.25.

The most important values in Table 2 are the averages at the bottom. Among these 20 structurally most similar groups, comprising on average 3% of all structures, the superiority of topomer similarity over Tanimoto 2D fingerprint coefficient in concentrating together biologically similar structures is substantial, on average (62%−48%)/48% or about 34% ($p \gg 0.999$, $t = 5.43$).

Table 3 again compares the concentrating tendencies of topomer similarity and Tanimoto 2D fingerprint coefficients, this time, however, comparing structures whose similarities are much less likely to be obvious on inspection. The baseline for Table 3 is a topomeric similarity "shell" bounded by values between 200 and 240, a degree of similarity that has been productive in proprietary "lead hopping" collaborations using topomers, as well as in the "lead hopping" prospective validation study above. By inspection it was found that the Tanimoto 2D fingerprint coefficient "shell" that contains similar numbers of neighbors, both inside of the shell and "within" the shell extent itself, is bounded by values between 0.65 and 0.55 (note the similarities in "average" values for "nbrs" counts, at the bottom of Table 3). The counts of neighbors and actives shown for the Tanimoto coefficient thresholds of 0.65 and of 0.55 and the topomer similarities of 200 and of 240 were established as described for Table 2. From these counts, the hit rates inside either of the threshold values immediately follow, and the hit rates within the shells bounded by these thresholds and beyond the outsides of those shells then require only suitable subtractions of the counts (including also the columns *n* from Table 2).

Although an analysis similar to that summarized in Table 2 using these larger radii yields similar topomer superiorities, there is a more interesting way to consider these data. The averaged hit rate of 2.2% for the Tanimoto 2D fingerprint coefficient range 0.55−0.65 is larger than the averaged hit rate of 1.5% for the topomer similarity range of 200−240. At first glance this comparison might suggest superiority for the Tanimoto coefficient. However, the better a structural descriptor is in closely concentrating biologically similar structures, the fewer the possible occurrences of biologically similar structures that remain at the greater structural dissimilarities. Accordingly Table 3 also shows that the hit rates inside and outside of these shells indicate a stronger concentrating influence by topomer similarity, with the averaged hit rate inside the shells favoring topomers by 25.3% vs 18.2% and outside disfavoring topomers by 0.3% vs 1.0%.

This concentrating tendency can also be expressed as the ratio of the inside the shell ("core") hit rate to the hit rate within the shell itself. Similarly, the ratio of the hit rate altogether outside the shell to the hit rate within the shell indicates how strongly a concentrating tendency persists for even larger structural dissimilarities. A higher value of either ratio indicates a stronger concentrating or enrichment performance. So division of a ratio for topomer concentration by the corresponding ratio for fingerprint concentration provides a direct comparison, where larger values favor topomer concen-

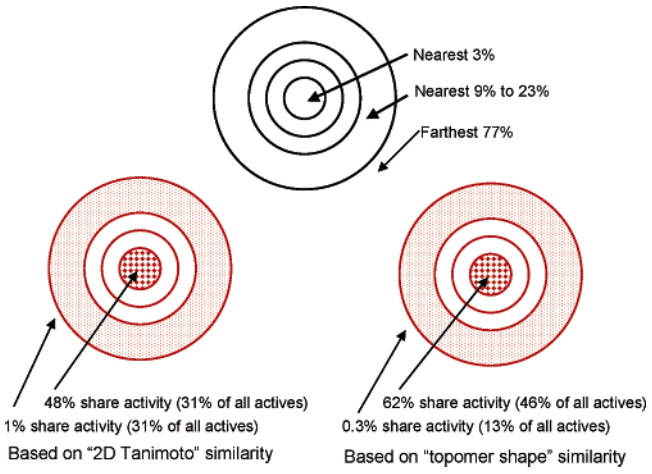


Figure 1. Summary comparison of 2D fingerprint Tanimoto coefficients to topomer similarity, in their abilities to "concentrate together" structures reportedly having the same biological activity. Shown for each of the descriptors and their respective ratings of "most similar" (3%) and "most dissimilar" (77%) structures are the averaged percentages of the rated structures that have the same activity and the averaged percentages of all active structures that have that rating. The underlying data are taken from Tables 2 and 3.

tration. The two rightmost columns of Table 3 show these "ratio of ratios" values, the "core/shell enrich" column comparing "inside shell" to "within shell" hit rates and the "shell/tail enrich" column comparing "within shell" to "outside shell". From the averages of these two columns it is apparent that topomers provide almost double the inside/within shell enrichment of the Tanimoto 2D fingerprint coefficient ($p \gg 0.995$, $t = 4.70$), and almost three times the within/outside shell enrichment ($p \gg 0.995$, $t = 3.31$).

Figure 1 is a cartoon that more simply though a bit simplistically summarizes the results shown in Tables 2 and 3. Imagine a "typical" active structure taken from the WDA. All the other structures in the WDA are then ranked by similarity to that structure, using either 2D fingerprint Tanimoto coefficients or topomers as the similarity measurement. Every one of the 8832 compounds lies somewhere on both of these targets, but the preferable similarity measurement would place nearer to the "bulls-eye" those compounds having the same activity as that typical compound. Tables 2 and 3 show that on average, when structural difference is defined by topomers instead of 2D fingerprints, the nearest 3% of other structures contains the higher frequency of the structures with the same activity (62% vs 48%) and then of course also the larger proportion of all the active compounds (46% vs 31%). Conversely, with topomers the most dissimilar 77% of structures includes the much lower frequency of activity and the much lower proportion of all the known inactive compounds.

However, please note that another impression this cartoon may create, that there is only one structural "center" for a particular activity, is false. In the experiments described here, to be located among the center 3%, a structure needs to have only one very close neighbor having the same activity, but that pair of structures may also be far distant from other structures having the same activity. Conversely a structure located among the most distant 77% may (and probably does) in fact have numerous structural neighbors with the same activity—but (for any of a number of valid reasons) these neighbors are not recorded in the WDA. Thus little confidence should be placed in the absolute values of the performances shown in Figure 1, being as much consequences of how the WDA dataset happens to have accumulated as of the similarity measurements. Nevertheless the superiority of topomers compared to 2D fingerprints is clearly established by Figure 1 and its underlying data, since presumably any biases from dataset assembly affect both measures equally.

Yet all these retrospective comparisons of average descriptor performance bear only indirectly and inferentially on the main objective of similarity selection. Is either of these descriptors more effective in "lead hopping"—in using nonobvious structural features to recognize possible biological similarities between pairs of molecules? Table 4 provides a few structural comparisons as a basis for this essentially subjective judgment. Topomer neighbor lists had been generated for each of the 8832 structural "queries". From each activity class, one or more of the longest neighbor lists were scanned to identify examples of structures that (1) belonged to the same activity class as that neighbor list's query and (2) were "distant neighbors", according to the Table 3 criteria for either 2D fingerprint Tanimotos or topomers, i.e., within either the 0.55 to 0.65 Tanimoto similarity shell or the 200 to 240 topomer similarity shell. Although it is not claimed that this manual selection of these example structures was unbiased, there were often so few possibilities that perhaps half the structures shown in Table 4 represent around half of the candidate structures for that query. The results of these selections make up the rows of structures in Table 4, the left-most structures being the query that generated the neighbor list, the next structure a distant neighbor according to 2D fingerprints, and the two right-hand structures distant neighbors according to topomers. All structures have been drawn to show the topomer shape correspondences, and, of course, the four structures within any row of Table 4 all belong to the same activity class or "mechanism". Under each structure is shown its World Drug Index name identifier and, for each of the three neighbors, its topomer similarity to that query structure and its Tanimoto 2D fingerprint similarity to that query structure.

Within Table 4, it is our perhaps subjective opinion that the apparent structural dissimilarities of a query from its right-hand pair of topomer distant neighbors are much greater than that from its Tanimoto distant neighbor. More objectively, the averages of the paired similarity values (bottom of Table 4) reiterate that the concentration of similarly active structures is much higher with topomer similarity than with Tanimoto 2D fingerprints. As illustrated in Figure 1, the structures in Table 4 are roughly equally distant from their query structures: they belong to similar "percentiles", in terms of the frequency distributions of their defining descriptor values. However, the Tanimoto similar structures shown are by topomer similarity placed much closer to their query structures (the average topomer dissimilarity value of 144 from the query for the Tanimoto similar column of structures being much less than 200). Conversely, the topomer similar structures would be placed

Table 4. Examples of Structures Having the Same Biological Mechanism as a Query Structure, with Distant Structural Similarity to That Query According to Either 2D Fingerprint (Tanimoto) or Topomer Shape (These Two Individual Similarity Values Are Shown under Each Neighbor Structure)

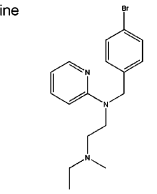
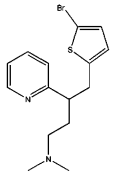
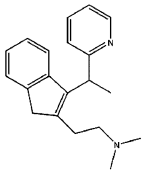
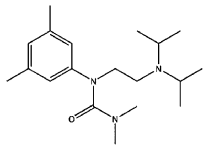
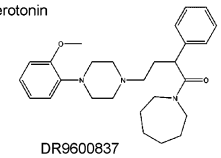
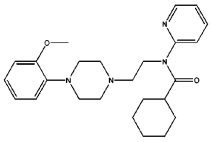
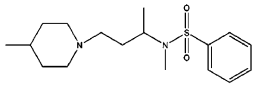
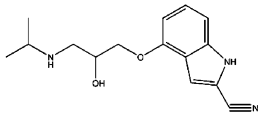
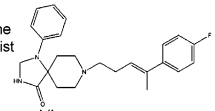
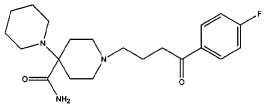
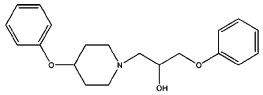
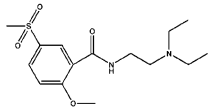
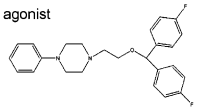
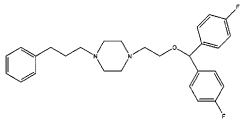
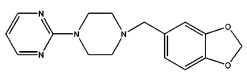
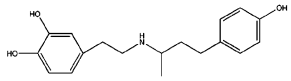
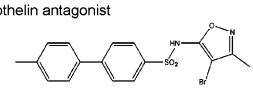
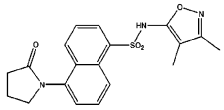
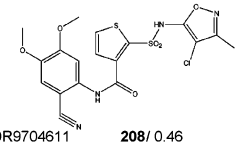
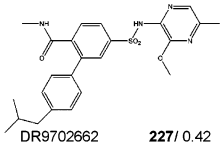
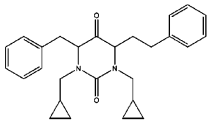
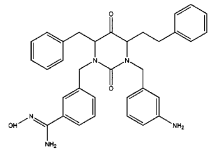
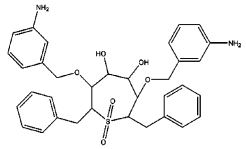
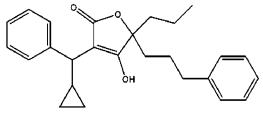
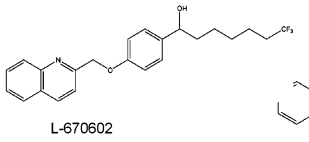
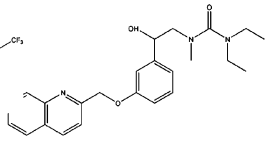
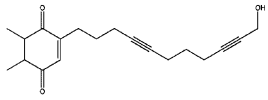
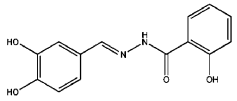
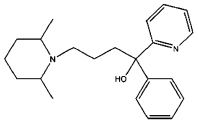
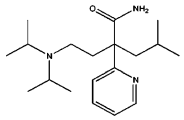
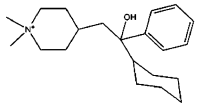
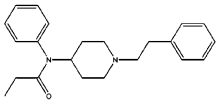
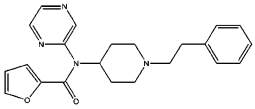
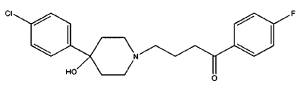
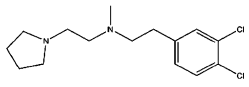
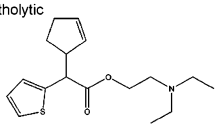
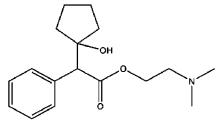
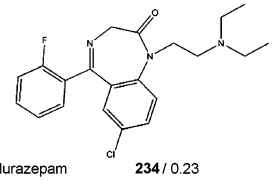
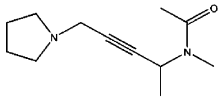
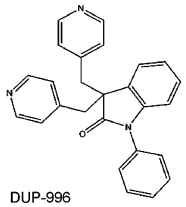
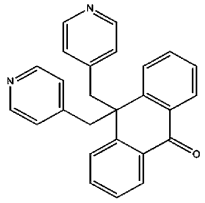
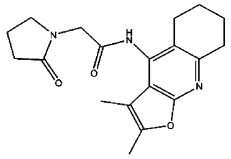
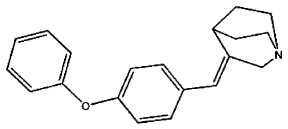
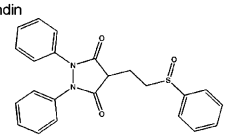
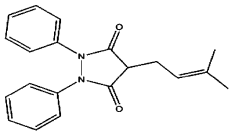
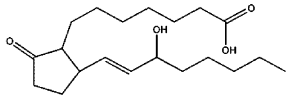
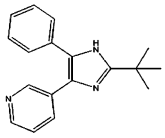
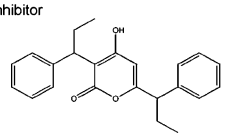
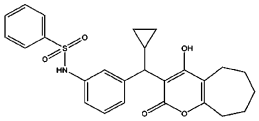
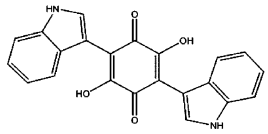
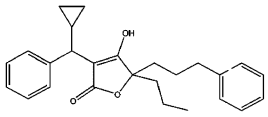
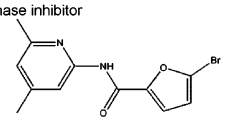
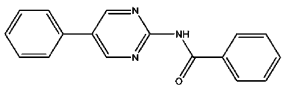
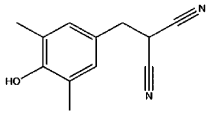
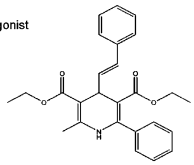
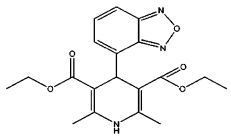
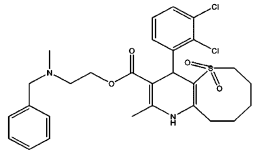
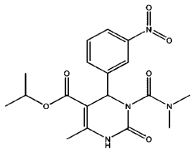
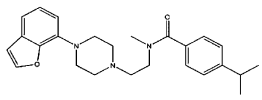
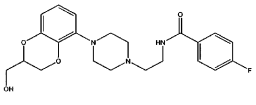
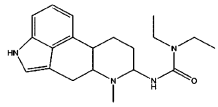
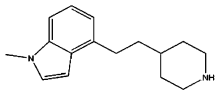
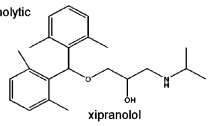
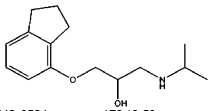
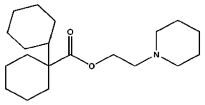
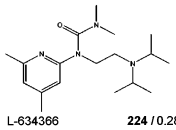
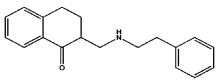
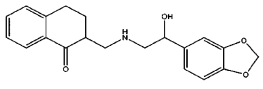
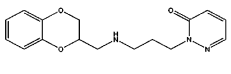
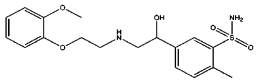
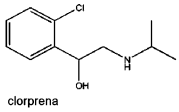
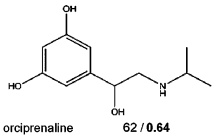
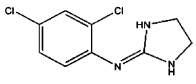
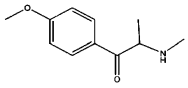
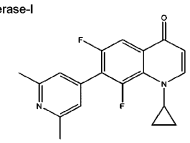
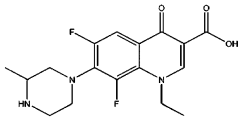
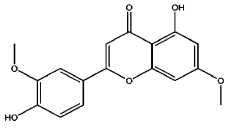
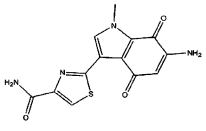
Mechanism	Query Structure	2D fingerprint (Tanimoto) Query neighbor	Topomer shape similarity Query neighbors	
antihistamine	 adeptolon	 bromopyramine 169 /**0.60**	 dimetindene **233**/ 0.28	 L634366 **216**/ 0.53
antiserotonin	 DR9600837	 WAY 100635 83 /**0.63**	 DR9802293 **209**/ 0.29	 cyanopindolol **235**/ 0.30
dopamine antagonist	 spirilene	 pipamperone 161 /**0.60**	 DR9703770 **203**/ 0.24	 tiapride **231**/ 0.26
dopamine agonist	 GBR13045	 BS 8824 150 /**0.69**	 piribedil **201**/ 0.28	 LY-206243 **239**/ 0.30
endothelin antagonist	 DR9505097	 DR9603054 161 /**0.58**	 DR9704611 **208**/ 0.46	 DR9702662 **227**/ 0.42
HIV	 DR9504899	 DR9606532 155 /**0.65**	 DR9607842 **217**/ 0.31	 DR9502809 **211**/ 0.31
leukotriene antagonist	 L-670602			
muscarinic	 pirmenol	no muscarinic structures with Tanimoto similarity > 0.52 to pirmenol	 propisomi **226** / 0.52	 hexcycliu **211** / 0.52
opioid	 fentanyl	 A-3508 134 /**0.45**	 haloperidol **236** / 0.38	 BD-1008 **221** / 0.44
parasympatholytic	 W-290	 cypentola 148 /**0.56**	 flurazepam **234** / 0.23	 BM-5 **239** / 0.18

Table 4 (Continued)

Mechanism	Query Structure	2D fingerprint (Tanimoto) Query neighbor	Topomer shape similarity Query neighbors	
parasympathomimetic	 DUP-996	 DR960418 166 / **0.42**	 MKC-231 **219** / 0.40	 KST-5452 **230** / 0.28
prostaglandin	 sulfinpyrazone	 feprazone 142 / **0.67**	 PGB **239** / 0.14	 GP-53633 **224** / 0.17
protease inhibitor	 U-96988	 DR9701077 237 / **0.64**	 DR9704658 **232** / 0.37	 DR9502809 **230** / 0.69
protein kinase inhibitor	 DR9511336	no protein kinase inhibitors with Tanimoto similarity >.37 to DR9511336 (except trivial -Br => -H)	 DR9604451 **235** / 0.37	 AG-17 **235** / 0.15
purine antagonist	 DR9605765	 darodipine 146 / **0.62**	 RWJ-22726 **217** / 0.51	 SQ-31298 **223** / 0.47
serotoninergic	 befiperide	 flesinoxan 137 / **0.48**	 lisuride **227** / 0.30	 LM-5005 **236** / 0.29
sympatholytic	 xipranolol	 USVC-6524 172 / **0.58**	 dihexyverine **211** / 0.30	 L-634366 **224** / 0.28
sympatholytic-alpha	 BE-1726	 BE-2637 100 / **0.65**	 GYKI12743 **213** / 0.20	 amosulalol **232** / 0.28
sympathomimetic	 clorprena	 orciprenaline 62 / **0.64**	 NC-8 **236** / 0.28	 meophedr **206** / 0.39
topoisomerase-I	 DR9505975	 lomefloxacin 174 / **0.59**	 DR9606128 **203** / 0.28	 BE-10988 **217** / 0.34
Average		145 / **0.59**	**223** / 0.33	

in more distant percentiles using the 2D fingerprint Tanimoto coefficient (indeed, the average Tanimoto 2D fingerprint coefficient of 0.33, between the query structure and the two topomer similar structures, coincides with the Tanimoto 2D fingerprint coefficient expected between random structures[33]).

It may be asked whether the active neighbor lists for topomer similarity are supersets of the active neighbor lists for Tanimoto 2D fingerprints, rather than disjoint sets. Although this question has not been investigated directly, the structures and similarity values shown in Table 4 strongly suggest that the Tanimoto lists are subsets of the topomer lists.

Discussion

The ~20% improvement in hit rates achieved with topomer similarity selection (bottom of Table 1) merits comparison with the best results reported using other virtual screening methodologies, especially when one also considers:

(1) The concurrent goal of "lead hopping". Indeed, on the basis of "lead hops" per query rather than actives per tested compound, our success rate is 11/13 or 85%! (Every cell but two in the "TopSim Hit" column of Table 1 contains a structure significantly different from the query structure that yet possesses the same biological activity.)

(2) The minimal starting information, simply the structure of a single active ligand.

(3) The number and variety of the biological systems and the chemical structures. When investigating unfamiliar biological systems, it would seem that the consistency or robustness that a compound selection methodology has already shown across many biological systems should be at least as important a consideration as its best individual performance.

Before comparing selection methodologies, however, two potential caveats with Table 1 must be addressed. First, the hit rates for randomly selected LeadQuest compounds, averaging 15%, 6%, and 2% for approximate IC_{50} thresholds of 40 μM, 10 μM, and 1 μM, respectively,[34] are much higher than the hit rate generally cited in HTS, 0.1% or less. This HTS conventional wisdom is supported in the published literature by individual reports of HTS hit rate of 0.02%[35] and $\leq$0.2%,[36] for IC_{50} thresholds around 100 μM. Thus the hit rates reported in Table 1 for random LeadQuest compounds are about 2 orders of magnitude greater than either of these two specific HTS hit rates. Informal discussions of this large difference with experienced screening scientists from several different organizations[37] have generated at least six explanatory hypotheses. (1) There has been experimental error. However, recall that six of our negative control activities (over 25% of all randomly chosen hits) were confirmed, with and without additional test sample purification, by a different testing organization. (2) The mix of biological assays differs, with enzymatic assays being common in HTS and less demanding receptor assays more common within Table 1. (3) The design characteristics of a high quality random screening library such as LeadQuest may provide more hits than do typical corporate collections. (4) Experimental variables may be tuned by independent screening organizations to favor more hits, but oppositely in HTS. (5) The original query selections, resulting from WDA reports of high activity and commercial availability of the assay, may bias the assay mix toward those yielding more hits. (6) LeadQuest might contain substantially more frequent and powerful aggregate formers than typical corporate collections;[38] however, see the footnote. Regardless of the true cause(s), we eventually decided simply to treat as an empirical fact this unexpectedly high frequency of activity among structures randomly chosen from LeadQuest and continue. Of course, these high random hit rates may instead be regarded only as a suspicious anomaly. However, the biasing effect of our treatment is conservative, depressing the lead hopping effectiveness that we are then able to report for topomers.

A second possible caveat is that the potencies of the TopSim hit structures in Table 1 average roughly 3 orders of magnitude lower than the Query potencies. However the Query structures have typically been optimized and the TopSim hits have not. The average 5 μM IC_{50} potencies for the eleven TopSim hits seem typical of starting points when optimizing any new lead series. And of course there may not in fact exist any nanomolar potent agents among the 80000 selection candidates.

So how does the ~20% hit rate enhancement that topomer selection provided in these "lead hops" compare with other virtual screening performances? Much the most popular virtual screening methodology is docking into receptor crystal structures, and several impressive docking results have recently appeared. Of 103 structures selected in part by docking to Checkpoint Kinase-1, 36 (35%) structures representing four novel and tractable chemical classes inhibited the enzyme with IC_{50} values between 68 and 0.11 μM.[39] Similarly, 127 of 365 (35%) docked structures inhibited protein tyrosine phosphatase-1B with IC_{50} values between 100 and 1.6 μM.[36] Even a docking hit rate of 100% has been reported, with 9 K_i's between 249 and 0.25 μM for binding to tRNA-guanine transglycosylase by 9 structures of differing chemotype selected from over 800000 candidates.[40] Docking has also produced extremely potent inhibitors of carbonic anhydrase,[41,42] although by structural changes that seem too modest to be called "lead hops". Coupling of docking to a de novo structure generator led to synthesis and testing of 18 novel HIV reverse transcriptase inhibitors, 10 of which (56%) had IC_{50}'s between 100 and 0.1 μM.[43] However, docking calculations are by now so widely practiced that it seems fair to assume that these publications represent the most favorable among a wide distribution of results, arguably comparable to the NMDA, p38a, serotonin 5HT1a, and SERT results within Table 1. Informal estimates by industry aficionados of average hit rates from docking calculations, 10% to 15%, resemble our average of ~20% hit rate enhancement for topomer similarity selection. Turning to factors other than prospective hit rates, docking and topomers also seem equivalently strong in their "lead hopping" capabilities. Comparing finally the breadth of their established biological applicability, on the one hand, the large number of docking research efforts has surely generated a far larger number of successes than for topomers. On the other hand, docking predictions are not even pos-

sible for most of the assays in Table 1, because no X-ray structure for the receptor is available.[44]

Selection of compounds by ligand "2D similarity" measures such as Tanimoto 2D fingerprints provides a well-established hit rate of at least 30%,[4] but with few if any occurrences of "lead hopping". There are scattered reports of moderately successful hit rate enhancements with various flavors of "3D (pharmacophoric) searching".[12,45] The only ligand-based selection methodologies reported to yield "lead hopping" predictive performances comparable to the best of those in Table 1 seem to be two emerging approaches. Consensus predictions from several QSAR models led to synthesis and testing of nine arguably novel "functionalized amino acids", seven (78%) showing moderate anticonvulsant activities.[46] A highly sophisticated shape similarity methodology has achieved retrospective classification accuracies of 60% among diverse serotonin, histamine, muscarinic, and $GABA_A$ ligands.[47] These predictive performances are similar to the best within Table 1. However, their capabilities for prospective "lead hopping" and their breadth of applicability have yet to be reported.

To summarize, prospective predictions of biological properties based on topomer similarity seem at least comparable in success rates and diversity of applicability to the best prospective predictions reported using other methodologies. This outcome mirrors the retrospective validation of topomer alignments in CoMFA, which for fifteen data sets yielded potency predictions whose average accuracy equaled that from the originally published models.[10]

The retrospective comparisons of topomer similarity and Tanimoto coefficients of 2D fingerprints in their propensity to concentrate together biologically similar structures, summarized in Tables 2 through 4, provide independent and indeed much stronger evidence for a superior biological predictiveness of topomer similarity. "Tanimoto 2D fingerprint similarity" is the most thoroughly validated[1−4] predictor of biological similarity and the most widely used, particularly when selecting structures to follow up initial hits. The 30+% hit rate for structures having a "≥0.85 Tanimoto" to another active structure,[4] though retrospective, is at least as high as and more widely confirmed than any other prospective hit rate. Yet the comparative results in Tables 2 and 3 indicate that, over a very wide range and diversity of compounds (8832) and "activity classes" (20), topomer similarity decisively outperforms "2D fingerprint Tanimotos".[48]

(1) When the 3% of structures "most similar" to structures having a particular activity are selected, either by topomers or 2D fingerprint Tanimotos, from a large pool of (biologically active) structures (Table 2), the topomer most similar 3% contains on average a third more other structures having the same activity (62%) than does the 2D fingerprint Tanimoto most similar 3% (48%). Thus the occurrence of "false positives" is substantially lower for topomers than for 2D fingerprint Tanimotos.

(2) The one-third topomer improvement in selectivity is quite consistent throughout the various activity classes. In only two of the twenty individual enrichment comparisons do Tanimoto 2D fingerprints exceed topomers ("dopaminergic" and "protein kinase"), by a meager 2% in both cases.

(3) Almost exactly the same comparison obtains when the most similar ~10%, rather than ~3%, structures are selected by topomers and by Tanimoto 2D fingerprints (Table 3), the largest difference being that "serotoninergic" replaces "dopaminergic" as one of the two Tanimoto-superior activity classes.

(4) Conversely, if one considers the 75% of compounds "least similar" to structures having a particular activity, within this large pool of biologically active structures (Table 3), either by topomers or Tanimoto 2D fingerprints, the Tanimoto-not-similar set contains three times the number of actives (1.0%) as does the topomer-not-similar set (0.3%). Thus the "false negative" rate is much lower for topomers than for Tanimoto 2D fingerprints.

(5) This 300% improvement by topomers in "separating the chaff from the wheat" is even more consistent than the 33% enrichment. All twenty of the individual activity classes leave fewer "structural outliers" in the topomer-not-selected 75% of compounds than in the Tanimoto-not-selected 75% of compounds.

Of course the major weakness of 2D fingerprints is the low probability that any "lead hops" will result from their use. There are many successful lead hops buried within these 8832 World Drug Index structures (indeed, from its method of construction, most structural lead hops that were successful enough to reach consideration for clinical evaluation should be present!). As described above, on the basis of unpublished lead hop successes such as those shown in Table 1, we chose a topomer similarity "shell" between 200 and 240 as being especially productive of unambiguous lead hops. Among this set of compounds, this topomer similarity "shell" happens to contain the structures that are between the 9th and the 23rd percentiles most topomerically similar to structures belonging to a particular activity class. The Tanimoto 2D fingerprint shell that was roughly comparable to this topomer shell in its percentile distribution of "similar" structures was bounded by Tanimoto coefficients of 0.55 and 0.65. (Table 3).

Which "shell" contains more acceptable "lead hops"? Table 4 provides a few structural data points that address this essentially subjective question. Shown for an arbitrarily chosen though relatively "central" query structure are "established lead hops", structures that were in the Tanimoto and/or topomer "lead hopping shells", relative to the query structure shown ("potential lead hops"), and which also belong to the same activity class as the query ("actual lead hops"). Subjectively it seems to us that as structural "lead hops" the topomer selections have decisively higher quality. More objectively, the "Tanimoto 2D fingerprint" selections are on average too topomer similar to the query to be convincing lead hops, while the "topomer" lead hops are on average too 2D-fingerprint-dissimilar to be Tanimoto-detectable (the average 0.33 Tanimoto coefficient for these topomer "lead hops" being equivalent to the average Tanimoto between randomly chosen structures).

Since it is well-known that 2D fingerprints are not very effective in "lead hopping", it may be asked how well other 3D based similarity approaches would have

performed in such a study. Though surprising when first observed,[1] it is by now well confirmed that in general 3D based similarity approaches (typically pharmacophoric) have not concentrated activity together even as well as have 2D fingerprints, while we have just seen that topomers instead concentrate activity much better than do 2D fingerprints. Yet, despite the unpromising performance of most 3D descriptors so far, the very superiority of the topomer approach constitutes proof that other superior 3D similarity descriptions may well exist, awaiting only discovery and/or validation.

Why does the topomer abstraction of physicochemical reality predict biological similarity at least as well as anything else, even "Tanimoto 2D fingerprints"? (Not, to be sure, that the reasons for the relative effectiveness of Tanimoto 2D fingerprints are well understood!) Our current thoughts on this subject are the following. The main emphasis in the development of computer-aided molecular design methodologies has, quite understandably, been increasingly accurate physicochemical modeling. However, as our understanding of the processes responsible for drug action slowly improves, it becomes increasingly clear how daunting is the molecular complexity that a rigorously physicochemical model of drug action should address. For example there are many conformational and tautomeric states of every potential drug molecule that might bind productively to any of the many more conformational and tautomeric states of their intended biomacromolecular target. The strengths of any interaction may be significantly modified by the many local configurations of solvent and other biochemical species. It is clear that only some small subset of the resulting combinatorial combinations of physicochemical states can practically be considered, that the choice of that subset must be at least somewhat arbitrary, and that even some relatively small subset of these possibilities for only one potential drug molecule will be computationally costly to analyze. But today hundreds of thousands of structures are routinely tested, so that many orders of magnitude more structures should be evaluated computationally as candidates for testing.

The topomer description of molecular structure resulted from acknowledgment of these frustrations inherent in the dominant purely physicochemical paradigm and a consequent decision to try an altogether different approach. The central strategy in the development of topomers has been to substitute relentless consistency, computational expedience, and lots of experimental data for physicochemical accuracy, hoping thereby for more productive comparisons of molecular shape. However, such an idiosyncratic approach is not easily justified except by empirical results such as these.

This superiority of topomer similarity in forecasting biological properties of structures is particularly useful when combined with the extraordinary speed of topomer similarity searching within ChemSpace virtual libraries. For example, all the "lead hops" illustrated in Table 1 could have been discovered directly, by experimentally screening all ~80000 compounds in all thirteen assays. On the other hand, complete topomer similarity searches of 2×10^{13} structures,[7] roughly a million times more compounds than have ever been registered into *Chemical Abstracts*, each readily synthesizable from commercially offered building blocks, routinely finish in a few hours.[49] And of course, when these 2×10^8 times more than 80000 candidates are the candidates, many of the potential lead hops retrieved are far more attractive than those reported in Table 1. The flexibility and responsiveness resulting from such a rapid, flexible, and accurate means of candidate structure selection would also facilitate other trends in drug discovery. For example, the growing emphasis on considering multiple lead series and multiple biological endpoints during earlier stages of drug discovery, reinforced by the plummeting costs of multiplexed data acquisition, would imply that combinatorial numbers of SAR may soon need to be simultaneously generated and optimized. We are not aware of any other methodologies as appropriate as the related "topomer CoMFA"[10] for addressing these cheminformatics challenges emerging from "systems biology" and "chemical genomics".

Conclusion

These two validation studies, seemingly among the largest and most complete to be published for any computational methodology, indicate that the topomer description of molecular structure is more effective in predicting "lead hops" (otherwise unexpected biological activity) than other ligand-based approaches, and at least comparable in its effectiveness to receptor-based approaches such as docking. Considering also the enormous speed superiority of topomer searching within virtual libraries,[7] which can make unprecedentedly vast structure spaces usefully accessible and the increasing biological complexities of drug discovery more tractable, it seems reasonable to maintain that topomer searching represents a significant addition to the methodologies potentially available for meeting growing demands on medicinal chemists.

Acknowledgment. We particularly thank Peter Hecht, Mike Lawless, Mark Allen, and John McAlister for their support of the "lead hopping" projects and the requisite negative control experiments, and the referees for making some useful suggestions.

References

(1) Brown, R. D.; Bures, M. G.; Martin, Y. C. Similarity and Clustering Analysis Applied to Molecular Diversity. *Chem. Abstr.* **1995**, *209*, 3-COMP.
(2) Patterson, D. E.; Cramer, R. D.; Ferguson, A. M.; Clark, R. D.; Weinberger, L. E. Neighborhood behavior: a useful concept for validation of molecular diversity descriptors. *J. Med. Chem.* **1996**, *39*, 3049−3060.
(3) Matter, H. Selecting Optimally Diverse Compounds from Structure Databases: A Validation Study of Two-Dimensional and Three-Dimensional Molecular Descriptors. *J. Med. Chem.* **1997**, *40*, 1219−1229.
(4) Martin, Y. C.; Kofron, J. L.; Traphagen, L. M. Do Structurally Similar Molecules Have Similar Biological Activity? *J. Med. Chem.* **2002**, *47*, 4350−4358 and other references therein.
(5) Cramer, R. D.; Redl, G.; Berkoff, C. E. Substructural Analysis. A Novel Approach to the Problem of Drug Design. *J. Med. Chem.* **1974**, *17*, 533−534.
(6) Cramer, R. D.; Clark, R. D.; Patterson, D. E.; Ferguson, A. M. Bioisosterism as a molecular diversity descriptor: steric fields of single topomeric conformers. *J. Med. Chem.* **1996**, *39*, 3060−3069.
(7) Andrews, K. M.; Cramer, R. D. Toward General Methods of Targeted Library Design: Topomer Shape Similarity Searching with Diverse Structures as Queries. *J. Med. Chem.* **2000**, *43*, 1723−1740.
(8) Cramer, R. D.; Patterson, D. E.; Clark, R. D.; Soltanshahi, F.; Lawless, M. S. Virtual libraries: a new approach to decision making in molecular discovery research. *J. Chem. Inf. Comput. Sci.* **1998**, *6*, 1010−1023.

(9) Cramer, R. D.; Poss, M. A.; Hermsmeier, M. A.; Caulfield, T. J.; Kowala, M. C.; Valentine, M. T. Prospective Identification of Biologically Active Structures by Topomer Shape Similarity Searching. *J. Med. Chem*. **1999**, *42*, 3919−3933.
(10) Cramer, R. D. Topomer CoMFA: A Design Methodology for Rapid Lead Optimization. *J. Med. Chem*. **2003**, *46*, 374−389.
(11) Jilek, R. J.; Cramer, R. D. Topomers: A Validated Protocol for their Self-Consistent Generation. *J. Chem. Inf. Comput. Sci*. **2004**, *44*, 1221−1227.
(12) Abrahamian, E.; Fox, P. C.; Naerum, L.; Christensen, I. T.; Thogersen, H.; Clark, R. D. Efficient Generation, Storage, and Manipulation of Fully Flexible Pharmacophore Multiplets and Their Use in 3-D Similarity Searching. *J. Chem. Inf. Comput. Sci*. **2003**, *43*, 458−468. This description of a recent advance in 3-D pharmacophore similarity recognition is accompanied by thorough discussions and references.
(13) CONCORD was developed by R. S. Pearlman, A. Rusinko, J. M. Skell, and R. Baducci at the University of Texas, Austin, and is distributed exclusively by Tripos, Inc., 1699 S. Hanley Rd., St. Louis, MO 63144.
(14) World Drug Alert, Thomson/Derwent, http://thomsonderwent.com/products/pca/worlddrugalerts/.
(15) Cramer, R. D.; Jilek, R. J.; Andrews, K. M. dbtop: Topomer Similarity Searching of Conventional Databases. *J. Mol. Graphics Modell.* **2002**, *20*, 447−462.
(16) Experimental details for the individual assays may be found at www.mdsps.com. The only exception to the 10 μM screening concentration was the first round of PDE4 testing, where the vendor chose a concentration of 100 μM in order to reproduce the reported effect of a control compound. For this publication these experimental 100 μM % binding responses were interpolated to 10 μM using the logit transform, yielding values whose accuracy was confirmed by retests of the most active hits at 10 μM.
(17) World Drug Index, Thompson/Derwent, http://thomsonderwent.com/products/lr/wdi/.
(18) Kuroda, S.; Takamura, F.; Tenda, Y.; Itani, H.; Tomishima, Y. Design, synthesis and biological evaluation of a novel series of potent, orally active adenosine A1 receptor antagonists with high blood-brain barrier permeability. *Chem. Pharm. Bull*. **2001**, *49*, 988−998.
(19) Khatuya, H.; Pulito, V. L.; Jolliffe, L. K.; Li, X.; Murray, W. V. Novel thiophene derivatives for the treatment of benign prostatic hyperplasia. *Bioorg. Med. Chem. Lett*. **2002**, *12*, 2145−2148.
(20) Finke, P. E.; Oates, B.; Mills, S. G.; MacCross, M.; Malkowitz, L. Antagonists of the human CCR5 receptor as anti-HIV-1 agents. Part 4: synthesis and structure-activity relationships for 1-[N-(methyl)-N-(phenylsulfonyl)amino]-2-(phenyl)-4-(4-N-(alkyl)-N-(benzyloxycarbonyl)amino)piperidin-1-yl)butanes. *Bioorg. Med. Chem. Lett*. **2001**, *11*, 2475−2479.
(21) Habeeb, A. G.; Praveen, R. P. N.; Knaus, E. E. Design and synthesis of 4,5-diphenyl-4-isoxazolines: novel inhibitors of cycloxygenase-2 with analgesic and antiinflammatory activity. *J. Med. Chem*. **2001**, *44*, 2921−2927.
(22) Weber, B.; Schlicker, E.; Sokoloff, P.; Stark, H. Identification of the dopamine autoreceptor in the guinea-pig retina as D2 receptor using novel subtype-selective antagonists. *Br. J. Pharmacol*. **2001**, *133*, 1243−1248.
(23) Lober, S.; Hubner, H.; Utz, W.; Gmeiner, P. Rationally based efficacy tuning of selective dopamine D4 receptor ligands leading to the complete antagonist 2-[4-(4-chlorophenyl)piperazin-1-ylmethyl]pyrazolo[1,5-*a*]pyridine (FAUC 213). *J. Med. Chem*. **2001**, *44*, 2691−2694.
(24) Enyedy, I. J.; Wang, J.; Zaman, W. A.; Johnson, K. M.; Wang, S. Discovery of substituted 3,4-diphenyl-thiazoles as a novel class of monoamine transporter inhibitors through 3-D pharmacophore search using a new pharmacophore model derived from mazindol. *Bioorg. Med. Chem. Lett*. **2002**, *12*, 1775−1778.
(25) Kornberg, B. E.; Lewthwaite, R. A.; Manning, D. D.; Nikam, S. S.; Scott, I. L. Piperidine derivatives are NMDA receptors useful for the treatment of e.g. stroke, cerebral ischemia, CNS disorders, depression, trauma, hypoglycemia, neurodegenerative disorders, Parkinson's disease, anxiety, migraine headache and convulsions. Patent WO200250070A2, issued 7/11/2002.
(26) Dumas, J.; Hatoum-Mokdad, H.; Sibley, R. N.; Smith, R. A.; Scott, W. J. Synthesis and pharmacological characterization of a potent, orally active p38 kinase inhibitor. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1559−1562.
(27) Van der May, M.; Hatzelmann, A.; Van der Laan, I. J.; Sterk, G. J.; Thibaut, U.; Timmerman, H. Novel selective PDE4 inhibitors. 1. Synthesis, structure−activity relationships, and molecular modeling of 4-(3,4-dimethoxyphenyl)-2*H*-phthalazin-1-ones and analogues. *J. Med. Chem*. **2001**, *44*, 2511−2522.
(28) Del Castillo Nieto, J. C.; Lasheras Aldaz, B. E.; Monge Vega, A.; Mourelle Mancini, M.; Del Rio Zambrana, J. New benzothiophene derivatives are 5-HT reuptake inhibitors and 5-HT1A receptor modulators useful for the treatment of *e.g.* neurological disorders, depression, psychosis and anxiety. Patent WO200244170A2, issued 6/27/2002.
(29) Nishiyama, A.; Bougauchi, M.; Minoguchi, M.; Morio, Y.; Horikawa, T. New piperidine compounds are 5-hydroxytryptamine-1A receptor antagonists and serotonin reuptake inhibitors useful for treating and preventing e.g. depression and CNS disorders. Patent WO200242297A1, issued 7/11/2002.
(30) Bender, S. L.; Bhumralkar, D.; Collins, M. R.; Cripps, S. J.; Deal, J. G. New amide derivatives are protein tyroinse kinase inhibitors useful for the treatment of e.g. tumor growth, cell proliferation and angiogenesis. Patent WO200153274A1, issued 8/9/2001.
(31) Downing, D.; Clark, J. *Statistics The Easy Way. Barron's Educational Services*; Hauppauge, NY, 1997. Provides excellent references for calculations of χ^2 (pp 208−211) and t (pp 202−206).
(32) Snedecor, G. W.; Cochran, W. G. *Statistical Methods*; Iowa State Press: Ames, IA, 1989; pp 117−120.
(33) Holliday, J. D.; Salim, H.; Whittle, M.; Willett, P. Analysis and Display of the Size Dependence of Chemical Similarity Coefficients. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 819−828. In agreement with another estimate: Godden, J. W.; Xue, L.; Bajorath, J. Combinatorial Preferences Affect Molecular Similarity/Diversity Calculations Using Binary Fingerprints and Tanimoto Coefficients. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 163−166.
(34) At the 10 μM concentration the ≥20% activity threshold, yielding an average 20% of actives among random selections as shown in Table 1, extrapolates to a 40 μM IC_{50} using the logit transform. The ≥90% activity threshold similarly corresponds to just over 1 μM.
(35) Doman, T. N.; McGovern, S. L.; Witherbee, B. J.; Kastern, T. P.; Kurumbail, R.; Stallings, W. C.; Connolly, D. T.; Shoichet, B. K. Molecular Docking and High Throughput Screening for Novel Inhibitors of Protein Tyrosine Phosphatase-1B. *J. Med. Chem*. **2002**, *45*, 2213−2221.
(36) Paiva, A. M.; Vanderwall, D. E.; Blanchard, J. S.; Kozarich, J. W.; Williamson, J. M.; Kelly, T. M. Inhibitors of dihydrodipicolinate reductase, a key enzyme of the diaminopimelate pathway of *Mycobacterium tuberculosis. Biochim. Biophys. Acta* **2001**, *1545*, 67−77.
(37) We also encountered scattered reports of other general hit rates much greater than 0.2%. In particular, the NovaScreen Receptor Selectivity Mapping Data Base (RSMDB) in one early snapshot comprised 250 compounds (all of established pharmacological interest) by 67 targets (mostly receptor/channel binding), totaling 16750 data points, of which 1500 to 1600 (9−10%) had IC_{50}/K_i determinations triggered by >75% response at 10 μM screening concentration (H. Chen, private communication). This RSMDB overall hit rate closely resembles our hit rates for random compound selections. One of us recalls a similar experience from an internal study by a former employer. NovaScreen reports similar hit rates in general screening, while the PanLabs MDS catalog reports even higher hit rates of 25%, but of course in both laboratories the identity of the outside samples screened is unknown and not likely to be random.
(38) However the available evidence completely contradicts the proposition that the activities summarized in Table 1 result mostly from aggregation. Indeed, the first round of negative controls was designed in part to address this issue, where the finding that the control hit rate was virtually identical to that obtained for randomly chosen compounds shows that such promiscuous activity is not significant. Furthermore, the screening concentration of 10 μM is rather low for aggregation, conventional dose−response curves were usually obtainable, and generally the structures themselves lack the features found most correlative with aggregate formation (high log *P*, low proportion of tertiary N, high proportion of COOH, and extensive conjugation). Seidler, J.; McGovern, S. L.; Doman, T, N.; Shoichet, B. K. Identification and Prediction of Promiscuous Aggregating Inhibitors among Known Drugs. *J. Med. Chem*. **2003**, *46*, 4477−4486.
(39) Lynn, P. D.; Kenny, P. W.; Cosgrove, D. A.; Deng, C.; Zabludoff, S.; Wendoloski, J. J.; Ashwell, S. Identification of Compounds with Nanomolar Binding Affinity for Checkpoint Kinase-1 Using Knowledge-Based Virtual Screening. *J. Med. Chem*. **2004**, *47*, 1962−1968.
(40) Brenk, R.; Naerum, L.; Gradler, U.; Gerber, H.-D.; Garcia, G. A.; Reuter, K.; Stubbs, M. T.; Klebe, G. Virtual Screening for Submicromolar Leads of tRNA-guanine Transglycosylase Based on a New Unexpected Binding Mode Detected by Crystal Structure Analysis. *J. Med. Chem*. **2003**, *46*, 1133−1143.
(41) Gruneberg, S.; Wendt, B.; Klebe, G. Subnanomolar Inhibitors from Computer Screening: A Model Study Using Human Carbonic Anhydrase II. *Angew. Chem., Int. Ed*. **2001**, *40*, 389−393.
(42) Grzybowski, B. A.; Ishchenko, A. V.; Kim, C.-Y.; Topalov, G.; Chapman, R.; Christiansen, D. W.; Whitesides, G. M.; Shakh-

novich, E. I. Combinatorial computational method gives new picomolar ligands for a known enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1270−1273.

(43) Vinkers, H. M.; de Jonge, M. R.; Daeyaert, F. F. D.; Heeres, J.; Koymans, L. M. H.; van Lenthe, J. H.; Lewi, P. J.; Timmerman, H.; Van Aeken, K.; Janssen, P. A. J. SYNOPSIS: SYNthesize and OPtimize System In Silico. *J. Med. Chem*. **2003**, *46*, 2765−2773.

(44) Docking into homology-built models of membrane-bound receptors, though often practiced, has not to our knowledge been prospectively validated with success rates comparable to those in Table 1.

(45) Wang, S.; Milne, G. W. A.; Yan, C.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. Discovery of Novel, Non-Peptide HIV-1 Protease Inhibitors by Pharmacophore Searching. *J. Med. Chem*. **1996**, *39*, 2047−2054. Wang, S.; Sakamuri, S.; Enyedy, I. J.; Kozikowski, A. P.; Deschaux, O.; Bandyopadhyay, B. C.; Tella, S. R.; Zaman, W. A.; Johnson, K. M. Discovery of a Novel Dopamine Transporter Inhibitor, 4-Hydroxy-1-methyl-4-(4-methylphenyl)-3-piperidyl 4-Methylphenyl Ketone, as a Potential Cocaine Antagonist through 3D-Database Pharmacophore Searching. Molecular Modeling, Structure−Activity Relationships, and Behavioral Pharmacological Studies. *J. Med. Chem*. **2000**, *43*, 351−360. Carlson, H. A.; Masukawa, K. M.; Rubins, K.; Bushman, F. D.; Jorgensen, W. L.; Lins, R. D.; Briggs, F. D.; McCammon, J. A. Developing a Dynamic Pharmacophore Model for HIV-1 Integrase. *J. Med. Chem*. **2000**, *43*, 2100−2114.

(46) Shen, M.; Beguin, C.; Golbraikh, A.; Stables, J. P.; Kohn, H.; Tropsha, A. Application of Predictive QSAR Models to Database Mining: Identification and Experimental Validation of Novel Anticonvulsant Compounds. *J. Med. Chem*. **2004**, *47*, 2356−2364.

(47) Jain, A. N. Ligand-Based Structural Hypotheses for Virtual Screening. *J. Med. Chem*. **2004**, *47*, 947−961.

(48) This particular data set continues to serve as the benchmark for improvements in topomer similarity searching. Recently, mainly by dynamically adjusting the relative weighting of shape and feature differences according to their relative importance within each query structure, substantial reductions have been achieved in the incidence of false positives at larger topomer distances. For example, referring to the "average" line at the bottom of Table 3, the "hit rates" value for "<200" increases from 25.3% to 34.3%, as the result of selecting only 70% as many structures while retaining 95% of the previous number of hits. However, when validation is the objective, it seems more appropriate to emphasize results achieved *before* the data set was used to improve the methodology being validated.

(49) Currently, development of the topomer technology is emphasizing the reduction of limitations on the diversity of structures amenable to rapid topomer similarity searching. A still-unnamed third generation of this technology, reduced to practice but not yet deployed, accesses structures whose synthesis involves more operations than combichem-style linkage of building blocks. Whereas less than 5% of structures discussed in the *Journal of Medicinal Chemistry* would typically be found among the 2×10^{13} structures in current ChemSpace libraries, it appears that these $\sim 10^{20}$ structures will include about half the structures in the medicinal chemistry literature, limited mostly by which reactions, starting materials, and numbers of synthetic steps are considered appropriate for use in "high throughput medicinal chemistry". Search speeds, always the primary topomer design consideration, are still projected to be faster than overnight.

JM049501B