JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

Discovery of α7-Nicotinic Receptor Ligands by Virtual Screening of the Chemical Universe Database GDB-13

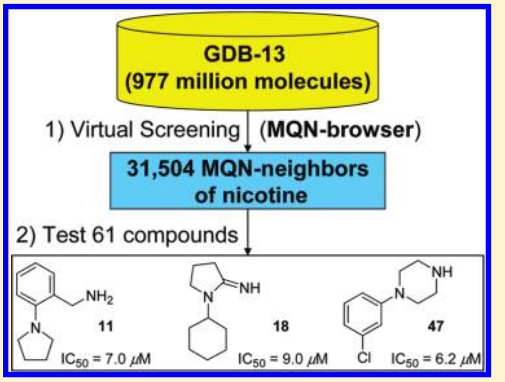
Lorenz C. Blum,[†] Ruud van Deursen,[†] Sonia Bertrand,[‡] Milena Mayer,[‡] Justus J. Bürgi,[†] Daniel Bertrand,[‡] and Jean-Louis Reymond*[,†]

[†]Department of Chemistry and Biochemistry, University of Berne, Freiestrasse 3, 3012 Berne, Switzerland

[‡]HiQScreen, 15 rue de l'Athénée, 1206 Geneva, Switzerland

Supporting Information

ABSTRACT: The chemical universe database GDB-13 enumerates 977 million organic molecules up to 13 atoms of C, N, O, Cl, and S that are virtually possible following simple rules for chemical stability and synthetic feasibility. Analogs of nicotine were identified in GDB-13 using the city-block distance in MQN-space (CBD_{MQN}) as a similarity measure, combined with a restriction eliminating problematic structural elements. The search was carried out with a Web browser available at www.gdb.unibe.ch. This virtual screening procedure selected 31 504 analogs of nicotine from GDB-13, from which 48 were known nicotinic ligands reported in Chembl. An additional 60 virtual screening hits were purchased and tested for modulation of the acetylcholine signal at the human α7 nAChR expressed in *Xenopus* oocytes, which led to the identification of three previously unknown inhibitors. These experiments demonstrate for the first time the use of GDB-13 for ligand discovery.



INTRODUCTION

In the context of exploring chemical space for drug discovery,[1−8] we recently reported the chemical universe databases GDB-11[9,10] and GDB-13,[11] which enumerate all organic molecules up to 11 atoms (26.4 million molecules), respectively, 13 atoms (977 million molecules) that are virtually possible following simple rules for chemical stability and synthetic feasibility. This publicly available resource represents a vast potential reservoir for new drugs. However virtual screening (VS) of these large databases is difficult because current scoring functions are relatively slow and limited to at best a few million structures.[12,13] For instance, our initial ligand discovery examples from the GDB-11 database used high-throughput docking to score virtual compounds and focused on subsets of only a few hundred thousand preselected structures due to the limitations of docking in terms of throughput.[14−17]

To address the problem of efficient VS of very large databases, such as GDB-13, we have recently developed a simple classification system for small organic molecules based on a 42-dimensional chemical space created by assigning dimensions to 42 integer value descriptors of molecular structure, called molecular quantum numbers (MQNs).[18] Most interestingly, similarity measures based on MQNs (cityblock distance CBD_{MQN} or the Tanimoto similarity coefficient T_{MQN})[19,20] relative to a reference bioactive compound can be used as scoring functions to perform ligand-based VS, which leads to enrichment factors similar to those obtained by substructure fingerprint similarity scoring (T_{SF} or CBD_{SF}) for recovering a series of known bioactives from PubChem[21] and the fragment subset of PubChem,[22] with the notable advantage that they allow lead-hopping relationships[23] between actives to be identified. The GDB-13 database can be searched within seconds using the CBD in MQN space (CBD_{MQN}) to a reference compound as similarity measures using a web-based application is freely available at www.gdb.unibe.ch.[24]

Herein we report the first example of ligand discovery from the GDB-13 database using its MQN-browsable version. Due to our interest in identifying new α7 nAChR modulators,[16] we have used this tool to search for new bioactive analogs of nicotine, a 12 atom natural product present in GDB-13. The MQN-browser served to select 31 504 analogs of nicotine, 48 of which had documented activity at the nicotinic receptor as listed in Chembl.[25] Another 60 ligands were acquired and tested by electrophysiology,[26] which led to the identification of **11**, **18**, and **47** as three previously unreported inhibitors of the α7 nicotinic acetyl choline receptor (α7 nAChR), a drug target of importance in neurological disorders, such as schizophrenia, Alzheimer's disease, Parkinson's disease, tobacco addiction, and cancer (Figure 1).[27,28]

RESULTS AND DISCUSSION

Virtual Screening. Analysis of the CBD_{MQN} distribution of the entire GDB-13 relative to nicotine showed a bell-shaped histogram with a maximum frequency at $CBD_{MQN} = 38.8 \pm 11.1$.

Received: September 1, 2011
Published: November 12, 2011

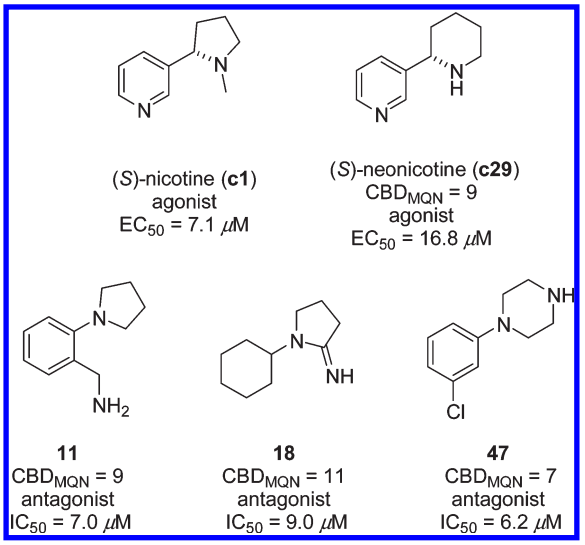


Figure 1. Nicotine and analogs identified by virtual screening and testing in GDB-13. EC_{50} and IC_{50} values are for the $\alpha 7$ nAChR in human (**c1**, **11**, **18**, and **47**) or rat (**c29**). Values for **c1** and **c29** from ref 29.

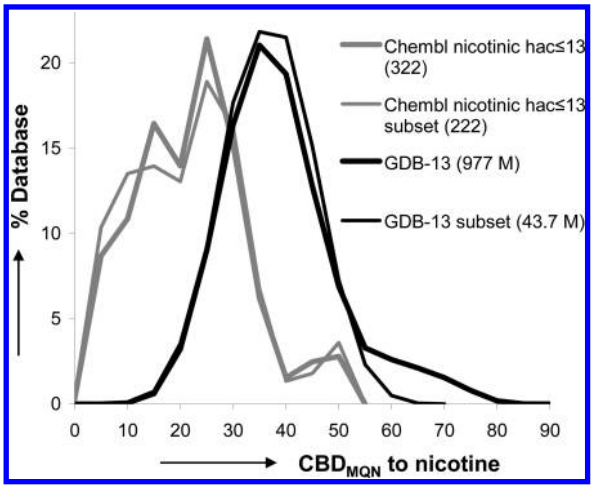


Figure 2. Histogram of CBD_{MQN} to nicotine for various subsets of GDB-13 and Chembl. The "subsets" refer to structures without nonaromatic CC unsaturations, nonaromatic NN or NO bonds, labile functional groups, or small rings.

By contrast, the 322 compounds up to 13 atoms of C, N, O, S, or Cl found in the 2445 unique compounds annotated with nicotinic acetylcholine receptor bioactivity (from any subtype, species, and assay) in the Chembl database[25] showed much shorter CBD_{MQN} distances to nicotine, with a maximum frequency at 22.8 ± 12.5.

Many molecules in GDB-13 present structural elements that are rare or undesirable in bioactive molecules, in particular hydrolytically or oxidatively labile functional groups (esters, carbonates, epoxides, aziridines, aldehydes, 29% of GDB-13), nonaromatic CC unsaturations (63% of GDB-13), nonaromatic NN and NO bonds (in oximes and hydrazones, 35% of GDB-13), and three- or four-membered rings (54% of GDB-13). The MQN browser allows one to exclude these structural features, which substantially reduces the number of compounds to be considered for VS. Eliminating compounds containing these structural elements reduced GDB-13 to 43.7 million molecules (4.5% of GDB-13). On the other hand, applying these restrictions to the 322 bioactive nicotinic ligands up to 13 atoms in Chembl retained 222 of them (68.9%). These 222 known nicotinic ligands were again much closer to nicotine (CBD_{MQN} = 22.3 ± 12.8) than the 43.7 million GDB-13 subset (CBD_{MQN} = 36.8 ± 8.4), showing that the higher MQN similarity of known nicotinic ligands compared to GDB-13 molecules was not caused by the presence of problematic structural elements in the latter (Figure 2).

The shorter CBD_{MQN} distance of known nicotinic ligands to nicotine compared to GDB-13 molecules suggested that GDB-13 molecules with a relatively short CBD_{MQN} distance to nicotine should have a higher probability of bioactivity against the nicotinic receptor compared to molecules taken randomly from GDB-13. This was also suggested by a previous analysis showing that CBD_{MQN}-neighbors of nicotine in GDB-13 have a higher shape similarity to nicotine as measured by the shape similarity score ROCS,[30] compared to randomly selected GDB-13 molecules, a phenomenon that was observed with ten different drugs up to 13 atoms.[24]

A cutoff distance of $CBD_{MQN} \leq 11$ from nicotine was applied to select compounds from the 43.7 million GDB-13 subset. The VS procedure was carried out using the Web browser at www.gdb.unibe.ch, which required 4.9 s server time and retrieved 31 504 molecules from GDB-13 (0.0032% of the entire database). Among these 31 504 molecules, 48 were found within the 2445 known actives retrieved from Chembl (compounds **c1**−**c48**, Figure 3). Searching for availability showed that 692 (2.2%) of the VS hits appeared in the ZINC database[31] (allowing any halogen to replace chlorine). To test if the MQN similarity search would allow to identify nicotinic ligands in addition to those reported in Chembl, 61 of the VS hits found in ZINC, which were available from one supplier, were acquired for activity testing. These 61 compounds covered the range $4 \leq CBD_{MQN} \leq 11$ and included the known active (±)-neonicotine **c29** as a positive control (Figure 4).

Diversity Analysis. *MQNs and Substructure Fingerprints.* The 31 504 VS hits from GDB-13 (in short: 31 504 VS hits), the 692 molecules found in ZINC (in short: 692 ZINC subset), and the 61 ligands selected for testing (in short: 61 purchased subset) have a similar distribution in terms of CBD_{MQN} to nicotine, showing that the selected compounds sampled the MQN space around nicotine similarly to GDB-13 itself (Figure 5A). These three series also showed comparably low similarity to nicotine as measured by the Tanimoto similarity coefficient of substructure fingerprint (T_{SF}). By contrast, a significant fraction of the 48 nicotinic active ligands in Chembl (in short: 48 Chembl subset) had a relatively high substructure similarity to nicotine caused by the frequent occurrence of the pyridine and *N*-methyl pyrrolidine substructure in this set (T_{SF} closer to 1, Figure 5B).

Measuring all pairwise T_{SF} within each subset showed a high degree of substructure diversity (low T_{SF} values) within the three GDB-13 derived subsets (31 405 VS hits, 692 ZINC, 61 purchased) but a somewhat higher substructure similarity (high T_{SF} values) within the known nicotinic ligands in the 48 Chembl subset (Figure 5C). The cross-similarity measures of each subset to the 48 Chembl also showed that the GDB-13 derived subsets were very dissimilar to the 48 Chembl subset in terms of T_{SF} (Figure 5D). This data showed that, as measured by MQN and substructure fingerprints, the selection of compounds for testing from ZINC did not significantly restrict structural diversity compared to the 31 504 VS hits and also that it did not introduce a bias toward known nicotinic ligands.

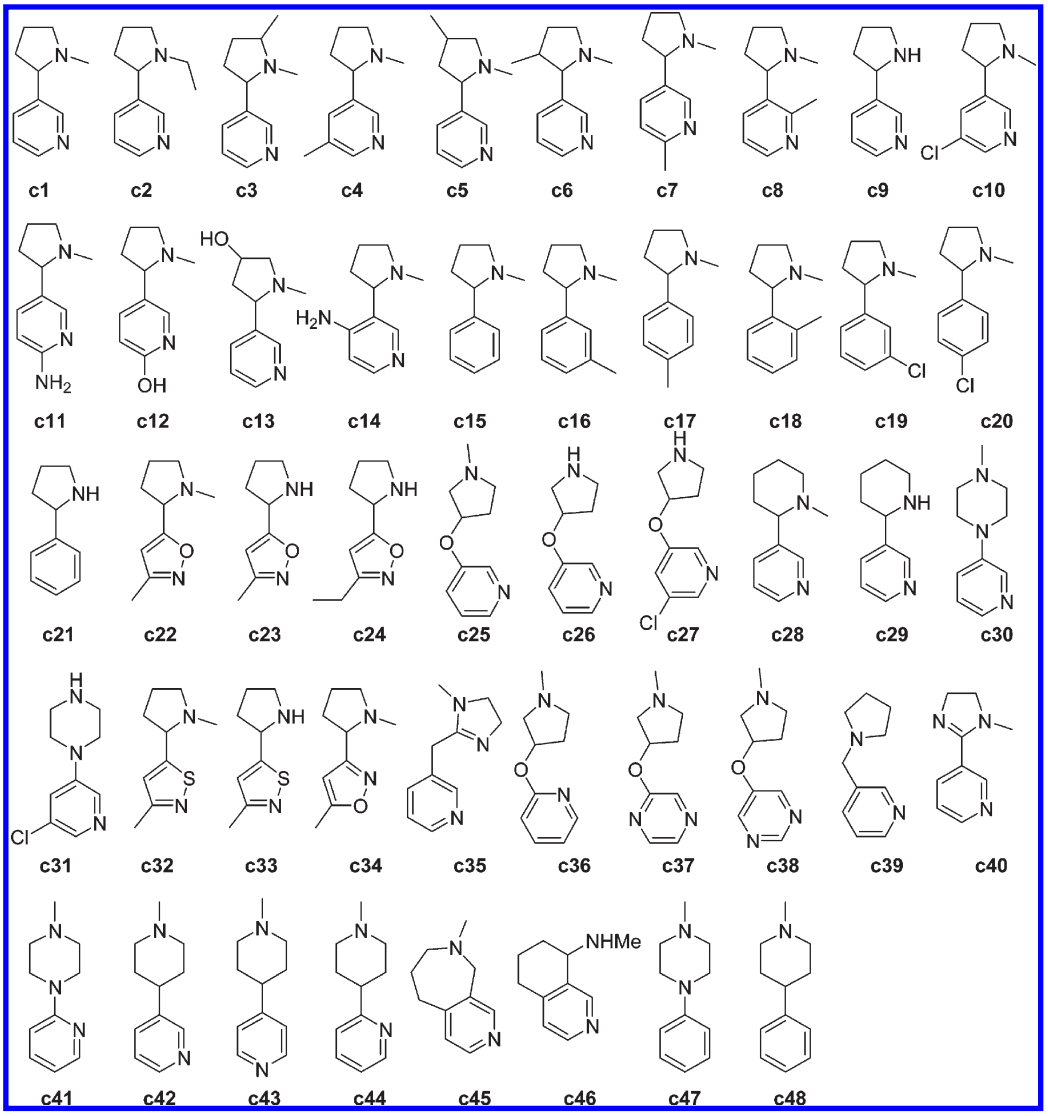


Figure 3. Structures of the 48 Chembl known actives. The compounds are sorted by scaffolds in the order of decreasing number of compound per scaffold (22 scaffolds with 1−14 molecules each, see text). SMILES annotated with T_{SF} and CBD_{MQN} to nicotine are available in the Supporting Information file 48Chembl.smi.

Scaffolds. The 31 504 VS hits were further analyzed in terms of scaffold composition. Among these virtual hits, 31 422 (99.74%) were bicyclic, 82 (0.26%) were monocyclic, and none were acyclic. Each of the 31 504 VS hits was simplified to a parent scaffold, defined here as the structure obtained by iteratively removing all terminal atoms until none are left and reducing all bonds to single bonds. This procedure reduced the data complexity and left 1899 different scaffolds, which represented 2.2% of the 88 249 mono- or bicyclic scaffolds occurring in the initially considered 43.7 million subset of GDB-13. Fourty-five of the 1899 scaffolds corresponded to more than 100 molecules in the 31 504 VS hits, including the nicotine scaffold as could be expected by CBD_{MQN} similarity search (Figure 6A and Figure S1, Supporting Information). On the other hand 332 scaffolds corresponded to only one molecule in the 31 504 VS hits.

The 692 ZINC subset contained 184 scaffolds exemplified by 1−57 molecules each and occurring in 9730 molecules within the 31 504 VS hits (30.88%). In the 692 ZINC subset, 279 compounds (40.3%) contained 25 of the 45 most frequently represented scaffolds in the 31 504 VS hits. The 48 Chembl subset contained 22 different scaffolds exemplified by 1−14 molecules each and occurring in 1158 molecules in the 31 504 VS hits (3.68%). In this case only two scaffolds belonged to the 45 most represented scaffolds, namely the scaffold of nicotine and the scaffold corresponding to molecules **c15**−**c21**. The 61 purchased subset contained 24 scaffolds exemplified by 1−10 molecules each and occurring in 2657 molecules from the 31 504 VS hits (8.4%). Nine of these 24 scaffolds belonged to the 45 most frequent scaffolds.

Ring Types. Further simplification to ring types (defined here as single rings without acyclic atoms and all bonds set to single bonds) showed that the 31 504 VS hits contained 66 (4.2%) of the 1561 ring types available in the 43.7 million GDB-13 subset (Figure 6B and Figure S2, Supporting Information). The 692 ZINC subset contained 28 ring types, which occurred in 21 696 (69%) of the 31 504 VS hits. The 48 Chembl subset contained 9 ring types, which occurred in 15 711 (50%) of the 31 504 VS hits. Finally the 61 purchased subset contained 10 ring types, which occurred in 16 756 (53%) of the 31 504 VS hits. Six ring types

dx.doi.org/10.1021/ci200410u |*J. Chem. Inf. Model.* 2011, 51, 3105–3112

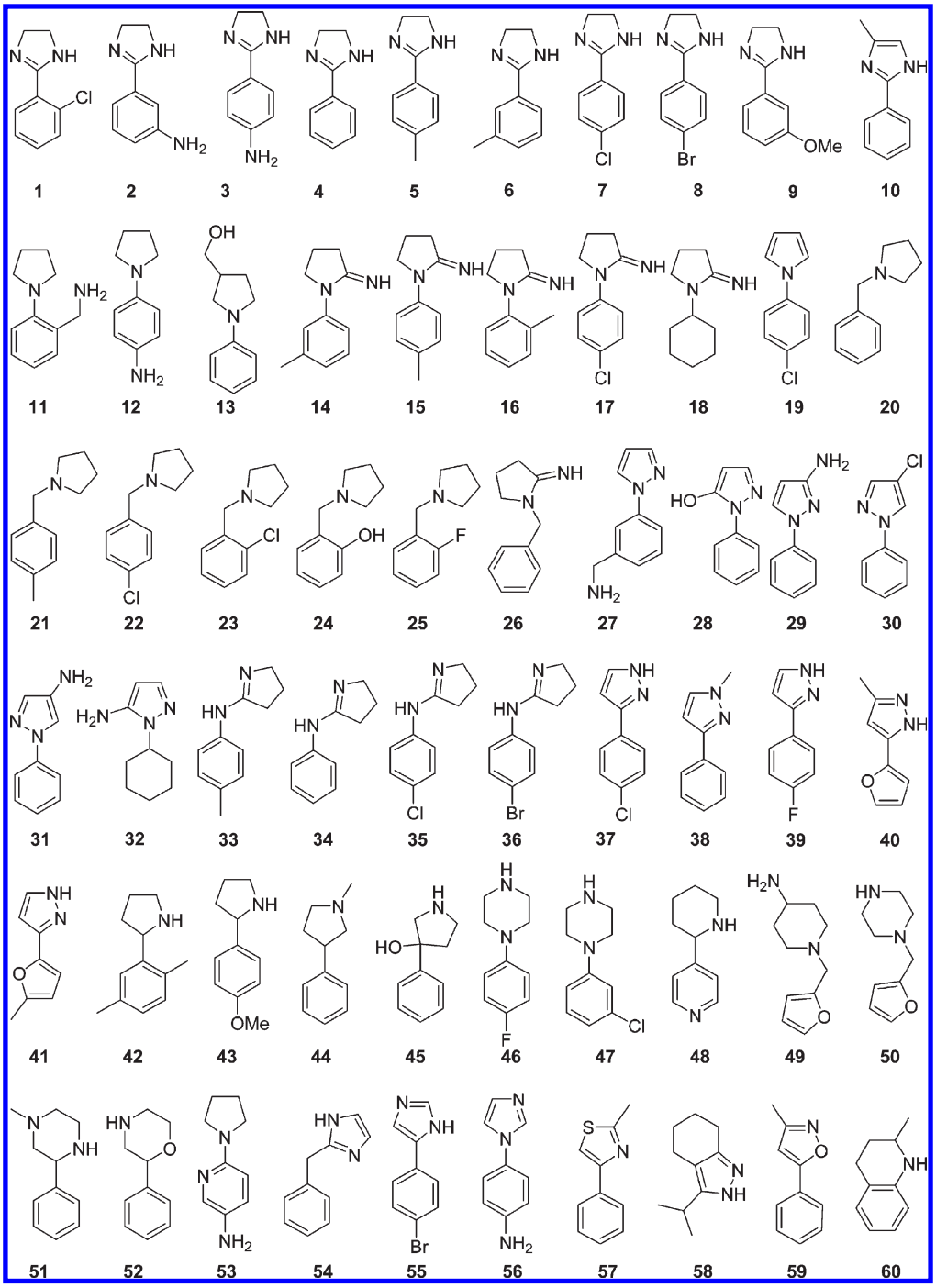


Figure 4. Structures of 60 purchased VS hits. The compounds are sorted by scaffolds in the order of decreasing number of compound per scaffold (24 scaffolds with 1−10 molecules each, see text). SMILES annotated with T_{SF} and CBD_{MQN} to nicotine is available in the Supporting Information file 61purchased.smi.

occurred in both the 48 Chembl subset and the 61 purchased subset, including the 2 most frequent ring types, which were a 5- and a 6-membered ring with one nitrogen found in 12 372, respectively, 11 581 molecules among the 31 504 VS hits. These two ring types were also the first and fourth most frequent ring type in the 43.7 million GDB-13 subset with 16.4 and 6.6 million molecules each. The ring types from the 31 504 VS hits that were not represented in the 692 ZINC or 48 Chembl subsets were almost all 8-membered rings and larger.

The scaffold and ring-type diversity analysis showed that the 31 504 VS hits were not only more numerous than the 692 ZINC, 61 purchased, or 48 Chembl subsets but also that they contained a much larger structural diversity in terms of scaffolds and ring types, which was not readily apparent from substructure fingerprint diversity analysis (Figure 5). Nevertheless, the 61 purchased compounds covered a reasonably broad area of diversity that was clearly different from the already known nicotinic ligands and was judged worth testing for activity.

dx.doi.org/10.1021/ci200410u |J. Chem. Inf. Model. 2011, 51, 3105–3112

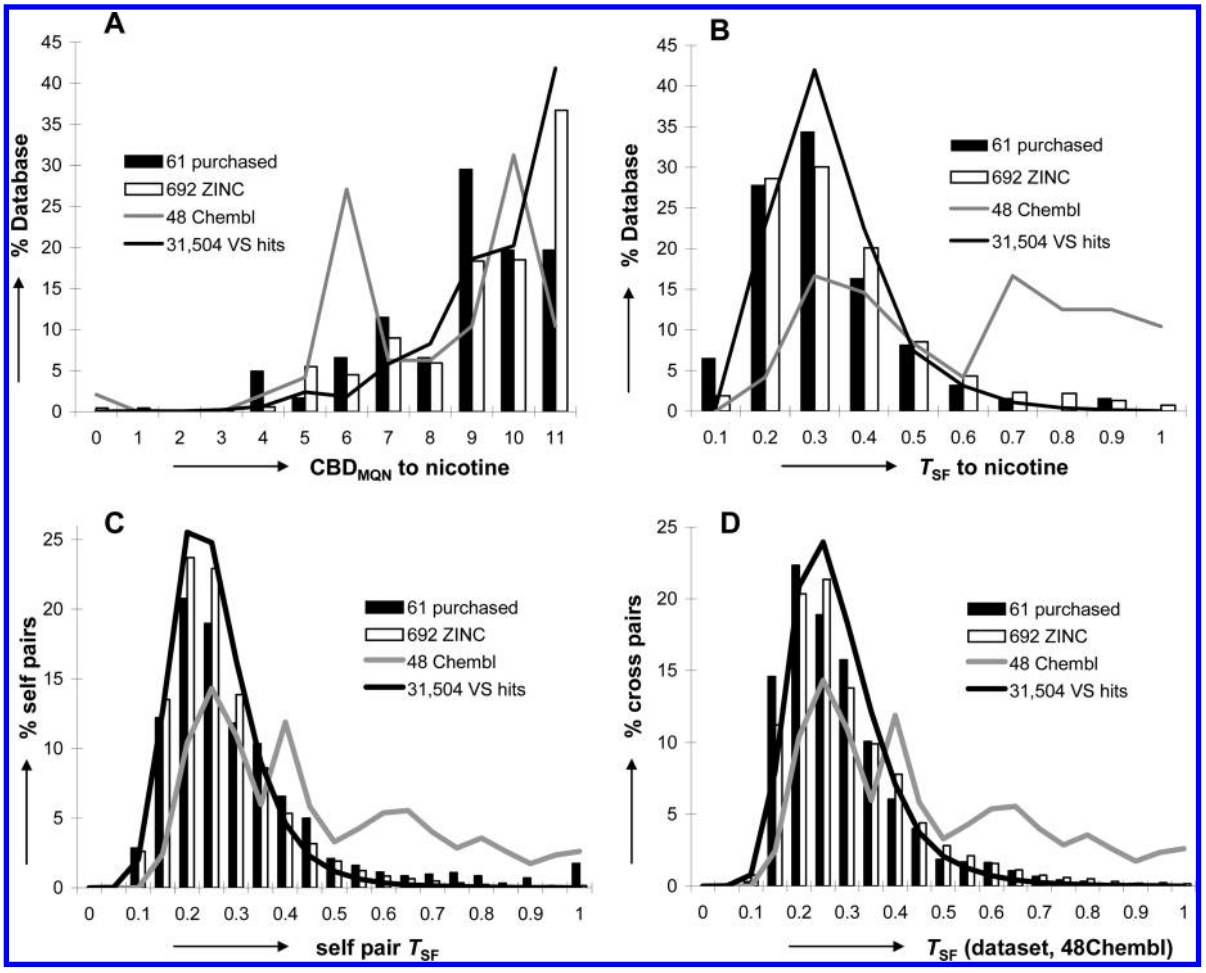


Figure 5. Similarity histograms. (A) CBD in MQN-space CBD_{MQN} to nicotine. (B) Tanimoto similarity coefficient of substructure fingerprint (T_{SF}) to nicotine. (C) Self-similarity of subsets measured by T_{SF} for all pairs within the data set (self-pairs). (D) Similarity of subsets to the 48 known nicotinic ligands in Chembl (48 Chembl) measured by T_{SF} for cross-pair pairs. See Figures 3 and 4 and Supporting Information .smi files for structural formulas of the molecules in each subset.

Electrophysiology. The 61 purchased compounds were tested for their effect on the human $\alpha 7$ nicotinic acetylcholine receptor recombinantly expressed in *Xenopus* oocytes using electrophysiology. To assess properties of compounds, the following experiments were designed: Cells expressing robust acetylcholine (ACh)-evoked currents (100 μM, 5 s) were challenged for 30 s to 10 μM of the test compound under scrutiny. Immediately following compound exposure, cells were exposed again to the same acetylcholine test pulse. This protocol allows examination of the agonist, antagonist, or modulatory effect on the receptors. Compounds acting as agonists should evoke a current during the 30 s exposure and are expected to reduce the ACh-evoked current by desensitization. Compounds evoking no current during the 30 s exposure might reduce the ACh-evoked current revealing a potential antagonist effect.[26]

The positive control **c29** showed a strong agonist effect as well as complete inhibition of the ACh-evoked signal during screening. The other 60 compounds showed a range of inhibitory effects but no significant agonist effect. Overall 11 compounds (18%) effected 60% inhibition of the ACh-evoked current or more (Figure 7). Three of the strongest inhibitors, compounds **11**, **18**, and **47**, were selected for closer investigation due to their interesting structure. Compound **11** possessed an unusual primary amine and a scaffold type both not found in the Chembl actives. Compound **18** contained an amidine functional group present in 19 of the 61 purchased compounds but absent from the Chembl actives and was also nonaromatic unlike most known nicotine analogs. Finally compound **47** was selected because it gave the strongest inhibition signal and was strikingly similar to the Chembl active **c31**. In addition **47** was documented in Chembl with strong activity on various 5-hydroxy-tryptamine receptors but no reported activity on nicotinic receptors. The three compounds were repurchased in larger quantities (25 mg each), purified if necessary, and characterized. Determination of the IC_{50} showed that all three compounds inhibited the $\alpha 7$ nAChR in the micromolar range. Compounds **11** and **47** behaved as competitive inhibitors of ACh, establishing that they bound at the nicotinic binding site. Compound **18** on the other hand showed a noncompetitive inhibition mode, most likely due to direct inhibition of the $\alpha 7$ nAChR ion channel by steric hindrance (Table 1).

Docking. A docking study was performed to test if the activity of compounds **11**, **18**, and **47** might have been predicted by structure-based VS. The compounds were docked to the nicotinic site of the ACh binding protein (1UW6.pdb),[32] using Autodock 3.0.5.[33] The program correctly positioned nicotine in its crystallographically observed pose (Figure 8A). Neonicotine was

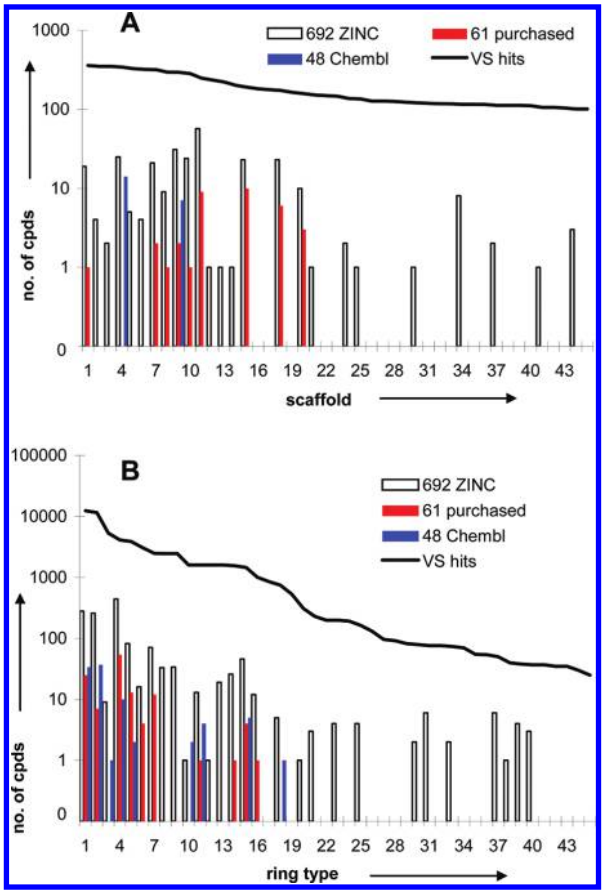


Figure 6. Scaffold and ring type diversity. (A) Number of compounds per scaffold in various subsets. The 45 most frequent scaffolds in the 31 504 VS hits are ordered by decreasing number of compounds per scaffold in the 31 504 VS hits. The scaffold is defined here as the structure obtained from a molecule after iteratively removing all terminal atoms until none is left and reducing all bonds to single bonds. (B) Number of compounds per ring type. The 45 most frequently occurring ring types in the 31 504 VS hits are ordered in decreasing order of compounds per ring type. Ring types are defined here as single rings without acyclic atoms, with in-ring heteroatoms, and all bonds set to single bonds. See Figures S1 and S2 and files scaffolds.top31504.smi and ringtypes.top31504.smi for structures and numbers in the Supporting Information.

Table 1. Inhibition of the α7 nAChR[a]

compound no.	IC_{50}[b]	inhibition mode[c]
11	7.0 ± 0.6 μM	competitive
18	9.0 ± 1.0 μM	noncompetitive
47	6.2 ± 0.4 μM	competitive

[a] Determined by electrophysiology in *Xenopus* oocytes expressing recombinant human α7 nAChR. [b] Inhibition of a pulse of 100 μM ACh using variable concentration of the inhibitor. [c] Measured at the IC_{50} of the inhibitor with variable acetylcholine concentration. Competitive inhibition (**11** and **47**) increases the EC_{50} of ACh but not its maximum signal, noncompetitive inhibition (**18**) decreases the maximum signal of acetylcholine but its EC_{50} is not affected. See Figure S5, Supporting Information.

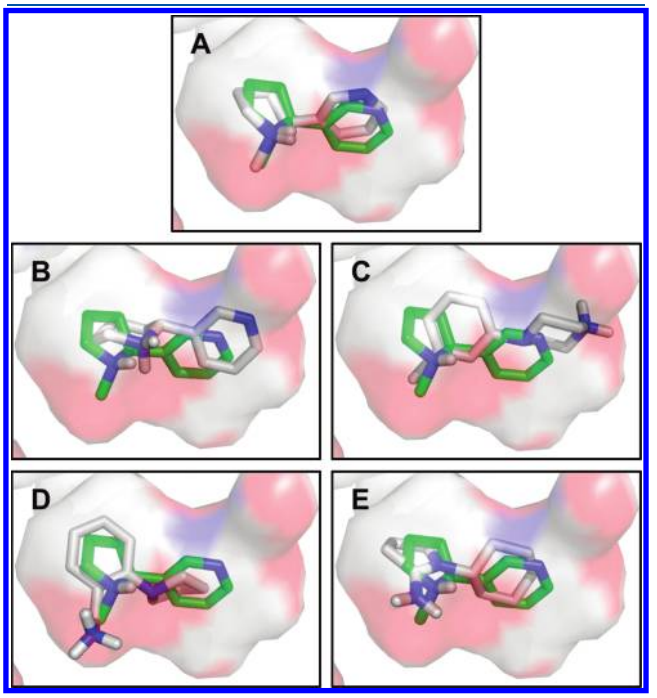


Figure 8. Docking poses of ligands overlaid with the crystallographic position of nicotine in the ACh binding protein AChBP (1uw6). (A) Nicotine, docking energy DE = −8.7 kcal/mol (B) Compound **c29** (neonicotine), docking energy DE = −8.7 kcal/mol. (C) Compound **47**, docking energy DE = −9.1 kcal/mol. (D) Compound **11**, docking energy DE = −9.1 kcal/mol. (E) compound **18**, docking energy DE = −8.6 kcal/mol. See also Figures S3 and S4, Supporting Information.

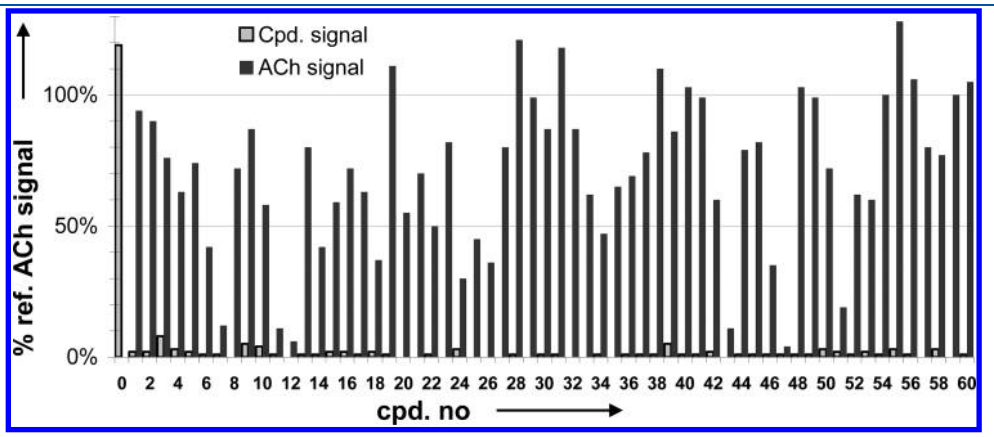


Figure 7. Electrophysiology screening of the 61 compounds selected from GDB-13. The peak intensity is reported as % of the full ACh agonist reference peak upon addition of the test compound (10 μM, agonist, gray bars) and upon addition of ACh (100 μM) after 30 s incubation in presence of the test compound (antagonist, black bars). Compound no. **0** is the positive control **c29** (Figure 3). Compounds **1**−**60** are shown in Figure 4.

positioned in a similar pose with overlapping pyridine and piperidine/pyrrolidine rings (Figure 8B). The newly identified inhibitors docked with poses comparable to that of nicotine (Figure 8C−E), with particularly strong predicted energies of binding (BE) compared to the VS hits and the ZINC, Chembl, and purchased subsets (BE = −9.3 to −6.0 kcal/mol, Figures S3 and S4, Supporting Information). This docking study showed that the ligands identified by MQN similarity could be considered as analogs of nicotine in terms of predicted binding to the nicotine binding site and that they would have been marked as potential analogs of nicotine by structure-based VS. In particular, a prediction would have been made for an agonist effect due to their small size considering that typical antagonists are often much larger than nicotine,[34,35] highlighting the difficulty to distinguish nicotinic agonists from antagonists by VS.

CONCLUSION

The experiments above provide the first demonstration of VS of the entire GDB-13 database leading to the identification of bioactive ligands. Besides the 48 known nicotinic ligands present among the 31 504 MQN neighbors of nicotine selected from GDB-13, two previously unknown competitive inhibitors and one noncompetitive inhibitors of the α7 nAChR were identified. Ligand-based VS by nearest neighbor relationship in MQN space combined with the elimination of problematic structural elements, as shown here, provides a general strategy to exploit GDB-13 for ligand discovery. Diversity analysis in terms of scaffolds and ring types shows that the MQN-similarity searching reaches across a large diversity of structural types. Indeed MQN-similarity searching generally allows one to identify analogs with low substructure similarity and hence lead-hopping relationships,[23] as was the case here for the three identified analogs (T_{SF}(**11**, nicotine) = 0.41, T_{SF}(**18**, nicotine) = 0.31, and T_{SF}(**47**, nicotine) = 0.19). While the present example focused on commercially available ligands within GDB-13 for in vitro testing, the approach should also be applicable to guide synthetic efforts toward new drugs. The MQN similarity search tool freely available at www.gdb.unibe.ch, including the restriction to subsets of GDB-13, provides the essential cheminformatic support for VS to identify diverse and interesting analogs of known ligands for further investigation, which might include additional, more refined VS, such as docking or shape comparisons for prioritization.

METHODS

Nearest Neighbor Search. The nearest neighbors of nicotine were retrieved from GDB-13 using a recently published search tool available on www.gdb.unibe.ch. To get all nicotine neighbors with a CBD up to 11, nicotine was drawn in the main window, and the property filters were set to filter out molecules with unstable functional groups, nonaromatic heteroatom−heteroatom bonds, and nonaromatic carbon−carbon unsaturations and rings with four or less atoms. The search took 4.9 s server time and resulted in a list of 31 504 structures in SMILES format.

Docking. The protonation state of the five compounds nicotine, neonicotine, **11**, **18**, and **47** was set at pH = 7.4 using the cxcalc tool from ChemAxon. The compounds were then expanded to low-energy 3D structures using CORINA. AutoDock Tools (version 3.0.5) were applied to prepare and parametrize the receptor protein and ligands in batch mode. All structures were docked into the target receptor protein using a set of precalculated AutoDock grid maps of interactions between each atom type of the ligands and each atom type of amino acid residues in the binding pocket. Fifty Larmackian genetic algorithm runs were set to output the docking conformation with the best estimated binding energy for each structure.

Structural Similarity. For substructure similarity calculations (T_{SF}) a Daylight-type 512 bit hashed fingerprint from ChemAxon was used.

Screening Solutions. The 61 compounds in Figure 4 were purchased from Princeton Molecular, Inc. as 1 mg solid sample each and dissolved in dimethylsulfoxide to obtain a 10 mM stock solution. Compounds **11**, **18**, and **47** were also purchased in 25 mg, purified (see Supporting Information) and similarly prepared as 10 mM stock solution in dimethylsulfoxide.

Electrophysiology. All experiments were carried out at the human α7 nAChR receptors expressed in *Xenopus* oocytes using the method of cDNA expression. *Xenopus* oocytes were prepared and injected using standard procedures. ACh-evoked currents were recorded using the standard two electrode voltage-clamp configuration (TVEC).

Briefly, ovaries were harvested from *Xenopus laevis* females that have been deeply anesthetized and pithed according to the animal rights rule from the Geneva canton. A small piece of ovary was isolated for immediate preparation, while the remaining part was placed at 4 °C in a sterile Barth solution containing in mM: NaCl 88, KCl 1, $NaHCO_3$ 2.4, HEPES 10, $MgSO_4 \cdot 7H_2O$ 0.82, $Ca(NO_3)_2 \cdot 4H_2O$ 0.33, $CaCl_2 \cdot 6H_2O$ 0.41, at pH 7.4, and supplemented with 20 μg/mL of kanamycine, 100 unit/ml of penicillin, and 100 μg/mL of streptomycin. On the second day following dissociation, oocytes were injected with 2 ng of cDNA per oocyte containing the genes encoding for the human α7 nAChR using an automated injector.[26] All recordings were performed at 18 °C and cells superfused with OR2 medium containing in mM: NaCl 82.5, KCl 2.5, HEPES 5, $CaCl_2 \cdot 2H_2O$ 1.8, at pH 7.4. Cells were held at −80 mV. Data will be filtered at 10 Hz, captured at 100 Hz, and analyzed using proprietary data acquisition and analysis software running under Matlab (Mathworks Inc.).

Concentration inhibition curves were fitted using the empirical Hill equation in the form:

$$y = 1/1 + (x/IC_{50})^{nH}$$

where: y = the fraction of evoked current, x = the antagonist concentration, IC_{50} = concentration for 50% inhibition and, nH = the apparent cooperativity.

Concentration activation curves were fitted using the empirical Hill equation:

$$y = 1/1 + (EC_{50}/x)nH$$

where: y = the fraction of ACh-evoked current, x = the agonist concentration, EC_{50} = the concentration causing 50% activation, and nH the apparent cooperativity.

ASSOCIATED CONTENT

Supporting Information. SMILES files of the 31 504 VS hits, 48 Chembl, 692 ZINC, 61 purchased, 1899 scaffolds, and 66 ring types. Lists of scaffolds and ring types. Details of docking study. Analytical data and electrophysiology data of compounds **11**, **18**, and **47**. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: jean-louis.reymond@ioc.unibe.ch.

ACKNOWLEDGMENT

This work was supported financially by the University of Berne, the Swiss National Science Foundation, the NCCR TransCure to J.-L.R., and by the EEC grant Neurocypres to D. B. We thank Claudette Duret for her contribution to electrophysiological measurements.

REFERENCES

(1) Bohacek, R. S.; McMartin, C.; Guida, W. C. The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.* **1996**, *16*, 3–50.

(2) Pearlman, R. S.; Smith, K. M. Novel software tools for chemical diversity. *Perspect. Drug Discovery Des.* **1998**, *9−11*, 339–353.

(3) Oprea, T. I.; Gottfries, J. Chemography: The art of navigating in chemical space. *J. Comb. Chem.* **2001**, *3*, 157–166.

(4) Dobson, C. M. Chemical space and biology. *Nature* **2004**, *432*, 824–828.

(5) Schneider, G.; Fechner, U. Computer-based de novo design of drug-like molecules. *Nat. Rev. Drug Discovery* **2005**, *4*, 649–63.

(6) Geppert, H.; Vogt, M.; Bajorath, J. Current trends in ligand-based virtual screening: molecular representations, data mining methods, new application areas, and performance evaluation. *J. Chem. Inf. Model.* **2010**, *50*, 205–216.

(7) Reymond, J. L.; Van Deursen, R.; Blum, L. C.; Ruddigkeit, L. Chemical space as a source for new drugs. *Med. Chem. Commun.* **2010**, *1*, 30–38.

(8) Hartenfeller, M.; Schneider, G. De novo drug design. *Methods Mol. Biol.* **2011**, *672*, 299–323.

(9) Fink, T.; Bruggesser, H.; Reymond, J. L. Virtual exploration of the small-molecule chemical universe below 160 Da. *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 1504–1508.

(10) Fink, T.; Reymond, J. L. Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: Assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discovery. *J. Chem. Inf. Model.* **2007**, *47*, 342–353.

(11) Blum, L. C.; Reymond, J. L. 970 million druglike small molecules for virtual screening in the chemical universe database GDB-13. *J. Am. Chem. Soc.* **2009**, *131*, 8732–8733.

(12) Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery Today* **2006**, *11*, 580–94.

(13) Kolb, P.; Ferreira, R. S.; Irwin, J. J.; Shoichet, B. K. Docking and chemoinformatic screens for new ligands and targets. *Curr. Opin. Biotechnol.* **2009**, *20*, 429–436.

(14) Nguyen, K. T.; Syed, S.; Urwyler, S.; Bertrand, S.; Bertrand, D.; Reymond, J. L. Discovery of NMDA glycine site inhibitors from the chemical universe database GDB. *ChemMedChem* **2008**, *3*, 1520–1524.

(15) Nguyen, K. T.; Luethi, E.; Syed, S.; Urwyler, S.; Bertrand, S.; Bertrand, D.; Reymond, J. L. 3-(aminomethyl)piperazine-2,5-dione as a novel NMDA glycine site inhibitor from the chemical universe database GDB. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3832–3835.

(16) Garcia-Delgado, N.; Bertrand, S.; Nguyen, K. T.; van Deursen, R.; Bertrand, D.; Reymond, J.-L. Exploring a7-Nicotinic Receptor Ligand Diversity by Scaffold Enumeration from the Chemical Universe Database GDB. *ACS Med. Chem. Lett.* **2010**, *1*, 422–426.

(17) Luethi, E.; Nguyen, K. T.; Burzle, M.; Blum, L. C.; Suzuki, Y.; Hediger, M.; Reymond, J. L. Identification of selective norbornane-type aspartate analogue inhibitors of the glutamate transporter 1 (GLT-1) from the chemical universe generated database (GDB). *J. Med. Chem.* **2010**, *53*, 7236–7250.

(18) Nguyen, K. T.; Blum, L. C.; van Deursen, R.; Reymond, J. L. Classification of organic molecules by molecular quantum numbers. *ChemMedChem* **2009**, *4*, 1803–1805.

(19) Willett, P.; Barnard, J. M.; Downs, G. M. Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983–996.

(20) Khalifa, A. A.; Haranczyk, M.; Holliday, J. Comparison of Nonbinary Similarity Coefficients for Similarity Searching, Clustering and Compound Selection. *J. Chem. Inf. Model.* **2009**, *49*, 1193–1201.

(21) van Deursen, R.; Blum, L. C.; Reymond, J. L. A searchable map of PubChem. *J. Chem. Inf. Model.* **2010**, *50*, 1924–1934.

(22) van Deursen, R.; Blum, L. C.; Reymond, J. L. Visualisation of the chemical space of fragments, lead-like and drug-like molecules in PubChem. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 649–662.

(23) Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G. "Scaffold-Hopping" by Topological Pharmacophore Search: A Contribution to Virtual Screening. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 2894–2896.

(24) Blum, L. C.; van Deursen, R.; Reymond, J. L. Visualisation and subsets of the chemical universe database GDB-13 for virtual screening. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 637–647.

(25) Warr, W. A. ChEMBL. An interview with John Overington, team leader, chemogenomics at the European Bioinformatics Institute Outstation of the European Molecular Biology Laboratory (EMBL-EBI). *J. Comput.-Aided Mol. Des.* **2009**, *23*, 195–198.

(26) Hogg, R. C.; Bandelier, F.; Benoit, A.; Dosch, R.; Bertrand, D. An automated system for intracellular and intranuclear injection. *J. Neurosci. Methods* **2008**, *169*, 65–75.

(27) Taly, A.; Corringer, P. J.; Guedin, D.; Lestage, P.; Changeux, J. P. Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system. *Nat. Rev. Drug Discovery* **2009**, *8*, 733–750.

(28) Paleari, L.; Cesario, A.; Fini, M.; Russo, P. alpha7-Nicotinic receptor antagonists at the beginning of a clinical era for NSCLC and Mesothelioma? *Drug Discovery Today* **2009**, *14*, 822–836.

(29) *Tocris Bioscience*; http://www.tocris.com (accessed October 17, 2011).

(30) Rush, T. S., III; Grant, J. A.; Mosyak, L.; Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* **2005**, *48*, 1489–1495.

(31) Irwin, J. J.; Shoichet, B. K. ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182.

(32) Celie, P. H.; van Rossum-Fikkert, S. E.; van Dijk, W. J.; Brejc, K.; Smit, A. B.; Sixma, T. K. Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* **2004**, *41*, 907–914.

(33) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.

(34) Hansen, S. B.; Sulzenbacher, G.; Huxford, T.; Marchot, P.; Taylor, P.; Bourne, Y. Structures of Aplysia AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *Embo J.* **2005**, *24*, 3635–3646.

(35) Reymond, J. L.; van Deursen, R.; Bertrand, D. What we have learned from crystal structures of proteins to receptor function. *Biochem. Pharmacol.* **2011**, *82*, 1521−1527.