Interactive evolutionary algorithms and data mining for drug design

Lameijer, E.M.W.

7 Designing Active Template Molecules by Combining Computational *De Novo* Design and Human Chemist's Expertise

Eric-Wubbo Lameijer, Reynier A. Tromp, Ronald F. Spanjersberg, Johannes Brussee, and Adriaan P. IJzerman
Leiden/Amsterdam Center for Drug Research, Division of Medicinal Chemistry, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands

This chapter was first published in the Journal of Medicinal Chemistry, reference: Lameijer, E.W.; Tromp, R.A.; Spanjersberg, R.F.; Brussee, J.; IJzerman, A.P. Designing Active Template Molecules by Combining Computational De Novo Design and Human Chemist's Expertise. *J. Med. Chem.*, **2007**, *50* (8), 1925–1932.

Abstract

We used a new software tool for *de novo* design, the "Molecule Evoluator", to generate a number of small molecules. Explicit constraints were a relatively low molecular weight and otherwise limited functionality, for example low numbers of hydrogen bond donors and acceptors, 1 or 2 aromatic rings, and a small number of rotatable bonds. In this way we obtained a collection of scaffold- or template-like molecules rather than fully "decorated" ones. We asked medicinal chemists to evaluate the suggested molecules for ease of synthesis and overall appeal, allowing them to make structural changes to the molecules for these reasons. On the basis of their recommendations we synthesized 8 molecules with an unprecedented (not patented) yet simple structure, which were subsequently tested in a screen of 83 drug targets, mostly G protein-coupled receptors. Four compounds showed affinity for biogenic amine targets (receptor, ion channel and transport protein), reflecting the training of the medicinal chemists involved. Apparently the generation of lead-like solutions helped

the medicinal chemists to select good starting points for future lead optimization, away from existing compound libraries.

Introduction

Chemical space is vast – the number of potential drug-like molecules has been estimated to be beyond the number of atoms in the universe.[1,2] This is in sharp contrast with the total count of molecules in large compound databases such as CAS, with approximately 25 million references to chemical compounds.[3] Hence, *de novo* design is crucial to cover more of the chemical universe. Computational methods are particularly suitable for this goal, as they can quickly generate and store thousands of putative structures. Currently, there are dozens of *de novo* design programs, many of which have been covered in a recent review.[4] For example, the program CoG (Compound Generator) of Brown et al.[5] constructs molecules based on atoms and fragments that have been given as input to the program, eventually yielding molecules that resemble a number of selected ligands. Other programs construct new molecules based on the structure of the target protein. For example, DycoBlock[6] takes a list of fragments and searches for their optimal position in the active site of the protein. Then it searches for combinations of building blocks that could be linked together to form a new molecule.

We have recently developed a software tool to help medicinal chemists in designing new active structures; we called it "The Molecule Evoluator" (see also chapter 3 of this thesis). The Molecule Evoluator constructs molecules from atoms and a limited number of predefined larger fragments (such as phenyl and carboxylic acid groups). The use of atoms and the ability to attach atoms to any other atom and make rings at all chemically valid positions of a molecule allows an exhaustive search of chemical space and fine–tuning of the molecular structure.

An important difference between the Molecule Evoluator and most other *de novo* design programs is the focus on interaction with the user to produce lead compounds. Instead of generating a large database which is then screened virtually by docking or molecule similarity calculations, it presents a number of molecules to the user, who selects and edits the molecules to make them more lead-like. This cycle of computer generation and user modification can go on for several rounds, hence the name "Molecule Evoluator". This user involvement was inspired by new approaches in computer science that stress the collaboration between computer and user, such as interactive evolutionary computing.[7] The user is able to use his implicit knowledge, e.g.

of synthetic feasibility, to eliminate structures suggested by the program that are difficult to make in the laboratory. The user may also bring in other areas of expertise, such as domain knowledge for a certain drug target, for example in the form of structure-activity relationships.

The aim of the present study was to determine whether combining computational inspiration with the domain knowledge of a number of medicinal chemists could produce novel, biologically active, lead-like structures. We used the Molecule Evoluator in a more constrained way than the usual cycle, in which the molecules modified by the user are fed back to the computer program to "breed" new molecules. Instead, we just created one database of molecules, the structures of which were refined by the medicinal chemists alone. For that we asked a panel of medicinal chemists to select, comment on and amend a limited number of compounds out of the library, which were subsequently checked for novelty. On the basis of their recommendations a limited number of the chosen and amended compounds, further simplified for reasons of chemical feasibility, was synthesized and tested on an array of drug targets. Half of the compounds synthesized possessed significant activity for biological targets, indicating that our combination of computer-based generation of molecules and chemist-based selection and modification can be useful to develop entirely novel lead structures.

Results

De novo design of template molecules

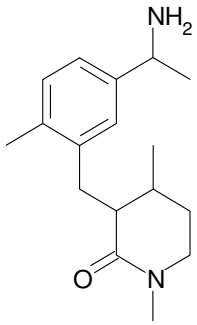
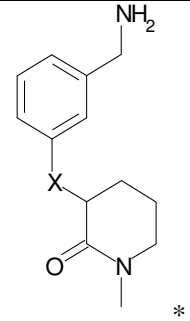
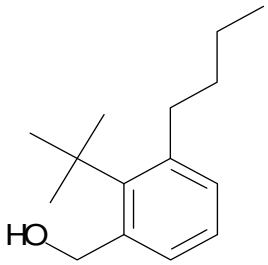
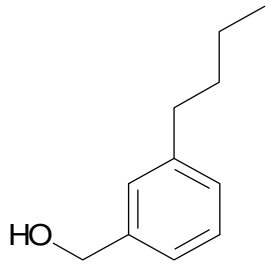
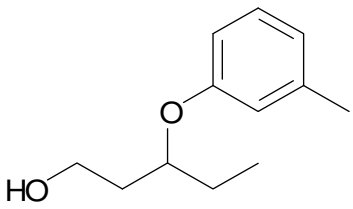
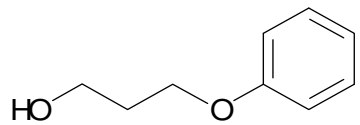
We used the Molecule Evoluator to generate a virtual library of 300 compounds according to a number of restrictions meant to produce template-like rather than drug-like molecules. These limitations, extended on but stricter than Lipinski's "rule of five",[8] were as follows:

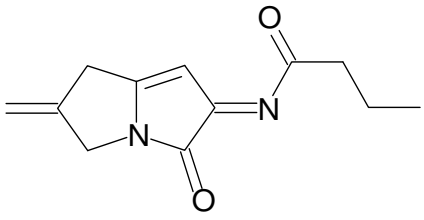
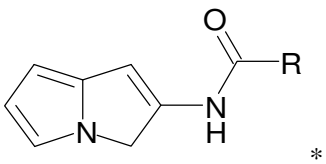
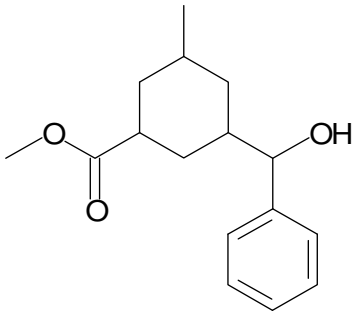
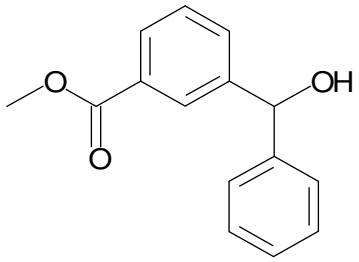
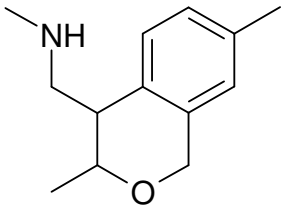
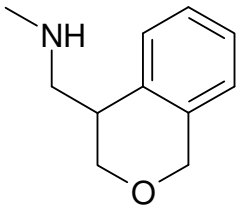
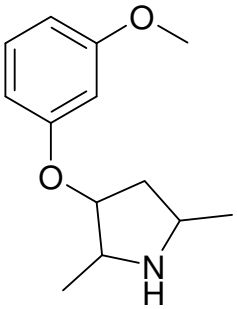
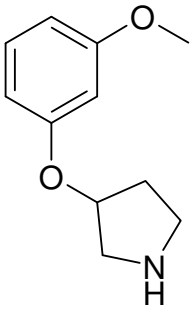
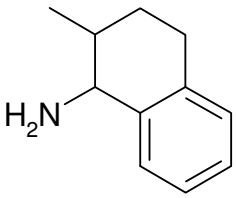
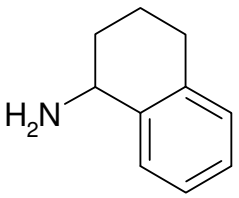
1) At least one and at most two aromatic systems
2) Polar surface area equal to or below 70 $Å^2$
3) A maximum number of two hydrogen bond donors and four hydrogen bond acceptors
4) Not more than five rotatable bonds

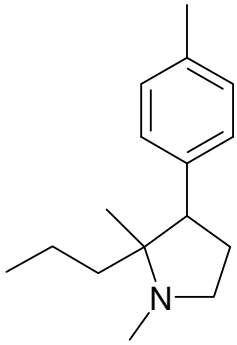
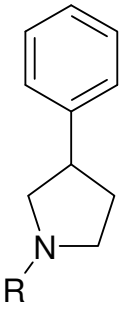
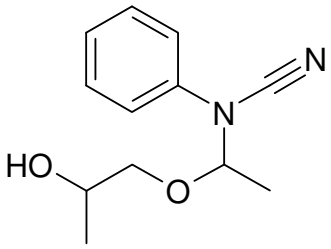
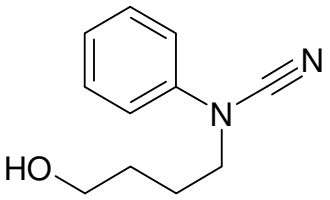
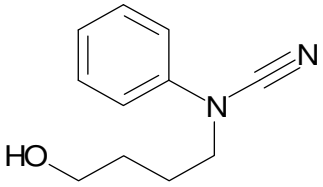
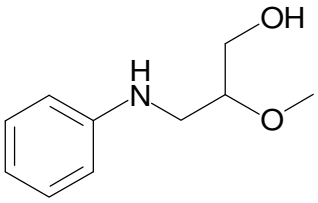
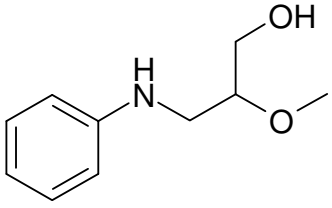
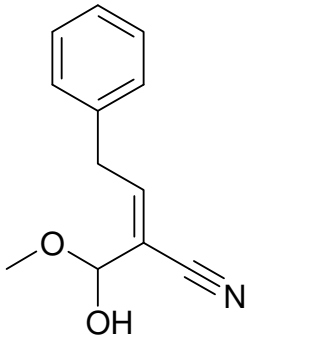
Although we also experimented with molecular weight restrictions we learned that the above four criteria invariably resulted in compounds with molecular weights lower than 400 D, hence lower than "Lipinski's" cut-off of 500 D.

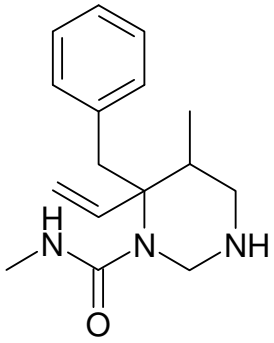
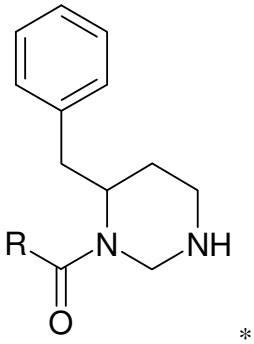
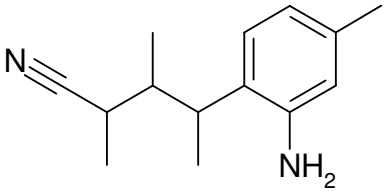
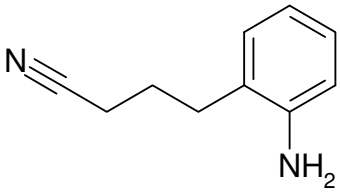
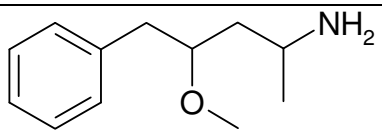
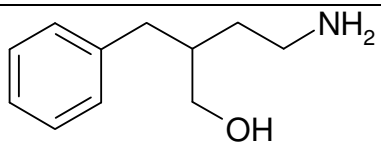
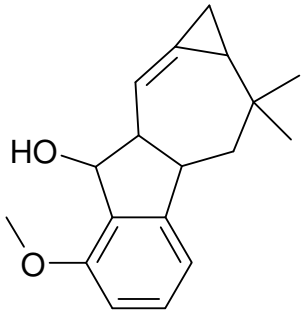
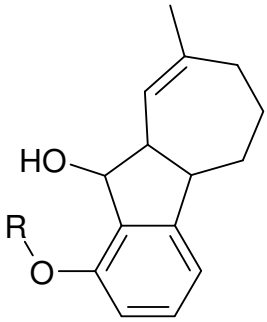
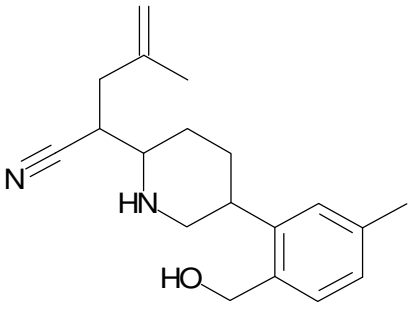
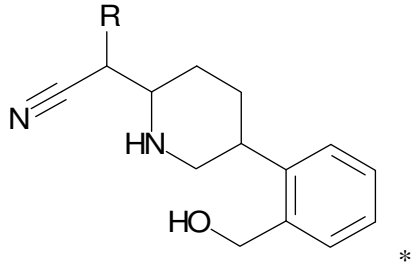
The 300 compounds were presented to a panel of five medicinal chemists with different backgrounds (chemistry of peptides, biogenic amines (2x), nucleosides/ nucleotides, and chiral synthesis). They were asked to select at least 10 compounds to their liking. Specifically, the selected compounds had to look drug-like and synthetically feasible, or at least be amenable to be changed into such compounds by minor modifications. This led to a total of 34 compounds (Table 7.1).

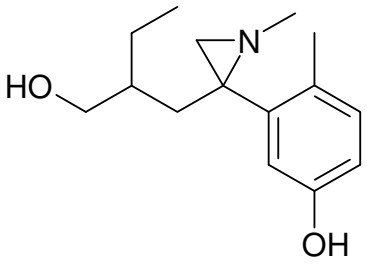
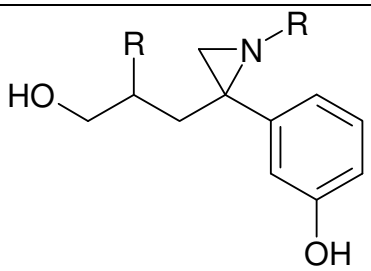
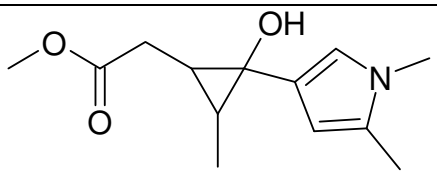
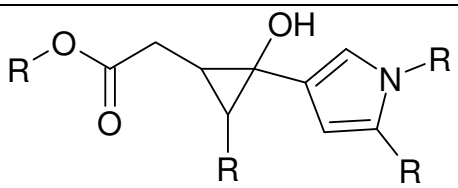
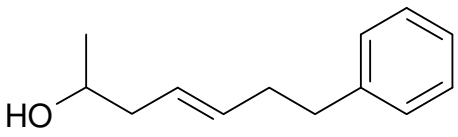
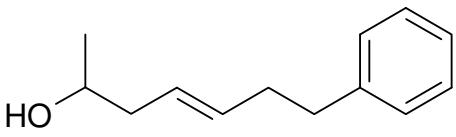
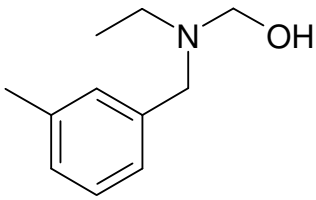
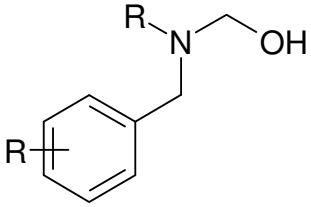
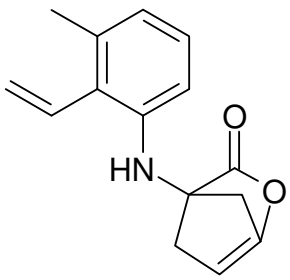
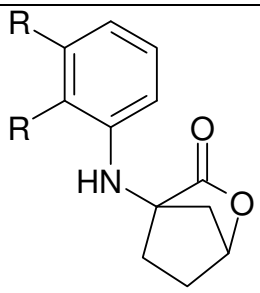
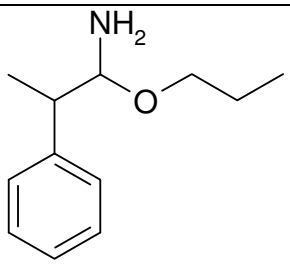
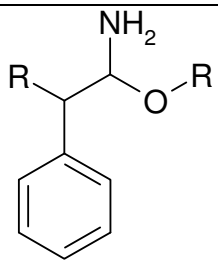
Table 7.1: The 34 compounds selected from the 300-member library generated by the Molecule Evoluator. The left-hand column shows the structures as generated, while the right-hand column lists the structures after initial amendment by the medicinal chemists. Those molecules marked with a star were selected for further amendment and synthesis.

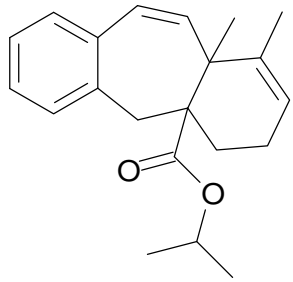
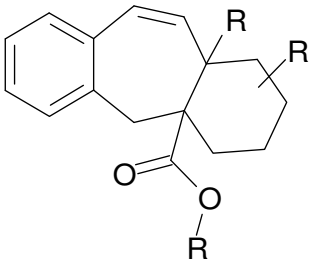
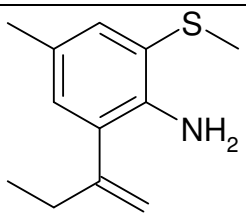
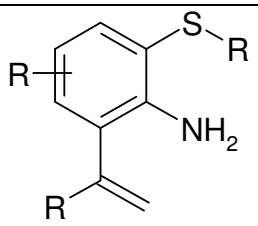
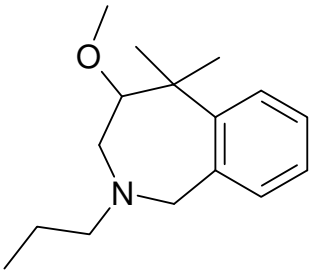
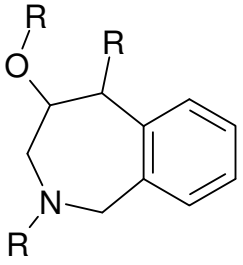
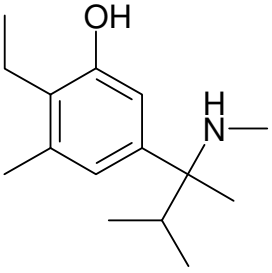
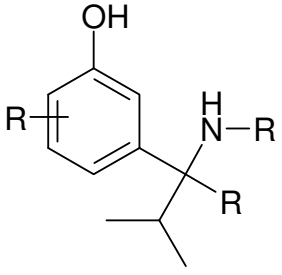
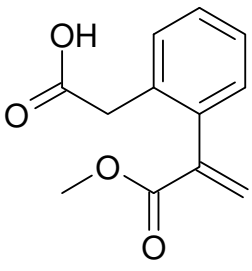
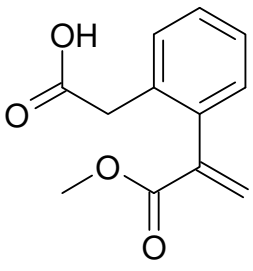
Original molecule	Modified molecule
	
	
	

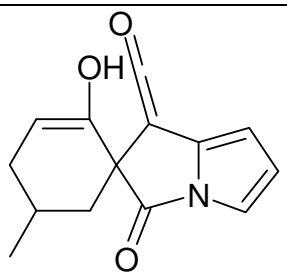
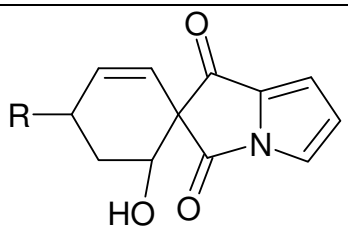
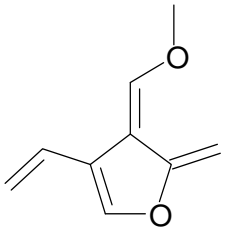
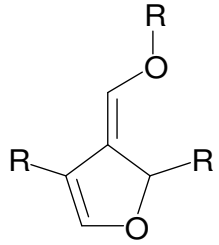
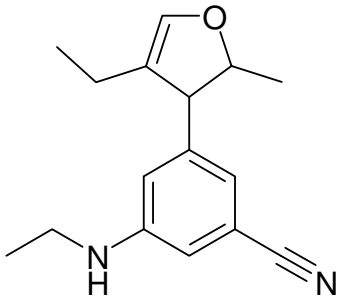
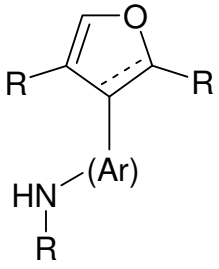
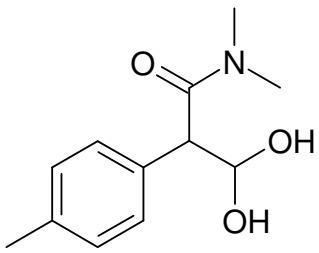
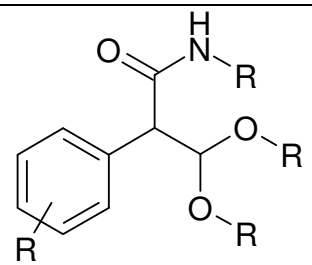
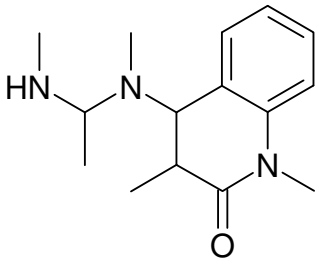
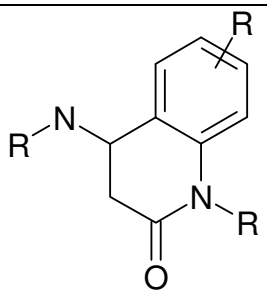
	 *
	
	
	
	

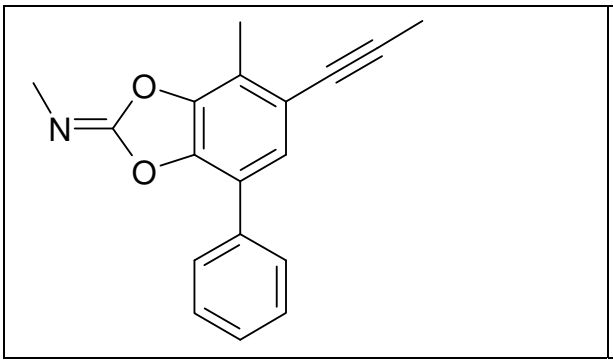
	
	
	
	

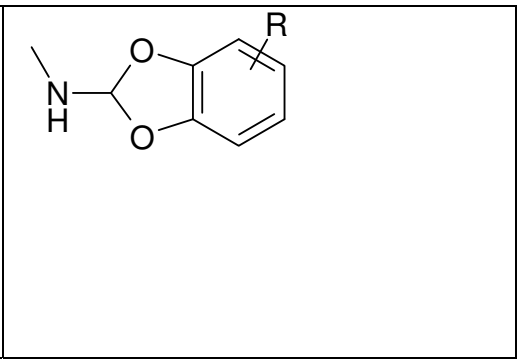
	 *
	
	
	
	 *





Our next step was to inspect the 34 molecules for novelty, ease of synthesis and drug-likeness. Novelty in this case was defined as absence from both the Beilstein and SciFinder databases either as a structure or substructure.[3,9] This process took place in March 2003; we did not check for later occurrence. For ease-of-synthesis we allowed the chemists to modify the suggested structures to reduce the anticipated number of synthetic steps (maximally 3 from a commercially available starting material). Drug-likeness was not only based on the filters that we already applied when the virtual library was generated, but also on the intuition of the individual medicinal chemist. All in all this led to a top-nine of compounds that formed the start for our synthetic program (see Figure 7.1). Two chemists (R.T. and R.S.) were allotted a restricted period of time to try and synthesize these compounds. It was decided to rapidly terminate a project whenever synthetic feasibility in practice was less than anticipated 'on paper'. This was particularly true for compounds **1f** and **1g**. It was also decided to allow further variations on the nine molecules presented in Figure 7.1 on the basis of experimental findings in the synthetic program. As a consequence the final series of compounds, although much inspired by the very first suggestions, deviated from the original structures. In general, the computer-generated molecules were simplified by eliminating most substituents, while the core structure was retained together with one or two of the most important or interesting substituents. Further variation was produced by making derivatives of the remaining substituents (such as oxidizing a CH_2OH group to COOH). Eventually we prepared and characterized eight compounds as represented in Figure 7.2. Their synthesis is outlined in the chemistry paragraph below and described in full detail in the Experimental Section.

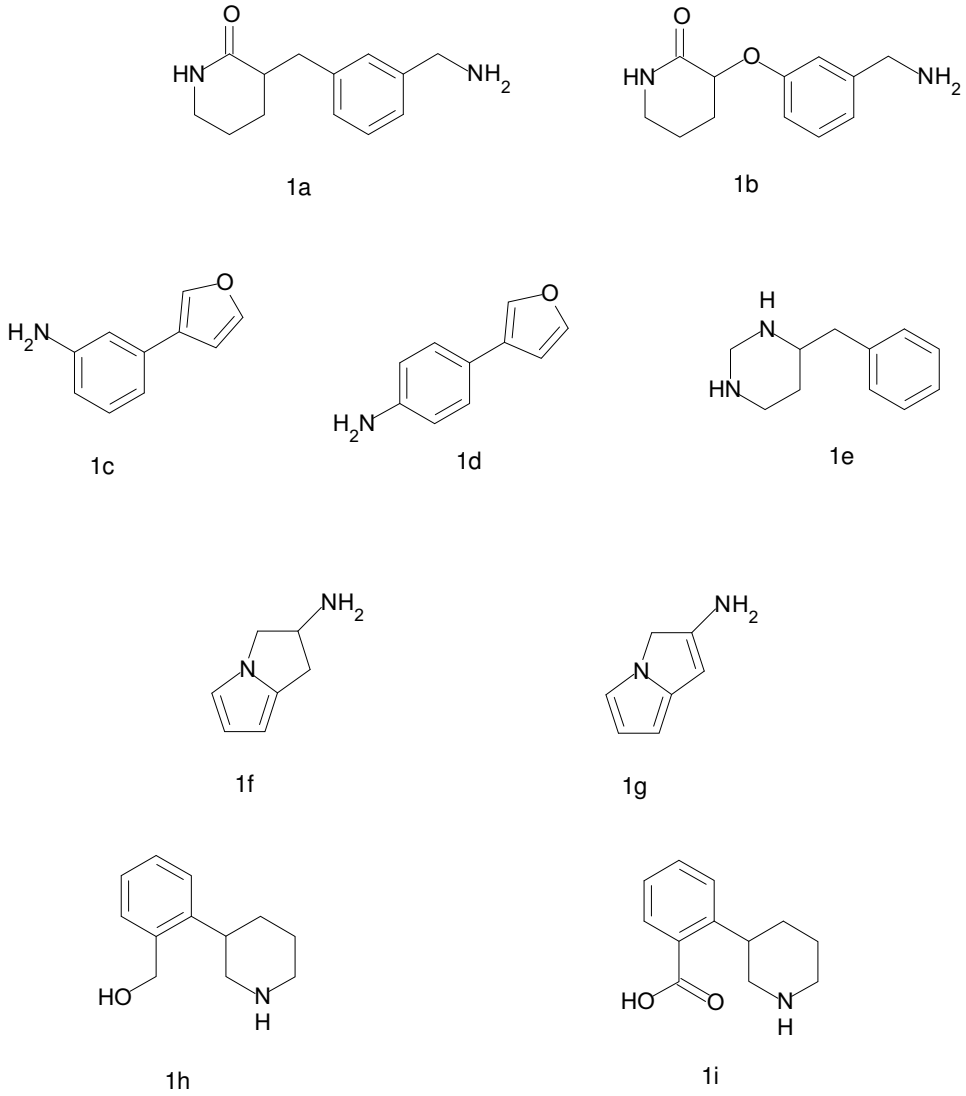
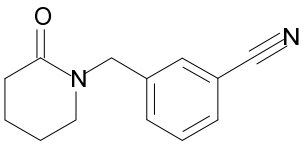
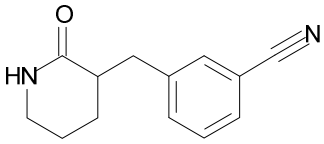


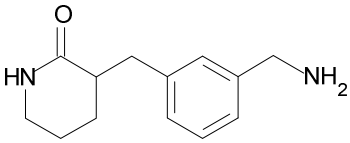
Figure 7.1: The final selection of nine compounds (1a-1i), amended by the panel of medicinal chemists.



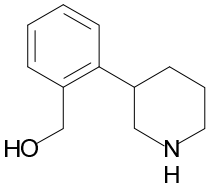
3



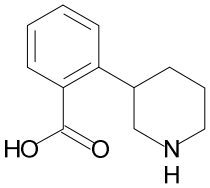
4



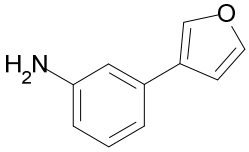
5 (1a)



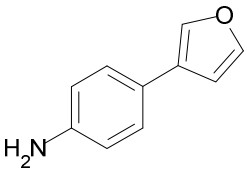
6 (1h)



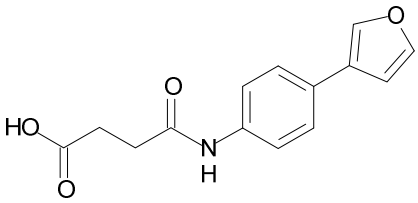
7 (1i)



8 (1c)



9 (1d)

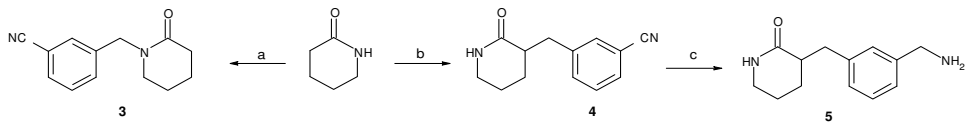


10

Figure 7.2: The eight compounds resulting from the synthetic program (3-10)

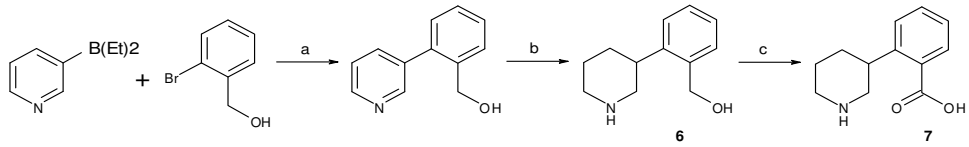
Chemistry

Compound **3** was prepared by substitution of 3-(bromomethyl)benzonitrile with 2-piperidinone which was deprotonated with one equivalent of butyllithium.[10] Synthesis of compound **4** was performed by alkylation at the 3-position of 2-piperidinone via the enolate anion in which the nitrogen atom was temporarily protected with TMS.[11] Hydrogenation of compound **4** with Pd/C as catalyst afforded the benzylamino compound **5** (Scheme 7.1).



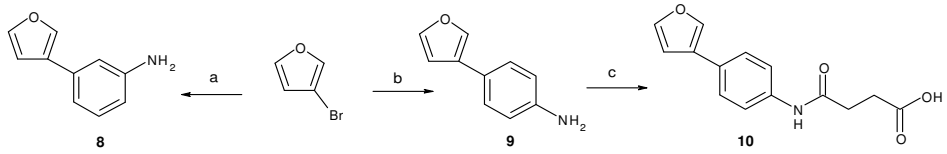
Scheme 7.1: a: 1 eq. *n*-BuLi, 3-(bromomethyl)benzonitrile. b: TMSCl, *n*-BuLi, 3-(bromomethyl)benzonitrile. c: Pd/C 10%, H_2.

Synthesis of 2-(3-piperidyl)-benzyl alcohol (**6**) was done by a two step reaction. First 2-(3-pyridyl)-benzyl alcohol was prepared by a Suzuki reaction of diethyl(3-pyridyl)borane and 2-bromobenzyl alcohol under microwave conditions.[12] The product of this reaction was hydrogenated under acid conditions with PtO_2 as catalyst and provided compound **6**. Benzyl alcohol **6** was oxidized with chromic acid and isolated as zwitterion. Purification was problematic, however preparative HPLC provided pure product **7** (Scheme 7.2).



Scheme 7.2: a: Na_2CO_3, TBAB, $(Ph_3P)_4Pd$, H_2O, MW. b: HCl, PtO_2/H_2. c: Jones' reagent.

The most straightforward way to prepare compounds **8** and **9** was the Suzuki coupling reaction of 3-bromofuran with boron derivatives of 3- and 4-aniline respectively, under microwave conditions. Compound **10** was prepared from **9** by reaction with succinic acid and crystallization from diethyl ether (Scheme 7.3).



Scheme 7.3: a: 3-aminophenyl-boranic acid, Na_2CO_3, TBAB, $(Ph_3P)_4Pd$, MW. b: 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)aniline, Na_2CO_3, TBAB, $(Ph_3P)_4Pd$, MW. c: succinic anhydride, 4-methyl-morpholine.

Biology

We tested the 8 compounds in a commercially available screening program. Radioligand binding and enzyme assays, 68 and 15, respectively, were the read-outs to probe the interaction of the individual compounds with this large collection (83) of drug targets. These included G protein-coupled receptors (rhodopsin-like, class A; metabotropic glutamate-like, class C), ion channels (for Na^+, K^+, Ca^{2+}), nuclear hormone receptors (e.g., estrogen, progesterone), transport proteins (e.g. for dopamine, norepinephrine, GABA), and enzymes (several phosphodiesterases, Na^+/K^+-ATPase, etc). All compounds were tested in duplicate at a single concentration of 10 μM. In Table 7.2 the percentage inhibition of specific radioligand binding to the indicated target (with a minimum of 30%) is shown. Negative values indicate an increase in specific binding. This might indicate an allosteric mechanism of enhancement,[13] but this was not investigated further. Four out of eight compounds displayed activity in a number of radioligand binding assays, while none of the compounds appeared active in the enzyme assays. Compounds **5** (imidazoline and muscarinic receptors), **6** (α-adrenergic receptors), and **8** and **9** (norepinephrine transport protein) caused approximately 50% radioligand displacement or more. It should be mentioned here that compound **7**, being inactive at all tested targets, appealed to one of the chemists for a different reason, i.e. it being an unnatural and new amino acid, which will be used for incorporation in modified peptides.

Table 7.2: Percentage inhibition of specific radioligand binding (min. 30%) to the indicated target by 10 µM of test compounds **3-10**. Negative values indicate an increase in specific binding.

Target		**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
CB_1				-33					
I_2				49	36		41	32	
M_{1-5}				50					
NAch					45		42		
NE transp.							80	77	
DA transp.							-62		
5-HT transp.							-31		
kainate								-37	
α_1-adrenerg.					47				
α_2-adrenerg.					62				
NK_1					-32				
opiate					35				
5-HT					39				

CB_1: cannibinoid receptor 1; I_2: imidazoline receptor 2; M_{1-5}: muscarinic receptors 1-5 in rat brain; NAch: nicotinic acetylcholine ion channel; NE transp: norepinephrine transport protein; DA transp: dopamine transport protein; 5-HT transp: serotonin transport protein; kainate: glutamate/kainate receptor; NK_1: neurokinin receptor 1; opiate: all opioid receptors in rat brain; 5-HT: serotonin receptors in rat brain.

Discussion

For a medicinal chemist, drug-likeness, synthetic feasibility and overall 'molecule appeal' are very important criteria in drug design. However, these features are very difficult to quantify, such that good 'scoring functions' are often lacking. For instance, computer-assisted organic synthesis was recently reviewed by Todd,[14] who concluded that available software invariably required human intervention to be useful. Similarly,

computational approaches to predict ligand binding affinity for a given target protein ("docking") are notoriously inaccurate. Aware of such considerations we decided to rely on the user as evaluator. A user cannot know the binding strength of a given molecule *a priori*, but we reasoned this defect may not be much worse than the inaccuracy of scoring functions. A definite advantage in letting the user choose would be that intensive feedback from a medicinal chemist would make the compounds easier to synthesize, and steer the idea generation away from areas which have already been explored. Furthermore, user feedback could still be combined with experimental results or advanced computed fitness functions if so desired. Considering these advantages we developed a software tool for *de novo* molecule design, called the Molecule Evoluator, which we recently described. It contains a graphical user interface and has options for directly editing the molecule, marking part of a molecule as conserved, and calculating relevant physicochemical parameters.[7]

It should be noted that the Molecule Evoluator mainly uses the atom-based approach to construct molecules, that is, a molecule is built from individual atoms and bonds, though some predefined fragments can be added. A number of other researchers have also constructed molecules in an atom-based way, for example Nachbar[15], Douguet et al.[16] and Brown et al.[5] Others construct molecules from a number of multi-atom fragments, such as Pegg et al.[17], Vinkers et al.[18], and Schneider et al.[19] The main difference between atom-based and fragment-based methods is not so much the size of the fragments used (atom-based methods often also use fragments, and vice versa) but the emphasis placed on synthetic feasibility. Atom-based methods such as ours sample the entire chemical space but also produce molecules of doubtful synthetic feasibility, and fragment-based methods like the one of Vinkers et al.[18] stress synthetic accessibility and therefore sample a much smaller part of chemical space, excluding hard-to-synthesize molecules but also many potential drugs. In the Molecule Evoluator, we have chosen for the flexibility of the atom-based approach, although we are aware of the sensitive issue of synthetic ease and have developed a number of features which allow the user to restrict the variety of molecules produced.[7]

In the present study we generated 300 molecules according to the criteria specified in the Results section. These criteria are well below the classic 'rule-of-five'[9] to largely yield template or scaffold-like molecules only. For example, the number of hydrogen bond donors was confined to a value of two, rather than five. Repeating the experiment would yield a largely different library of molecules due to the random-number generator in our algorithm. While by setting the criteria identical to our original experiment the average physicochemical properties of such molecules would be similar

to those in the original library, the enormous number of molecules possible with a molecular weight between, say, 150 and 300, would ensure that there would be barely any molecules in common between the two libraries. Changing the parameters would force the algorithm to sample another part of chemical space, however the physicochemical properties of the lead structures might not change as greatly as the parameters since the chemists generally adapt the molecules to the complexity of their taste.

The 300 molecules were shown to a panel of medicinal chemists. They examined them for drug-likeness, synthetic feasibility, and overall appeal as mentioned above, and identified their preferences. It should be noted that human judgment is not unequivocal. In a study by Takaoka and coworkers, five chemists judged a collection of almost 4000 molecules in a Japanese corporate database for their drug-likeness and ease of synthesis. Their scores showed considerable variation.[20] A similar inconsistency was noted among 13 medicinal chemists at a US-based company when asked to reject compounds with undesirable properties from one or more lists of 2000 compounds each.[21] Apparently unanimity among medicinal chemists is not self-evident. On a more positive note, their diversity in opinion may in fact constitute an important and discriminative asset for a research group. While our computational generation of the library benefited from human intervention, the chemists themselves also found that the computational generation of molecules added value. They appreciated the many choices possible, which emphasizes that it is easier to *recognize* a "good" structure than to *invent* one.

We did not give the chemists explicit instructions on which molecules should be chosen or rejected, other than that the molecules should seem lead-like and not too difficult to synthesize. Analyzing their choices in retrospect, it became clear that the chemists did use some general "implicit" rules for molecule choice. Molecules without heteroatoms (or with only one heteroatom if that was a nitrogen) were almost always rejected, as were molecules with more than two ring systems, cyclophanes (having a bridged benzene ring), molecules with odd or unwanted groups like halogen atoms or nitro groups, and molecules with many alkyl substituents. However, if anything, these rules seemed more like a weighing of attractive and inattractive features than a black-or-white approval or elimination. For example, one molecule with only one heteroatom, a nitrogen, was nevertheless selected, probably because the nitrogen was in a two-ring system instead of somewhere in a substituent. Occasionally the chemists disagreed about the appropriateness of a certain selection. So instead of general rules one could say that the chemists used general guidelines, which were weighed according to

individual experience and taste.

The compounds that were suggested (Figure 7.1) and eventually synthesized (Figure 7.2) all had a relatively small number of hydrogen bond donors and/or acceptors, next to their low molecular weight, as a logical consequence of the strict criteria imposed. They largely adhere to a recently proposed "rule-of-three" for fragment-based lead discovery, in which molecular weight is <300, the number of hydrogen bond donors is ≤3, the number of hydrogen bond acceptors is ≤3 and the calculated logP value is ≤3[22], and can be considered leads[23] or fragments rather than potential drugs. In this view fragments should have features that, when combined, still adhere to Lipinski's "rule-of-five". The differences between "rule-of-three" and "rule-of-five" allow a further "decoration" of our compounds. At the same time, fragments tend to have very low affinity for a given target, in view of the limited options for interaction.[22] Surprisingly, quite a few of our compounds displayed affinities that allowed them to be recognized in conventional radioligand binding assays, as opposed to more sophisticated and demanding NMR- or X-ray-based screening that is generally applied in fragment-based approaches.

It appeared that most of our ligands intervened with targets for biogenic amines (e.g., adrenergic, muscarinic and serotonin receptors, norepinephrine transport protein, nicotinic acetylcholine ion channel). Interestingly, the background, education and training of some of our medicinal chemists involved in the selection of the compounds had been focused on this important ligand class, suggesting that medicinal chemists can indeed develop a "feel" for a certain target or family of targets.

The chemical structures of the suggested molecules as well as those synthesized are simple, or, as some medicinal chemists put it, "quite boring". Apparently, chemical space is vast, but also nearby, i.e. entirely novel structures can be far from exotic. It suggests that medicinal chemists when asked tend to prefer more uncommon structures. Interestingly, it has been shown on a number of occasions that currently available drugs in fact have low diversity.[24,25] In a recent analysis of the NCI database harboring over 250,000 molecules tested for biological activity we learned that in it 80% of all ring systems found in molecules belonged to one out of the 66 "top" ring systems – which was only 0.5% of the total variety in ring systems in the database. The same analysis taught us that a phenyl ring was present in almost half of the compounds, whereas the next most prevalent (pyridine) ring occurred in less than 3% of the molecules,[27] "quite boring" indeed. The reason may be that exotic ring systems and substituents have undesirable synthetic or biological properties. It emphasizes that our method of template development, which puts "ordinary" parts in novel combinations,

may actually be quite suitable for drug design.

Conclusion

Computational generation of novel molecules, as implemented in the Molecule Evoluator, appeared useful in *de novo* template and scaffold design. It helped a panel of medicinal chemists in generating, amending and selecting a number of 'simple' yet novel chemical entities. A number of low-molecular weight compounds was eventually synthesized and tested on a diverse panel of drug targets. Some of the compounds proved to be active, mainly on targets for biogenic amines, in line with the background and expertise of some of the medicinal chemists. It seems that nearby chemical space still offers substantial room for drug design, and that simple structures can be very attractive.

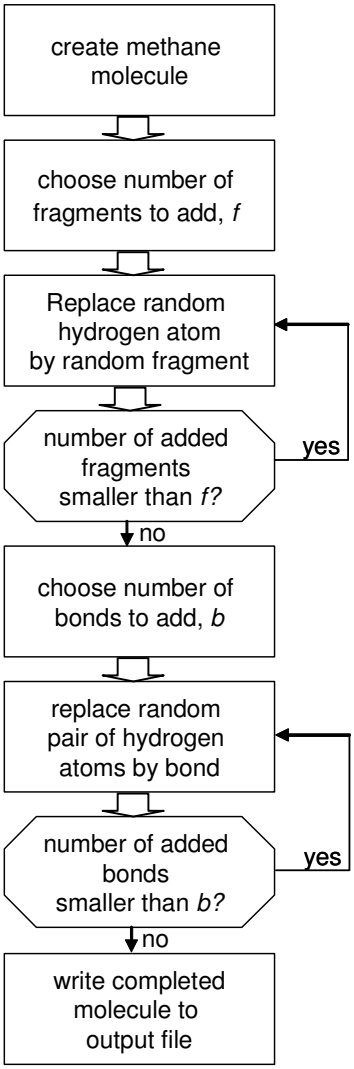
Experimental

De novo design algorithm

The 300 molecules were generated by taking a methane molecule, and growing the molecule for a number of iterations by attaching atoms to it at random positions, and adding double bonds and rings. The algorithm is shown in Figure 7.3.

If a molecule did not obey pre-set criteria (at least one and at most two aromatic systems, polar surface area (calculated according to Ertl et al.[26]) equal to or below 70 $Å^2$, a maximum number of two hydrogen bond donors and four hydrogen bond acceptors, not more than five rotatable bonds) it was discarded and a new molecule was generated, until we had 300 molecules with the desired physicochemical properties.

Figure 7.3: Flowchart of the *de novo* design algorithm. A molecule is generated by adding a random number of fragments (varying from 1 to 16) to a methane molecule, and subsequently adding bonds, thereby creating double bonds and rings. The exact number of rings and double bonds is determined by a weighted probability table, as is the ring size (so a 5-membered ring is more frequent than an 8-membered ring, like in normal chemical databases).



Chemistry

Microwave reactions were performed in an Emrys^{TM} Optimizer (Biotage AB). Wattage was automatically adjusted so as to maintain the desired temperature. Column chromatography was performed on Baker Silica Gel (0.063-0.200 mm). For TLC analysis, Schleicher and Schuell F1500/LS 254 silica plates were used. Spots were visualised with ultraviolet light. ^{1}H NMR and ^{13}C NMR were recorded with a Bruker AC 200 spectrometer at room temperature. Tetramethylsilane was used as internal standard; δ in ppm, *J* in Hz. Melting points were determined with a Büchi melting point apparatus and are uncorrected. High Resolution Mass spectroscopy was performed on a PE-Sciex API Qstar instrument. Elemental analyses were within 0.4% of the theoretical values.

1-(3-Cyanobenzyl)-piperidin-2-one (3)

A solution of 2-piperidinone (5 mmol) in THF (25 mL) was stirred for 1 h at 0 °C before 1 eq. of *n*-BuLi (5 mmol, 3.2 mL of a 1.6 M solution in hexane) were added dropwise. After stirring for another hour at 0 °C 1 eq. of 3-(bromomethyl)benzonitrile (5 mmol) was added rapidly. The mixture was allowed to warm slowly to room temperature and stirred overnight. After quenching by adding 15 mL of brine, the solvent layers were separated. To the aqueous layer was added 20 mL of water. After extraction of the water layer with CH_2Cl_2 the combined organic layers were dried (Na_2SO_4), filtered and the solvents evaporated. The product was purified by column chromatography (eluent: CH_2Cl_2/MeOH, 99/1→97.5/2.5 v/v). Yield: 24%. White solid. M.p.: 53-55 °C. Anal. ($C_{13}H_{14}N_2O$) C, H, N.

3-(3-Cyanobenzyl)-piperidin-2-one (4)

To a solution of 2-piperidinone (10 mmol) in THF (15 mL) was added at -78 °C 1 eq. of *n*-BuLi (10 mmol; 6.3 mL of a 1.6 M solution in hexane). After stirring for 15 minutes at -78 °C 1.1 eq. of TMSCl was added and the solution was allowed to warm to room temperature and left to stir for 45 min. The resulting solution was added at -78 °C to a solution of 11 mmol of 1,1,1,3,3,3-hexamethyldisilazane and 11 mmol of *n*-BuLi (6.9 mL of a 1.6 M solution in hexane) in 20 mL of THF. After stirring for 15 min 3-(bromomethyl)benzonitrile (11 mmol) was added and the mixture was allowed to warm slowly to -25°C, before the reaction was quenched by adding an aqueous NH_4Cl (sat.) solution. After extraction with diethyl ether the combined organic layers were washed with a saturated NH_4Cl (aq.) solution, a saturated $NaHCO_3$ (aq.) solution, dried ($MgSO_4$) and the solvents removed by evaporation. The product was purified by

column chromatography (eluent: CH_2Cl_2/MeOH, 99/1→98/2 v/v). Yield: 47%. White crystals. M.p.: 95-96 °C. Anal. ($C_{13}H_{14}N_2O$) C, H, N.

3-(3-Benzylamino)-piperidin-2-one (5)

Compound **4** (2 mmol) was dissolved in methanol, and 2 mmol of concentrated HCl and 100 mg of Pd/C 10% were added. The mixture was hydrogenated at 3 atm for 3 h. After the catalyst was filtered off and the methanol was evaporated the residue was dissolved in water and the pH was adjusted to 4. This solution was washed with ether and the water layer was adjusted with 0.1 M NaOH to pH 12. The free amine was extracted with CH_2Cl_2, dried (Na_2SO_4) and the solvent evaporated. White powder. Yield: 31%. M.p.: 114-116 °C. Anal. ($C_{13}H_{18}N_2O$) C, H, N.

2-(3-Pyridyl)-benzyl alcohol

A suspension of 2-bromobenzyl alcohol (1 mmol), and diethyl(3-pyridyl)borane (1 mmol), Na_2CO_3 (3.8 mmol), TBAB (1 mmol), and $(Ph_3P)_4Pd$ (3%) in 2.5 mL of water was heated in a microwave for 12 min at 150 °C. The product was extracted with ethyl acetate. The combined organic layers were dried ($MgSO_4$), filtered and the solvent was evaporated. The product was purified by flash column chromatography. Eluent column: CH_2Cl_2/MeOH, 99/1→96/4 v/v. Yield: 79%. Oil.

2-(3-Piperidyl)-benzyl alcohol (6)

A mixture of 5.77 mmol of 2-(3-pyridyl)-benzyl alcohol, HCl (5.77 mmol) and PtO_2 (0.38 mmol) in 46 mL of absolute ethanol was placed in a Parr apparatus under H_2 (3 atm) for 3 days. The catalyst was filtered off and the solvent evaporated. After addition of water to the residue the pH was adjusted to 12 and the product was extracted with ethyl acetate. The combined organic layers were dried ($MgSO_4$) and the solvent was evaporated. Recrystallisation from ethyl acetate provided the pure product. Yield: 27%. White needles. M.p.: 135 °C. Anal. ($C_{12}H_{17}NO$) C, H, N.

2-(3-Piperidyl)-benzylic acid (7)

Compound **6** (2 mmol) was dissolved in 50 mL of acetone. Jones' reagent (chromic acid) was added slowly until the orange colour persisted. The pH of the mixture was adjusted to 7 with 1 M NaOH and the product was extracted with ethyl acetate. The combined organic layers were dried ($MgSO_4$), filtered and the solvent was evaporated. The product was purified by preparative HPLC. Anal. ($C_{12}H_{15}NO_2$) C, H, N.

3-(3'-Furyl)-aniline (8)

A suspension of 3-bromofuran (1 mmol), 3-aminophenyl-boranic acid (1 mmol), Na_2CO_3 (3.8 mmol), tetrabutylammonium bromide (1 mmol), and $(Ph_3P)_4Pd$ in 2.5 mL of water was heated in a microwave for 12 min at 150 °C. The product was extracted with ethyl acetate. The combined organic layers were dried ($MgSO_4$), filtered and the solvent was evaporated. The product was purified by flash column chromatography. Eluent: CH_2Cl_2. Yield: 74%. Yellowish solid. M.p.: 73-74 °C. Anal. ($C_{10}H_9NO$) C, H, N.

4-(3'-Furyl)-aniline (9)

A suspension of 3-bromofuran (1 mmol), 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)aniline (1 mmol), Na_2CO_3 (3.8 mmol), tetrabutylammonium bromide (1 mmol), and $(Ph_3P)_4Pd$ in 2.5 mL of water was heated in a microwave for 12 min at 150 °C. The product was extracted with ethyl acetate. The combined organic layers were dried ($MgSO_4$), filtered and the solvent was evaporated. The product was purified by flash column chromatography. Eluent: CH_2Cl_2. Yield: 78%. Yellow solid. M.p.: 92-93 °C. Anal. ($C_{10}H_9NO$) C, H, N.

4-Oxo-4-[4-(3'-furyl)-phenylamino]-butanoic acid (10)

To a solution of 0.63 mmol 4-(3'-furyl)-aniline (**8**) in 10.5 mL of CH_2Cl_2 were added succinic anhydride (0.63 mmol) and 4-methylmorpholine (0.63 mmol). After stirring for 4.5 h the mixture was filtered, the residue washed with CH_2Cl_2 and the filtrate evaporated to dryness. The product was purified by chromatography (eluent: CH_2Cl_2/MeOH, 9/1 v/v). Yield: 27%. Yellow solid. M.p.: 198 °C (dec.). Anal. ($C_{14}H_{13}NO_4 \bullet 0.3CH_3OH$) C, H, N.

Biology

The final compounds (Figure 7.2) were tested at one concentration (10 μM) in duplicate in the Diversity Profile program, including 68 receptors and 15 enzymes, at Cerep (Paris, France).

Software

For the template design we used the Molecule Evoluator software package (Cidrux Pharminformatics, Haarlem, the Netherlands, www.cidrux.com).

Acknowledgements

The authors thank Lisa Chang, Gijs van der Marel and Marc Overhand for their involvement in the template selection process. Gruenenthal GmbH is gratefully acknowledged for their financial contribution, as is Biopartner for support from its First Stage Grant program.

References

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

[2] Dobson, C.M.; Chemical space and biology. *Nature* **2004**, *432*, 824-828.

[3] as taken from http://www.cas.org/chemplus/chemplus1.html and http://www.cas.org/ SCIFINDER/ataglance.html (accessed on April 26, 2006)

[4] Schneider, G.; Fechner, U. Computer-based *de novo* design of drug-like molecules. *Nature Rev. Drug Disc.* **2005**, *4*, 649-663.

[5] Brown, N.; McKay, B.; Gilardoni, F.; Gasteiger, J. A Graph-Based Genetic Algorithm and Its Application to the Multiobjective Evolution of Median Molecules. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1079-1087.

[6] Zhu, J.; Yu, H.; Fan, H.; Liu H.; Shi Y. Design of new selective inhibitors of cyclooxygenase-2 by dynamic assembly of molecular building blocks. *J. Comput. Aid. Mol. Des.* **2001**, *15*, 447-463.

[7] Takagi, H. Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation. In: *Proceedings of the IEEE*, **2001**, *89*, 1275-1296.

[8] Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3-25.

[9] http://www.mdl.com/products/knowledge/crossfire_beilstein/

[10] Cossy, J.; de Filippis, A.; Gomez Pardo, D. Palladium-catalyzed intermolecular α-arylation of N-protected 2-piperidinones. *Org. Lett.* **2003**, *5*, 3037-3039.

[11] Brimble, M.A.; Trzoss, M. A double alkylation-ring closing metathesis approach to spiroimines. *Tetrahedron* **2004**, *60*, 5613-5622.

[12] Miyaura, N.; Suzuki, A. Palladium catalyzed cross-coupling reaction of organoboron compounds. *Chem. Rev.* **1995**, *95*, 2457-2483.

[13] Christopoulos, A.; Kenakin, T. G protein-coupled receptor allosterism and complexing. *Pharmacol. Rev*. **2002**, *54*, 323-374.

[14] Todd, M. Computer-aided organic synthesis. *Chem. Soc. Rev.* **2005**, *34*, 247-266.

[15] Nachbar, R. Molecular Evolution: A Hierarchical Representation for Chemical Topology and its Automated Manipulation. In *Proceedings of the Third Annual Genetic Programming Conference;* Madison, Wisconsin, **1998**, 246-253.

[16] Douguet, D.; Thoreau, E.; Grassy, G. A genetic algorithm for the automated generation of small organic molecules: Drug design using an evolutionary algorithm. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 449-466.

[17] Pegg, S.C.; Haresco, J.; Kuntz, I. A genetic algorithm for structure-based de novo design. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 911-933.

[18] Vinkers, H.; De Jonge, M.; Daeyaert, F.; Heeres, J.; Koymans, L.; Van Lenthe, J.; Lewi, P.; Timmerman, H.; Van Aken, K.; Janssen, P. SYNOPSIS: SYNthesize and OPtimize System in Silico. *J. Med. Chem.* **2003**, *46*, 2765-2773.

[19] Schneider, G.; Lee, M.; Stahl, M.; Schneider, P. De novo design of molecular architectures by evolutionary assembly of drug-derived building blocks. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 487-494.

[20] Takaoka, Y.; Endo, Y.; Yamanobe, S.; Kakinuma, H.; Okubo, T.; Shimazaki, Y.; Ota, T.; Sumiya, S.; Yoshikawa, K. Development of a method for evaluation of drug-likeness and ease of synthesis using a data set in which compounds are assigned scores based on chemists' intuition. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1269-1275.

[21] Lajiness, M.S.; Maggiora, G.M.; Shanmugasundaram, V. Assessment of consistency of medicinal chemists in reviewing sets of compounds. J. Med. Chem. **2004**, *47*, 4891-4896.

[22] Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A "rule of three" for fragment-based lead discovery? *Drug Discovery Today* **2003**, *8*, 876-877.

[23] Oprea, T.I. Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308–1315.

[24] Bemis, G.W.; Murcko, M.A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* **1996**, *39*, 2887-2893.

[25] Bemis, G.W.; Murcko, M.A. Properties of known drugs. 2. Side Chains. *J. Med. Chem.* **1999**, *42*, 5095-5099.

[26] Ertl, P.; Rohde, B.; Selzer, P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. *J. Med. Chem.*, **2000**, *43*, 3714-3717.