Property-Based Optimization of Pharmaceutical Lead Compounds

Sanjivanjit K. Bhal and Kiril Lanevskij

Advanced Chemistry Development, Inc.
Toronto, ON, Canada
www.acdlabs.com

Introduction

While the majority of synthetic medicinal chemists strive to produce compounds in accordance with ADME-Tox and physicochemical property requirements, achieving this while simultaneously striving for maximum potency is a challenging and labor-intensive task. The synthesis of hundreds or thousands of analogs aimed at identifying the optimal drug candidate involves not only synthetic chemistry but also various specialist groups that support synthetic efforts with compound testing and evaluation. Helping to focus synthetic efforts on analogs likely to meet requirements is useful not only to the chemist at the bench but the organization as a whole, resulting in less labor intensive trial-and-error research.

Furthermore, over recent years the pharmaceutical and bio-tech industry has lost experience through lay-offs and retirements of senior scientists. These experienced chemists who would enrich a project through their breadth and depth of knowledge are no longer an asset for their colleagues.

Through ACD/Structure Design Engine,[1] we offer medicinal chemists a software tool that can aid lead optimization efforts, allowing them to:

- Test structure modification ideas prior to the synthesis of new analogs
- Prioritize synthetic modifications based on full and reliable physicochemical and ADME-Tox profiles
- Ensure that project teams do not miss the most promising candidates

In this document we work through a published lead optimization example to demonstrate how the structure design engine can assist in early discovery.

Case Study: Optimizing retro bis-aminopyrrolidine urea derived MCH-R1 antagonists

Research into a new anti-obesity drug led to structure activity relationships (SAR) around a series of bis-aminopyrrolidine urea derived MCH-R1 antagonists to be investigated. In a 2006 publication, Rowbottom *et al.* discuss optimization of compound series (**1**) for potency.[2] Their colleagues Hudson *et al.* discuss optimization of the same series for better brain penetration[3] using compound **2** as a starting point (which had been optimized for potency). Hudson *et al.* synthesized more than 30 analogs in the second round of optimization efforts.

1 **2**
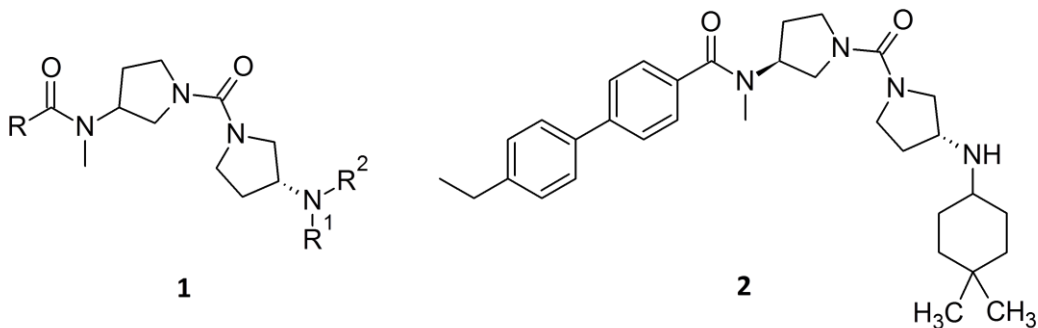
Applying the Structure Design Engine in Lead Optimization
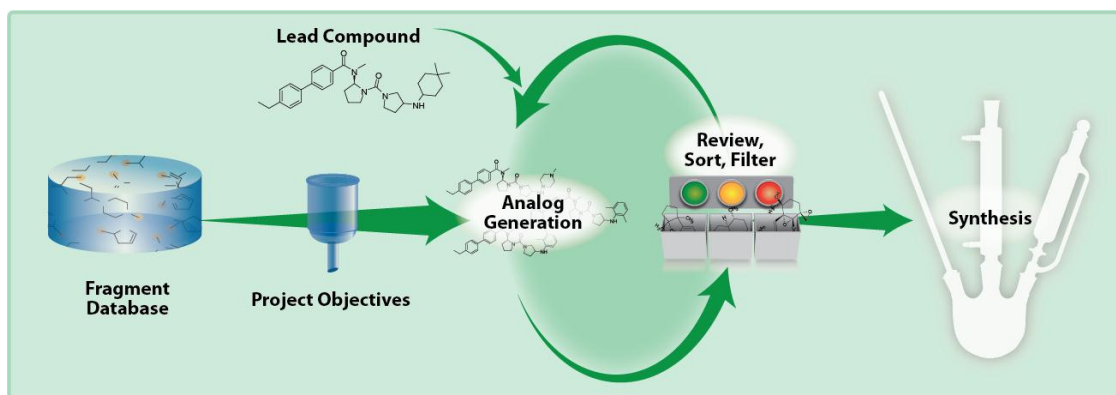
Results from running this example through the software are very encouraging, producing a number of the analogs investigated by the scientists and other suggestions. For our investigation, we will focus on central nervous system (CNS) exposure, and also use compound **2** as our starting point.

First we filter the >16,000 substituent fragment library to a more manageable number by considering only cyclic substituents with one ring. Not only will this help retain the overall shape of the molecule but also restrict increased molecular weight. (While we use the built-in substituent database in this optimization, you can choose to use a proprietary compound, or alternative fragment database.)

The Optimization Strategy
We select the portion of the compound to be optimized and calculate its ADME profile. According to PhysChem-based predictions, compound **2** is too lipophilic and too large for rapid permeation across the blood-brain barrier. It is also extensively plasma protein-bound precluding efficient distribution into the brain (see "CNS non-penetrant" prediction and other properties highlighted red in Fig. 1).

The strategy to improve CNS penetration will therefore be to reduce log*P* (to the 2.0–5.0 optimal range for CNS penetration), restrict molecular weight, and not allow a substantial increase in H-bonding, particularly H-bond donors since these parameters have a profound influence on passive transport of compounds across the blood brain barrier.

Figure 1

Setting up the desired property profile for analogs in the ACD/Structure Design Engine.

Note: Slider bars on the right of the dialog box help the user select appropriate property ranges to achieve the desired profile change.

Indicate the property(ies) of interest (in this case CNS) through star ratings and adjust sliders, striving for green areas for optimal results and avoiding red.

We select the right-hand portion of the molecule (the 4,4-dimethylcyclohexanamine) for optimization. The fragment/substituent database (right-hand side of Fig. 2) can be further refined to select the most appropriate substituents. For example, we can restrict the database to include only substituents of particular ionization state(s), apply the similarity threshold so that only substituents similar to the initial one are listed, as well as adjust the desired ranges of electronic effects of substituents (resonance and inductive) according to the requirements of the project.
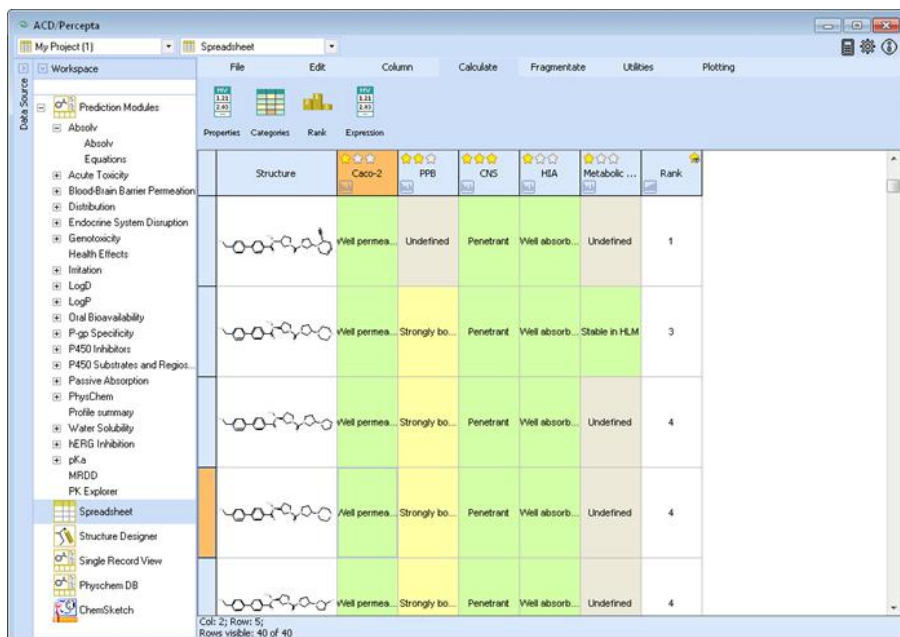
Figure 2

ACD/Structure Design Engine interface—the portion of the molecule to be optimized is selected. The fragment database may also be further refined to focus on the most suitable substituents.
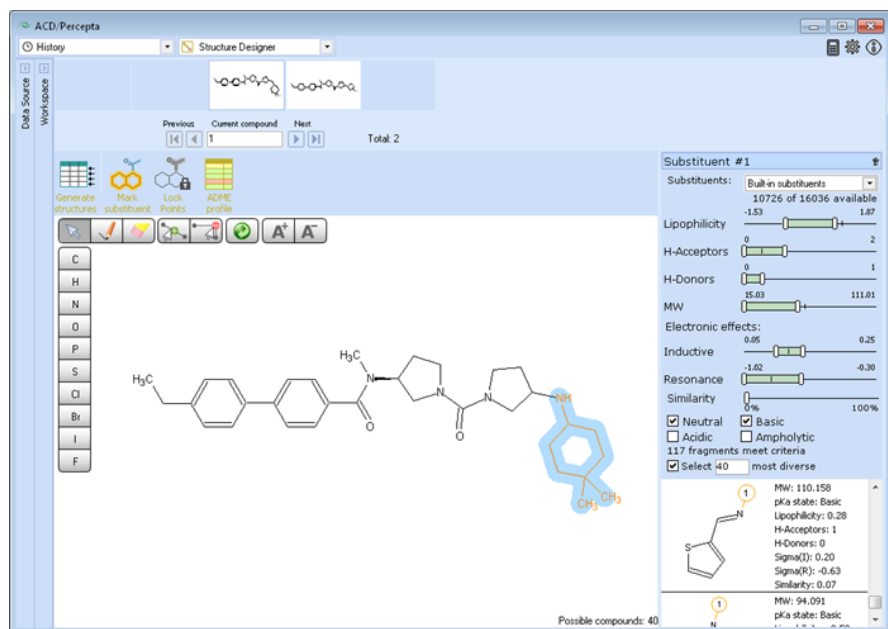
Finally, we select that the software reported 40 of the most diverse analogs so we can examine the available chemical space of manageable proportions. This strategy can also prove useful in a multi-step optimization process when you search for the most favorable region of chemical space in the first round of optimization, and then subject the best analog to further optimization and thoroughly explore the local chemical space within that region.

Results

Figure 3

Results of the optimization can be sorted, filtered, and ranked in spreadsheet view. The full range of physicochemical and ADME-Tox properties can also be calculated to examine full property profiles and further refine results.

Of the 40 analogs generated by the software, several synthesized and tested by the Hudson group are among the results. In fact, compound **3** is their best analog and **4** is very similar to their second best compound (with slight differences in ring connectivity and an alternative left-hand aryl substituent not considered in this study). In our results, however, these compounds only achieve average rank. The morpholine derivative (**5**) also investigated by the group ranks number four, but was disregarded due to a drop in potency.
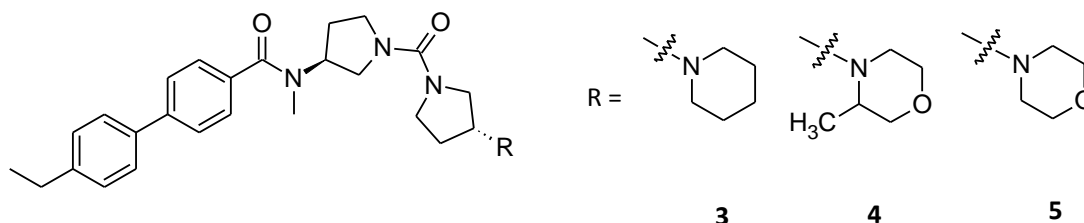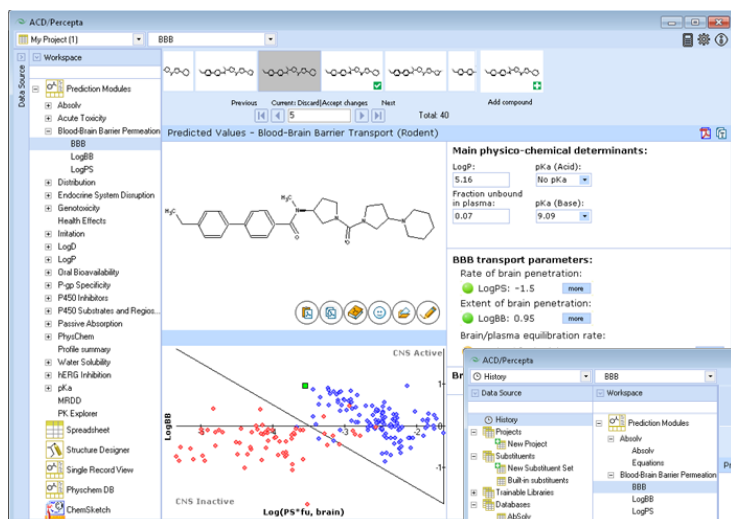
R =

3   **4**   **5**

Table 1: Comparison of predicted and experimental results for analogs suggested by optimization of compound **2** using the Structure Design Engine, and investigated by Hudson *et al.*

Analog		Rank	Exp. logBB	Predicted logBB	CNS Permeation	Predicted Plasma Protein Binding
5		4	—	0.83	Penetrant	Strong (80-90%)
3		12	0.81	0.95	Penetrant	Extensive (> 90%)
4		12	0.53 (for similar compound)	0.23	Weak Penetrant	Extensive (> 90%)

When we look at the results in detail we learn that experimental logBB values and CNS permeation for **3** and **4** agree well with those predicted by the software (see Table 1). The reason that other analogs rank higher than **3** and **4** is that the overall profile of the compounds is better. While compound **3** is sufficiently accessible to CNS, it is extensively bound to protein plasma leaving lower free drug concentration. Compound **5** is predicted to exhibit weaker plasma protein binding, leaving a larger free drug concentration for action at the target site.

Access to CNS—Looking at the Bigger Picture
While the Hudson group only looked at logBB when considering CNS access, the blood brain barrier (BBB) penetration module of the ACD/Percepta[4] platform also takes into account the rate of equilibration between the brain and plasma. The BBB module gives more detailed information that can help the chemist decide on the best candidates to synthesize. As demonstrated in Fig. 4, compound **3** has high logBB, but equilibrates between brain and plasma at an average rate. It is, therefore, likely that the initial compound (**2**) would be poorly accessible to CNS not due to low logBB, but due to a low brain/plasma equilibration rate.

Figure 4

BBB Module of ACD/Percepta
As well as providing rate and extent of brain penetration, the predictor also gives plasma/brain equilibration rates and plots the predicted compound vs. known CNS active and inactive drugs.

See the predicted compounds (plotted green) in the CNS activity graphs.

Top: Compound **3** is predicted to have sufficient brain penetration for CNS activity.

Bottom: Compound **2** is likely to be CNS inactive due to a poor brain/plasma equilibration rate.
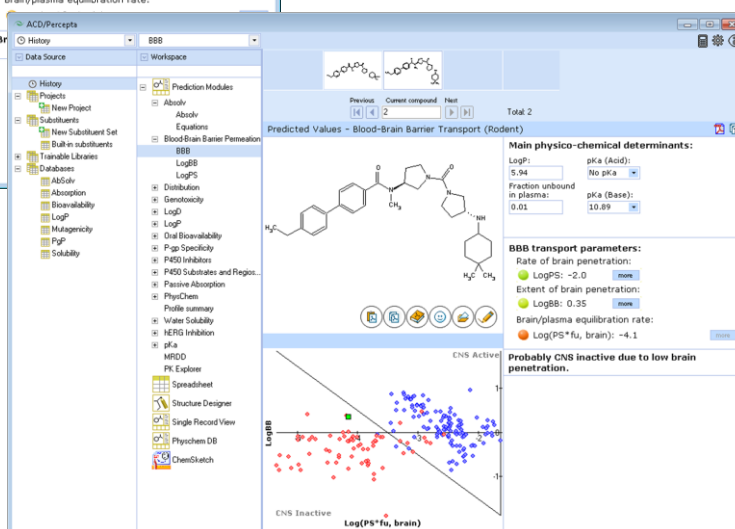
Metabolic Stability

A major problem identified with the optimized series was susceptibility of these analogs to biotransformation by human liver microsomal (HLM) enzymes, resulting in poor metabolic stability. The HLM Regioselectivity module in Percepta correctly predicts the metabolic soft spot that was introduced into this series. With early prediction, therefore, such problems could be avoided and the project could be shifted to more metabolically stable compounds earlier, saving time and effort.
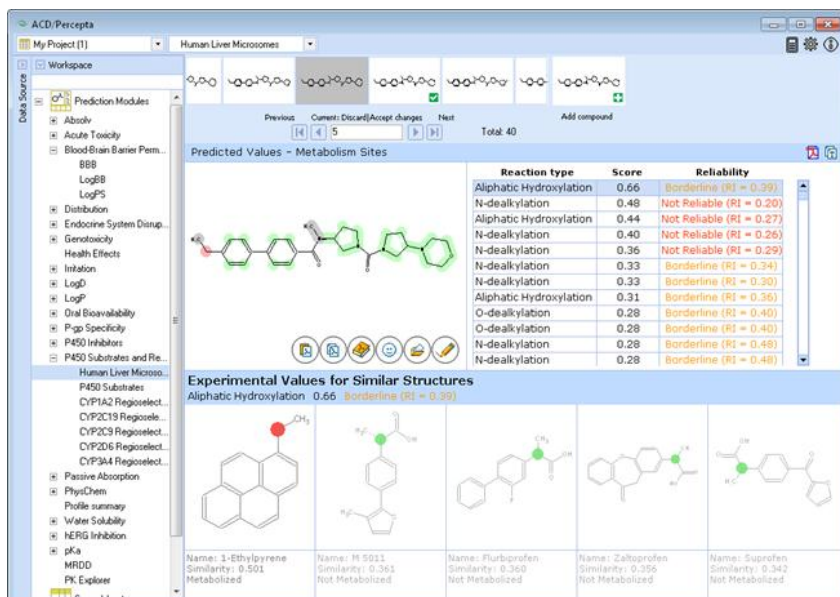

Figure 5

The HLM Regioselectivity module in ACD/Percepta clearly identifies the metabolic soft spot in compound **2**.

Potency

Since increasing or retaining potency is a major factor in any optimization project, a favorable ADME profile must be balanced with good potency. Compound **5** is a good example of how a molecule with a good ADME profile will not always be the best candidate. MCH-R1 binding of this morpholine derivative was found to be relatively poor compared with the piperidine derivative (**3**).

While the structure design engine does not predict potency, where available, data can be imported into the software to help rank compounds.

Conclusion

While physicochemical and ADME data can be invaluable in the lead optimization process, this information is often elusive and difficult to apply in the practice of directing structure optimization. In this paper we demonstrated how the ACD/Structure Design Engine can aid the lead optimization process by helping the chemist to apply physicochemical and ADME knowledge and focus synthetic efforts on compounds that meet project requirements. The software helps chemists brainstorm chemical transformations based on desired physicochemical and ADME properties to quickly produce a concise list of analogs with the required enhancements. Consequently, chemists no longer have to rely solely on their understanding and knowledge of chemical classes, or spend hours consulting colleagues and literature to make the critical decision of how to best modify their lead.

Furthermore, the software can be customized by plugging-in a variety of prediction modules that provide detailed information about the behavior a compound is likely to exhibit in the human body. Early knowledge can help shorten the discovery timeframe and put knowledge in the hands of those at the forefront of the research project.

References

1. ACD/Structure Design Engine www.acdlabs.com/sde, Advanced Chemistry Development, Inc., Toronto, Canada, December 20, 2011.
2. Martin W. Rowbottom, Troy D. Vickers, Brian Dyck, Jonathan Grey, Junko Tamiya, Mingzhu Zhang, Mehrak Kiankarimi, Dongpei Wu, Wesley Dwight, Warren S. Wade, David Schwarz, Christopher E. Heise, Ajay Madan, Andrew Fisher, Robert Petroski and Val S. Goodfellow, *Bioorg. Med. Chem. Lett.*, **2006**, 16: 4450–4457.
3. Sarah Hudson, Mehrak Kiankarimi, Martin W. Rowbottom, Troy D. Vickers, Dongpei Wu, Joseph Pontillo, Brett Ching, Wesley Dwight, Val S. Goodfellow, David Schwarz, Christopher E. Heise, Ajay Madan, Jenny Wen, William Ban, Hua Wang and Warren S. Wade, *Bioorg. Med. Chem. Lett.*, **2006**, 16: 4922–4930.
4. ACD/Percepta www.acdlabs.com/percepta, Advanced Chemistry Development, Inc., Toronto, Canada, December 20, 2011.