Reducing the Synthetic Burdens of Lead Structure Optimization: A Novel Software-Aided Approach

Karim Kassam, Ryan Sasaki, and Michel R.J. Hachey
Advanced Chemistry Development, Inc. (ACD/Labs), Toronto, Ontario, Canada

Introduction

The in-vivo biological activity of a drug is contingent on its ability to get to the intended site of action, commonly through a transport process that involves barrier penetration and crossing yet is hindered by various forms of binding that can limit efficacy. Plasma protein binding, tissue partitioning, and blood brain barrier permeability are all influenced by the physicochemical properties of a drug (log*P*, log*D*, pK_a, aqueous solubility, etc.). After activity screening it is often beneficial to optimize the properties of selected lead compounds to improve their deliverability by making small structural changes. A major caveat and challenge to this process is that only modifications that optimize the desired properties without introducing negative effects on the lead's activity or toxicity profiles are acceptable. Trial-and-error like efforts are inefficient and costly, and there is a need for a systematic approach to guide the structural optimization process. By combining physicochemical property predictors and a critically evaluated database of biologically-acceptable substituents, the medicinal chemist can reduce the number of analogs that need to be synthesized to achieve optimal drug-like properties. A software-assisted approach will be outlined here using acetyl sulfadiazine as an example.

Description of the Software

ACD/Structure Design Suite builds on the industry standard predictions of molecular physical properties to offer a structure design studio that allows the rapid identification and assessment of chemically relevant modifications in terms of their ability to aid in the optimization of molecular physical properties. The technology within this suite of ACD/Labs software products assists by:

- Suggesting substituents to modify log*P*, log*D*, solubility, or single pK_a values in a desired direction.
- Minimizing possible negative effects by ensuring that suggested substituent changes have been previously evaluated in other successful clinical settings.

Method and Theory

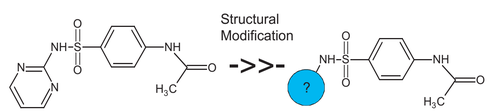
The detailed theoretical underpinning of this software-guided approach is given elsewhere and will only be summarized here (1–4). The software exploits empirical trends that are readily observed between structure, ionization, or lipophilicity, and their correlation with desired endpoints such as solubility, permeability, blood brain barrier penetration, etc. Varying physicochemical properties affect the endpoint. For example, solubility can be increased within medicinal chemistry series by increasing ionization and reducing lipophilicity. While it is advantageous to understand how pK_a, log*P*, and log*D* affect solubility, and how to assess and identify creative substituents for replacement, this knowledge can be readily incorporated in expert structure design algorithms that can exploit highly accurate prediction algorithms for pK_a, log*P*, log*D*, and solubility. These expert algorithms allow medicinal chemists to focus their planning mainly on the desired end point and reaction feasibility without requiring in-depth knowledge of the intricacies of physicochemical trends.

Challenges Associated with Structural Optimization

After having identified a promising compound, there are many criteria that can be used to guide further structural optimization. For example, medicinal chemists can plan their chemical modifications in a way to increase or decrease specific physicochemical parameters toward more drug-like values for one or all of the following properties:

- Solubility
- Log*D*
- Log*P*
- pK_a

Let's assume the acetyl sulfadiazine is our lead candidate and that its solubility is too low. Figure 1.a shows the structure and its calculated properties at pH 5.5—the prevalent pH for orally absorbable drugs in the GI tract.



Lead Candidate Calculated Properties	Next Analog: Substituent Group Modification Aiming to Achieve Optimal Properties?
Solubility at pH 5.5 = 1.7×10^{-3} mol/L	Solubility at pH 5.5 = …
Log*D* at pH 5.5 = 1.7×10^{-3} mol/L	Log*D* at pH 5.5 = …
Log*P* = 0.34	Log*P* = …
pK_a = 14.51, 6.30, -0.03	pK_a = …

Figure 1a Selected acetyl sulfadiazine lead compound to be optimized by further structural modification.

As seen in Figure 1b, it is rather trivial to use current software to calculate key physicochemical properties for a lead compound or multiple analogs. Indeed batch mode calculations readily allow one to explore behaviors of large series of compounds. However, with an infinite number of possible changes, it is often not practical nor truly viable to blindly test any and all possible modifications. Knowledge of physicochemical trends and experience with various substituent suggestions is still very much required to quickly arrive at a compound with optimized physicochemical properties. While experts in the field might be able to narrow down the viable modifications based on their own experience, a systematic application of structure-property relationships is often beneficial for choosing the best possible analogs to synthesize first. ACD/Structure Design Suite is built with the end point in mind: its purpose is to narrow the number of analogs that need to be considered by leveraging knowledge and existing experience in a systematic and synergistic computer-assisted fashion.

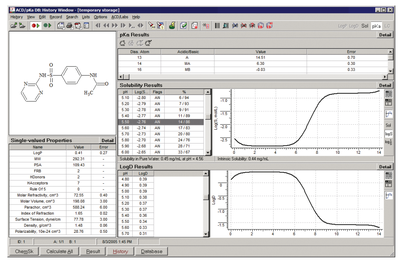


Figure 1b Calculated physicochemical properties for acetyl sulfadiazine.

Details of Substituent Database

A key part of ACD/Structure Design Suite is its substituent database, which captures some of the accumulated experience on 10,000 substituents used in the pharmaceutical industry and other fields. For each chemical fragment, the database includes the pK_a value, molar volume, molar refractivity, Hansch constant (π), and a variety of Hammett sigma (σ) electronic substituent constants. These properties aid in rationalizing experimental trends and correlating to new properties.

To minimize the chance of introducing potentially toxic fragments into the design of a drug, database substituent searches can be limited to only those substituents previously used in clinical evaluations. The 2000 drug-like fragments included in the database provide an indirect indication of potentially lower toxicity risk.

Computer Assisted Structure Design: A Simple Workflow

Once the lead compound to be optimized is drawn or imported into the ACD/Structure Designer module, a wizard interface guides the user in defining the appropriate parameters for optimization and the necessary end point conditions.

1) Select Property for Optimization: The program first asks whether an overall increase or decrease in the aqueous solubility, log*D*, log*P*, or pK_a properties values is desired. For this example, an increase in solubility will be chosen.

2) Choose the Site of Modification and pH: As shown in Figure 2, the medicinal chemists are asked to lasso the site they wish to modify in order to modulate the physicochemical properties selected. Care must be exercised to avoid selecting regions that could be part of the pharmacophore since changes to this environment could adversely affect binding to a receptor site. The pH value is set to 5.5, which is a reasonable value for the GI tract.

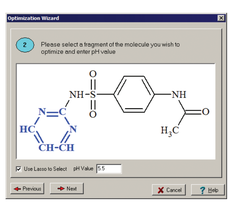


Figure 2 Selection of a fragment for modification in acetyl sulfadiazine.

3) Choose Optimization Parameters: Since solubility is a function of both ionization (pK_a) and lipophilicity (log*P*), the user is given a choice to optimize according to pK_a, log*P*, or both. Optimization with respect to both was selected in this case.

4) Refine Substituent Choice: In most cases, medicinal chemists will want substituents that preserve the acid/base characteristics of the original compound. The substituent selection available from the database of ACD/Structure Design Suite, or added by the user, can therefore be narrowed to match the general class of substituent being replaced. The choices are acidic, basic, zwitterionic, ampholytic, or neutral. Since the diazine ring on the lead compound is neutral at pH 5.5, the neutral option is selected for this example.

5) Search for Suitable Modifications: The series of steps taken so far have resulted in the automatic creation of a query according to the choices that were made. A search at this point would yield 93 hits, with various substituents identified as potential replacements. However, remembering that keeping similar overall topology for the parent is important, the search query can be narrowed by adding a constraint that the substituent must contain a six-membered ring as present in its parent. The new query is shown in Figure 3, and now yields 35 hits.

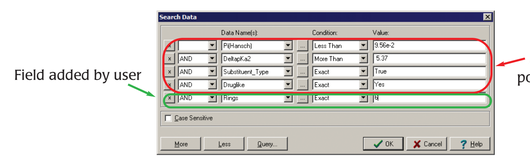


Figure 3 Automated query creation modified to narrow the search to six-membered ring substituents.

6) Refining the Hit List: At this point, it is suggested that hits be ranked according to similarity to the original through a structural similarity search. In our case, a similarity search on the original diazine ring allows five substituents to be identified that are very similar in structural nature as shown in Figure 4. Using the check box (top right hand corner of each structure pane) allows interactive cherry picking of four fragments that look most promising for final evaluation.

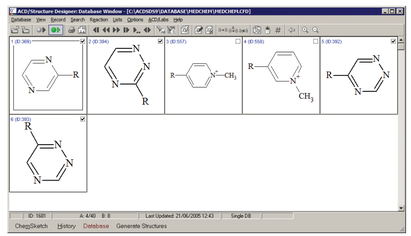


Figure 4 Selected drug-like substituents that show desired solubility increase trends while retaining high structural similarity to the parent.

7) The Final Evaluation: The calculated solubility values for each of the suggested forms can be reviewed in a convenient graph format, shown in Figure 5. The parent molecule is shown first in the lower left corner, and provides a reference point with respect to the analogs.

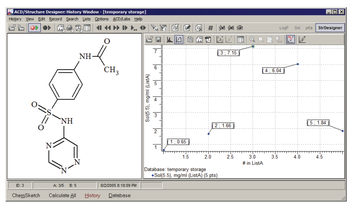


Figure 5 Solubility graph for selected molecules.

Clicking data points on the graphs automatically shows the relevant structure associated with a data point. Figure 6 shows that a small modification to the diazine ring with respect to the parent (highlighted in a red box) yielded a desired improvement in solubility. Since the modifications are relatively minor and do not include the hypothesized pharmacophore, it is reasonable that binding of the drug to the target enzyme will not be compromised.

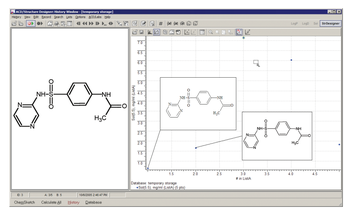


Figure 6 Easy comparison of analogs with the parent molecule.

Other properties, such as polar surface area, freely rotatable bonds, log*D*, and rule of 5 filters can all be plotted to allow desired compounds to be selected in an intuitive and visual manner. This allows users to very quickly reduce the number of analogues needed to be synthesized, thereby saving an organization significant resources, materials and, overall, offering more opportunities to populate the pipeline for a pharmaceutical company.

Conclusion

ACD/Structure Design Suite permits a chemist to rapidly determine a series of possible alterations to a lead compound in order to improve the physicochemical properties of the molecule, whether it is improving log*P*, log*D*, pK_a, or solubility. A user friendly interface offers a means of expediting workflow to help medicinal chemists increase their efficiency and bring products to the market faster.

Acknowledgments

Ed Kolovanov, Fedor Gorohov, Daria Jouravleva, and Antony Williams.

References

1. R. DeWitte, F. Gorohov, and E. Kolovanov, *Property-Based Design Methodology I: Parameters Influencing Solubility*, ADMET-1 Conference, San Diego, CA, USA (2004).
2. R. DeWitte, F. Gorohov, and E. Kolovanov, *Property-Based Design Methodology II: Assessing Elementary Substituents*, ADMET-1 Conference, San Diego, CA, USA (2004).
3. R. DeWitte, F. Gorohov, and E. Kolovanov, *Property-Based Design Methodology III: Identifying Creative Substituents*, ADMET-1 Conference, San Diego, CA, USA (2004).
4. R. DeWitte, F. Gorohov, and E. Kolovanov, *Property-Based Design Methodology IV: Heterocycle Replacement*, ADMET-1 Conference, San Diego, CA, USA (2004).

Toll Free: 1-800-304-3988
Email: info@acdlabs.com
www.acdlabs.com