JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

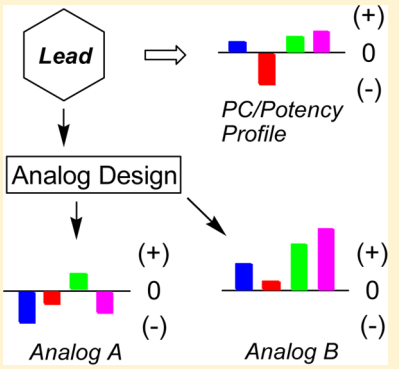
A Prospective Method To Guide Small Molecule Drug Design

Alan T. Johnson*

Department of Chemistry, Hawaii Biotech Inc., Aiea, Hawaii 96701, United States

Supporting Information

ABSTRACT: At present, small molecule drug design follows a retrospective path when considering what analogs are to be made around a current hit or lead molecule with the focus often on identifying a compound with higher intrinsic potency. What this approach overlooks is the simultaneous need to also improve the physicochemical (PC) and pharmacokinetic (PK) properties of these compounds, and illustrates the multivariate problem the chemist must face when targeting new analogs for synthesis. To address this problem, a simple method is presented which allows the chemist to integrate PC properties into small molecule drug design in a prospective manner, prioritize new target molecules for synthesis, and potentially shorten the path to the clinic. This simple method also provides a tool for the student of medicinal chemistry to see how changes in PC properties and intrinsic potency can influence drug-like properties of small molecules during the drug discovery process.



KEYWORDS: *Graduate Education/Research, Organic Chemistry, Inquiry-Based/Discovery Learning, Medicinal Chemistry, Drugs/Pharmaceuticals, Physical Properties, Synthesis*

INTRODUCTION

One question often asked by the medicinal chemist engaged in small molecule drug discovery is what compound shall I make next? The answer is usually based on the desire to achieve greater intrinsic potency, filling a gap in a Structure Activity Relationship (SAR) table, or what can be easily made using the established synthetic route for the current lead series. Although this approach has led to the discovery of many FDA approved drugs, an over dependence on intrinsic potency for compound selection during the drug discovery process has not improved the success rate for compounds entering clinical trials.[1] Very often many analogs are made which possess high intrinsic potency but prove to be unsuitable for pharmacokinetic (PK), early *in vivo* efficacy, and toxicity studies due to poor drug-like properties.

In recent years, chemists have asked the question; what PC properties are found in good hits, leads, and oral drugs?[2−8] Early studies led to the groundbreaking Lipinski's rule of five[9] to guide analog synthesis and the rule of three[10] for selection hit compounds. More recently, others have provided detailed analyses of structural features and their relation to drug-like properties which have led to additional guidelines for the medicinal chemist to consider when involved in analog synthesis.[11−13] Although this has helped to some degree in guiding chemists with new analog design in the hit to lead and lead optimization stages of a project, it still has not made the process more efficient or decrease the burden of making large numbers of analogs in the pursuit of preclinical candidates that have a high probability of success in the clinic.

What is needed is a change in the way a chemist thinks about small molecule drug design which can increase the probability of identifying a compound with suitable drug-like properties, decreases the number of analogs to be made, and can be easily applied by the practicing medicinal chemist on a day to day basis—a more prospective approach. To address this problem, a method is presented which provides anyone with access to simple computational tools to evaluate molecules for their potential to be good drugs prior to actual synthesis and testing. In addition, this method can be used to introduce students in chemistry to the concepts involved with small molecule drug design by having them follow the progress from hit to clinical candidate of successful drug discovery projects published in medicinal chemistry journals. To assist readers unfamiliar with the subject, a glossary of terms and phases commonly used in medicinal chemistry is provided in the Supporting Information (SI).

AN OUTLINE OF THE PROBLEM

To illustrate the problem, consider **1** as a hit or early lead compound (Figure 1). The good intrinsic potency (IC_{50} value) and modular nature of the core structure makes it an attractive starting point for analog synthesis. In an effort to improve the intrinsic potency and PC properties of this compound, the chemist has several avenues of analog synthesis to pursue. To improve potency, a common practice is to make changes to the perimeter of the core structure[14] altering the type, number, and position of standard R-groups (R = alkyl, halogen, O-, N-, or S-alkyl). To improve PC properties, minor changes to the core structure (Figure 1) could be made such as replacement of one or both of the phenyl rings in **1** with heteroaryl ring systems. Exploring major changes to the core structure has the potential to improve both potency and PC properties, but is often

dx.doi.org/10.1021/ed5002653 | J. Chem. Educ. XXXX, XXX, XXX−XXX

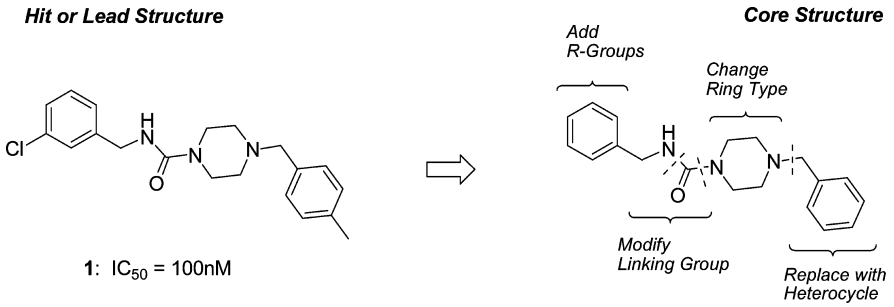


Figure 1. Analysis of a hit or early lead compound as a starting point for analog synthesis.

Table 1. Selected Molecular Parameters, Their Relation to PC, PK, and Pharmacodynamic (PD) Properties, and the Target Value Range Taken from the Literature and Derived from a Set of 89 FDA Approved Small Molecule Drugs

Molecular Parameter[a]	PC/PK/PD	Target Value/Range	New Drugs Median Values (mean)[f]	References
BEI = $\frac{\mathrm{pIC_{50}}\left(\text{or } \mathrm{p}K_i\right)}{\mathrm{MW\ (kDa)}}$	solubility, permeability	20.9 to 23.6[b]	20.7 (22.1)	23, 18
FPA = $\frac{\sum(\mathrm{N}+\mathrm{O})}{\text{total \# HAs}}$	permeability, oral absorption, toxicity	0.21[c]	0.21 (0.22)	8, 29 , 30
LLE = $\mathrm{pIC_{50}}\left(\text{or } \mathrm{p}K_i\right) - \mathrm{CLogP\ (or\ LogD)}$	solubility, permeability, absorption, PPB[g], metabolism, toxicity	5 to 7[d]	5.4 (5.7)	11, 17
$\mathrm{Fsp^3}$ = $\frac{\sum \mathrm{sp^3\ C\ atoms}}{\text{total \# C atoms}}$	solubility, permeability, PPB[g]	>0.31[e]	0.38 (0.41)	8, 25, 26

[a] $\mathrm{pIC_{50}} = -\log(\mathrm{IC_{50}})$ and $\mathrm{p}K_i = -\log(K_i)$ where the units for each constant ($\mathrm{IC_{50}}$ or K_i) are molarity; HA = Heavy Atom. [b] Derived using $\mathrm{XC_{50}}$ values of 10 and 1 nM and the mean MW = 382 for oral drugs from 1993 to 2002 (ref 8, Table 4). [c] Equivalent to the fPSA given by Meanwell (ref 8, Table 3); [d] Range provided by Leeson et al. (ref 11). [e] $\mathrm{Fsp^3}$ = 0.31 correlates with logS = −6 (ref 8, Table 20); [f] See analysis of FDA approved drugs from 2005 to 2011 in Supporting Information, [g] PPB = plasma protein binding.

considered to carry higher risk and therefore is usually not the first path pursued in the early stages of a drug discovery project. In the case of **1**, converting the urea to an amide or the piperazine ring to a phenyl or heteroaryl ring are just some of the many possibilities which can be envisioned. Given all of the possible changes which can be made, where is the best place to begin? One aspect of analog design always considered by medicinal chemists when prioritizing targets is the ease of synthesis and the commercial availability of the required reagents. In the current example, there are three obvious disconnections leading to very common and readily available reagents. But is this really the best place to start analog synthesis?

An additional problem inherent to this process is that experimental data for a specific compound needed to help guide new analog design becomes available only *after* the compound is made. Further, this information may come from many sources, such as *in vitro* Absorption, Distribution, Metabolism, and Excretion (ADME), PK, or early efficacy studies, and often cannot be easily translated by the chemist into simple changes in molecular structure when considering the design of the next analog in a compound series. When compared to improving the intrinsic potency of a compound *in vitro*, designing a drug that can reach and maintain an effective concentration at the target site *in vivo* is a significantly more complex problem for the practicing medicinal chemist to address, especially on a daily basis. As a result, it is lead optimization chemistry which lies at the heart of the small molecule drug discovery process.[15]

PC PROPERTIES AND SMALL MOLECULE DRUGS

To further complicate matters, each drug target is related to a specific disease. As a result, a unique set of selection criteria leading to the identification of a drug candidate will emerge as the project moves forward and the screening cascade becomes more tailored to the disease indication and the specific target for intervention. This will also define the drug-like properties for a successful clinical candidate against this disease. In essence, it is this selection process which forces a change in the PC properties to new analogs as they move down the screening cascade, and not necessarily the foresight of the chemist. Since it is not currently possible for a chemist to design specific molecular properties into an analog molecule which will ensure an improved PK, safety profile, or better *in vivo* efficacy relative to the lead, the most a chemist can hope to do is avoid making analogs which have no chance of succeeding in the clinic due to obvious metabolic liabilities or poor PC properties. Indeed, several authors have recently stated the need for medicinal chemists to find a balance between good PC properties and potency in new ligand design.[16−18] But how can a chemist integrate PC properties into new ligand design to make analogs which are more likely to demonstrate improved drug-like properties relative to the lead molecule?

dx.doi.org/10.1021/ed5002653 | J. Chem. Educ. XXXX, XXX, XXX−XXX

MOLECULAR PARAMETERS AND SMALL MOLECULE DRUG DESIGN

Fortunately, many PC properties can be accurately calculated from the 2D-structure of a molecule. These include molecular weight, ClogP, polar surface area (PSA), the Fraction of sp^3 hybridized atoms (Fsp^3), hydrogen bond donor number (HBD), hydrogen bond acceptor number (HBA), pK_a, and the number of rotatable bonds, all of which have been related to favorable drug-like properties of molecules which have made it to market.[8,12,19] Several molecular parameters (metrics) based upon PC properties have been introduced which incorporate the key aspects of these design variables into simple to understand target values.[20,21] Use of these parameters has led to a number of methods being devised in an attempt to help medicinal chemists make better decisions during the drug discovery process.[22] Although each has merit, all rely on using data generated *after* the new target compound has been synthesized. For the discussion which follows, four of several possible molecular parameters[8,20] have been selected (Table 1) to illustrate the concept behind a simple method which integrates PC properties into small molecule drug design.

Two of the most often used molecular parameters are the binding efficiency index (BEI)[23] and the lipophilic ligand efficiency (LLE).[11] Both of these parameters have been associated with successful drug discovery projects[24] and are of interest to the chemist since they combine intrinsic potency, the most common parameter driving new ligand design, with a PC property (MW and logP) to give a single molecular parameter.

A third molecular parameter which has been associated with good drug-like properties is the Fraction of sp^3-hybridized carbon atoms (Fsp^3) in a molecule.[25,26] As pointed out in these early articles and more recent studies,[8,27] aqueous solubility increases dramatically for molecules possessing values of $Fsp^3 > 0.30$, a concept which is also consistent with the influence of the number of aromatic rings on drug-like properties.[28]

The final molecular parameter of interest for this discussion is related to the PSA of a molecule, a parameter which has been correlated with various properties such as permeability,[29] bioavailability in rats,[19,30] and potential for toxicity.[31] While the importance of this PC property is well accepted, how to use the PSA of a molecule to guide new drug design is considerably more complicated.[32] Recently, it was noted that the mean fraction of PSA (fPSA = PSA/total SA) for all orally available drugs reaching the market up to 2002 was 0.21,[8] suggesting that this number may be used as a target value for new analog design. Although the calculation of fPSA is not difficult for chemists having access to the appropriate software, the fraction of polar atoms (FPA), defined simply as the total number of nitrogen and oxygen atoms divided by the total number of heavy atoms (Table 1), can be obtained knowing simply the molecular formula of the molecule. Given this parameter shows a very strong positive correlation ($r = 0.910$) to the fPSA (see Figure SI-12), it too can be considered a parameter suitable to represent changes of PSA in a molecule.

TARGET VALUES FOR SELECTED MOLECULAR PARAMETERS

A prerequisite for the medicinal chemist involved in new ligand design is knowledge of what the desired drug-like properties of the target molecules should be. As mentioned above, these may be target specific and often emerge during the optimization process from the assay data obtained following a customized screening cascade for the project. Of those usually considered, only PC properties can be estimated simply from the structure of the target molecule with any level of certainty, and therefore they are the only design elements which can be controlled by the chemist prior to actual synthesis and testing of a compound. To assess whether a proposed structural modification to a lead compound is worthy of becoming a target for synthesis, target values for the molecular parameters being used to guide new ligand design are required. For the purpose of this discussion, the target values or the range of values for each of the four molecular parameters of interest were identified from literature reports and are given in Table 1. In addition, median values derived from recent (2005 to 2011) FDA approved drugs (Tables SI-25 and SI-26) are also shown in Table 1 which validate the literature target values/range and may serve as a secondary guide to the chemist involved in new ligand design as discussed in the example below. Since the target values (range and median) for these molecular parameters were derived from the properties of FDA approved drugs, molecules designed to meet these criteria are more likely to possess good drug-like properties relative to analogs with molecular parameters further away from these values.

PC/POTENCY PROFILES

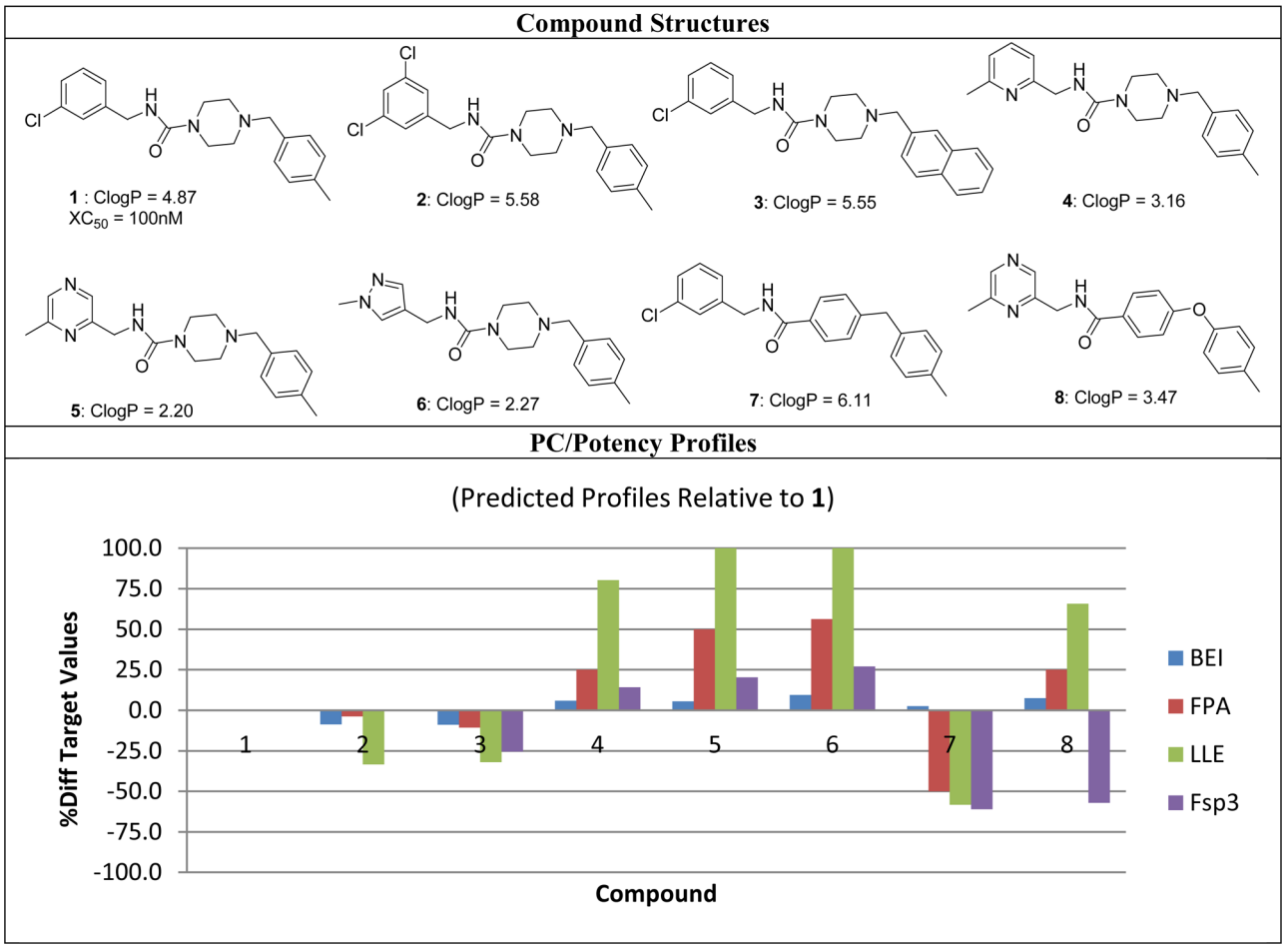
To illustrate how all four of the molecular parameters selected (Table 1) can be used simultaneously by the chemist to guide new ligand design, the changes in these parameters resulting from structural modifications to the lead compound compared to each target compound will be examined by calculating their PC/potency profiles (see SI). These profiles are the graphical display of the percent difference values for each molecular parameter of interest relative to those of the lead or desired target values for each compound targeted by the chemist for synthesis. Use of percent difference values allows the chemist to instantly see if the changes represented in the 2-D structure of their proposed analogs are an improvement relative to the lead molecule, and the actual values resulting from this change provides a quantitative measure which can be used to further refine their analysis. In addition, it will be shown that changes in the intrinsic potency of targeted analogs which are usually considered to be significant (>10-fold), are not always sufficient to alter the priority for synthesis as reflected in changes to their PC/potency profiles, a result which deemphasizes the role of intrinsic potency as the primary guide for new ligand design.[18] This finding also suggests that the most efficient approach to identifying a preclinical candidate should focus more on optimizing the drug-like properties of the lead[24] rather than the intrinsic potency early in the project as opposed to waiting to the later stages of the discovery process.

A PROSPECTIVE METHOD FOR SMALL MOLECULE DRUG DESIGN

The example presented below introduces a simple method to help guide the chemist to be more efficient in the design of new target molecules by using PC/potency profiles both as a guide for analog synthesis, and to suggest when changes in the core structure may be appropriate. This approach is easy to use (see SI) and allows the chemist to prioritize analogs for synthesis with the information available from a simple 2-D structure of the target molecule.

dx.doi.org/10.1021/ed5002653 | J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 2. Structures of 1 and Possible Target Compounds, Their PC/Potency Profiles, and PC Data



	Compound Data[a]						% Difference from Lead Values[b]			
Compd	MW	XC_{50}	BEI	FPA	LLE	Fsp^3	BEI	FPA	LLE	Fsp^3
1	357.88	100.00	19.6	0.16	2.1	0.35	0.0	0.0	0.0	0.0
2	392.32	100.00	17.9	0.15	1.4	0.35	−8.7	−3.8	−33.3	0.0
3	393.91	100.00	17.8	0.14	1.5	0.26	−8.9	−12.5	−31.9	−25.7
4	338.45	100.00	20.7	0.20	3.8	0.40	5.9	25.0	80.3	14.3
5	339.43	100.00	20.6	0.24	4.8	0.42	5.6	50.0	125.4	20.3
6	327.42	100.00	21.4	0.25	4.7	0.44	9.5	56.3	122.1	27.0
7	349.85	100.00	20.1	0.08	0.9	0.14	2.6	−50.0	−58.2	−61.0
8	333.38	100.00	21.0	0.20	3.5	0.15	7.5	25.0	65.7	−57.1

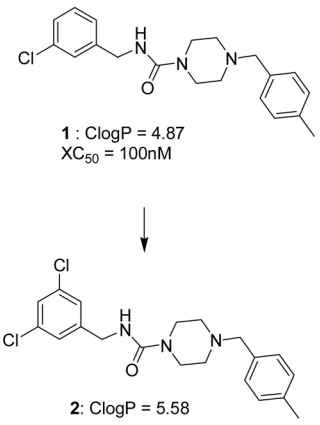
[a] XC_{50} = IC_{50}, EC_{50}, or K_i value (nM); BEI = pXC_{50}/MW (kDaltons) FPA = Σ N + O/total # HA; LLE = pXC_{50} - ClogP; Fsp^3 = Σ sp^3 C atoms/total # C atoms. [b] determined using the PC/potency spreadsheet PC_potency_templ.xlsx.

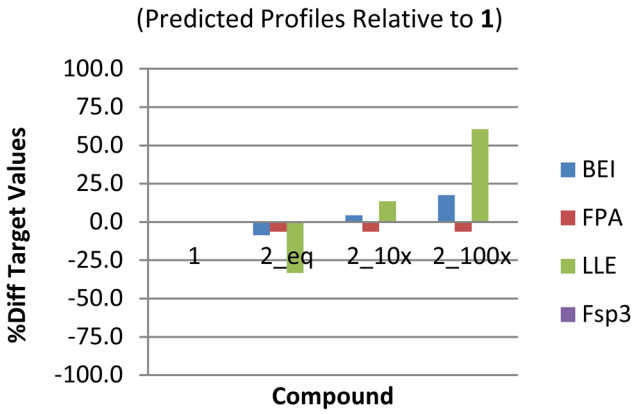
Assume compound **1** (Figure 1 and Table 2) is a newly discovered lead and displays an intrinsic potency (XC_{50})[33] of 100 nM against the target of interest, and the remaining molecules **2** to **8** are potential targets being considered by the chemist for synthesis as analogs of **1** (Table 2). Analysis of the molecular parameters for **1** indicates this compound has a good BEI value of 19.6, however, the LLE of 2.1 is well below the desired target value/range of 5−7 given in Table 1. Given the XC_{50} = 100 nM, a common first strategy would be to modify the perimeter of the core structure present in **1** to improve the intrinsic potency. Compounds **2** and **3** represent simple modifications of **1** that could be prepared to test this approach. Alternatively, if the initial goal is to improve PC properties, target molecules **4**, **5**, and **6** are examples of exploring minor changes to the core structure of the lead (**1**) by replacement of the phenyl ring of the *N*-benzyl group with various heteraromatic rings. Finally, the remaining two compounds **7** and **8**, represent a major change in the core structure from the lead molecule **1** where the urea group has been replaced by an amide, and the piperazine ring with a phenyl ring.

We can now return to the first question posed in this discussion; which of these molecules should the chemist target first for synthesis? Assuming the goal is to identify analogs of **1** with improved drug-like properties (molecular parameter values closer to those given in Table 1), we can begin by comparing the PC/potency profiles of the proposed target molecules relative to **1** using their calculated BEI, FPA, LLE, and Fsp^3 values. Once this is done, analogs that display improved profiles relative to **1** can be considered for synthesis. Since the lead optimization process necessarily depends on improving the

dx.doi.org/10.1021/ed5002653 | J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 3. Analysis of Compound 2 at Intrinsic Potency Values Equal to or 10 to 100-fold Higher Than 1





Compound Data[a]							% Difference from Lead Values[b]			
Compd	MW	XC_{50}	BEI	FPA	LLE	Fsp^3	BEI	FPA	LLE	Fsp^3
1	357.88	100.00	19.6	0.16	2.1	0.35	0.0	0.0	0.0	0.0
2-eq	392.32	100.00	17.9	0.15	1.4	0.35	-8.7	-6.3	-33.3	0.0
2-10x	392.32	10.00	20.4	0.15	2.4	0.35	4.4	-6.3	13.6	0.0
2_100x	392.32	1.00	23.0	0.15	3.4	0.35	17.4	-6.3	60.6	0.0

[a] XC_{50} = IC_{50}, EC_{50}, or K_i value (nM); BEI = pXC_{50}/MW (kDa) FPA = $\sum$ N + O/total # HA; LLE = pXC_{50} − ClogP; Fsp^3 = $\sum$ sp^3 C atoms/total # C atoms. [b] determined using the PC/potency spreadsheet PC_potency_templ.xlsx.

properties of analogs relative to the designated lead molecule (**1**), the molecular properties of the lead can be used to obtain the seed values (see Table SI-1) needed for comparing target molecules **2** to **8** to the lead based on their PC/potency profiles. Also, given that the usual goal is for the XC_{50} value of the new analogs to equal or exceed that of the lead compound, we will use the XC_{50} of the lead molecule **1** (100 nM) for the calculation of the BEI and LLE values for **2**−**8** (see the SI for details of this process).

As shown in Table 2, the predicted profiles for compounds **2** and **3** appear less favorable relative to the lead compound (**1**). The small decrease in BEI values (<10%) seen for both analogs while still close to the acceptable range, is accompanied by a >30% change in their LLE values in the wrong direction. At this point the question may be asked, if the XC_{50} values of **2** or **3** were found to be 10 or 100-fold higher relative to **1** after actual synthesis and testing, would this finding support giving these analogs a higher priority for synthesis? This possibility can be easily explored by calculating their PC/potency profiles at these higher intrinsic potencies. Shown in Table 3 are the results from this type of analysis using compound **2** as the example.

From this analysis it can be seen from the PC/potency profiles of **2** at the new XC_{50} values of 10 and 1.0 nM, that the BEI and LLE values do become positive relative to the lead compound **1**. Closer inspection of the percent difference values, however, shows that even with a 100-fold increase in the intrinsic potency (XC_{50} = 1.0 nM), compound **2** would attain only a 60% increase in the LLE value to 3.4, a value still well below the LLE target value/range (5−7) and the median value (5.4) seen for recently approved FDA drugs (Table 1). This prospective analysis clearly shows that any attempt to discover a better drug-like molecule relying on the improvement of intrinsic potency alone could fail since it may not lead to an improvement in molecular parameters to levels expected of a compound with good drug-like properties. If the chemist was still interested in pursuing analogs in this series by modifications to the perimeter of this core structure, the introduction of polar functionality to both increase intrinsic potency and alter PC properties may be an option. Again, this possibility can be easily explored prior to synthesis by generating the PC/potency profiles for the newly proposed target molecules bearing polar functionality, and analyzing the results in the same manner as illustrated with **2** above (see Figures SI-6 and -7 and Tables SI-14 and -15).

Returning to the target compounds **4**, **5**, and **6** (Table 2), their PC/potency profiles are clearly improved relative to **1**, with each analog displaying positive a value for all four parameters. The greater than 100% increase (>2-fold) seen in LLE values for **5** (4.8) and **6** (4.7) when assuming an intrinsic potency equivalent to that of the lead compound (**1**) are just outside the target range (Table 1) and would support the placement of these compounds on the top of the priority list for synthesis. To provide further support for this conclusion, calculation of the PC/potency profiles (Figure SI-5) for these compounds by assuming a 10-fold increase in potency (Table SI-11 and -12) results in all four of the molecular parameters for **5** and **6** fall within the desired target value/range (Table 1) and therefore these compounds are likely to display improved drug-like properties relative to **1**.

A major change in the core structure of **1** can be represented by analogs **7** and **8**. The modifications found in **7** results in a very unfavorable PC/potency profile (Table 2). This simple benzamide based core structure appears to be a less favorable starting point for analog synthesis compared to the initial lead **1**. For example, simple changes to the perimeter of **7** analogous to those proposed above for **1** (e.g., **2** and **3**) are likely to produce analogs with even less favorable PC/potency profiles relative to **2** and **3**. In contrast, exploring more innovative

dx.doi.org/10.1021/ed5002653 | J. Chem. Educ. XXXX, XXX, XXX−XXX

changes to this same benzamide core represented by target compound **8**, produces a more favorable PC/potency profile with regard to three of the four parameters and would warrant a higher place on the list of candidates for synthesis. Based upon this simple analysis, a rank order of the proposed target molecules (Table 2) for actual synthesis would place **4**, **5**, **6**, and **8** on top of the list and brings into question the need to even prepare the remaining targets **2**, **3**, and **7**.

As a secondary check prior to synthesis, the proposed target molecules can be further assessed by examining their PC/potency profiles relative to any user defined target values, such as the median values for properties associated with known drugs against a specific therapeutic category (*eg.* CNS agents or kinase inhibitors) and not necessarily only those of the lead compound. As an illustration, values derived from a set of FDA approved drugs (Table 1, and Table SI-25) were used to analyze compounds **1** to **8** (Tables SI-4 to 6 and Figure SI-3) and shows that the drug-like properties of **5** and **6** are closest to the median values given in Table 1, once again supporting a position high on the list for synthesis.

The next step in this approach, assuming in this case that one of the proposed target molecules **4**, **5**, **6**, or **8**, displays improved activity relative to **1** and becomes the new lead, would be to use the experimentally determined XC_{50} and PC property values of this molecule in the next round of ligand design following a similar analysis as described above for **1** (see Figure SI-9, Tables SI-20, SI-21, and Figure SI-10). With the use of this method, any number of new analogs can be designed, limited only by the imagination of the chemist, and quickly checked for their potential to move the project forward by comparing their PC/potency profiles against that of the new lead.

Finally, the PC/potency profiles generated with the actual XC_{50} values and experimentally determined PC data can be used to generate multiparameter SAR tables. This data, when used in combination with results from PK and *in vivo* efficacy studies. may help fine-tune the selection criteria of the screening cascade for the project. Molecular parameters selected the start of a project which are found later to be of minor importance toward improving the drug-like properties of the lead series against the target, can be dropped or replaced by a new parameter[20] for future rounds of lead optimization using PC/potency profiles.

SUMMARY

The primary purpose of the method outlined above is to improve efficiency during hit-to-lead and the lead optimization stages of small molecule drug discovery by helping the chemist prioritize new target molecules for synthesis. The examples given illustrate the procedure to achieve this goal by using one of many possible sets of PC properties shown to be important for successful drugs, combined with intrinsic potencies of lead molecules, to generate PC/potency profiles which can guide new ligand design in a prospective manner. Results from *in vitro* ADME, PK, PD, and safety studies will still be the deciding factor in the selection of preclinical candidates; however, by using this method, the drug-like properties of the compounds selected for development may be expected to improve. This method also provides a useful tool for students of medicinal chemistry to integrate PC properties and intrinsic potency into small molecule drug design, and allow them to visually analyze how their proposed changes to the structure of a lead molecule will affect PC/potency profiles and therefore potential drug-like properties of target compounds. It is hoped that by implementing this easy to use method on a regular basis, chemists can identify preclinical candidates with the highest potential for success while making the fewest number of analogs along the way.

ASSOCIATED CONTENT

Supporting Information

Information on how to compute and graph compound PC/potency profiles using an accompanying Excel spreadsheet, along with data tables for all of the example compounds discussed in this article has been supplied as Supporting Information. In addition, the molecular property data for FDA approved drugs (2005 to 2011) along with the statistical analysis which produced the median values given in Table 1 is also provided. A brief analysis of several successful drug discovery projects using PC/potency profiles is presented as additional support for the potential use of this method and to illustrate a potential use of this tool for teaching medicinal chemistry. Finally, a glossary of terms and phases appearing in this article and commonly used in medicinal chemistry is also provided to assist the reader new to this subject.This material is available via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: johnson@hibiotech.com.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The author would like to thank the members of the Medicinal Chemistry Group at Hawaii Biotech for discussions and comments on the use of the compound PC/potency profiles. Their questions and critical viewpoints have helped in the preparation of this manuscript. The author would also like to thank the medicinal chemistry community whose curiosity and desire to understand what makes a good small molecule drug led to many of the useful articles cited here, and laid the foundation for the discussion on the use of compound PC/potency profiles to help guide new ligand design.

REFERENCES

(1) Gleeson, M. P.; Hersey, A.; Montanari, D.; Overington, J. Probing the links between in vitro potency, ADMET and physicochemical parameters. *Nat. Rev. Drug Discovery* **2011**, *10*, 197−208.

(2) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308−1315.

(3) Proudfoot, J. R. Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1647−1650.

(4) Vieth, M.; Siegel, M. G.; Higgs, R. E.; Watson, I. A.; Robertson, D. H.; Savin, K. A.; Durst, G. L.; Hipskind, P. A. Characteristic physical properties and structural fragments of marketed oral drugs. *J. Med. Chem.* **2004**, *47*, 224−232.

(5) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. The Design of Leadlike Combinatorial Libraries. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 3743−3748.

(6) Wenlock, M. C.; Austin, R. P.; Barton, P.; Davis, A. M.; Leeson, P. D. A comparison of physiochemical property profiles of

development and marketed oral drugs. *J. Med. Chem.* **2003**, *46*, 1250−1256.
(7) Leeson, P. D.; Davis, A. M. Time-related differences in the physical property profiles of oral drugs. *J. Med. Chem.* **2004**, *47*, 6338−6348.
(8) Meanwell, N. A. Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. *Chem. Res. Toxicol.* **2011**, *24*, 1420−1456.
(9) 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 Delivery Rev.* **1997**, *23*, 3−25.
(10) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'rule of three' for fragment-based lead discovery? *Drug Discovery Today* **2003**, *8*, 876−877.
(11) Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discovery* **2007**, *6*, 881−890.
(12) Gleeson, M. P. Generation of a set of simple, interpretable ADMET rules of thumb. *J. Med. Chem.* **2008**, *51*, 817−834.
(13) Muchmore, S. W.; Edmunds, J. J.; Stewart, K. D.; Hajduk, P. J. Cheminformatic tools for medicinal chemists. *J. Med. Chem.* **2010**, *53*, 4830−4841.
(14) Leung, C. S.; Leung, S. S.; Tirado-Rives, J.; Jorgensen, W. L. Methyl effects on protein-ligand binding. *J. Med. Chem.* **2012**, *55*, 4489−4500.
(15) Kubinyi, H. Drug research: myths, hype and reality. *Nat. Rev. Drug Discovery* **2003**, *2*, 665−668.
(16) Wager, T. T.; Chandrasekaran, R. Y.; Hou, X.; Troutman, M. D.; Verhoest, P. R.; Villalobos, A.; Will, Y. Defining Desirable Central Nervous System Drug Space through the Alignment of Molecular Properties, in Vitro ADME, and Safety Attributes. *ACS Chem. Neurosci.* **2010**, *1*, 420−434.
(17) Tarcsay, A.; Keseru, G. M. Contributions of molecular properties to drug promiscuity. *J. Med. Chem.* **2013**, *56*, 1789−1795.
(18) Hann, M. M. Molecular obesity, potency and other addictions in drug discovery. *MedChemComm* **2011**, *2*, 349−355.
(19) Verber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties that Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615−2623.
(20) Hann, M. M.; Keseru, G. M. Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat. Rev. Drug Discovery* **2012**, *11*, 355−365.
(21) Hopkins, A. L.; Keseru, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discovery* **2014**, *13*, 105−121.
(22) Ritchie, T. J.; Ertl, P.; Lewis, R. The graphical representation of ADME-related molecule properties for medicinal chemists. *Drug Discovery Today* **2011**, *16*, 65−72.
(23) Abad-Zapatero, C.; Metz, J. T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discovery Today* **2005**, *10*, 464−469.
(24) Perola, E. An analysis of the binding efficiencies of drugs and their leads in successful drug discovery programs. *J. Med. Chem.* **2010**, *53*, 2986−2997.
(25) Yan, A.; Gasteiger, J. Prediction of aqueous solubility of organic compounds based on a 3D structure representation. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 429−434.
(26) Lovering, F.; Bikker, J.; Humblet, C. Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* **2009**, *52*, 6752−6756.
(27) Yang, Y.; Engkvist, O.; Llinas, A.; Chen, H. Beyond size, ionization state, and lipophilicity: influence of molecular topology on absorption, distribution, metabolism, excretion, and toxicity for druglike compounds. *J. Med. Chem.* **2012**, *55*, 3667−3677.
(28) Ritchie, T. J.; Macdonald, S. J. The impact of aromatic ring count on compound developability—are too many aromatic rings a liability in drug design? *Drug Discovery Today* **2009**, *14*, 1011−1020.
(29) Kelder, J.; Grootenhuis, P. D.; Bayada, D. M.; Delbressine, L. P.; Ploemen, J. P. Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. *Pharm. Res.* **1999**, *16*, 1514−1519.
(30) Lu, J. J.; Crimin, K.; Goodwin, J. T.; Crivori, P.; Orrenius, C.; Xing, L.; Tandler, P. J.; Vidmar, T. J.; Amore, B. M.; Wilson, A. G.; Stouten, P. F.; Burton, P. S. Influence of molecular flexibility and polar surface area metrics on oral bioavailability in the rat. *J. Med. Chem.* **2004**, *47*, 6104−6107.
(31) Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; Decrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; Greene, N.; Huang, E.; Krieger-Burke, T.; Loesel, J.; Wager, T.; Whiteley, L.; Zhang, Y. Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4872−4875.
(32) Martin, Y. C. A bioavailability score. *J. Med. Chem.* **2005**, *48*, 3164−3170.
(33) XC50 value; the definition used in this case follows that of Gleeson et al. (ref 1) XC50 is "the concentration of a substance that results in 50% of the maximal effect of the substance on a biochemical function".