1

INTRODUCTION

The search for new drugs is a long process. Attrition is high and the costs keep escalating (now perhaps as high as $2 billion per marketed drug). The traditional discovery–development models are undergoing change, as many pharmaceutical companies reign in the R&D costs, by consolidating research sites, downsizing research staff, engaging in more outside collaborations, and outsourcing.

1.1 BULLDOZER SEARCHING FOR A NEEDLE IN A HAYSTACK?

Although the last decade has led to improvements in attrition due to poor pharmacokinetic profiles of discovery compounds, drug absorption continues to be an important issue in modern pharmaceutical research and development. The search for new drugs is daunting, expensive, and highly risky, but potentially highly rewarding.

If chemicals were confined to molecular weights of less than 600 Da and consisted of common atoms, the chemistry space is estimated to contain 10^{40} to 10^{100} molecules, an impossibly large space to search for potential drugs [1]. To address this limitation of vastness, "maximal chemical diversity" [2] was applied in constructing large experimental screening libraries. It's now widely accepted that the quality of leads is more important than the quantity. Traditionally, large compound libraries have been directed at biological "targets"

Absorption and Drug Development: Solubility, Permeability, and Charge State, Second Edition. Alex Avdeef.

to identify active molecules, with the hope that some of these "hits" may someday become drugs. The pre-genomic era target space was relatively small: Less than 500 targets had been used to discover the known drugs [3]. This number may expand to several thousand in the next few years as genomics-based technologies and better understanding of protein–protein interactions uncover new target opportunities [4, 5]. Of the estimated 3000 new targets, only about 20% are commercially exploited [5]. Due to unforeseen complexities of the genome and biologic systems, it is taking a lot longer and is more expensive to exploit the new opportunities than originally thought [5–8].

Although screening throughputs have massively increased over the past 20 years (at great cost in set up and run), lead discovery productivity has not necessarily increased accordingly [5–8]. C. Lipinski has suggested that maximal chemical diversity is an *inefficient* library design strategy, given the enormous size of the chemistry space, and especially that clinically useful drugs appear to exist as small tight clusters in chemistry space: ". . . one can make the argument that screening truly diverse libraries for drug activity is the fastest way for a company to go bankrupt because the screening yield will be so low" [1]. Hits *are* made in pharmaceutical companies, but this is because the most effective (not necessarily the largest) screening libraries are *highly focused*, to reflect the putative tight clustering. Looking for ways to reduce the number of tests, to make the screens "smarter," has an enormous cost reduction implication.

Figure 1.1 sketches out the process of drug exploration, discovery, and development followed at several pharmaceutical companies in the early 2000s [9–12]. A large pharmaceutical company may screen 100,000 to 1,000,000 molecules for biological activity each year. Some 3000–10,000 hits are made. Most of these molecules, however potent, do not have the right physicochemical, stability, and safety properties. Large pharmaceutical companies promote about 12 molecules into preclinical development each year. Only about 5 in 12 candidates survive after Phase I (Figure 1.1). A good year sees perhaps just one molecule reach the product stage after 9 molecules enter first-in-man clinical testing [6]. For that molecule, the start-to-finish may have taken 14 years (Figure 1.1).

The molecules that fail have "off-target" activity or poor side effects profiles. Unfortunately, animal models have been weak predictors of efficacy and/or safety in humans [7]. The adverse reactions in humans are sometimes not discovered until the drug is on the market in large-scale use in humans.

In 2001, a drug product cost about $880 million to bring out to market—which included the costs of numerous failures (Figure 1.1). In 2010, the cost was closer to $2 billion/approval [7]. It has been estimated that about 33% of the molecules that reach preclinical development are eventually rejected due to ADME (absorption, distribution, metabolism, excretion) problems. Other attrition causes are lack of efficacy (33%) and toxicity (34%). Much more money is spent on compounds that fail than on those that succeed. The industry has started to respond by attempting to screen out those molecules with

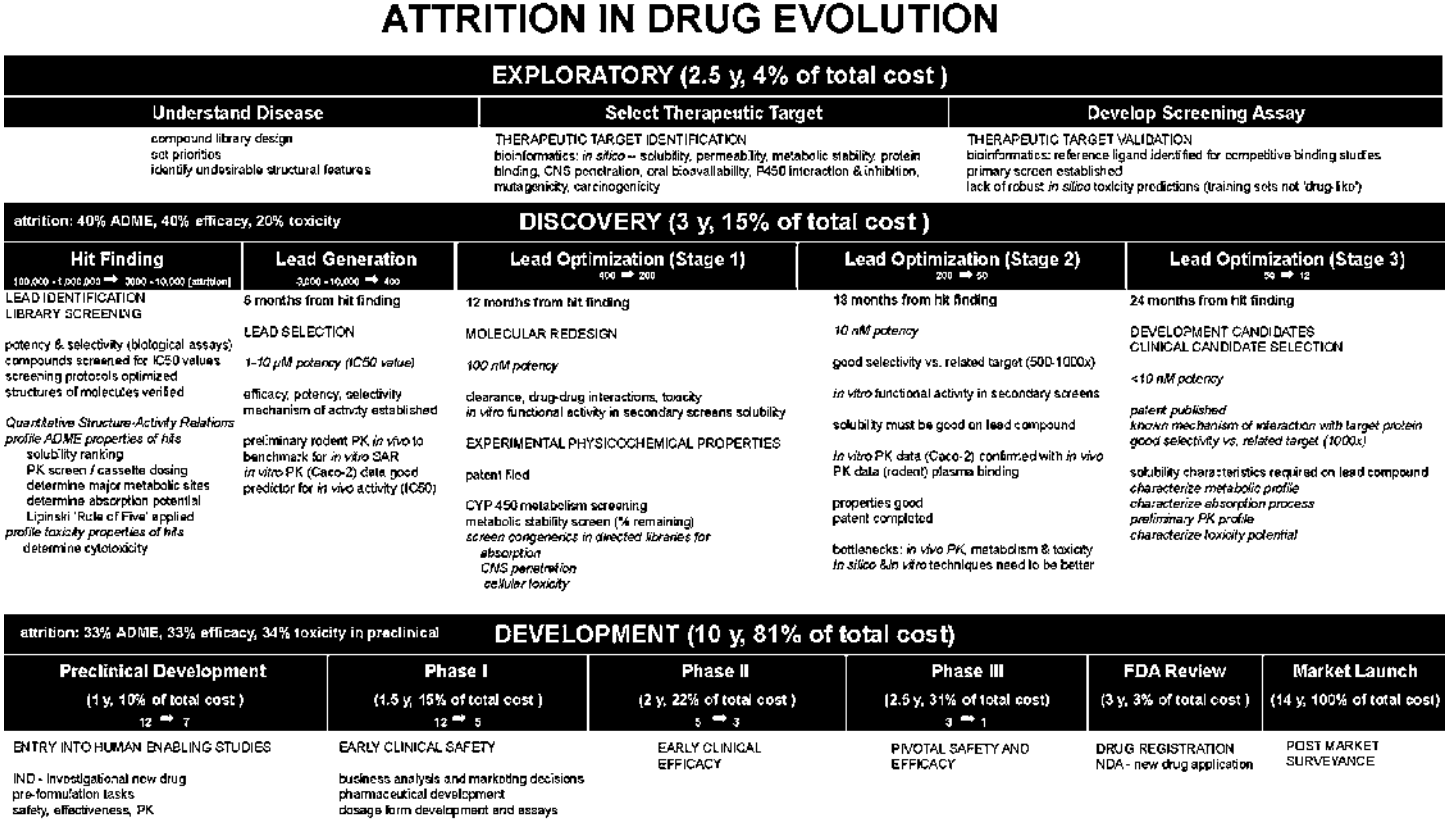

Figure 1.1 Chart summarizing the various stages in the evolution of a drug product. Included are estimated times at each stage, the cost of each stage, and the relative attrition rates. Based on studies taken from multiple sources [7–10].

poor ADME properties during discovery, before the molecules reach development. However, that has led to another challenge: how to do the additional screening quickly enough [13]. An undesirable consequence of cheap and quick assays it that their quality is low [5].

Combinatorial chemistry programs have tended to select for higher-molecular-weight molecules, predictably low in solubility. "Early warning" tools, such as Lipinski's "Rule of Five" [1] and simple computer programs that predict solubility and other properties from 2-D structure [14, 15], attempt to weed out such molecules early in discovery programs. Still, many solubility-problematic molecules remain unrecognized in early studies, due to the overly simplistic methods used to measure solubility in discovery [16]. More accurate (but still fast) solubility [16–19] (Chapter 6) and artificial membrane permeability [20–24] (Chapter 7) methods in the candidate selection stage in pharmaceutical R&D have proven to be particularly helpful for recognizing at a much earlier time the truly problematic molecules. It had even been suggested that screening for future formulation efficacy (pH and excipient effects on solubility and permeability) of candidates could be justified, if the methods were fast, compound-sparing, cost effective, and reasonably accurate [16, 18].

1.2 AS THE PARADIGM TURNS

As a consequence of the increased and unsustainable cost of bringing out a therapeutic product, many pharmaceutical companies have begun to change the way discovery and development are done [5]:

- Size and scope of internal research capabilities are decreasing, as more outsourcing is considered, not only in discovery, but also in development.
- Several companies have rearranged internal structures to be smaller "biotech-like" units.
- External collaborations with small biotech companies and academia have increased.
- Many in the industry predict that more biologic therapies will emerge (which have lower Phase II attrition [6]), and the emphasis on small molecules may decrease.

Strategies of discovery are changing [7]:

- Development of multitargeted therapeutics will increase.
- Whole pathway approaches, drawing on increasing understanding of protein–protein interactions, will be increasingly explored.
- Biology-driven drug discovery, starting with a specific disease model and a pathway, benefitting from external collaborations with academic groups.

- Analysis of multigenic complex diseases.
- Network pharmacology.
- Obtaining early proof of concepts, with small clinical studies and/or applying microdosing.

The "open innovation model" (OIM) [8] involves the progression of discovery and development that's different from that depicted in Figure 1.1. An attrition "funnel" will start with many test compounds. Even at the early stage, ideas and technologies may be either in-licensed or out-licensed. At later optimization stages, two-way collaborations with academic labs will play an increasing role. Product in-licenses will be considered. Near the product launch stage, line extensions via partners and joint ventures will become increasingly popular. In the OIM, intellectual property would be selectively distributed and proactively managed and shared to create value that otherwise would not surface.

1.3 SCREEN FOR THE TARGET OR ADME FIRST?

Most commercial combinatorial libraries, some of which are very large and may be diverse, have a very small proportion of drug-like molecules [1]. Should only the small drug-like fraction be used to test against the targets? The existing practice is to screen for the receptor activity before "drug-likeness." The reasoning is that structural features in molecules rejected for poor ADME properties may be critical to biological activity related to the target. It is believed that active molecules with liabilities can be modified later by medicinal chemists, with minimal compromise to potency. Lipinski [1] suggested that the order of testing may change in the near future, for economic reasons. He adds that looking at data already available from previous successes and failures may help to derive a set of guidelines to apply to new compounds. When a truly new biological therapeutic target is examined, nothing may be known about the structural requirements for ligand binding to the target. Screening may start as more or less a random process. A library of compounds is tested for activity. Then computational models are constructed based on the results, and the process is repeated with newly synthesized molecules, perhaps many times, before adequately promising compounds are revealed. With large numbers of molecules, the process can be costly. If the company's library is first screened for ADME properties, that screening is done only once. The same molecules may be recycled against existing or future targets many times, with knowledge of drug-likeness to fine-tune the optimization process. If some of the molecules with very poor ADME properties are judiciously filtered out, the biological activity testing process would be less costly. But the order of testing (activity versus ADME) is likely to continue to be the subject of future debates [1].

1.4 ADME AND MULTIMECHANISM SCREENS

In silico property prediction is needed more than ever to cope with the screening overload [14, 15]. Improved prediction technologies are continuing to emerge. However, reliably measured physicochemical properties to use as "training sets" for new target applications have not kept pace with the *in silico* methodologies.

Prediction of ADME properties should be simple, since the number of descriptors underlying the properties is relatively small, compared to the number associated with effective drug-receptor binding space. In fact, prediction of ADME is difficult. The current ADME experimental data reflects a multiplicity of mechanisms, making prediction uncertain. Screening systems for biological activity are typically single mechanisms, where computational models are easier to develop [1].

For example, aqueous solubility is a multimechanism system. It is affected by lipophilicity, H-bonding between solute and solvent, intra- and intermolecular H-bonding, electrostatic bonding (crystal lattice forces), and charge state of the molecule. When the molecule is charged, the counterions in solution may affect the measured solubility of the compound. Solution microequilibria occur in parallel, affecting the solubility. Many of these physicochemical factors are not well understood by medicinal chemists, who are charged with making new molecules that overcome ADME liabilities without losing potency.

Another example of a multimechanistic probe is the Caco-2 permeability assay (Chapter 8). Molecules can be transported across the Caco-2 monolayer by several mechanisms operating simultaneously, but to varying degrees: transcellular passive diffusion, paracellular passive diffusion, lateral passive diffusion, active influx or/and efflux mediated by transporters, passive transport mediated by membrane-bound proteins, receptor-mediated endocytosis, pH-gradient- and electrostatic-gradient-driven mechanisms, and so on (Chapter 2). The P-glycoprotein (Pgp) efflux transporter can be saturated if the solute concentration is high enough during the assay. If the substance concentration is very low (perhaps because not enough of the compound is available during discovery, or due to low solubility), the importance of efflux transporters in gastrointestinal tract (GIT) absorption can be overestimated, providing the medicinal chemist with an overly pessimistic prediction of intestinal permeability [1, 25]. Drug metabolism in some *in vitro* cellular systems can further complicate the assay outcome.

Compounds from traditional drug space ("common drugs"—readily available from chemical suppliers), often chosen for studies by academic laboratories for assay validation and computational model-building purposes, can lead to misleading conclusions when the results of such models are applied to "real" [12] discovery compounds, which most often have extremely low solubilities [25].

Computational models for single-mechanism assays (e.g., biological receptor affinity) get better as more data are accumulated [1]. Computational

models for multimechanism assays (e.g., solubility, permeability, charge state), in contrast, get worse as more measurements are accumulated [1]. Predictions of human oral absorption using Caco-2 permeability values can look very impressive when only a small number of molecules is considered. However, good correlations deteriorate as more molecules are included in the plot, and predictivity soon becomes tenuous. "The solution to this dilemma is to carry out single-mechanism ADME experimental assays and to construct single-mechanism ADME computational models. The ADME area is at least 5 or more years behind the biology therapeutic target area in this respect" [1].

1.5 ADME AND THE MEDICINAL CHEMIST

Although ADME assays are usually performed by analytical chemists, medicinal chemists—the molecule makers—need to have some understanding of the physicochemical processes in which the molecules participate.

> It is now almost a century since Overton and Meyer first demonstrated the existence of a relationship between the biological activity of a series of compounds and some simple physical property common to its members. In the intervening years the germ of their discovery has grown into an understanding whose ramifications extend into medicinal chemistry, agrochemical and pesticide research, environmental pollution, and even, by a curious reinvention of familiar territory, some areas basic to the science of chemistry itself. Yet its further exploitation was long delayed. It was 40 years later that Ferguson at ICI [AstraZeneca] applied similar principles to a rationalization of the comparative activity of gaseous anaesthetics, and 20 more were to pass before the next crucial step was formulated in the mind of Hansch. . . . Without any doubt, one major factor [for delay] was compartmentalism. The various branches of science were much more separate then than now. It has become almost trite to claim that the major advances in science take place along the borders between its disciplines, but in truth this happened in the case of what we now call Hansch analysis, combining as it did aspects of pharmacy, pharmacology, statistics, and physical organic chemistry. Yet there was another feature that is not so often remarked, and one with a much more direct contemporary implication. The physical and physical organic chemistry of equilibrium processes—solubility, partitioning, hydrogen bonding, etc.—is not a glamorous subject. It seems too simple. Even though the specialist may detect an enormous information content in an assemblage of such numbers, to synthetic chemists used to thinking in three-dimensional terms they appear structureless, with no immediate meaning that they can *visually* grasp. Fifty years ago it was the siren call of Ehrlich's lock-and-key theory that deflected medicinal chemists from a physical understanding that might otherwise have been attained much earlier. Today it is glamour of the television screen. No matter that what is on display may sometimes possess all the profundity of a five-finger exercise. It is visual and therefore more comfortable and easier to assimilate. Similarly, MO theory in its resurgent phase combines the exotic appeal of a mystery religion with a new-found instinct for three-dimensional colour projection which

> really can give the ingénue the impression that he understands what it is all about. There are great advances and great opportunities in all this, but nevertheless [there is] a concomitant danger that medicinal chemists may forget or pay insufficient attention to hurdles the drug molecule will face if it is actually to perform the clever docking routine they have just tried out: hurdles of solubilization, penetration, distribution, metabolism and finally of its nonspecific interactions in the vicinity of the active site, all of them the result of physical principles on which computer graphics has nothing to say. Such a tendency has been sharply exacerbated by the recent trend, for reasons of cost as much as of humanity, to throw the emphasis upon *in vitro* testing. All too often, chemists are disconcerted to discover that the activity they are so pleased with *in vitro* entirely fails to translate to the *in vivo* situation. Very often, a simple appreciation of basic physical principles would have spared them this disappointment; better, [it] could have suggested in advance how they might avoid it. We are still not so far down the path of this enlightenment as we ought to be. What is more, there seems a risk that some of it may fade if the balance between a burgeoning receptor science and these more down-to-earth physical principles is not properly kept.—Peter Taylor [26].*

In 1990, Taylor [26] described physicochemical profiling in a comprehensive and richly descriptive way, but much has happened since then. Then, instrument companies took no visible interest in making pK_a (Chapter 3), log P (Chapters 4 and 5), or solubility (Chapter 6) analyzers; it did not occur to anyone to do PAMPA (Chapter 7). Combinatorial chemistry, HTS, Caco-2 (Chapter 8), IAM, and CE were largely unknown. Thus it is a good time to take stock of what can be learned from the work of the last two decades.

1.6 THE "ABSORPTION" IN ADME

This book focuses on physicochemical profiling in support of improved prediction methods for the "absorption" in ADME. Metabolism and other components of ADME will be beyond the scope of this book. Furthermore, properties related to *passive* absorption will be the focus, and active transport mechanisms will be considered only indirectly. The most important physicochemical parameters associated with passive absorption are *acid–base* character (which determines the charge state of a molecule in a solution of a particular pH), *lipophilicity* (which determines distribution of a molecule between the aqueous and the lipid environments), *solubility* (which limits the concentration that a dosage form of a molecule can present to the solution and the rate at which the molecule dissolves from the solid form), and membrane *permeability* (which determines how quickly molecules can cross membrane barriers). Current state of the art in measurement of these properties, as the ever important function of pH, will be discussed in depth in this book.

*This excerpt was published in Comprehensive Medicinal Chemistry, Vol. 4, Peter J. Taylor, Hydrophobic Properties of Drugs, pp. 241–294, Copyright Elsevier (1990). Reproduced with permission from Elsevier.

1.7 IT IS NOT JUST A NUMBER, IT IS A MULTIMECHANISM

Drugs exert their therapeutic effects through reactions with specific receptors. Drug-receptor binding depends on the concentration of the drug near the receptor. Its form and concentration near the receptor depend on its physical properties. Orally administered drugs need to be dissolved at the site of absorption in the GIT, and need to traverse several membrane barriers before receptor interactions can commence. As the drug distributes into the various compartments of the body, a certain (small) portion finds itself at the receptor site. Transport and distribution of most drugs are affected by passive diffusion, which depends on lipophilicity, since lipid barriers need to be crossed [27]. Passive transport is well described by the principles of physical chemistry [27–29].

The goal of this book is to examine the components of the multimechanistic processes related to charge state: the pK_a of molecules (Chapter 3), lipophilicity (Chapters 4 and 5), solubility (Chapter 6), and permeability (Chapters 7–9), with the aim of advancing improved strategies for *in vitro* assays related to drug absorption. In high-throughput screening (HTS) these parameters are sometimes viewed simply as numbers, quickly and roughly determined, to be used to rank molecules into "good" and "bad" classes. An attempt will be made to examine this important aspect. In addition, how fundamental, molecular-level interpretations of the physical measurements can help to improve the design of the profiling assays will be examined, with the aim of promoting the data fodder of HTS to a higher level of quality, without compromising the need for high speed [16–24]. Quality measurements in large quantities will lead to improved *in silico* methods. Simple rules (presented in visually appealing ways), in the spirit of Lipinski's rule of fives, will be sought, of use not only to medicinal chemists but also to preformulators. This book attempts to make easier the dialog between the medicinal chemists charged with modifying test compounds and the pharmaceutical scientists charged with physicochemical profiling, who need to communicate assay results in an optimally effective manner.

REFERENCES

1. Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **44**, 235–249 (2000).
2. Martin, E. J.; Blaney, J. M.; Siani, M. A.; Spellmeyer, D. C.; Wong, A. K.; Moos, W. H. Measuring diversity: Experimental design of combinatorial libraries for drug discovery. *J. Med. Chem.* **38**, 1431–1436 (1995).
3. Drews, J. Drug discovery: A historical perspective. *Science* **287**, 1960–1963 (2000).
4. Pickering, L. Developing drugs to counter disease. *Drug Discov. Dev.* **Feb.**, 44–47 (2001).

5. Perrior, T. Overcoming bottlenecks in drug discovery. *Drug Discov. World* 29–33 (Fall 2010).
6. Kola, I.; Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nature Rev. Drug Discov*. **3**, 711–715 (2004).
7. Haberman, A. B. Overcoming phase II attrition problem. *Gen. Eng. Biotech. News* **29**, 63–67 (2009).
8. Hunter, J. Is the pharmaceutical industry open for innovation? *Drug Discov. World* **Fall**, 9–14 (2010).
9. Allan, E.-L. Balancing quantity and quality in drug discovery. *Drug Discov. World* **Winter**, 71–75 (2002/2003).
10. Browne, L. J.; Taylor, L. L. *Drug Discov. World* **Fall**, 71–77 (2002).
11. Kerns, E. H.; Di, L. *Drug-like Properties: Concepts, Structure Design and Methods*, Academic Press, Amsterdam, 2008.
12. Rydzewski, R. M. *Real World Drug Discovery—A Chemist Is Guide to Biotech and Pharmaceutical Research*, Elsevier, Amsterdam, 2008.
13. 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.* **23**, 3–25 (1997).
14. Algorithm Builder v1.8; ADME Boxes v4.9; ACD/pK_a Database in ACD/ChemSketch v3.0; ACD/Solubility DB. Advanced Chemistry Development Inc., Toronto, Canada (www.ACD/Labs.com).
15. MarvinSketch v5.3.7. ChemAxon, Budapest, Hungary (www.chemaxon.com).
16. Glomme, A.; März, J.; Dressman, J. B. Comparison of a miniaturized shake-flask solubility method with automated potentiometric acid/base titrations and calculated solubilities. *J. Pharm. Sci.* **94**, 1–16 (2005).
17. Bergström, C. A. S.; Luthman, K.; Artursson, P. Accuracy of calculated pH-dependent aqueous drug solubility. *Eur. J. Pharm. Sci.* **22**, 387–398 (2004).
18. Avdeef, A.; Bendels, S.; Tsinman, O.; Kansy, M. Solubility—Excipient classification gradient maps. *Pharm. Res.* **24**, 530–545 (2007).
19. Avdeef, A. Solubility of sparingly-soluble drugs. [Dressman, J; Reppas, C. (eds.). Special issue: The Importance of Drug Solubility]. *Adv. Drug Deliv. Rev.* **59**, 568–590 (2007).
20. Kansy, M.; Avdeef, A.; Fischer, H. Advances in screening for membrane permeability: High-resolution PAMPA for medicinal chemists. *Drug Discov. Today: Technologies* **1**, 349–355 (2005).
21. Avdeef, A.; Artursson, P.; Neuhoff, S.; Lazarova, L.; Gräsjö, J.; Tavelin, S. Caco-2 permeability of weakly basic drugs predicted with the double-sink PAMPA pK_a^{flux} method. *Eur. J. Pharm. Sci.* **24**, 333–349 (2005).
22. Avdeef, A. The rise of PAMPA. *Expert Opinion Drug Metab. Toxicol.* **1**, 325–342 (2005).
23. Avdeef, A.; Bendels, S.; Di, L.; Faller, B.; Kansy, M.; Sugano, K.; Yamauchi, Y. PAMPA—A useful tool in drug discovery. *J. Pharm. Sci.* **96**, 2893–2909 (2007).
24. Sugano, K.; Kansy, M.; Artursson, P.; Avdeef, A.; Bendels, S.; Di, L.; Ecker, G. F.; Faller, B.; Fischer, H.; Gerebtzoff, G.; Lennernäs, H.; Senner, F. Coexistence of passive and active carrier-mediated uptake processes in drug transport: A more balanced view. *Nature Rev. Drug Discov.* **9**, 597–614 (2010).

25. Lipinski, C. A. Avoiding investment in doomed drugs—Is solubility an industry wide problem? *Curr. Drug Discov.* **Apr**, 17–19 (2001).
26. Taylor, P. J. Hydrophobic properties of drugs. In: Hansch, C.; Sammes, P. G.; Taylor, J. B. (eds.). *Comprehensive Medicinal Chemistry*, Vol. 4, Pergamon, Oxford, 1990, pp. 241–294.
27. Kubinyi, H. Lipophilicity and biological activity. *Arzneim.-Forsch./Drug Res.* **29**, 1067–1080 (1979).
28. van de Waterbeemd, H.; Smith, D. A.; Jones, B. C. Lipophilicity in PK design: Methyl, ethyl, futile. *J. Comp.-Aided Molec. Design* **15**, 273–286 (2001).
29. van de Waterbeemd, H.; Smith, D. A.; Beaumont, K.; Walker, D. K. Property-based design: Optimization of drug absorption and pharmacokinetics. *J. Med. Chem.* **44**, 1313–1333 (2001).