Editorial

Solubility: A speed–breaker on the drug discovery highway

Volume 3 Issue 3 - 2017

Aravindhan Ganesan, Khaled Barakat
University of Alberta, Canada.

Correspondence: Khaled Barakat, Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, Edmonton, Alberta, Canada, Tel 1780 492 5783, Email kbarakat@ualberta.ca

Received: April 17, 2017 | **Published:** April 25, 2017

Introduction

Drug development remains to be a lengthy, painful and expensive process. According to a very recent report by the Tufts Center for the Study of Drug Development (CSDD), the overall development cycle of a new market-approved drug could cost a whopping $2.6 billion and approximately 14 years. The past two decades have seen the emergence of more sophisticated experimental techniques, advanced computational methods and complementary technologies in the field of drug design and development.[1,2] Such advances have not led to the increase in the number of drugs in the market though. For instance, in the year 2016, the U.S. Food and Drug Administration (FDA) approved only 22 new drugs. The reasons for the slow progress can be attributed to attrition of lead candidates along the drug discovery pipeline. It has been estimated that ~40% of the attrition rate of candidate-drugs has been associated with poor pharmacokinetic features and toxicity. Poor solubility is, particularly, a very significant impediment in the drug development efforts.

The solubility of a drug molecule is vital for its' bioavailability. If an orally administered drug is not sufficiently soluble, then it could not be fully absorbed into the blood circulation and will be expelled from the gastrointestinal tract before reaching its' site of action. Nevertheless, hydrophobicity and innate low–water solubility are gradually becoming rather unsurprising characteristics of early hits, lead compounds and even in some market–approved drugs.[3] Nearly 60%–90% of the compounds that are currently being developed exhibit poor water–soluble features[3,4] and categorized under the Biopharmaceutical Classification System (BCS) classes II (low solubility and high permeability) and IV (low solubility and low permeability).[4–6] This is mainly because, most of the ligand-binding sites in the target proteins are secluded from the aqueous environment and hence hydrophobic compounds are generally preferred (or developed) in order to gain high binding affinity and activity against the target(s).[3,7] In addition, recently, the drug discovery research is also seeing a paradigm shift from enzymes to more complicated therapeutic targets, such as ion channels, protein-protein interfaces, kinases and nuclear receptors.[1,5,8–10] Such challenging targets generally demand more lipophilic drugs, which also involves high-crystal energies from strong intermolecular interactions.[5] These factors altogether tend to adversely impact the solubility of compounds.

Tackling solubility concerns in market–approved drugs

In spite of poor solubility, if a compound exhibits desirable activity (against the target protein), then it could be formulated into successful drugs with the help of various drug delivery technologies. This includes pH modification technology, co–solvent and surfactant solubilization, nanoparticle technology and micro emulsion drug delivery systems.[11,12] For example, ciprofloxacin, a current fluoroquinolone-based antimicrobial agent, is weakly basic and have poor water solubility at neutral pH.[11] But the solubility of this agent increases when the pH increases. As such, the currently administered formulations of ciprofloxacin contain lactic acid as a solubilizing agent (or pH modifier) and hydrochloric acid for pH adjustment.[11] Imatinib, a protein-tyrosine kinase inhibitor for targeting cancers, is another example for pH-dependent soluble drugs. This drug displays very poor solubility; nevertheless, its' β-polymorphic form (i.e., imatinib salt) has showed to be very soluble at pH< 5.5.[11,13] Hence, Imatinib mesylate has been orally administered for a targeted cancer therapy.[11,13] Inclusion of co-solvents is also employed as a successful strategy for the formulation of insoluble drugs in the market.[14–16] Ethanol, polyethylene glycol (PEG 300 or 400), dimethylacetamide, dimethyl sulfoxide (DMSO) and polysorbates are commonly used co-solvents.[3] For instance, docetaxel, an anti-cancer chemotherapy drug, is poorly soluble and found to show only ~8% oral bioavailability. In order to improve the solubility of the drug, the new formulations are administered with PEG 300 and polysorbate 80.[11] Thus, insoluble drug delivery technologies provide a great opportunity to save the poorly soluble drugs (but with high activity) that passed the most-time consuming and expensive clinical phases.[11,12] At the same time, it is important to acknowledge that the success is not always guaranteed with the drug delivery technologies[3] and they also involve additional costs and time to identify and establish a suitable technology.

Solubility challenges in biochemical assays

Apart from bioavailability concerns, the low solubility of compounds poses a significant challenge for *in vitro* and *in vivo* assays during the hit-to-lead stage of drug discovery. In particular, the low solubility of compounds can impact the biological assays in a number of ways, including underrated activity, variable data, inaccurate structure–activity relationships (SAR), reduced HTS-hit rates, and inaccurate toxicity estimations. For instance, the low–solubility of compounds in DMSO causes variable data from bioassay experiments. It is a common practice to dissolve each of the compounds at a concentration of 10–30 mM in DMSO and an aliquot is employed in experiments. And it is understood that ~10–20% of discovery compounds are usually not soluble in DMSO at high concentrations.[17] If the initial concentration of the compound is not correct, then this will lead to inaccurate concentrations in all of the following dilutions. In most bioassays, the dose–dependent activity of the test compounds over a range of concentrations will be carried out to determine the concentration of half-maximal inhibition (or IC50). Again in this case, if the compounds are unable to solubilise

at the highest concentration, this will shift the dose-curve and lead to inaccurate prediction of IC50 values for the compounds. Furthermore, the physicochemical properties of low soluble compounds also tend to increase the possibility of non-specific adsorption of these compounds to the surfaces of different equipment involved in the experiments, such as pipettes, syringes, multiwell plates, etc.[3] As a consequence, this leads to reduced drug concentration in solution, thus leading to erroneous interpretation of results. Di et al.[17] discussed a range of solubility challenges for biological assays and various strategies for optimizing the assay systems. Some of their suggestions to enhance bioassay performance for low-solubility compounds include, serial dilution in DMSO (not in buffers), screening at lower concentrations, in-well sonication, dissolve salts in 1:1 DMSO: water solvent and use of solid arrays to store compounds. In their recent pharmacological review, Williams et al.[3] have also presented extensive discussions of various strategies for addressing low drug solubility in drug discovery and development. Although different techniques (and strategies) are being continuously developed and refined to address the low solubility of compounds in biochemical assays, the solubility challenges are only constantly increasing.

Methods for solubility estimation of drug–like compounds

There has been wide agreement on the fact that it is hugely beneficial (in terms of money, time and labour) if the solubility of compounds can be determined at the early stages of drug development, i.e., even during the 'hit-search' phase. This realization has led to the developments of various experimental techniques and computational approaches for solubility estimations of drug (or drug-like) molecules. Experimental methods for measuring the solubility of compounds include kinetic, semi-equilibrium and equilibrium methods. The former two approaches are mostly employed as a high-throughput assay during the early stage of drug discovery; while the equilibrium method is low-throughput and hence is used at advanced stages of development. Major pharma industries, such as AstraZeneca, Novartis, Roche and Pfizer have been reported to employ semi–equilibrium techniques for high–throughput solubility measurements.[5] On the other hand, kinetic method has been employed by Boehringer Ingelheim, GSK, Pfizer & Warner–Lambert.[5] Different analytical methods, such as light scattering, UV plate reader, LC-MS and LC-UV are employed in these solubility assays and each of them has its own merits and demerits that also impact the results obtained. Di et al.[5] has reviewed in detail about the various solubility measurement approaches, their usefulness and limitations in drug discovery and development.

On the other end of the spectrum, there were also significant efforts focused on developing computational models as low–cost tools to predict the solubility of drugs and drug-like molecules. Obviously, the expectation here is that a computation tool could make quick and accurate estimations of aqueous solubility of compounds based on their molecular structures, before testing them in more expensive and complex biochemical assays. Quantitative structure property relationship (QSPR) approach has been most popularly employed for this purpose. The QSPR approach generally aims to ascertain a mathematical relationship (or correlation) between the solubility property of the compounds and the molecular descriptors obtained from their 1D, 2D or 3D structures. Different methods such as multiple linear regression (MLR), artificial neural networks (ANN), support vector machines and random forest (RF) and Partial–Least Squares (PLS) are employed for developing QSPR models.[7,18–21] Until recently, several QSPR models have been developed for predicting the solubility of drugs and most of them are able to perform with an accuracy of 0.6–0.9 log solubility.[18] For example, Palmer et al.[19] has developed QSPR models using RF, SVM, ANN and PLS for the prediction of aqueous solubility of drugs. These models were developed based on the experimental data for 988 organic compounds and found the RF model to predict the aqueous solubility of compounds with a root mean square error (RMSE) of 0.69 log S units, which was better than the other models in this study. Hughes et al.[20] compared the different QSPR models for predicting the aqueous solubility (logS), melting point (Tm) and octanol–water partition coefficient (LogP) of drug–like compounds. The results found that, the QSPR model for LogP made predictions with an R^2=0.87, while the R^2 values for the predictions of Tm and LogS were predicted with 0.49 and 0.79, respectively. Most QSPR models perform better for general organic compounds than drugs or drug–like molecules. The reasons for these are due to the quantity and quality of experimental data available for drugs. Most QSPR models are generally developed using the available experimental data for the chosen training set; however such measured data are not available for large number of drugs but only for organic compounds.[18–20] Even the available experimental values are gathered from different sources which means that the actual measurements of the data sets could have been carried out under different experimental conditions and different assays.[21] Hence, the QSPR models, for that case any informatics-based model generally tend to perform poorly when it comes to drugs.

Such drawbacks from QSPR models can be overcome by employing computational chemistry (or theoretical chemistry) approaches, such as classical simulations based on empirical force fields or using quantum mechanics. These approaches attempt to model the real system used in bioassays, thus giving much better understanding about the solubility and aggregation of drugs under prescribed conditions. Nevertheless, these approaches are in general very computationally intensive and do not bring the high-throughput value of informatics-based methods. Palmer et al.[18] employed ab initio calculations to predict the intrinsic aqueous solubility of some organic compounds and found that the accuracy of predictions was still not high. So, the authors combined informatics and computational chemistry-based approaches to establish a thermodynamic cycle and predict the solubility, which resulted in predictions with RMSE=0.71. Recently, a research team conducted an open challenge to predict the solubility of 32 compounds that they have determined experimentally.[22,23] For this purpose, the authors also provided the experimental measurements for 100 drug-like molecules (with wide range of MW and pKa values) for using as a test set (or training set). This open solubility challenge[22,23] received ~100 participants, and each of them employed a variety of computational models at their disposal (which were not revealed in the paper) to predict the solubility of the 32 compounds. The results[23] from this challenge show that the predictions for ~28 compounds were made with an R^2 between 0.0 to 0.642. Although, what computational methods were employed for predictions and what was the expertise levels of the participants were not revealed in the paper,[23] one thing was clear; the current computational methods still needs many improvements for accurate estimation of solubility of drugs. Particularly, more experimental data about drug solubility needs to be made available, which then can be used to either develop new predictive models or improve the existing ones. Further, there are also some important challenges such as

Citation: Ganesan A, Barakat K. Solubility: A speed–breaker on the drug discovery highway. MOJ Bioequiv Availab. 2017;3(3):56–58.
DOI: 10.15406/mojbb.2017.03.00033

accounting solubility of drugs at different pH conditions and different (or a mixture) of solvent conditions employed in the assays, while making predictions. Such challenging questions can be addressed by employing molecular modelling and dynamics simulations for at least select set of compounds in different scaffolds, which can then drive the experiments. For example, a mixture of a desired drug, tween80 and DMSO at a concentrations employed in a bioassay can be modelled and simulated quickly (using classical MD) to see if the drug forms any aggregates. This will not only be useful to predict solubility of the compound, but also provide an overall understanding for the molecular basis to drug solubility. Having said that, there is always a trade-off and in this case, it is the escalation of the computational costs to gain better accuracy, which we think is worth the price. Finally, the rule of 'one-size-does not-fit-all' still holds true in the case of computational prediction of aqueous solubility. One needs to choose one or a combination of methods to gain the optimal accuracy levels which can complement (and sometimes guide) further experiments and assist in delivering safe drugs to the market.

Acknowledgements

None.

Conflict of interest

The author declares no conflict of interest.

References

1. Ganesan A, Barakat K. Applications of computer–aided approaches in the development of hepatitis C antiviral agents. *Expert Opin Drug Discov*. 2017;12(4):407–425.
2. Ganesan A, Coote ML, Barakat K. Molecular dynamics–driven drug discovery:leaping forward with confidence. *Drug Discov Today*. 2017;22(2):249–269.
3. Williams HD, Trevaskis NL, Charman SA, et al. Strategies to Address Low Drug Solubility in Discovery and Development. *Pharmacol Rev*. 2013;65(1):315–499.
4. Di L, Kerns EH, Carter GT. Drug–Like Property Concepts in Pharmaceutical Design. *Curr Pharm Des*. 2009;15(19):2184–2194.
5. Di L, Fish PV, Mano T. Bridging solubility between drug discovery and development. *Drug Discov Today*. 2012;17(9–10):486–495.
6. Thapa RK, Choi HG, Kim JO. Analysis and optimization of drug solubility to improve pharmacokinetics. *Journal of Pharmaceutical Investigation*. 2017;47(2):95–110.
7. Cisneros JA. Systematic Study of Effects of Structural Modifications on the Aqueous Solubility of Drug–like Molecules. *ACS Med Chem Lett*. 2017;8(1):124–127.
8. Anwar Mohamed A, Barakat KH, Bhat R, et al. A human ether–á–go–go–related (hERG) ion channel atomistic model generated by long supercomputer molecular dynamics simulations and its use in predicting drug cardiotoxicity. *Toxicol Lett*. 2014;230(3):382–392.
9. Jordheim LP, Barakat KH, Heinrich Balard L, et al. Small Molecule Inhibitors of ERCC1–XPF Protein–Protein Interaction Synergize Alkylating Agents in Cancer Cells. *Mol Pharmacol*. 2013;84(1):12–24.
10. Wells JA, Mc Clendon CL. Reaching for high–hanging fruit in drug discovery at protein–protein interfaces. *Nature*. 2007;450(7172):1001–1009.
11. Kalepu S, V Nekkanti. Insoluble drug delivery strategies: review of recent advances and business prospects. *Acta Pharm Sin B*. 2015;5(5):442–453.
12. Savjani KT, Gajjar AK, Savjani JK. Drug Solubility: Importance and Enhancement Techniques. *ISRN Pharm*. 2012;195727.
13. Béni S, Szakács Z, Csernák O, et al. Cyclodextrin/imatinib complexation: Binding mode and charge dependent stabilities. *Eur J Pharm Sci*. 2007;30(2):167–174.
14. Hennenfent KL, Govindan R. Novel formulations of taxanes: a review. Old wine in a new bottle? *Ann Oncol*. 2006;17(5):735–749.
15. Lee YC, Zocharski PD, Samas B. An intravenous formulation decision tree for discovery compound formulation development. *Int J Pharm*. 2003;253(1–2):111–119.
16. Kawakami K, Oda N, Miyoshi K, et al. Solubilization behaviour of a poorly soluble drug under combined use of surfactants and cosolvents. *Eur J Pharm Sci*. 2006;28(1–2):7–14.
17. Di L, Kerns EH. Biological assay challenges from compound solubility: strategies for bioassay optimization. *Drug Discov Today*. 2006;11(9–10):446–451.
18. Palmer DS, Llinàs A, Morao I, et al. Predicting Intrinsic Aqueous Solubility by a Thermodynamic Cycle. *Mol Pharm*. 2008;5(2):266–279.
19. Palmer DS, O Boyle NM, Glen RC, et al. Random Forest Models to Predict Aqueous Solubility. *J Chem Inf Model*. 2007;47(1):150–158.
20. Hughes LD, Palmer DS, Nigsch F, et al. Why Are Some Properties More Difficult To Predict than Others? A Study of QSPR Models of Solubility, Melting Point, and Log P. *J Chem Inf Model*. 2008;48(1):220–232.
21. Bergström CA. In silico Predictions of Drug Solubility and Permeability: Two Rate–limiting Barriers to Oral Drug Absorption. *Basic Clin Pharmacol Toxicol*. 2005;96(3):156–161.
22. Llinàs A, Glen RC, Goodman JM. Solubility Challenge: Can You Predict Solubilities of 32 Molecules Using a Database of 100 Reliable Measurements? *J Chem Inf Model*. 2008;48(7):1289–1303.
23. Hopfinger AJ, Esposito EX, Llinàs A, et al. Findings of the challenge to predict aqueous solubility. *J Chem Inf Model*. 2009;49(1):1–5.