Enabling Technology

Molecular Modeling and Data Science for De-risking and Cost-saving of Pharmaceutical Development

Dr. Yuriy Abramov

Introduction

With advances in computational materials science and process simulation, computational applications are becoming an integral part of the drug development and manufacturing workflow.[1,2] These methods facilitate easier navigation through the complex multidimensional space of pharmaceutical process development tasks. By combining computational and experimental approaches, pharmaceutical development can be de-risked and lead to significant time and cost savings. Molecular modeling and data science are among the most popular computational methods, providing a rational approach and guidance to the drug development workflow, as discussed in this publication. This rational design approach helps de-risk experimental work by prioritizing the most promising solutions for targeted experimental follow-up.

Computational Capabilities

At J-Star-Research / Porton USA we utilize a wide range of computational methods - from molecular mechanics to quantum chemistry and data science - to study systems ranging from small molecules to new modalities in gas, liquid, and solid phases (Figure 1). Selection of the method depends on a compromise between the expected accuracy of predictions and the computational feasibility of the application; and the available amount of historical data for mining and analysis and potential artificial intelligence/machine learning (AI/ML) model building.

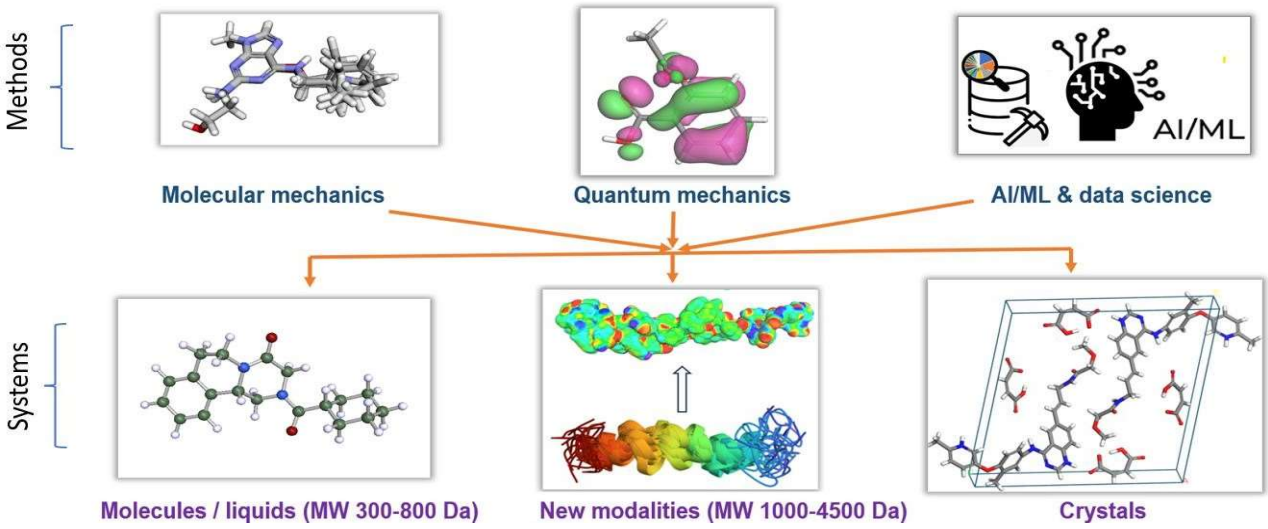

Figure 1. Computational capabilities at J-Star Research/Porton in terms of computational levels of theory and application systems.

Computational approaches

Molecular simulation and data science approaches for combined computational and experimental project support can be divided into four categories: virtual screening approaches, properties characterization and optimization, chemical reaction and reactivity predictions, and AI/ML-based DoE calculations. All applications were validated on multiple internal and external projects.

Virtual screening approaches

Virtual screening is a computational technique used to rank a database of compounds (for example, solvents, coformers, counterions or excipients) to identify the most promising solutions for a specific task (Table 1, Figure 2).[3-5] A subset of the most favorable compounds is recommended for a targeted experimental follow up.

Virtual solvent and coformer/counterion screening for crystallization requires only API molecular structure as an input to screen 62 class 2 and 3 solvents and over 100 coformers and counterions. The estimated average performance of the solvent and coformer virtual screening is above 0.85 out of 1.0. Accuracy of counterion virtual screening is slightly lower.[4] The average performance of the best solvent system selection for impurity rejection is above 0.85 out of 1.0.

Approach	Tasks	Input	Output
Virtual solvent screening for pharmaceutical crystallization[3,4]	Solubility and solvate propensity prediction; solvent selection for salt/cocrystal congruent crystallization	API molecular structure	~25 most promising solvent systems
Virtual coformer/counterion screening for cocrystal/salt crystallization[3,4]	Cocrystal/salt screening	API molecular structure	~25 most promising counterions or coformers
Virtual solvent and solid form screening for impurity(ies) rejection[5]	Solvent and solid form selection for efficient rejection of impurity(ies)	Molecular structures of API and impurity(ies); API crystal structure (optional)	Best solvent and solid forms for efficient impurity(ies) rejection
Virtual excipient screening	Excipient selection for maximum API load in amorphous phase	API molecular structure	Few most promising excipient choices

Table 1. Virtual screening approaches

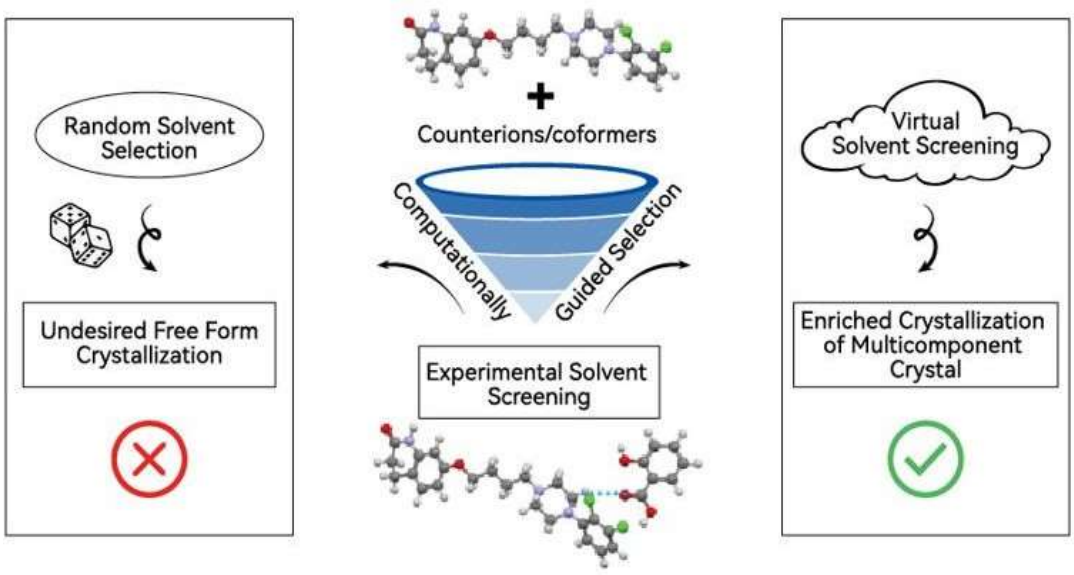

Figure 2. Combined computational and experimental coformer/counterion and solvent screening for multicomponent solid forms crystallization utilizing various virtual screening approaches.[4]

Properties characterization and optimization approaches

Computational approaches for properties characterization and optimization are presented in Table 2. The primary goal of these approaches is to predict the characteristic properties of solid forms or molecules either before or instead of experimental measurements. This is particularly important when experimental evaluation is hindered by physical phenomena (e.g., poor X-ray scattering by polar hydrogens, which may complicate distinguishing between a salt and a cocrystal solid form using SCXRD) or by insufficient material for measurements (e.g., impurities). Most property characterization approaches involve quantum chemistry calculations, which can be, and have been, applied to the characterization of new modality systems.

Approach	Tasks	Input	Output
Observed crystal form characterization along salt-cocrystal spectrum[6]	Determine whether the solid form is a salt or a cocrystal	X-ray crystal structure	Definitive salt vs cocrystal determination
Crystal shape prediction and optimization	Optimization of a poor crystal shape (needle or plate) by solvent virtual screening	Crystal structure	Crystal shape prediction in various solvent systems
Mechanical properties prediction	Predict plasticity and stiffness of the solid form(s)	Crystal structure	Plasticity and rigidity determination by predicted Young's modulus and Hardness
Molecular and crystal form analytical properties prediction	Prediction and/or interpretation of analytical properties, particularly in cases where measurements are not feasible	Molecular or crystal structure	(ss)NMR, UV-vis, Raman, IR, ECD, or VCD spectrum
Chemometric analysis of spectral data	Extraction of meaningful information from complex spectroscopic datasets	Complex (hyper) spectroscopic data (e.g., FTIR, Raman, etc)	Patterns and qualitative or quantitative relationships; clusters, PCA analysis and/or machine learning models

Table 2. Properties characterization and optimization approaches

Chemical reaction and reactivity predictions

Quantum mechanical calculations are conducted to predict heat and free energy of reactions in solvent systems of interest. The regio- and stereoselectivity of reactions are also investigated through reactivity calculations. Transition state searches can also be performed to determine the activation barriers of reactions. These calculations can help identify the most effective synthetic pathways before experimentation, provide insight into reaction mechanisms, or predict reaction properties needed for chemical engineering computations.[7]

AI/ML-based DoE

A typical process design task is performed in a multidimensional parameter space (Pj) and requires the optimization of one or more target properties (Tj). Statistical DoE is a traditional approach to process optimization, typically requiring extensive experimental measurements. Instead, we apply an AI/ML-DoE platform,[8] which reduces the number of experiments needed to optimize multidimensional process development for crystallization, process chemistry, analytical chemistry, and more. The AI/ML-DoE platform employs a Bayesian optimization approach, beginning with a limited number of initial measurements (five or more) to guide subsequent experimentation. This method facilitates rapid and reliable convergence to the parameter space for achieving optimal target properties within just a few iterations (Figure 3).

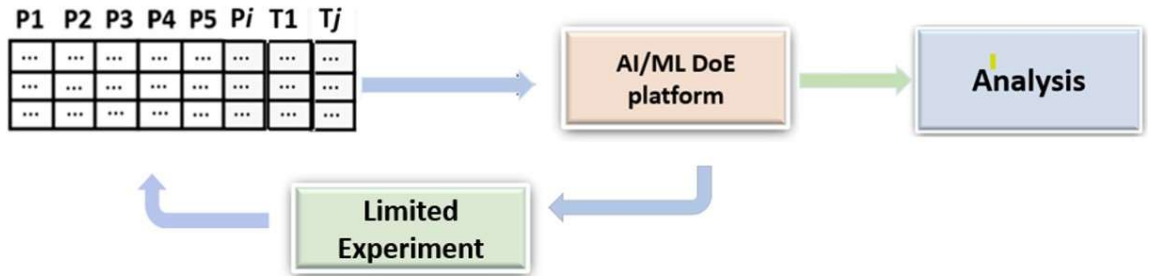

Figure 3. Workflow of AI/ML-DoE process optimization.

Conclusions

This publication provided an overview of the advanced molecular modeling and data science technologies that enable the pharmaceutical development process. Despite the high level of theory involved, the computational approaches provide real-time support for fast-paced projects. The combined computational and experimental approaches are instrumental in accelerating the pharmaceutical development process, enhancing accuracy, and reducing both risk and cost.

Dr. Yuriy Abramov

Executive Director, Computational Chemistry and Data Science of PORTON J-STAR

23+ Years Exp. in Pharmaceutical Industries, supporting Drug Discovery, Material Sciences, DP Formulation, Process, and Analytical Chemistry; Development and Application of Multiple Novel Computational Methodologies.

yuriy.abramov@jstar-research.com
Click here to follow the author on LinkedIn

References

1. Abramov, Y.A., Ed. Computational Pharmaceutical Solid State Chemistry, John Wiley & Sons, 2016.

2. Abramov, Y. A.; Sun, G.; Zeng, Q. Emerging Landscape of Computational Modeling in Pharmaceutical Development. J. Chem. Inf. Model. 2022, 62, 1160.

3. Abramov, Y.A.; Loschen, C.; Klamt, A. Rational Coformer or Solvent Selection for Pharmaceutical Cocrystallization or Desolvation. J. Pharm. Sci., 2012, 101, 3687.

4. Shah, H.S.; Michelle, C.; Xie, T.; Chaturvedi, K.; Kuang. S.; Abramov, Y.A. Computational and Experimental Screening Approaches to Aripiprazole Salt Crystallization. Pharm Res. 2023, 40, 2779.

5. Abramov, Y.A.; Zelellow, A.; Chen, C.-Y.; Wang,J.; Sekharan, S. Novel Computational Approach to Guide Impurities Rejection by Crystallization: A Case Study of MRTX849 Impurities. Cryst. Growth Des. 2022, 12, 6844.

6. Abramov, Y.A.; Wang, J. Is It Salt, Cocrystal, or Continuum? Successes and Limitations of Computationally Efficient Periodic System Calculations. Cryst. Growth Des. 2024, 24, 4017.

7. Lam, Y.; Abramov, Y.A.; Ananthula, R.S., et al. Applications of Quantum Chemistry in Pharmaceutical Process Development: Current State and Opportunities. Org. Process Res. Dev. 2020, 24, 1496.

8. https://sunthetics.io/