

Schrödinger Invests Deeply in RDKit

8th RDKit UGM, 2019
University of Hamburg, Germany

Kyle D. Konze, Pieter H. Bos, Markus K. Dahlgren, Karl Leswing, Ivan Tubert-Brohman, Andrea Bortolato, Braxton Robbason, Robert Abel, and Sathesh Bhat
Schrödinger Inc, 120 West 45th St, 17th floor, New York, New York, 10036

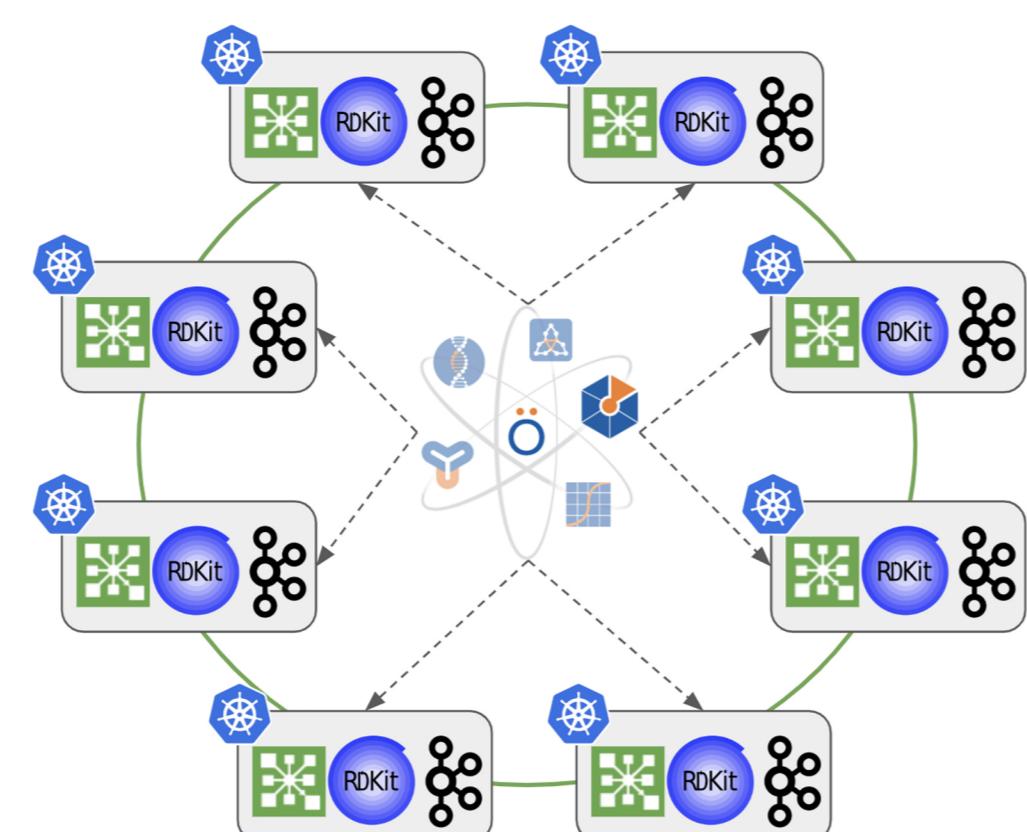
INTEGRATING RDKIT INTO SCHRÖDINGER RESEARCH

Using RDKit and tools based on the RDKit, Schrödinger has been able to extend the scope of applicability of its physics-based platform for drug discovery.

Schrödinger has contributed a 2D coordinate generation tool into the latest release of RDKit. Schrödinger has also developed an open source GPU-enabled similarity search tool, and will continue to support and improve both of these currently available tools. Schrödinger's open source product, PyMOL, now contains RDKit-based plugins, with plans for further integration. Schrödinger is also proud to contribute to the open source Python library toolchain DeepChem, a project developed through academic and industrial collaborators, which heavily uses RDKit.

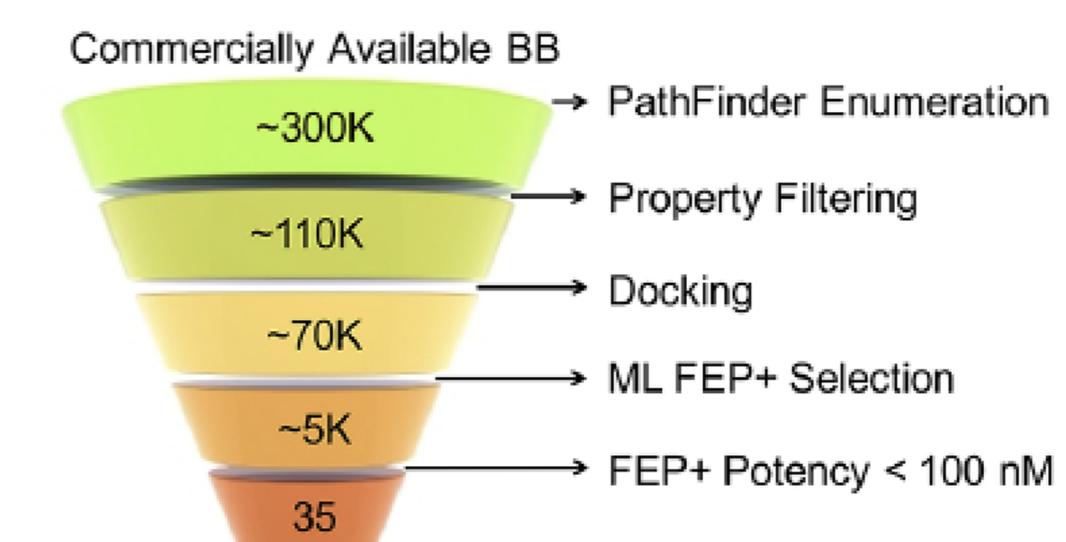
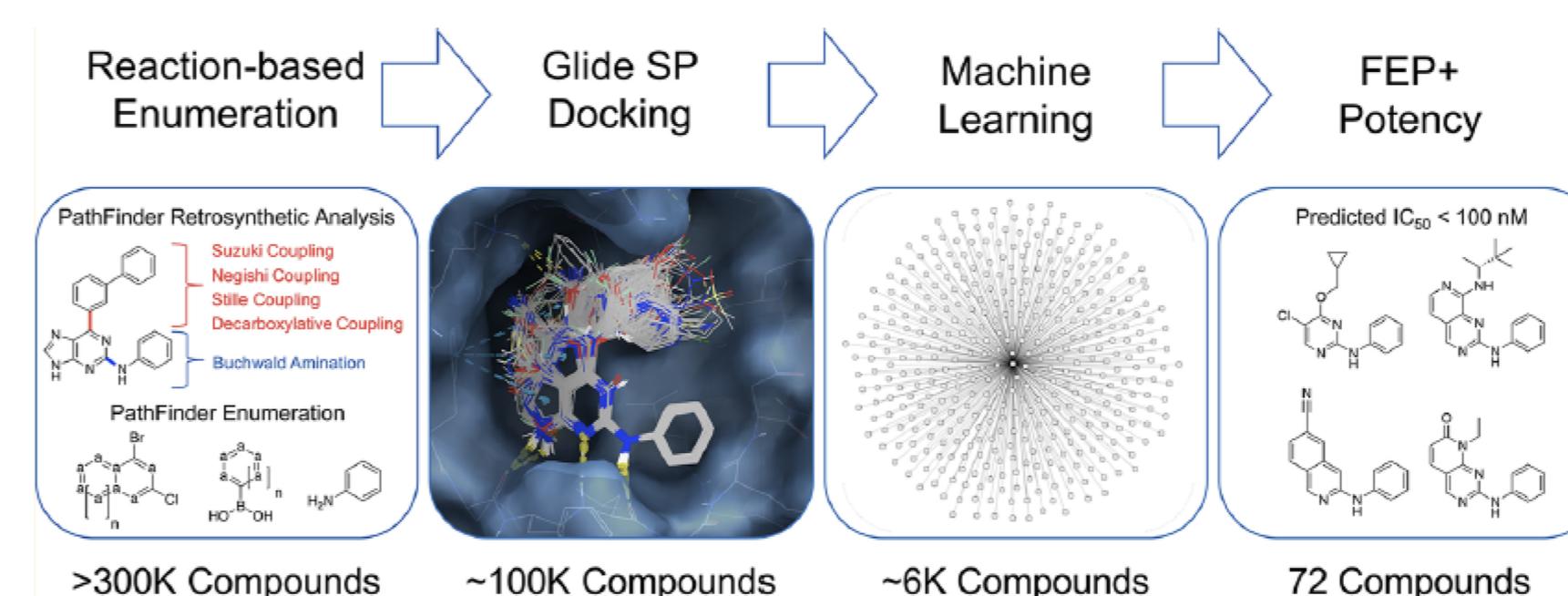
INFINITE ON-DEMAND SCALABILITY

Infinite on-demand scalability of RDKit-powered Schrodinger Platform using the latest Kubernetes container-orchestration technologies. High throughput performance is achieved by running a microservices based LiveDesign architecture powered by Apache Kafka and RDKit.



INTEGRATING RDKIT INTO SCHRÖDINGER RESEARCH

Reaction-Based Enumeration, Active Learning, and Free Energy Calculations To Rapidly Explore Synthetically Tractable Chemical Space and Optimize Potency of Cyclin-Dependent Kinase 2 Inhibitors

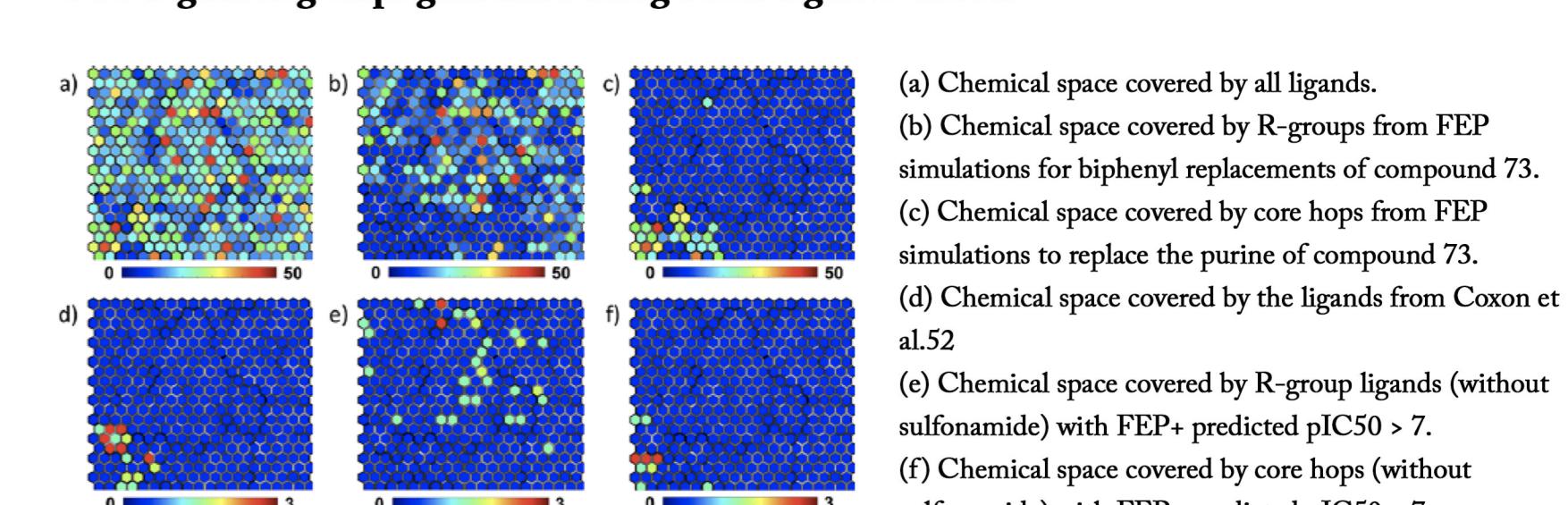


The hit-to-lead and lead optimization processes usually involve the design, synthesis, and profiling of thousands of analogs prior to clinical candidate nomination. A hit finding campaign may begin with a virtual screen that explores millions of compounds, if not more. However, this scale of computational profiling is not frequently performed in the hit-to-lead or lead optimization phases of drug discovery. This is likely due to the lack of appropriate computational tools to generate synthetically tractable lead-like compounds in silico, and a lack of computational methods to accurately profile compounds prospectively on a large scale. Recent advances in computational power and methods provide the ability to profile much larger libraries of ligands than previously possible. Herein, we report a new computational technique, referred to as "PathFinder", that uses retrosynthetic analysis followed by combinatorial synthesis to generate novel compounds in synthetically accessible chemical space. In this work, the integration of PathFinder-driven compound generation, cloud-based FEP simulations, and active learning are used to rapidly optimize R-groups, and generate new cores for inhibitors of cyclin-dependent kinase 2 (CDK2). Using this approach, we explored >300 000 ideas, performed >5000 FEP simulations, and identified >100 ligands with a predicted $IC_{50} < 100$ nM, including four unique cores. To our knowledge, this is the largest set of FEP calculations disclosed in the literature to date. The rapid turnaround time, and scale of chemical exploration, suggests that this is a useful approach to accelerate the discovery of novel chemical matter in drug discovery campaigns.

R-Groups Identified with PathFinder Combined with ML/FEP Workflow

R ₁	R ₂										
	C12	C13	C73	C2	C3	2	3	4	5	6	7
H	H										
SO ₂ NH ₂		SO ₂ NH ₂		H	SO ₂ NH ₂						
Exp. pIC ₅₀	4.21	5.82	7.36	6.01	8.30	ND	ND	ND	ND	ND	ND
LE	0.36	0.40	0.31	0.34	0.40	ND	ND	ND	ND	ND	ND
FEP+ pIC ₅₀ *	4.87	6.14	7.31	5.72	7.36	7.81	7.76	7.39	7.36	7.11	6.41
LE _{FEP+}	0.42	0.42	0.31	0.33	0.36	0.37	0.38	0.37	0.36	0.37	0.36

Self-organizing maps generated using Schrödinger's Canvas



REFERENCE

Kyle D. Konze, Pieter H. Bos, Markus K. Dahlgren, Karl Leswing, Ivan Tubert-Brohman, Andrea Bortolato, Braxton Robbason, Robert Abel, and Sathesh Bhat