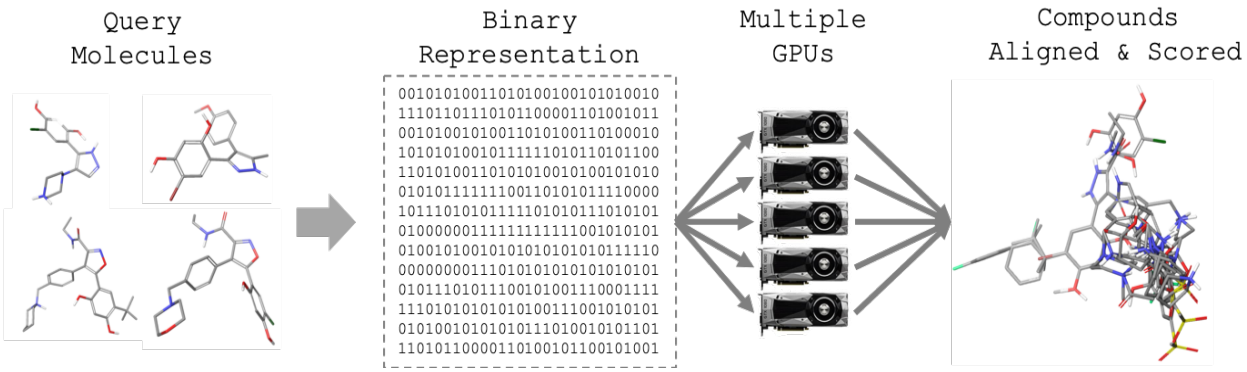
Performing Hit Identification and Lead Optimization at Very Large Scale

Dan Cannon

Schrödinger GmbH, Q7 23, 68161, Mannheim, Germany



It has been estimated there are around 10^60 potential organic molecules with MW < 500, considering only basic structural rules, yet traditional virtual screening decks include up to tens of millions of compounds and design-make-test-analyze (DMTA) cycles in lead optimization generally evaluate tens to low thousands of compounds. Bridging these gaps are on-demand synthesizable screening libraries and synthetically tractable enumerated libraries on the order of millions to low billions of compounds. To virtually screen such massive libraries of compounds using 3D information within project timelines of drug discovery projects, GPU-based shape screening methods have demonstrated considerable utility. They often outperform pharmacophore methods while being easier to apply to more diverse collections of known hits. In this presentation we introduce a GPU enabled algorithm based on the previously described CPU implementation [1] and describe how it can impact lead discovery projects using your existing local or cloud GPU resources, even when screening >1 billion compounds.

[1] G.M. Sastry, S.L. Dixon, W.J. Sherman, *J. Chem. Inf. Model.*, **2011**, *51*, 2455-66.