Design in 2D, model in 3D: Live 3D pose generation from 2D sketches

P. Tosco, M. D. Mackey

Cresset, Cambridgeshire, UK paolo@cresset-group.com https://cresset-group.com

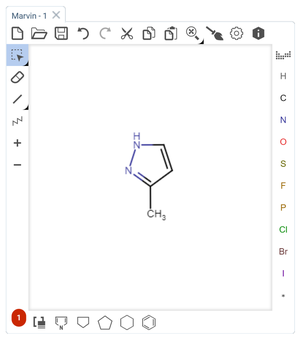
Idea

The main idea which sparked the *grow3D* project was: how nice would it be to draw a molecule in your favorite 2D sketcher and simply see it grow sensibly within the binding site in the 3D viewport, naturally obeying to the constraints imposed by the active site, with little or no intervention from the user?

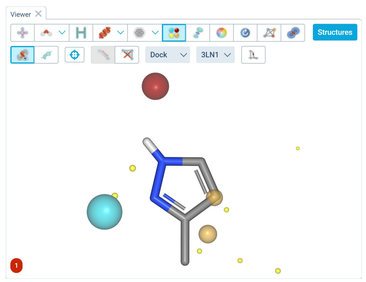
How hard can it be?

In a nutshell:

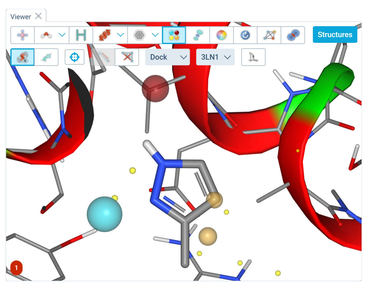
- You sketch some starter structure in the 2D editor (or upload it from your hard drive, or copy/paste it from any source, such as your ELN)



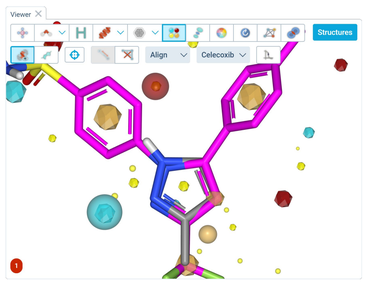
- The largest 2D fragment is popped to a 3D conformation…



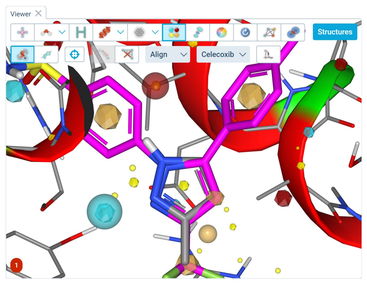
- … and docked into the protein's active site



- …or 3D field/shape aligned against a reference…

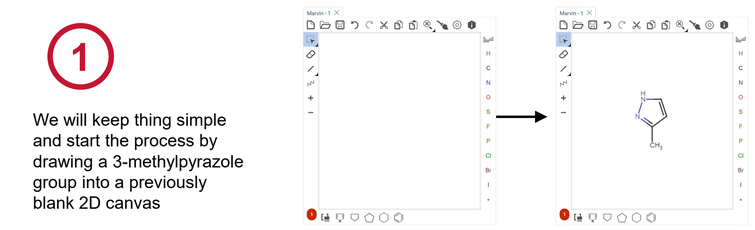


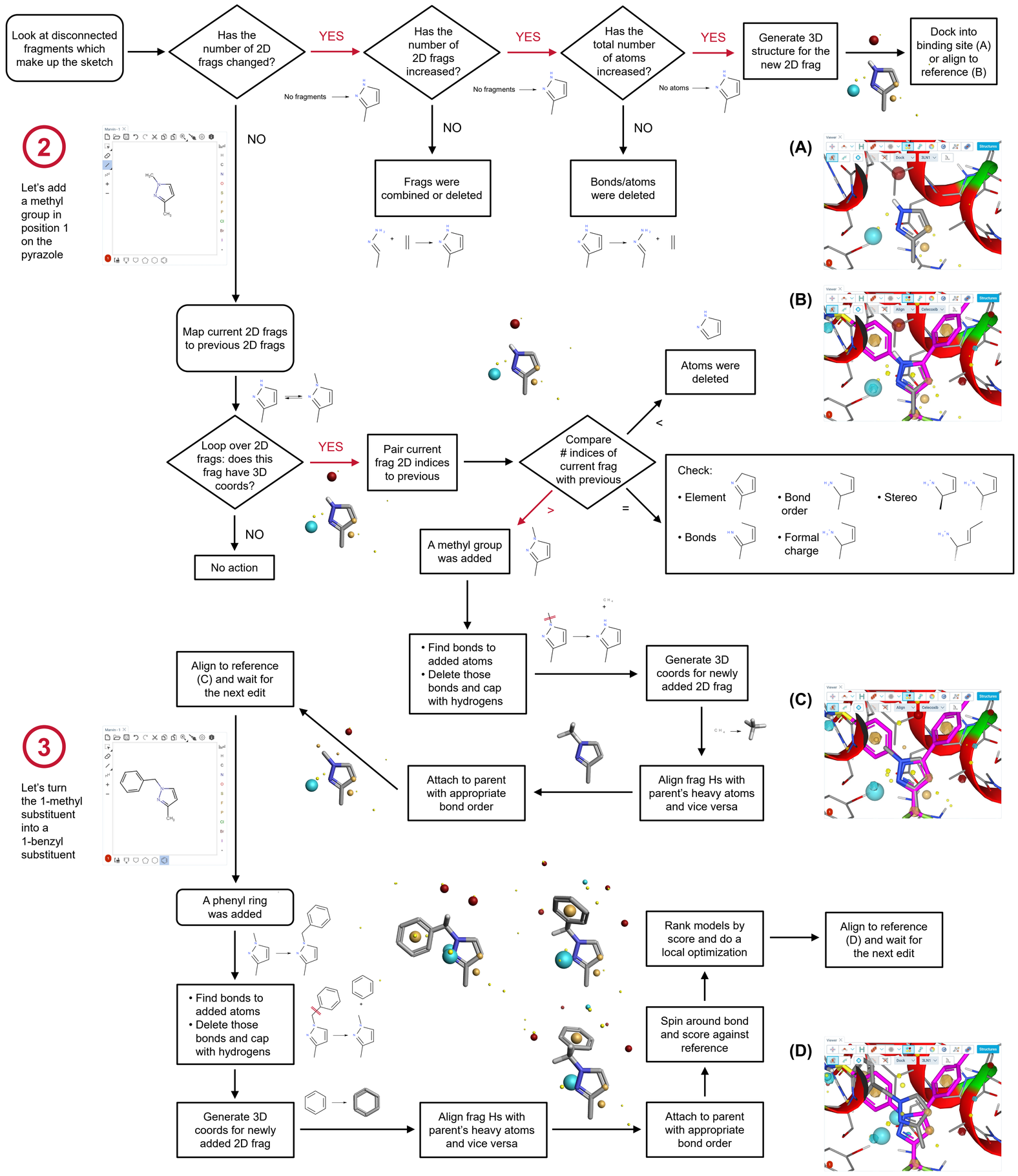
- …using the protein as excluded volume (if available)



- then the *grow3D* process begins: changes to the 2D sketch are applied to the 3D viewport intelligently, in order to avoid a full re-docking/re-alignment and allow real-time 3D feedback on the 2D design

Flow diagram for *grow3D*

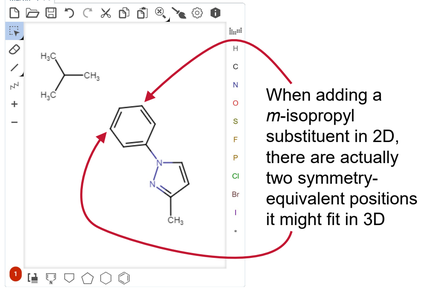




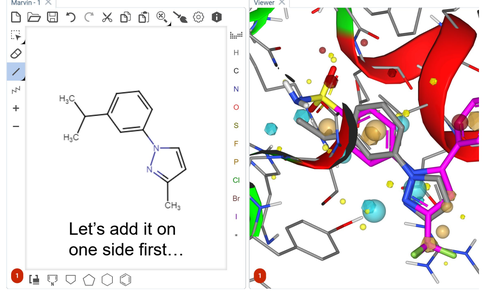
RDKit
Open-Source Cheminformatics and Machine Learning

The RDKit is extensively used throughout the *grow3d* workflow for substructure matching, MCS finding, file parsing, etc.

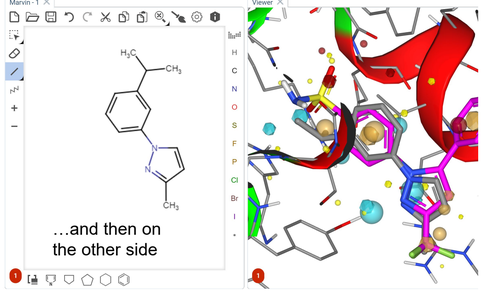
The devil is in the details (1): symmetries



When adding a *m*-isopropyl substituent in 2D, there are actually two symmetry-equivalent positions it might fit in 3D



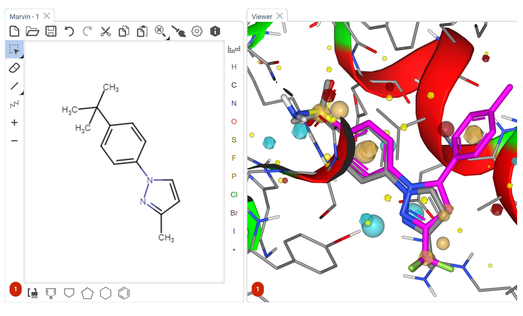
Let's add it on one side first…



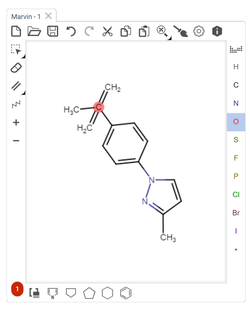
…and then on the other side

No matter which side the isopropyl substituent is drawn in 2D, the *grow3D* algorithm chooses the one which fits the active site best

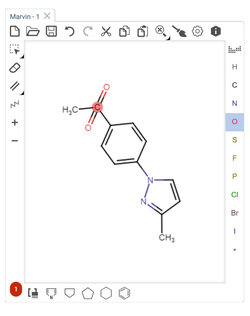
The devil is in the details (2): chemically invalid states



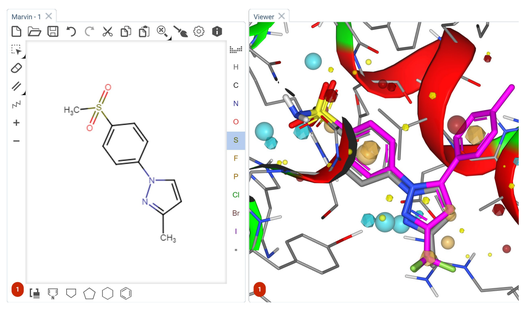
When we draw a sulfone-containing functional group in 2D, we tend to go through a number of chemically invalid states. We keep adding methyl groups until we get to a *t*-butyl, which is chemically valid…



…but then we may simply keep adding bonds creating a chemically invalid *gem-bis*-methylidene intermediate…



…then turn it into an equally invalid *gem*-dicarbonyl beast…



…eventually getting, at last, to a chemically valid sulfone. *grow3D* copes with the invalid intermediates and only converts to 3D the chemically valid molecules.

Conclusions

- *grow3D* is an algorithm which generates sensible 3D poses in response to edits to a 2D sketch
- Poses are scored against their ability to fit a binding site and/or mimic electrostatics and shape of a reference ligand
- The algorithm is enough quick and accurate to enable real time assessment of design ideas
- *grow3D* is part of Torx™, the new web-based Cresset platform for collaborative small molecule discovery chemistry