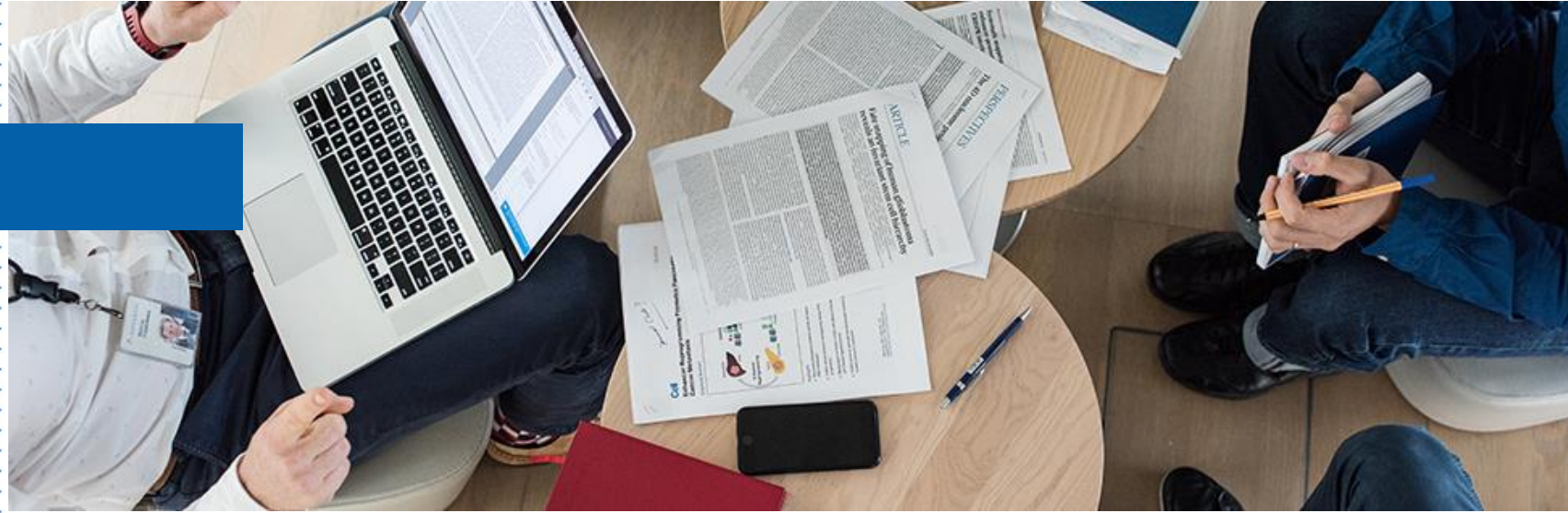
GDC-CADD



An RDKit-based JS component for molecule visualization

Paolo Tosco, Nico Pulver, Nik Stiefl

10th RDKit UGM

October 15, 2021

NOVARTIS | Reimagining Medicine

Agenda

Motivation

Implementation

Live demo

Outlook

Motivation

Visualize 2D molecules in the browser: server-based

- At NIBR we have several web apps that require 2D molecule visualization
- Currently this is handled by a React component
- The web client requests a 2D layout to the server (SMILES or MDL molblock)
- The server generates the 2D layout and sends it back to the web client as a PNG image
- A number of configuration options is available
  - Width, height
  - Align the layout to a scaffold (SMILES or MDL molblock)
  - Font, coloring, etc.

Visualize 2D molecules in the browser: client-based

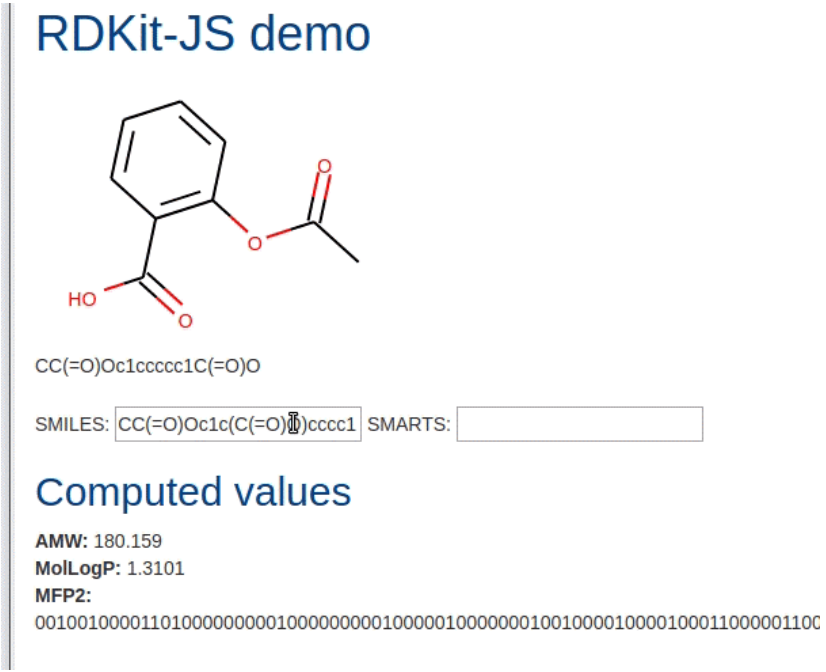
- We would like to switch from a server-based app to a client-based app for performance reasons
- We would like to move from a static PNG image to a vector image that can be easily rescaled without losing definition (HTML5 canvas, SVG)
- We would like the app to be written in vanilla JavaScript to be independent from a specific technology (i.e., not be bound to React)
- Wrappers for different technologies can be easily built around this low-level app

RDKit in the browser

- 2016: at the UGM Guillaume Godin presented the first JavaScript port of RDKit through `emscripten`
- 2019: Greg Landrum introduced the new RDKit JavaScript wrappers

  http://rdkit.blogspot.com/2019/11/introducing-new-rdkit-javascript.html

  A subset of RDKit was ported to WebAssembly through `emscripten`

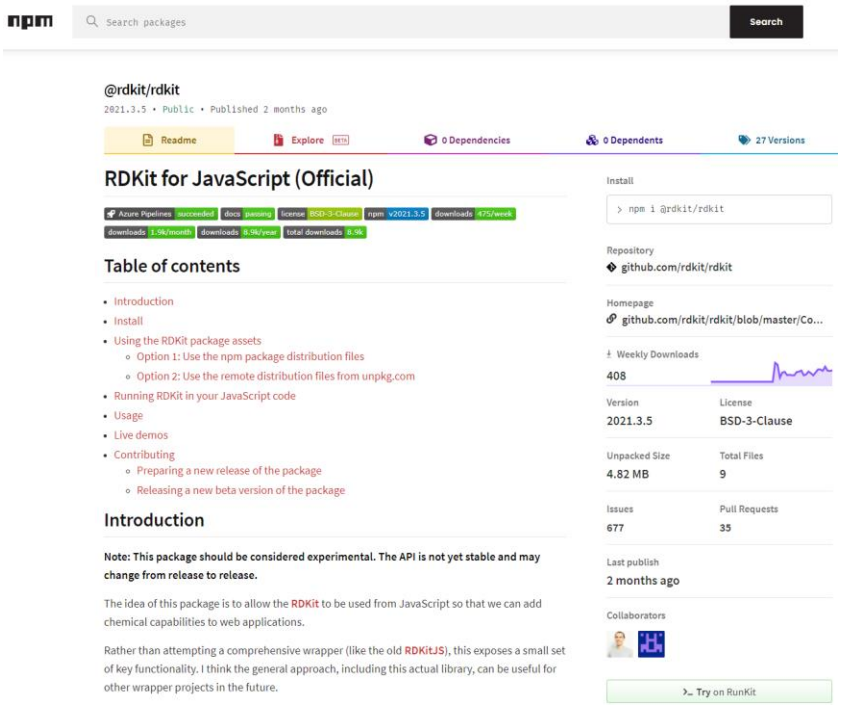


RDKit in the browser

- 2021: Michel Moreau made the RDKit JavaScript MinimalLib available as a `npm` package

  https://www.npmjs.com/package/@rdkit/rdkit

  This further lowers the barrier to access RDKit functionality in the browser

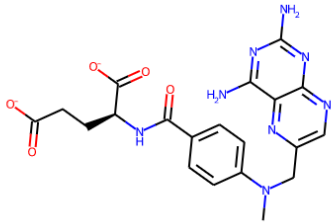


Web apps using RDKit MinimalLib

- Building web apps dealing with chemistry becomes easier

  http://chembl.blogspot.com/2021/03/target-predictions-in-browser-with.html

Multitask Target prediction with RDKit MinimalLib (JS) and ONNX.js



SMILES

CN(Cc1cnc2nc(N)nc(N)c2n1)c1ccc(C(=O)N[C@@H](CCC(=O)[O-])C(=O)[O-])cc1

Predictions

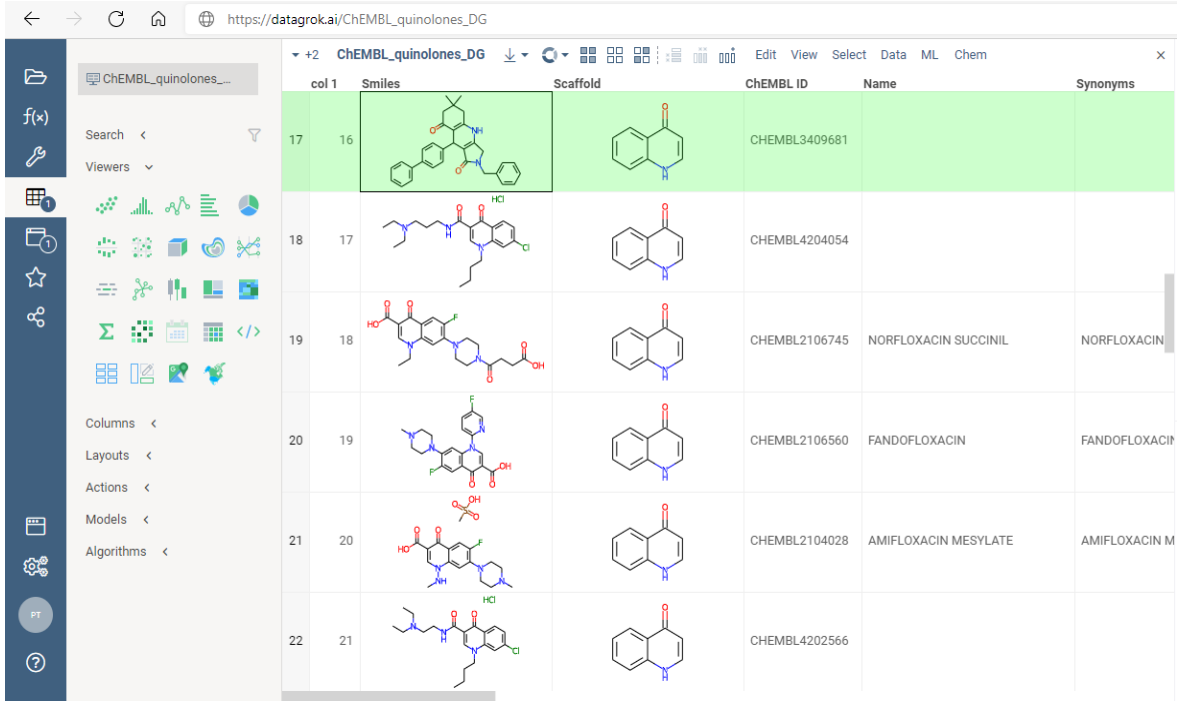
Search

Target ChEMBL id	proba ↓
CHEMBL4409	0.9893653988838196
CHEMBL4191	0.9790000319480896
CHEMBL1744525	0.9765774607658386

Web apps using RDKit MinimalLib

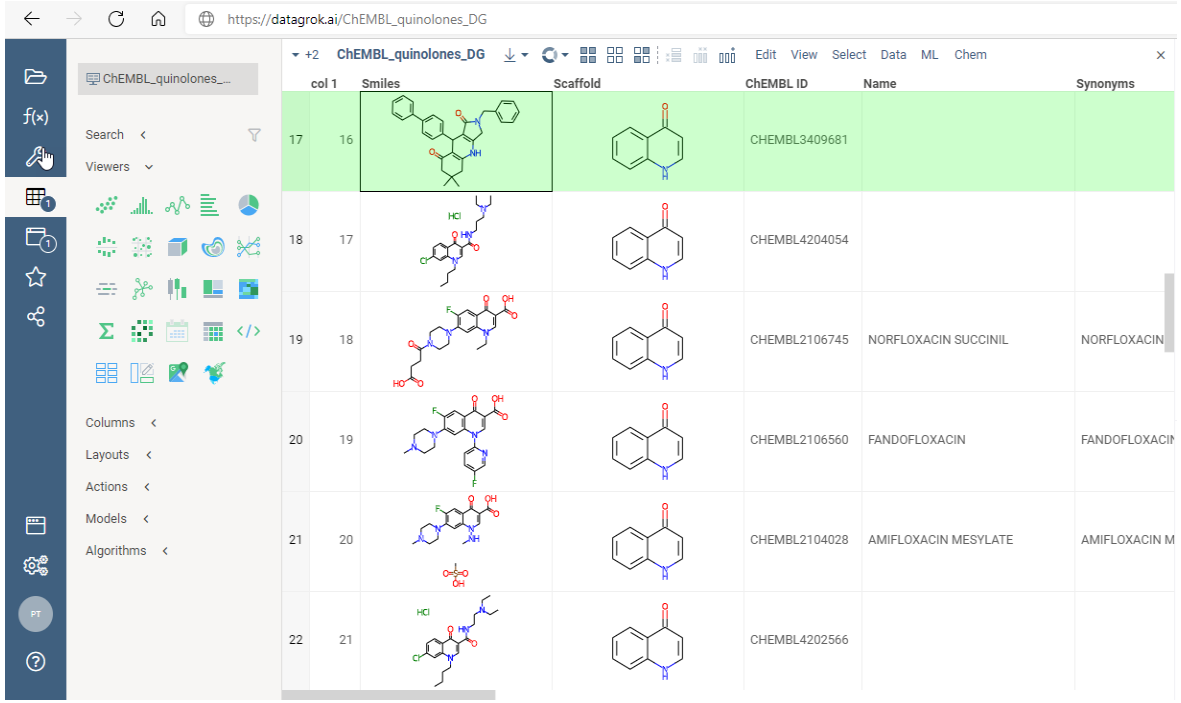
- Datagrok
- Web-based data analysis tool
- Uses RDKit MinimalLib for 2D molecule visualization

https://datagrok.ai



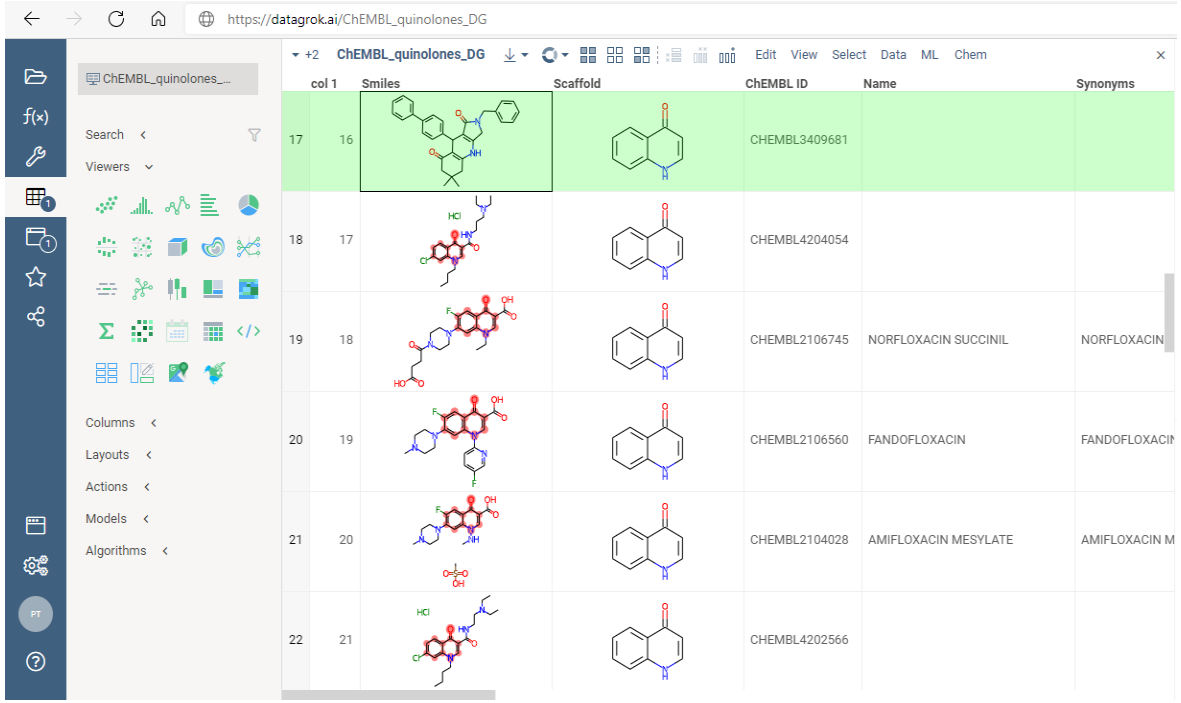
Web apps using RDKit MinimalLib

- Molecules can be aligned to a scaffold if the scaffold matches



Web apps using RDKit MinimalLib

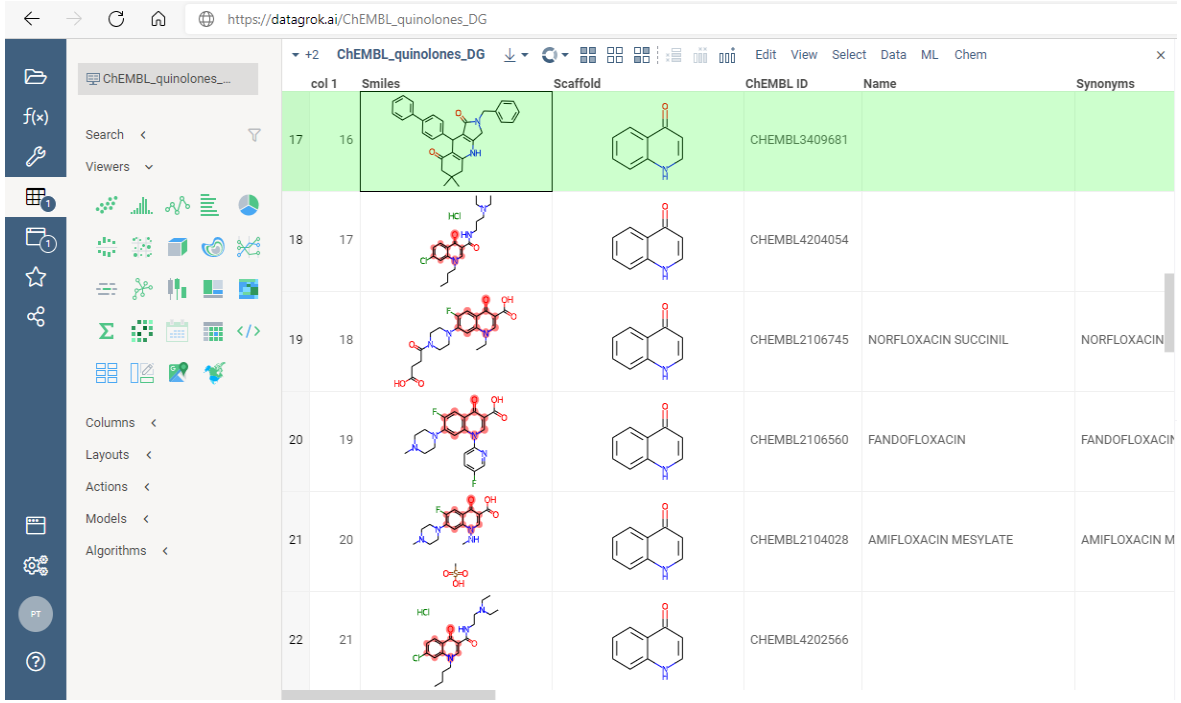
- Molecules can be aligned to a scaffold if the scaffold matches
- The scaffold atoms can be highlighted



NOVARTIS | Reimagining Medicine

Web apps using RDKit MinimalLib

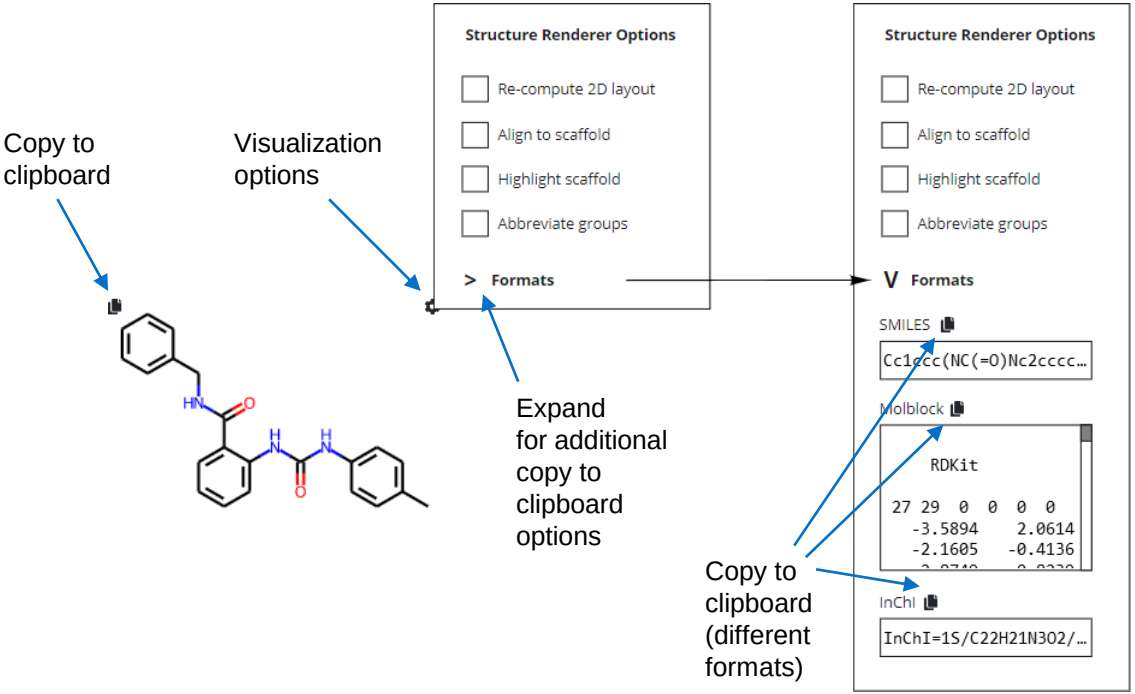
- Have better granularity
- Align or highlight the scaffold only for selected molecules
- Communicate the per-molecule setting to the parent app for storage



Implementation

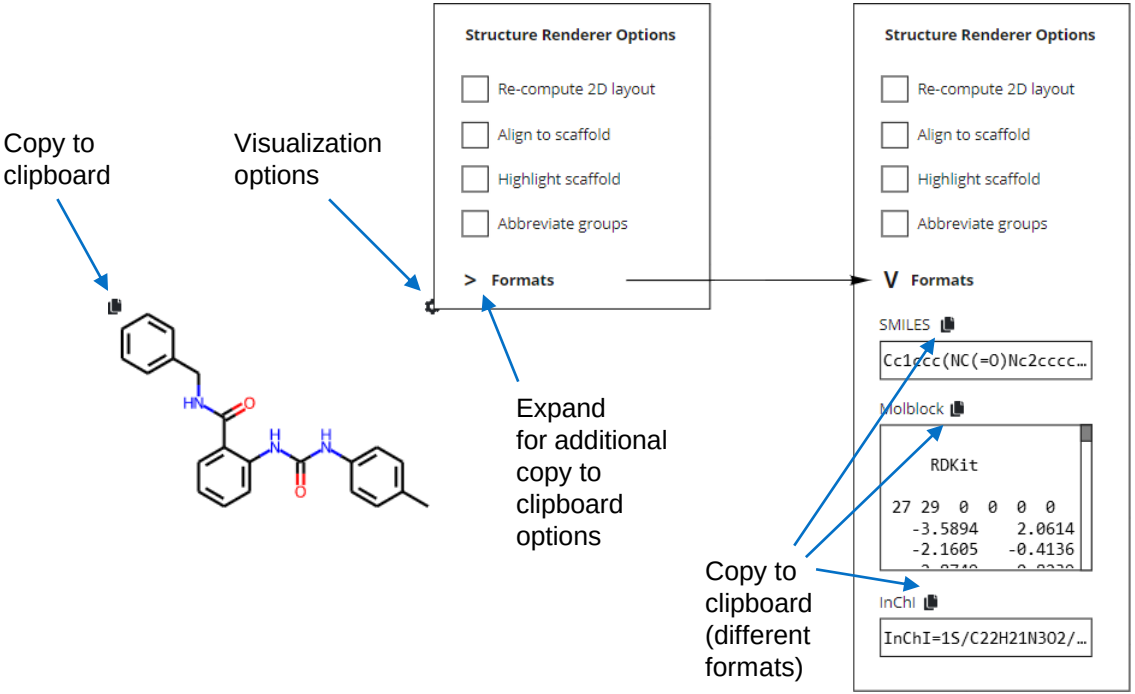
StructureRenderer.js design

- Each 2D layout is decorated with two icons that appear upon hovering
- Multiple checkboxes with visualization options
- Multiple formats can be copied to clipboard



StructureRenderer.js design

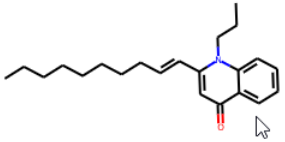
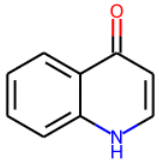
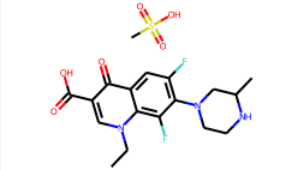
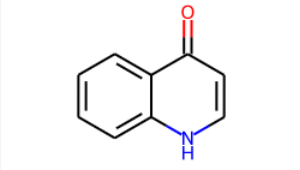
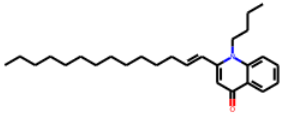
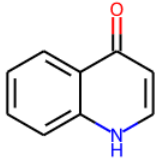
- Having no JS dependencies other than the RDKit MinimalLib, this should work in something as simple as, e.g., a Jupyter Notebook



Pandas DataFrame in Jupyter

- Most of us are familiar with the visualization of a `pandas DataFrame` containing RDKit molecules in Jupyter

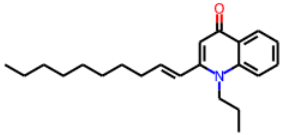
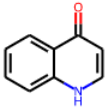
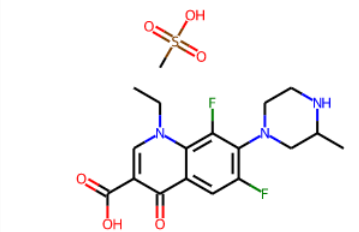
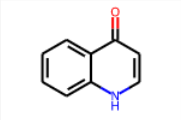
```
PandasTools.ChangeMoleculeRendering(df_quinolones)
show(df_quinolones)
```

Smiles	Scaffold	ChEMBL ID	Name	Synonyms
		CHEMBL1782626	NaN	NaN
		CHEMBL2106827	LOMEFLOXACIN MESYLATE	LOMEFLOXACIN MESYLATE|SC-47111B
		CHEMBL1782648	NaN	NaN

Pandas DataFrame in Jupyter with StructRenderer.js

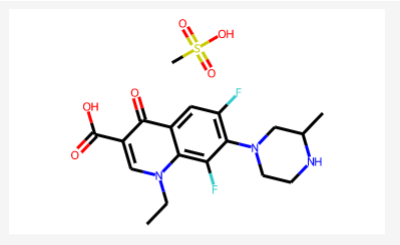
- Where's the difference?
- They look pretty much the same

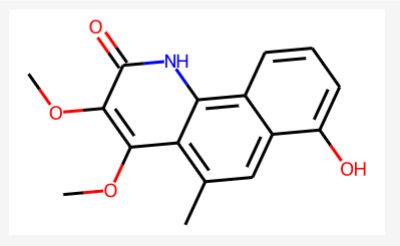
```
PandasTools.ChangeMoleculeRendering(df_quinolones)
show(df_quinolones)
```

	Smiles	Scaffold	ChEMBL ID	Name	Synonyms
			CHEMBL1782626	NaN	NaN
			CHEMBL2106827	LOMEFLOXACIN MESYLATE	LOMEFLOXACIN MESYLATE|SC-47111B

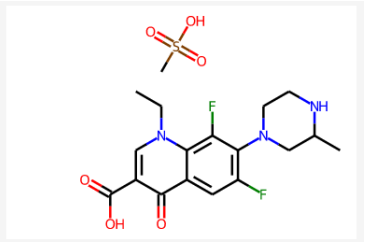
Molecules are rotated to look prettier

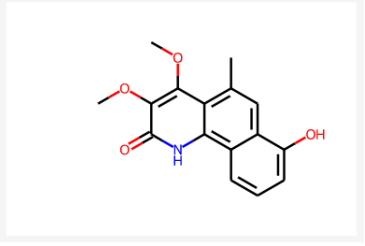
RDKit aligns principal axes to Cartesian axes to exploit available space at best



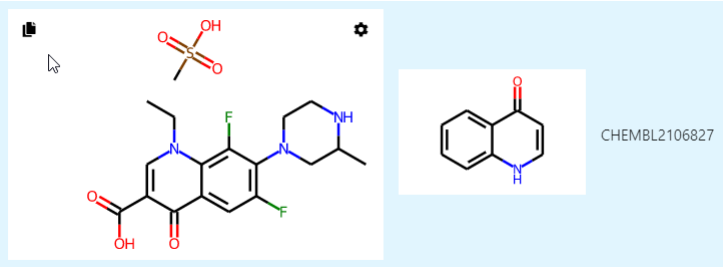


`StructureRenderer.js` canonicalizes 2D coordinates and then rotates molecules such that most bonds have a 30-degree angle with the X axis

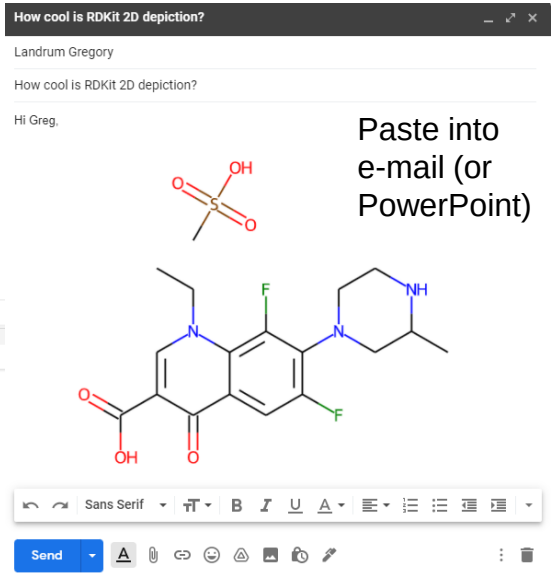




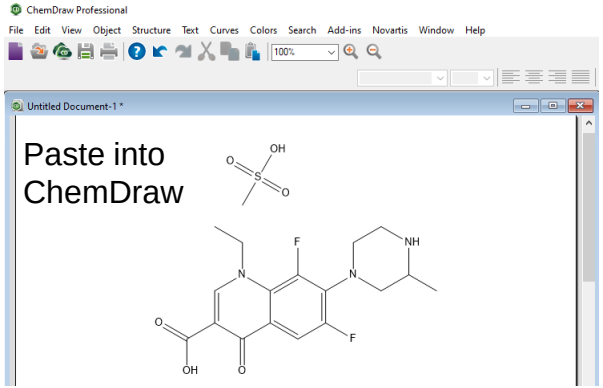
Additional controls appear on hover



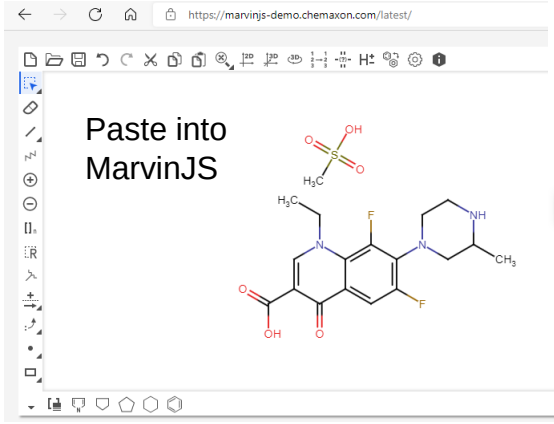
Copy to clipboard from Jupyter and paste into different tools



Paste into e-mail (or PowerPoint)



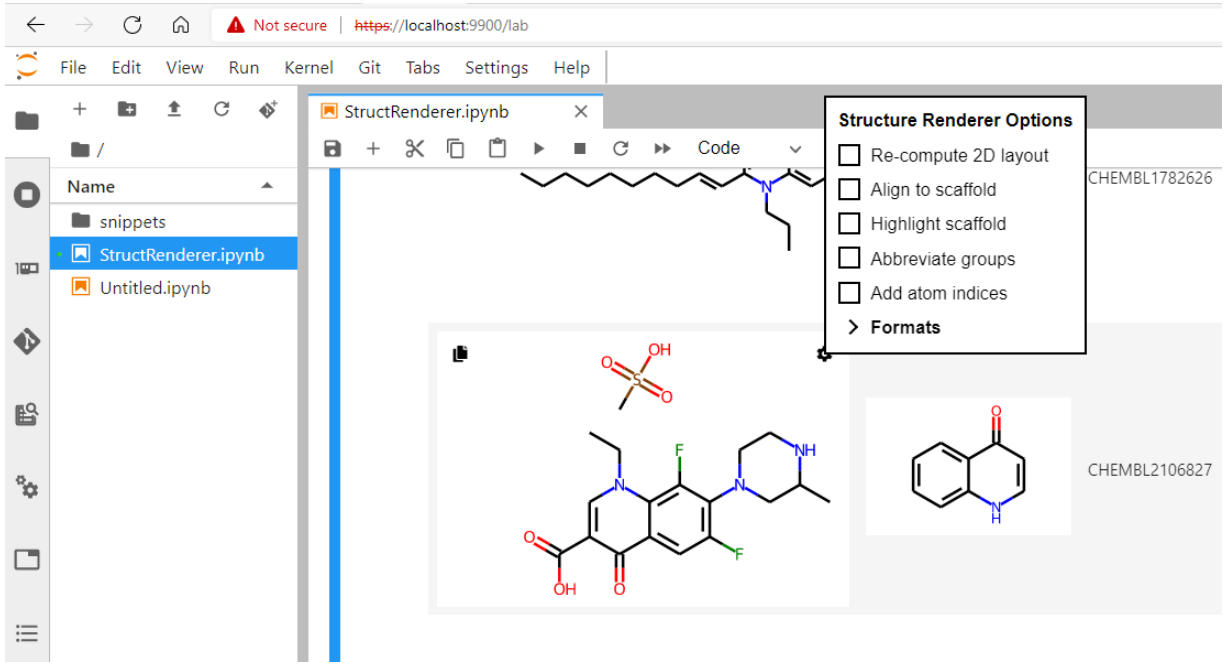
Paste into ChemDraw



Paste into MarvinJS

Tailor the appearance of the 2D layout

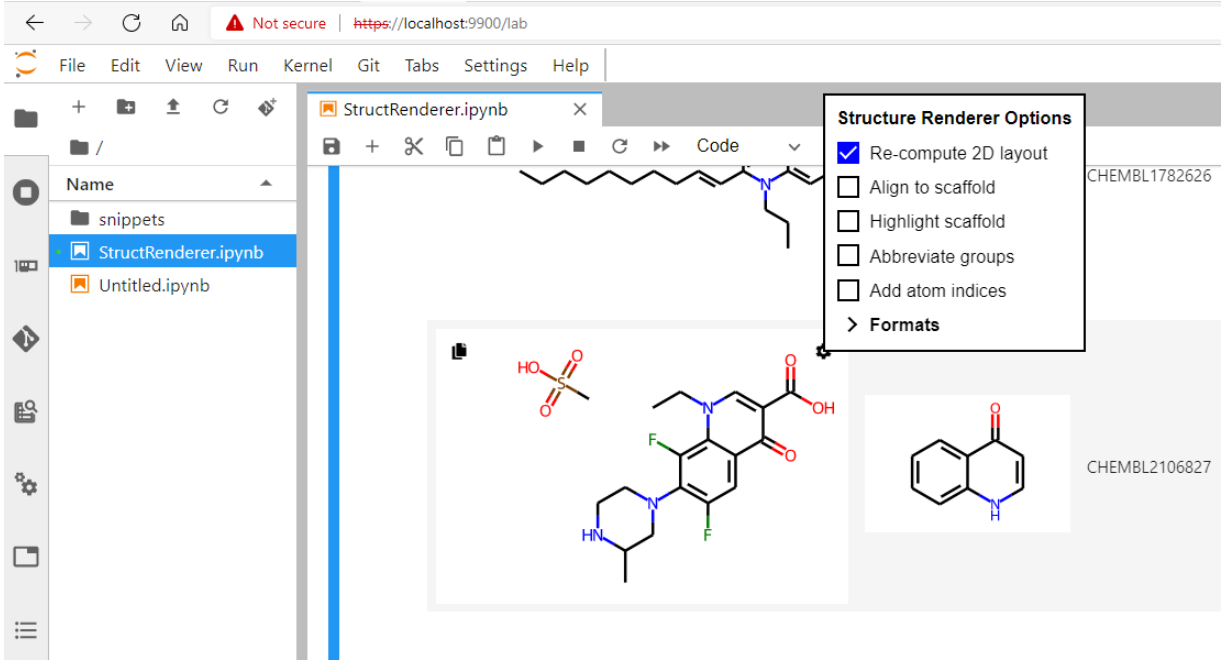
- The appearance of the 2D layout can be tailored through multiple, non-exclusive options



NOVARTIS | Reimagining Medicine

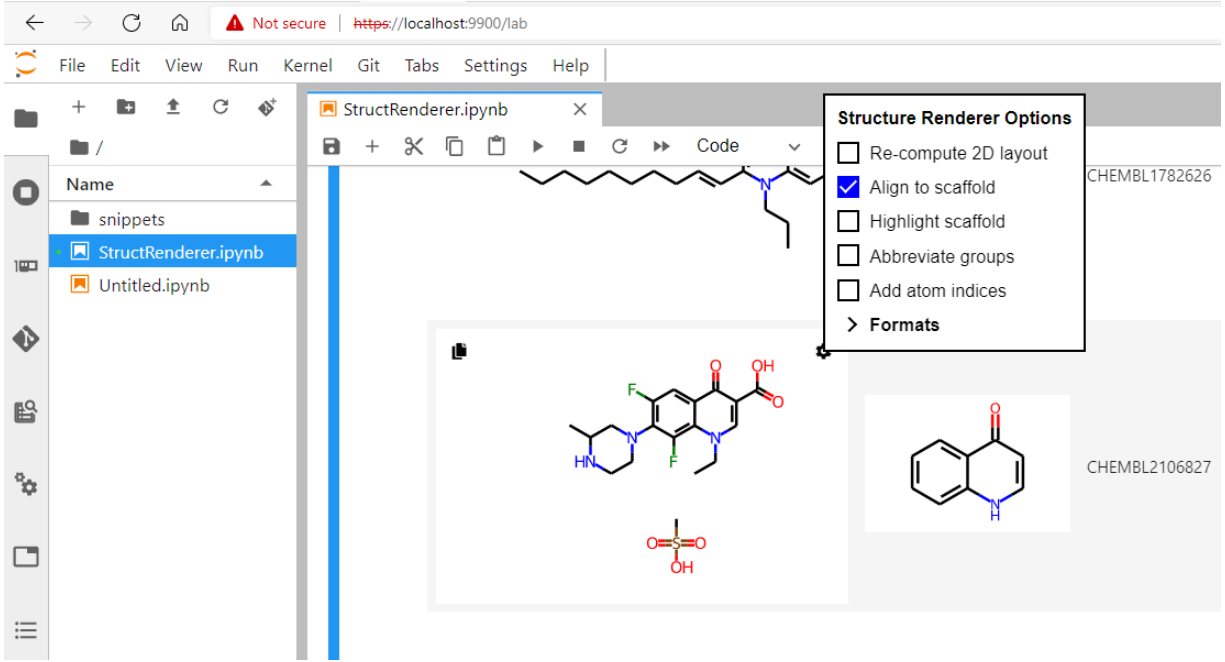
Re-compute 2D layout

- The layout can be re-computed using CoordGen (useful in case of, e.g., macrocycles or poor initial coordinates if the molecule was generated from a MDL molblock)



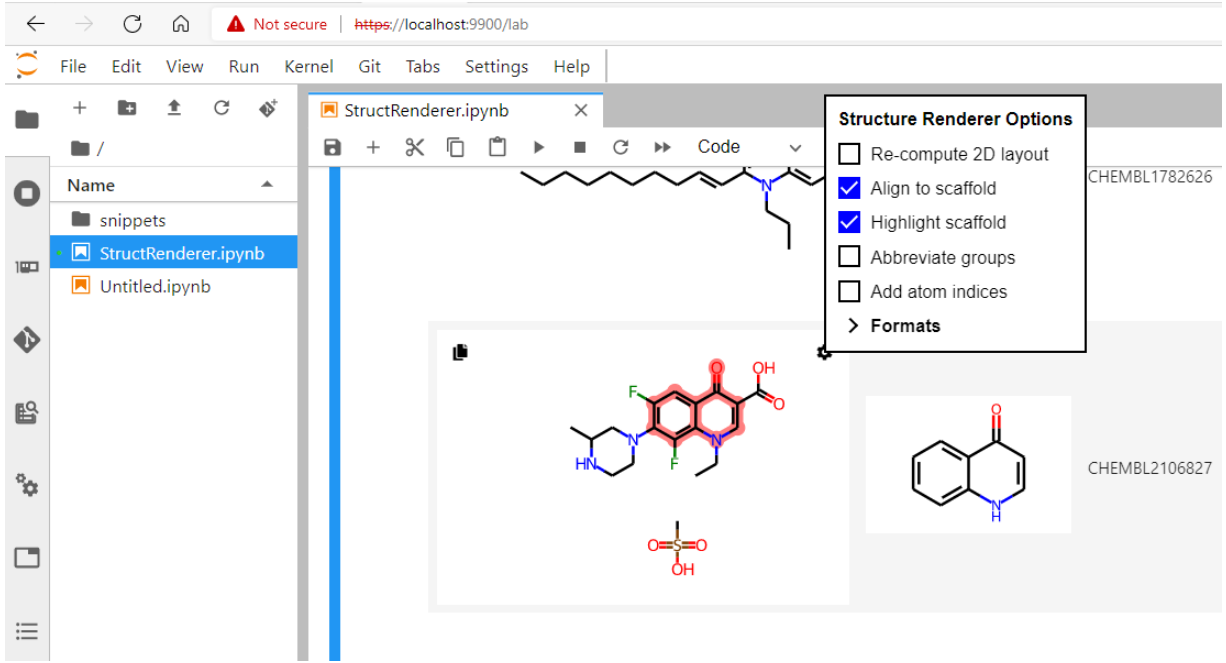
Align to scaffold

- The layout can be aligned to a scaffold (as previously seen in the Datagrok example)



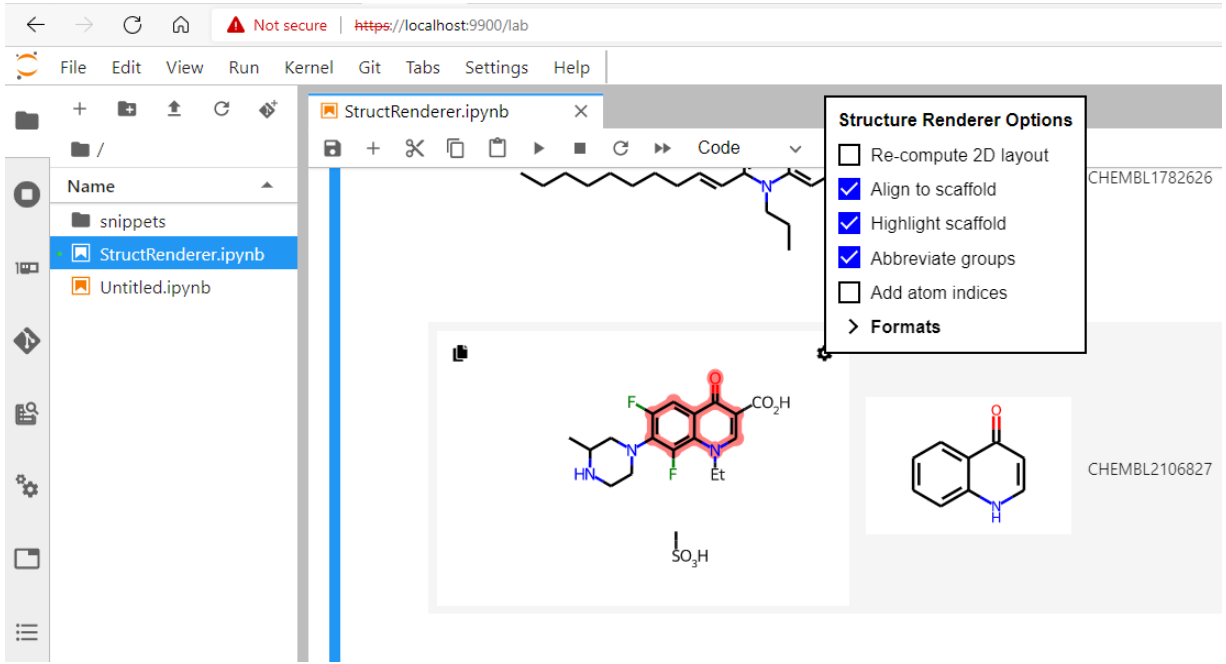
Align to scaffold with highlighting

- We can also highlight the scaffold atoms in the molecule



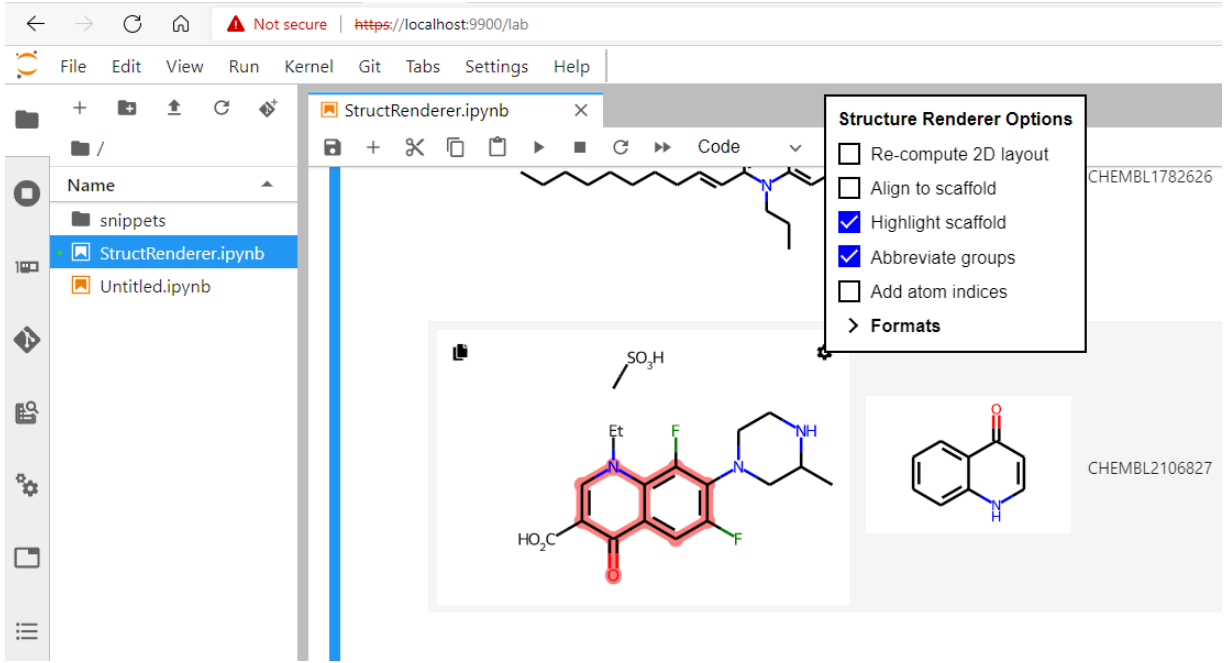
Display abbreviated functional groups

- Functional groups can be displayed in their abbreviated form



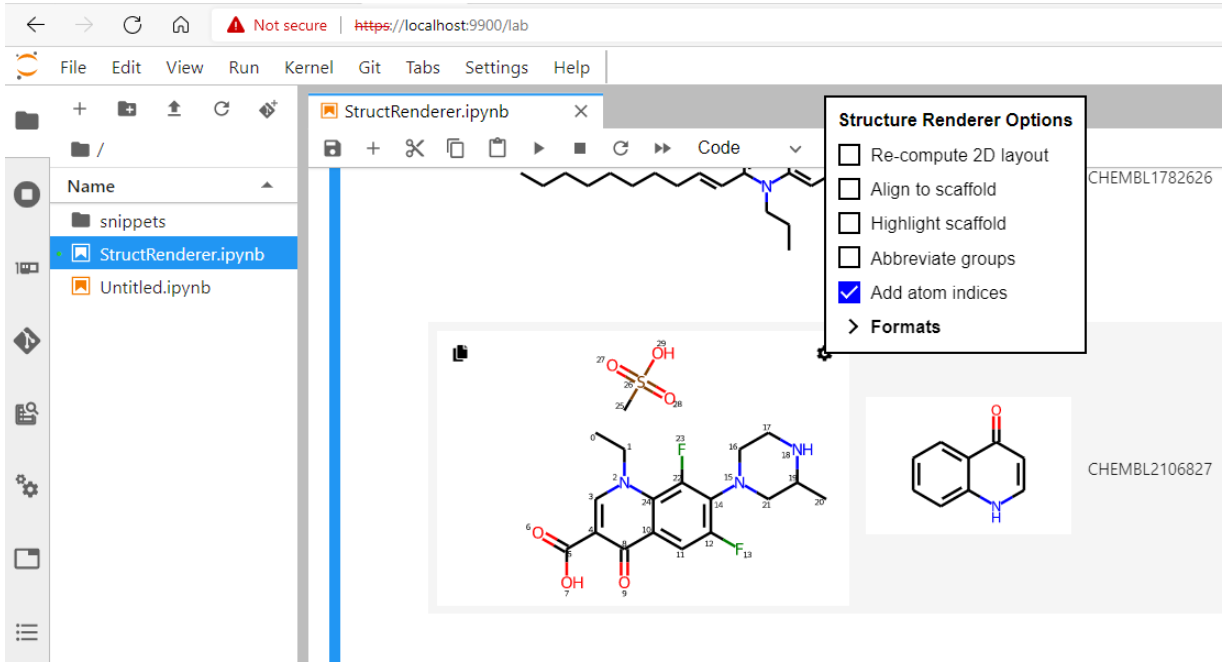
Unchecking options works

- Unchecking individual options triggers a layout update to only apply the relevant settings



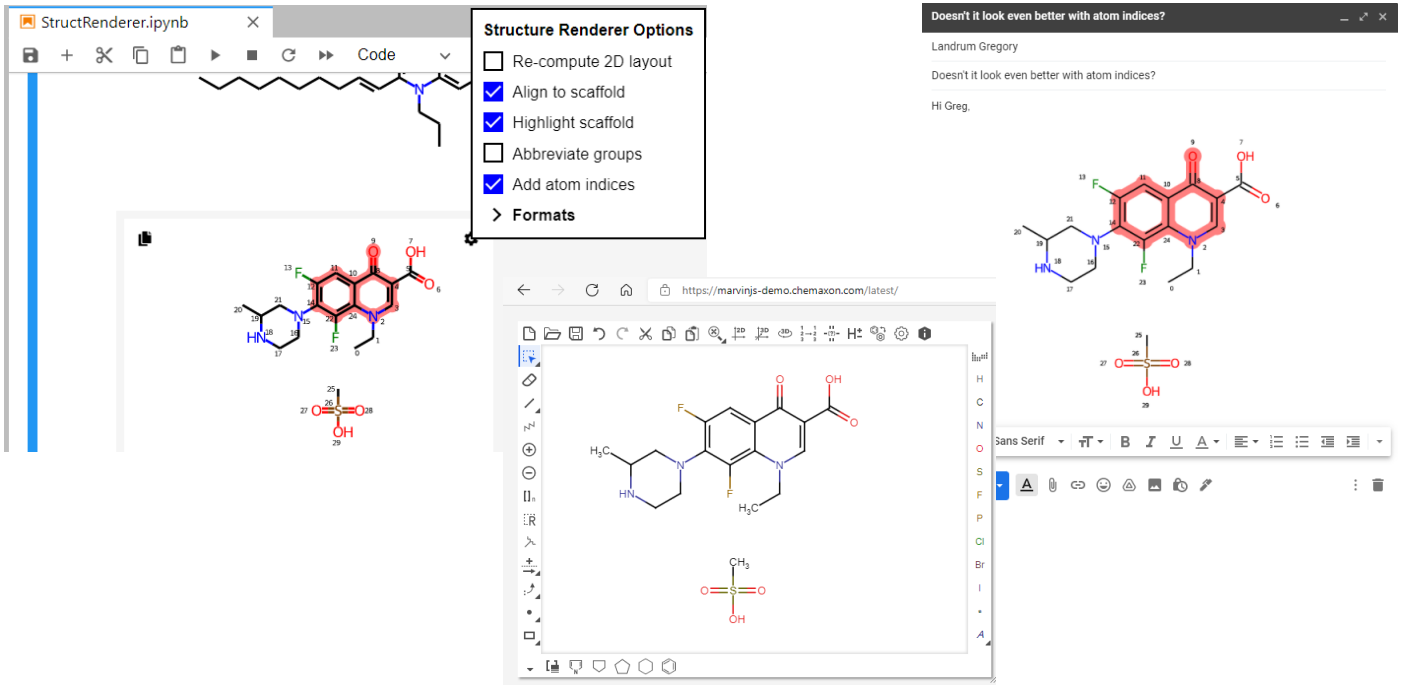
Bonus option: add atom indices

- Compared to the original design I added the "Add atom indices" option that I find very handy when working with molecules in Jupyter



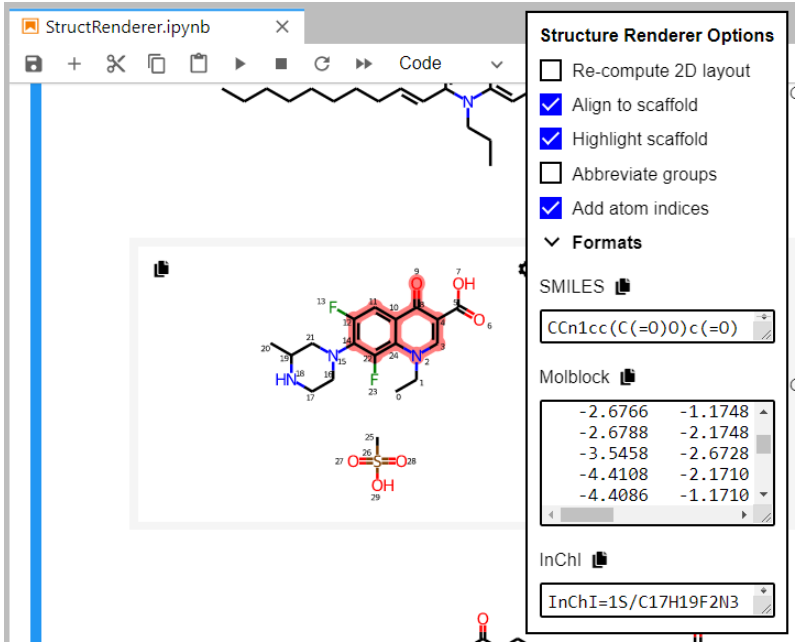
Copy follows WYSYWYG paradigm

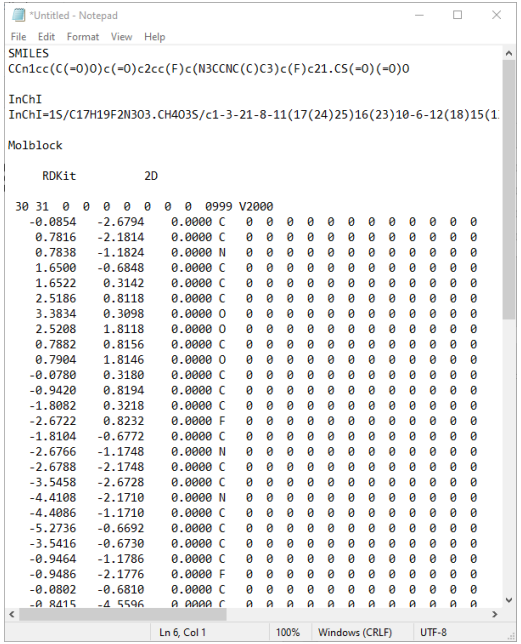
- The copy functionality puts on the clipboard the same image/ coordinates as shown in Jupyter



Additional paste formats available

- Expanding "Formats" gives access to further paste formats (SMILES, InChI, MDL molblock)





Live demo

Outlook

Outlook

- We have realized a prototype vanilla JS app to support RDKit-based 2D molecule visualization in the browser
- The only dependency is on RDKit MinimalLib
- This makes the app lightweight and allows adoption in many different contexts
- We are planning to release the app as open-source code on GitHub

Acknowledgments

- NIBR
  - Grégori Gerebtzoff
  - Andreas Gasser
  - Nik Clare
  - Robert Strulak
  - Riccardo Vianello
  - Jimmy Kromann
  - [...]
- EPAM
  - Dmitrii Makarkin
  - Ekaterina Stepanova
  - Anastasia Lebedeva
  - [...]
- RDKit
  - Greg Landrum
  - David Cosgrove
  - Brian Kelley
  - [...]
- Datagrok
  - Dan Skatov
  - Alex Paramanov
  - Andrew Skalkin
  - [...]
- Schrödinger
  - Nic Zonta
  - Rachel Walker
  - Ricardo Rodríguez
  - [...]

Thank you