A New Application for Discussing Chemical Ideas

By Andras Stracz, Ivan Solt, ChemAxon

Abstract:

In this article, we look at a new application called Marvin Live, developed by ChemAxon for online collaboration. Marvin Live offers a novel approach to designing molecules and leading chemical discussions among remote participants.

We will describe the pharmaceutical industry and academic needs that led to these developments and the workflows made possible, all the way from simple chemical reaction drawing tasks, to connecting globally distributed research teams for project design meetings. We look at use cases that enable their use for tutoring and mentoring, and their capabilities to automate the otherwise time-consuming aspects of meeting preparations and note-taking.

Background

Commercial, as well as academic researchers have to face the problem of communication between remote participants, as discovery processes get more and more distributed[1]. Design and project meetings involving chemical structures in the discussions can be difficult, unless people taking part sit in the same room and use the same surface for drawing, evaluating, and discussing their ideas.

This applies not only to projects where contract research organizations or consultants are involved, but also when academic research groups, based in different universities are collaborating with each other[2]. In these projects sharing ideas, reflecting on problems and collecting the necessary feedback is essentially a burden.

Similarly, in the case of teaching and tutoring, ad hoc discussions between a mentor and a student can be essential, but take a high activation energy. A professor's office hours are limited; a face to face meeting is not always possible, especially with a group of students.

Over time, commonalities surfaced to how people approach this issue. In lieu of a face to face meeting, one of the most common methods is sharing data via emails either directly or to project/course specific mailing lists. This however comes with the difficulty of forming all relevant information into plain text or email attachments, and each collaborator requiring access to software that handles these attachments.

With the recent increase in availability of free video messaging solutions[3], their use in scientific discussions also increased, either to share the desktop of one researcher, or to show the drawings on a whiteboard using webcams. These methods however make it very difficult for remote collaborators to actively participate, as they need to create local

copies of chemical drawings, before adding an idea of their own, or storing them for reports.

Given these problems, we started to investigate if an alternative solution is possible. The requirements were to:

1) use digital representations of chemical structures in common file formats,
2) enable both data driven and discussion driven sessions, and
3) possess a simple and focused user interface where first time users can be just as effective as others.

Use cases

Based on qualitative feedback we identified two basic types of use in this context and developed Marvin Live to support them.

1) Usage in early phase drug discovery

In the pharmaceutical industry, this application can be used for chemical design and project discussions among collaborators in different physical locations. Individual researchers can collect and refine their ideas in non-shared rooms in advance, where they prepare and refine interesting molecules, eventually sharing them with the team. Medicinal chemists, computational chemists and biochemists can join a single virtual meeting room, to analyze, amend and make decisions about chemical ideas for a given project.

The designs during these discussions are influenced by immediate feedback on the molecules provided in the application. For this, common data sources are:

1) calculated physico-chemical properties and a report on how these change upon modifications,
2) the closest patented molecules or citations from publications,
3) biological data on similar or identical structures,
4) 3D overlays on reference molecules.

A plugin system was developed to address these needs, and several plugins were developed for demonstration purposes in these categories, using ChemAxon's property predictors, EBI's various datasets including SureChemBL[4], ChemBL[5], and UniChem[6]; PubChem Compound[7]; Elsevier's Reaxys[8], eMolecules[9] compound registration systems, electronic lab notebooks, and support for KNIME Server[10] and Pipeline Pilot Server[11]. This plugin system helps researchers to analyze the chemical space they are discussing using their trusted data sources, modify the molecules to achieve the desired properties, and save important milestones for further in-silico analysis and synthesis.

Marvin Live keeps track of authorship of a chemical structure, and any attached comments and metadata used during design. At the end of each meeting, this is used to create reports that assist with further sharing of the status of the project. As various

states of a design process can be retraced, the application can also be used as an idea repository for molecules.

In these cases the layout of the user interface has a medium-sized editor in the center, a list of saved or uploaded molecules on one side, and the plugins providing feedback on the other. Below the editor is a chat area for easy commenting on the structures (*see fig. 1, fig. 2.*).
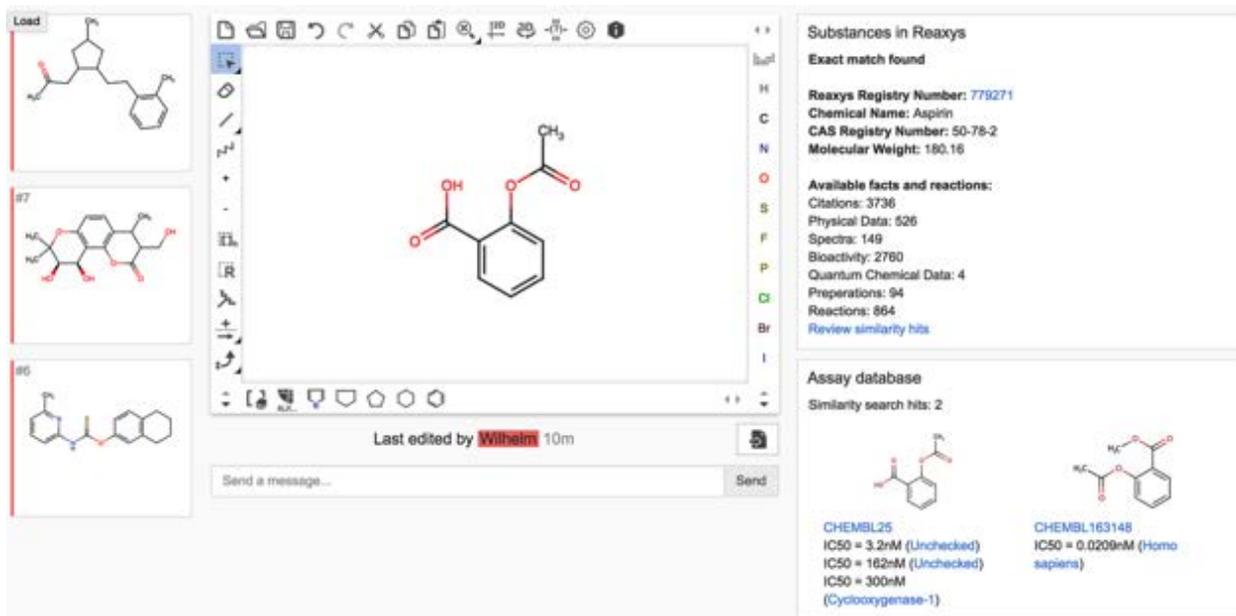

Fig 1: Marvin Live used to mine available chemical and biological data
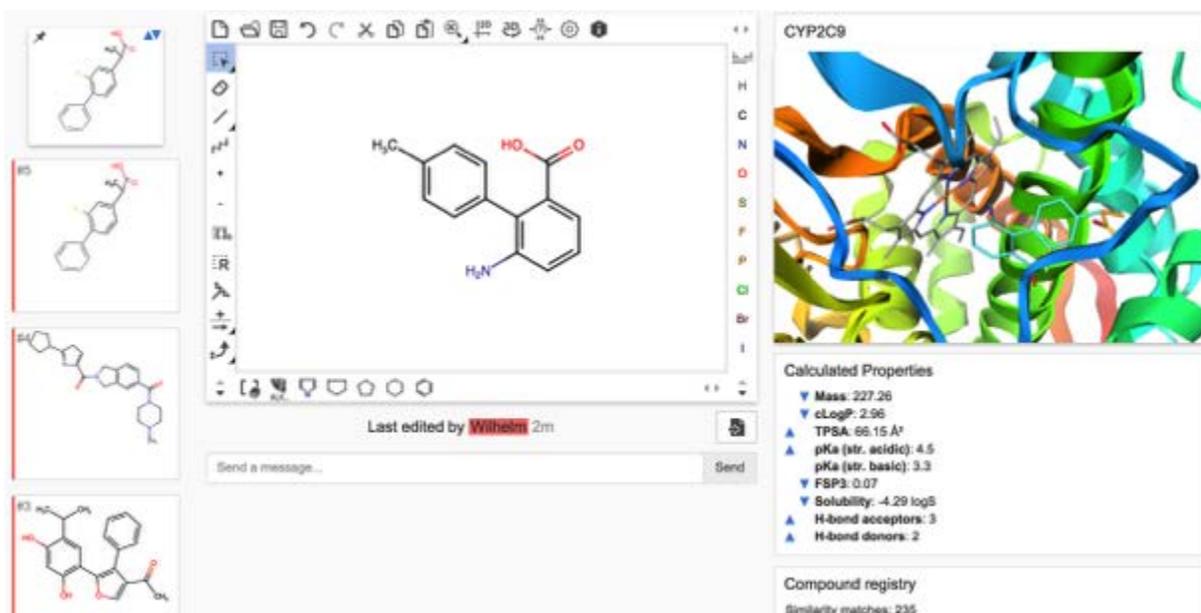

Fig 2: Marvin Live used to explore property space of a compound

2) Usage in academic teaching and tutoring

In the academia, Marvin Live can also be used by professors or educators and their students.

We identified typical uses for mentoring and general guidance, where students ask questions and discuss problems they face during their studies, or when writing papers. As in many cases subtle aspects of a molecule or a synthetic process are in the center of the discussion, the chemical editor has a more important role on the interface and takes up most of its area.

A slightly different scenario involving group discussions, for example students forming study groups to prepare for exams and discuss problems has also come in focus, where chat gains more importance.

To support these needs, we developed customization capabilities and an alternative user interface, that allocates more area to the chemical editor and chat, and hides options rarely relevant for these uses (*see fig. 3*). The application also supports tweaking the chemical editor's toolbar and switching to presets designed for education uses.
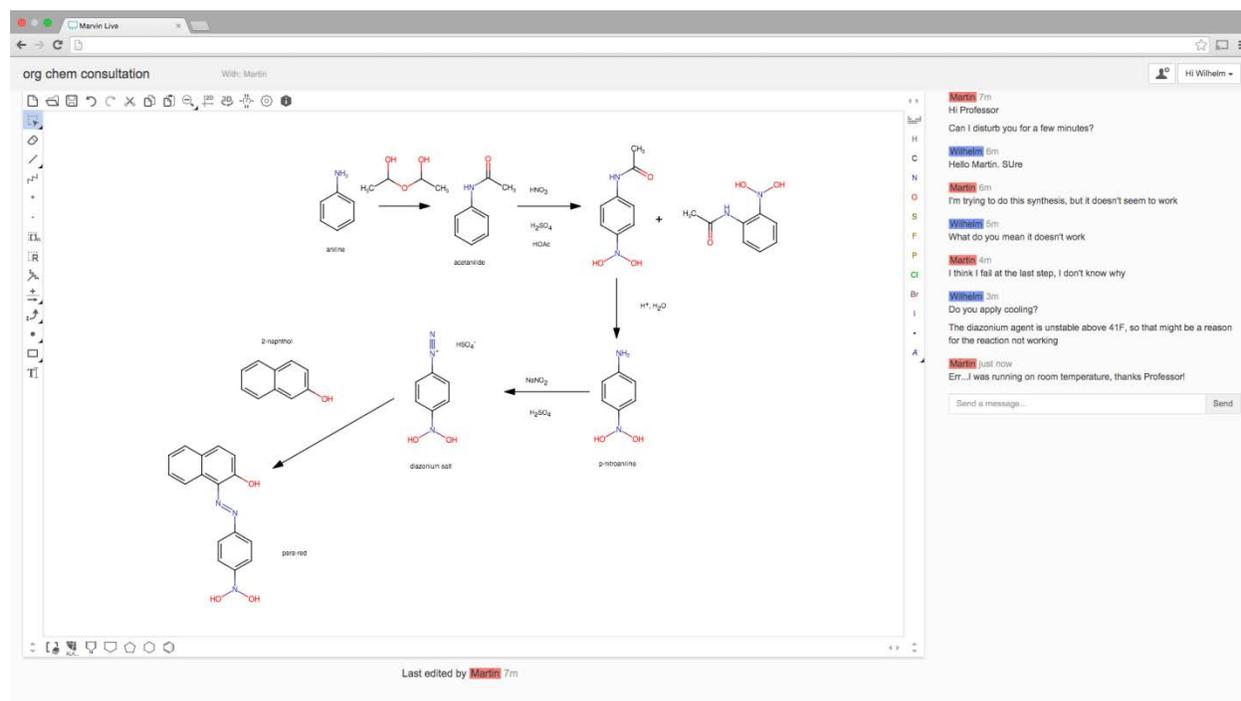

Fig 3: Marvin Live used for tutoring in an "office hours" setting

Technology

Marvin Live is a server based application accessible through a web browser. This lets first time users immediately use the entire application, without installing any application locally. The security system supports the most common type of authentication protocols: LDAP[12], Active Directory[13] and SAML based single sign on[14].

If the project or course benefits from database cross-checks or calculations, a JavaScript based plugin system is available to connect Marvin Live to public or custom web services. This lets administrators specify where a web service is located, what is the syntax used for queries, and how the results should be presented.

Chemical ideas can be loaded in an embedded chemical drawing application called Marvin JS[15]. To cover the most common research areas and teacher-student discussion topics, we made sure it supports not only the drawing of simple chemical structures, but also more advanced chemical features[16], such as:

- generic groups (i.e. R-groups in Markush structures) to describe collections of molecules in a compact representation
- simple and complex reactions for the discussion of potential synthetic approaches
- electron flow arrows for discussing reaction mechanisms
- and organo-metallic compounds.

The editor also supports the calculation of the absolute stereo configuration of chiral atoms and automatic mapping of reactions. With the option to alter the general layout and switch toolbar presets, the editor can display and support complex synthetic pathways.

This editor is extended with a mechanism that in real time updates the view for everyone (*see fig. 4*).
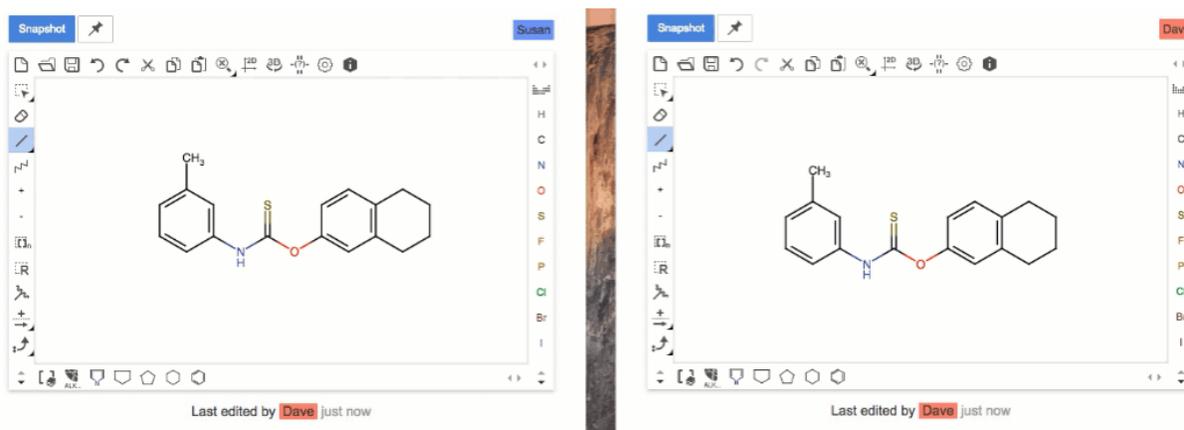

Fig 4: real time synchronization of chemical editors

Online tryout

Instead of a typical "Conclusions" section, we provide the following sandbox to readers of this newsletter to investigate the usefulness of this approach and experience the possibilities first hand. The sandbox is customized for the academic use cases described above.

URL: https://officehours-us.chemaxon.com/login/uni-nowhere
Username: professor1
Password: prof

On this sandbox, parallel discussions can be created and all rooms are open to join by anyone. Please note whether other participants are actively working when you are entering a room.

References:

1) What does the future hold for laboratory informatics? - Michael Elliott, CEO, Atrium Research https://www.youtube.com/watch?v=aT7T6rwdk2s
2) e.g: http://cen.acs.org/articles/93/i39/New-Collaborative-Approach-Chemists.html
3) https://en.wikipedia.org/wiki/List_of_video_telecommunication_services_and_product_brands#Software_clients_.E2.80.93_free_instant_messaging_video_programs
4) SureChemBL - Chemical Structure Information in Patents: https://www.surechembl.org/search/
5) ChemBL - chemical database of bioactive molecule: https://www.ebi.ac.uk/chembl/
6) UniChem - cross-reference between chemical structure identifiers from different data sources: https://www.ebi.ac.uk/unichem/
7) PubChem - https://pubchem.ncbi.nlm.nih.gov/
8) Reaxys - retrieval of chemistry information and data from published literature: https://www.reaxys.com/
9) eMolecules - database of screening compounds and building blocks: https://www.emolecules.com/
10) KNIME Server - https://www.knime.org/knime-server
11) Pipeline Pilot - http://accelrys.com/products/collaborative-science/biovia-pipeline-pilot/
12) LDAP: https://en.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol
13) ADFS: https://en.wikipedia.org/wiki/Active_Directory
14) SAML: https://en.wikipedia.org/wiki/Security_Assertion_Markup_Language
15) Marvin JS: https://www.chemaxon.com/products/marvin/marvin-js/
16) Marvin JS User's Guide: https://docs.chemaxon.com/display/marvinjs/Marvin+JS+User's+Guide