MolView: an attempt to get the cloud into chemistry classrooms

Herman Bergwerf*

Bachelor student Nanobiology, TU Delft, Delft, Netherlands

E-mail: hermanbergwerf@gmail.com

Abstract

MolView is a web application which helps students and teachers to visualize molecular structures and view their properties. There are numerous databases publicly available to provide the required data such as PubChem, ChemSpider, ChEMBL, DrugBank, the Crystallography Open Database,[1] and many more.[2 3] Currently, MolView uses PubChem, RCSB, the Crystallography Open Database, the Chemical Identifier Resolver,[4] and the NIST WebBook[5] to retrieve data. MolView offers a simple search interface to find small molecules, proteins and crystal structures in these databases. MolView uses JavaScript libraries, that use modern web technologies such as WebGL, to visualize these structures. In the past year I have designed a new version of this application from the ground up to facilitate the implementation of new databases and tools. Along with a new architecture and user interface, this version will include internationalization, interactive instructions, advanced search tools, more import/export tools, and more presentation tools. I also intend to include more computational chemistry tools to make the analysis and processing of complex data easier and more fun.

Introduction

In recent years web technologies have come a long way. The usage of browser plugins, such as Adobe Flash and Java, for the development of advanced applications in websites, has been replaced by built-in JavaScript APIs such as WebGL[6] and Web Workers.[7] The market of cloud computing[1] has skyrocketed, partially due to the vast increase of big data[2].[8][9] New standards for database APIs[3], such as REST and JSON[4], have emerged and are now used by almost all public API's on the internet including some major chemical databases like PubChem, ChEMBL and PDBe. I think these developments have opened many new opportunities in cheminformatics and bioinformatics.

History

MolView started out as 2D to 3D structure converter where the user can draw a structural formula and view a generated 3D conformer that was generated using the Chemical Identifier Resolver[5]. The ability to search by name using the Chemical Identifier Resolver was added later as an experimental feature. This turned out so well I added integrations for PubChem, RCSB and the Crystallography Open Database. Now MolView has become an example of what happens when you bring modern web browser technologies and online scientific data resources together. MolView is available free of charge on `http://molview.org` since 1 July 2014. The user base is still rather small (¿10k sessions per month) but is growing rapidly.

[1] Using a network of remote servers for computational tasks rather than a local server.

[2] The term 'big data' is often used to describe datasets that are too large to be processed by traditional applications or infrastructures.

[3] In this case the API (short for Application Programming Interface) is the interface offered by the database that can be used by other applications to exchange data.

[4] A format to encode data (like XML) that is based on the way JavaScript objects are encoded (JSON stands for JavaScript Object Notation), making it a suitable format for web applications

[5] In case you want to read learn more, you can have a look at this article: `http://blog.molview.org/posts/2015/06/01/looking-back/`

Use cases

You might be wondering what you can use MolView for today. The best way to find out is, of course, to visit `http://molview.org` and try it out for yourself. To get you started, here is a list of things you might want to try. Each section is also demonstrated in a YouTube video.

View 3D structure of organic molecules

You can draw organic molecules in the sketcher on the left side. By default the sketcher shows the structural formula of caffeine. Clear the sketcher by clicking the trash icon in the top-left corner. Then draw a new structure using the sketch tools. You can for example draw a benzene ring, an individual atom or a bond. When you are done drawing the structure, convert the structure to 3D using the '2D to 3D' button in the top-right corner of the sketcher. You can now view and interact with the 3D structure. Watch a video tutorial at: `https://www.youtube.com/watch?v=sUlEsncsQeQ`

Measure distances, angles and torsion angles in small molecules

You can also search for chemical structures. Type a molecule name into the search input in the top-left corner of the window. The search input will display a list of suggestions from PubChem, the Crystallography Open Database and RCSB. When you have loaded a structure, you can enable a measurement tool via the *Jmol* menu. MolView uses the JavaScript variant of Jmol (JSmol) for the measurement tools.[10] When you have enabled a measurement tool, you can click atoms in the viewing window to measure distances, angles and torsion angles. Watch a video tutorial at: `https://www.youtube.com/watch?v=ZN75lKSsBFc`

Study electron distributions in small molecules

Jmol has many powerful computational tools and a few are currently directly accessible in MolView. You can for example render a Molecular Electrostatic Potential surface of the loaded 3D molecule. If you connect a fluorine atom to a hydrogen atom and render a translucent MEP surface, you can clearly see that the fluorine atom attracts electrons much stronger than the hydrogen atom. You can also run an energy minimization via the *Jmol* menu. This can be useful when the loaded structure is resolved using the Chemical Identifier Resolver. The Chemical Identifier Resolver uses CORINA, a program that splits the molecule into ensembles, looks these ensembles up in a database and assembles them back together.[11] Watch a video tutorial at: `https://www.youtube.com/watch?v=yo2E3Ftv4rE`

Study small crystal structures

You can load small crystal structures from the Crystallography Open Database via the search interface. The blue suggestions from the search input are mineral names from the Crystallography Open Database. Additionally you can search trough the entire Crystallography Open Database via the search menu. After you have loaded a crystal structure, you can render a supercell model via the *Model* menu. Watch a video tutorial at: `https://www.youtube.com/watch?v=ELIIynwpfes`

Visualize biological macromolecule's

Just like small molecules and crystal structures, macromolecule's can also be loaded into MolView. MolView can retrieve biological macromolecule's from RCSB. You can switch between different color schemes and protein structure representations via the *Protein* menu. Watch a video tutorial at: `https://www.youtube.com/watch?v=jkWXMy8xQGM`

View spectra

You can also view certain spectra. To do this, you first have to load or draw a molecule. Then you can open the spectrum viewer via the *Tools* menu where you can choose from different spectra. IR and mass spectra are fetched from the NIST Chemistry WebBook. A H1-NMR prediction is calculated using *nmrdb.org.*[12] When you select a spectrum, it will be loaded into the interactive spectrum viewer where you can read out the values. Watch a video tutorial at: `https://www.youtube.com/watch?v=Yjaoy3zk0sg`

Embed interactive 3D models in websites

If you write web articles that involve chemistry, the embedding tool from MolView might be just what you need. You can get an embed code for every 3D structure in the viewer by opening the embedding dialog via the *Tools* menu. You can then paste this embed HTML code into your web page. In future versions you will be able to embed more content such as the spectrum viewer and the sketcher.

Prospects

Over the past years numerous ideas for new features have come up. I'm now writing a new version with a much higher level of modularity than the current one. This will help me to integrate new exiting stuff. Apart from quite a number of new database integrations, I want to focus on the three other subjects: visualization tools, sketch tools and tools for teachers. The next three sections will explain these subjects in more detail.

Visualization tools

Currently MolView uses GLmol, JSmol and ChemDoodle Web for 3D visualization.[13][14] In the last few years more mature web viewers for molecules have been developed that will replace the current viewers (except for JSmol of course). These viewers include 3Dmol.js

(organic molecules, fork of GLmol), PV (simple viewer for proteins) and NGL (advanced viewer for proteins).[15][16][17] But viewing 3D models is not the only visualization tool I intend to offer. An interactive NMR viewer like the one on *nmrdb.org* might finds its way into MolView pretty soon. And perhaps even a DNA sequence viewer/explorer.

Sketch tools

The sketcher is the only visualization tool I've actually written myself. I created my own sketcher because it is a very critical component and I was not satisfied with the web-based, open-source sketchers that are currently available. I've added some basic features to the sketcher already and I intend to add quite a number of new features to make it even more powerful. Apart from obvious features such as functional groups, dot structures, reaction arrows and annotations, I've been thinking about a tool for automatically depicting the structural formula using a certain projection such as the Newman projection or the Fisher projection.[18]

Tools for teachers

I want teachers to use MolView, and therefore making MolView more powerful for teachers is a top priority. In the past year I received quite some requests for new features from teachers who tried MolView. The new architecture will open up the way for quite a number of these features. It have built-in support for embedding any view, anyone will be able to create a step-by-step guide for tutorials in MolView, and it will have mature support for saving and sharing files. I'm especially exited about the step-by-step guides. These can be used to learn new users how to use MolView, but it could for example also be used for a tutorial to explain the difference between geometric isomers.

Get involved

You can start using MolView today on `http://molview.org`. Your feedback is very valuable! don't hesitate to contact me if you have questions or ideas. MolView is growing faster than ever and will require an increasing amount of server resources. A dutch hosting company called PCextreme granted me free access to their infrastructure so I can continue providing this application for free! In the future I want to add more heavy computational tools to MolView. Perhaps we can connect MolView to your computing grid to make this possible!

References

(1) Saulius Graulis, Daniel Chateigner, Robert T. Downs, et al. (August 2009). *Crystallography Open Database an open-access collection of crystal structures.* doi:10.1107/S0021889809016690

(2) Guy R. Cochrane and Michael Y. Galperin. (January 2011). *The 2011 Nucleic Acids Research Database Issue and the online Molecular Biology Database Collection.* Nucleic Acids Res January 1, 2011 vol. 39 no. suppl 1 D1-D6. doi:10.1093/nar/gkq1243

(3) Richard L. Apodaca. (October 2011). *Sixty-Four Free Chemistry Databases.* `http://depth-first.com/articles/2011/10/12/sixty-four-free-chemistry-databases/`. (accessed Nov 2, 2015).

(4) Markus Sitzmann. (March 2012). *ACS San Diego, March 2012, InChI Symposium.* `http://www.slideshare.net/sitzmann/acs-san-diego-march-2012-inchi-symposium`. (accessed Nov 2, 2015).

(5) P.J. Linstrom and W.G. Mallard, Eds. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69.* National Institute of Standards and Technology, Gaithersburg MD, 20899. `http://webbook.nist.gov`. (accessed Nov 27, 2015).

(6) Khronos Group. *OpenGL ES 2.0 for the Web*. https://www.khronos.org/webgl/. (accessed Nov 2, 2015).

(7) W3C. (September 2015). *Web Workers - W3C Working Draft 24 September 2015*. http://www.w3.org/TR/workers/. (accessed Nov 28, 2015).

(8) Louis Columbus. (September 2015). *Roundup Of Cloud Computing Forecasts And Market Estimates Q3 Update, 2015*. Forbes. http://www.forbes.com/sites/louiscolumbus/2015/09/27/roundup-of-cloud-computing-forecasts-and-market-estimates-q3-update-2015/. (accessed Nov 2, 2015).

(9) James Manyika, Michael Chui, Brad Brown, et al. (May 2011). *Big data: The next frontier for innovation, competition, and productivity*. http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation. (accessed Nov 2, 2015).

(10) Hanson, R. M., Prilusky, J., Renjian, Z., Nakane, T., & Sussman, J. L. (2013). *JSmol and the Next-Generation Web-Based Representation of 3D Molecular Structure as Applied to Proteopedia*. Israel Journal Of Chemistry, 53(3-4), 207-216. doi:10.1002/ijch.201300024

(11) (a) Sadowski, J.; Gasteiger, J.; Klebe, G. *Comparison of Automatic Three-Dimensional Model Builders Using 639 X-Ray Structures*. J. Chem. Inf. Comput. Sci. 1994, 34, 1000-1008. (http://dx.doi.org/10.1021/ci00020a039) (b) The 3D structure generator CORINA is available from Molecular Networks GmbH, Erlangen, Germany (http://www.molecular-networks.com).

(12) Banfi, D., & Patiny, L. (2008). *Www.nmrdb.org: Resurrecting and Processing NMR Spectra On-line*. CHIMIA -ZURICH-, 62(4), 280-281.

(13) Takanori Nakane. *GLmol - Molecular Viewer on WebGL/Javascript*. `http://webglmol.osdn.jp/index-en.html`. (accessed Nov 27, 2015).

(14) Burger MC. (2015). *ChemDoodle Web Components: HTML5 toolkit for chemical graphics, interfaces, and informatics*. Journal Of Cheminformatics, 7, 35. doi:10.1186/s13321-015-0085-3

(15) Rego N, & Koes D. (2015). *3Dmol.js: molecular visualization with WebGL*. Bioinformatics (Oxford, England), 31(8), 1322-4. doi:10.1093/bioinformatics/btu829

(16) Marco Biasini. *PV - WebGL-based protein viewer*. doi: 10.5281/zenodo.12620

(17) Rose AS, & Hildebrand PW. (2015). *NGL Viewer: a web application for molecular visualization*. Nucleic Acids Research, 43(W1), W576-9. doi:10.1093/nar/gkv402

(18) *Three-Dimensional Representations*. Chemical Education Digital Library. `http://www.chemeddl.org/resources/stereochem/threed1.htm`. (accessed Nov 27, 2015).