J Comput Aided Mol Des (2012) 26:135–136
DOI 10.1007/s10822-011-9513-2

PERSPECTIVE

# The perspectives of computational chemistry modeling

**Igor V. Tetko**

Received: 15 November 2011 / Accepted: 2 December 2011 / Published online: 11 December 2011


**Abstract** The on-line tools for computational chemistry modeling will be increasingly used in the future. This will bring the advantages both for the authors and the readers.

**Keywords** Web tool · On-line modelling · QSAR · Open access · On-line publishing · Model repository

The recent decade has been characterized by an appearance and fast development of open databases and tools for analysis of chemical data on the web. In brief, such online services can be categorized as data collections (e.g. PubChem [1], ChemSpider [2], CheMBL [3]) and tools for development and publishing of predictive models (e.g. OpenTox [4], Chembench [5], OCHEM [6]). The former resources have brought to the public enormous and previously unimaginable amount of data with various biological and physicochemical properties of molecules. The latter resources enabled an easy development of models and publishing them on the web for public access by the scientific community. Both types of resources significantly contribute to the development of computational chemistry.

This tendency will gain a momentum in the future and will lead to adoption of new standards for publishing of methodological and research articles in the computational chemistry. The imperative part of any published computational model will be the possibility to access and run it on the web as a free or pay-for-access service. The web services will replace standalone software and will become a standard for sharing and dissemination of results of developed approaches. The developed models will also include data (in case of non-industry submissions) and will be submitted to specialized software repositories. They will be accessible during peer-review process and permanently supported after publication. This will become possible thanks to development of new standards for data exchange and publishing of models, development of electronic archives to store published models.

The advantages of such publishing options are obvious. Out of 20 top most cited articles published in JCAMD, more than half are directly related to software developments, which are now widely known and accepted tools in the computational chemistry, e.g. MOPAC [7], DOCK [8], Molden [9], LUDI [10], VCCLAB [11]. Thus allowing readers to use the developed methodology as software programs has a great advantage to the authors of the algorithms. The advantages of making methodological developments and models available for use as software tools will only increase in the future. Previously, support and development of the software required significant knowledge of programming languages, efforts and resources. The new technologies and standards will make it easy and transparent. The new model development tools will include ports to directly submit models and data to computational chemistry repositories or/and models can be developed and published directly on the web.

The requirements to publish models on-line and availability of software platforms to support it would allow the re-use of knowledge, data and will boost the field of computational chemistry and drug discovery. Thus, instead of just reading about a model, the future readers will be able to access it, verify its statistics and data, apply it to their own data and compare the results with the previously

I. V. Tetko (✉)
Helmholtz Zentrum Muenchen, German Research Center for Environmental Health and eADMET GmbH, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany
e-mail: itetko@vcclab.org
URL: http://ochem.eu

published models in just few clicks. Moreover, the user can use the published models to do virtual screening of new molecules or apply them to predict physico-chemical parameters and biological activities. The aforementioned predictive modeling tools already provide such possibilities for traditional QSAR models developed using machine-learning algorithms. The future developments of these and other tools will include wide support of quantum chemistry, molecular dynamics and docking approaches. The complicated models could be shared and published using workflow and pipe-line tools, such as KNIME [12] or PipeLine pilot [13].

Thus, the future of scientific articles will be quite different from now: instead of just reading, the future readers will be trying and testing the resources directly accessible on the web with a complete hands-on experience. I think this future will become the reality within the next decade.

## References


1. Wang Y, Bolton E, Dracheva S, Karapetyan K, Shoemaker BA, Suzek TO, Wang J, Xiao J, Zhang J, Bryant SH (2010) Nucleic Acids Res 38(Database issue):D255
2. Williams AJ (2008) Drug Discov Today 13(11–12):502
3. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP (2011) Nucleic Acids Res. doi:10.1093/nar/gkr777
4. Hardy B, Douglas N, Helma C, Rautenberg M, Jeliazkova N, Jeliazkov V, Nikolova I, Benigni R, Tcheremenskaia O, Kramer S, Girschick T, Buchwald F, Wicker J, Karwath A, Gutlein M, Maunz A, Sarimveis H, Melagraki G, Afantitis A, Sopasakis P, Gallagher D, Poroikov V, Filimonov D, Zakharov A, Lagunin A, Gloriozova T, Novikov S, Skvortsova N, Druzhilovsky D, Chawla S, Ghosh I, Ray S, Patel H, Escher S (2010) J Cheminform 2(1):7
5. Walker T, Grulke CM, Pozefsky D, Tropsha A (2010) Bioinformatics 26(23):3000
6. Sushko I, Novotarskyi S, Korner R, Pandey AK, Rupp M, Teetz W, Brandmaier S, Abdelaziz A, Prokopenko VV, Tanchuk VY, Todeschini R, Varnek A, Marcou G, Ertl P, Potemkin V, Grishina M, Gasteiger J, Schwab C, Baskin II, Palyulin VA, Radchenko EV, Welsh WJ, Kholodovych V, Chekmarev D, Cherkasov A, Aires-de-Sousa J, Zhang QY, Bender A, Nigsch F, Patiny L, Williams A, Tkachenko V, Tetko IV (2011) J Comput Aided Mol Des 25(6):533
7. Stewart JJ (1990) J Comput Aided Mol Des 4(1):1
8. Ewing TJ, Makino S, Skillman AG, Kuntz ID (2001) J Comput Aided Mol Des 15(5):411
9. Schaftenaar G, Noordik JH (2000) J Comput Aided Mol Des 14(2):123
10. Bohm HJ (1992) J Comput Aided Mol Des 6(1):61
11. Tetko IV, Gasteiger J, Todeschini R, Mauri A, Livingstone D, Ertl P, Palyulin VA, Radchenko EV, Zefirov NS, Makarenko AS, Tanchuk VY, Prokopenko VV (2005) J Comput Aided Mol Des 19(6):453
12. Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, Ohl P, Sieb C, Thiel K, Wiswedel B (2008) Data analysis machine learning and applications. In: Preisach C, Burkhardt H, Schmidt-Thieme L, Decker R (eds) KNIME: the Konstanz information miner. Springer, Berlin, p 319
13. http://accelrys.com/products/pipeline-pilot/. Accelrys Inc. Accessed 08 Dec 2011