COMPUTER SOFTWARE REVIEWS

ACD/I-Lab 4.5: An Internet Service Review

Artëm Masunov†

Department of Chemistry, City College, City University of New York, Convent Avenue at 138th Street, New York, New York 10031

Introduction. Prediction of the properties and structure of the chemical substances is one of the ultimate goals of Computational Chemistry. Although ab initio and semi-empirical methods are speeding progress toward this goal, for now bench chemists will have to rely on faster and computationally affordable empirical methods to fulfill their practical needs. Many software developers now offer a variety of prediction tools and databases for the chemist with a PC. Unfortunately, substantial cost and reasonable skepticism toward new commercial products may be preventing many practicing chemists from using these software packages. Fortunately, the Internet revolution brings us an alternative. An Internet-based service would give instant access to property prediction programs and databases on known compounds. A user would have the choice to pay-per-prediction or to subscribe to one or more services for a defined period of time. It would be no longer necessary to purchase the software package and install frequent updates. Service could run on virtually any platform and would require comparatively small storage capacity and clock speed. Interactive Laboratory (I-Lab) from ACD Labs (http://www.acdlabs.com/ilab) is exactly this kind of service. The present review is intended to critically describe the capabilities of I-Lab service,as well as its limitations and to provide some examples of its use.

Installation and Capabilities. After opening an account, a user has two options to access the service: Java applet or Windows application. Java applet is available for anyone who uses Netscape 4.5, Internet Explorer 5.0, or any later version. It does not require installation and opens at a click on the web page. I found no technical glitches when testing the service on Linux/Alpha with Netscape 4.73 and on Windows 2000/PIII with IE 5.5, Netscape 4.7, and (my personal favorite) lightweight browser Enigma 3.6.[1] Optionally, the user can download the applet locally for faster start-up, but I did not find this necessary. After a two-dimensional molecular structure is sketched in the applet window it can be idealized, imported to the webpage, and converted into SMILES description.[2] With one click on the menu a request is submitted to the I-Lab server and progress is displayed. During off-peak hours the calculation and download of the results usually takes less than a minute. Network jams during regular business hours made my experience less enjoyable.

Windows users have also an option to install ChemSketch. It comes in three parts, which are available for free download and require 11, 5, and 7 Mb of disk space. ChemSketch is an integrated graphical package for 2D-structure drawing and text manipulations, 2D/3D conversion with optimization, and export/import for a variety of file formats. Its convenience of use and functionality exceeds that of ChemOffice—a popular suite from CambridgeSoft.[3] The only shortcoming of ChemSketch seems to be the absence of geometry optimization with popular force fields (MM3, Amber) and quantum chemical methods. As a result, supramolecular complexes cannot be optimized. ChemSketch, however, offers numerous templates, a tautomers generator, a chemical spell checker, and a macrocomand language called ChemBasic (installed separately). Among the preprogrammed macros are peptide, carbohydrate, and DNA/RNA builders, along with the export to HTML and VRML formats for instant publication on the Web. Although ChemSketch is provided with a large manual, self-explanatory menus and detailed online help eliminate the need to read it, unless one wants to learn ChemBasic. One of the help options (Journal Submission Guidelines) came in handy in preparation of this manuscript. The older versions of ChemSketch were reviewed elsewhere.[4,5] The current version can calculate liquid properties without an Internet connection. The list includes the following: molar refractivity and volume, parachor, surface tension, density, dielectric constant, polarizability, and refractive index. Interface to I-Lab (a third component in the installation) is necessary for IUPAC and CAS name generation and for more advanced property prediction, such as aqueous solubility, boiling point, vapor pressure, enthalpy of vaporization, bioconcentration factor, and adsorption coefficient. ^{1}H, ^{13}C, ^{19}F, ^{31}P NMR spectra, pK_a, and LogP prediction and database searches also require I-Lab login. Finally, I-Lab can name the molecule using IUPAC and CAS conventions or build molecular structure based on the chemical name (including trivial and brand names). All this functionality (except for 3D optimization, graphics manipulations, and macrocomands) is available in the browser-based service, so later on both flavors of I-Lab will be discussed together. One of the advantages of the ChemSketch over Java applet is that the user can submit all the prediction requests concurrently and work on the other tasks. A pop-up window will inform about the arrival of the results. This is especially useful when Internet connection is slow. Web interface has an option to access the results of the requests submitted for the past 24 h. Similar functionality is available in ChemSketch via Task Manager window.

Windows users with Internet Explorer have the choice of a third, hybrid interface. It is based on Active X control technology and is presently in beta-testing. The interface is browser based, but ChemSketch is used to edit the structure. The screen layout looks nicer, but I found this option less convenient in cases where editing was needed. It took longer for the heavy ChemSketch to open than it would for Java applet to download from the server.

NMR Prediction. After a structure is built, a click on the menu produces a realistic picture of the predicted NMR spectrum in a special window and can be inserted into a

† Corresponding author e-mail: artem.masunov@usa.net.

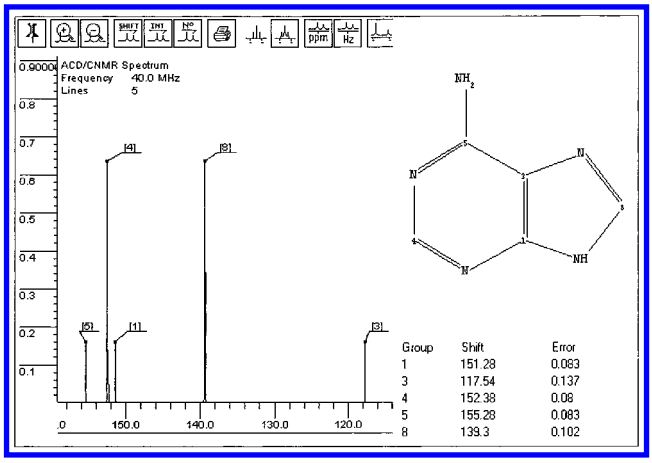

Figure 1. A spectrum generated for adenine by I-Lab/C NMR prediction service using Java applet.

separate page of ChemSketch or the Web browser. A ^{13}C spectrum for adenine prepared using the Java applet is shown on Figure 1. Windows application creates an even more professional looking result. But how precise is it? Comparisons tend to favor knowledge-based algorithms for calculation of chemical shifts and coupling constants, as implemented in ACD Labs/NMR.[6,7] A detailed comparison between empirical (ChemDraw) and ab initio (Gaussian 98) prediction of NMR spectra was recently published on the Web.[8] The general conclusion was that average errors of empirical method could be smaller (although not systematic) in the common cases. In turn, ab initio methods performed better in unusual cases, like propellane. Since systematic comparison of accuracy was not the goal of this review, I tested I-Lab/CNMR predictor only on the examples considered previously.[8] For strained systems I-Lab gave large errors, similar to ChemDraw (better for bicyclobutane, worse for propellane). For natural compounds adenine and taxol I-Lab outperformed both ChemDraw and Gaussian, giving nearly experimental values for all atoms. Precise prediction of the chemical shift may indicate an efficient incorporation of experimental data into the predictive algorithm. Unfortunately, the actual experimental values were not reported in the output. Moreover, taxol was not found in the I-Lab/CNMR database search. This could indicate either an imperfection in a search algorithm or the fact that taxol is missing from the database and the chemical shift values are derived based on analogous compounds.

The quality of empirical prediction always improves when more analogues are added to the database. Unfortunately, user-supplied database extensions available as an option for traditional ACD Labs software is not yet implemented in the Web service as well as two-dimensional NMR spectra. The reverse problem, elucidation of the molecular formula based on NMR spectra, is probably the most challenging and useful for a synthetic chemist. It is also unavailable from the I-Lab service, so that the interested users have to buy and install the appropriate ACD Labs module locally.

NMR Database. I-Lab/NMR DB allows search by exact molecular structure or by single fragments. If desired, the user can specify a range of brutto-formulas, molecular weights, chemical shifts, part of chemical name, or literature reference (with or without chemical structure or substructure). The results contain IUPAC name, structure with numbered atoms, and values of the chemical shifts with references. It can be printed or inserted into the open ChemSketch document.

Unfortunately, the search did not work as smoothly as it was designed to. For adenine both the search by exact structure and the search by name (both trivial and IUPAC) failed to produce any hits. Brutto-formula and substructure search alone gave multiple hits, but only the combination of the last two correctly located two entries (pure adenine and its hydrochloride). Apart from this glitch in the search engine, I found a number of minor inconveniences. When multiple hits were found, they were presented as a list of brutto-formulas with systematic names in no apparent order, up to 15 hits at a time. To see the structure of each hit, the user has to click on it and wait for the web server to respond (ChemSketch gives the result instantly). A simultaneous display of structural formulas on the web page for all the found hits as soon as they are downloaded from the server would be a nice improvement in future releases. In ChemSketch the list of hits disappears after one of them is saved. To select another hit the user has to open Task Manager window from the I-Lab menu. Also, it is impossible to look at the previous pages of the current document while viewing the list of hits, which makes the comparison difficult. Finally, the option to export the results to a spreadsheet application via both Windows and Java-based interfaces would be very helpful.

Physicochemical Properties. The database search options for LogP and pK_a are similar to those for NMR DB: structure/substructure, formula, and target value. For pK_a the temperature and ionic strength must be specified. These values replace the name and the reference fields in the search request. In the case of pK_a the database seems to be incomplete: for all three isomers of cyclopropanedicarboxylic acid only the first dissociation constant is given, while the standard reference book[9] gives both the first and the second constants for all three. Also, most hits appear in the list with no name shown, and some have typos. Apparently, ACD Labs did not apply their own name generator to their databases.

It is necessary to note, that predicting algorithm for pK_a calculations is based on molecular topology, and not on three-dimensional structure. As a result, this algorithm does not distinguish cis and trans isomers. Also, pK_a predictions for polypeptides are not probably reliable for that reason.

As in the case of NMR spectra, the databases form the foundation for empirical predictions. The accuracy of the incremental approach implemented in the LogP module from ACD Labs was addressed earlier.[3] A definite plus of the LogP prediction service is that experimental value with the reference is reported when available. pH dependence for octanol/water partition coefficients LogD also has this feature, but it was not available in pK_a and NMR predictors.

A click on the menu allows the user to predict the variety of other properties listed above. Aqueous solubility requires melting point as an input parameter, while vapor pressure and boiling point need external conditions (temperature value and the pressure, accordingly). The other properties do not need any extra input. The user can select different units, including Rankine for the temperature. Unlike locally installed version of ACD Labs software, I-Lab has no option to return the results of vapor pressure and boiling point

predictions in the form of tables or graphs, so that the user has to build them point by point. Batch submissions of the multiple compounds are also absent from the web service.

Naming. The name to structure function is my favorite part of the I-Lab service. It accepts any trivial name present in the vast dictionary as well as systematic or semisystematic names. This service makes it much easier for a chemist to use of the computer by eliminating the need to draw the structure. It handles ionic structures and exotic functional groups and warns about ambiguity in most cases.

The conversion of chemical structure into the IUPAC name is a difficult task. The recent reviews of ACD Labs/Name product[10,11] found it superior to current naming software from Belstein in both limits (maximum number of atoms and cycles) and observance of IUPAC rules. In case of some natural compounds, like peptides and polysaccharides, I-Lab/Name generates more compact trivial-based names.

Pricing. All software reviewed in this article is free for download from ACD Labs website. Predictions of liquid properties, which are implemented in ChemSketch and do not require Internet connection, are also free. Generation of IUPAC names is free with certain limitations. The other predictions and database searches are $1.5−$4 per transaction. Unlimited use is $20 daily or $95 monthly. Site licenses are also available. New users can request a fully functional trial access, which is free for 2 weeks. ACD Labs also offers an affiliate program, allowing access to calculation engines from the third-party websites.

REFERENCES AND NOTES

(1) http://www.enigmabrowser.com.
(2) http://www.daylight.com/dayhtml/smiles/.
(3) ChemOffice 4.0 Ultra; CambridgeSoft Corporation: 875 Massachusetts Ave, Cambridge, MA 02139, U.S.A.; http//www.camsoft.com.
(4) Spessard, G. O. ACD Labs/LogP dB 3.5 and ChemSketch 3.5. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1250−1253.
(5) Gunda, T. A. Chemical Drawing Programs Review − The Comparison of ISIS/Draw, ChemDraw, ChemWindow, ChemSketch and Chemistry 4-D Draw. *Internet J. Chem.* **2000**, *3*, L.25; http://www.klte.hu/~gundat/rajzprogramok/dprog.html.
(6) Thomas, S.; Bruhl, I.; Heilmann, D.; Kleinpeter, E. ^{13}C NMR Chemical Shift Calculations for Some Substituted Pyridines: A Comparative Consideration. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 726−730.
(7) Bramer, S. ACD/CNMR and ACD/HNMR: A Software Review. *Concepts Magn. Reson.* **1997**, *9*, 271−273.
(8) Cheeseman, J. R., Frisch, A. Predicting Magnetic Properties with ChemDraw and Gaussian. http://www.Gaussian.com/nmrcomp.htm.
(9) Dean, J. A.; *Lange's Handbook of Chemistry*; McGraw-Hill: New York, 1992; p 528.
(10) Ryan, S.; Verlag, G. T. ACD/IUPAC Name Pro 3.6. *Synthesis* **1999**, *5*, 904.
(11) Koch, R.; Lemmler, M. Nomenklatur: Die nächste Generation. *Nachrichten Chemie* **2000**, *48*, 642; Engl. Ed.: http://www.acdlabs.com/download/publ/nametrans_e.pdf.

CI010400L