

Online services for your desktop on a pay-as-you-surf principle—ACD/Labs I-Lab

Antony N. Davies

ISAS, Institute of Spectrochemistry and Applied Spectroscopy, Dortmund, Germany

In recent years a new company out of Toronto, Canada, has been causing quite a stir with innovative products and an aggressive marketing style. Many ideas which were being discussed in different quarters have found their realisation in the catalogues of Advanced Chemistry Development, ACD/Labs Inc. In the final article of this series on new Internet-based services for the chemist/spectroscopist, I will be taking a quick look at the ACD/Labs I-Lab online web-based program range.

Introduction

The idea behind ACD originated with a small team of independently minded post-doctoral researchers working together at Moscow State University, Moscow, Russia. In 1990, these scientists brought together their talents in the areas of chemistry, spectroscopy and computer science in order to develop chemical software tools focusing on prediction of chemical properties. From the original group of seven at Moscow State University, ACD now has a staff of eighty in Moscow. The Toronto headquarters has a staff of 27 providing sales, marketing, project management and technical support. The I-Lab product is one of a range and the most easily accessible to the average spectroscopist. The term I-Lab would, you assume, refer to Internet Laboratory but perversely, or perhaps not, it really stands for Interactive Laboratory! The I-Lab has a large number of services, mainly based on physical chemistry properties or nuclear magnetic resonance spectroscopy, as well as the structure naming services. A complete list is shown in Table 1. As can be seen, the list is large and in the space available it would be impossible to do the whole I-Lab service justice, so I will endeavor

Table 1. This table gives some idea of the features available through the ACD/Labs I-Lab server system.

Spectroscopy	
<i>NMR Prediction:</i>	H-NMR, C-NMR, P-NMR, F-NMR
Structure elucidator	
Database searching	H-NMR, C-NMR, P-NMR, F-NMR
Physical chemistry	
<i>Properties Prediction (Calculation)</i>	BP, LogP, LogD, pKa, H(vap), Vapour Pressure, Aqueous Solubility
<i>Database Searching</i>	pKa, LogP
“Environmental”	BCF, Adsorption Coefficient

our to concentrate on those features which I feel will be of interest to the average spectroscopist. This will include a brief description of the various chemical structure naming programs available either for free or at a fixed payment through the I-Lab server.

Initial access

System requirements

There are two separate routes into the I-Lab system.

1. The first, and the one I would recommend, is the installation of the ChemSketch program which can be downloaded from their website at <http://www.acdlabs.com/download>. I tested the systems with ChemSketch 4.01 and 4.55. You will also need the I-Lab extension software for ChemSketch which is available from the same site and which enables direct access to the ACD/Labs servers without the need to work with a browser. The 4.55 download is 5.8 Mbytes and the I-Lab add-on is a further 3.5 MB. Obviously a direct Internet connection

is required although there is no need for fast links. Once the software had been obtained and installed I worked quite happily from home over a slow modem link without problems.

2. There is also a direct route into the I-Lab services through a purely web-based environment where a structure applet called MolX is used to input the chemical information. Downloading the applet takes time and may be useful for those not allowed to install new software on their own PCs. But there again, these users may well be forced to work in a Java-free environment where the applet won't work either! To circumvent these problems the ability does exist to install the applets locally, requiring that they be downloaded once only. I ranked this option a poor second to working through ChemSketch.

Registration

In order to carry out the required downloads you will have to provide your email address as the system username and a password of your choice. There is a European mirror site so there shouldn't be any problems with

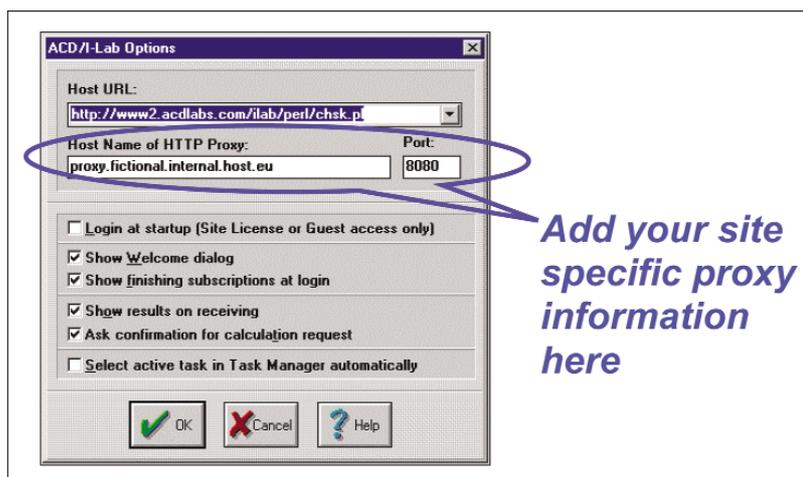


Figure 1. In order to use the I-Lab features direct from your desktop through ChemSketch you will need to input your own network proxy information as shown.

bandwidth, although I often find accessing the North American sites in the morning when they are asleep often provides better results.

Once the software has been installed on your system you will need to ensure that the network access information is correct, otherwise the direct access to the ACDLabs servers won't work.

Start up the ChemSketch program, wait for the annoying opening screen to time out its advertisement and click on the drop-down "I-Lab" menu.

If you don't work with proxy servers then this step can be skipped. Once you have added the required information, it is now possible to test the link by using the guest registration button on the I-Lab login menu.

As I-Lab is a mixture of free and pay-by-result services, only certain features are available to the guest login account. One of these which many people may find extremely useful is the IUPAC Name Free. This is a limited version of the full product.

Data entry

The initial impact of the ChemSketch user interface can be somewhat scary as there are a multitude of buttons and icons as well as two completely different types of screen depending on whether the "Draw" or "Structure" button in the top left-hand part of the screen is pressed. By switching between these two options the main central display does not change but the drawing tools alternate between somewhat more familiar and self-explanatory icons of a structure drawing package and those of a text/graphic editor.

For those who cannot face learning another user interface the software is

written with desktop integration in mind and intelligent data import through cut-and-paste from your usual structure drawing package—for this article I was using a version of ISIS-Draw from MDL, for example. One thing you will need to note, however, is to opt for "paste structure" rather than the default paste option, otherwise the structure will exist in ChemSketch as an ISIS-Draw object which will not allow itself to be uploaded into I-Lab and double clicking will start ISIS-Draw. Figure 2 shows in the upper trace that the two imported structures look identical but in reality only the right-hand trace is present as a chemical structure. In "Draw" mode using the EDIT / SELECT ALL option gives the result shown in the centre trace B indicating that two objects are present in the window but by switching to the structure window shows that the right-hand molecule exists as individual atoms and bonds whereas the left-hand molecule is still a singular object (C).

So assuming that we now have a chemical structure correctly entered into ChemSketch, either through cut-and-paste or by using the structure drawing tools, and the proxy settings correctly entered, we can now make use of the opportunities available in I-Lab.

NMR prediction

The most obvious tool for spectroscopists is the ability to predict NMR spectra for four different nuclei. This is a very rapid service, the example shown in Figure 3 was obtained over a modem connection to the internet in around 50 seconds.

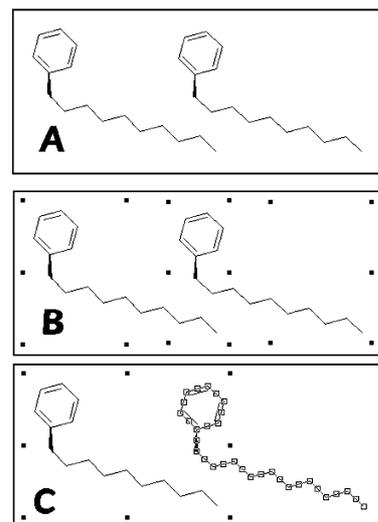


Figure 2. The same molecular structure imported into ChemSketch as a graphical object in the left hand traces and correctly as a chemical structure in the left hand traces. A is without any selection, B follows SELECT ALL in DRAW mode and C STRUCTURE mode.

This display is from an additional freeware product called CNMR Viewer, which links seamlessly into the main ChemSketch program. The viewer has many features normally only available in a full NMR spectra data handling package and to start to describe them would expand this article way outside our page limits. These options are also available for ¹H-NMR spectrum prediction but for the other nuclei there are currently no spectra manipulation packages and the results are displayed as peak table data (Figure 4).

This data can then be pasted into the ChemSketch reporting package using the "insert" button. The individual items are all cut-and-paste-able and as such can be transferred to your standard reporting package without any problems. Figure 5 displays how they then appear.

Database searching

There is also the capability within I-Lab to search for spectra within the ACD/Labs databases. This database facility can be a useful alternative to having large in-house reference databases but the same drawbacks will apply to those industrial chemists who

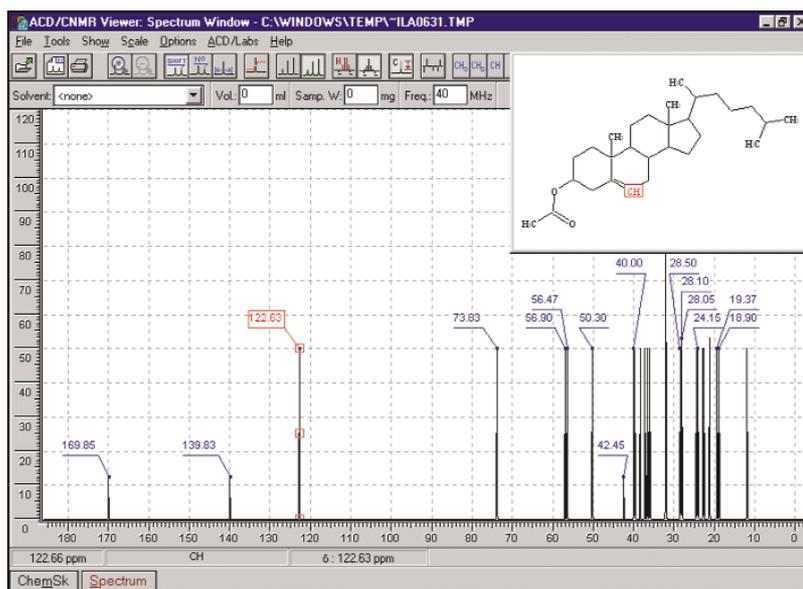


Figure 3. A ^{13}C -NMR spectrum predicted and displayed in the CNMR Viewer freeware package for manipulation before importing into the reporting package.

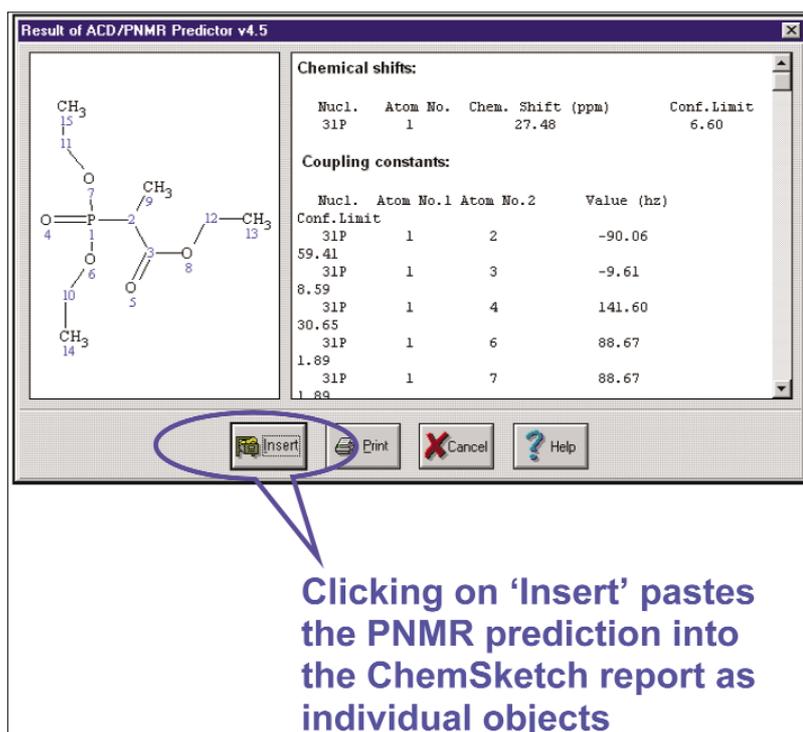


Figure 4. A ^{31}P -NMR spectrum predicted as a table of shifts and coupling constants.

cannot send their test data through the corporate firewall. In-house intranet versions are available for those wishing to test the systems on a more serious basis. According to the information on the I-Lab services page, <http://www.acdlabs.com/ilab/services.html>, they now have over 100,000 structures providing data for the ^{13}C and ^1H NMR database as well as 11,400 fluorine and more than 19,000 phosphorous NMR structures.

Spec Viewer

Another freeware product which isn't directly linked to the I-Lab service but which integrates into the ChemSketch report manager is Spec Viewer. This allows for your own data files to be loaded. I tested version 4.53 and had trouble with the stability of the product when relatively large NMR JCAMP-DX files had been imported for display.

Naming services

Another feature of I-Lab which spectroscopists are likely to enjoy is the various structure naming services which are available.

IUPAC NAME and IUPAC NAME FREE

Called ACD/IUPAC NAME, this option provides a translator from the inputted chemical structure to the IUPAC name. Due to the collaboration with IUPAC a cut-down version of the IUPAC naming software is available for free through the I-Lab interface although the range of substances whose nomenclature is available is restricted.

We found this service extremely useful when a particular reviewer recently demanded that a list of substances determined by pyrolysis GC/MS should conform to IUPAC nomenclature rather than the rather rough mixture of generic and CAS names we had submitted.

Chemical Abstracts Service nomenclature

Under the name ACD/INDEX NAME resides a service to provide Chemical Abstracts Service type names, using their systematic nomenclature approach, for the submitted structures. A typical example would be the difference between CAS (cholest-5-en-3-ol, acetate) and IUPAC (cholest-5-en-3-yl acetate).

Other services

The remaining services available through I-Lab can be seen in Table 1. How useful these will turn out to be depends entirely on your personal interests. The theoretical basis for each of the predictions is well described in the Help files, and it would be pointless to reproduce it here. I tested each of the services in turn with the same structure, which only went to prove I didn't fully understand what I was doing. I picked one of our NMR round robin test molecules, ethyl 2-(diethoxyphosphoryl)propanoate (Figure 6).

The first service I tried out was the pKa predict which failed with the error message "Error: Cannot calculate pKa value because the structure does not contain ionization centers". Well 1 : 0 to I-Lab! A search of the pKa database not surprisingly gave no hits either in full structure or substructure mode. The

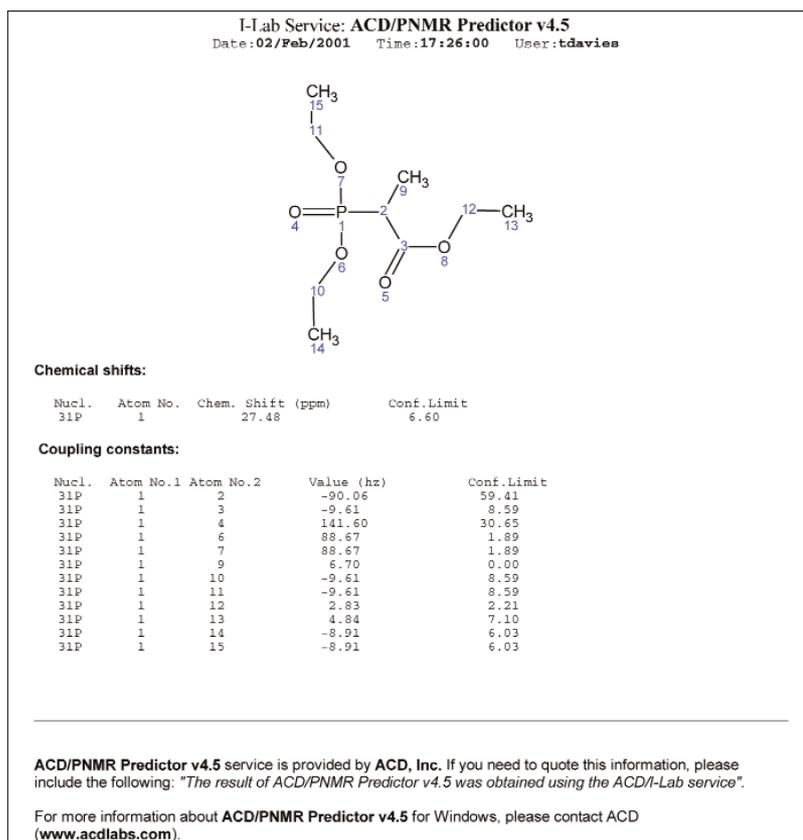


Figure 5. PNMR results are imported directly into ChemSketch as no manipulation package exists at this time for this nucleus.

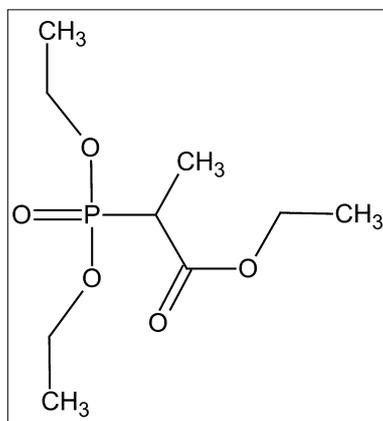


Figure 6. Ethyl 2-(diethoxyphosphoryl)propanoate (ACD/IUPAC Name v4.5) or propanoic acid, 2-(diethoxyphosphinyl)-, ethyl ester (ACD/Index Name v4.5).

absorption coefficient (K) provided a constant 116 regardless of pH as did the predicted $\log D$ values (1.3 ± 1.0).

I then decided to opt for a more chemically interesting molecule from the round robin group, D-gluconic acid sodium salt (527-07-1). This also failed to interest I-Lab producing naming problems "Cannot name structure for the following reason: Structure contains non-supported carbohydrate derivative!". When

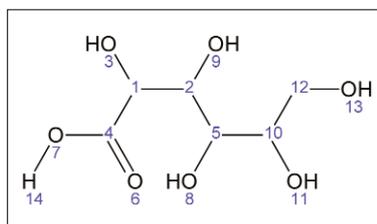


Figure 7. The acidified form of our final test molecule.

the boiling point predictor fell over with the message "Error: Molecule is charged!" and similarly LogP with "Error: Structure is charged!" and $\log D$ "Cannot calculate LogD! LogP cannot be calculated, Structure is charged" I slowly began to realise my mistake. So back to the structure editor and try again with the acid form and not the salt (Figure 7).

This time things worked. I didn't recognise the name thrown up by all three predictors "hexonic acid" but this is probably just my ignorance. Anyway it now started to throw out results one after another:

- BP ($^{\circ}\text{C}$ at 760.00 mmHg) = 673.59 ± 55.00
- Calculated LogP: -3.18 ± 0.85 (experimental data not available)

- A table of $\log D$ values varying with pH
- Six different pKa values
- $\Delta_{\text{vap}} H(T_{\text{boiling point}}) = 113.20 \pm 6.0 \text{ kJ mol}^{-1}$
- $\log_{10}(p) = -20.299 \pm 4.69$ (p in mmHg) p (at 25.00°C) = $0.0000000000 \text{ mm Hg}$
- the pKa Database search produced: Formula: $\text{C}_6\text{H}_{12}\text{O}_7$, Molecule Weight: 196.16, Ionic Form: H6L, [1] pKa(H6L/H+H5L)=3.18 at $t = 20.00$, Reference: Hirsch P. "Recueil trav. chim.", 1952, 71, 999
- The adsorption coefficients and the BCF bio concentration factors were all ~ 1 regardless of pH!

The method of obtaining the free services kept me guessing for quite some time until I gave up and decided to us the Help files(!). These parameters will be calculated from a completely different drop-down menu "TOOLS" and then only when in the STRUCTURE mode. The results are below.

Molecular formula = $\text{C}_6\text{H}_{12}\text{O}_7$
Molecular weight = 196.155
Composition = C(36.74%) H(6.17%) O(57.10%)
Molar refractivity = $38.91 \pm 0.3 \text{ cm}^3$
Molar volume = $111.2 \pm 3.0 \text{ cm}^3$
Parachor = $365.8 \pm 4.0 \text{ cm}^3$
Index of refraction = 1.616 ± 0.02
Surface tension = $116.9 \pm 3.0 \text{ dyne cm}^{-1}$
Density = $1.763 \pm 0.06 \text{ g cm}^{-3}$
Dielectric constant = Not available
Polarisability = $15.42 \pm 0.5 \cdot 10^{-24} \text{ cm}^3$

I don't have the time to check through these values but would certainly be interested to hear from anyone with more experience of using these values in their daily work.

Summing up, we have checked through all of the features available and can say that we failed to bring the system down. What we have not attempted to date is to assess the accuracy of the results of the predictions. Well, that is not strictly true as we have extensively tested the ^{13}C -NMR prediction power – but that is for another day!

Acknowledgements

The author would like to thank ACD/Labs for full access to their I-Lab services. I would also like to thank the Bundesministerium für Bildung und Forschung, and the Ministerium für Schule und Weiterbildung, Wissenschaft und Forschung des Landes Nordrhein-Westfalen, Germany, for financial support.