COMPUTER SOFTWARE REVIEWS

Software Review: ChemOffice 2005 Pro by CambridgeSoft

John J. Irwin*

Department of Pharmaceutical Chemistry, University of California San Francisco, 1700 4th Street, San Francisco, California 94143-2550

Received July 14, 2005

As suggested by its name, ChemOffice 2005 aims to be to chemistry what Microsoft Office and its imitators are to the office. Like that ubiquitous package, ChemOffice is composed of several standalone programs, which interoperate with each other as well as with Microsoft Office. The flagship program is ChemDraw, now "Ultra", for drawing and working with chemistry in 2D. For 3D chemical modeling there is Chem3D Pro, and for working with databases there is ChemFinder. The CD and DVD are packed with additional versions, features, and databases. There is a ChemDraw plugin for use in Web pages, an ActiveX control version of Chem3D for use in Web pages, PowerPoint, and other applications. There is a software development kit (SDK) to allow programmers to integrate ChemOffice features into proprietary code. There is a copy of Microsoft SQL database desktop server that provides a database server for use with ChemFinder. There is E-notebook, an electronic notebook system. Finally, there are 11 demonstration databases, including subsets of the National Cancer Institute database and chemical reactions from ISI. The package comes with three manuals, which seemed to be fine, although I usually prefer to try online help first and paper manuals only as a last resort.

In my work on virtual screening and databases of druglike compounds, I generally find that I need access to many different programs to get the results I want. So I was delighted to discover ChemOffice, a software environment that offers so many features in a single integrated package. This suite of software contains many, many features. The program generally performed well, although as I discuss below, I had a few complaints, largely due to inflated expectations of a fully integrated chemical environment. It is exceedingly difficult to make software as complex as ChemOffice, that handles 2D and 3D objects, interoperates with Microsoft Office, and handles all the complex rules of chemistry. To test the software, I tried a combination of tasks, some of which would be routine for my work; others I performed just because they were available in ChemOffice and I wanted to see how they worked.

ChemDraw has been the industry standard for drawing and representing molecules in 2D for as long as I can remember. Most chemists and many biologists will be familiar with its non-nonsense interface, the essential features of which are still recognizable from an early Macintosh version that I used nearly 20 years ago. The program's perception of chemistry, and its sophistication at representing chemical concepts in a drawing context, is unparalleled. The program checks chemical syntax as you draw, highlighting possible problems, for example, if the valence of an atom is exceeded. But ChemDraw is just the tip of the iceberg of ChemOffice. For example, one of the really great features is the ability to import a spreadsheet of chemical data, including SMILES, into Excel using the ChemSAR wizard that is installed as an Add-in. ChemSAR allows you to get data from just about anywhere in tab-delimited format and then to work with it in Excel, using all the features and chemical perception of ChemOffice. For this feature alone, this package is worth getting.

ChemDraw Ultra contains many new features that looked intriguing, such as systematic naming, a TLC plate tool, improved proton NMR predictions, mass fragmentation, more graphics file format support, and chemical markup language support among others. I tried out the improved "name to compound" and "compound to name" features. I typed in "pteridine" and asked for the structure—it was correct, except that it was upside down from the usual convention—normally N1 is in the bottom left—ChemDraw puts it in the top right. This sort of thing should be easy to fix—just orient named rings in the standard convention. Cutting and pasting of ChemDraw objects into other applications such as Word and Powerpoint were generally trouble free, although Word crashed on me a few times. I could not convince myself that it was ChemOffice's fault, but it was a little unnerving. When I typed "Pteridinium", ChemDraw protonated the pteridine on N1, as one sees in methotrexate bound to DHFR.

There were some cases where the name to structure translation was less than perfect, but I did not figure out why. So for example, whereas imidazole and imidazolium were rendered correctly, ChemDraw offered identical structures for phenylsulfonamide and phenylsulfonamidate. It also charmingly announced for charged compounds, "Caution: a charge appears to be present". These structures could generally be cut and pasted into Chem3D to yield a reasonable 3D geometry. Curiously, the protons on the nitrogen atoms of imidazolium in 3D were considerably smaller than the protons on the carbons when rendered as balls and sticks. Annoyingly, the default meaning of the left mouse button in Chem3D is "add atoms" so I found myself trying to rotate the molecule and instead adding methyl groups. Fortunately, I could undo this easily and change modes. Looking under preferences, I was unable to find out how to change the default mode.

When I brought up folate, one carboxylate was anionic, whereas the other was curiously protonated. Folate was represented in the keto form, and ChemDraw gave no hint of the enol form. When rendered into 3D by cutting and

* Corresponding author e-mail: jji@cgl.ucsf.edu.

pasting into Chem3D, it is easy to compute properties such as calculated LogP, polar surface area, net charge, and so on. Unfortunately, unless the form rendered is the biologically relevant one, computing these properties may be quite misleading. The word tautomer is not even in the ChemDraw manual.

I looked up "Viagra" and pasted it into Chem3D. The structure had a serious internal collision involving the ethoxy group and the 1H-pyrazolo[4,3-d]pyrimidin-7(6H)-one ring, as named by ChemDraw. This was a little worrying, as the collision was so bad, you would not want to use properties you had calculated on it for anything. I wanted to figure out how close the nearly touching atoms actually were. I clicked on "measure distances" and was offered a table of all pairwise distances in the molecule. Clicking on the atoms in the molecule did not take me to the relevant distances, so I had to scan through the list by eye—not entirely satisfactory! The clash was quickly resolved by running a few hundred steps of energy minimization. ChemDraw added lone pairs but not protons to amines, so I ended up minimizing a structure that I would have to fix later by hand.

Chemdraw does not seem to know that neutral amino acids are normally zwitterions under physiological conditions. So requesting "alanine" yields a structure with an amine and a carboxylic acid. When we paste this into Chem3D, we see the structure, we can calculate a LogP and other properties, but there seems to be a whole layer of chemical intelligence missing. Perhaps I just have not managed my expectations. After all, ChemDraw is a chemical drawing program—perhaps it really is the user's job to get the protonation and charge correct by himself/herself. But the counterargument is, if you are going to offer to calculate parameters used for drug discovery, like PSA and LogP, then you really need to get the representation of the molecule right first! Also, if ChemDraw is going to used in high throughput or semihigh-throughput, it must come to terms with the automatic treatment of charge and tautomers.

In addition to displaying individual or a pair of molecules, Chem3D can handle small lists of ligands. This is a welcome extension, as we ourselves often have hit lists to review from docking, lists of annotated ligands, or other subsets of the ZINC database. Although Chem3D did not seem to know about the convention of using TER records to separate ligands from docking in PDB format, it was quite adept at reading small lists of molecules in SD format. On my 1.5 GHz Pentium4M laptop with 512 MB of memory, 500 molecules were a little too many. There is the unfortunate default of displaying all 500 ligands at once, which is rarely what one wants in the first instance. Again, I was unable to find a way to change this default behavior.

ChemOffice is one of those packages that promises a lot. Some of the core functionality, such as drawing and manipulation of chemical structures, is better than ever. There really is no better program. The Excel spreadsheet feature "ChemSAR" is also very powerful and can be easily integrated into diverse computing environments. The activeX control for Chem3D, the plugin for browsers, and the SDK for programming give enormous flexibility to integrate ChemOffice features into your own applications.

ChemOffice Pro 2005 is not cheap. For companies, a single seat will set you back $2330.00, while academics can get it for a bargain $990.00. There is special volume pricing too. Is it worth it? Everyone's situation will be different. If you review and prioritize lists of compounds for ligand discovery or optimization projects, you need a lot of the features in ChemOffice Pro 2005. We certainly use it, as do many of our colleagues.

CI050286X

10.1021/ci050286x CCC: $30.25
Published on Web 08/19/2005