J Comput Aided Mol Des (2011) 25:103–106
DOI 10.1007/s10822-011-9411-7

WARR'S PIECE

Guiding effective decisions: an interview with Matthew Segall, CEO of Optibrium

Wendy A. Warr

Received: 30 December 2010 / Accepted: 2 January 2011 / Published online: 12 January 2011

Matt Segall has an M. Sc. in computation from the University of Oxford and a Ph.D. in theoretical physics from the University of Cambridge. As Associate Director at Camitro (UK), ArQule Inc. and then Inpharmatica, he led a team developing predictive ADME models and state-of-the-art, intuitive, decision-support and visualization tools for drug discovery. In January 2006, he became responsible for management of Inpharmatica's ADME business, including experimental ADME services and the StarDrop [1] software platform. Following acquisition of Inpharmatica, Matt became Senior Director responsible for BioFocus's ADMET division and in 2009 led a management buyout of the StarDrop business to found Optibrium.

W. A. Warr (✉)
Wendy Warr & Associates, Cheshire CW4 7HZ, UK
e-mail: wendy@warr.com

Interview

WAW: Optibrium has been here at this very attractive science park (IQ Cambridge) for about 18 months now. What's the history?

MS: Ed Champness (our CSO) and I founded Optibrium in 2009 as a spin-out of BioFocus. We had been responsible for the development of StarDrop [1] in Inpharmatica from 2003; Inpharmatica's ADMEnsa Interactive evolved into StarDrop. Inpharmatica was acquired by BioFocus, the service division of Gapalagos in 2006. BioFocus is a successful CRO, but by 2009 it was clear that investment in software R&D wasn't a good fit with its business model. Its predictive drug discovery databases business moved to the European Bioinformatics Institute (EMBL-EBI) in July 2008, funded by a grant from the Wellcome Trust [2] and in May 2009 Optibrium was formed to acquire the StarDrop business from BioFocus.

WAW: Who uses StarDrop?

MS: We have doubled our customers in 18 months and will have more than 30 by year-end. We have a global customer base ranging from top-ten pharma companies to small biotechs and academic groups. Some of them are listed on our web site [Pfizer, J&J, Merck Serono, BioFocus, Galapagos, Kyorin, The Campbell Family Institute for Breast Cancer Research, University of Dundee, Teijin, MRC Technology, PharmaDesign, Heptares Therapeutics] but some organizations don't like to be named, of course. We should have at least 10 more customers by the end of 2011.

WAW: How will Optibrium itself have to grow to achieve that?

MS: Our current head count is six. Growth is steady: we may recruit one or two more people in 2011. I do most of

the sales and marketing and that keeps me constantly on the move, but I do like to contribute to the R&D effort occasionally. We also have agents in Japan (PharmaDesign) who help us with sales and marketing and support our customers there. We are not a bootstrap company; we had angel funding for the start-up. We don't require additional funding to fund our continued organic growth and at the moment 50–60% of our revenues go into R&D.

WAW: **Does Optibrium have partners**?

MS: Xemistry's CACTVS toolkit [3] provides a base for perception and display of chemical structures and that enables StarDrop to provide interactive access to chemical information. Molecular Network's algorithms [4] can be accessed through StarDrop and CORINA provides high quality 3D structures as input to the P450 metabolism models. Optibrium is also a participant in Accelrys' ISV partner program and uses Pipeline Pilot [5] to create components that integrate StarDrop with Accelrys' technologies.

WAW: **That brings us on to the features of the software itself**. **Give me an outline**.

MS: StarDrop is designed to guide decisions involving complex, uncertain data in an intuitive way. It helps scientists in the design and selection of high quality potential drug molecules and is aimed at the hit-to-lead and lead optimization stages of drug discovery. Multi-parameter optimization is the challenge we address. The objectives are to identify chemistries with an optimal balance of properties, or to identify quickly situations when such a balance is not possible in order to "fail fast and fail cheap". Key to this is the principle that the user can make these decisions with confidence. So, the key features of StarDrop are probabilistic scoring, to prioritize compounds based on a user-defined profile of property criteria and weights, while allowing for uncertainty in the underlying data; and the ability to view property distributions across chemical space, selecting compounds with a balance between quality and diversity. Also, the "Glowing Molecule" feature allows users to link structures of compounds with their predicted properties and explore new ideas with immediate feedback. In other words, StarDrop's proprietary visualization allows researchers to redesign compounds interactively.

WAW: **Looking at StarDrop in action**, **I can see that a useful add-on would be something like Abbott's Drug Guru** [6], **to suggest suitable transformations corresponding to medchem design rules of thumb**.

MS: In fact we intend to do precisely that: in early 2011 the new release of StarDrop will have a proprietary module generating ideas for possible compounds, to be used either at the start of a project or during redesign. There are, in fact, many other useful tools in drug discovery and so part of our principle has been to ensure that it's easy to connect these with StarDrop. The StarDrop client has a Python script interface, linking to ODBC, SQL etc. or Web Services for Pipeline Pilot. SDfiles, SMILES, molfiles and CSV files can be opened and saved. In addition, you can plug in your own, in-house QSAR models.

WAW: **What else does StarDrop do**?

MS: There are three optional modules: the ADME QSAR module, the P450 module, and the Auto-Modeller module. The ADME QSAR module provides a suite of validated, global QSAR models for key ADME properties. The P450 module identifies sites of metabolism for major P450 enzymes, based on quantum mechanical simulation of the reaction mechanisms: AM1 semi-empirical calculations are used, although post-hoc corrections based on ab initio calculations are applied to improve the accuracy. Potential sites of metabolism are labeled with a percentage value, indicating the predicted regioselectivity of metabolism by a P450 isoform. The software also calculates site lability (a measure of a site's vulnerability to metabolism) to guide redesign, which is displayed as a "metabolic landscape" shown in red, yellow, green, and blue (where red means that the site will be most efficiently metabolized). This detailed view of the reactivity of the site is a unique selling point of the software for which quantum mechanical calculations are needed. You can also drill down and study the impact of steric accessibility, orientation etc. For example, it is well known that the ASP301 residue of 2D6 binds basic nitrogens and orients them away from the oxidation site.

WAW: **You have quite a few competitors in the ADMET space**: **products such as Schrödinger's QikProp** [7], **Simulations Plus' ADMET Predictor** [8], **ACD/ADME Suite** [9] **Molecular Discovery's MetaSite** [10], **Strand's SARchitect** [11], **tools in MOE** [12]. **What are StarDrop's unique selling points**?

MS: No one of those products has the combination of features offered by StarDrop integrated in an intuitive, user-friendly environment. Interestingly, MetaSite and StarDrop have almost opposite approaches for calculating P450 regioselectivity. MetaSite uses a detailed grid-based description of the binding pockets of the different P450s and an approximate model of the "electronic" reactivity of each site. StarDrop uses a complex quantum mechanical approach to calculating the electronic reactivity of each site and ligand-based descriptors to capture steric and orientation effects of the enzyme active site on the pattern of metabolism. The advantage of this approach is that StarDrop can calculate the site lability and composite site

lability, measures of the efficiency and hence vulnerability to metabolism, which is additional information over the regioselectivity. In terms of overall accuracy in predicting sites of metabolism, based on the published results for MetaSite, StarDrop and MetaSite are similar. Most of our customers have evaluated both StarDrop and MetaSite and some of them use both, as the different approaches can be complementary.

WAW: Lhasa Limited' METEOR is different, being an expert system, but what about Tripos [13] and Accelrys [14]?

MS: As you say, METEOR is an expert system and has quite a different objective. It addresses the question of all the possible metabolites arising from a molecule. It doesn't provide a specific prediction targeted at a chemist trying to redesign a molecule. The VolSurf component of Tripos' platform provides some ADME predictions, but this isn't really a central feature of their platform. Pipeline Pilot does have an ADMET component, but this is not a central selling point of their platform either.

WAW: Now tell me about the Optibrium Auto-Modeller.

MS: The Auto-Modeller builds QSAR models tailored to the user's own data. Many of StarDrop's users are not computational scientists and so it can be used, as the name suggests, in automatic mode, but if you are familiar with QSAR modeling then you can configure it as you prefer. For example, splitting the data set into three subsets (training, validation, and test sets) can be done automatically or that can be controlled by the user. The split method options are clustering, y-based sampling, or random. There are lots of descriptors to choose from and you can also input your own. Multiple modeling techniques are provided, including PLS, radial basis functions and six ways of training Gaussian Processes models (a Bayesian, non-linear method [15]). If you do not select specific modeling methods, all of them will be used to build models based on the training set. The models will be compared and validated using the independent validation and test sets. After the method is run, a plot is shown and you can click on any point and look at the related chemical structure. An expert can drill down to examine all the underlying details.

WAW: Can models be shared?

MS: It's very easy to save and share the final model for use in StarDrop, at which point all of its capabilities, e.g., the Glowing Molecule, can be applied to the output. Also, the Optibrium online community [16] provides a forum for sharing ideas and best practices in drug optimization. Users of StarDrop can download free add-ons including new models and scripts. The community pages are also used to make available papers and presentations that the company and others have given (not only about StarDrop) and our desktop app, StarVue is a free download from the community.

WAW: Is StarVue popular?

MS: We have given away about 800 copies on CD so far and there have been a few hundred downloads.

WAW: It doesn't have the glowing molecule feature, does it?

MS: No, but it is a very nice, free browser for chemical structures and related data. It gives you two ways to look at data: a traditional table that you can customize, or our proprietary Molecule View which summarizes all of the data for a compound on a single screen. It's a really innovative visual interface that allows you to scroll through the data set smoothly and visualize the molecule in a unique way.

WAW: There are other visualization tools out there, like Spotfire [17], Dotmatics' Vortex [18], and Synaptic Science's SEURAT [19]. How does StarDrop compare?

MS: There is a difference between *supporting* decision making and *proactively guiding* decision making. Good data visualization is necessary to support decisions, but visualization alone is not sufficient to meet our objective to guide decisions. Another example of this relates to ADME models; former ADME modeling companies such as Camitro (which was acquired by ArQule), Lion and Amedis provided predictive models, but models are not enough: you need software that helps you to use this information to make effective decisions and quickly reach the goal of your project.

WAW: What are the hurdles to getting people to let an algorithm help them make decisions?

MS: When researchers see StarDrop they want it: we have a very high success rate in converting trials into sales. The main challenge that StarDrop overcomes is to present a complex analysis in a very user friendly and intuitive way. Scientists are dealing with complex and uncertain data; they are doing multi-parameter optimization in the face of uncertainty. An algorithm can't make a decision for them; instead, it should provide a rigorous approach to assessing all of the available information and help to focus scientists' efforts on those potential solutions that are most likely to achieve their goals. A paper in *DDT* [20] discusses cognitive biases: the reproducible biases affecting human decision making. It shows how feedback on problem solving performance in simulated environments, and computational tools that encourage objective consideration of all of the available information, could help teams

improve their selection of compounds and effective screening sequences.

WAW: Terry Stouch said I should ask what primary functions are used most: decision making tools or the ADME tools?

MS: I don't think that there is a clear distinction. You can only get the full value from ADME predictions if you have a rigorous framework to use the data they provide to guide good decisions. So, I'd say that they are used hand-in-hand.

WAW: One final question (maybe it should have been the first): why "StarDrop"?

MS: StarDrop is an acronym for "Structure Activity Relationships for Drug Optimization"!

References

1. StarDrop. Optibrium Ltd., 7226 IQ Cambridge, Beach Drive, Cambridge, CB25 9TL, UK. http://www.optibrium.com/. Accessed Dec 2010
2. Warr WA (2009) ChEMBL. An interview with John Overington, team leader, chemogenomics at the European Bioinformatics Institute Outstation of the European Molecular Biology Laboratory (EMBL-EBI). J Comput Aided Mol Des 23(4):195–198
3. CACTVS. Xemistry GmbH, Hainholzweg 11, D-61462 Königstein, Germany. http://85.214.71.72/. Accessed Dec 2010
4. Molecular Networks. Molecular Networks GmbH - Computerchemie, Henkestrasse 91, 91052 Erlangen, Germany. http://www.molecular-networks.com/. Accessed Dec 2010
5. Pipeline Pilot. Accelrys, Inc., 10188 Telesis Court, Suite 100, San Diego, CA 92121, USA. http://accelrys.com/products/pipeline-pilot/. Accessed Dec 2010
6. Stewart KD, Shiroda M, James CA (2006) Drug Guru: a computer software program for drug design using medicinal chemistry rules. Bioorg Med Chem 14(20):7011–7022
7. QikProp. 120 West 45th Street, 17th Floor, Tower 45, New York, NY 10036-4041, USA. http://www.schrodinger.com/products/14/17/. Accessed Dec 2010
8. ADMET Predictor. Simulations Plus, 42505 10th Street West, Lancaster, CA 93534-7059, USA. http://www.simulations-plus.com/Products.aspx?grpID=1&cID=11&pID=13&gclid=CKSsv4G91aUCFUYe4QodOzOclA. Accessed Dec 2010
9. ACD/ADME Suite. Advanced Chemistry Development, Inc., 110 Yonge Street, 14th floor, Toronto, Ontario, Canada M5C 1T4. http://www.acdlabs.com/products/pc_admet/. Accessed Dec 2010
10. MetaSite. Molecular Discovery, Via Stoppani, 38, 06135 Ponte San Giovanni, Perugia, Italy. http://www.moldiscovery.com/soft_metasite.php. Accessed Dec 2010
11. SARchitect. Strand Life Sciences, 5th Floor, Kirloskar Business Park, Bellary Road, Hebbal, Bangalore, 560024, India. http://www.strandls.com/Sarchitect. Accessed Dec 2010
12. MOE. Chemical Computing Group, 1010 Sherbrooke St. W, Suite 910, Montreal, Quebec, Canada H3A 2R7. http://www.chemcomp.com/. Accessed Dec 2010
13. SYBYL-X. Tripos Inc., Tripos, 1699 South Hanley Road, St. Louis, MO 63144-2319, USA. http://www.tripos.com/index.php?family=modules,SimplePage,,,&page=comp_informatics&s=0. Accessed Dec 2010
14. Discovery Studio. Accelrys, Inc., 10188 Telesis Court, Suite 100, San Diego, CA 92121, USA. http://accelrys.com/products/discovery-studio/admet.html. Accessed Dec 2010
15. Obrezanova O, Gola JMR, Champness E, Segall MD (2008) Automatic QSAR modeling of ADME properties: blood-brain barrier penetration and aqueous solubility. J Comput Aided Mol Des 22(6–7):431–440
16. The Optibrium community. http://www.optibrium.com/community. Accessed Dec 2010
17. TIBCO Spotfire. Tibco Software, Inc., 212 Elm Street, Somerville, MA 02144, USA. http://spotfire.tibco.com/. Accessed Dec 2010
18. Vortex. Dotmatics Ltd.,The Old Monastery, Windhill, Bishops Stortford, Herts CM23 2ND, UK. http://www.dotmatics.com/products_vortex.jsp. Accessed Dec 2010
19. SEURAT. Synaptic Science LLC, 155 Gibbs St, Suite 430, Rockville, MD 20850-0353, USA. http://www.synapticscience.com/seurat/. Accessed Dec 2010
20. Chadwick A, Segall M (2010) Overcoming psychological barriers to good discovery decisions. Drug Discov Today 15(13/14):561–569