

## Wendy Warr Interviews Tudor Oprea



Tudor I. Oprea (TIO) graduated from the “Victor Babeş” University of Medicine and Pharmacy, Timișoara, Romania with an M.D. in general medicine in 1990; he gained a Ph.D. in molecular physiology in 1992. Past appointments include: Visiting Scientist, University of Utrecht, Holland (1990-1991); post-doctoral researcher at Washington University, St. Louis (1992-1994), and at Los Alamos National Laboratory (1994-1996); then Research Scientist (1996-2000), and Associate Director (2000-2002) at AstraZeneca R&D Mölndal, Sweden. Since August 2002, Dr. Oprea has been Professor of Biochemistry and Molecular Biology, and Chief, Division of Biocomputing, at the University of New Mexico (UNM) School of Medicine in Albuquerque, USA. Prof. Oprea has two decades of experience in cheminformatics, chemical space navigation, virtual screening, and lead identification. He has edited one book, and co-authored 19 book chapters, more than 85 peer-reviewed papers and several patents. He received the 2002 Hansch Award from the QSAR and Molecular Modeling Society. Since 2005 he has served as

Chair of the Cheminformatics and QSAR Society. He is Founder and CEO of Sunset Molecular Discovery LLC, which specializes in drug informatics and medicinal chemistry databases. He has served on the Scientific Advisory Board of ChemDiv, Inc. (Moscow and San Diego) since 2003. During his 2009 sabbatical leave from UNM, he is hosted by the Department of Systems Biology, Danish Technical University, Lyngby, Denmark (January-March) and by the Department of Chemistry, University of Perugia, Italy (April-December).

### Interview

#### Wendy A. Warr: What are your drivers in cheminformatics?

Tudor Oprea: Moving up the ladder was never the big driver for me. So I would say that my interest has clearly become “team science”. Nearly half of my papers in the last four years are biology-centric, and I am often one of the middle authors working with teams of biologists and chemists. Truth be told, I am not one to invent “yet another file format”. We have SDfile, and Mol2, and InChI and SMILES; these are all important, but what matters equally is how you apply cheminformatics technology to produce meaningful advancements in discovery. To that end I want to make cheminformatics useful to society; to focus on delivery and to provide a collaborative, supportive environment for chemistry, biology and medicine. Other people may dedicate their efforts for software development, but someone has to highlight the utility of the tools we create. This is why I support ventures such as Chris Steinbeck’s CDK [1] and Overington’s ChEMBL [2], that parallel the NIH molecular libraries program [3] and PubChem [4], in creating a meaningful open source, open access resource that fosters the use of available technologies. We need to take the tools that we have, since many of them are quite good, and apply them before we abandon them for the next fashion. In a way, the world is tired of the cheminformatics promise. Look at CADD (computer-aided drug design). It was ambitiously touted over 20 years ago. Maybe 20 to 30 drugs can by now be attributed to it, but then the world moved on to be enthusiastic about HTS (high throughput screening) and other tools. Perhaps an advancement could mean the return to earlier days, a combination of modern and “old” technologies in an appropriate mix. For example, I myself favor intensive data mining combined with medium throughput, backed up by *in vivo* experiments.

#### WAW: How did WOMBAT (World of Molecular Bioactivity) begin?

TIO: Moving to Astra in Sweden from Los Alamos was a change of climate in all respects, but I had a two year honeymoon before the merger with Zeneca. WOMBAT [5] was born while I was there. I kept using MDDR (MDL Drug Data Report), as reflected in what later became my most cited paper (over 200 citations) as single author [6], but over time I became increasingly frustrated that I could not get real numbers for biological activity. I decided to do something about it. One Christmas vacation I went to Timișoara and enlisted a team of collaborators out of my friends and former colleagues. Shortly afterwards, AstraZeneca agreed to fund my colleagues at Timișoara for one year, in an effort that led to the first version of WOMBAT. The project was terminated after that year but by then a new opportunity at New Mexico had surfaced and as I moved to the United States I also took over WOMBAT. I had to spend out of my own pocket for six months, until I made the first sale. From then on, WOMBAT has been profitable. It's a thrifty operation with small overhead and no dedicated sales staff: a labor of love. Of course I am grateful to AstraZeneca for opening up this opportunity.

**WAW: What is the future of WOMBAT?**

TIO: Recently I met Luis Suarez, who has a PhD in engineering and much experience in technology companies; he has become my business advisor at Sunset, and helped set up the new company Drug4cast [7]. This is a Delaware corporation with sales and marketing offices in Albuquerque, New Mexico, USA. We are driven by a desire for better tools to empower scientific research in the area of lead and drug discovery. We want to provide a visual as well as conceptual framework for navigating across annotated chemical libraries, and to merge chemical and biological spaces by enabling fast data mining at the chemistry-biology interface. There are existing tools for mining databases for structures and maybe activities but you cannot see the hidden answers: which structures are the more interesting answers, which structures are promiscuous, and so on. What we really want to know is what drives activity, selectivity and promiscuity.

**WAW: What are your plans for the Cheminformatics and QSAR Society (formerly the QSAR and Molecular Modeling Society)?**

TIO: About a year after taking over the society, I began to lose some faith in predictive sciences as a route to anything of true novelty. We continue to deal with “unknown unknowns”, and therefore should continuously explore new ways. Truth is, all predictions are rooted in the past, since we derive predictions only based on something we have previously encountered. By now much time and effort has gone into analyzing the past by way of machine learning and QSAR, and not enough novelty is coming from using QSAR based predictions, unless one enables multiple safety checks, such as those pioneered in Alex Tropsha's group at UNC. Hence, at this time I am not excited about QSAR-centric predictions, except for drug safety and toxicology. The fact is projects are well advanced by the time one can perform a QSAR. As far as novelty goes, it is a bit late in the game by then. My bottom line here is that there is no magic bullet in drug discovery and a multi-prong approach makes best sense. And so as I continue this soul-searching quest for drug discovery, with new tools and perspectives, I am also trying to integrate the Cheminformatics and QSAR Society into this bigger picture, in particular by way of joint, cross-disciplinary meetings focused on translational sciences. Indeed, for the first time in its 20-year history, the Society is a co-organizer of EuroQSAR in 2010 [8]. Also, Wiley-VCH is closing down the journal *QSAR and Combinatorial Science* and is in the process of starting a new journal, called *Molecular Informatics* with Gisbert Schneider, Knut Baumann, Jordi Mestres and Gerhardt Ecker as editors, in 2010.

**WAW: You have a number of recent publications involving flow cytometry**

My interest in high throughput flow cytometry, HyperCyt, goes back to 2001, when I met Larry Sklar. As I joined UNM, their flow cytometric throughput was 50 compounds a week. I encouraged them by purchasing the Prestwick library, while at the same time bringing ChemDiv [9] on-board. Some exciting results later, including a first agonist on a new estrogen receptor [10],

Larry was successful in establishing UNM as one of 10 national NIH centers for HTS, based on the flow-cytometric approach. With Cristian Bologna, my team supported the HyperCyt effort with cheminformatics. Meanwhile the chemistry department at UNM was renamed as the Department of Chemistry and Chemical Biology and we hope to foster new collaborations over there; we continue to support HyperCyt using “SAR by Commerce”. The HyperCyt-based center worked on improving these technologies from 2005-2008, and by now Sklar was able to renew the NIH grant for another six years, as of 2008. We are now one of the nine molecular libraries screening centers looking for new chemical probes.

Biology is where it is all happening now, but a good thing about QSAR is that it paved the way to virtual screening. Indeed, the combination of virtual and biomolecular screening can deliver something useful from scratch to sub-micromolar in 3-6 months. UNM has already delivered at least seven probes, the worst of them around 500 nM, and several of them double- or single-digit nanomolar. An integral part of this effort was our collaboration with Jeff Arterburn, a chemistry professor at New Mexico State University [11]. Between 2004 and 2008, the NIH produced 64 chemical probes, which we recently discussed with a team of experts [12].

I am currently on sabbatical leave in Europe, developing an extension of this HyperCyt work that's currently NIH funded. For example, the ATP binding cassette transporters are a multiplex target in the Sklar group. There are 48 such transporters, with different effects related to drug resistance in e.g., bacteria and cancer cells. I am working with Gabriele Cruciani and Fabio Brocatelli in this area, at the University of Perugia, where we conduct virtual screening; at the same time, UNM is running HyperCyt studies on several ABC transporters. Earlier this year I was in Denmark, working with Søren Brunak and Olivier Taboureau on endocrine disruptors and infertility: the Danish team does the systems biology while we jointly do the cheminformatics.

**WAW: It has been claimed that fragment-based drug design helps with the “druggability” of new targets. Do you agree?**

No, this is a case of *nomen est omen*, since fragment-based drug discovery does precisely what its name implies: it uses fragments to identify ligands. Will they be drugs? Only time and clinical trials can tell. New targets are indeed a risk since we lack the ability to identify new targets. To ensure druggability you need clinical observation: you need to evaluate the effects of target perturbation under clinical conditions. And even if the target is the right one, it might still not lead to economic success, although it might meet a critical need. This recently happened to Pfizer, who launched Selzentry, an anti-HIV drug that blocks the CCR5 chemokine viral co-receptors. The drug works very well on CCR5-tropic HIV patients, but has not witnessed commercial success. We need to ask ourselves whether we really need a fifth or sixth target or drug class for high blood pressure, when we should be focusing on other, less popular needs. Just a few weeks ago, the NIH launched the TRND program, embarking on full-fledged drug discovery for rare, orphan and neglected diseases [13].

**WAW: Where is the discipline of cheminformatics heading?**

TIO: I believe that the future of cheminformatics lies in its ability to provide an integrative, predictive framework that links biological sciences, in particular systems chemical biology [14]. There is ample evidence that network biology will soon be coupled with systemic simulation tools at the molecular level, and chemistry cognizance is one of the missing links. In a sense, the age of cheminformatics is just dawning, and its future is related to chemical biology. We need to create an integrated framework, and add a temporal dimension to our models for perturbation of biological processes *via* small molecules. Not only chronic perturbations, such as upon ingesting drugs, but also acute, rapidly evolving phenomena should be subject to systems chemical biology simulations. Thus, whether we're going to study metabolic, signaling and regulatory pathways, or chronic toxicity, the field is wide open.

**References**

1. Chemistry Development Kit. <http://www.qsarworld.com/qsar-workflow1.php> (accessed May 8, 2009).
2. ChEMBL. <http://www.ebi.ac.uk/chembl/> (accessed May 29, 2009).
3. Molecular Libraries Program. <http://mli.nih.gov/mli/> (accessed May 29, 2009).
4. PubChem. <http://pubchem.ncbi.nlm.nih.gov/> (accessed May 29, 2009).
5. Olah, M.; Ramona, R.; Ostopovici, L.; Bora, A.; Hadaruga, N.; Hadaruga, D.; Moldovan, R.; Fulas, A.; Mracec, M.; Oprea, T. I. WOMBAT and WOMBAT-PK: bioactivity databases for lead and drug discovery. *Chemical Biology* **2007**, *2*, 760-786.
6. Oprea, T. I. Property distribution of drug-related chemical databases. *J. Comput.-Aided Mol. Des.* **2000**, *14*(3), 251-264.
7. Drug4cast. <http://drug4cast.com> (accessed May 8, 2009).
8. EuroQSAR. <http://www.euroqsar2010.gr> (accessed May 29, 2009).
9. ChemDiv. <http://www.chemdiv.com> (accessed May 28, 2009).
10. Bologna, C.; Revankar, C. M.; Young, S. M.; Edwards, B. S.; Arterburn, J. B.; Parker, M. A.; Tkachenko, S. E.; Savchuck, N. P.; Sklar, L. A.; Oprea, T. I.; Prossnitz, E. R. Virtual and Biomolecular Screening Converge on a Selective Agonist for GPR30. *Nat. Chem. Biol.* **2006**, *2*, 207-212.
11. Dennis, M.; Burai, R.; Ramesh, C.; Petrie, W.; Alcon, S.; Nayak, T.; Bologna, C.; Leitão, A.; Brailoui, E.; Deliu, N. S.; Dun, N. S.; Sklar, L. A.; Hathaway, H.; Arterburn, J. B.; Oprea, T. I.; Prossnitz, E. R. *In vivo* effects of a GPR30 antagonist. *Nat. Chem. Biol.* **2009**, *5*, 421-427.
12. Oprea, T. I.; Bologna, C. G.; Boyer, S.; Curpan, R. F.; Glen, R. C.; Hopkins, A. L.; Lipinski, C. A.; Marshall, G. R.; Martin, Y. C.; Ostopovici-Halip, L.; Rishton, G.; Shoichet, B. K.; Ursu, O.; Vaz, R. J.; Waller, C. L.; Waldmann, H.; Sklar, L. A. A crowdsourcing evaluation of the NIH chemical probes. *Nat. Chem. Biol.* **2009**, in press.
13. NIH Therapeutics for Rare and Neglected Diseases program. <http://www.nih.gov/news/health/may2009/nhgri-20.htm> (accessed May 29, 2009).
14. Oprea, T. I.; Tropsha, A.; Faulon, J.-L.; Rintoul, M. D. Systems Chemical Biology. *Nat. Chem. Biol.* **2007**, *3*, 447-450.

Dr. Wendy A. Warr, Wendy Warr & Associates ([wendy@warr.com](mailto:wendy@warr.com), <http://www.warr.com>), May 2009