



Published in final edited form as:

J Mol Graph Model. 2012 September ; 38: 360–362. doi:10.1016/j.jmglm.2012.07.007.

Teach-Discover-Treat (TDT): Collaborative Computational Drug Discovery for Neglected Diseases

Johanna M. Jansen¹, Wendy Cornell², Y. Jane Tseng³, and Rommie E. Amaro⁴

¹Novartis Institutes for Biomedical Research, Global Discovery Chemistry / Oncology & Exploratory Chemistry, Emeryville, CA 94608

²Merck Research Labs, Knowledge Discovery/Knowledge Management, Rahway, NJ 07059

³Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan, 106

⁴Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093

Abstract

Teach – Discover – Treat (TDT) is an initiative to promote the development and sharing of computational tools solicited through a competition with the aim to impact education and collaborative drug discovery for neglected diseases. Collaboration, multidisciplinary integration, and innovation are essential for successful drug discovery. This requires a workforce that is trained in state-of-the-art workflows and equipped with the ability to collaborate on platforms that are accessible and free. The TDT competition solicits high quality computational workflows for neglected disease targets, using freely available, open access tools.

Keywords

Neglected diseases; drug discovery; computational workflows; computational tools; Teach-Discover-Treat

Introduction

The drug discovery endeavor is influenced by a dynamic interplay between advances in science and technology on one hand and global economic forces on the other hand, which together impact the development of clinically useful compounds. Industrial and academic labs are under increased pressure to foster successful collaborations in order to maximize innovation and translational impact. Correspondingly, there is a growing need for enhanced integration across relevant disciplines, from basic biological sciences, to chemical sciences, to understanding the role of emerging new physical experiments, and how animal and clinical studies are carried out under constantly changing regulatory guidelines.

For decades, drug discovery has been uniquely suited to industry, but academic and government labs are becoming increasingly more active in drug discovery efforts. Much of

© 2012 Elsevier Inc. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

the knowledge and experience of how to develop drugs has been gathered within proprietary boundaries, where specialized groups employ state-of-the-art chemical and biological tools and technologies to identify promising small molecule compounds and manipulate them for specificity, absorption-distribution-metabolism-excretion (ADME), and toxicology properties. The recent increase in drug discovery focus in academia has highlighted the fact that students, and even those with advanced degrees, are sorely lacking in the practical know-how required to meaningfully participate in these endeavors. The need to educate drug discovery scientists in a multi-disciplinary way, and in particular, to teach students the real-world tools that are being used in actual drug discovery pipelines, presents a clear challenge to the community.

At the same time, increased globalization has underscored the need to develop drugs for the developing world. Discovering treatments for the so-called neglected diseases provides a platform to explore new ways of collaboration between industry, academia, foundations, governments and international agencies. Neglected diseases are those which afflict the world's poorest populations who cannot afford to pay for treatment, and thus they provide little incentive for development of treatments in the free market. The World Health Organization (WHO) estimates that one sixth of the world population is afflicted with one or more neglected diseases¹. Many of these diseases are neglected tropical diseases (NTDs), prevalent in areas where the hot climate and dense population promote the growth and transmission of infective organisms. Despite the lack of economic incentive, there is an urgent unmet need for medicines.

A new initiative is underway to address these multi-faceted challenges within the computational chemistry community. Teach – Discover – Treat (TDT) is both an effort to promote the development and sharing of computational tools and a challenge to the community of computational chemists to impact education and collaborative drug discovery for neglected diseases.

TDT - challenge to the computational chemistry community

Our challenge to the computational chemistry community addresses the need for better collaborative and educational tools to train a more sophisticated drug discovery workforce and to improve and advance our field. To this end, in conjunction with the American Chemical Society's Division of Computers in Chemistry (COMP), we have launched a competition to drive the development and sharing of computational models and tutorials for drug discovery against neglected diseases. These materials will be solicited through a competition and made available through the initiative's website <http://www.teach-discover-treat.org/>.

The competition requires the use of open access computational tools to encourage the wide use of the best practices and provide the opportunity for further improvements and innovation through collaboration. A critical component of the submissions to the competition is the tutorial material. Submissions will be judged based on the quality of the tutorials as well as on the quality of the computational models (readers are referred to the website for more detailed information on the judging criteria). All the submissions that meet basic requirements regarding clarity of tutorials, quality of models, and availability of computational tools will be made available through the TDT website for use in teaching and as starting points for innovation and collaboration.

There are four categories in the competition, which are summarized in Table 1. Three of the categories are focused on specific neglected diseases (malaria and sleeping sickness) for which data sets are provided through the website. Specific requirements for the workflows are outlined in the readme files that are part of the data download packages. Real-life impact

on these three challenges will be realized through experimental follow-up on the winning submissions, which is being made possible through partnerships for compound acquisition, synthesis, and biochemical testing. A fourth "open innovation" category seeks innovative drug discovery workflows that are either exemplified on a neglected disease project or adaptable to a neglected disease application.

The timeline for the competition is presented in Table 2. Four winners will be selected and receive a travel award to cover up to \$1,500 in travel expenses to attend the Spring 2013 ACS National Meeting in New Orleans and present their work in the COMP TDT symposium. Winners in the first two categories will have the opportunity to have at least 100 of their 1000-top ranking compounds purchased by COMP and tested by the academic partners, Anang Shelat and Kip Guy at St. Jude and Margaret Phillips at University of Texas Southwestern Medical Center. The winning submission in the third category will be used to prioritize synthesis efforts in the lab of Michael Pollastri at Northeastern University.

Future Prospects for Collaborative Drug Discovery

In the 2010 WHO report "First report on neglected tropical diseases: Working to overcome the global impact of neglected tropical diseases,"² Director-General Margaret Chan described a change in strategy: "The logic has changed: instead of waiting for these diseases to gradually disappear as countries develop and living conditions improve, a deliberate effort to make them disappear is now viewed as a route to poverty alleviation that can itself spur socioeconomic development." Her introductory note presented and ended with a positive message, "Aiming at their complete control and even elimination is fully justified, and this report sets out the solid evidence needed to achieve control. Above all, it makes the case for doing more, as an international community, to relieve hidden misery, on a grand scale, among people who would otherwise suffer in silence."²

The public, private and non-profit sectors have responded jointly to this gap by expanding resources for neglected disease research. In January 2012 thirteen pharmaceutical companies, the U.S., U.K. and U.A.E governments, the Bill & Melinda Gates Foundation, the World Bank and other global health organisations committed to working together to eradicate or control ten NTDs by 2020.³

This focus on neglected disease research comes at a time when there is increased support for drug discovery efforts in both the public and non-profit sectors. NIH supports academic drug discovery efforts and Francis Collins has described "four NIH-funded facilities that collectively have the capacity of a midsized pharmaceutical company to do high-throughput screening, assay development and medicinal chemistry."⁴ Further, in 2012 the NIH established the Center for Advancing Translational Science (NCATS), which aims to position academicians more firmly in the "bench-to-bedside" pipeline. NCATS consolidates existing programs such as Therapeutics for Rare and Neglected Diseases (TRND - established 2008), Components of the Molecular Libraries Program, the Clinical and Translational Science Awards (CTSA) program, and the new Cures Acceleration Network (CAN) under a single umbrella supported by \$760 million in annual research funding.⁵

The recent roadmap established by the WHO combined with the increasing availability of relevant and sophisticated drug discovery tools in the public sector makes this an ideal time to pursue neglected disease research. The TDT initiative is a call to action for the computational chemistry community to develop high quality models for neglected disease targets, using freely available, open access software, and furthermore, to share those models and methods with the global drug discovery community. We hope that our reward-based competitions sponsored through the ACS COMP division will help nucleate initial efforts. With such traction established, it is our expectation and ultimate goal that TDT will form

one of the cornerstones of computational drug discovery education and outreach for the community. If successful, we anticipate a substantial positive impact for both educating a new student workforce, as well as towards making inroads into developing treatments for diseases that affect the world's most vulnerable populations. We invite you to join us in this endeavor.

Reference

1. World Health Organization. Geneva, Switzerland: 2007. Global Plan To Combat Neglected Tropical Diseases 2008–2015.
2. World Health Organization. Geneva, Switzerland: 2010. First WHO report on neglected tropical diseases: working to overcome the global impact of neglected tropical diseases.
3. Bill & Melinda Gates Foundation. Private and Public Partners Unite to Combat 10 Neglected Tropical Diseases by 2020. <http://www.gatesfoundation.org/press-releases/pages/combating-10-neglected-tropical-diseases-120130.aspx>.
4. Mullard A. An audience with Francis Collins. *Nat. Rev. Drug Discov.* 2011; 10:14. [PubMed: 21151031]
5. Collins FS. Reengineering translational science: the time is right. *Sci. Transl. Med.* 2011; 3:90.

- We present a call-to-action for the community in the form of the Teach-Discover-Treat (TDT) initiative.
- TDT will sponsor a 4-category competition to develop state-of-the-art CADD workflows.
- Workflows must use open-access software in order to promote collaboration and innovation.
- We seek high-quality drug discovery tutorials to educate a new cadre of scientists.
- Tutorials will describe CADD workflows developed around targets in neglected diseases.

Table 1

Competition Categories and Data Sets

#	Challenge	Provided to contest entrants*	To be submitted by entrants*
1	Workflow to analyze HTS data & build models for further hit finding Partnership - Anang Shelat, Kip Guy, St. Jude	<ol style="list-style-type: none"> 1 Single point percent inhibition data for two primary screens on 305,568 compounds 2 Structures for 1,056 compounds in external, held-out test set 	<ol style="list-style-type: none"> 1 Tutorial including ligand and data preparation, hitlist triaging, hit selection, model building & validation with internal and external test sets 2 Predictions for external test set of compounds 3 Rank-ordered list of top-1000 commercially available compounds from eMolecules predicted to be active
2	Structure-based design workflow, new chemotypes Partnership - Margaret Phillips, University of Texas Southwestern Medical Center	<ol style="list-style-type: none"> 1 IC50 data for 193 triazolopyrimidine compounds against malarial dihydroorotate dehydrogenase, DHODH 2 Structure of compound with unpublished binding mode <p>Available in the public domain: multiple liganded crystal structures</p>	<ol style="list-style-type: none"> 1 Tutorial including protein and ligand preparation, binding model generation, and SAR rationalization 2 Predicted binding mode for new chemotype (an unpublished crystal structure is available as a held-out test) 3 Rank-ordered list of top-1000 commercially available compounds from eMolecules predicted to be active
3	Structure-based design workflow, medicinal chemistry strategy Partnership - Michael Pollastri, Northeastern University	<ol style="list-style-type: none"> 1 IC50 and % inhibition data for 35 compounds against Tbr phosphodiesterase, TbrPDEB1 2 Virtual library of 1.8M compounds proposed as potential TbrPDEB1 inhibitors <p>Available in the public domain: crystal structures of homologous proteins</p>	<ol style="list-style-type: none"> 1 Tutorial including sequence alignment, template selection, protein preparation, homology model construction, SAR rationalization, and compound prioritization using a multi-objective function 2 PDB file for trypanosomal TbrPDEB1 homology model 3 Rank-ordered list of top-50 preferred candidates for synthesis from the virtual library
4	Innovative drug discovery workflows, e.g. target ID, scaffold-hopping, property predictions, and target selection / druggability assessment	No explicit data sets provided Can use data from challenges 1–3 or other relevant datasets, including proprietary data	<ol style="list-style-type: none"> 1 Tutorial for a computational chemistry workflow relevant to drug discovery and either exemplified on a neglected disease project or adaptable to a neglected disease application. <p>Note: If models are built on proprietary data, the proprietary data need not be released as long as the model and software needed to run the model are freely available to all and example data files are provided.</p>

* Identifiers and canonical smiles strings are provided for each compound. Full details on provided data and data required for submissions are provided on the TDT web site.

Table 2

Timeline

Launch of initiative	March 25, 2012 (ACS Natl Mtg)
Data sets available for download	May 2, 2012
Close of submission site	September 5, 2012
Announcement of winners & start of drug discovery efforts	October 17, 2012
Award presentations	April 7–11, 2013 (ACS Natl Mtg)