

# Computational tools and interdisciplinary ingenuity for accelerating the development of new medicine

Matthew E. Welsch<sup>1\*</sup> and Sotirios G. Stergiopoulos<sup>1,2</sup>

<sup>1</sup>A2A Pharmaceuticals Inc., San Francisco, USA

<sup>2</sup>Department of Internal Medicine, Jacobi Medical Center, Albert Einstein College of Medicine, Bronx, New York, USA

Advances in big data and therapeutic target validation technology has yielded a multitude of new targets that would constitute first-in-class medicines [1]. Despite this deluge of information identifying safer and more effective pharmacological strategies, the targets of nearly half of all new molecular entities belong to three classes (GPCRs, ion channels, nuclear receptors) [2].

The rising costs of research and development, coupled with the low probability of reaching approval, have resulted in many big pharma's being risk adverse and less likely to pursue novel targets. Research and development to point of marketing costs have been increasing at an annual rate of 8.5% above inflation [3]. On the order of 90% of the compounds entering clinical trials do not meet the requirements for FDA approval [1]. Insufficient safety and efficacy are estimated to account for 66% and 21% of failures at phase III, respectively [4]. Taking into consideration the associated costs of these program failures, a new prescription drug approval is estimated to require an \$2.588 billion USD [3].

How can the cost of development be lowered and the risk mitigated? Common strategies are centered around taking the road as extensively traveled as possible. Introducing minor modifications to existing drugs has proven to be a much less rocky path to development. This also allows for the revenue streams from blockbuster drugs to be retained instead of 'falling off the patent cliff' to the generics market, which is estimated to nearly double from 2010 to 2017 [4,5]. But do these new variants constitute an improvement? An analysis of 122 drugs approved between 1999 and 2005 indicated that only 13 were superior to already available drugs [6]. The drawbacks of this approach are evident in the field of antibiotic discovery. Drug resistance has rapidly emerged to third- and fourth-generation cephalosporins [7] leading to fifth-generation drugs, such as ceftobiprole. Other approaches involve not creating a new drug at all, these include: utilizing new combinations of existing therapeutics, new formulations, or drug repurposing for a new indication [8].

There has been a gravitation away from higher risk (and higher reward) pharmacological targets and creating truly novel chemical entities. This can be attributed, in part, to the challenges encountered by major research initiatives within big pharma. Proteins that elicit their physiological effect through non-enzymatic interactions with other proteins have been considered by many to be the most exigent of the untapped target classes [9]. In examining the high-throughput screening program success rates of AstraZeneca (2004-2008) by target type, of the 16 different classes only protein-protein interaction targeted programs (5 total) failed to provide a lead candidate [10]. Another example are the efforts in antibiotic development by GlaxoSmithKline from 1995-2001 [11]. To extend beyond the confines of these "me too"

antibiotics, they evaluated over 300 genes for essentiality in bacterial strains for new drug targets. This led to 70 high-throughput screening campaigns, the majority of which failed to even yield any hits [11].

While these discouraging results lend credence to pursuing lower risk alternatives; where there are challenges, there are opportunities.

Analysis of the sizes of screening collections for GlaxoSmithKline, Novartis, Sanofi-Aventis and Wyeth ranged from roughly 300,000 to 2,000,000 molecules [12]. The consensus estimate of the chemical space (all possible stable molecules) for compounds having a molecular weight of less than 500 Da is on the order of 1060 different molecules [10]. To put these numbers into context, as of November 11, 2016 there are 92,366,303 molecules reported in the PubChem database – a minuscule fraction of the possibilities.

To successfully modulate these desirable untapped targets, we need to extend beyond the confines of existing libraries. This necessitates exploring chemical space in a manner that hones in on molecules matching target specific features, with physicochemical properties/structural characteristics that will enable successful development. Diversity oriented synthesis is one approach to build libraries with members in new regions of chemical space, wherein the strategy is to access as many structurally distinct architectures in as few reactions as possible [13]. This, however, is limited by the availability of resources for compound synthesis.

Recently, computational enumeration of every stable combination up to just 17 atoms of C, N, O, S and halogens yielded 166,443,860,262 distinct molecules (GDB17 database) [14]. Many of these compounds in the GDB17 have exotic ring systems and multiple stereogenic centers, and are extremely difficult synthetic feats. The challenge becomes how to most accurately and efficiently parse through billions of molecules to the synthetically accessible, and therapeutically relevant structures.

Computational tools used for structure-based drug design have played a role in the discovery and/or optimization of over a dozen small molecules that have entered the clinic [15]. Of the numerous computational resources available, molecular docking programs are the most widely utilized [15]. Docking programs use a selected three-dimensional target structure, and provide a calculation of affinity

---

**Correspondence to:** Matthew Welsch, PhD, Consultant and Founder, A2A Pharmaceuticals Inc., San Francisco, USA, Tel: +216-240-5151; E-mail: matthew.welsch@atoms2applications.com or DrStergiopoulos@aol.com

**Received:** November 05, 2016; **Accepted:** November 19, 2016; **Published:** November 21, 2016

(score) and predicted binding orientation (pose) for a ligand input. These programs are commonly used in the virtual high-throughput screening of available libraries, for the selection of only minor subset for physical testing in laboratory assays. Performing docking prior to testing has been shown to dramatically improve hit rates compared to conventional HTS [16]. However, scores are notoriously uncorrelated with observed affinity, attributed to multiple potential points of error (scoring function accuracy, physiological relevance of the structure used, solvation effects, etc.) [17]. This approach has proven to be the most valuable for narrowing down the realm of possibilities, as opposed to precise affinity comparisons; facilitating more focused testing. The accuracy for hit identification is increased substantially through using intricate knowledge of the target biochemistry and combining multiple computational tools [18], which are selected based on the specific characteristics of the target and ligand set.

Other approaches have been fruitful in recent years for yielding new drugs are those involving fragment-based drug discovery (FBDD) [19]. FBDD has led to more than 30 compounds in clinical trials, including two FDA approvals [19]. The general concept is simple – start small and work your way up. This entails the initial screening of low-molecular weight compounds (< 300 Da) at high concentrations (> 500  $\mu$ M) [19]. The generally very weak target binders identified in fragment screens undergo iterative rounds of analoging and assay evaluation until a potent, drug-like lead is produced for further testing. While typically labor intensive, FBDD approaches extend beyond the confines of chemical space encompassed by in stock libraries, to help yield novel compounds with high specificity. The exemplification of this is with the breakthrough drug venetoclax for the treatment of chronic lymphocytic leukemia [19]. This orally bioavailable, 868 Da small molecule is a completely novel structure that began with fragments as low as 150 Da [19].

While R&D costs continue to rise with a low probability of program success, advancing technologies and new strategies are showing great promise for drug discovery. How can productivity and innovation be further improved? – Interdisciplinary cross-talk.

There needs to be a transition out of one size fits all “drug-like”, and tailor nascent discovery approaches to the biology of the specific indication. This is particularly pressing for antibiotic development [20]. There has been a prohibitive disconnect between biochemical assay activity to efficacy against the live bacterium [20]. This has been largely attributed to the unique prokaryotic physiology and lack of an understanding of what makes compounds “antibiotic-like” [20]. Overcoming biological challenges, such as cell wall permeation and drug extrusion by efflux pumps, can most productively facilitated through a close concerted effort by biochemists, microbiologists and synthetic chemists.

In general, physiochemical properties of drugs and the regions of occupied chemical space can differ significantly depending on the indication [21,22]. Discovery programs would benefit from a more extensive molecular descriptor profile, and incorporating better predictive models for toxicology and other factors associated with attrition [23,24]. The earlier in a program an interdisciplinary framework for discovery can be implemented, the greater the opportunity for innovative new drugs.

We propose that success can be improved by beginning with even smaller starting points (<100 Da) and utilizing computational enumeration with retrosynthetic analysis, to create extensive target-specific libraries. These same principles that have led to success in

fragment screening can be streamlined using computational tools, and with the three-dimensional structure of the target and thorough knowledge of its biochemistry as a guide. Starting with sub-structures, such as ring systems in successful drugs [25], has the advantage of more space on a molecular canvas. This opens the door for flexibility in design to complement the unique features of the target biochemistry; while incorporating better predictive models of the factors associated with attrition in the iterative design and selection process.

Meeting these challenges necessitates a much more cohesive integration of emerging computational technologies with multiple disciplines, to form a facile framework for the early discovery process.

## References

1. Plenge RM, Scolnick EM, Altshuler D (2013) Validating therapeutic targets through human genetics. *Nat Rev Drug Discov* 12: 581-594. [[Crossref](#)]
2. Kinch MS, Hoyer D, Patridge E, Plummer M (2015) Target selection for FDA-approved medicines. *Drug Discov Today* 20: 784-789. [[Crossref](#)]
3. DiMasi JA, Grabowski HG, Hansen RW (2016) Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ* 47: 20-33. [[Crossref](#)]
4. Mignani S, Huber S, Tomás H, Rodrigues J, Majoral JP (2016) Why and how have drug discovery strategies in pharma changed? What are the new mindsets? *Drug Discov Today* 21: 239-249. [[Crossref](#)]
5. Light DW, Lexchin JR (2012) Pharmaceutical research and development: what do we get for all that money? *BMJ* 345: e4348. [[Crossref](#)]
6. van Luijn JC, Gribnau FW, Leufkens HG (2010) Superior efficacy of new medicines? *Eur J Clin Pharmacol* 66: 445-448. [[Crossref](#)]
7. Yoshino Y, Okugawa S, Kimura S, Makita E, Seo K, Koga I, et al. (2015) Infective endocarditis due to *Enterobacter cloacae* resistant to third- and fourth-generation cephalosporins. *J Microbiol Immunol Infect* 48: 226-228. [[Crossref](#)]
8. Gupta SC, Sung B, Prasad S, Webb LJ, Aggarwal BB (2013) Cancer drug discovery by repurposing: teaching new tricks to old dogs. *Trends Pharmacol Sci* 34: 508-517. [[Crossref](#)]
9. Arkin MR, Tang Y, Wells JA (2004) Small-molecule inhibitors of protein-protein interactions: Progressing towards the reality. *Chem Biol* 3: 1102-1104. [[Crossref](#)]
10. Barker A, Kettle JG, Nowak T, Pease JE (2013) Expanding medicinal chemistry space. *Drug Discov Today* 18: 298-304. [[Crossref](#)]
11. Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* 6: 29-40. [[Crossref](#)]
12. Macarron R, Banks MN, Bojanic D, Burns DJ, Cirovic DA, et al. (2011) Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov* 10: 188-195. [[Crossref](#)]
13. Schreiber SL (2000) Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 287: 1964-1969. [[Crossref](#)]
14. Ruddigkeit L, van Deursen R, Blum LC, Reymond JL (2012) Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. *J Chem Inf Model* 52: 2864-2875. [[Crossref](#)]
15. Talele TT, Khedkar SA, Rigby AC (2010) Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. *Curr Top Med Chem* 10: 127-141. [[Crossref](#)]
16. Kapetanovic IM (2008) Computer-aided drug discovery and development (CADD): in silico-chemico-biological approach. *Chem Biol Interact* 171: 165-176. [[Crossref](#)]
17. Scior T, Bender A, Tresadern G, Medina-Franco JL, Martínez-Mayorga K, et al. (2012) Recognizing pitfalls in virtual screening: a critical review. *J Chem Inf Model* 52: 867-881. [[Crossref](#)]
18. Gao C, Thorsteinson N, Watson I, Wang J, Vieth M (2015) Knowledge-Based Strategy to Improve Ligand Pose Prediction Accuracy for Lead Optimization. *J Chem Inf Model* 55: 1460-1468. [[Crossref](#)]
19. Erlanson DA, Fesik SW, Hubbard RE, Jahnke W, et al. (2016) Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov* 15: 605-619. [[Crossref](#)]

20. Lewis K (2013) Platforms for antibiotic discovery. *Nat Rev Drug Discov* 12: 371-387. [[Crossref](#)]
21. Vieth M, Siegel MG, Higgs RE, Watson IA, Robertson DH, et al. (2004) Characteristic physical properties and structural fragments of marketed oral drugs. *J Med Chem* 47: 224-232. [[Crossref](#)]
22. Lipinski C, Hopkins A (2004) Navigating chemical space for biology and medicine. *Nature* 432: 855-861. [[Crossref](#)]
23. Zhu X, Kruhlak NL (2014) Construction and analysis of a human hepatotoxicity database suitable for QSAR modeling using post-market safety data. *Toxicology* 321: 62-72. [[Crossref](#)]
24. Waring MJ, Arrowsmith J, Leach AR, Leeson PD, Mandrell S, et al. (2015) An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat Rev Drug Discov* 14: 475-486. [[Crossref](#)]
25. Taylor RD, MacCoss M, Lawson AD (2014) Rings in drugs. *J Med Chem* 57: 5845-5859. [[Crossref](#)]

**Copyright:** ©2016 Welsch ME. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.