
COMPUTING TO THE TARGET: ACCELERATING ORPHAN DRUG DISCOVERY

BY KARA JIA

When most people think about pharmaceutical drugs, the first ones that come to mind are usually popular over-the-counter (OTC) drugs such as Tylenol, Advil, and Zyrtec. These medications are seemingly omnipresent; they can be found in ordinary places, from local pharmacies to convenience stores. While OTC drugs are among the most accessible medications to all consumers, they constitute only the tip of the iceberg of all existing and potential pharmaceutical drugs. Prescription drugs are required to treat most medical conditions apart from the common cold. New drugs are constantly under development for major diseases. Yet while treatments for rare diseases inflicting a much smaller population may be in the stages of drug development, many more remain non-existent.

What exactly sets rare diseases—also known as orphan diseases—apart from common ones? Orphan diseases affect a small subset of a population. In United States, such conditions typically affect less than 1 out of 200,000 people, and currently there are about 7,000 such diseases affecting approximately 25 million patients.⁶ These medical conditions generally lack attention from pharmaceutical companies as they do not yield large commercial success. Government efforts to increase economic incentives led to the passing of the Orphan Drug Act (ODA) in 1983, which encourages development and marketing of drugs for rare diseases. Among the main incentives for orphan drug research and development are grants, research design support, and a seven-year exclusive orphan drug marketing.⁴

Since the enactment of ODA, the number of orphan drugs produced has significantly increased. There is a perpetual need for orphan drug development as about 250 new rare diseases are reported each year.⁴ At present, drugs for rare cancer subtypes represent the majority (31%) of approved orphan drugs.⁶ Fabrazyme, an enzyme replacement therapy, is an example of an approved orphan drug used to treat Fabry's disease, an ultra-rare X-linked genetic disorder caused by deficiency in the enzyme alpha-galactosidase A. Under this disease, gastrointestinal symptoms, ophthalmological symptoms, and neuropathic pain manifest in

early childhood due to the accumulation of glycosphingolipids—lipids that are part of the cell membrane and crucial for cell-cell interactions—in different organs.¹⁰

Developing drugs such as Fabrazyme entails a time-consuming, costly, and difficult process. Basic research first generates the necessary data that initiates and fuels the discovery phase of drug development. During the discovery phase, potential candidate molecules—or leads—are selected. A lead is a chemical compound with pharmacological or biological activity that is potentially therapeutically useful. Lead compounds enter a pre-clinical validation stage; selected candidate leads are then permitted to enter clinical phases. If the clinical trials are successful, the FDA filing process for the manufacturing of an orphan drug begins.

Drug discovery is initiated for diseases that do not yet have suitable medical products. Identifying a target molecule is the start of a lengthy process, which involves a concerted effort between academic research institutions and the pharmaceutical industry. The drug discovery process can be divided into the following stages: initial target identification and validation, assay development, high throughput screening, hit identification, lead optimization, and selection of a lead molecule for clinical development. Only 10% of potential leads actually make it to clinical trials.¹ Because there is a high attrition rate for compounds entering the clinical phase, novel techniques could significantly help optimize the drug discovery process. A failed drug not only impedes the goal of bringing effective treatment to patients but also leads to higher financial consequences for pharmaceutical companies, thus slowing down the whole drug development process. With an average of \$2.6 billion invested in developing a marketable drug, finding new techniques that would help reform the drug discovery pipeline thus remains a major priority. The efficiency of computers has made *in silico* methods an indis-



“While treatments for rare diseases inflicting a much smaller population may be in the stages of drug development, many more remain nonexistent.”

pensable tool for expediting the process of finding promising and novel drug leads.⁵

Compared to traditional methods such as high-throughput screening (HTS), computer-aided drug design (CADD) tools have been shown to be more efficient and complementary.² Two main types of CADD are structure-based drug design (SBDD) and ligand-based drug design (LBDD). SBDD is a method focusing on predicting and analyzing 3D structures of target proteins (i.e. finding a molecule with the best fit in an active site). Prediction of target structure is based on the assumption that “proteins with similar sequences have similar structures.”² Molecular dynamics simulations are used in SBDD to reveal the various pathways for ligands to interact with target proteins as well as the different possible target conformations.⁷ Accumulation of biological data makes SBDD studies particularly favorable.⁹ On the other

hand, LBDD is used to analyze possible ligands that interact with the target of interest to determine whether the small molecule increases or decreases target activity. It is used when the protein structure cannot be experimentally or computationally determined and relies on information about known active ligands.⁷

Other than finding novel drug leads, repositioning approved drugs is another possible approach to orphan drug development. Repurposing FDA-approved drugs is relatively inexpensive compared to developing de novo drugs.⁸ Duloxetine is a serotonin and norepinephrine reuptake inhibitor that has been approved to treat depression. Since serotonin and norepinephrine were found to be key neurotransmitters in fibromyalgia (a central nervous system disorder) and chronic musculoskeletal pain, the duloxetine pathway was successfully repositioned

to treat these disorders.⁶ One way to discover novel properties of approved drugs is through the use of computational methods; these insights could then be used to develop drugs for orphan diseases.⁸ For this purpose, machine learning and text mining are two promising computational approaches.

One proposed machine learning model creates algorithms devised to help predict unknown drug-disease associations.⁸ This method integrates genome, phenome, and chemical structure information into a computational framework that extracts a drug similarity matrix and a disease similarity matrix. Insights gained from this approach could potentially aid drug repositioning by revealing novel drug indications.⁹

Taking a more indirect approach, text mining can be used to extract particular terms and phrases from electronically-stored literature. Given a set biological ontology, text mining could then be used



“The efficiency of computational methods has made in silico methods an indispensable tool for expediting the process of finding promising novel drug leads.”

Figure 1: Fabrazyme is an approved orphan drug used to treat Fabry's disease.¹¹

for retrieval of relevant information on a particular drug in literature which could yield novel indications for existing drugs.⁸ Different institutions produce data organized in different ways; thus, text mining could also prove useful in standardizing heterogeneous data.³

Combined with economic incentives and efficient computational methods, more opportunities are created for orphan drug discovery as well as general drug development. While each computational method has its own set of limitations and challenges, the drug discovery process benefits from any new, potentially useful insights that can minimize drug discovery time for the many orphan diseases which are, as of yet, still incurable.

REFERENCES

1. Hughes, J., Rees, S., Kalindjian, S., & Philpott, K. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, 162(6), 1239–1249.
2. Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014, January). *Computational Methods in Drug Discovery*.
3. Yao, L., Evans, J. A., & Rzhetsky, A. (2010, April). Novel opportunities for computational biology and sociology in drug discovery.
4. Seoane-Vazquez, E., Rodriguez-Monguio, R., Szeinbach, S. L., & Visaria, J. (2008, December 16). Incentives for orphan drug research and development in the United States.
5. Molecular dynamics-driven drug discovery: leaping forward with confidence. (2016, November 25).
6. Li, Y. Y., & Jones, S. J. (2012, March 30). Drug repositioning for personalized medicine.
7. Leelananda, S. P., & Lindert, S. (2016, December 12). *Computational methods in drug discovery*.
8. Muthyala, Ramaiah. Orphan/rare drug discovery through drug repositioning. (2011, November 04).
9. Li, J., Zheng, S., Chen, B., Butte, A. J., Swamidass, S. J., & Lu, Z. (2015, March 31). A survey of current trends in computational drug repositioning.
10. Karpman, D., & Höglund, P. (2016, October 13). Orphan drug policies and use in pediatric nephrology.

IMAGE SOURCES

11. <https://upload.wikimedia.org/wikipedia/commons/e/e2/FabrazymeProductPicture.jpg>.
12. https://upload.wikimedia.org/wikipedia/commons/thumb/9/97/Docking_representation_2.png/1200px-Docking_representation_2.png.

Figure 2: The docking of a ligand to a protein target leading to the formation of a protein-ligand complex.¹²

