OIDD SCREENING: *IN SILICO* EVALUATION OF CHEMICAL STRUCTURES

1. OVERVIEW

The objective of the OIDD Screening program is the identification of attractive, drug-like molecules for *in vitro* testing, thus maximizing the chances of finding a structural class previously untapped at Lilly that warrants further exploration regarding its therapeutic potential.

This is the desired endpoint for both Lilly and the OIDD participants. However, in consistency with our commitment to safeguarding our partners' intellectual property, Lilly does not obtain direct access to the actual chemical structures submitted to OIDD. Instead, a completely automated set of computational tools collectively referred to as the OIDD cheminformatics evaluation is utilized to identify which molecular structures are eligible for biological evaluation in the OIDD *in vitro* screening modules without accessing the actual chemical structure.

The cheminformatics evaluation is accomplished in two phases.

- In the first phase, the initial structure-related processing takes place within the DMZ, a secured, neutral area of the OIDD platform, which is outside the Lilly internal facing network.
- In the second phase, additional informatics processing takes place inside the Lilly network.

Moreover, the Lilly scientific community **does not** have access to the chemical structure. The entire evaluation system is designed to ensure that only a limited number of Lilly OIDD IT personnel, out of necessity, have access all aspects of the system.

Figure 1 provides a schematic representation of the cheminformatics evaluation process and its most important features.

2. EVALUATION IN THE DMZ

The DMZ, or demilitarized zone, is an IT-industry term referring to a neutral zone within a system or platform which protects information from being accessed either from the inside as well as the outside of an organization. The initial stages of the cheminformatics evaluation of OIDD submitted structures take place inside the system's DMZ, *but outside the Lilly internal-facing firewall*.

At this stage, each molecular structure gets converted to a set of descriptors or "fingerprint" which is used henceforth for necessary calculations and analyses in lieu of the structure itself. Each newly generated fingerprint is then compared to the full set of all fingerprints previously submitted to the OIDD program, which is stored in the DMZ in order to facilitate the comparisons while protecting the confidentiality of the submission. Molecules that are found to be more than 85% similar to a previous submission will be discarded. Next, the Lilly MedChem Rules are used to identify and discard pharmaceutically undesirable motifs[1], in a manner similar to what is done at many pharmaceutical companies[2].

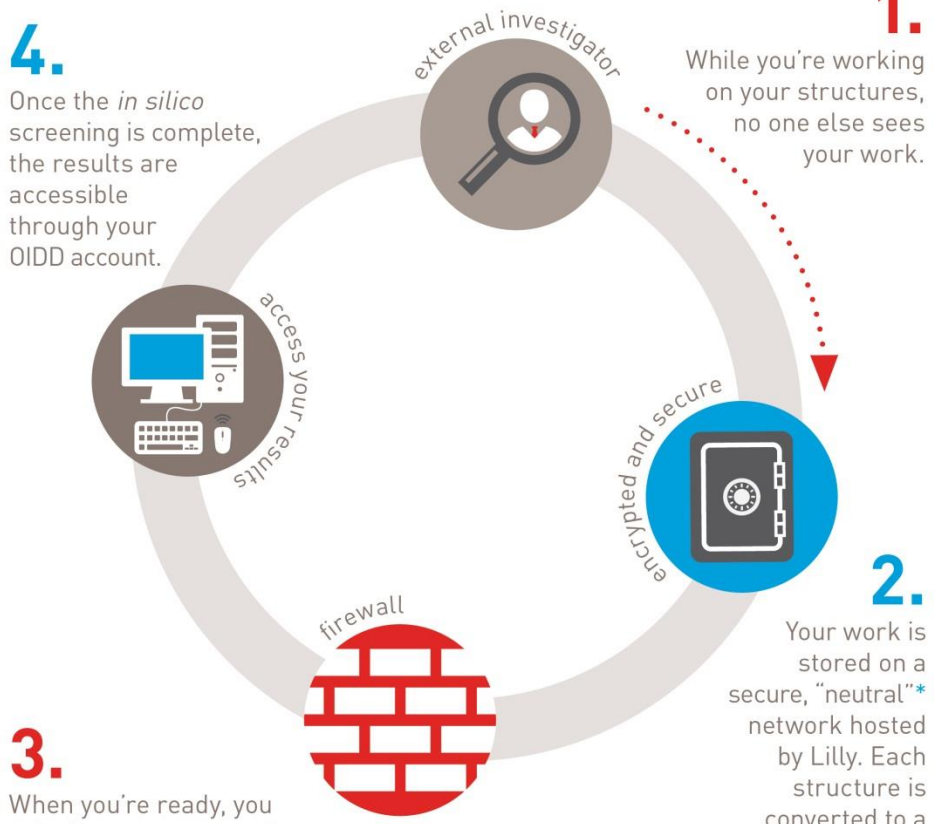
[1] Bruns, R.F.; Watson, I.A. "Rules for Identifying Potentially Reactive or Promiscuous Compounds," *J. Med. Chem*, **2012**, 55(22), 9763-9772

[2] I. Muegge, "Selection criteria for drug-like compounds," *Medical Research Reviews*, **2003**, 23(3), 302-32 [see for example references 28, 29]

OPEN INNOVATION DRUG DISCOVERY STRUCTURE SECURITY PROVISIONS

Making sure your discoveries stay yours.

The Lilly OIDD Security Provisions ensure a secure environment for you to design and investigate drug-like molecules without compromising the confidentiality of your structures.



1. While you're working on your structures, no one else sees your work.

2. Your work is stored on a secure, "neutral"* network hosted by Lilly. Each structure is converted to a set of calculated descriptors to be used in further analysis.

3. When you're ready, you submit your structure for *in silico* screening. Only the corresponding descriptor profile is sent across the firewall to the Cheminformatics Lilly servers.

4. Once the *in silico* screening is complete, the results are accessible through your OIDD account.

*The neutral zone, often referred to in the IT industry as a "demilitarized zone," or "DMZ," protects information from access inside or outside Lilly.

Figure 1

Once a submitted structure is processed, these fingerprints are the only information that crosses the Lilly internal-facing firewall into the internal cheminformatics server. Fingerprints that cross the internal-facing firewall do not contain any identifying information that would enable the discovery of a link to an outside entity. They are accompanied only by a system-generated random identifier.

None of the dictionary-based fingerprints are used, where each bit would correspond to a known molecular feature. All of the fingerprints are algorithmically computed, offer complete coverage of the molecule and are subject to collisions. This results in stronger obfuscation of the structure and enforces the confidentiality built into the OIDD process.

3. EVALUATION INSIDE THE LILLY NETWORK

For molecules passing the initial filters in the DMZ, their fingerprints are passed through the Lilly internal-facing firewall in order to undergo a series of similarity checks against several molecular collections. While all these similarity checks could determine the identity of submitted molecules in case of an *exact match*, the computations are set up to *only record a match above a threshold* – not to retain that similarity value. This feature of the evaluation algorithm is an additional design element intended to afford additional protection to the OIDD participants' intellectual property.

The fingerprint similarity evaluations are as follows:

a) Comparison with Controlled Substances
b) Comparison with Known Drugs
c) Comparison with Lilly Collection
d) Comparison with Pubchem Collection

3.1. Comparison with Controlled Substances

Lilly does not wish to receive controlled or illegal substances. We perform a similarity check against a list of more than 2,100 controlled substances. When a fingerprint exhibits more than 90% similarity to a known controlled substance, the molecule is discarded. We cannot guarantee that our list of controlled substances is complete for any particular jurisdiction, so OIDD participants are required to observe their local laws.

3.2. Comparison with Known Drugs

Submitted fingerprints are checked against a set of known drug molecules. Fingerprints more than 85% similar to a known drug will be discarded.

3.3. Comparison with Lilly

Fingerprints are compared against the Lilly Collection and any that are similar to an existing Lilly molecule will be discarded.

3.4. Comparison with Pubchem

Fingerprints are compared against a Lilly-hosted copy of the Pubchem database. During this comparison, fingerprints and/or original structures *are not sent outside Lilly.* Any fingerprints that are more than 85% similar to something in Pubchem will not be discarded and this score will not be used as selection criteria, but simply as a profiling feature.

4. CONCLUSIONS

The key objective of the cheminformatics evaluation process is to reduce or minimize the non-interesting, non-drug-like molecules submitted for *in vitro* screening. This, in turn, maximizes the chances of screening compounds that will provide the biological profile(s) targeted for further study. In the process of performing these evaluations, we take great care to protect the confidentiality of the molecular structures submitted by OIDD participants.

Due to the inherent capacity constraints of the OIDD screening panel, it is expected that Lilly will not be able to accept all submitted compounds for screening. Passing the cheminformatics evaluation step does not guarantee that a compound will be accepted for screening.

As the OIDD process evolves, the associated cheminformatics evaluations will also evolve to increase the chance of discovering interesting compounds that form the basis of a collaboration agreement between Lilly and the OIDD external investigator. Your contributions to the early drug discovery process are invaluable. We look forward to the opportunity to work with you as we collaborate to advance science.