Drug Repurposing for Neglected Diseases: Leishmaniasis

Hao Xu, Xinran Li, Jonathan Smith, Jahir Gutierrez, Laurence Yang

Machine Learning in Computational Biology Workshop 2021

Steps in identifying good drugs:

1. **Identify a protein target** for a drug.
   a. The protein's function will be inhibited or activated by the drug so that it triggers a cellular response (e.g. cell death, nutrient starvation)
   b. Proteins may participate in a functional pathway essential for survival of the parasite. When the pathway function is compromised, it may lead to growth inhibition or death in the parasite.
2. **Identify a drug molecule** that best binds with the target protein
3. Satisfy **constraints on design of the drug** molecule
   a. Minimizing cost of synthesis
   b. Minimizing likelihood of side-effects (eg. toxicity)
   c. Maximizing likelihood of oral administration
4. **Identify drug combinations** that boost efficacy (synergistic polypharmacy)

Neglected disease because:

- Existing drugs are quite toxic and costly.
- Most investment avoids research for these diseases (due to market size).
- Leishmaniasis has wide-reaching impact (>1 million diagnoses each year)

Goal to develop new treatments: easily administered, cheap, with few side-effects.

We show that existing binding affinity models can be used to select drugs in the molecule selection stage for repurposing: lacosomide shows promise as a drug that targets the Leishmania enzyme sterol 14-alpha demethylase, an essential component for membrane biogenesis.

- Final selected model predictions had high correlation with molecular docking simulations.

We also show that certain molecular descriptors correlate with strong molecular docking for drugs targeting leishmania proteins: VSA (Surface Area) in particular.

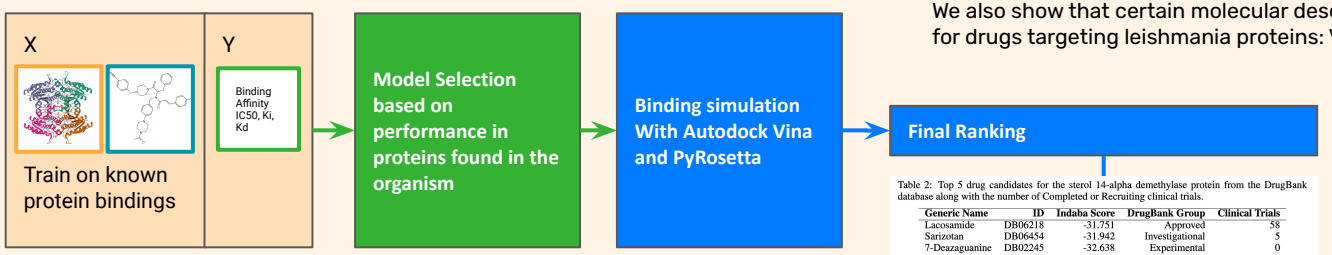

Table 2: Top 5 drug candidates for the sterol 14-alpha demethylase protein from the DrugBank database along with the number of Completed or Recruiting clinical trials.

Generic Name	ID	Indaba Score	DrugBank Group	Clinical Trials
Lacosamide	DB06218	-31.751	Approved	58
Sarizotan	DB06454	-31.942	Investigational	5
7-Deazaguanine	DB02245	-32.638	Experimental	0
Roxadustat	DB04847	-33.086	Investigational	43
Hydracarbazine	DB09243	-34.847	Experimental	0

Table 3: Spearman correlation of predicted binding affinity with label.

Molecule Encoder / Protein Encoder	CNN / CNN	Daylight / AAC	Morgan / AAC	Morgan / CNN	MPNN / CNN	Sample size
Leish (Kd)	0.6713	0.7342	0.5244	0.4056	-0.0349	12
Non-Leish (Kd)	0.6502	0.7077	0.6244	0.6459	0.0950	1200
Leish (IC50)	0.3739	0.4593	0.0948	0.1481	0.0895	156
Human (IC50)	0.1617	0.4218	0.2351	-0.1707	0.3179	438
Leish (Ki)	0.3004	0.1688	0.2697	0.3686	0.0455	99
Human (Ki)	-0.0613	-0.0007	0.0409	0.2007	-0.0234	404
Rand. Non-Leish (Ki)	0.0765	0.1584	0.1581	0.1819	0.0488	100

Table 1: Pearson Correlations between methods on the top 597 pairs.

Pearson Correlation	Vina	DeepPurpose	Indaba Score	MONN
Vina	-	0.3412	0.1936	0.3999
DeepPurpose	-	-	0.5512	-
Indaba Score	-	-	-	0.3684
MONN	-	-	-	-

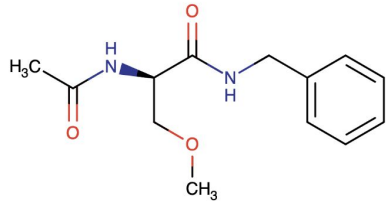

Figure 1: Chemical structure of Lacosamide