# *In silico* design of a novel SMYD3 inhibitor with Darwin

François Charih[1,2,3], Skye Wang[2], Kyle K. Biggar[2,3], James R. Green[1,3]

[1]Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada
[2]Institute of Biochemistry, Carleton University, Ottawa, ON, Canada
[3]NuvoBio, Ottawa, ON, Canada

Let's get in touch!

Carleton University

Carleton University Biomedical Informatics Collaboratory

NuvoBio
Intelligent peptide design

T-K-B Biggar Lab

## Background

- Lysine methylation is an important post-translational modification of proteins carried out by lysine methyltransferases (KMTs).
- The methylation status of a protein has consequences on its subcellular localization, stability, interactions with other proteins, among others.
- Overexpression of the KMT **SMYD3** is observed in several cancers, including pancreatic, lung, gastric, breast, ovarian, and esophageal cancer [1]. As such, SMYD3 constitutes a prime target for inhibition.
- Few peptide drugs have been approved for use in the clinic, but they are becoming increasingly promising, as they can be designed to disrupt protein-protein interactions or inhibit enzymes with high specificity.
- Designing peptide inhibitors is challenging due to the vastness of the synthesizable peptide space. Moreover, only an infinitesimal fraction of that space corresponds to chemically relevant peptides, *i.e.* that can bind and/or inhibit a specific protein target.
- The *In Silico Peptide Synthesizer* (InSiPS) [2] is a genetic algorithm-based program that optimizes a peptide *binder* sequence, using the PIPE interaction score of the sequence with the target as a measure of fitness.
- InSiPS was shown to have the ability to generate peptide binders for yeast and viral targets [2,3], but the computational requirements (*i.e.* thousands of CPUs) make InSiPS difficult to deploy at larger scales and in more complex organisms such as humans.

## Objectives

- Implement a more *computationally efficient* sequence-based peptide binder design algorithm, using InSiPS as a starting point.
- Design an effective peptide inhibitor against SMYD3 to demonstrate Darwin's ability to design a bioactive inhibitor for a human protein target *in silico*, using sequence information only.

## Darwin for *in silico* peptide design

- Similar to InSiPS, our implementation, named ***Darwin***, uses a genetic algorithm-based strategy to optimize the binder sequence.
- We designed a new, proprietary fitness function that favours high SPRINT interaction score [4] for the peptide-target pair and penalizes high off-target interaction scores.

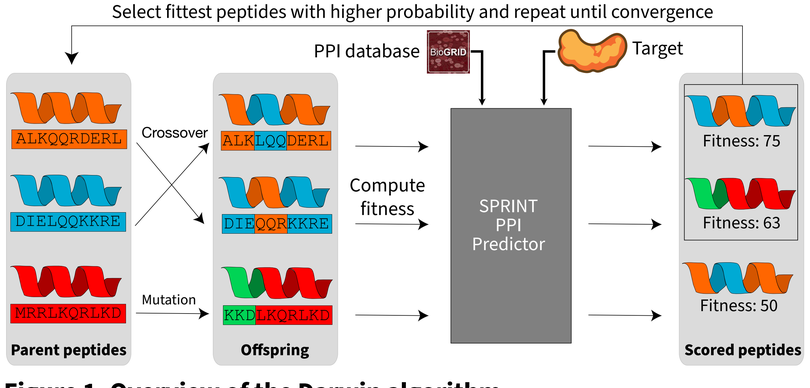



**Figure 1. Overview of the Darwin algorithm.**

The Darwin algorithm is strongly inspired by InSiPS, and uses a genetic algorithm optimization strategy to produce peptide binders using a protein-protein interaction score as a measure of peptide fitness.

## Design of anti-SMYD3 inhibitors with Darwin

We ran Darwin to design short peptides (20-mers) against the SMYD3 methyltransferase with a population size of 2000. We considered 20 iterations without improvement in maximum fitness to indicate convergence.

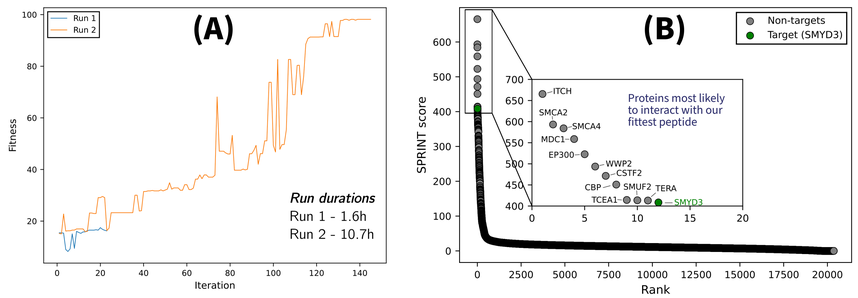



**Figure 2. *In silico* design of anti-SMYD3 peptide inhibitors.**

**(A)** Evolution curve for 2 Darwin runs against SMYD3 on a 11-node computer cluster with 440 Intel Xeon 2.40GHz CPUs. **(B)** One-to-all curve showing the SPRINT interaction scores between the fittest peptide and all proteins exposed to it.

## Assessment of our anti-SMYD3 peptides binding profiles

We investigated whether the peptides generated with Darwin actually interact with SMYD3 *in vitro*. To this end, we printed peptides on SPOT arrays, and incubated the peptide with the target protein and a non-target protein to assess interaction specificity.

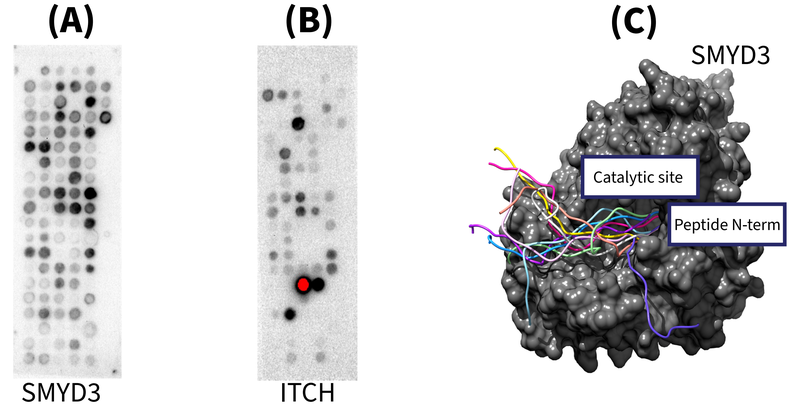



**Figure 3. Assessment of the binding activity of Darwin-generated peptides.**

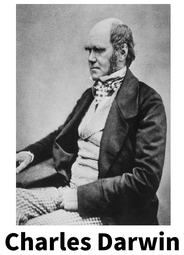
**(A)** We printed 96 peptides on SPOT arrays, and incubated the peptides with our target SMYD3 to identify actual binders. **(B)** We repeated the experiment with ITCH (ubiquitin protein ligase), a high-ranking off-target protein to assess off-target interactions. **(C)** Using HPEPDOCK [5], we docked our lead peptide in SMYD3 (PDB: 7O2C; ligand removed). Shown are the the 10 most thermodynamically-favoured poses predicted.

## Darwin: behind the name

You may wonder why we named our algorithm *Darwin*...

Genetic algorithms, the optimization strategy that underpins our work relies heavily on a core concept described by Charles Darwin's theory of evolution: *survival of the fittest*.

Our algorithm iteratively *evolves* peptides. The "fittest" (most likely to specifically bind the target protein) peptides in a given generation are more likely to be selected for the next generation, and to subsequently give rise to fitter peptides when subjected to mutation and/or crossover operations.



**Charles Darwin**
(1809-1882)

## Evaluation of inhibition activity

Given that Darwin only aims to produce peptides that bind to a selected target, we sought to determine whether the observed binding is associated with a reduction in the enzymatic activity of the target.

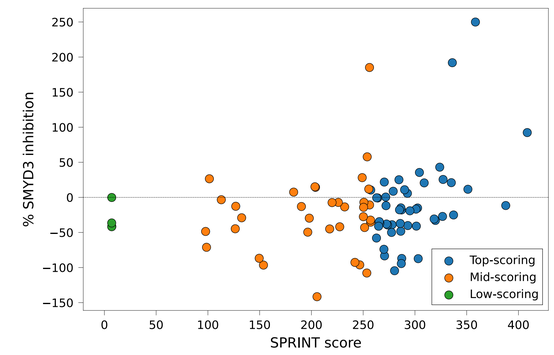



**Figure 4. *In vitro* assessment of enzyme inhibition for 96 peptides with a luminescence-based methyltransferase assay.**

Using a MAP3K2 fragment known to be a SMYD3 substrate, we measured the reduction in methylation signal upon incubation of SMYD3 with the substrate and our 96 Darwin-generated peptide binders.

## References

**[1]** Fasano, C. *et al.* Identifying novel SMYD3 interactors on the trail of cancer hallmarks. *Comput Struct Biotechnology J* 20, 1860–1875 (2022).

**[2]** Burnside, D. *et al.* In silico engineering of synthetic binding proteins from random amino acid sequences. *iScence* 11, 375–387 (2018).

**[3]** Hajikarimlou, M. et al. A computational approach to rapidly design peptides that detect SARS-CoV-2 surface protein S. *Nar Genom Bioinform* 4, lqac058 (2022).

**[4]** Li, Y. & Ilie, L. SPRINT: ultrafast protein-protein interaction prediction of the entire human interactome. *BMC Bioinformatics* 18, 485 (2017).

**[5]** Zhou, P., Jin, B., Li, H. & Huang, S.-Y. HPEPDOCK: a web server for blind peptide–protein docking based on a hierarchical algorithm. *Nucleic Acids Res* 46, gky357- (2018).

## Acknowledgements

This work was made possible by financial support and equipment generously provided by the following organization:

NSERC CRSNG

cancer computer
#ComputingForTheCure

## Highlights

- We have shown that Darwin can generate peptide binders for human enzyme targets *in silico*, using sequence only, and with little computational resources.
- Some of the peptide binders generated with Darwin for the SMYD3 enzyme are bioactive, *i.e.* achieve enzyme inhibition.
- For inquiries or collaboration opportunities, please visit https://nuvobio.com.