Novel In Silico Generation of a Synthetic Peptide Library Derived from Biological Complexes for Protein and Vector Engineering

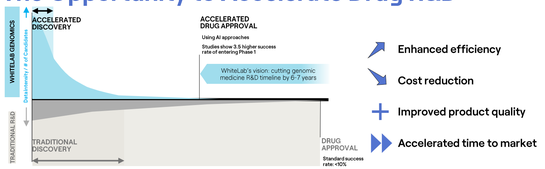
B. Dafniet[1]*, C. Alliot[1], Y. Habtoun[1], Y. Vander Meersche[1], C. Colas[1], J. Maes[1], D. Del Bourgo[1], J. Cottineau[1], D. Serillon[1]
[1] WhiteLab Genomics, Future4Care, 8 rue Jean Antoine de Baïf, 75013 Paris, France (*) presenting author

WHITELAB GENOMICS
Unleash the Potential of Genomic Medicine using AI

Genomic Medicine's Principal Challenges

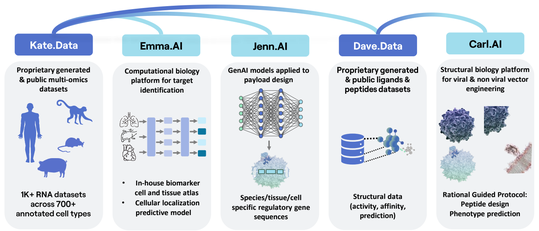
Specificity | Immunogenicity | Manufacturing complexity

The Opportunity to Accelerate Drug R&D



- Enhanced efficiency
- Cost reduction
- Improved product quality
- Accelerated time to market

ALFRED, Proprietary & Modular AI Platform



About WhiteLab Genomics

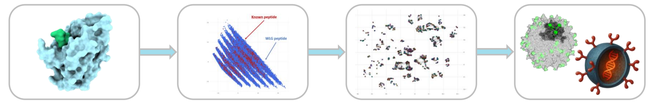
WhiteLab Genomics stands at the convergence of AI and biology. Founded in 2019, backed by Y-Combinator, WhiteLab is pioneering the accelerated development of life-saving genomic medicines. By leveraging their proprietary technology, WhiteLab analyzes complex biological data powered by AI to significantly reduce development timelines and mitigate associated risks. Based on exhaustive datasets, the platform provides in-silico simulations to discover and design optimized payloads and vectors.

45+ specialists in AI and biology, including 20+ PhDs and PharmDs, in Paris and Boston.

bdafniet@whitelabgx.com

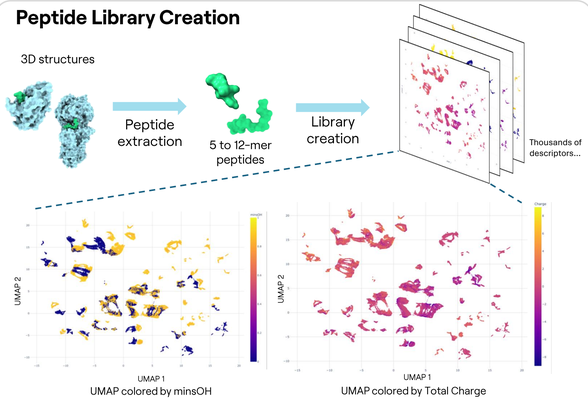
Introduction

The generation of chemical libraries, specifically peptide libraries, has been an ongoing topic for decades. Peptide design remains a key focus due to peptides' desirable properties, such as stability and binding specificity, as therapeutic agents and, in cell and gene therapy (CGT), for their targeting capabilities. Despite progress, peptide libraries are still less common and diverse than their small-molecule counterparts. To bridge this gap, WhiteLab Genomics has developed a proprietary library based on structurally resolved protein–peptide complex data, enhancing both data-driven and rational peptide and vector design.
More importantly, a dedicated "Diversity" sub-library was built to accelerate therapeutic peptide discovery and vector engineering. Based on key peptides, it allows exploration of broad chemical space with minimal redundancy, boosting efficiency in identifying optimal candidates.



Methods

Peptide Library Creation

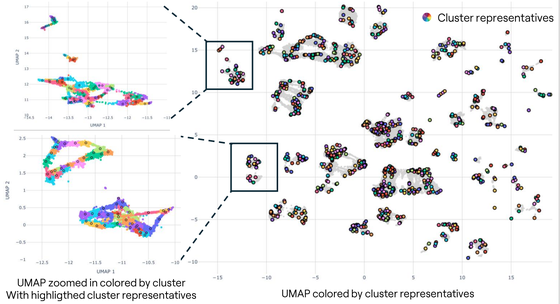


- Generation of more than a million unique structures
- Derived from active biological complexes
- Library generated from a diverse range of 2D/3D descriptors of interest for active targetting
- Useful for tailored machine-learning or artificial intelligence models

Library Content

- Mapping of all possible conformations for exhaustive similarity search
- Cyclized peptides can:
  - Mimic AAV insertion constraints
  - Optimise pharmacokinetics and pharmacodynamics properties

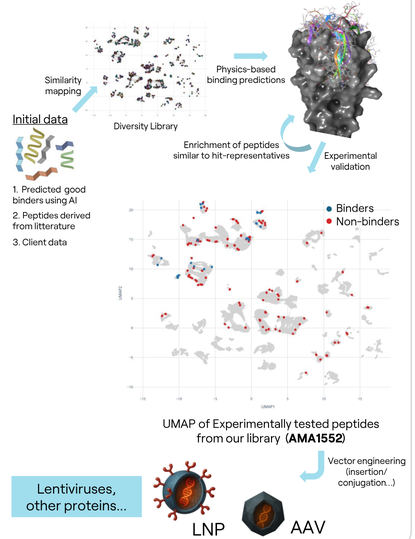
Peptide Diversity Library



Clusterisation performed using k-means algorithm:

- Iterations of cluster centroids mapping neighbouring peptides $\arg\min_{c_i \in C} dist(c_i, x)^2$ for point assignment *dist* being the euclidean distance, c_i being the cluster centroids and x the point to assign
- Since the first centroids are put at random an interation step is needed backed by $c_i = \frac{1}{|S_i|} \sum_{x_i \in Si} x_i$ with S_i the set of all points assigned to the *ith* cluster

Rational and AI-Guided Vector Design



Conclusion

By creating this peptide library, we were able to map and increase the chemical space of all active biological complexes, leading to a leap in active targeting efficiency. Exploring multiple types of 3D conformations and using cyclized peptides sustains a flexible and exhaustive approach, its capabilities enabling it to target any extracellular receptor with both viral and non-viral vectors.

Next steps

The library will be expanded to tackle broader challenges, such as incorporating longer peptides for complex binding, and extended beyond peptides to include larger, functional biomolecules like antibodies and nanobodies for advanced therapeutic applications.