Predicting peptides permeability of complex peptides

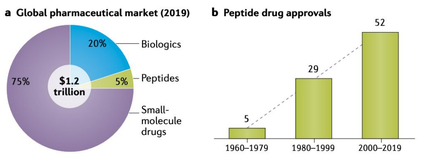
Selma Bengaouer[1], Julian Streit[1], Eleonora Serra[2], Massina Abderrahmane[1], Brice Hoffmann[1], Nicolas Devaux[1], Maud Jusot[1]

1. Iktos SAS, Paris, France ; 2. Istituto Italiano di Tecnologia, Genoa, Italia

- ❖ Predicting peptide properties using machine learning methods has gained interest in recent years, including permeability.
- ❖ Existing predictive approaches have shown interesting performance, but also limitations (restricted to natural amino acids).
- ❖ To address these challenges, new representations of peptides have been developed including multiple peptide fingerprints.
- ❖ Additionally, the incorporation of 3D descriptors related to flexibility through enhanced sampling methods is being explored.

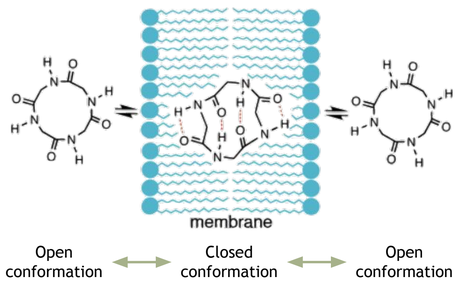
Peptides as therapeutic drugs

Muttenthaler *et al.*, Nature reviews Drug discovery 20.4 (2021): 309-325.



Advantages	Drawbacks
High affinity High specificity Low toxicity Diverse applications	Degradation / Stability Permeability Flexibility Synthesis can be challenging

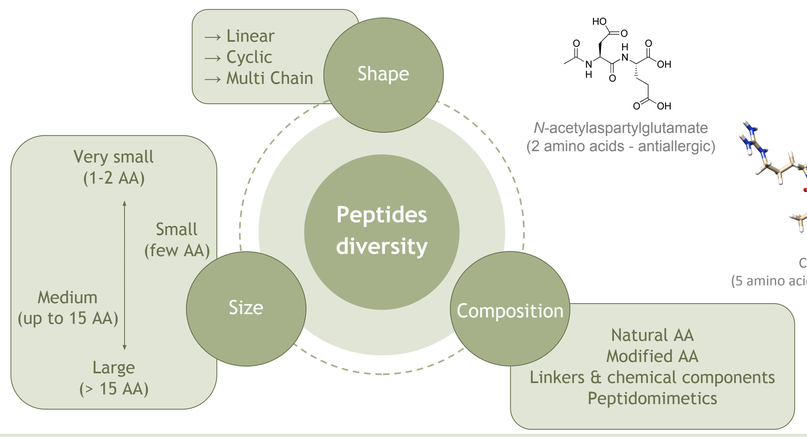
Rezai *et al. JACS* 128.8 (2006): 2510-2511.



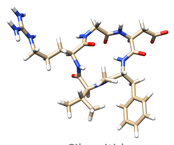
- ❖ Peptides are promising therapeutics drugs for undruggable target like protein-protein interaction.
- ❖ The majority of peptides are impermeable to cell membrane, preventing application against intracellular targets.
- ❖ Cyclization increases membrane permeability by eliminating charged termini and helping internal hydrogen bonds formation (closed conformation).

Peptides diversity

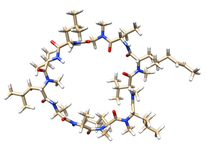
How to create a versatile technology able to handle this diversity ?



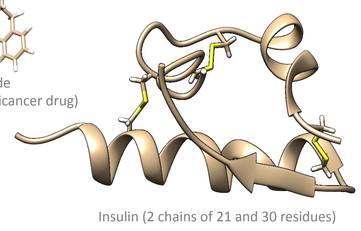
N-acetylaspartylglutamate
(2 amino acids - antiallergic)



Cilengitide
(5 amino acids - anticancer drug)

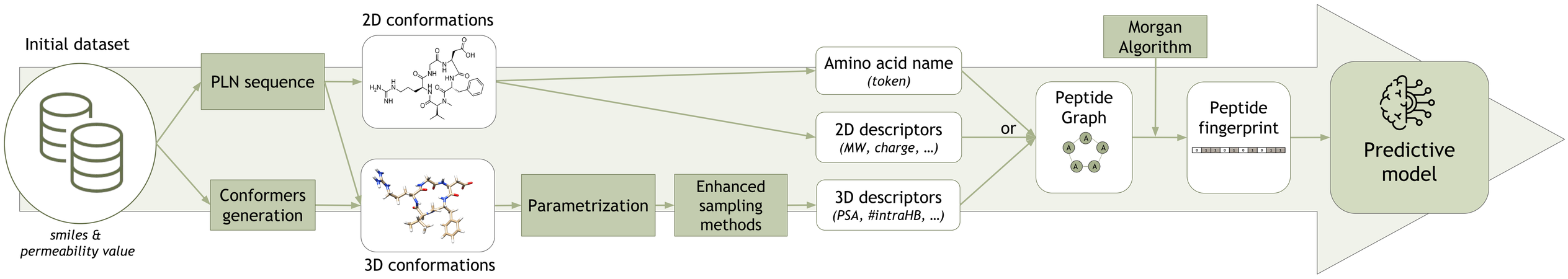


Cyclosporin
(11 amino acids - calcineurin inhibitor drug)



Insulin (2 chains of 21 and 30 residues)

Pipeline



Peptide representation

Iktos technology is versatile and handles the diversity of peptide drugs with different length, shape and include natural and modified residues.

SMILES

CC[C@H](C)[C@@H](C(=O)NCC(=O)N[C@@H](CCCCN)C(=O)N[C@@H](CC1=CC=CC=C1)C(=O)N[C@@H](CC(C)C)C(=O)N[C@@H](CC2=CNC=N2)C(=O)N[C@@H](C)C(=O)N[C@@H](C)C(=O)N[C@@H](CCCCN)C(=O)N[C@@H](CCCCN)C(=O)N[C@@H](CC3=CC=CC=C3)C(=O)N[C@@H](C)C(=O)N[C@@H](CCCCN)C(=O)N[C@@H](C)C(=O)N[C@@H](CC4=CC=CC=C4)C(=O)N[C@@H](C(C)C)C(=O)N[C@@H](C)C(=O)N[C@@H](CCC(=O)O)C(=O)N[C@@H](C@@H](C)CC)C(=O)N[C@@H](CCSC)C(=O)N[C@@H](C@@H](C)CC(=O)N)C(=O)N[C@@H](CO)C(=O)N)NC(=O)CN

- Atom level representation
- Long sequence
- Stereoisomerism adds complexity
- Conversion from SMILES to higher level representations is not well supported

Amino acid sequence

GIGKFLHAAKKFAKAFVAEIMNS

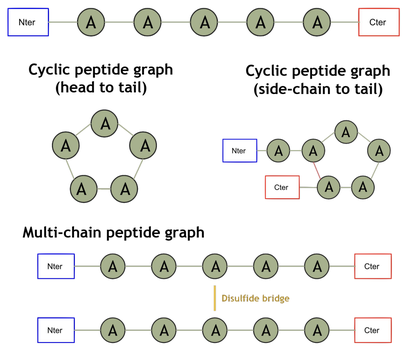
- AA level representation
- Sequences can only be used for natural AA
- Does not handle peptides with linkers and cyclic patterns

PLN and HELM

H-GIGKFLHAA[meV]KKFAKAFVAEIMNS-[NH2]

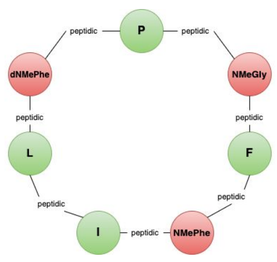
{G.I.G.K.F.L.H.A.A.[meV].K.K.F.A.K.A.F.V.A.E.I.M.N.S}

- AA level representation
- Handles modified AA & non-peptidic components
- Handles peptides with linkers & cyclic patterns
- Iktos developed proprietary tools to convert PLN sequences into graphs

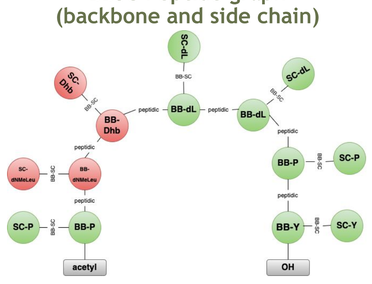


➢ Different kind of peptide graphs…

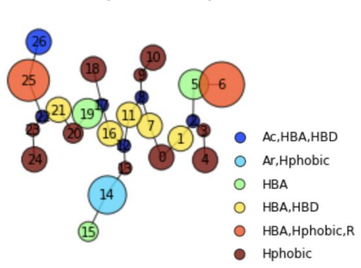
Simple Peptide Graph



BBSC Peptide graph (backbone and side chain)



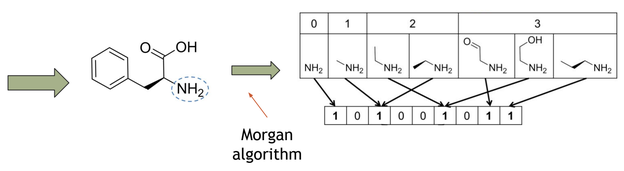
Pharmacophore Graph



➢ … leading to different peptide fingerprints

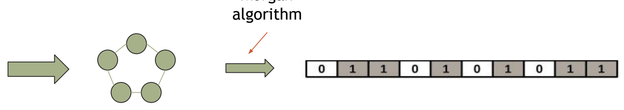
Morgan fingerprints

SMILES:
CC1=C(C(=CC=C1)C)NC(=O)CN2



Peptides fingerprints

PLN :
(cyclo)-FFPP[NMeAla]-(cyclo)



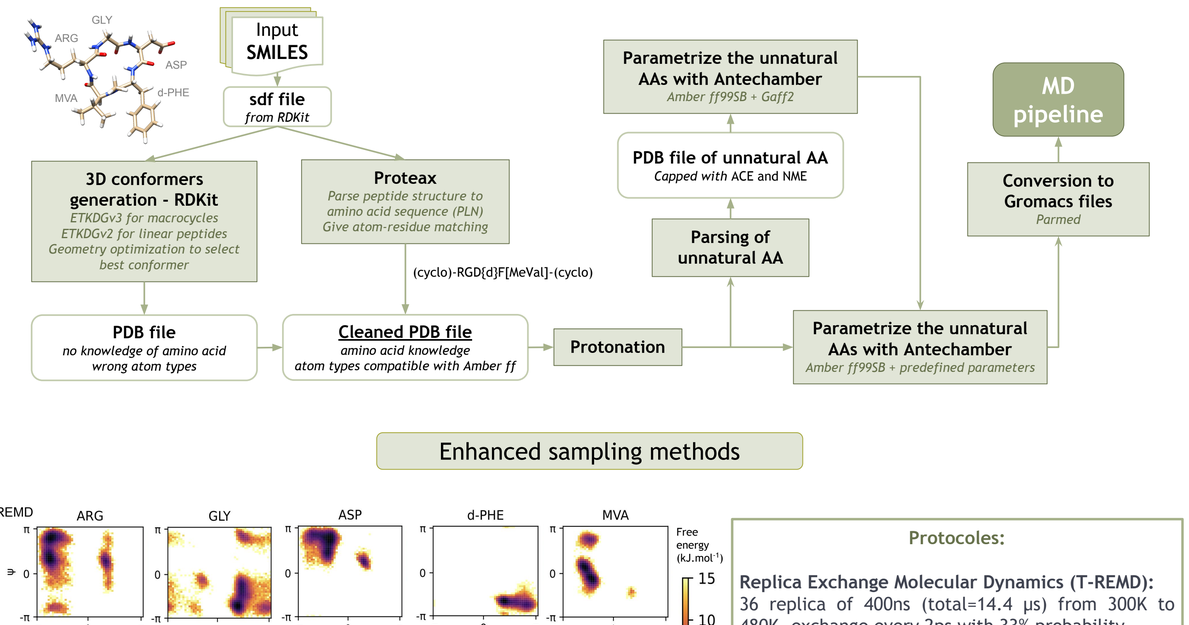
Definition of invariants:

1. **Amino acid tokens (names):** P , Y, NMeAla, Dhb…
2. **Amino acid descriptors**

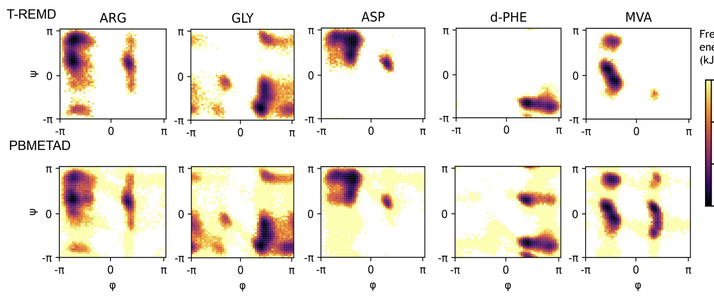
mw	rb	tpsa	logp	charge
1	1	2	3	1

Molecular Dynamics pipeline challenges

- ❖ The creation of an automated Molecular Dynamics (MD) pipeline handling the diversity of peptides present multiple challenges to address with among them:
  - ➢ the PDB reconstruction from a smiles and PLN sequence.
  - ➢ the parametrization of unnatural residues.
  - ➢ the sampling of the conformational landscape.



Enhanced sampling methods

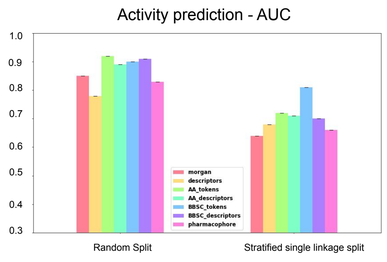


Protocoles:

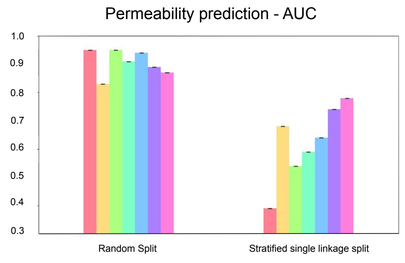
Replica Exchange Molecular Dynamics (T-REMD): 36 replica of 400ns (total=14.4 µs) from 300K to 480K, exchange every 2ps with 33% probability

Parallel bias metadynamics (PBMETAD): 25 CVs (all dihedral angles), 500 ns at 300K. Reweighted at the end of the simulation. Gaussian height=1.25, σ=0.2rad, bf=24, pace 0.5ps

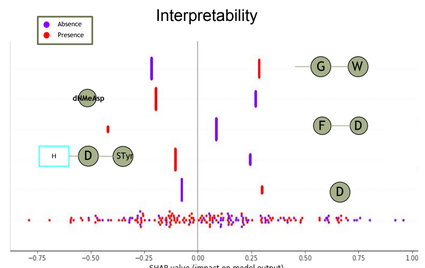
Predictive models



189 small linear peptides with activity values
Fixed length & modified AA



5252 cyclic peptides with PAMPA values
Length from 6 to 11AA & modified AA

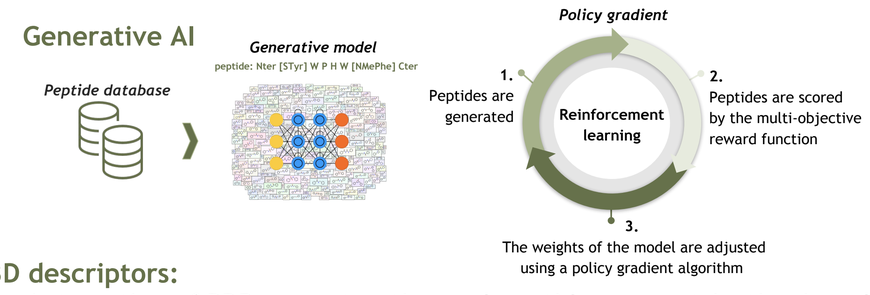


- ❖ On activity prediction, peptide fingerprints outperform baseline and generalize better.
- ❖ On permeability prediction, performance on random split gives the peptide fingerprints equivalent to the Morgan fingerprint. However, Morgan fingerprint performs very poorly on stratified single linkage split while the pharmacophore graph shows the best performance.
- ❖ Performance of the models really depends on the dataset and the split used, showing how challenging the prediction on peptides is.
- ❖ Only 2D descriptors have been used as 3D pipeline is currently under construction.
- ❖ Peptide fingerprint interpretability is more representative for peptides than Morgan fingerprint to understand important amino-acid sequence patterns for target activity.

On-going work & next steps

2D descriptors & peptides fingerprints:

- ❖ Our predictive models have been used in generative pipeline to help peptide design.



Objectives:

Predictive models (activity, permeability)

Metrics (similarity or quality scores)

Checkers (phys-chem properties, IP space)

3D descriptors:

- ❖ Automation of PDB reconstruction and peptide parametrization is well progressed.
- ❖ PBMETAD is currently being optimized and its ability to generalize to different systems is under investigation.
- ❖ Once the PBMETAD pipeline will be automatized, we would be able to apply it on the permeability dataset to compute 3D descriptors from the trajectories.
- ❖ Global and local 3D descriptors will be explored and evaluate for permeability prediction.

Contact

Maud Jusot: maud.jusot@iktos.com