

Concevoir un médicament en quelques heures : la révolution computationnelle

Simulation moléculaire, docking et machine learning pour la découverte de médicaments

Pinel Baudelaire T. DAHOUI

Data Scientist & IA | Ingénieur en Travaux Statistiques
Master Statistiques Appliquées aux Sciences du Vivant
tpineldahoui@gmail.com | github.com/Baudelaire12

CIPMA-CHAIRE UNESCO | Bénin

Vendredi 7 Mars 2025

Série LinkedIn STEM au Quotidien — Jour 5/40

Plan de la présentation

- 1 Mise en contexte
- 2 Concepts clés
- 3 Développement mathématique
- 4 Cas pratique
- 5 Conclusion
- 6 Références

Développer un médicament : une course contre la montre et le coût

Chiffre clé

Développer un médicament coûte en moyenne **2,6 milliards de dollars** et prend **10 à 15 ans**, avec un taux d'échec de plus de **90%** en essais cliniques (DiMasi et al., 2016).

- L'espace chimique des molécules potentiellement médicamenteuses contient entre 10^{23} et 10^{60} composés — impossible à explorer expérimentalement.
- **AlphaFold2** (DeepMind, 2021) a prédit la structure 3D de **200 millions de protéines** en quelques mois, une tâche qui aurait pris des siècles de cristallographie.
- En 2020, un candidat médicament contre le COVID-19 a été identifié en **4 jours** grâce à la simulation (Stokes et al., 2020).

Question centrale

Comment les algorithmes et la simulation permettent-ils de réduire de 10 ans à quelques mois la phase de découverte d'un médicament ?

Du gène à la molécule : pipeline de découverte computationnelle

Étapes du drug discovery computationnel :

- ➊ **Identification de la cible** : quel récepteur ou enzyme cibler dans la maladie ?
- ➋ **Criblage virtuel** (virtual screening) : sélectionner les candidats parmi des millions de molécules
- ➌ **Docking moléculaire** : simuler l'interaction 3D molécule-protéine
- ➍ **Optimisation par ML** : prédire et maximiser l'affinité de liaison
- ➎ **Simulation de dynamique moléculaire** : valider la stabilité de la liaison dans le temps
- ➏ **Prédiction ADMET** : Absorption, Distribution, Métabolisme, Élimination, Toxicité

Outil de référence

AutoDock Vina (Trott & Olson, 2010) : logiciel open-source de docking moléculaire, utilisé dans des milliers d'études publiées.

Énergie de liaison : la fonction de score du docking

Énergie libre de liaison estimée par :

$$\Delta G_{\text{bind}} = \Delta G_{\text{VdW}} + \Delta G_{\text{élec}} + \Delta G_{\text{H-bond}} + \Delta G_{\text{désolv}} + \Delta G_{\text{rot}}$$

Potentiel de Lennard-Jones (interactions de Van der Waals) :

$$V_{\text{LJ}}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$$

où r est la distance entre deux atomes, ε l'énergie au minimum et σ le diamètre de collision.

Constante d'inhibition (liée à ΔG) :

$$K_i = e^{\Delta G_{\text{bind}}/RT}$$

Plus K_i est faible, plus la molécule est un inhibiteur puissant.

Graph Neural Networks pour la prédiction de propriétés moléculaires

Représentation d'une molécule comme graphe :

- **Nœuds** v_i : atomes (caractérisés par leur numéro atomique, charge, hybridation...)
- **Arêtes** e_{ij} : liaisons chimiques (simple, double, aromatique...)

Message passing dans un GNN :

$$h_v^{(k)} = \text{UPDATE} \left(h_v^{(k-1)}, \text{AGG} \left(\left\{ h_u^{(k-1)} : u \in \mathcal{N}(v) \right\} \right) \right)$$

Application : DeepChem

La librairie **DeepChem** (Bharath Ramsundar et al., 2019) implémente des GNN pour prédire la toxicité, la solubilité et l'activité biologique de molécules directement depuis leur structure SMILES.

AlphaFold2 et la révolution de la structure des protéines

Percée historique

En 2021, **AlphaFold2** (Jumper et al., 2021 — DeepMind) a résolu le problème de repliement des protéines avec une précision atomique, remportant CASP14 avec un score GDT de **92.4/100**.

Impact quantifié :

Indicateur	Avant AlphaFold	Après AlphaFold2
Structures connues	~170 000 (PDB)	+ 200 millions (AF DB)
Coût cristallographie	50 000–300 000\$	~ 0\$ (prédit)
Délai de prédiction	1 mois à 10 ans	quelques minutes
Précision (GDT score)	< 60 (avant 2020)	92.4 (CASP14, 2021)

Application Afrique : Des chercheurs utilisent AlphaFold2 pour modéliser des protéines du *Plasmodium falciparum* (paludisme) en vue de découvrir de nouvelles cibles thérapeutiques (Bhatt et al., 2023 [Source à vérifier]).

Synthèse

La combinaison de la simulation moléculaire, des GNN et de modèles comme AlphaFold2 transforme la découverte de médicaments. Ce qui prenait 15 ans peut maintenant être réduit à 2 à 4 ans pour les premières phases de développement, avec un coût drastiquement réduit.

Perspectives :

- Conception *de novo* de médicaments par IA générative (Generative Drug Design)
- Application à des maladies négligées affectant l'Afrique : paludisme, tuberculose, trypanosomiase
- Prédiction de la résistance aux antimicrobiens par apprentissage automatique
- Intégration des données omiques dans les pipelines computationnels africains

« Nous pouvons maintenant concevoir des molécules comme des ingénieurs conçoivent des ponts — avec des calculs précis avant de construire. »

— Adapté de David Baker, Prix Nobel de Chimie 2024

Références bibliographiques

-  Jumper, J., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583–589.
-  DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry : New estimates of R&D costs. *Journal of Health Economics*, 47, 20–33.
-  Stokes, J. M., et al. (2020). A deep learning approach to antibiotic discovery. *Cell*, 180(4), 688–702.
-  Trott, O., & Olson, A. J. (2010). AutoDock Vina : Improving the speed and accuracy of docking. *Journal of Computational Chemistry*, 31(2), 455–461.
-  Ramsundar, B., Eastman, P., Walters, P., & Pande, V. (2019). *Deep Learning for the Life Sciences*. O'Reilly Media.
-  Gilmer, J., Schütt, A. T., Matyasko, J., Hamprecht, F., & Müller, K.-R. (2017). Neural message passing for quantum chemistry. *ICML 2017*, 70, 1263–1272.

Merci pour votre attention

Pinel Baudelaire T. DAHOUI

Data Scientist & IA | CIPMA-CHAIRE UNESCO, Bénin

`tpineldahoui@gmail.com`

`github.com/Baudelaire12`

LinkedIn : @Pinel Baudelaire DAHOUI

Série STEM au Quotidien — 40 jours pour comprendre le monde par les données