

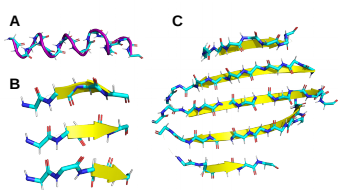
# Sammy Khalife

Séquences en Bioinformatique, Informatique et Mathématiques (Seqbim)  
22 Novembre 2020

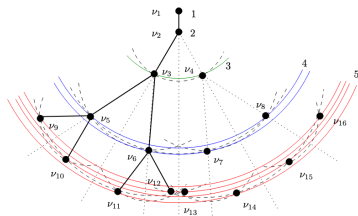
- ▶ Acutelement post-doctorant (2020/2021) au LIX, Ecole Polytechnique
- ▶ Thèse : (Défendue le 31 août 2020) Graphes, géométrie et représentations pour le langage et les réseaux d'entités
  - Questions combinatoires et problème inverse. Séquence  $\mapsto$  graphe? (en partie présentation de demain)
  - Propriétés géométrique du langage (lien entre sémantique et géométrie?)
  - Géométrie des distances : algorithmes efficaces pour les représentations en apprentissage
- ▶ Nouveau dans la communauté Seqbim

## ■ Sujet de recherche actuel (post-doctorat) : Approches géométriques multi-résolution et multi-échelle pour la détermination de la structure de bio-molécules

- Nouveaux descripteurs géométriques à l'aide de méthodes d'apprentissage avec information limitée (e.g. seulement à partir de certaines distances inter-atomiques, sans connaître la séquence)
- Modèles génératifs entre grandeurs géométriques des protéines : dépendance entre les angles diédraux et grandeurs de liaisons



(a) Détermination de (A)  $\alpha$  hélice (B) Feuillet  $\beta$  parallèle (C) Feuillet  $\beta$  anti-parallèle à l'aide de modèle d'apprentissage



(b) Utilisation de l'information géométrique dans des approches de type "branch-and-bound" pour la conformation

■ Autres travaux de recherche :

- Etude complète de la complexité de problème inverse de graphes de mots (présentation de demain)
- Fondement théorique des relations entre quantités géométriques de représentations (embeddings) et information mutuelle jointe (PMI) pour le traitement du langage

■ Liste de (pré-)publications :

<http://www.lix.polytechnique.fr/Labo/Sammy.Khalife/papers.html>

■ Email : [khalife@lix.polytechnique.fr](mailto:khalife@lix.polytechnique.fr)