

Генеративные модели для структуры белков и их комплексов с другими молекулами.

Выполнил: Клыпа Р.С.

Научный руководитель: Грудинин С.В.

МФТИ

18 мая 2024 г.

Цели исследования

Цель работы

Генерация структуры молекулярных соединений (например, РНК) при взаимодействии с протеином - $P(\text{geom}_{RNA} | \text{geom}_{pr}, \text{seq}_{pr}, \text{seq}_{RNA})$.

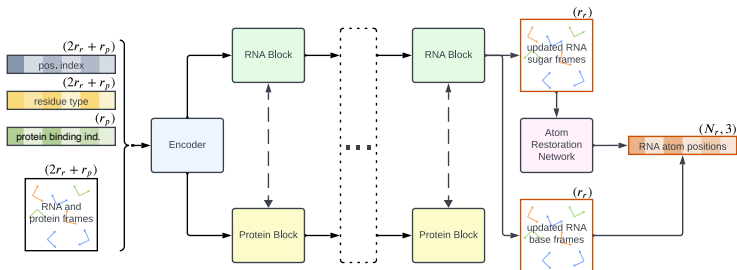
Проблема

Существующие решения обладают рядом технических и теоретических недостатков.

Метод решения

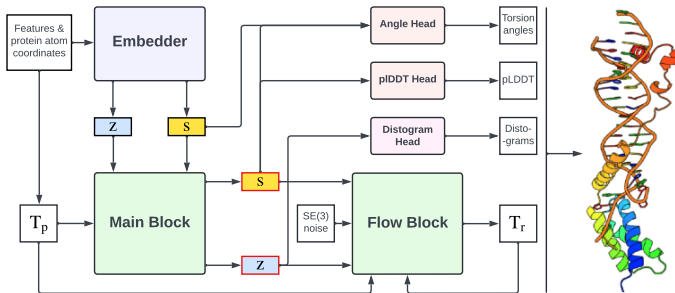
Дизайн архитектуры модели и генеративного процесса с учетом биологических и физико-химических особенностей задачи.

MolBindDif-ba и MolBindDif-bb



- ▶ Генеративный процесс - диффузия, Riemannian Score Matching.
- ▶ Существенно превосходит текущий SOTA (RoseTTaFold2NA) на всех метриках.
- ▶ Имеет ряд теоретических недостатков.

MolBindFlow



- ▶ Генеративный процесс - flow matching на SE(3).
- ▶ Имеет ряд технических преимуществ перед прошлыми моделями.
- ▶ Теоретический аппарат корректен и прост в использовании.

- ▶ Для решения задачи были разработаны модели, превосходящие по результатам прошлые работы.
- ▶ Было выполнено исследование теории генеративных процессов в контексте группы $SE(3)$ и их применения в молекулярно-структурной области.
- ▶ Планируется рассмотреть другие генеративные задачи молекулярной биологии.