

Лекция 5. Дизайн остова белка

Курс: Методы машинного обучения в дизайне белков

Головин А.В.¹

¹МГУ им М.В. Ломоносова, Факультет Биоинженерии и Биоинформатики

Москва, 2023

» Дизайн остова или формы белка

- * **Локальный дизайн:** Вставки и делеции для достижения необходимой формы локального окружения
- * **Глобальный дизайн формы:** Подгон последовательности под фолд и предсказание структуры из последовательности,



» Основные проблемы:

- * Монте-Карло: 100 а.к. 3N степеней свободы, получаем 10^{48} конформаций.
- * **Парадокс Левинталя:** "Промежуток времени, за который полипептид приходит к своему скрученному состоянию, на много порядков меньше, чем если бы полипептид просто перебирал все возможные конфигурации".
- * Для решения разумно использовать накопленные знания для моделирования.



» Построение остава

- * Генерируем координаты остава желаемого белка из каких-то элементов.
- * Не обязательно использовать координаты, могут подойти дистанционные ограничения.
- * На основании ограничений можно построить структуру для polyA



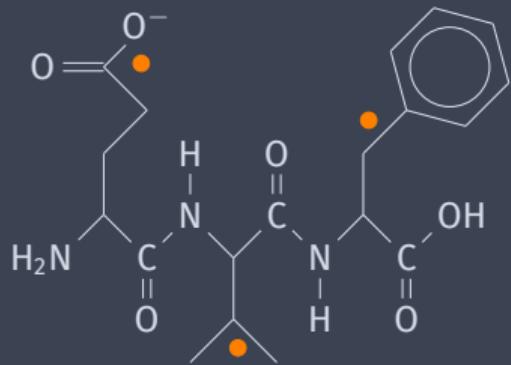
» Моделирование петель

- * Эмпирическое моделирование:
 - * Поиск подходящего фрагмента по PDB
 - * Использовать базы данных (LIP, etc..)
- * Молекулярная механика.
- * Монте-Карло.
- * Rosetta:
 - * Поиск фрагментов близких по последовательности.
 - * Комбинирование результатов поиска с помощью Монте-Карло.

Комбинации выше перечисленных.



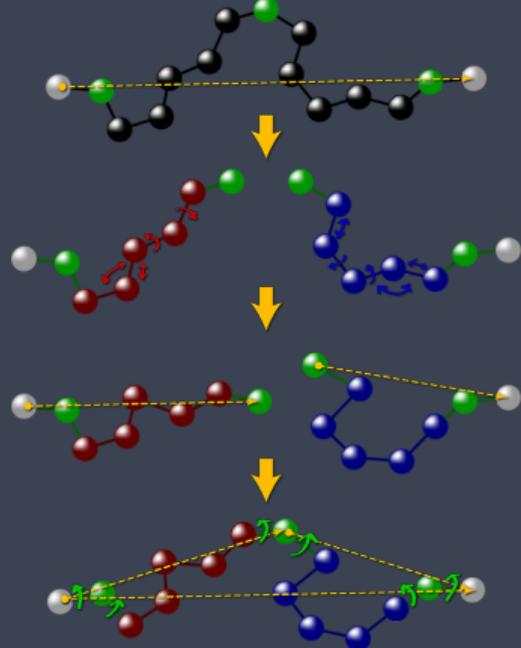
» Низкое разрешение в Rosetta



» Моделирование петель, CCD



» Моделирование петель, GeneralizedKIC



Делим петлю на две части

Меняем Φ, Ψ в фрагментах

Оптимизируем фрагменты для закрытия петли



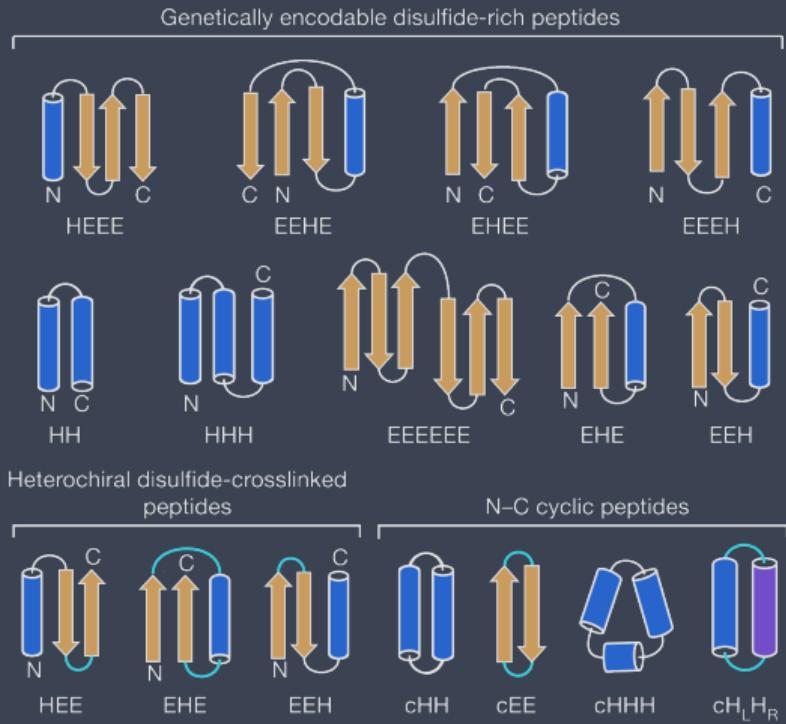
» Поиск последовательности

FastDesign

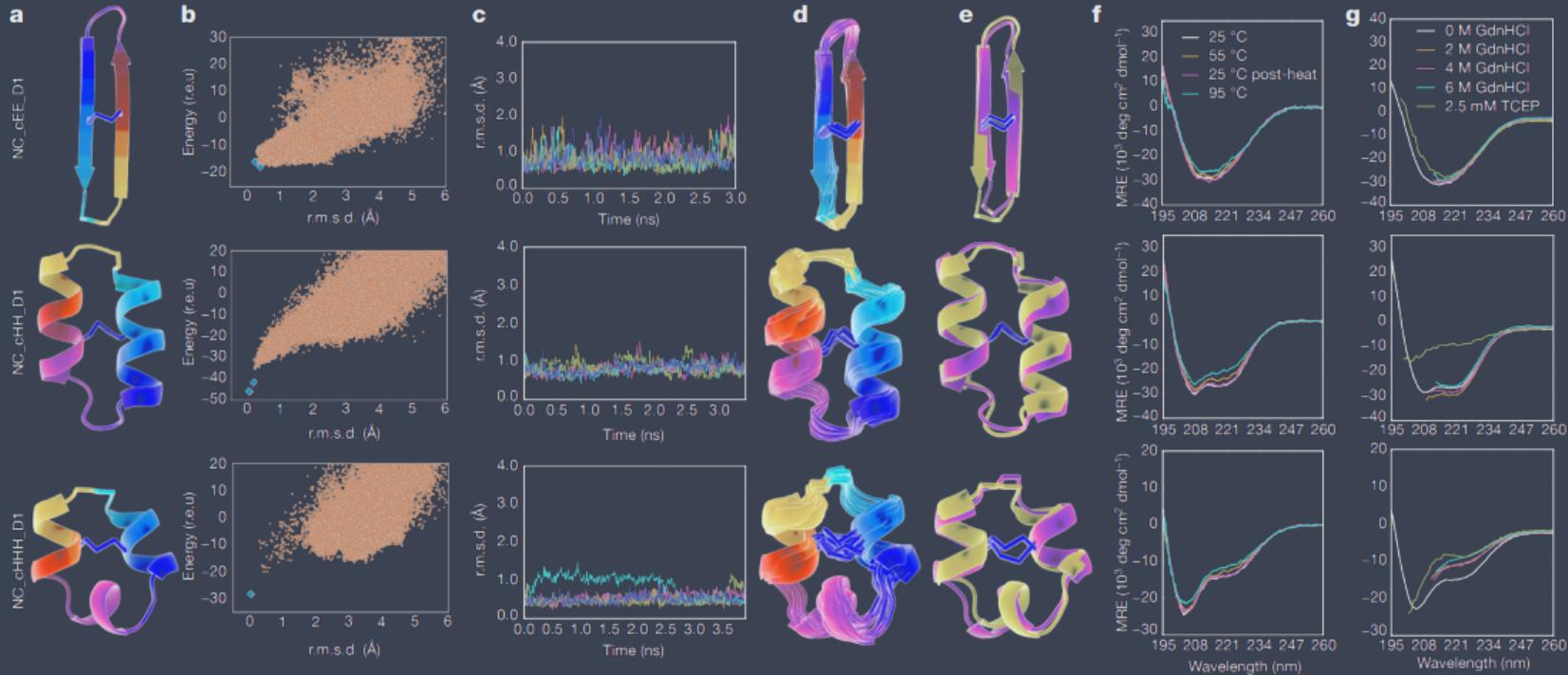
- * Перебор ротамеров в 4 раундах
- * Оптимизация геометрии
- * 12 раундов ротамеры/оптимизация
- * Разделение позиций : 'core', 'boundary', 'surface
- * 80000 вариантов для топологии и сортировка по энергии



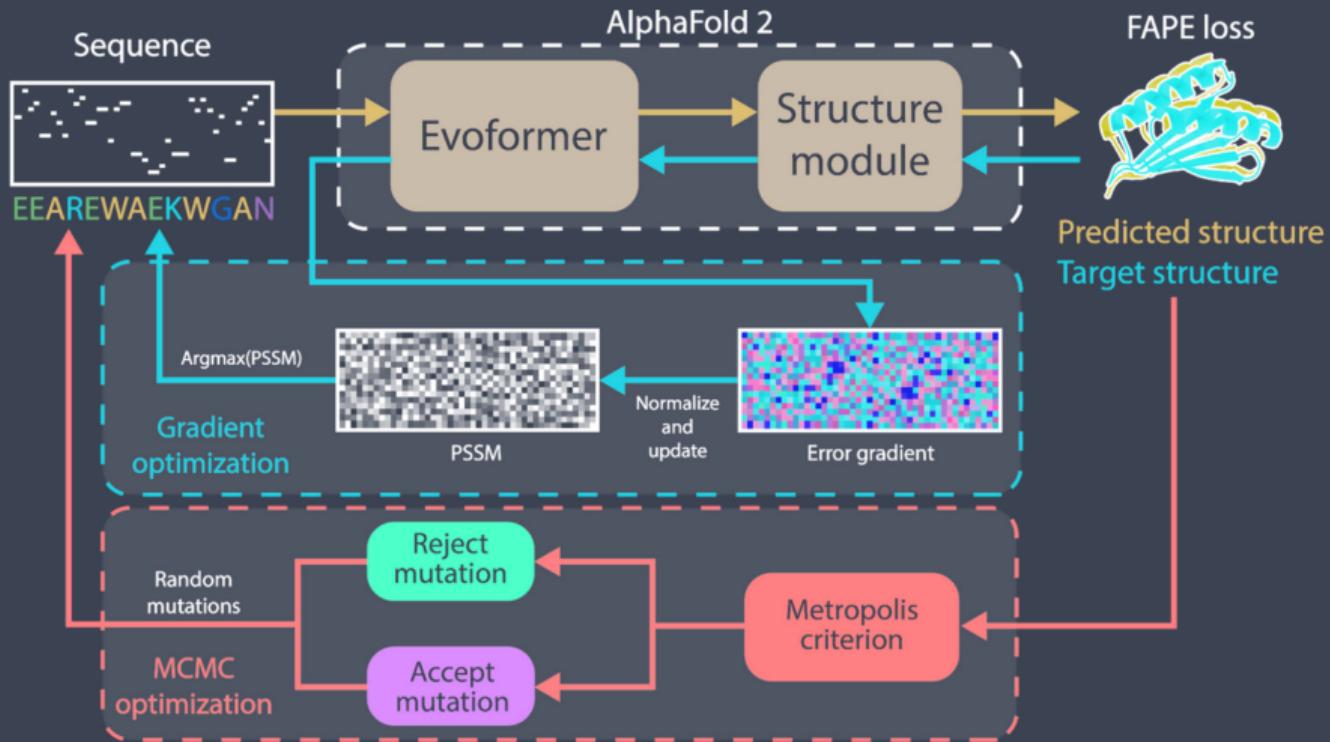
» Дизайн циклических пептидов



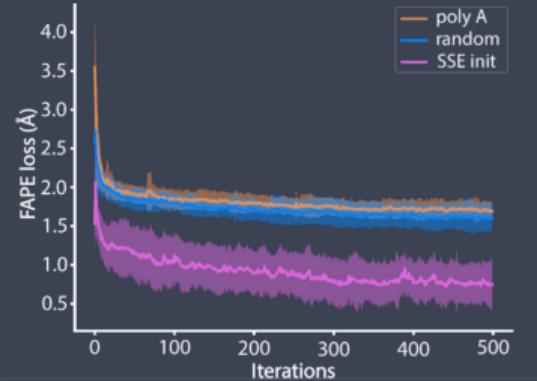
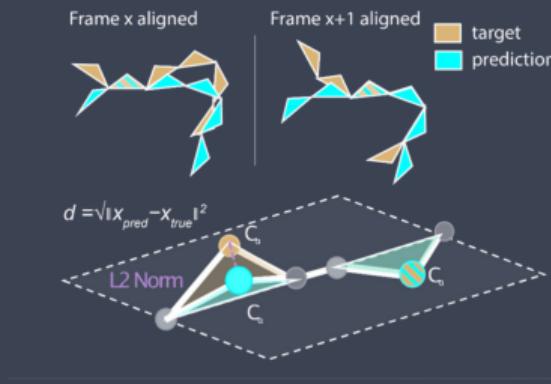
» дизайн циклических пептидов



» Дизайн по варианту укладки с AlphaFold



» Дизайн по варианту укладки с AlphaFold

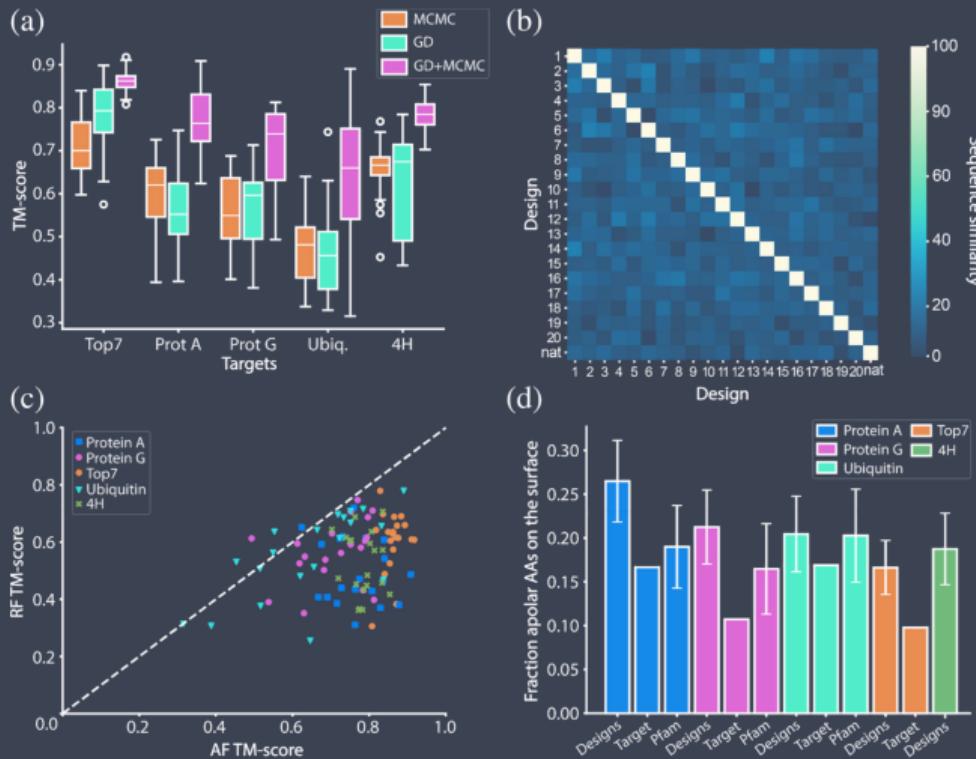


» Словами

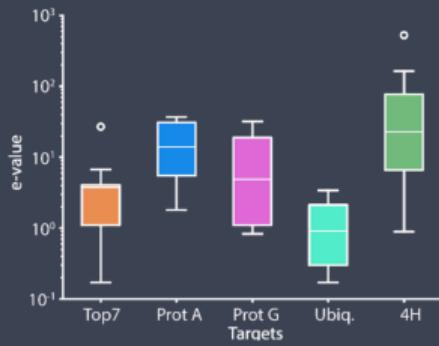
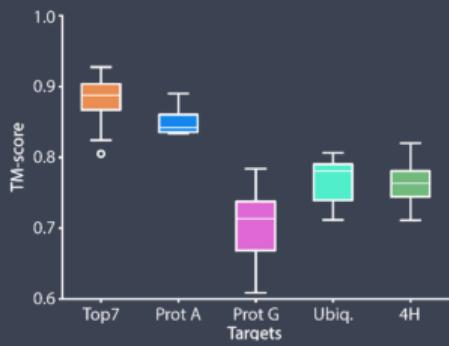
- * Два модуля оптимизации последовательности GD и МСМС (Марковские цепи в Монте-Карло)
- * В модуле GD дизайн начинается с последовательностей характерных для SS
- * "Обратный проброс ошибки" для структуры из дизайна от желаемой и оптимизация PSSM
- * После нескольких раундов последовательность "мутируется" МСМС



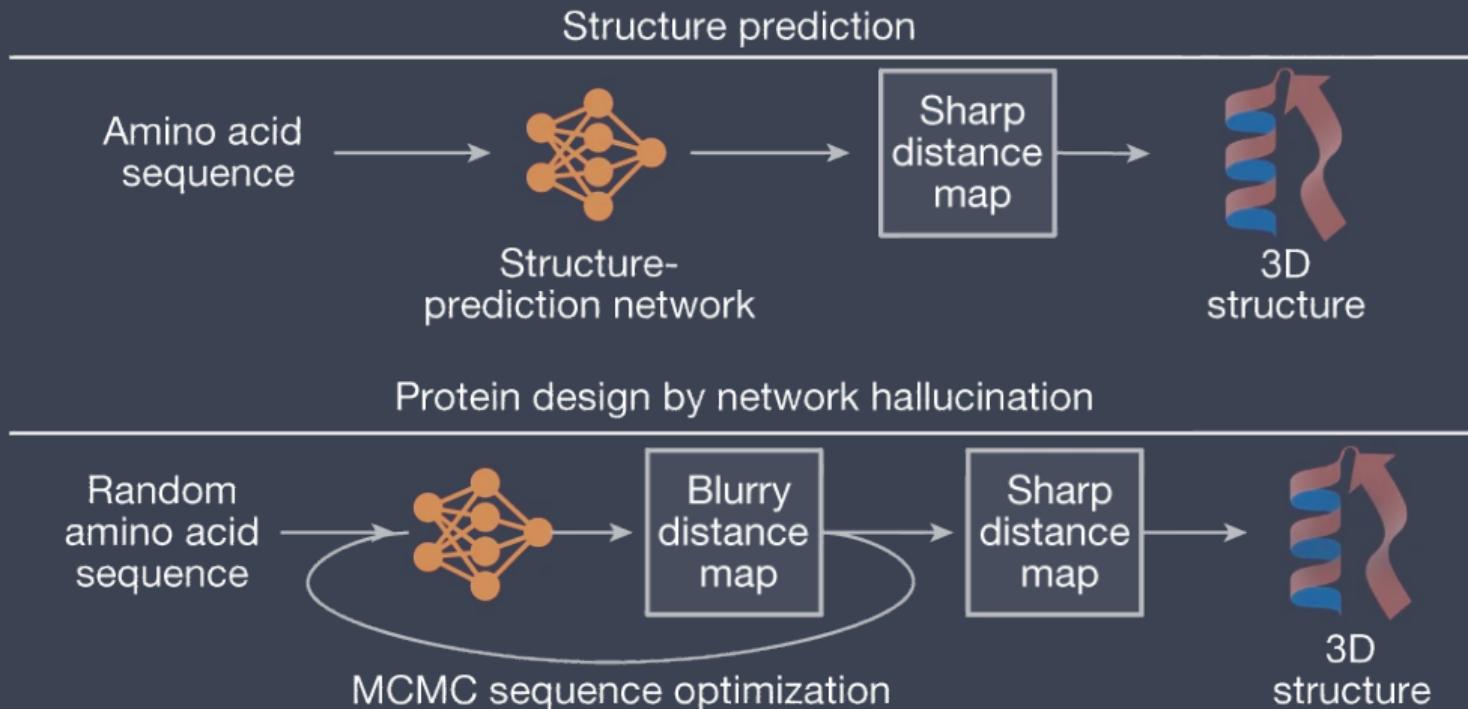
» Результат Alpha Design



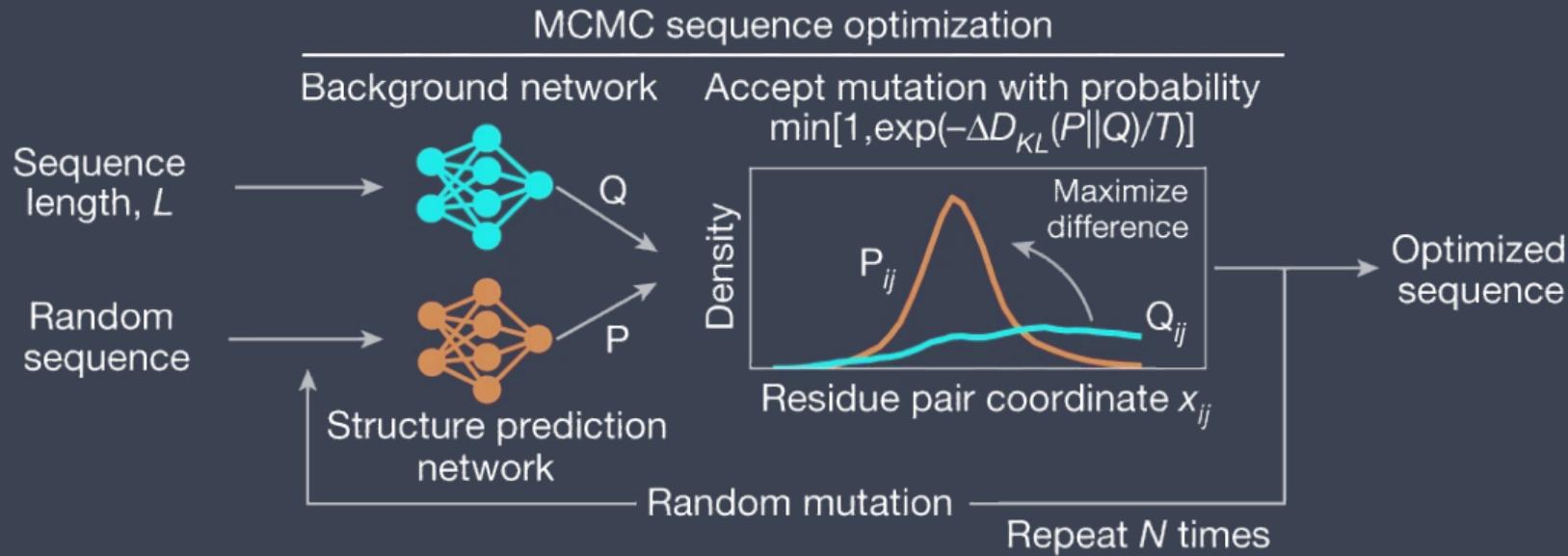
» Результат Alpha Design



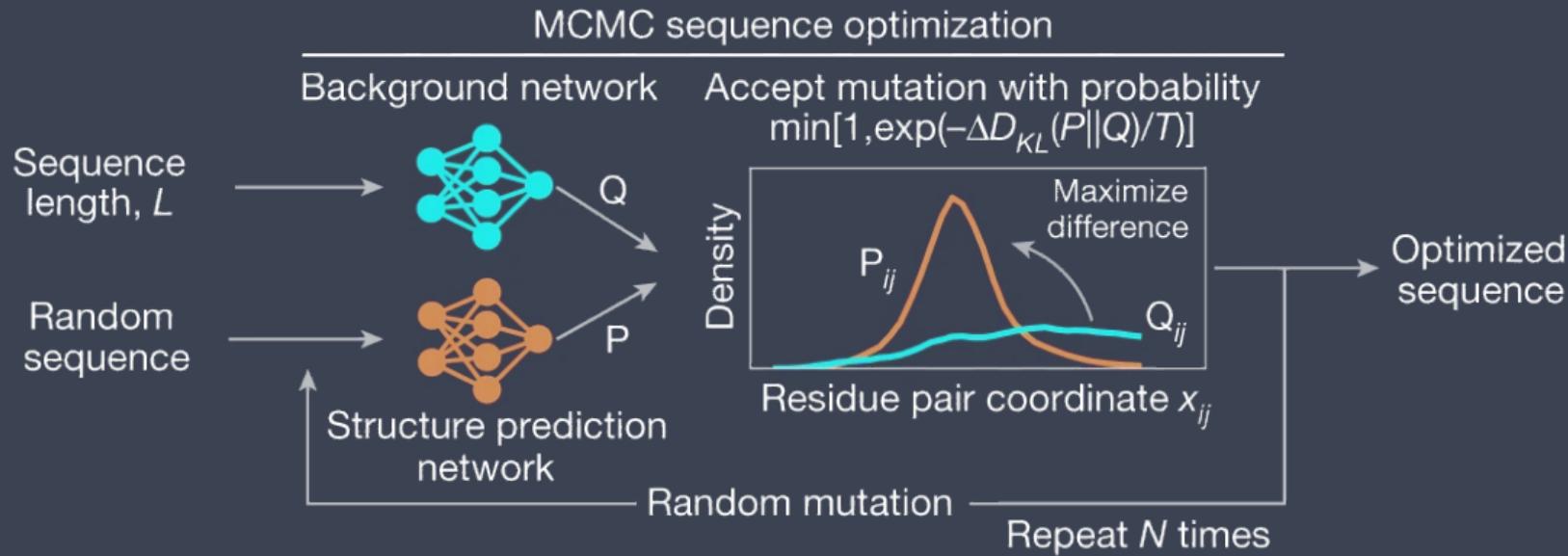
» "network hallucination"



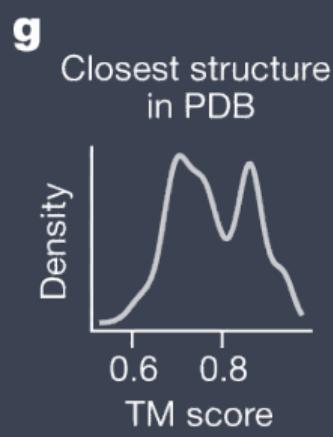
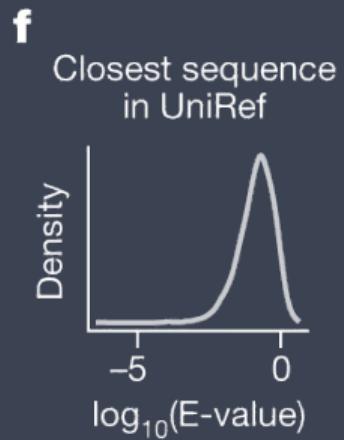
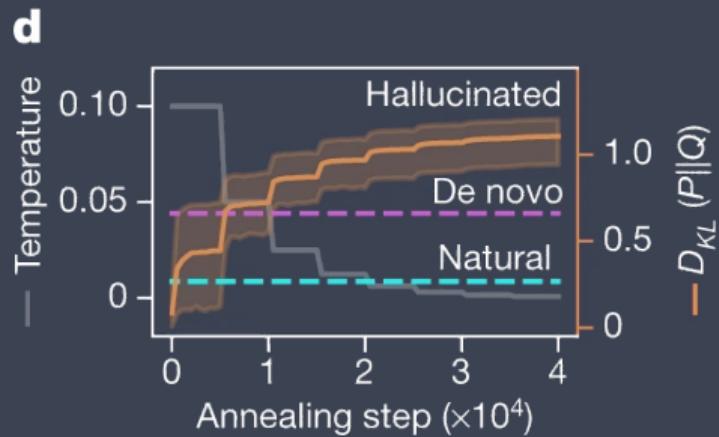
» "network hallucination"



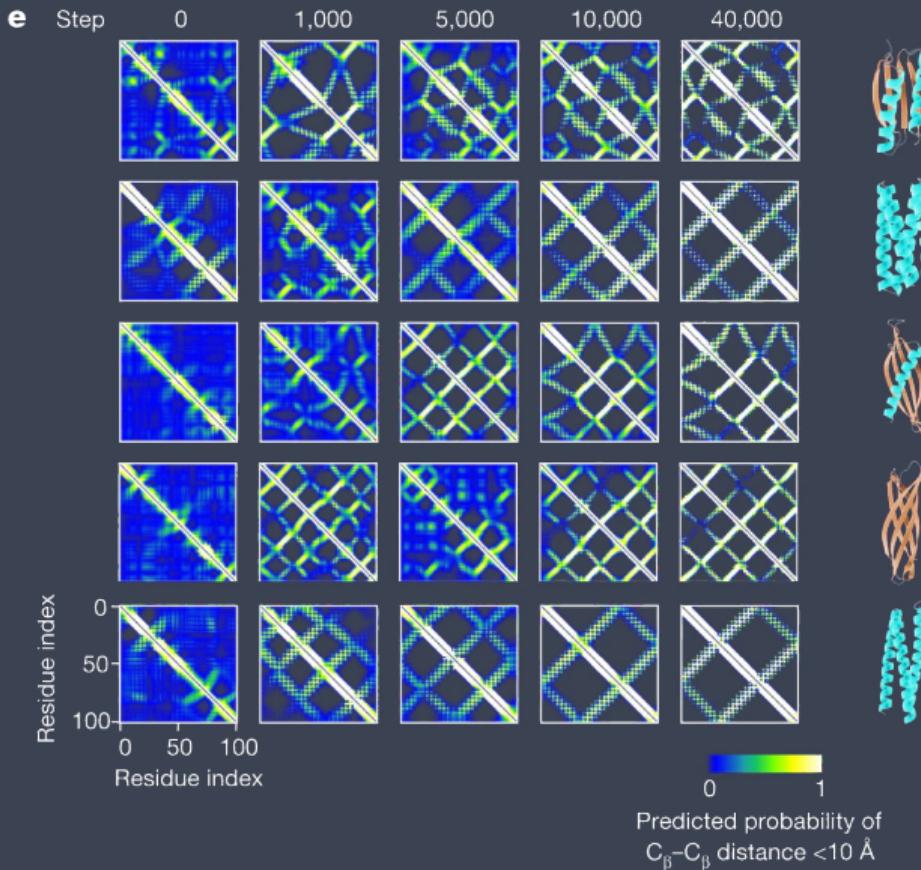
» "network hallucination"



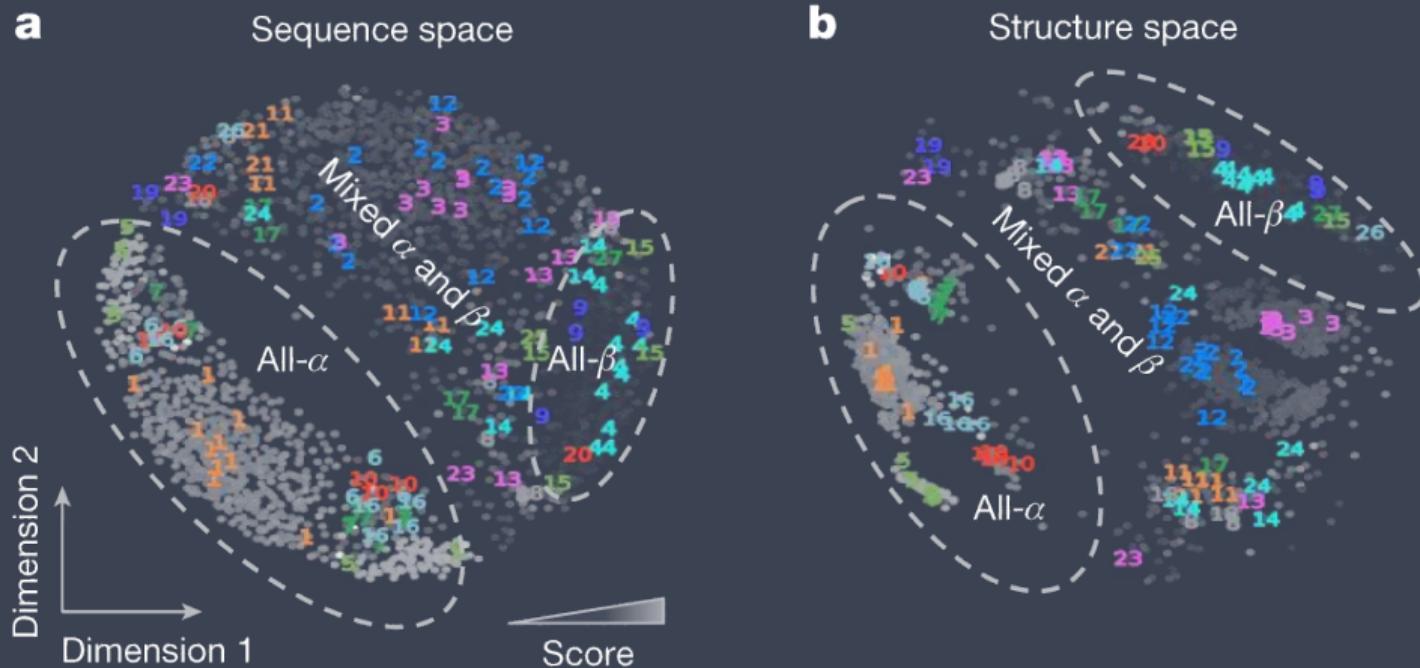
» Галлюцинации идеальных белков



» Галлюцинации идеальных белков



» Галлюцинации идеальных белков



129 белков проверено - 27 кластеров



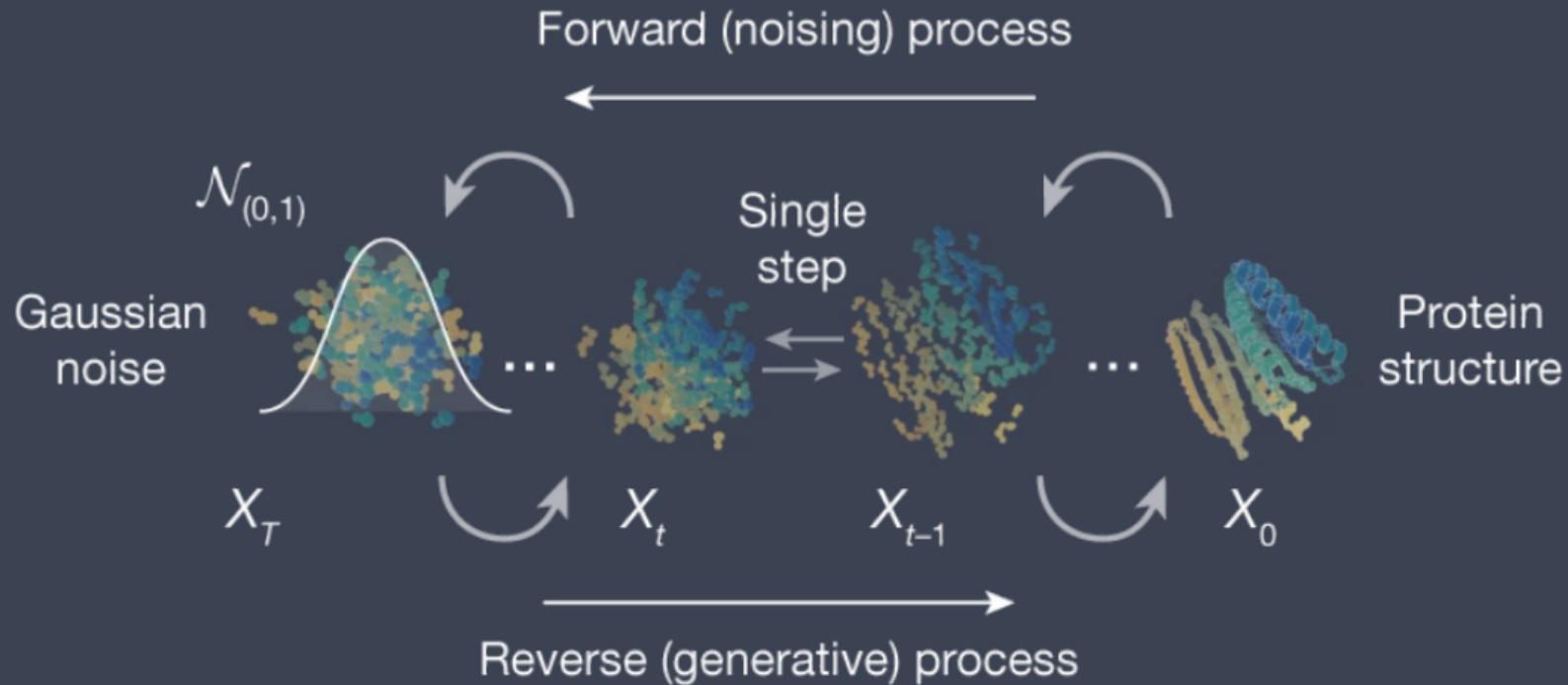
» Заключение

- * Из модели для нативных получили новые белки
- * Высокая точность совпадения предсказаний и эксперимента
- * Регулярность структур, короткие петли, они идеальны :)
- * Функция потерь может включать в себя желания пользователя (прошлая лекция)



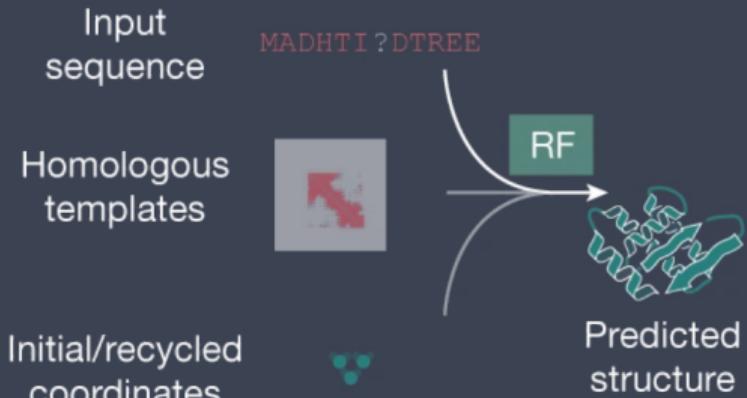
» Диффузия

Diffusion model



» Диффузия

RoseTTAFold

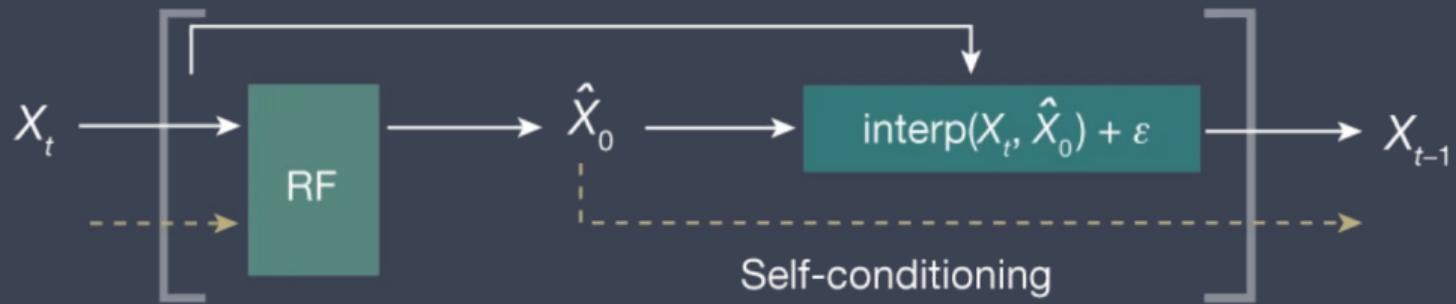


RFdiffusion



» Диффузия

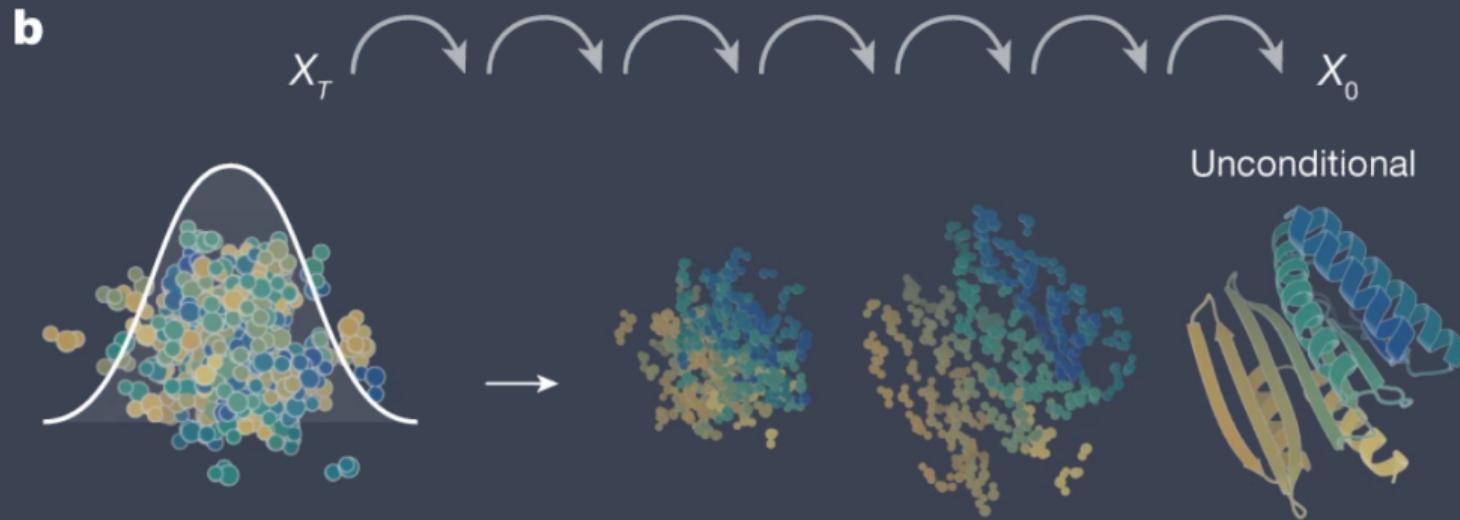
Single RFdiffusion step



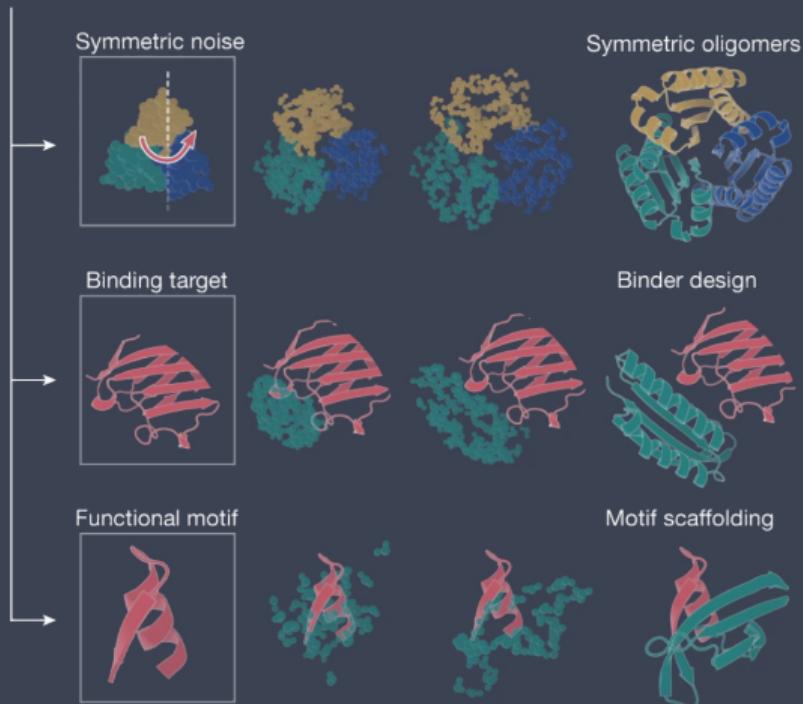
На каждом шаге t модель берет $X(t+1)$, а затем прогнозирует обновленную структуру $X(0)$. Следующий ввод координат $X(t-1)$ в модель генерируется путем интерполяции с шумом в направлении $X(0)$.



» Диффузия, без условий



» Диффузия, с условиями

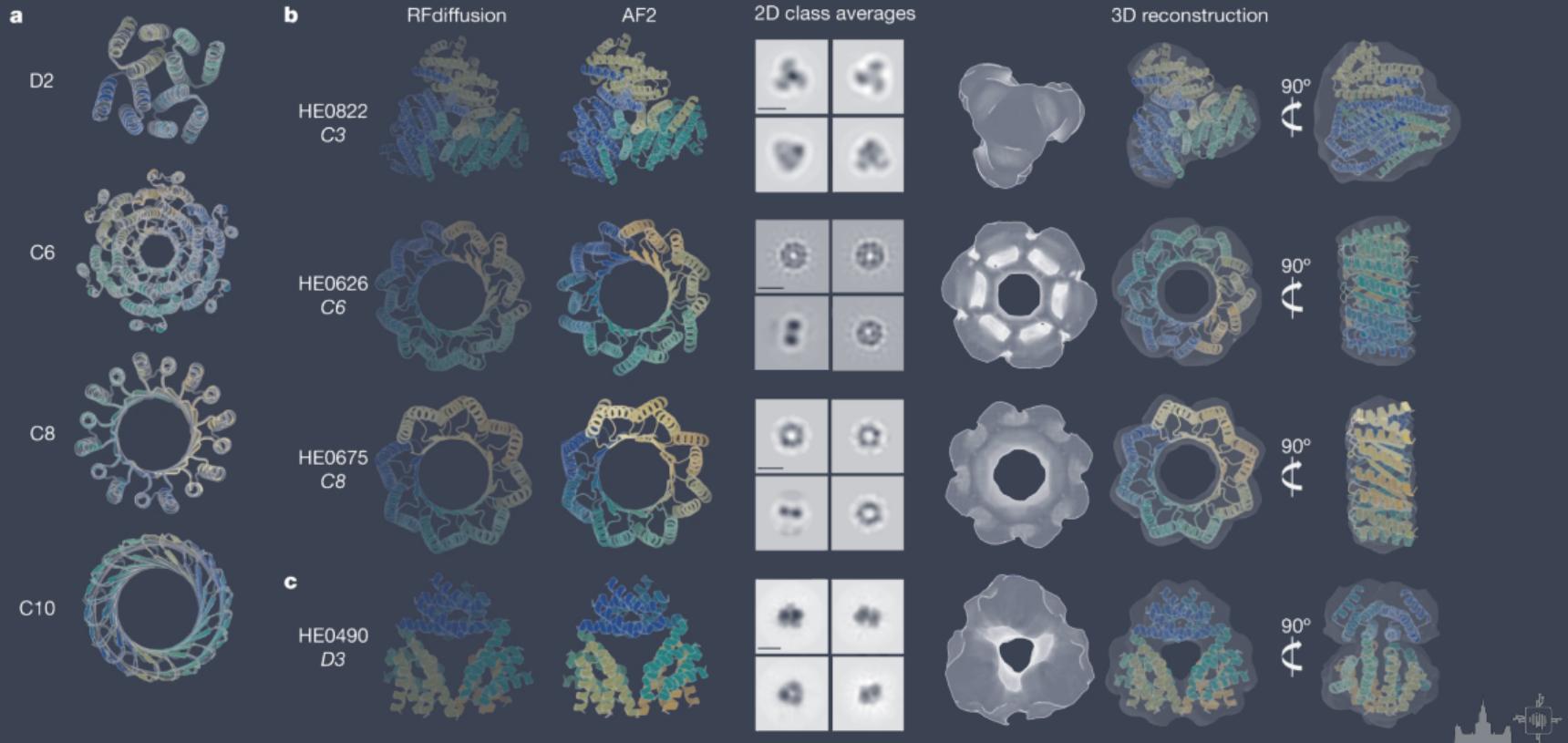


» При производительность

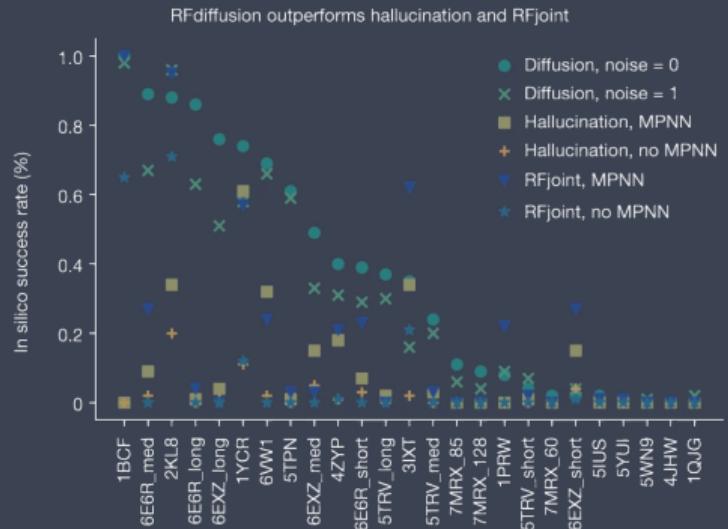
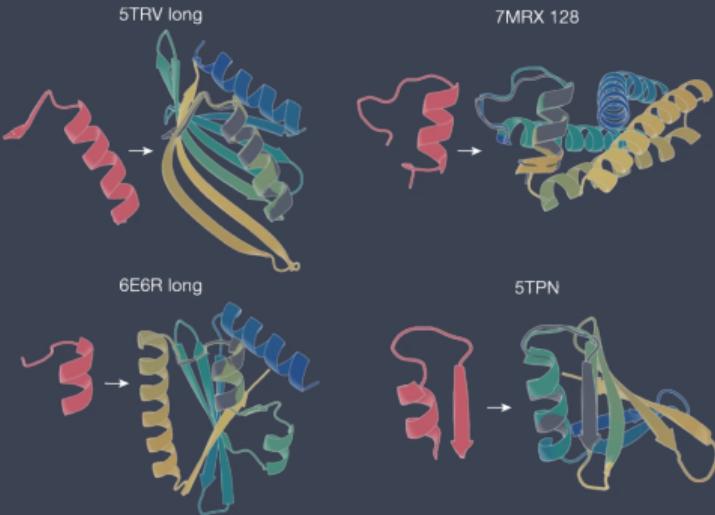
- * RF-диффузия значительно превосходит (AF2 score) галлюцинацию с RF
- * Генерация RF-диффузии также более эффективна в вычислениях, чем неограниченная галлюцинация с RF
- * Белок из 100 остатков может быть создан всего за 11 секунд на графическом процессоре NVIDIA RTX A4000, в отличие от RF Hallucination, который занимает около 8,5 минут.



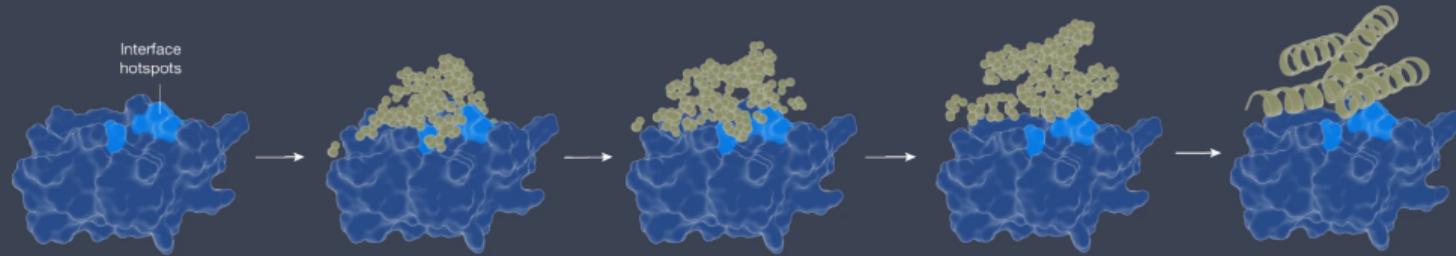
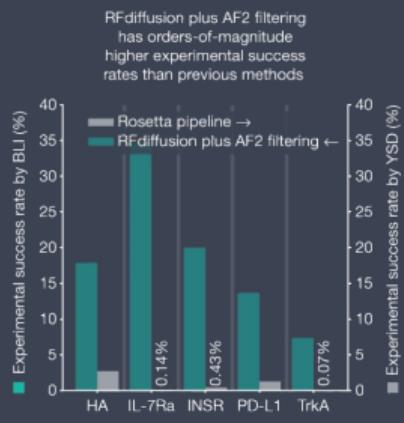
» Диффузия, с условиями



» Диффузия, с условиями

a**b**

» Дизайн "байндера"

a**b****c**