

Лекция 7. Дизайн ферментов

Курс: Методы машинного обучения в дизайне белков

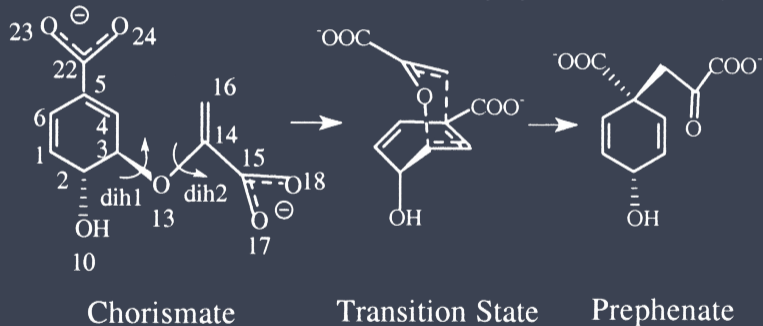
Головин А.В. ¹

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

Москва, 2023

» Ферменты это биокатализаторы

Катализатор – вещество, ускоряющее химическую реакцию. Биокатализатор – такое вещество биологической природы (белок, нуклеиновая кислота).



Фермент хоризмат мутаза ускоряет эту реакцию в 10^6 раз



» Биокатализаторы в промышленности и медицине

Пищевая индустрия:

- * Ферментация чего угодно
- * Разрушение глютена
- * Осветление напитков
- * Безлактозное молоко
- * Синтез ароматизаторов
- * Удаление неприятных ароматов

Не пищевая индустрия:

- * Производство биотоплива
- * Деградация пластиков
- * Очистка сточных вод
- * Улучшение тканей
- * Стиральные порошки
- * Зубные пасты
- * Производство бумаги
- * Производство чистых аминокислот

Медицина:

- * Синтез полусинтетических препаратов
- * Синтез прекурсоров
- * Терапевтические ферменты

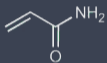


Factors that enable biocatalysis

- **Reaction design**
Manual, automatic
- **Choice of biocatalyst**
Genomic database, de novo design
- **Biocatalyst optimization**
Rational engineering, directed evolution
- **Bioprocess development**



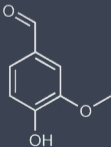
Materials and polymers



Acrylamide



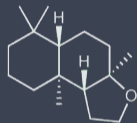
Fine chemicals



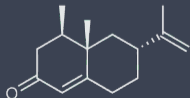
Vanillin



Fragrances and flavours



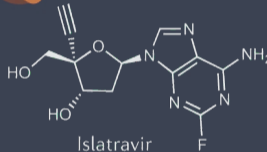
(-)-Ambrox
(fragrance intermediate)



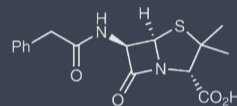
Nootkatone



Pharmaceuticals



Islatravir



Penicillin G



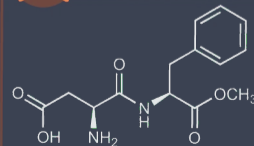
Biopharmaceuticals/
new modalities



Antibody–drug
conjugates



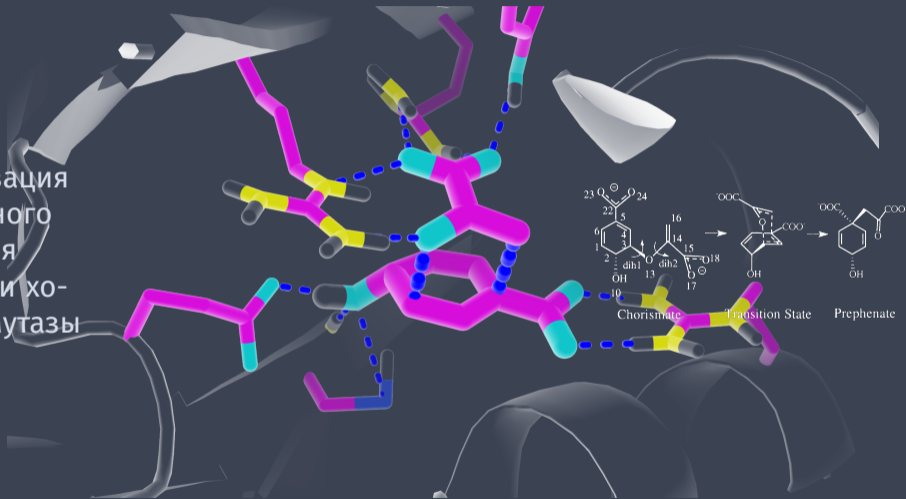
Food

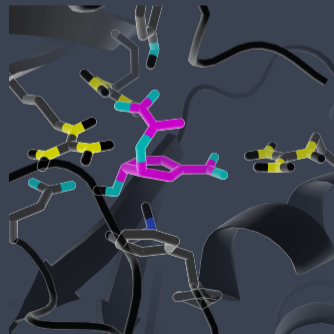
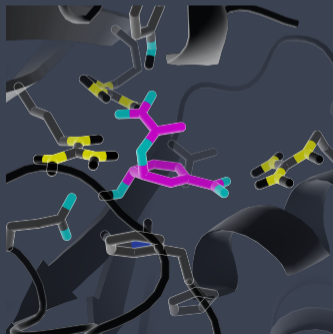
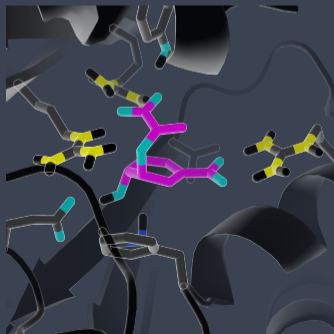
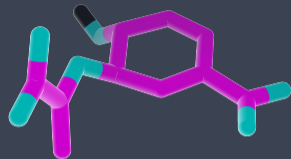
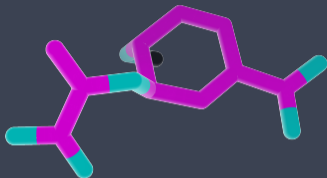
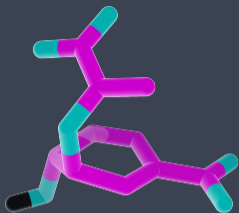


Aspartame

» Как они это делают

Стабилизация
переходного
состояния
остатками хо-
ризмат мутазы



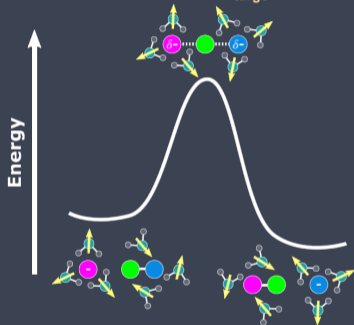


» Как они это делают

Reaction in Solvent

$$\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$$

large

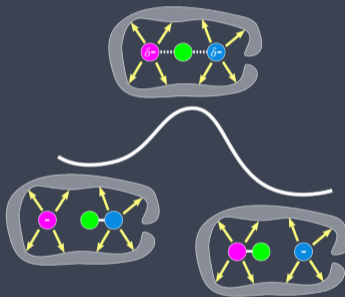


Reorganization of solvent in reactant, TS, and product

Reaction in Enzyme

$$\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$$

small small

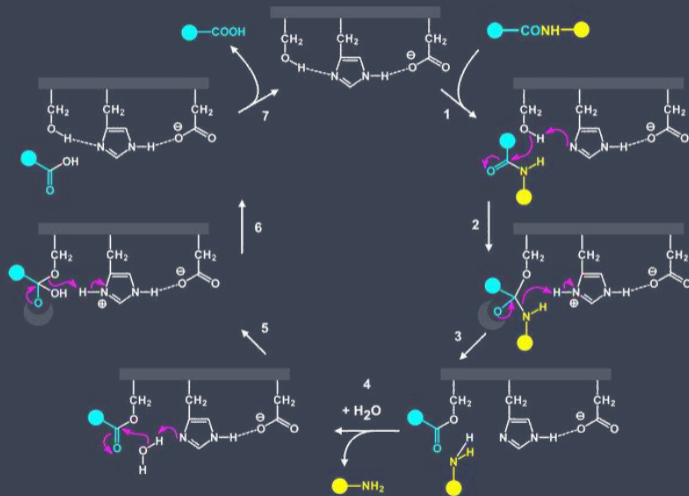


Protein does not reorganize along reaction pathway

Current Opinion in Structural Biology



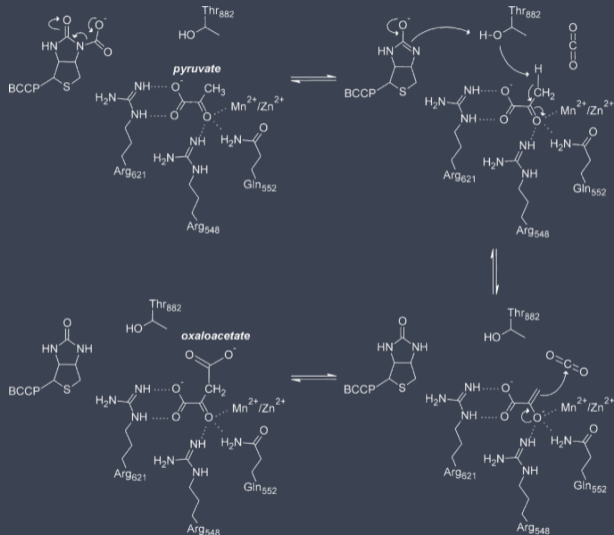
» Чаще всего все сложнее



Многостадийный
катализ сериновой
протеазы с
ковалентными
конъюгатами
субстрат-фермент



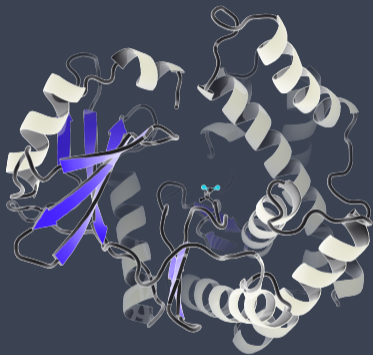
» Чаще всего все сложнее



Многостадийный катализ пируват карбоксилазы с кофактором-катионом металла и коферментом-биотином



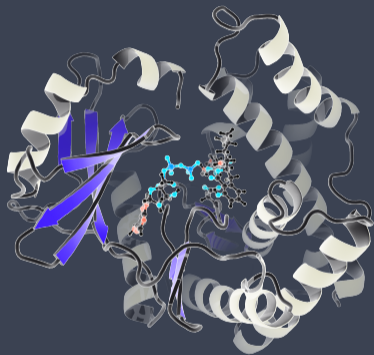
» Структурные особенности



Непосредственно в реакцию вовлечены только 1-3 остатка. Они называются активным центром фермента.
Конкретно в этом ферменте в реакции участвует только этот аспарат



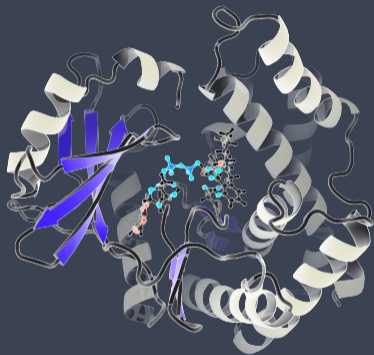
» Структурные особенности



Остатки активного центра
непосредственно участвуют в реакции
с реагентами



» Структурные особенности

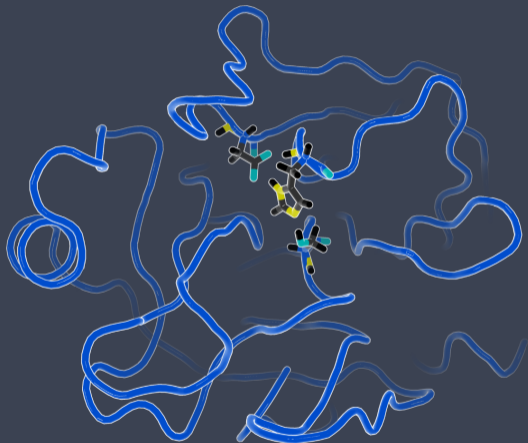


Остатки активного центра
непосредственно участвуют в реакции
с реагентами
Зачем нужен весь остальной белок?

Для начала реакции нам надо поймать
и зафиксировать реагенты в
правильном положении
Остатки, которые это делают - это
карман связывания



» Активный центр и карманы связывания протеазы



Без триады не будет реакции Какой именно пептид будет разрезан, определяют структуры карманов P1-P3, P1'-P3'



» Структурные особенности

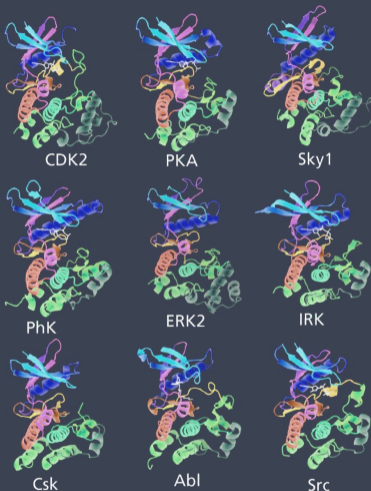


Figure 16.12a The Biology of Cancer (© Garland Science 2014)

Модульность: 9 эукариотических киназ (мутации всех замешаны в онкологии).

- * Активные сайты одинаковые
- * Карманы связывания АТФ одинаковые
- * Открытие/закрытие одинаковое
- * Карманы связывания субстратов разные



» Дизайн и его валидация

Дизайн – создание объекта с заданными функциями

- * Создание белка
- * Создание гена
- * Создание сиквенса

Инструменты дизайна – то, что делает создание сиквенса не случайным

Имеем сиквенс, выполняет ли соответствующий объект функцию?

- * Сразу в экспериментальный assay
- * Имитация эксперимента – молекулярное моделирование/ML

Инструмент валидации – то, что сопоставляет дизайну оценку его функции (с некоторой ошибкой)



» Дизайн и его валидация

Тривиальный вычислительный дизайн

Генерация случайных гипотез и их вычислительная валидация

- * Вычислительная валидация хоть и дешевле экспериментальной, но тоже трудозатратна (при отлаженном и предсказуемом пайплайне валидация процессивности одного варианта фермента – неделя на современной рабочей станции)
 - * Нужно генерировать гипотезы не случайно и иметь многоуровневый процесс валидации от дешевого и неточного этапа до дорогого и точного
- Создание белка



» Задачи дизайна ферментов

(До)-дизайн (тюнинг)

Уже есть стартовый вариант, он качественно делает то, что нужно. Хотим количественного улучшения

- * **Редизайн** : Есть стартовый вариант, он делает что-то похожее, но не совсем то, что нужно
- * **Де-ново дизайн** Нет никакого стартового варианта

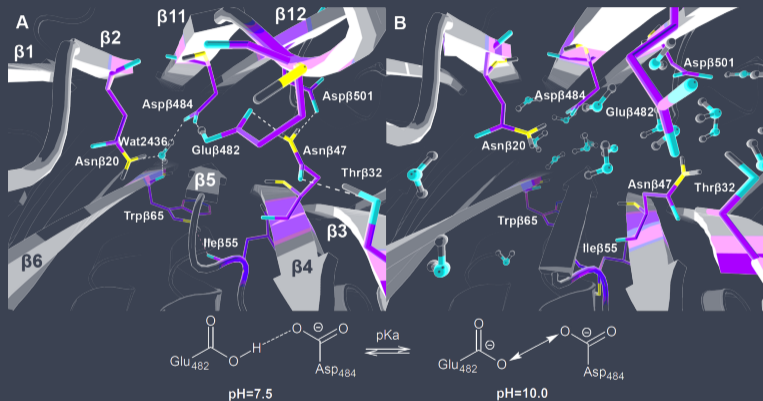


» ТЮНИНГ

- * Стабильность
- * Растворимость
- * рН-зависимость стабильности
- * рН-оптимум активности
- * Улучшение аффинности связывания
- * Улучшение константы реакции



» рН-зависимость стабильности



В пенициллин ацилазе в щелочных условиях разрушается водородная связь между E482 и D484. Замена D484N повысила стабильность в щелочной среде в 9 раз.

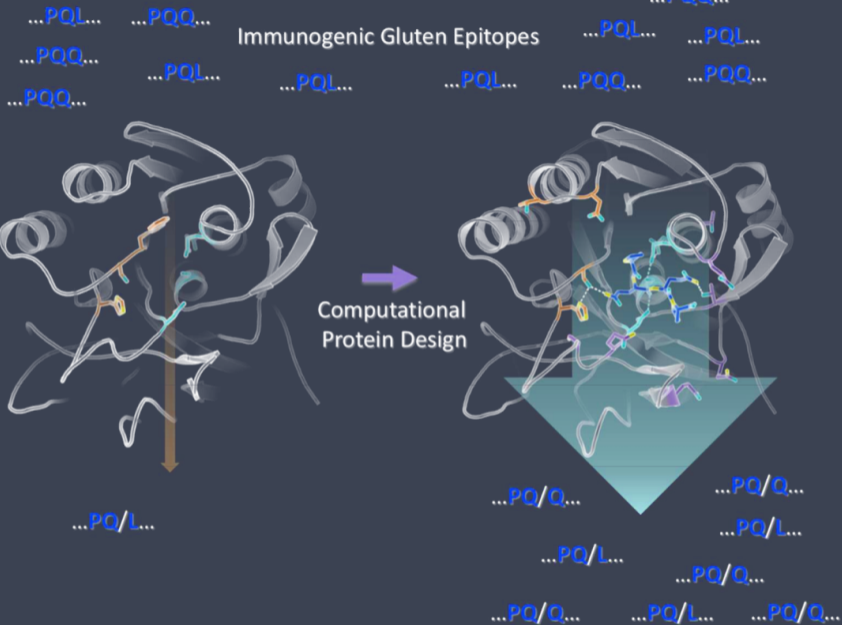


» Аффинность

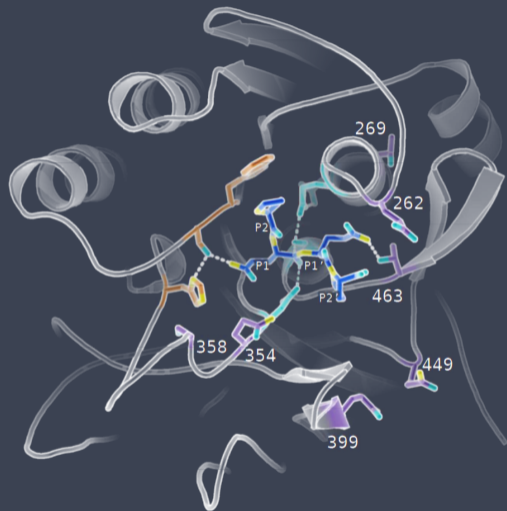
Задача ничем не отличается от улучшения неферментативного кармана связывания (за исключением строгой немутабельности каталитических остатков).

- * **Как решаем:** МС перебор с помощью Rosetta
- * **Нюанс:** Для ферментов такая задача ставится редко. Гораздо выгоднее понизить барьер активации.





» Оптимизация кумамолизина

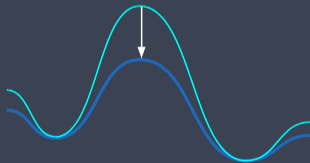


Стартовая идея о том, как должен располагаться нужный субстрат. Раунды МС перебора остатков фермента вокруг. Допуск небольшой подвижности остова улучшает результаты, но сильно дороже вычислительно.



» Константа

- * Стабилизация переходного состояния – дизайн непосредственно контактирующих остатков
- * Оптимизация путей перераспределения заряда – дизайн электростатического поля, может достигаться заменами в остатках на удалении от субстрата
- * Тривиальный дизайн с быстрой валидацией (замена вносится в состояние реагентов, моделируется реакция с оценкой барьера)



» Стратегия 1

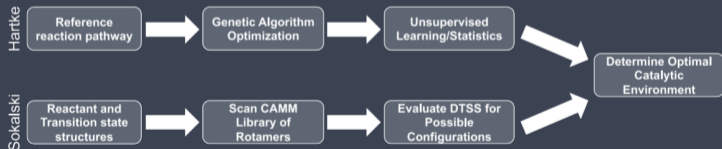
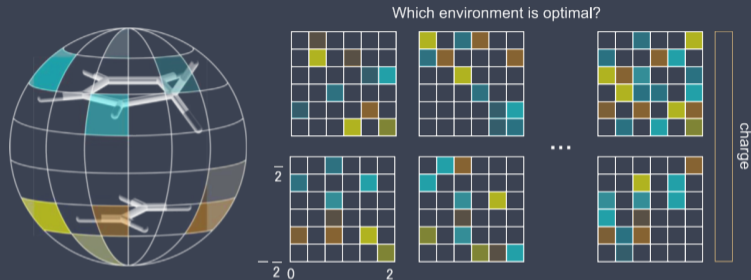
Стратегия 1 крайне редко используется для тюнинга

- * природные сайты уже максимально оптимизированы эволюционно
- * даже если есть теоретическая возможность улучшения, ее размер меньше типичной ошибки силового поля Rosetta

Где используется – редизайн и де-ново дизайн (обсудим дальше).



» Стратегия 2

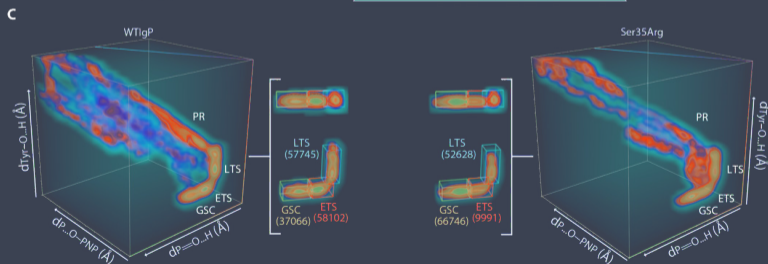
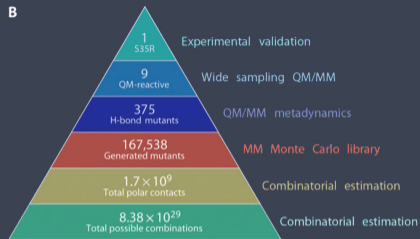
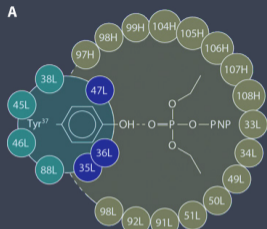


Current Opinion in Structural Biology

Стратегия “снизу-вверх”. Попытка реконструировать необходимое поле от структуры переходного состояния или от пути реакции. Предложена недавно, показала применимость на модельных системах. Однако пока не применялась для создания нового продукта.



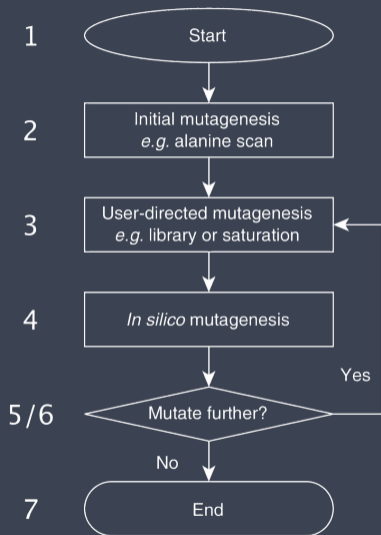
» Стратегия 3



Гипотезы о вариантах:
рациональный ручной
выбор
Валидация: QM/MM
метадинамика



» Стратегия 3: CADEE



Гипотезы о вариантах:
Ala-сканирование для
выявления хотспотов с
вычислением барьера с
помощью EVB, мутация на все,
снова вычисление барьера с
помощью EVB
Валидация: вшита в генерацию
гипотез (EVB)

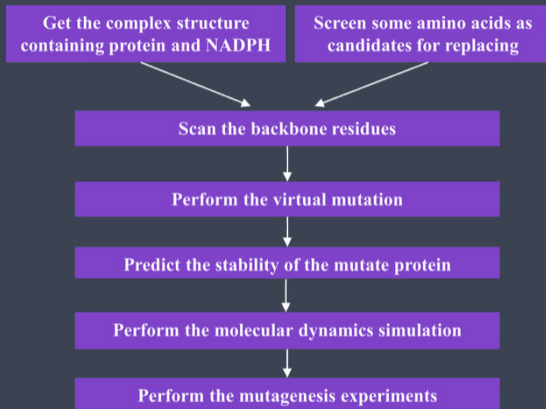


» Редизайн

- * **Как решаем:** Кофакторная специфичность
- * **Как решаем:** Субстратная специфичность
- * **Как решаем:** Предпочтительный механизм



» Кофакторная специфичность

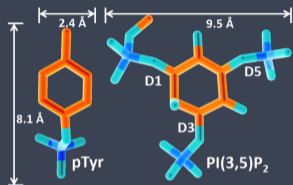


Концептуально то же самое, что дизайн кармана связывания
Кейс: фермент работает с НАДФ, а мы хотим заставить работать с НАД, так как в клетке его больше. Или наоборот. По меньшей мере 33 успешные работы к данному моменту для пары НАД/НАДФ. Для чего-то еще – исчезающие количества.



» Субстратная специфичность

А

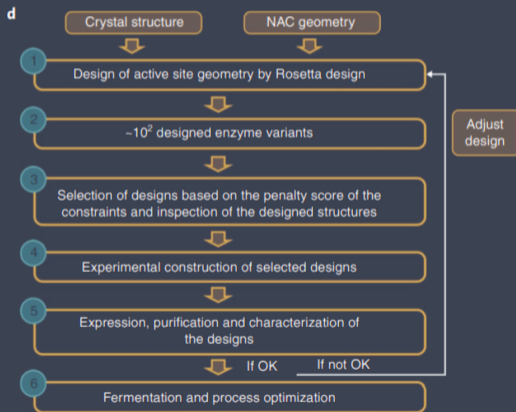
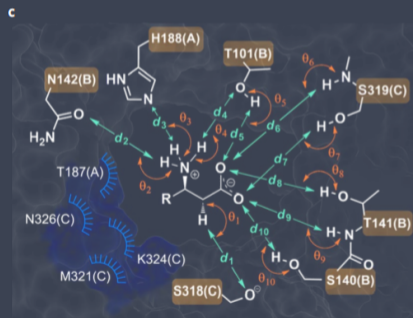
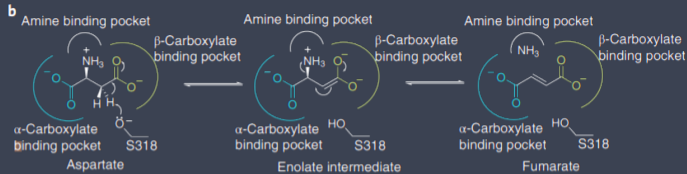
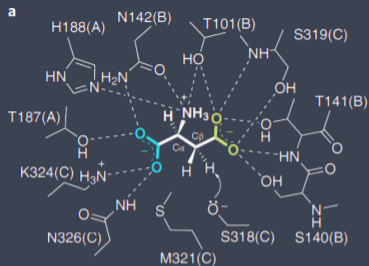


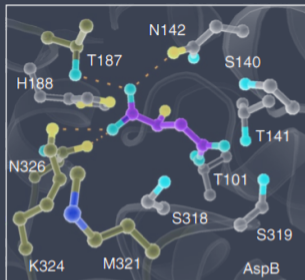
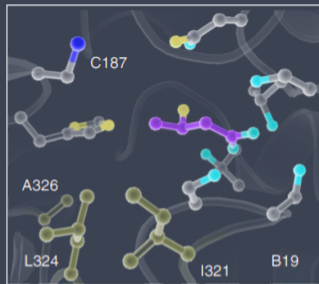
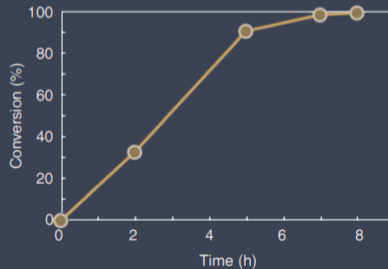
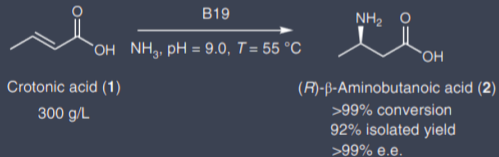
Пример: тирозин-фосфатаза снимает фосфат с тирозина. Хотим, чтобы снимала с инозитол-фосфата. Реакция одна и та же – нужно добиться хорошего позиционирования нового субстрата.

Что делали: Rosetta

Что получили: дизайны с k_{cat}/K_m на 1 порядок хуже, чем для пары дикий тип-фосфотирозин (K_m при этом лучше, k_{cat} сильно хуже)

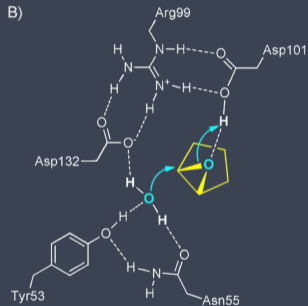




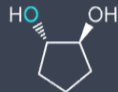
b**c****d**

» Региоспецифичность

Кейс: энантиоселективные варианты эпоксид гидролазы

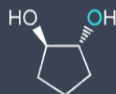


1a



(S,S)-1b

+



(R,R)-1b

Для каждого варианта нужно обеспечить связывание, приводящее к предпочтительной атаке по нужному атому углерода. Положение атакующей воды лучше не менять, ибо оно уже оптимизировано для низких барьеров.

Как: Rosetta + валидация дизайнов короткими MD

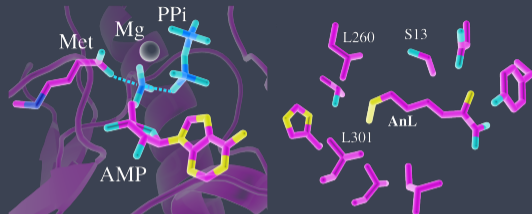


» Дизайн субстратной специфичности от структуры TS

Можем заглянуть немного в будущее и оптимизировать непосредственно связывание переходного состояния под версию субстрата.

Пример: заставить метионил-тРНК синтетазу пришивать вместо метионина что-то другое.

Как: Proteus (Wang-Landau MC).



В MC мы ориентируемся на скор, если он улучшается, мы принимаем изменение. Положительное изменение скор не обязательно значит образование взаимодействий с лигандом/TS. Можем сделать предварительный MC скан на аро-ферменте и сконструировать поправку к скору, делающую все замены равнозначными, чтобы отбирать только те замены, которые повышают скор исключительно благодаря взаимодействиям с лигандом/TS.



Де-ново дизайн

Новая молекула с заданной ферментативной функцией

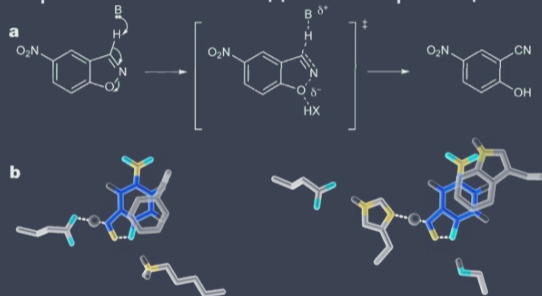
» Переходное состояние как лиганд

Допустим, мы уже умеем дизайннить карманы связывания под малые молекулы. Нам нужно

- * Описать переходное состояние субстрата как малую молекулу
- * Зафиксировать непосредственно участвующие в реакции остатки
- * Остатки + переходное состояние = теозим

Задача сводится к поиску места для размещения такого теозима в т.н. структурных матрицах (скэффолдах)

Варианты теозима для этой реакции:

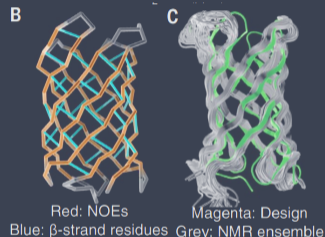


» Откуда брать скэффолды?

Допустим, мы уже умеем дизайнить карманы связывания под малые молекулы. Нам нужно

- * Весь PDB
- * Весь метапротеом + AlphaFold2
- * Искусственные скэффолды
Пока умеем делать что-то простое: фолды из элементов надвторичной структуры, простые альфа-пучки, бета-бочонки
- * Предковая реконструкция
Идея: современные ферменты – оптимизированные специалисты. Если начнем что-то менять, скорее всего сломаем. Но они произошли от предка-генералиста. Попробуем предсказать его последовательность, получим структуру и возьмем ее за старт наших дизайнов. Описать переходное состояние субстрата как малую молекулу

Хотелось бы воображать скэффолд под конкретный активный центр, но этого пока даже близко нет



» Дизайн ферментов: стабилизация TS

Удалось сделать несколько де-ново ферментов (для реакций, не имеющих своего фермента в природе): Кемп-Элиминаза, Дильс-Альдераза

- * Только 1 стадия
- * Любое число > 0 = успех
- * Структура TS получена из KM моделирования Весь PDB

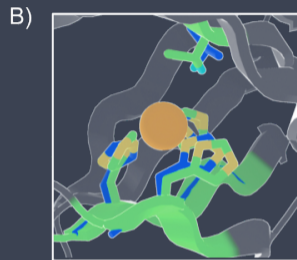
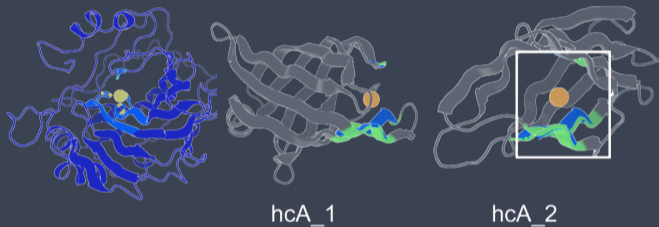
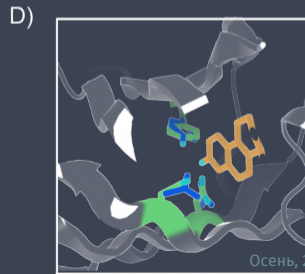
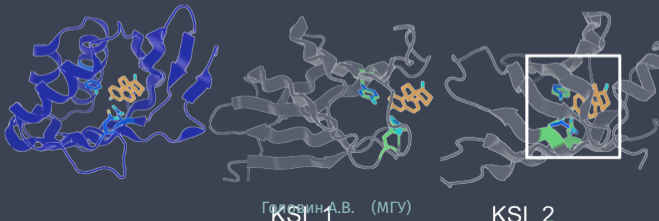
Не известны работы по успешному де-ново дизайну ферментов для сложных реакций (больше 1 стадии, наличие кофакторов/коферментов)

Не удается улучшить уже имеющиеся ферменты

- * Ошибка в вычислении скара выше, чем ожидаемый результат
- * Имеющиеся ферменты это уже итог долгой эволюции – простые решения (например, единичные замены) исчерпаны. Требуется масштабный редизайн
- * Очень просто сломать



» RF-design

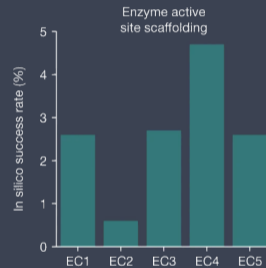
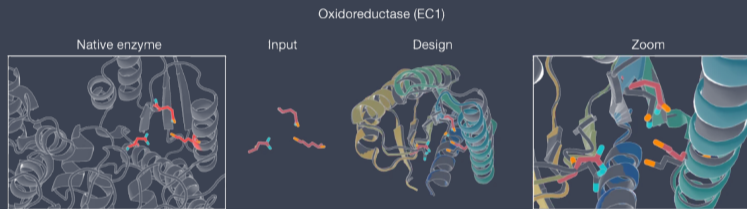
Carbonic Anhydrase II (*in silico*)KSI (*in silico*)

» Tricks

- * **Неудачно:** градиентный спуск путем обратного распространения ошибки через АФ, плохо ложилась боковые цепи
- * **Лучше:** Двухэтапный подход с использованием как АФ, так и trRosetta (для сглаживания ландшафта loss функции) и описания активного сайта на уровне остова. Второй этап это оптимизация последовательности с АФ, при сохранении хода остова (fixed bb design)



» RF-diffusion



» External potential

Algorithm 5 Generation with guidance

```

1: function SAMPLEGUIDED(L, P, GuideScale)
2:   ▷ Generation of  $L$ -residue backbone structure, guided by potential  $P$ 
3:    $x^{(T)} = \text{SampleReference}(M)$ 
4:   for  $t = T, \dots, 1$  do
5:      $\hat{x}^{(0)} = \text{RFDiffusion}(x^{(t)})$ 
6:      $x^{(t-1)} = \text{ReverseStep}(x^{(t)}, \hat{x}^{(0)})$ 
7:      $x^{(t-1)} = x^{(t-1)} + \text{GuideScale}(t) \nabla_{x^{(t)}} P(x^{(t)})$            ▷ Apply guidance
8:   end for
9:   return  $\hat{x}^{(0)}$ 
10: end function

```





Стабильность и растворимость

Субстратная специфичность

Региоспецифичность

Де-ново

pH

Улучшение кинетики

Внедрение регуляции