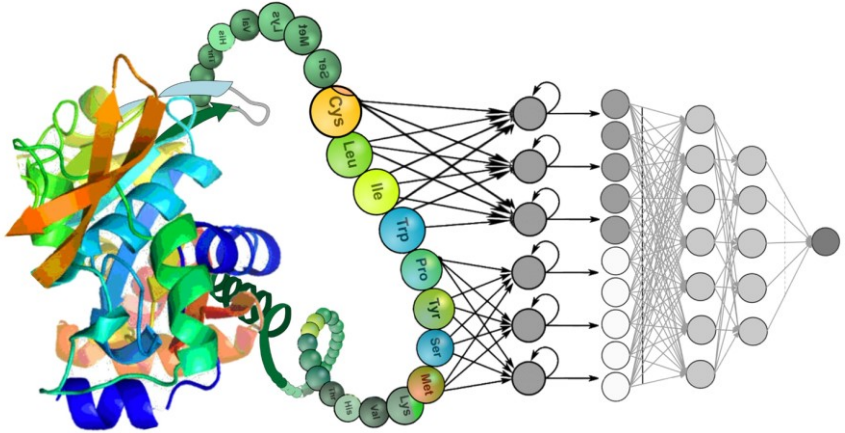
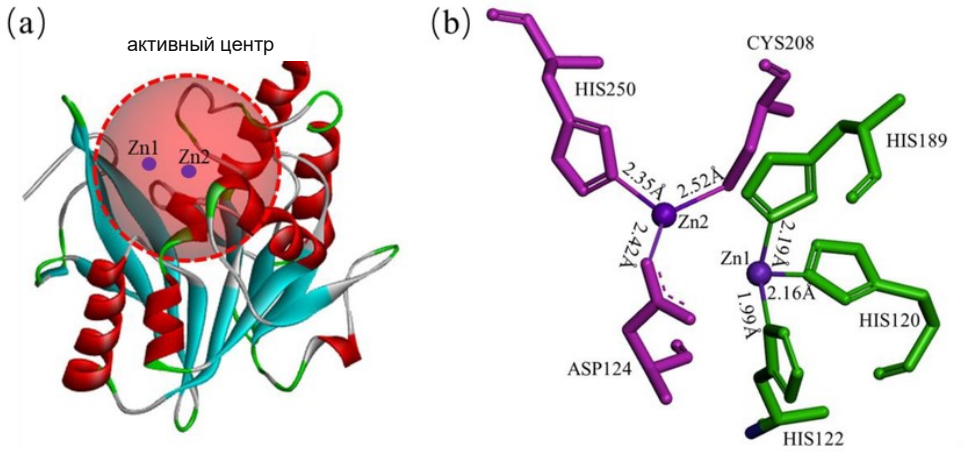
# DE NOVO ДИЗАЙН КАНДИДАТОВ В ИНГИБИТОРЫ NDM-1 ПОД УПРАВЛЕНИЕМ ПРЕДИКТОРА С КОНТРОЛЕМ ЦИНК-СВЯЗЫВАЮЩИХ ГРУПП




Иван Буров

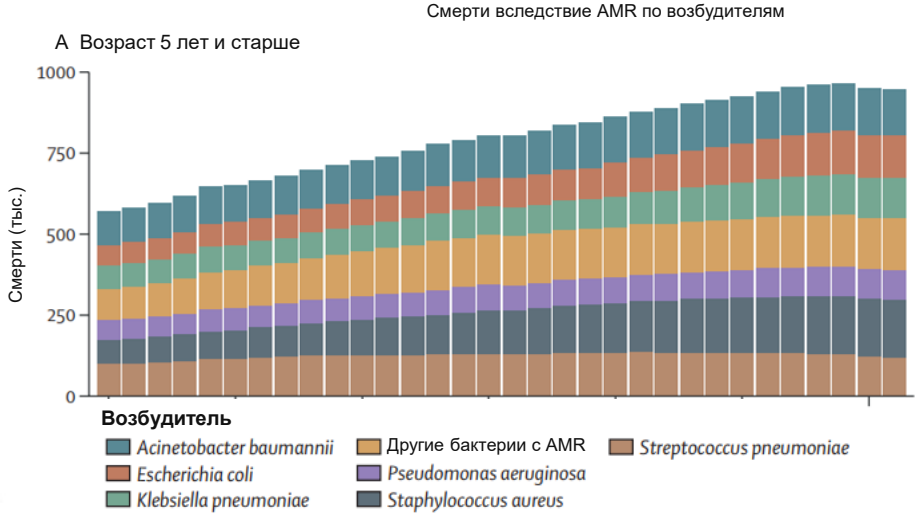
МГУ имени М. В. Ломоносова
Москва, Россия

# Антибиотикорезистентность и NDM-1




**Активный центр NDM-1.**

***Li, X., Zhao, D., Li, W., Sun, J., & Zhang, X. (2021). Enzyme Inhibitors: The Best Strategy to Tackle Superbug NDM-1 and Its Variants. International journal of molecular sciences, 23(1), 197. https://doi.org/10.3390/ijms23010197***




**Смертность, связанная с устойчивостью к антибиотикам.**

***GBD 2021 Antimicrobial Resistance Collaborators (2024). Global burden of bacterial antimicrobial resistance 1990-2021: a systematic analysis with forecasts to 2050. Lancet (London, England), 404(10459), 1199–1226. https://doi.org/10.1016/S0140-6736(24)01867-1***

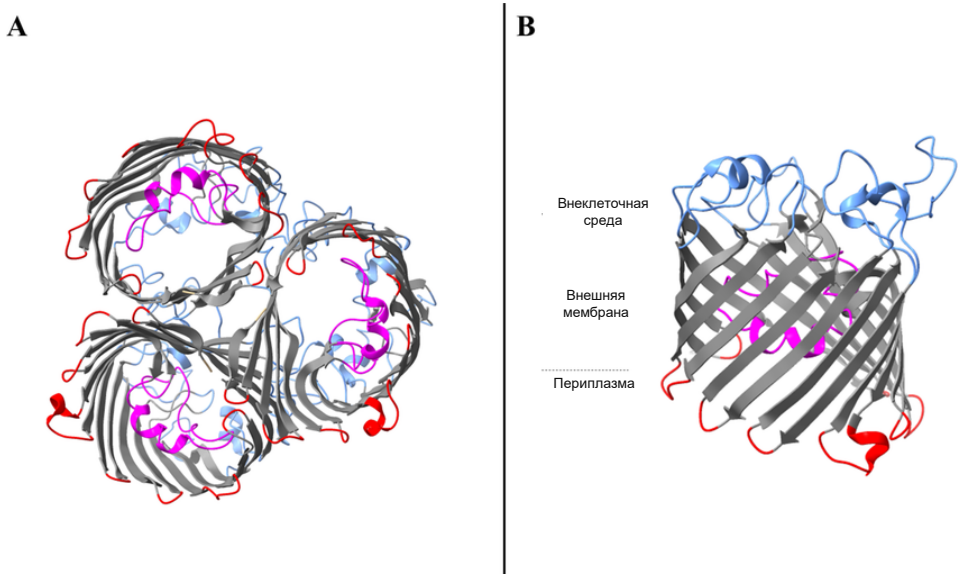
# Барьеры грамотрицательных бактерий

## Требования к молекулам

Ингибирование NDM-1

Проникновение через оболочку грамотрицательных бактерий

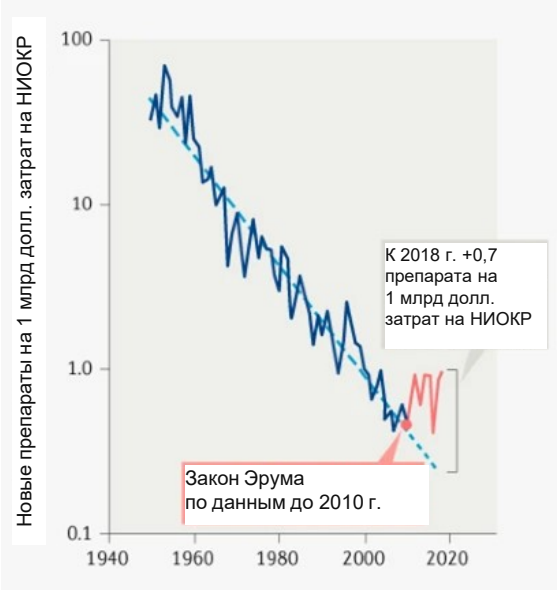
Внешняя мембрана, пептидогликан и внутренняя мембрана




**Структура порина OmpF Escherichia coli.**

***Mammeri, H., Sereme, Y., Toumi, E., Faury, H., & Skurnik, D. (2025). Interplay between porin deficiency, fitness, and virulence in carbapenem-non-susceptible Pseudomonas aeruginosa and Enterobacteriaceae. PLoS pathogens, 21(2), e1012902. https://doi.org/10.1371/journal.ppat.1012902***

# Машинное обучение в разработке лекарств

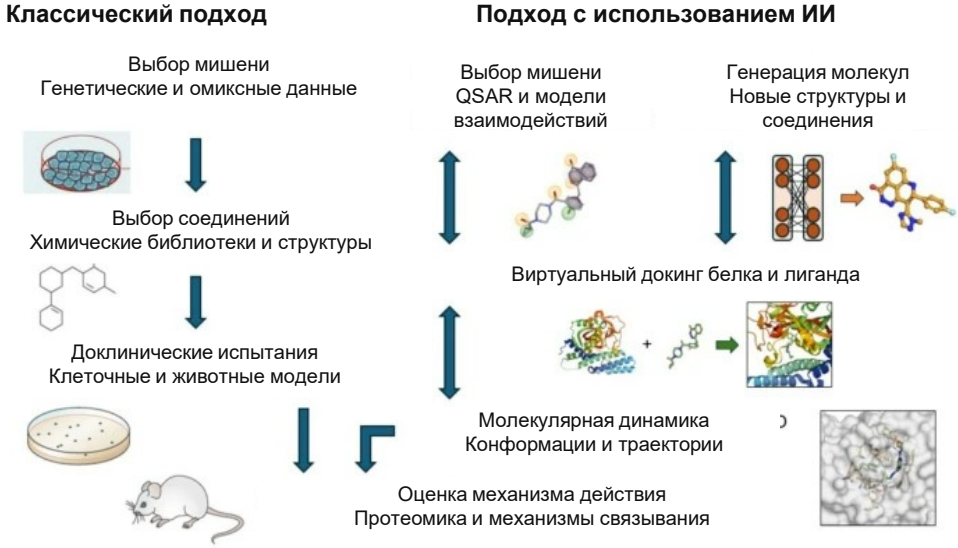



**Эффективность разработки лекарств**
***Ringel et al. (2020), Nat. Rev. Drug Discov. 19, 833–834***
***doi:10.1038/d41573-020-00059-3***




**Итеративный скрининг с машинным обучением**
***Müller et al. (2025), J. Med. Chem. 68, 14806–14817***
***doi:10.1021/acs.jmedchem.5c00943***

# Цель и план исследования




**Дизайн молекул с помощью ИИ**

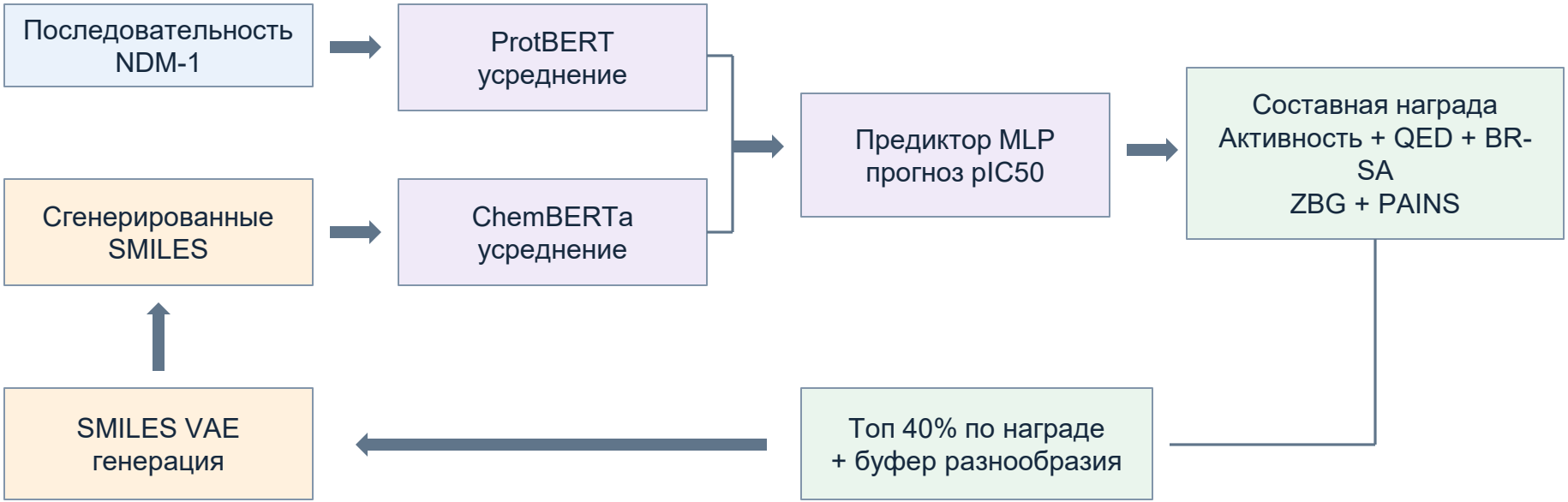
Цель
Получить кандидаты в ингибиторы NDM-1 с заданным числом цинк-связывающих групп.

Подход
Предсказать pIC50 по представлениям белка и молекулы.

Направлять генерацию VAE с учётом активности, свойств и числа ZBG.

Отбирать кандидатов с помощью калиброванного протокола докинга.

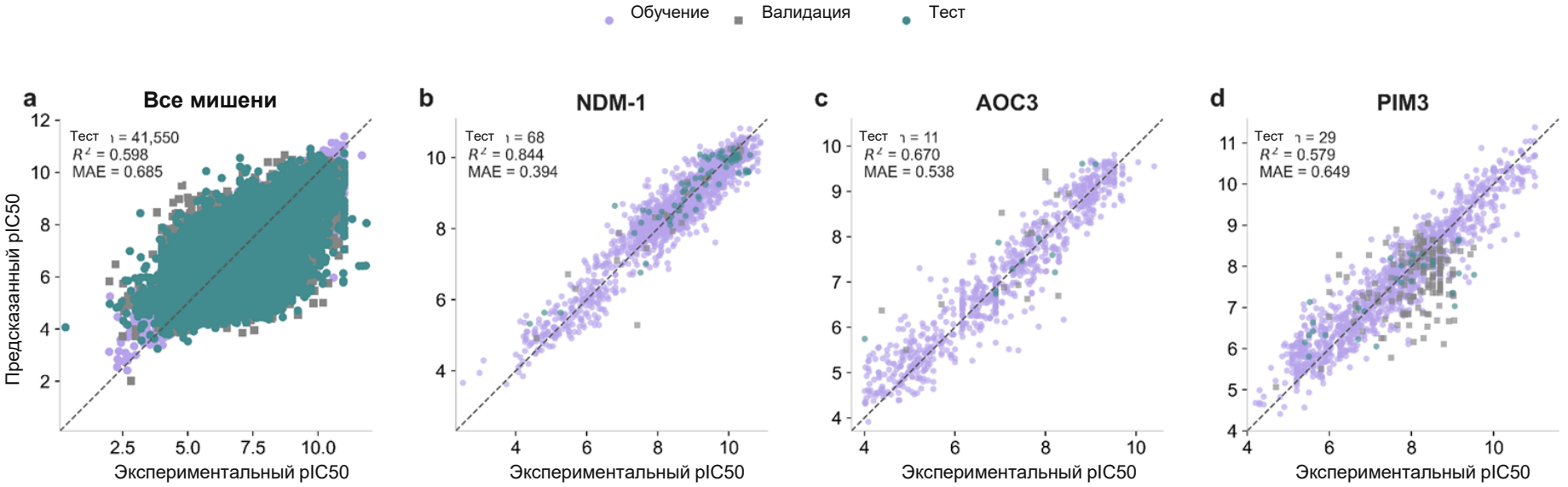
# Генерация под управлением предиктора




Дообучение VAE на референсных и отобранных молекулах

$$L = 0{,}8\ L_VAE(референсные) + 4\ L_VAE(отобранные)$$

# Результаты предиктора




**Внутренний тест NDM-1: n = 68, MAE = 0,394, $R^2$ = 0,844**

Внутренний тест не подтверждает активность сгенерированных молекул.

# Генерация при разных значениях ZBG

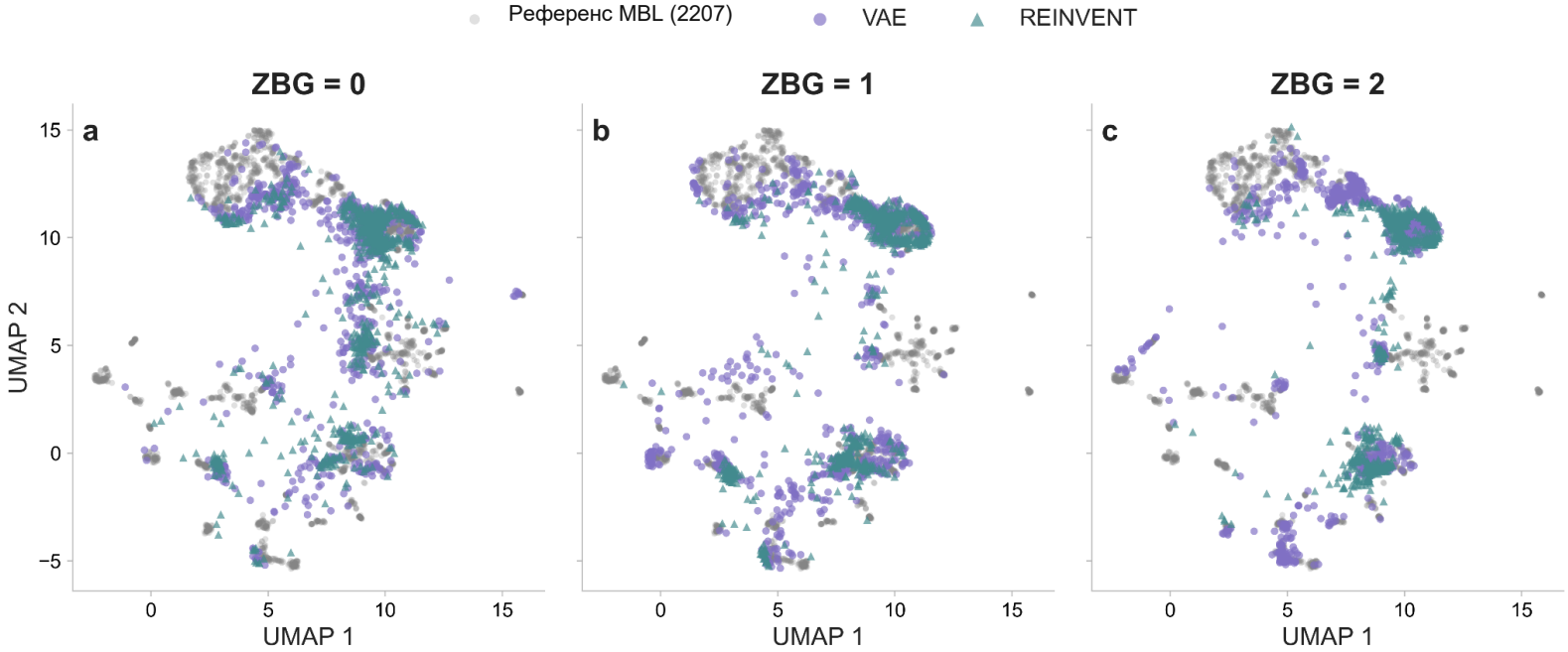
BR-SAScore и обязательный фильтр PAINS

| Число ZBG | Кандидаты | Шаги | Время |
|---|---|---|---|
| 0 | 1002 | 585 | 1 ч 10 мин |
| 1 | 1000 | 2077 | 3 ч 01 мин |
| 2 | 1000 | 3098 | 3 ч 34 мин |

Отбор: прогноз pIC50 ≥ 8, QED > 0,6, BR-SA < 4
Заданное число ZBG и отсутствие совпадений с PAINS

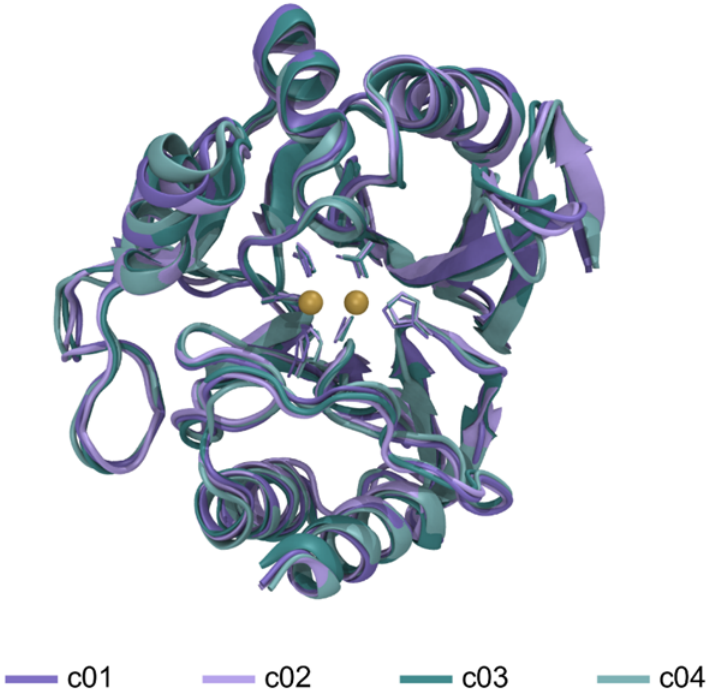
Число кандидатов не подтверждает ингибирование. Запуски не дают строгого сравнения эффективности.

# Химическое пространство




**VAE и REINVENT исследуют разные области химического пространства референсного набора MBL**

# Молекулярная динамика и докинг




Четыре МД-конформации, совмещённые по Cα
Золотые сферы: ионы Zn (c01)

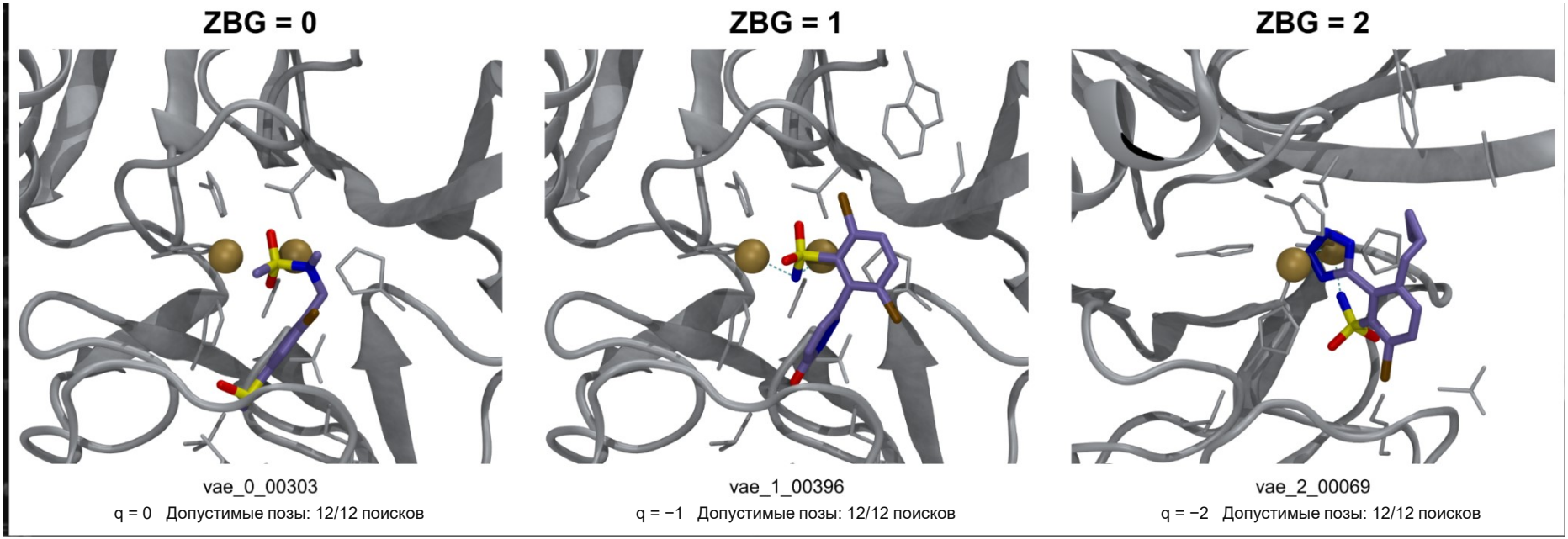
Молекулярная динамика
5ZR8 · CHARMM36 / TIP3P
300 K · 1 атм · последние 500 нс

Ансамбль конформаций
PCA углов φ/ψ основной цепи и кластеризация
Четыре репрезентативные конформации

Докинг с учётом металла
Калиброванные потенциалы AutoDock4Zn
Сетка 90 × 90 × 90 с шагом 0,25 Å

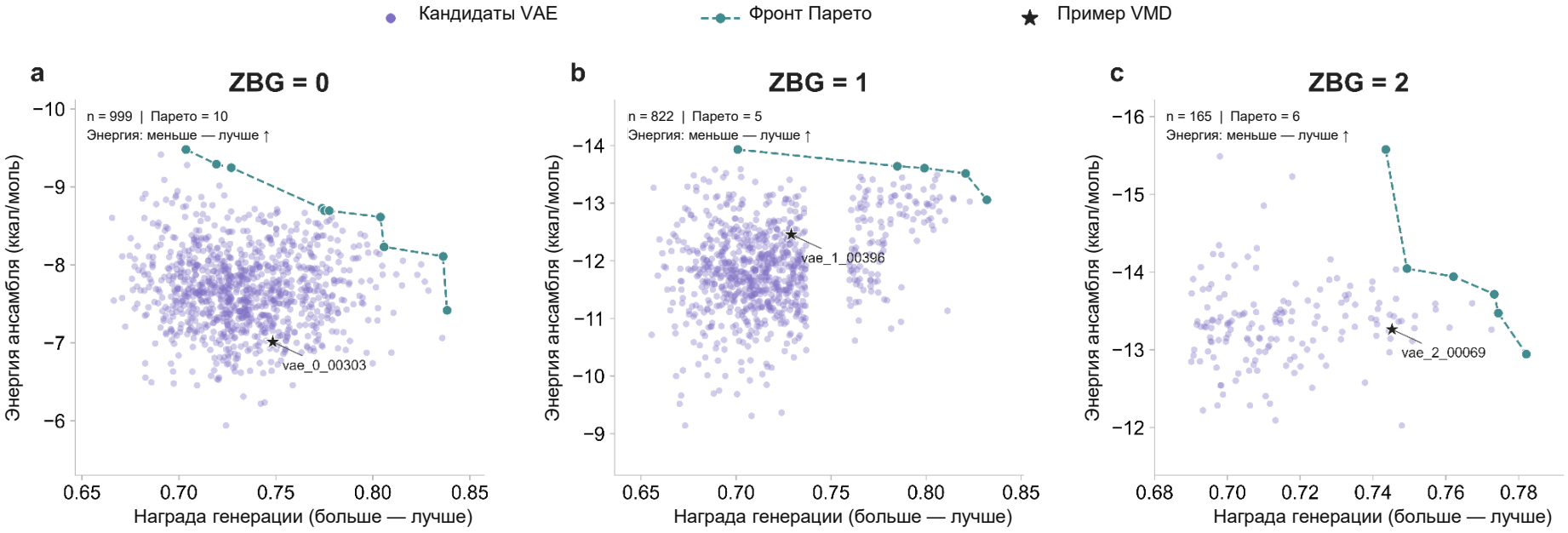
Поиск и отбор поз
3 seed × 10 запусков ГА на конформацию
2,5 × $10^6$ оценок на запуск ГА
Все позы сохраняются, отбор по геометрии

# Докинг на ансамбле конформаций




**У выбранных кандидатов есть допустимые позы во всех 12 поисках для каждой молекулы**

# Графики Парето




**Награда генерации и энергия докинга на ансамбле**
**Следующий этап: оценка выбранных кандидатов методом QM/MM**

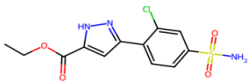
# Выводы и дальнейшие шаги

Предиктор с учётом последовательности белка направляет генерацию VAE для NDM-1.

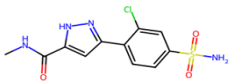
Контроль ZBG и фильтры дают около 1000 кандидатов для каждого режима.

Далее: QM/MM, оценка возможности синтеза и биохимические испытания.

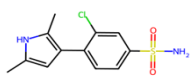
Примеры из исходного набора генерации. Активность пока только предсказана.



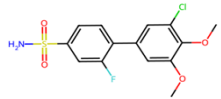
#274 p50:8.8 MW:330 SA:2.3



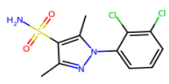
#412 p50:8.7 MW:315 SA:2.4



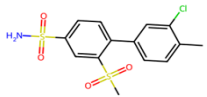
#969 p50:8.1 MW:285 SA:2.4



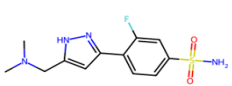
#255 p50:8.1 MW:346 SA:2.1



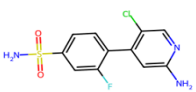
#52 p50:8.1 MW:320 SA:2.4



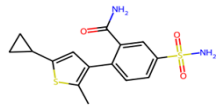
#438 p50:8.2 MW:360 SA:2.1



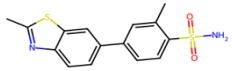
#59 p50:8.2 MW:298 SA:2.5



#316 p50:8.1 MW:302 SA:2.3



#329 p50:8.4 MW:336 SA:2.6



#156 p50:8.4 MW:318 SA:2.2

# Дополнение: функция награды

$$\mathbf{R_{total} = R_{ZBG} \times R_{PAINS} \times (w_{act} R_{act} + w_{qed} R_{QED} + w_{sa} R_{SA})}$$

$$R_{act} = \max\left(0, \min\left(1, \frac{\widehat{pIC_{50}} - 4.0}{6.0}\right)\right)$$

$$R_{QED} = \mathrm{QED(mol)}$$

$$R_{SA} = \max\left(0, \min\left(1, \frac{10 - \mathrm{BR{-}SA(mol)}}{9}\right)\right)$$

$R_{ZBG}(n, t; k)$

n: число ZBG в молекуле; k: целевое число

| $k$ | $n=0$ | $n=1$ | $n=2$ | $n \geq 3$ |
|---|---|---|---|---|
| 0 | 1.0 | 0.05 | 0 | 0 |
| 1 | $P_{zero}(t)$ | 1.0 | 0.05 | 0 |
| 2 | 0 | 0.05 | 1.0 | 0 |

R_PAINS = 0 при наличии PAINS; R_PAINS = 1 в остальных случаях.

Веса и штраф за отсутствие ZBG на разных шагах обучения

| Шаг обучения t | $w_{act}$ | $w_{qed}$ | $w_{sa}$ | $P_{zero}(t)$ |
|---|---|---|---|---|
| $t < 1000$ | 0.6 | 0.2 | 0.2 | 0.4 |
| $1000 \leq t < 3000$ | 0.5 | 0.3 | 0.2 | 0.2 |
| $t \geq 3000$ | 0.4 | 0.3 | 0.3 | 0.05 |

# Дополнение: протокол докинга

Типизация атомов
Депротонированный N сульфонамида, N тетразолата, карбоксилатный O и сульфонильный O.

Потенциалы взаимодействия с Zn
Специальные потенциалы донор–Zn, без TZ-псевдоатомов.
Сульфонил-O–Zn: $C_6$ × 0,50; $C_{12}$ без изменений.
Для депротонированных сульфонамидов.

Калибровка
Редокинг кристаллических лигандов, исходные и перевёрнутые позы, перенос на 4 МД-конформации.

**Восстановление поз в МД-контролях**

Исходный потенциал



**26/36**

Сульфонил-O–Zn $C_6$ × 0,50



**32/36**

3 лиганда × 4 МД-конформации × 3 seed
= 36 контрольных комбинаций

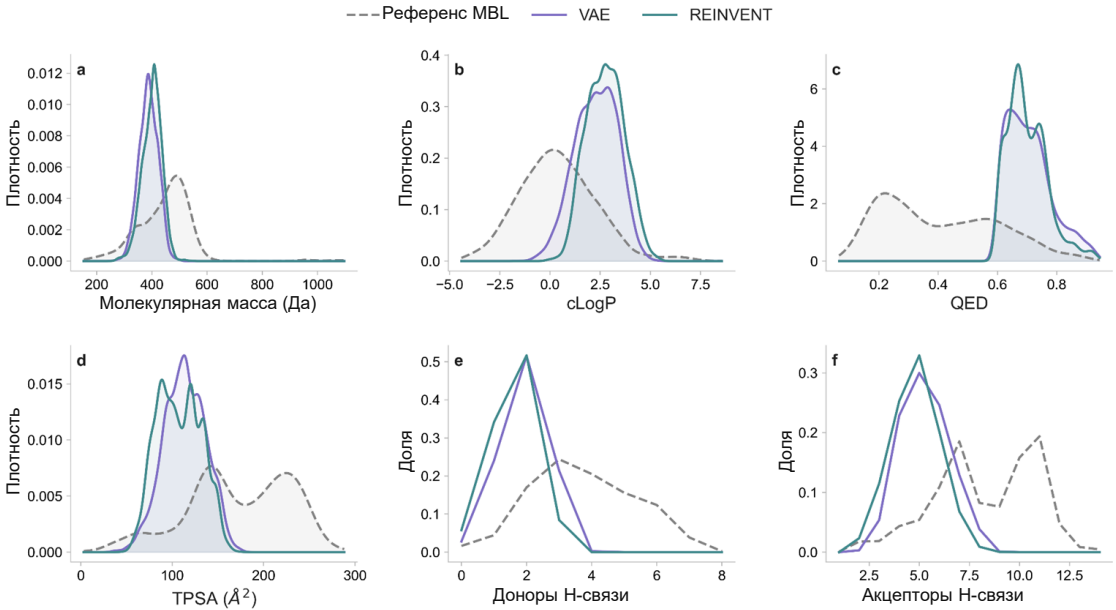
Успех: ≥1 допустимой позы среди 10.
Две разные ZBG контактируют с разными Zn.
Расстояния 1,8–2,8 Å, без слишком коротких контактов.

Референс: потенциал typed-NZ/NT/OX (λ = 2).

**Сохранение всех поз и отбор по геометрии**

Числа отражают наличие допустимых поз, а не успех лучшей позы по скорингу.

# Дополнение: свойства молекул




Референсный набор MBL, VAE и REINVENT

MW: молекулярная масса
cLogP: липофильность
QED: лекарственноподобие
TPSA: полярная поверхность
HBD / HBA: доноры / акцепторы водородной связи

Профили свойств помогают отбирать кандидатов.