

СУПЕРКОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ БИОМОЛЕКУЛЯРНЫХ ПРОЦЕССОВ, СВЯЗАННЫХ С КОРОНАВИРУСОМ

Часть I

Немухин А. В., Григоренко Б.Л., Луцкекина С.В., Поляков И.В., Хренова М.Г.

Химический факультет МГУ

Институт биохимической физики РАН

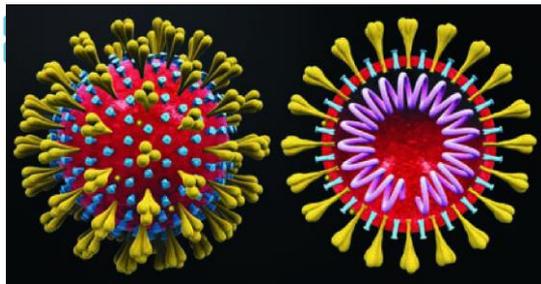
Институт биохимии РАН

24 ноября 2020 г.

Звездный час биомолекулярного компьютерного моделирования



27/03/2020
Российский суперкомпьютер поможет ученым в создании препаратов для борьбы с коронавирусом COVID-19



1184

Суперкомпьютер нашел лекарство против коронавируса

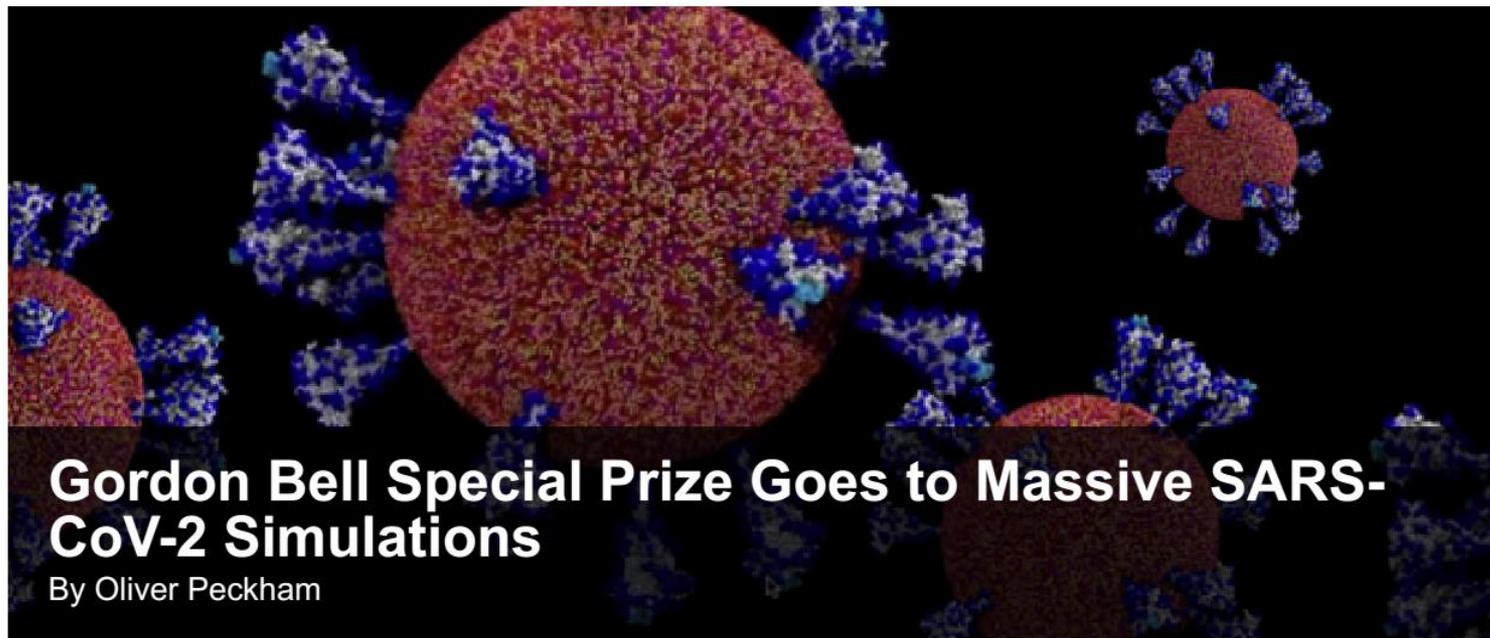
Facebook 3 VK 3 Odnoklassniki 2 Twitter Telegram Добавить в «Мою Ленту»



Фото: megware.com

Компьютерные технологии против коронавируса: первые результаты

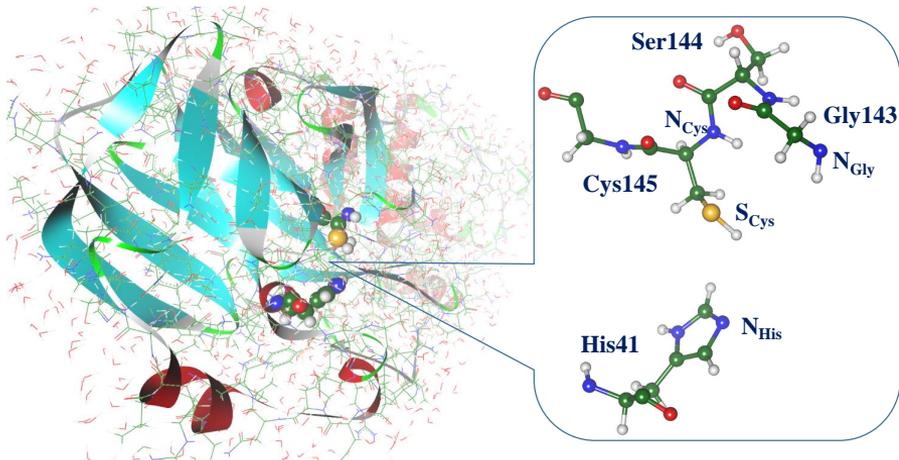
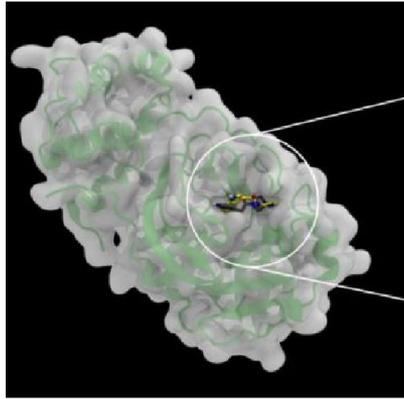
Звездный час биомолекулярного компьютерного моделирования - продолжение



November 19, 2020

2020 has proven a harrowing year – but it has produced remarkable heroes.

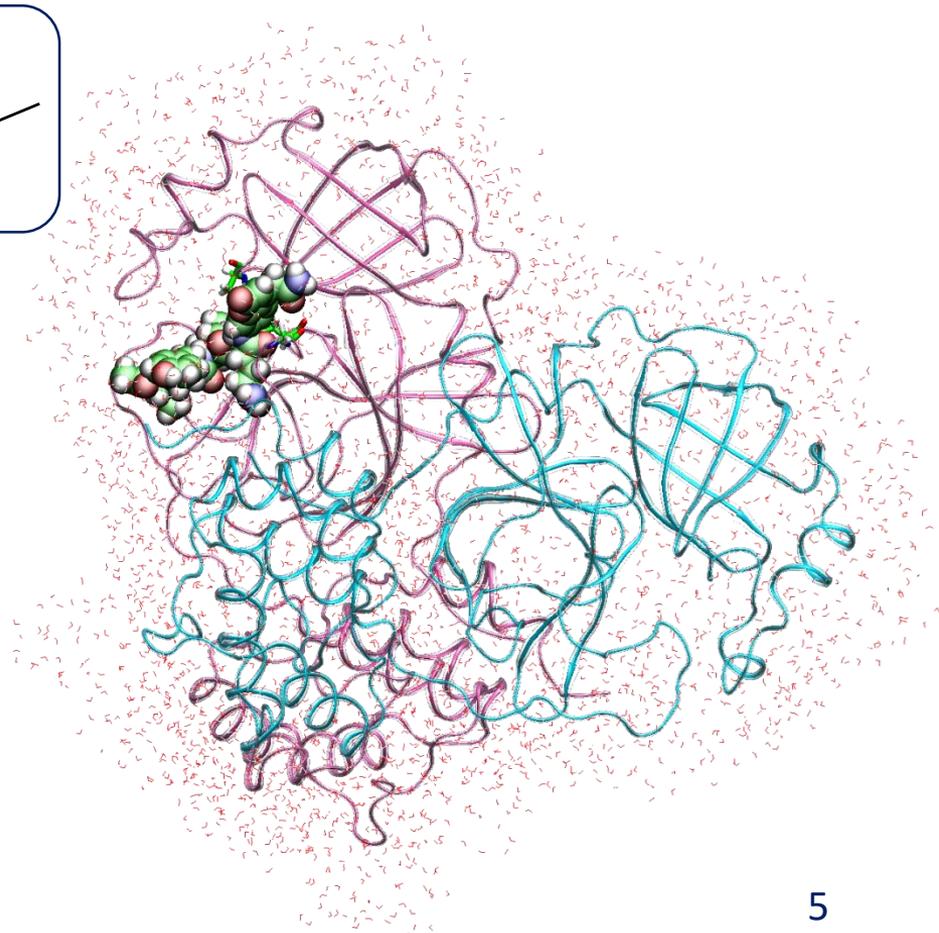
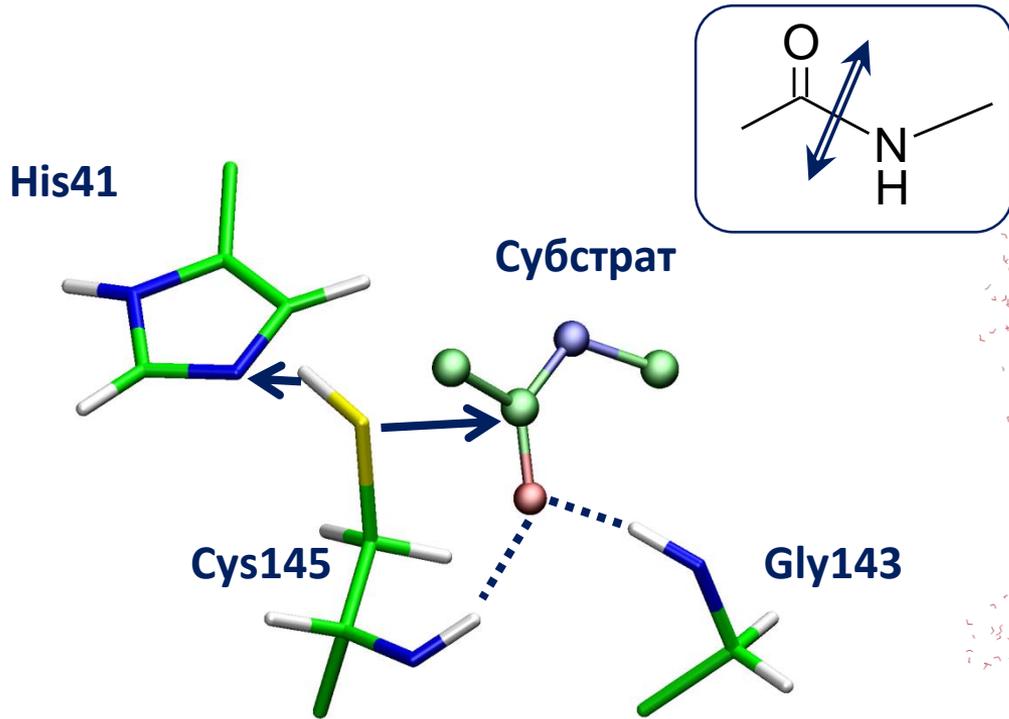
Мишени для лекарств – белки-ферменты



SARS-CoV-2

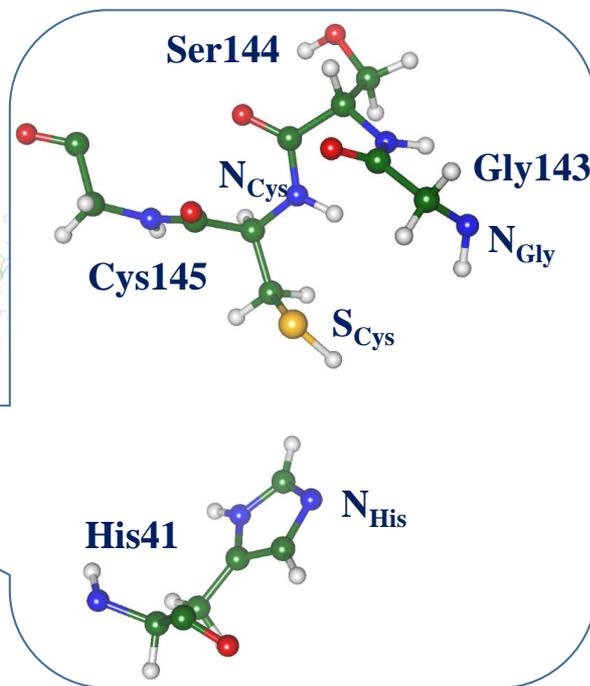
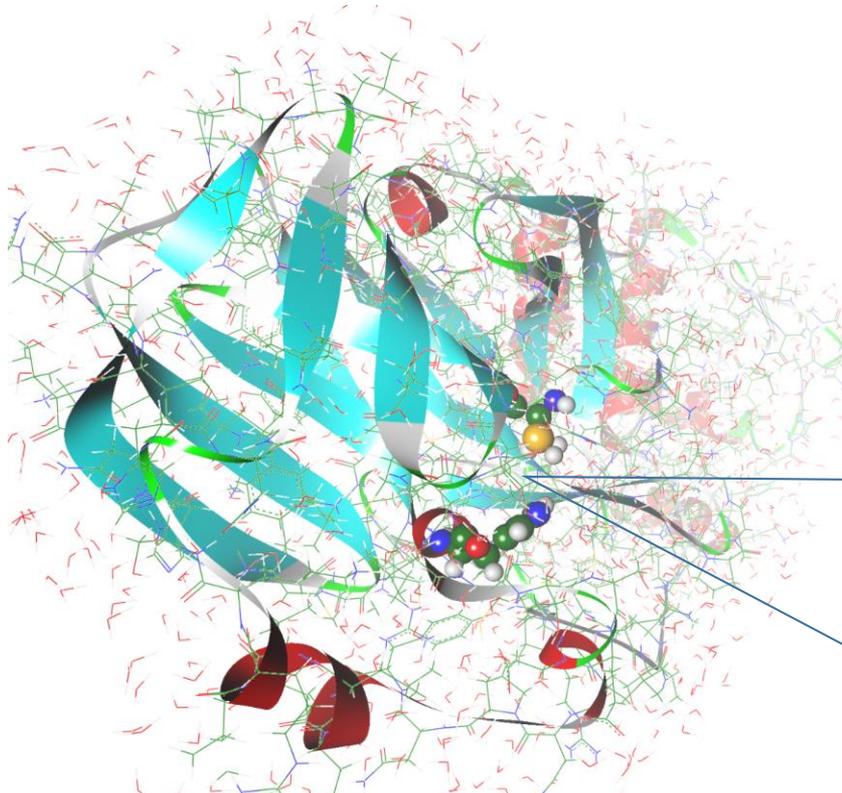
- ✓ “Spike protein”
белок-шип
- ✓ “Main protease/M^{Pro}”
основная протеаза
- ✓ “RNA polymerase”
РНК полимераза
- ✓ ...

Основная протеаза M^{Pro} вируса SARS-CoV-2



Лекарственные вещества – ингибиторы ферментов (нековалентные и ковалентные)

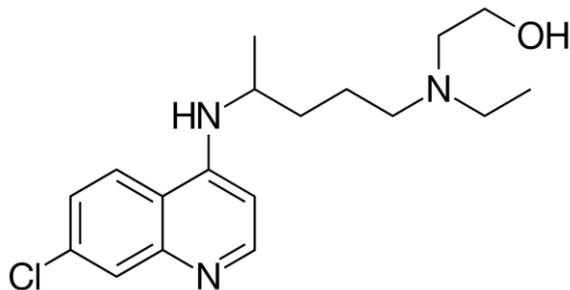
Основная протеаза M^{Pro} вируса SARS-CoV-2



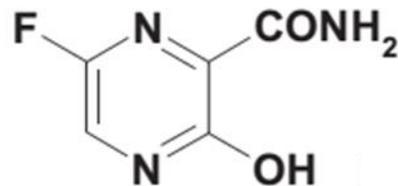
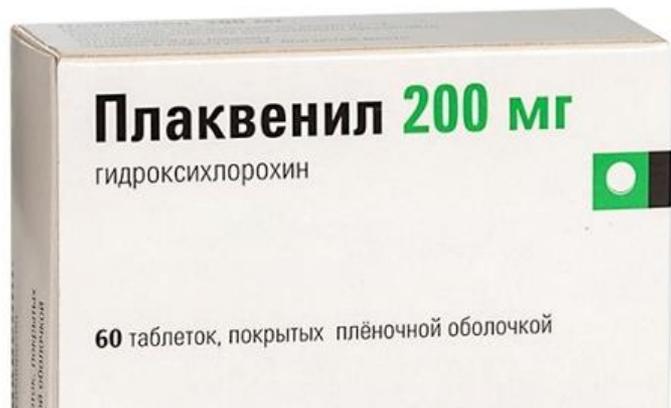
Лекарственные вещества – ингибиторы ферментов

Компьютеры (суперкомпьютеры) необходимы, чтобы построить модели белков, модели молекул потенциальных лекарственных веществ, проанализировать их взаимодействия

Компьютеры необходимы, чтобы построить модели молекул потенциальных лекарственных веществ



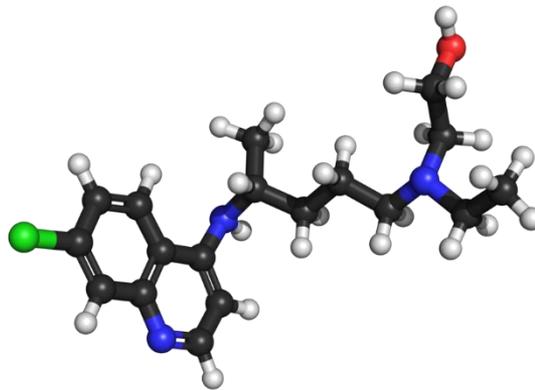
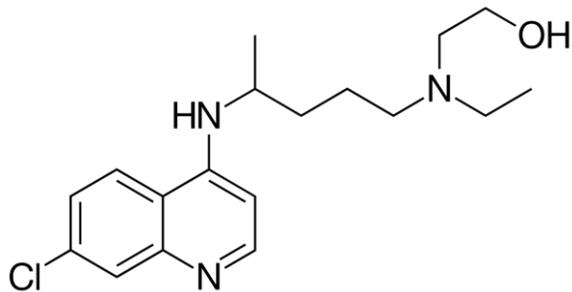
Гидроксихлорохин



Фавипиравир



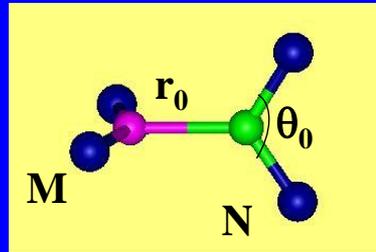
Компьютеры необходимы, чтобы построить модели молекул потенциальных лекарственных веществ



Методы молекулярной механики и квантовой химии

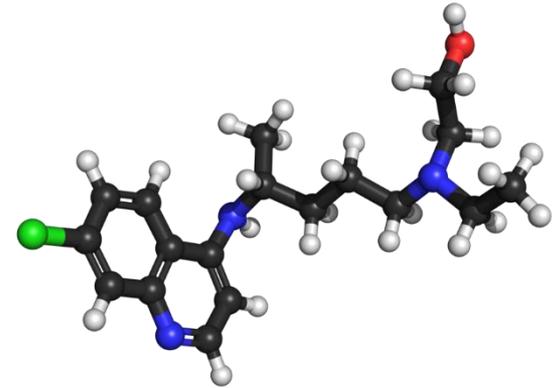
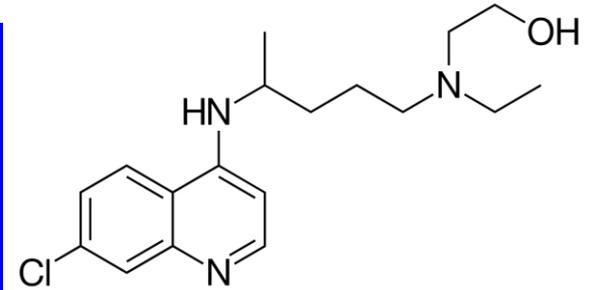
Методы молекулярной механики

Модель - атомы



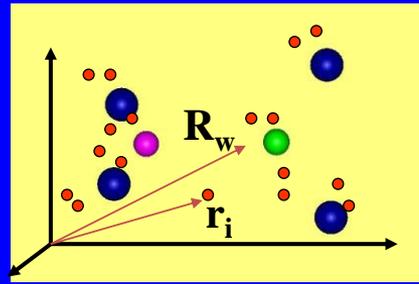
$$E_{MM} = \sum_{i=\text{stretches}} k_i (r_i - r_{i0})^2 + \sum_{j=\text{bends}} k_j (\theta_j - \theta_{j0})^2 + \sum_{l=\text{torsions}} \left[\frac{V_{l,1}}{2} (1 + \cos \phi_l) + \frac{V_{l,2}}{2} (1 + \cos 2\phi_l) + \dots \right] + \sum_{MN} \frac{q_M q_N}{r_{MN}} + \sum_{MN} 4\epsilon_{MN} \left[\left(\frac{\sigma_{MN}}{r_{MN}} \right)^{12} - \left(\frac{\sigma_{MN}}{r_{MN}} \right)^6 \right]$$

$k_i, r_{i0}, k_j, \theta_{j0}, V_{l,1}, V_{l,2}, q_M, \epsilon_{MN}, \sigma_{MN}$ - Параметры



Методы квантовой химии

Модель – ядра и электроны



$$\hat{H}\Psi = E_{QM}\Psi$$

$$\hat{H} = \sum_i \left(-\frac{1}{2} \Delta_{r_i} \right) + \sum_w \left(-\frac{1}{2M_w} \Delta_{R_w} \right) + \sum_{i < j} \frac{1}{r_{ij}} + \sum_{w < v} \frac{Z_w Z_v}{R_{wv}} - \sum_{iw} \frac{Z_w}{|R_w - r_i|}$$

$$HC = ESC$$

Численное решение матричных уравнений

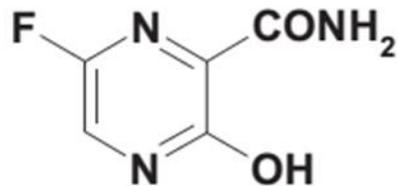
Варианты теории функционала электронной плотности (DFT).
Программы:
Gaussian, Gamess, ORCA, NWChem, Firefly,

Моделирование лекарств методами квантовой химии

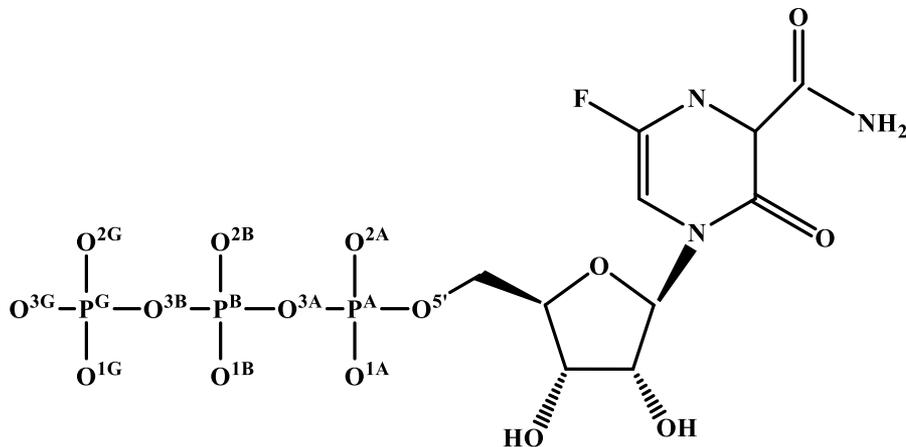
Варианты теории функционала электронной плотности (DFT).

Программы: Gaussian, Gamess, ORCA, NWChem, Firefly,

Проблемы не только вычислительные:



Фавипиравир



Компьютеры (суперкомпьютеры) необходимы, чтобы построить модели белков

Для построения трехмерной полноатомной модели нужны структурные шаблоны: лучше всего - подходящие конструкции из базы данных белковых структур (Protein Data Bank (PDB)).

Если подходящая структура в PDB есть, то дальше модель достраивается методами молекулярной механики (ММ),
**молекулярной динамики (МД), квантовой
механики/молекулярной механики (КМ/ММ).**

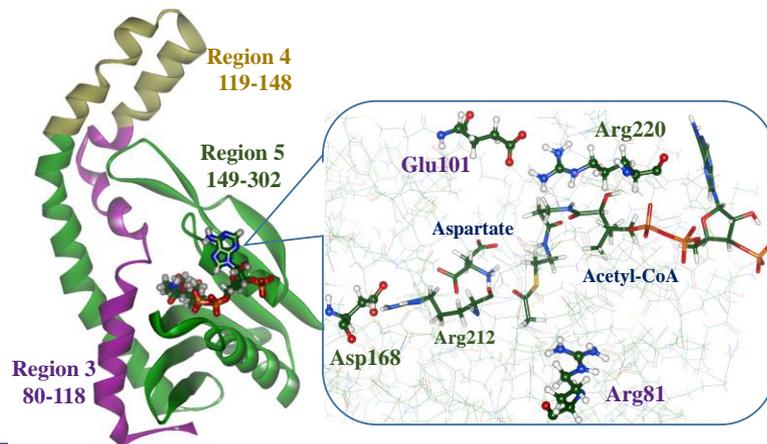
Компьютеры (суперкомпьютеры) необходимы, чтобы построить модели белков

Для построения трехмерной полноатомной модели нужны структурные шаблоны: лучше всего - подходящие конструкции из базы данных белковых структур (Protein Data Bank (PDB)). Если подходящая структура в PDB есть, то дальше модель достраивается методами молекулярной механики (ММ), молекулярной динамики (МД).

Для M^{pro} - в PDB сотни структур, на сайте для молекулярного моделирования [<https://covid.mossi.org>] – 8 файлов готовы для дальнейшей работы

Компьютеры (суперкомпьютеры) необходимы, чтобы построить модели белков

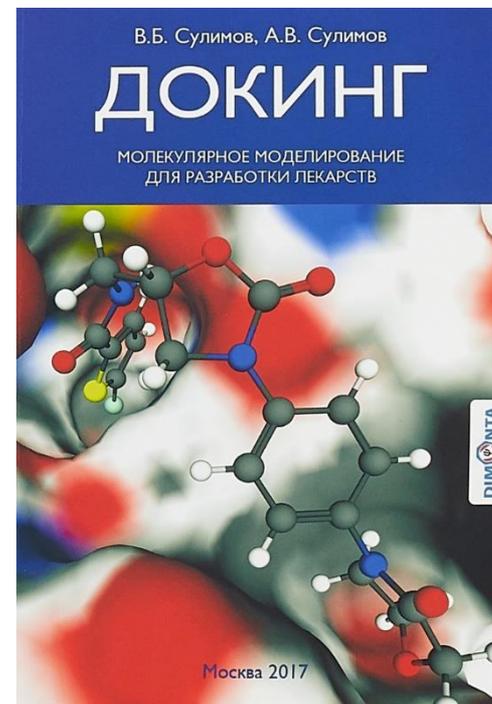
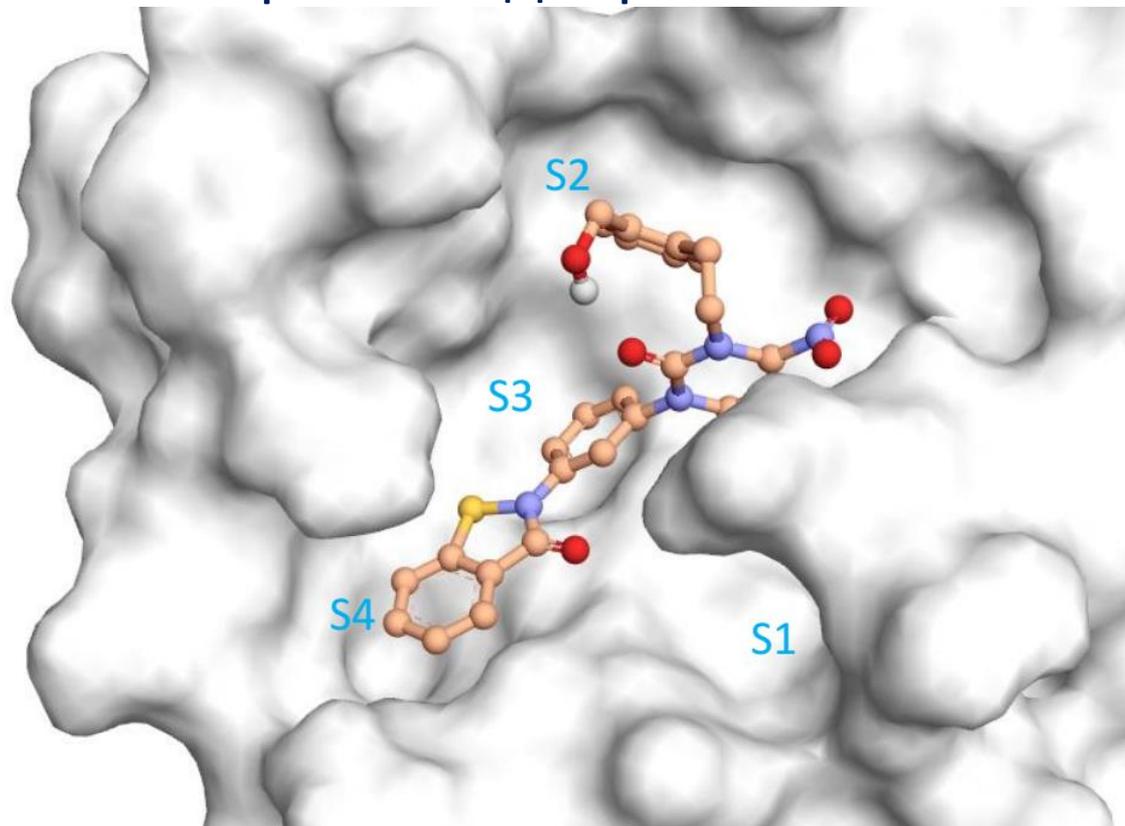
Если структурного шаблона нет, то сначала надо собирать модели, например, методами биоинформатики, а уже затем - манипуляции методами ММ, МД, КМ/ММ



Наш опыт –
построение модели
фермента (NAT8L) по
первичной
последовательности
аминокислотных
ОСТАТКОВ
(Rosetta)

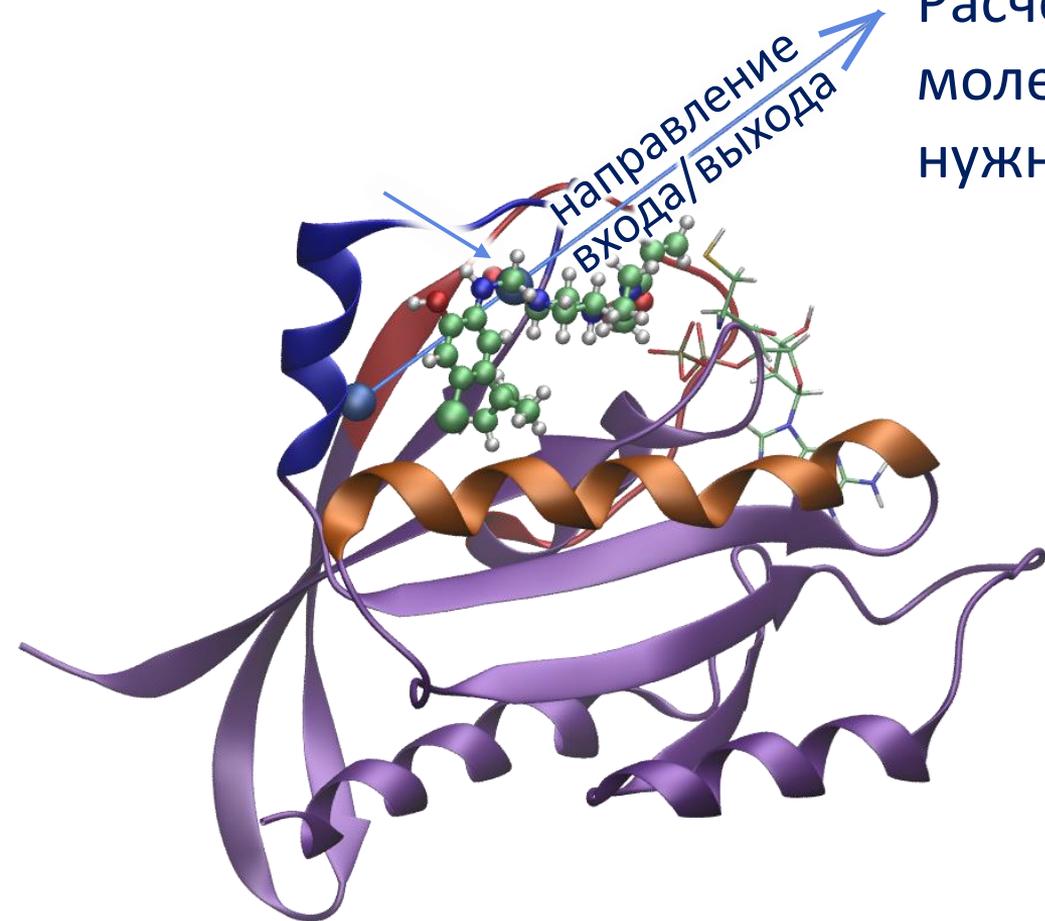
Суперкомпьютеры необходимы, чтобы моделировать взаимодействие белков с «малыми» молекулами

Компьютерный подбор нековалентных ингибиторов - докинг

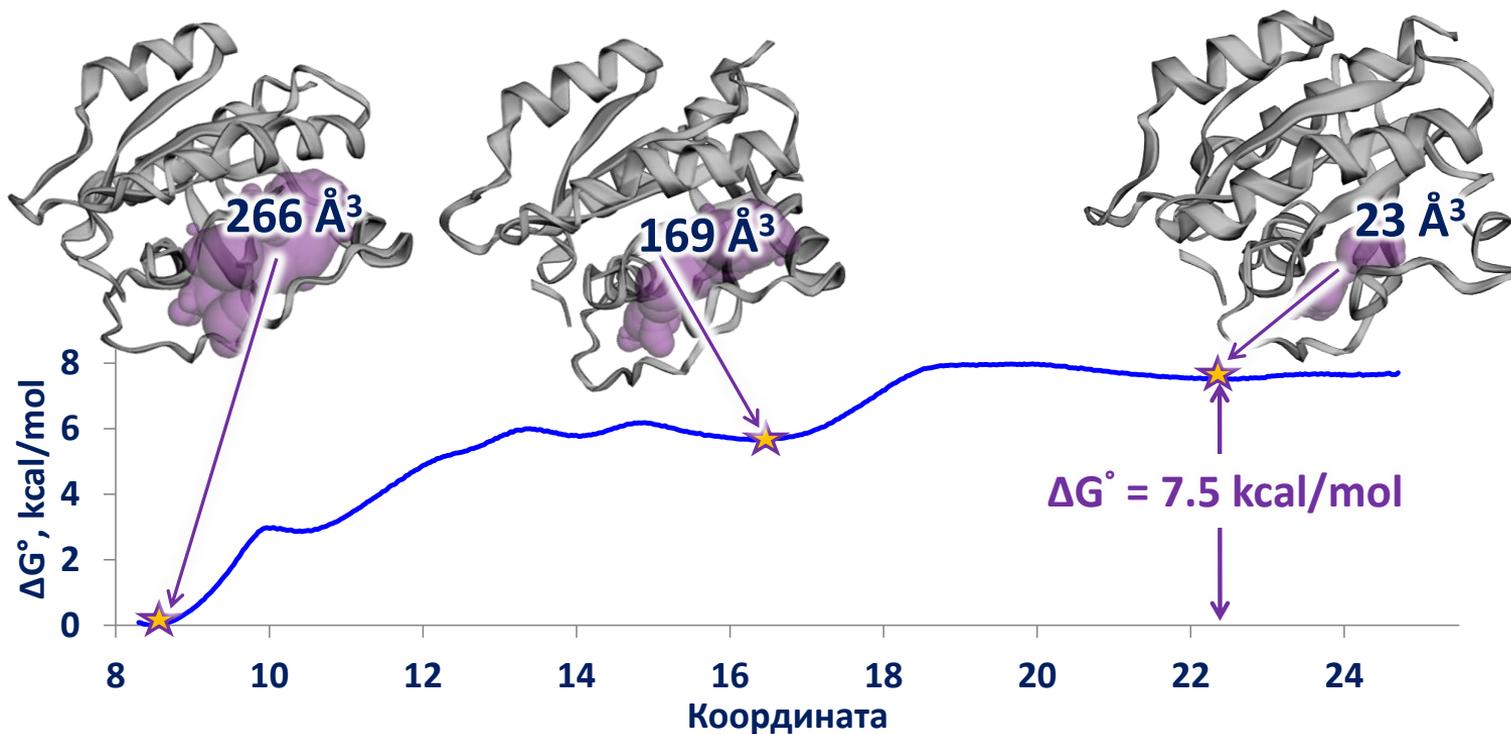


Моделирование нековалентного связывания

Расчёты методом классической молекулярной динамики (NAMD) нужны длинные траектории \sim мкс



Моделирование нековалентного связывания



Суперкомпьютеры необходимы для виртуального скрининга баз данных разрешенных лекарственных веществ, чтобы с построенными моделями белков проанализировать взаимодействия с «малыми» молекулами

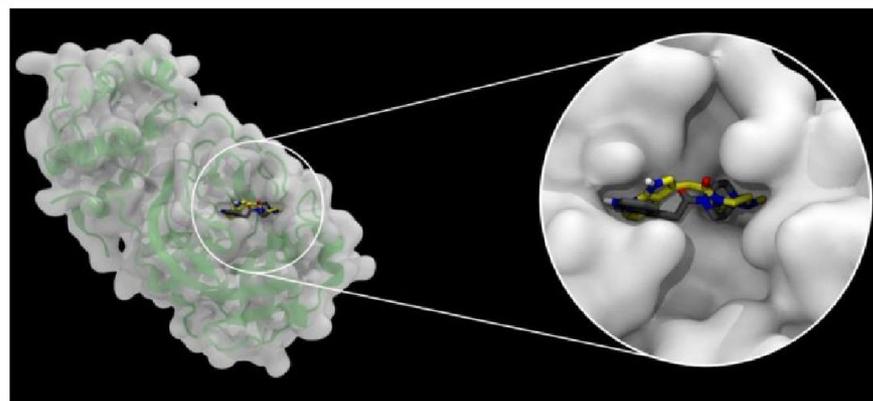
Суперкомпьютер нашел лекарство против коронавируса

 3  3  2    Добавить в «Мую Ленту»

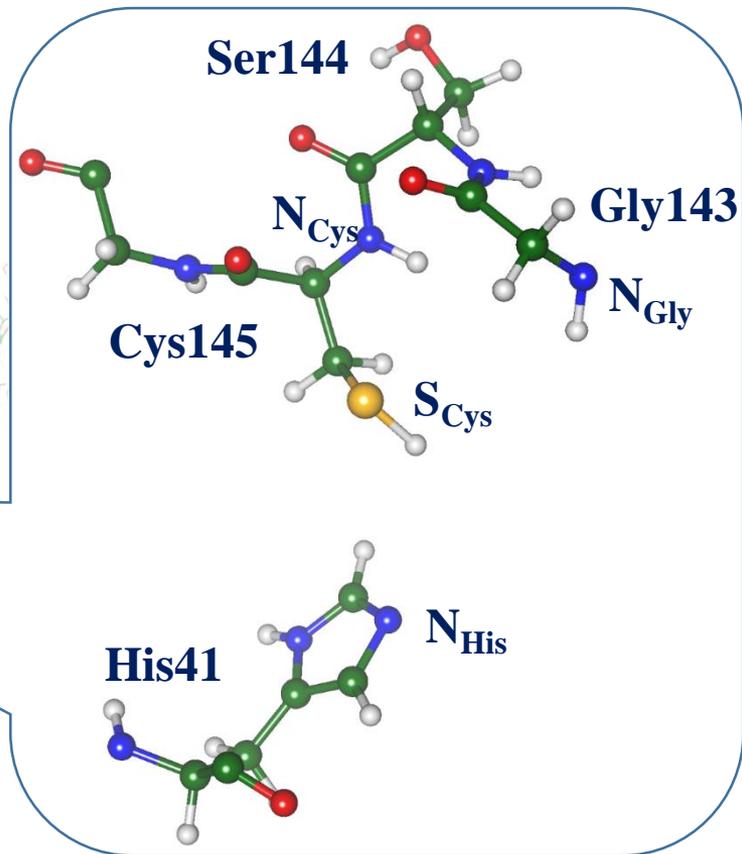
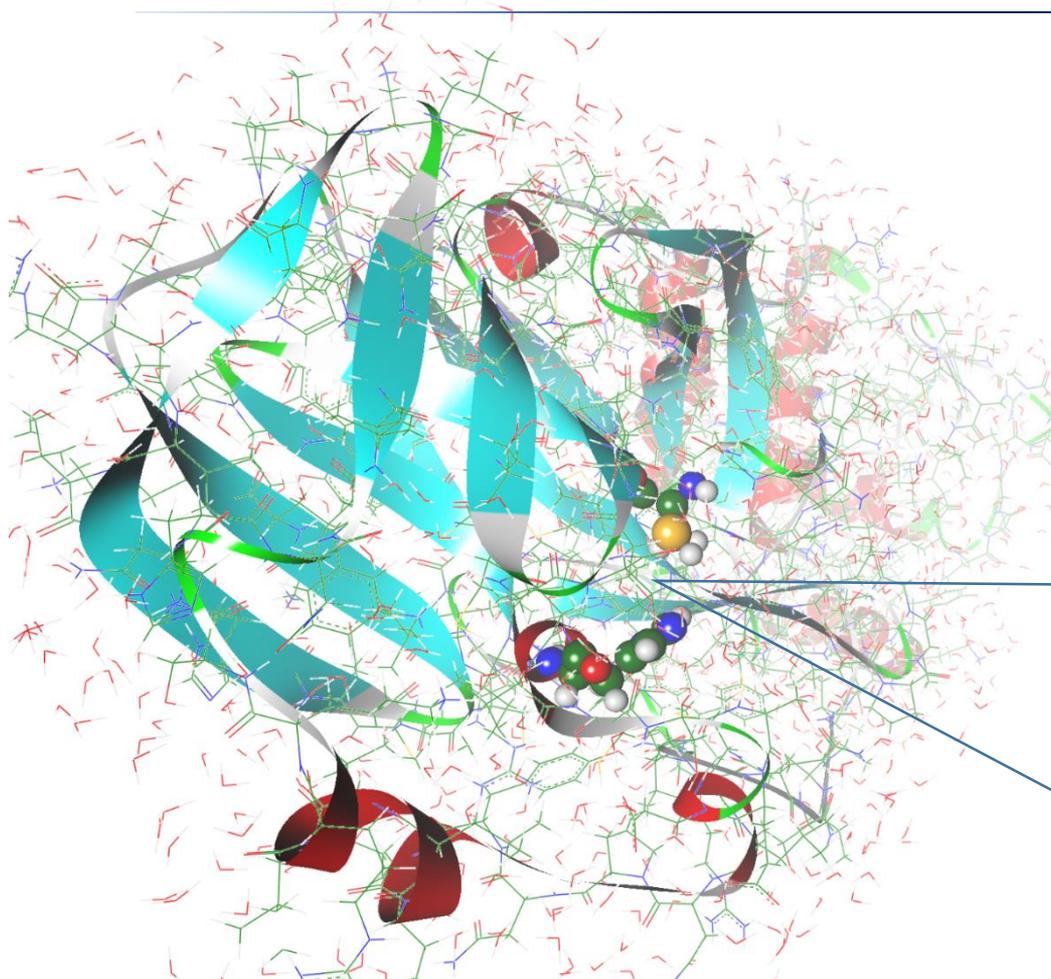


Фото: megware.com

“When we were using Summit, we were docking 20,000 compounds a second,” said Ada Sedova, a biophysicist in the Molecular Biophysics Group within ORNL’s Biosciences Division and co-lead of the project. “We have done this in 24 hours with full optimization of these poses, the way people would normally do at the small scale. To be able to do this on a billion compounds would have taken months on even the largest academic clusters without the optimizations of AutoDock-GPU for Summit.”



Ковалентное ингибирование M^{Pro} вируса SARS-CoV-2



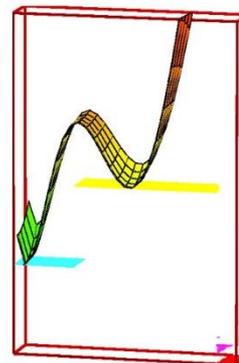
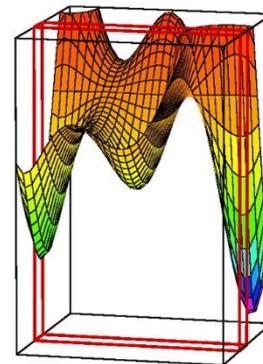
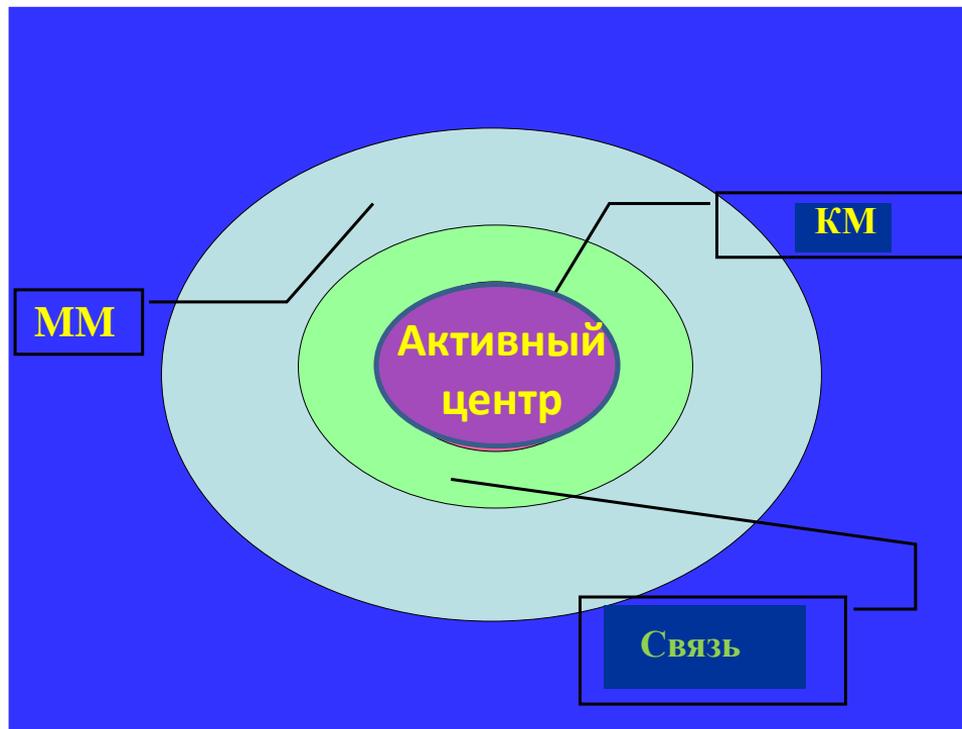
Суперкомпьютеры необходимы, чтобы рассчитать энергетический профиль реакции ковалентного ингибирования

Методы КМ/ММ

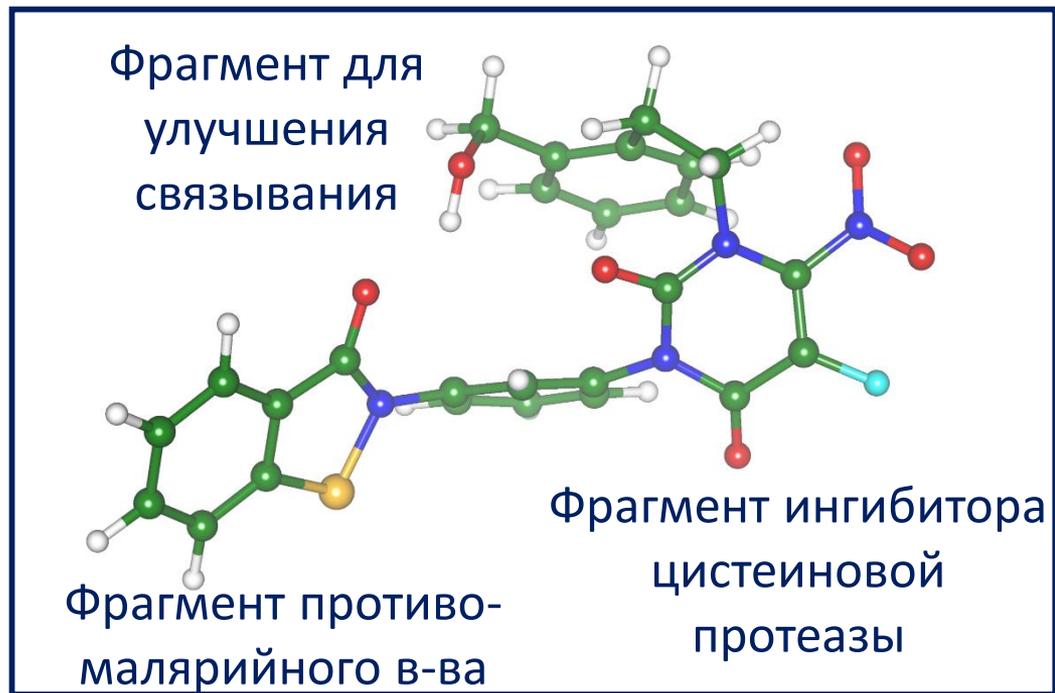
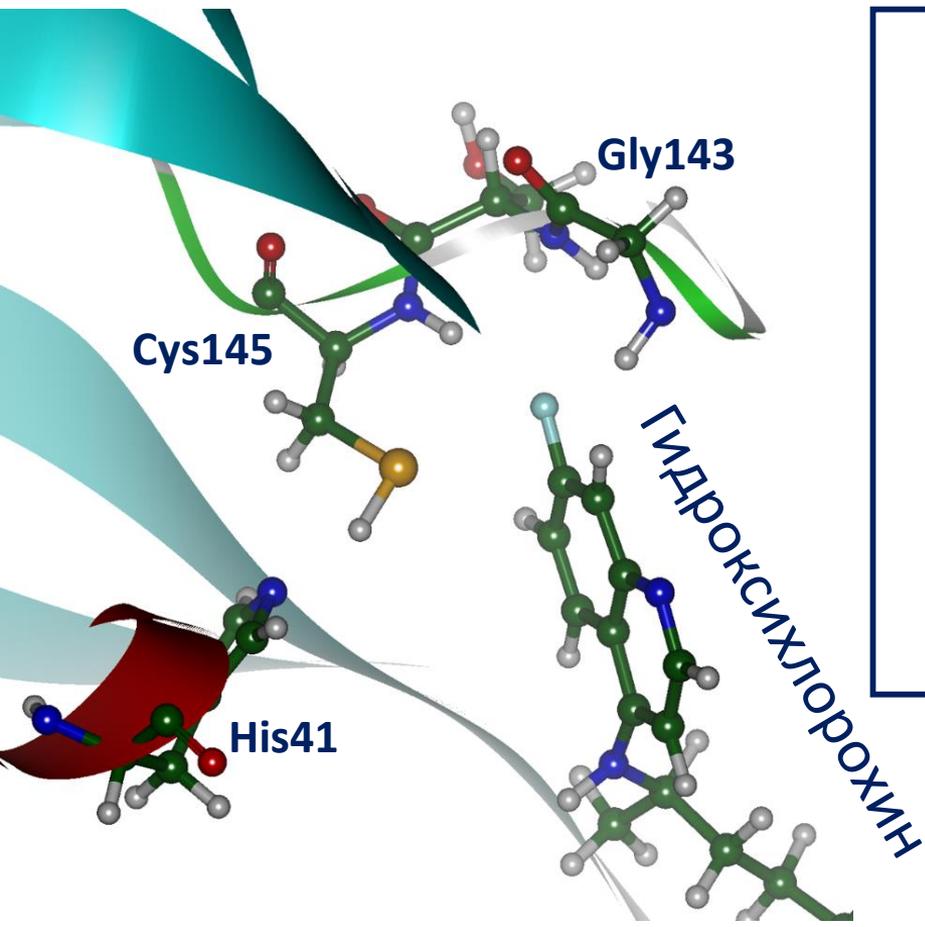
DFT в КМ,
AMBER,
CHARMM в ММ

Программы

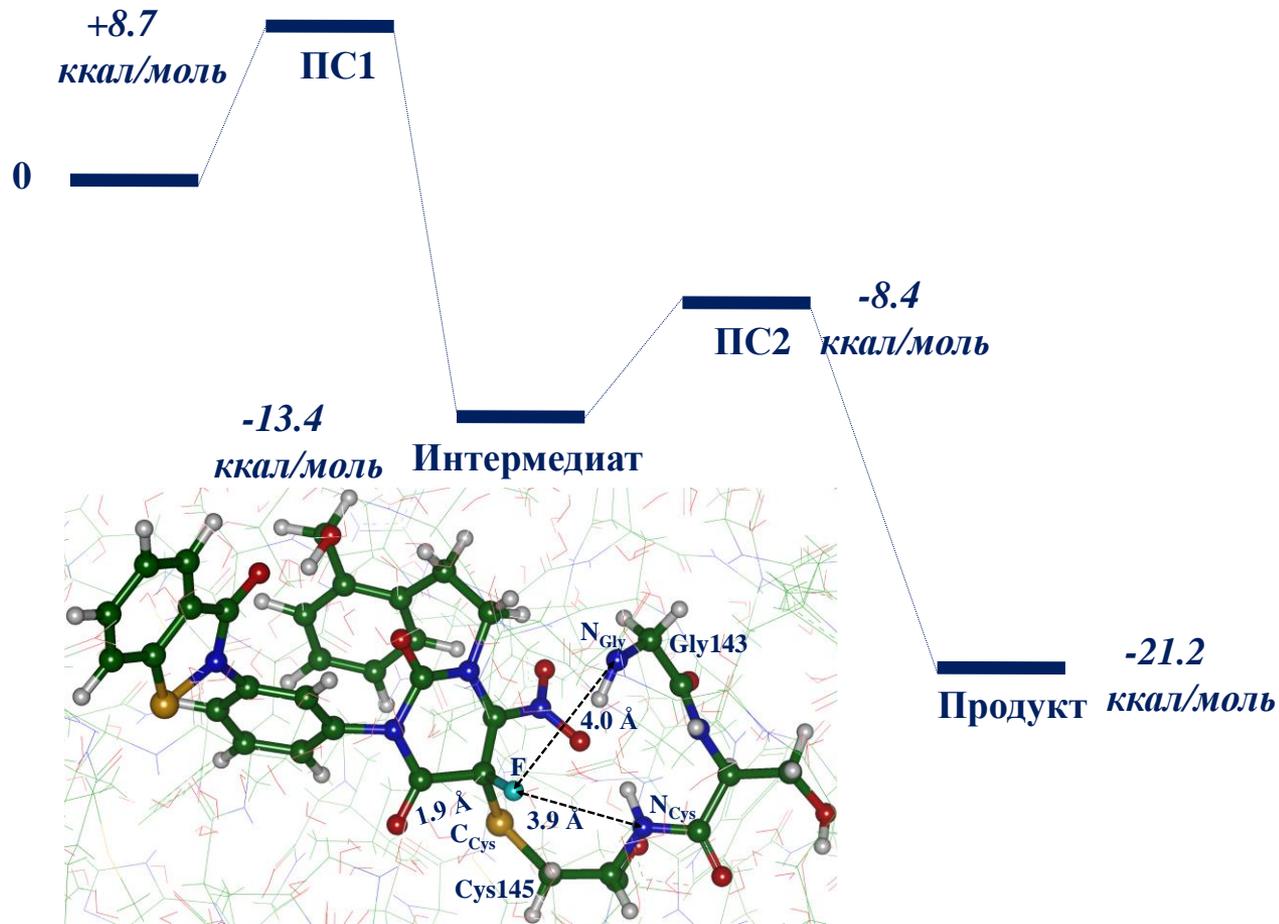
NWChem
QChem



Выбор кандидатов для блокирования M^{Pro}



Результаты расчетов: КМ(PBE0/6-31G^{*})/ММ(AMBER)



Стоимость расчетов: КМ(РВЕ0/6-31G^{*})/ММ(АМВЕР)

Программа NWChem, «Ломоносов-2»

1250 базисных функций, расчет 1 точки

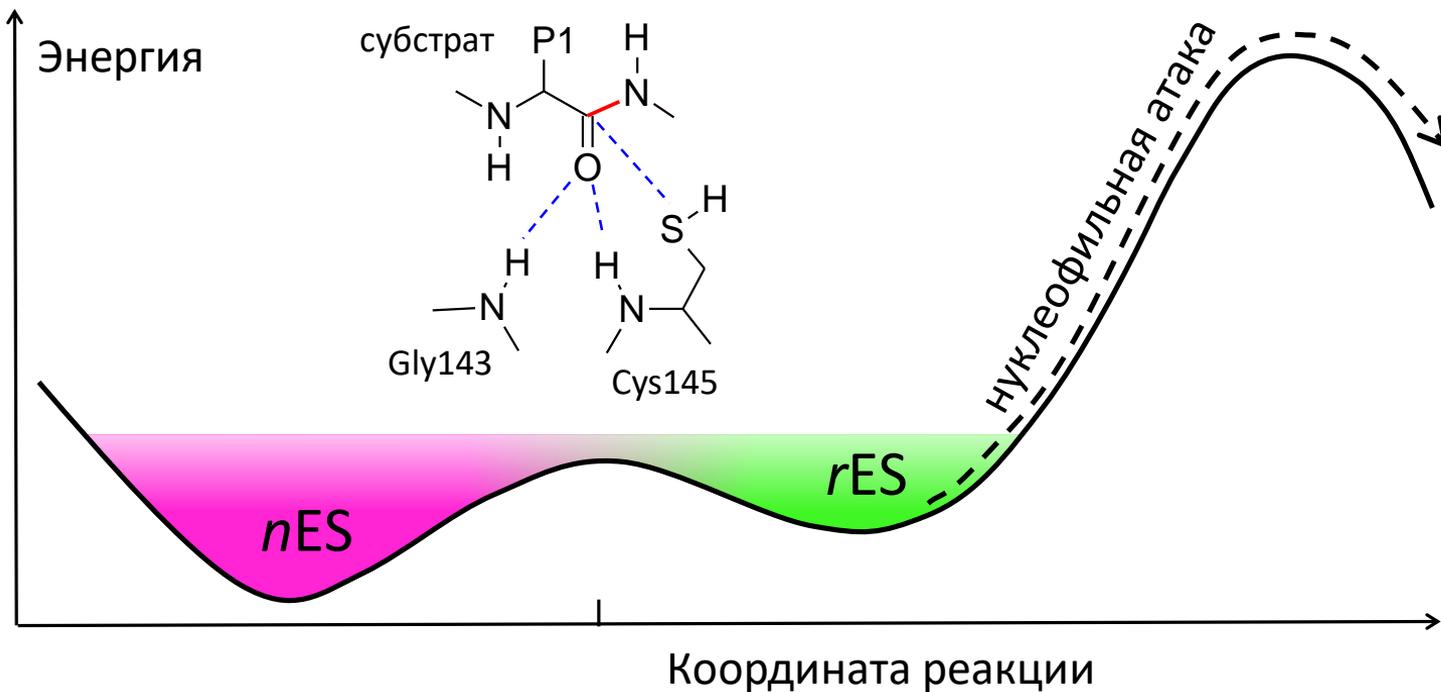
~ 50 циклов оптимизации (5 КМ, 100 ММ) ~ 1344 CPU*H

Весь профиль ~ 150000 CPU*H



Чем обусловлена субстратная специфичность?

Различная эффективность активации субстрата



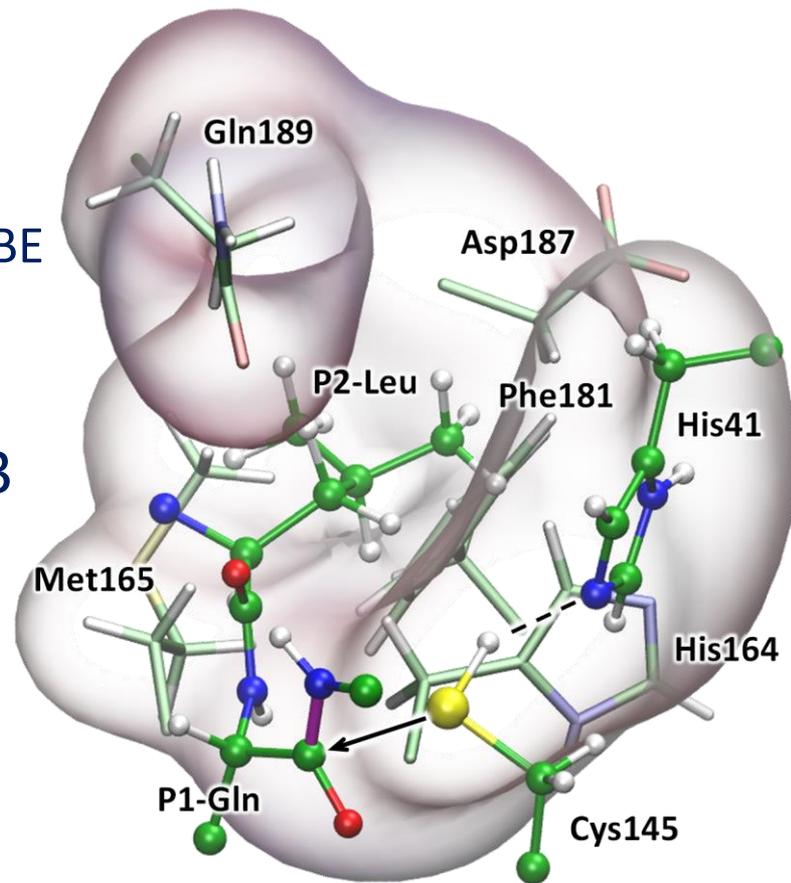
Выбор метода

✓ Методы описания КМ подсистемы

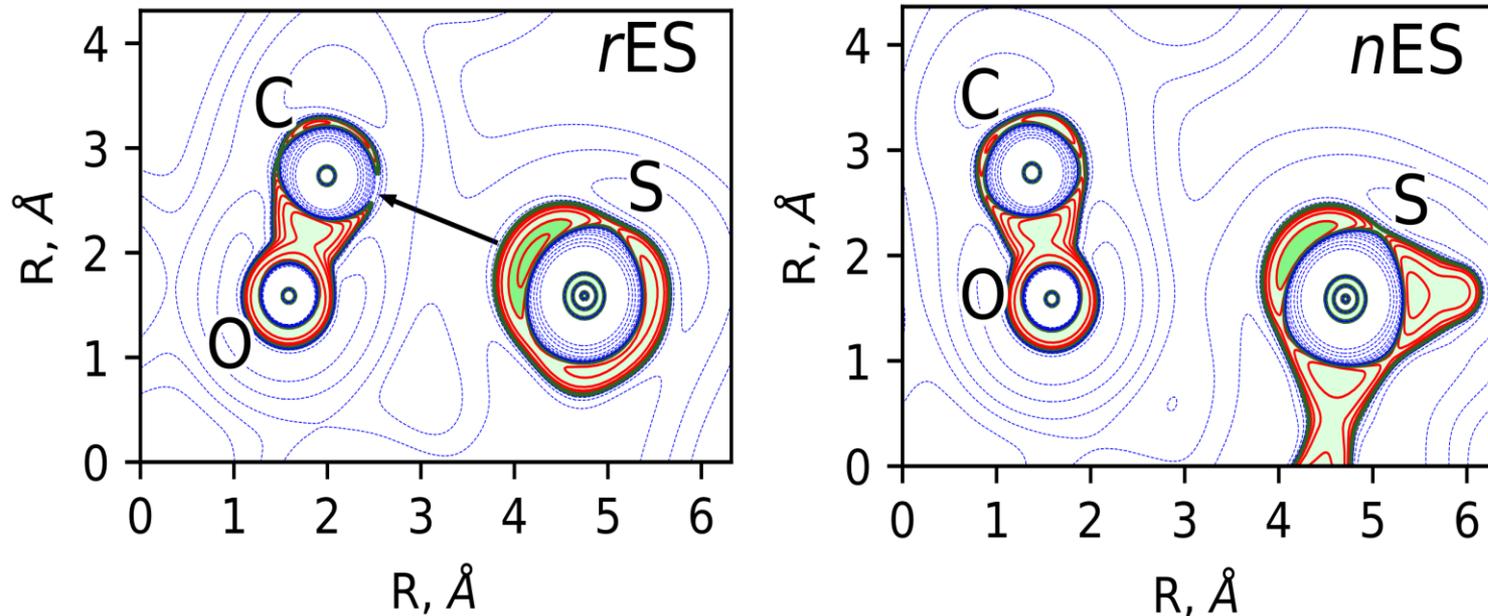
- ✓ Гибридный функционал PBE0
- ✓ Обобщенно-градиентный функционал PBE

✓ Тестирование проводилось для фермент-субстратных комплексов с 3 субстратами, различающимися а/к остатком в положении P2:

- ✓ S-P2Leu (относительная $k_{\text{cat}} = 1$)
- ✓ S-P2Ile (относительная $k_{\text{cat}} = 0.45$)
- ✓ S-P2Ala (относительная $k_{\text{cat}} < 0.1$)



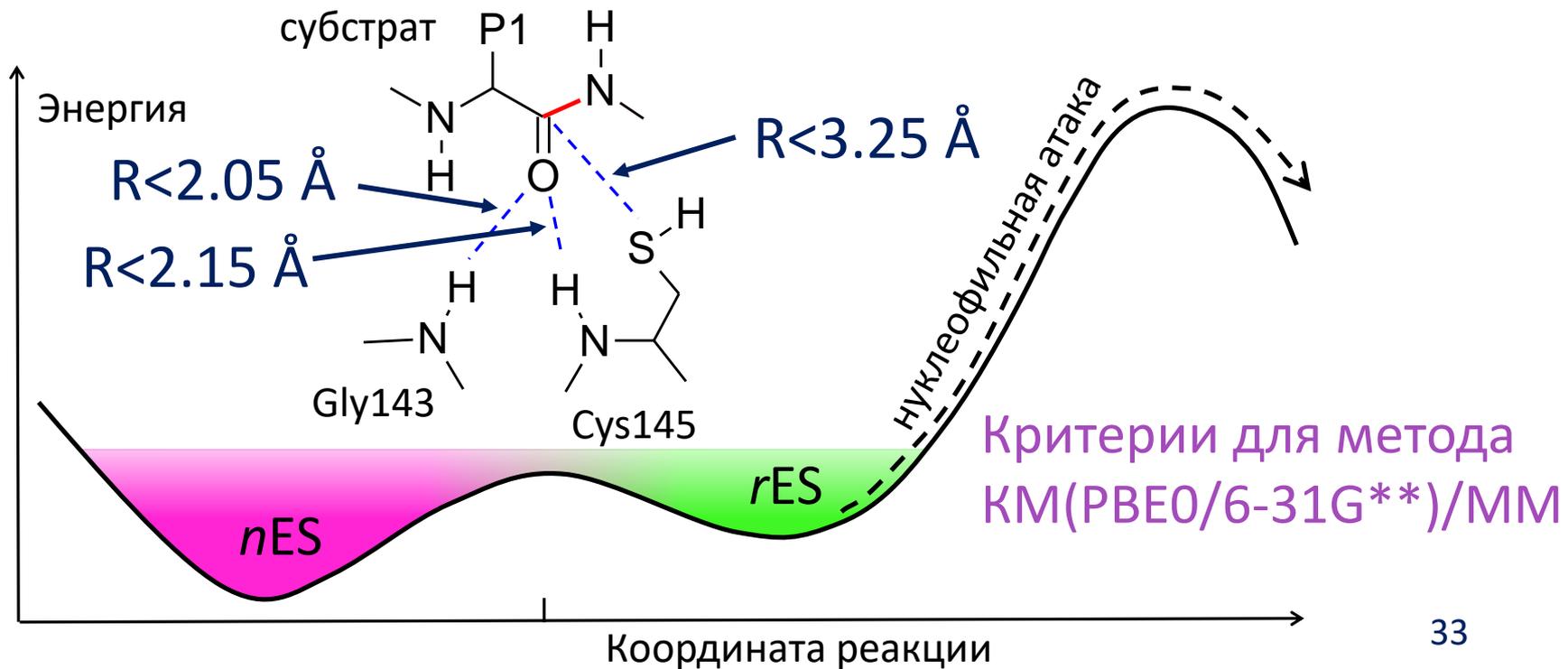
Критерии отбора реакционных конформаций



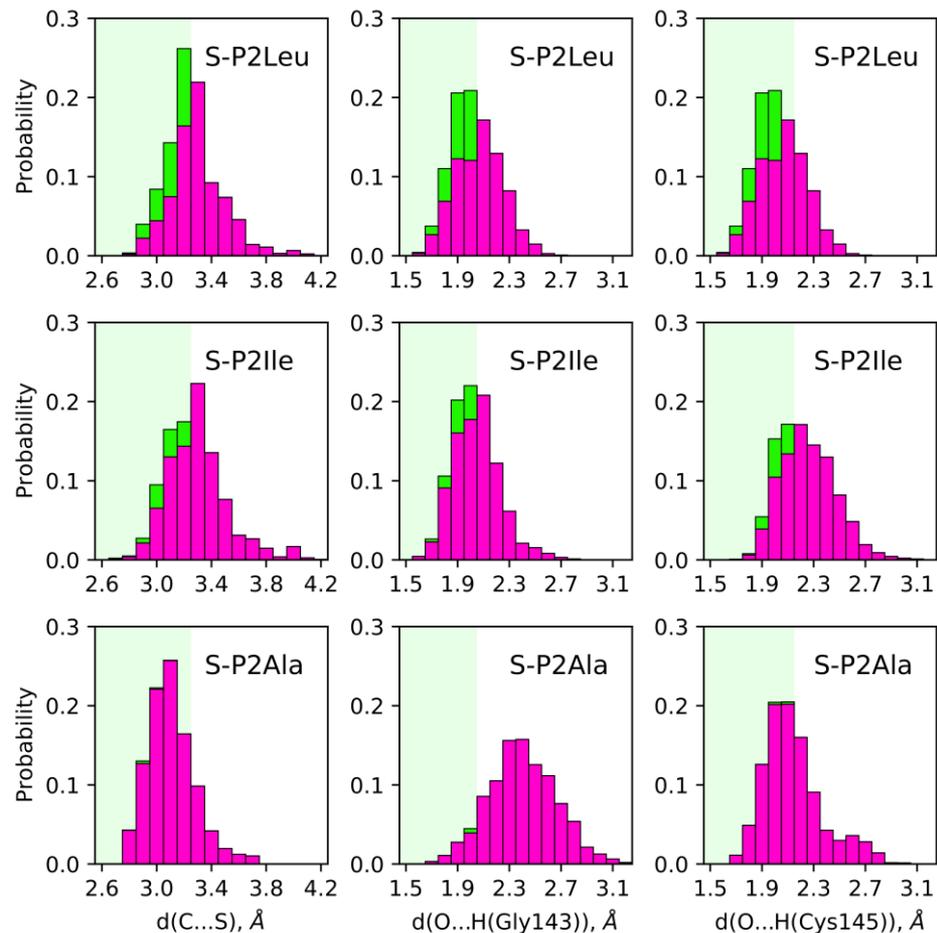
Карты Лапласиана электронной плотности ($\nabla^2\rho(\mathbf{r})$) в плоскости $S\dots C=O$:
Синие изолинии соответствуют областям деконцентрации ЭП, $\nabla^2\rho(\mathbf{r}) > 0$
Красные изолинии соответствуют областям концентрации ЭП, $\nabla^2\rho(\mathbf{r}) < 0$

Критерии отбора реакционных конформаций

Кадр траектории относится к реакционной конформации, если выполняются все три геометрических критерия одновременно



Результаты определения субстратной специфичности



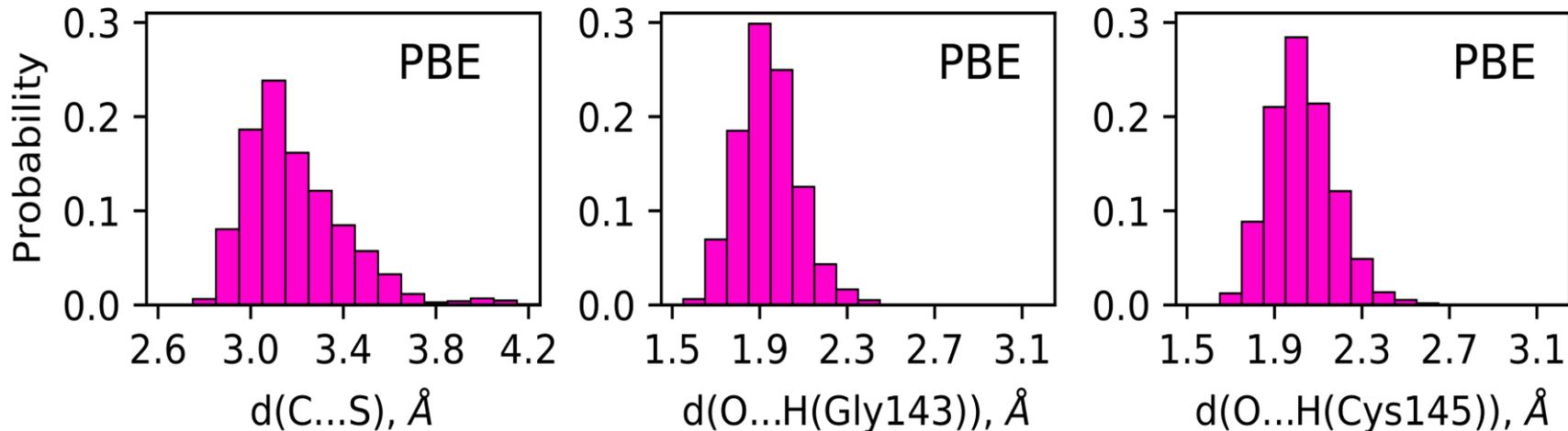
Субстрат	χ , %	k_{cat} (расч.)	k_{cat} (эксп.)
S-P2Leu	22.4	1	1
S-P2Ile	10.2	0.46	0.45
S-P2Ala	0.6	0.03	<0.1

* Значения относительно субстрата S-P2Leu

Результаты, полученные методом
КМ(PBE0/6-31G**)/ММ

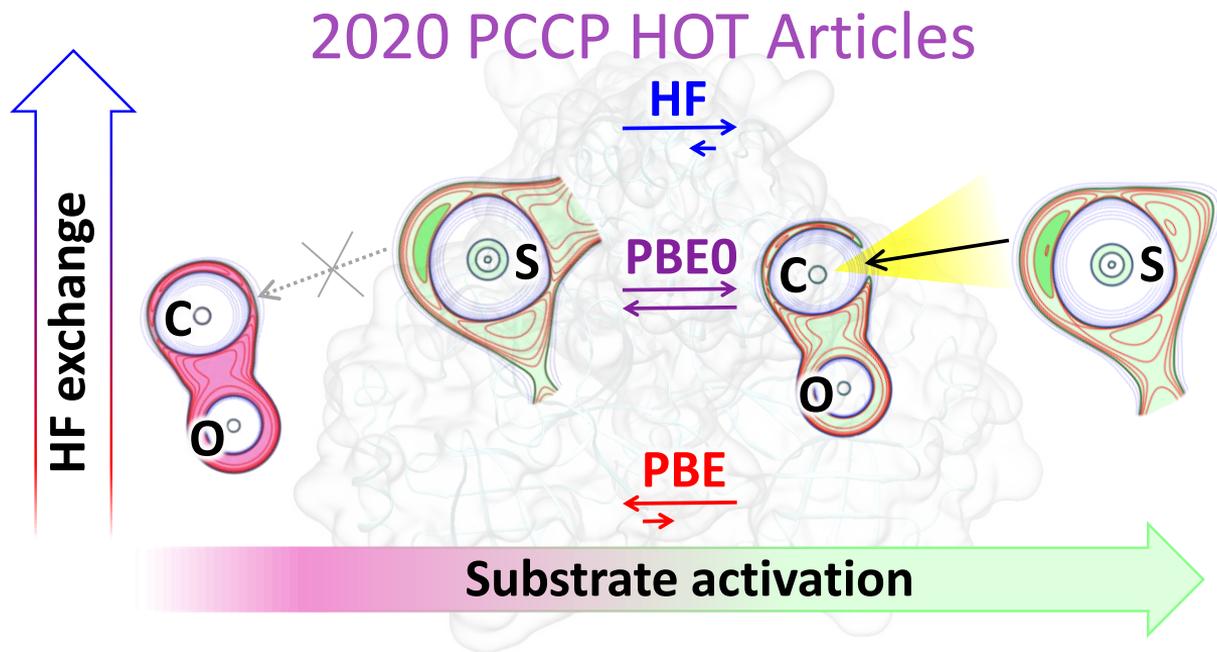
О важности выбора адекватного метода

- ✓ Реакционные конформации отсутствуют при расчёте с обобщенно-градиентным функционалом.



Результаты, полученные методом КМ(PBE/6-31G^{**})/ММ для системы с субстратом S-P2Leu

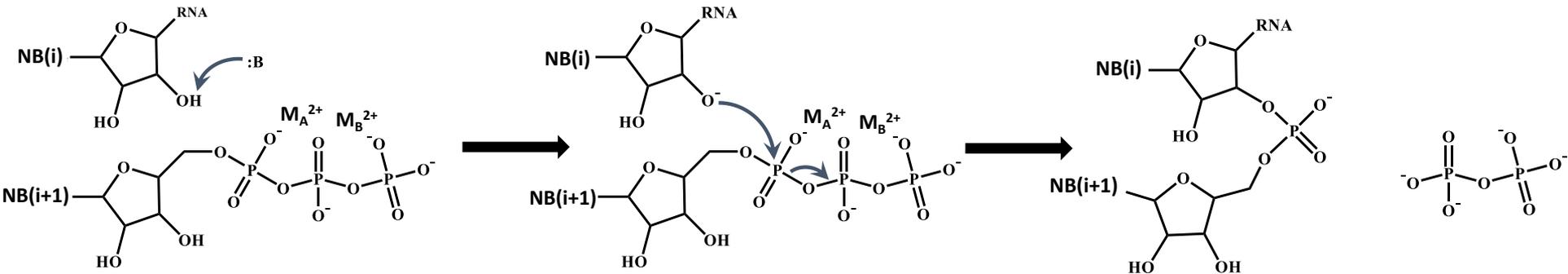
Предложен метод распознавания реакционных конформаций в ферментативных реакциях нуклеофильного присоединения, позволяющий интерпретировать наблюдаемую субстратную специфичность основной протеазы SARS-CoV-2



M.G. Khrenova, V.G. Tsirelson, A.V. Nemukhin, Physical Chemistry Chemical Physics (PCCP) 22,19069 (2020)

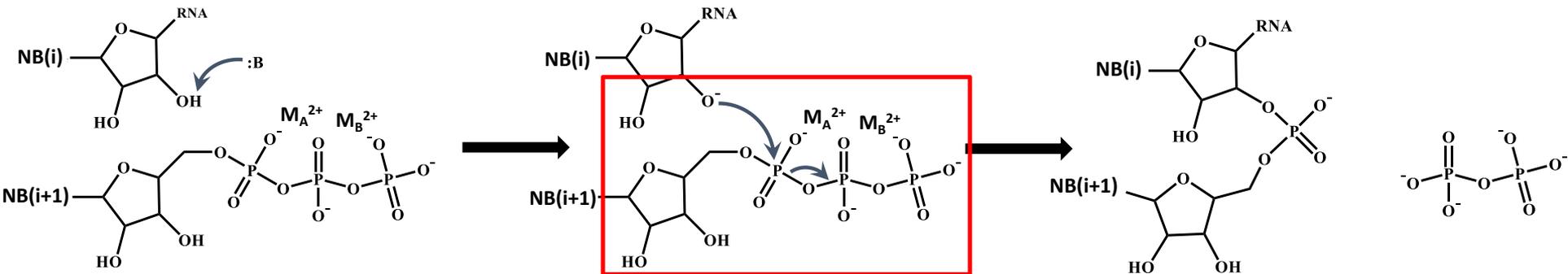
РНК полимераза: Первые шаги при моделировании

Химическая реакция, катализируемая ферментом РНК-полимеразой

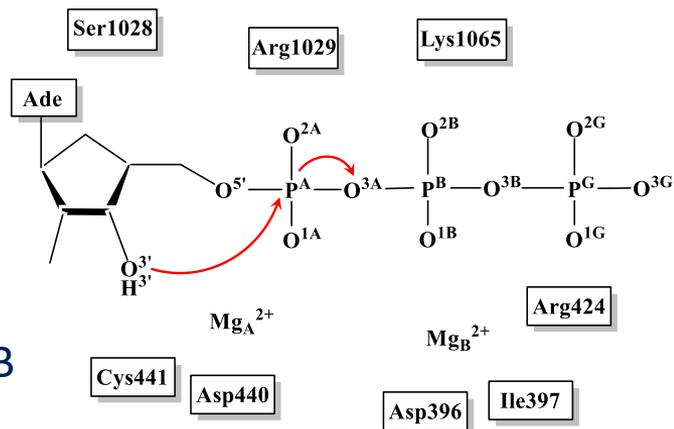


РНК полимераза: Первые шаги при моделировании

Химическая реакция, катализируемая ферментом РНК-полимеразой



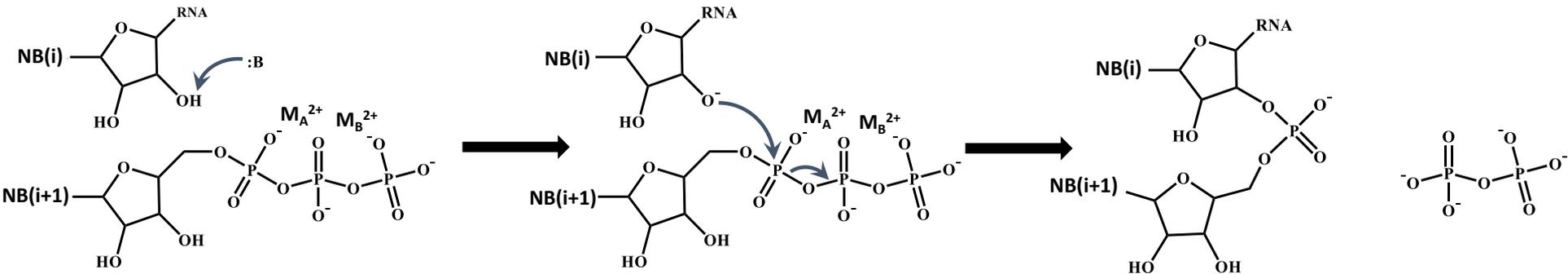
Наш предшествующий опыт
с другими ферментами
(аденилат-циклазы)



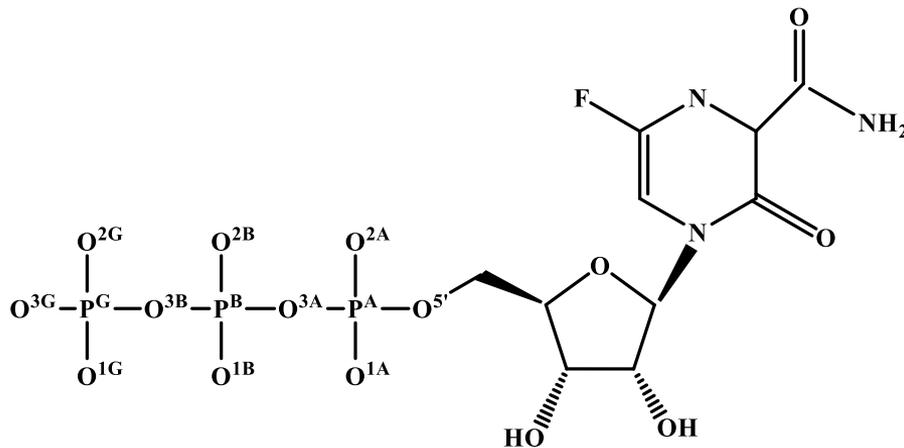
V.L. Grigorenko, I.V. Polyakov,
A.V. Nemukhin, J. Physical Chemistry B
124, 451 (2020)

РНК полимераза: Первые шаги при моделировании

Химическая реакция, катализируемая ферментом РНК-полимеразой

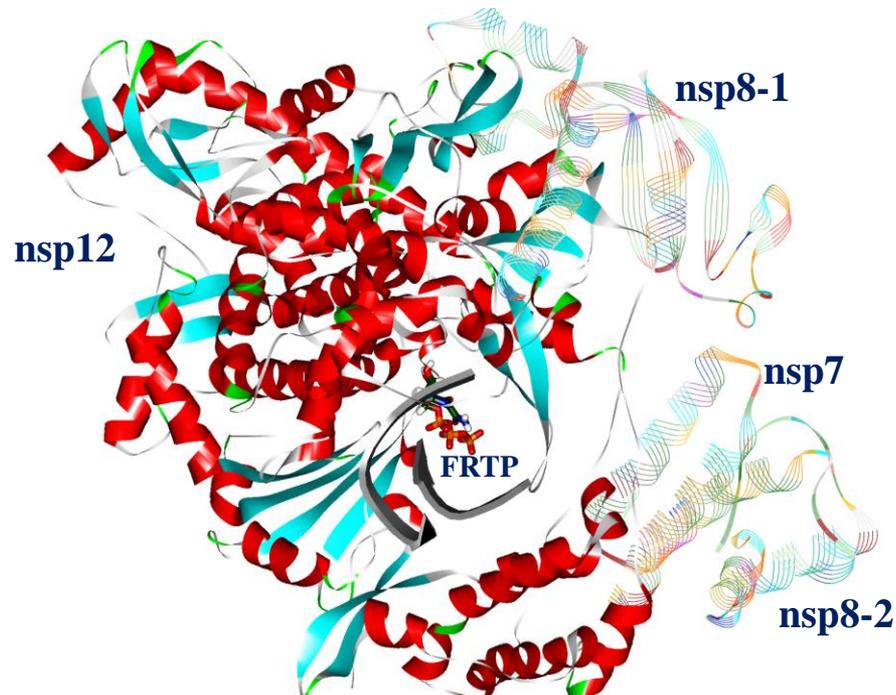
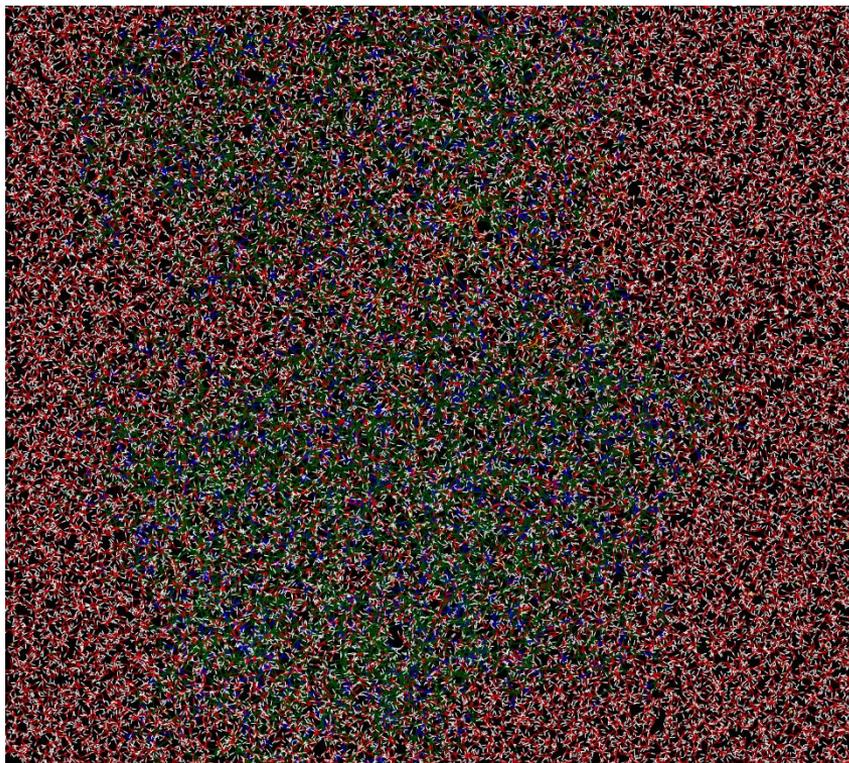


Причем здесь фавипиравир?



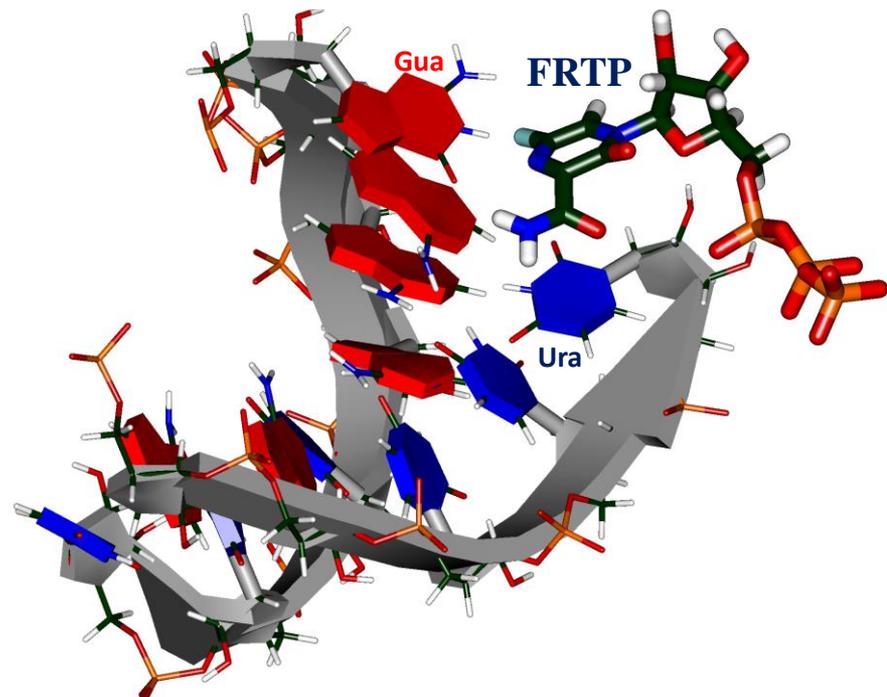
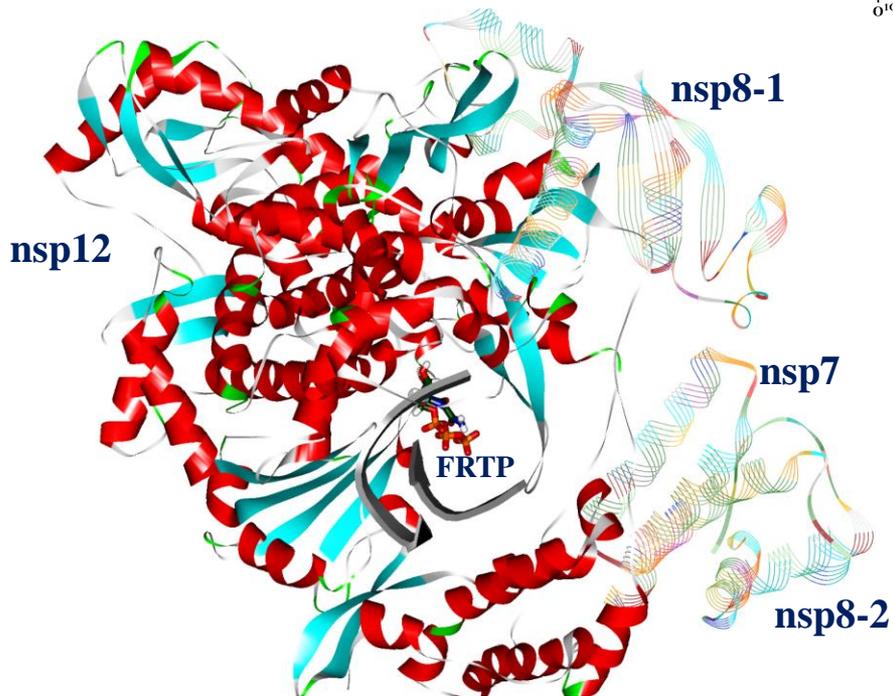
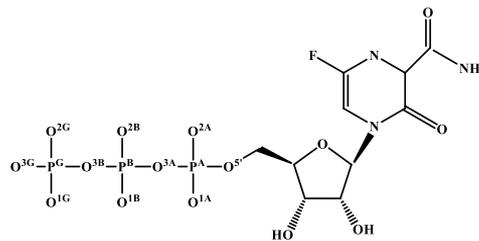
РНК-полимераза: Первые шаги при моделировании

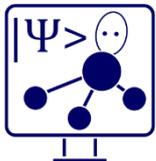
Молекулярные модели – нереально анализировать без суперкомпьютеров



РНК полимераза: Первые шаги при моделировании

Что делает фавипиравир





Базы данных

Интермедиаты реакций ферментативного катализа

[Main](#) [About Database](#) [Help](#)

Intermediates of reactions in proteins active sites

[Simple](#) [Advanced](#) [History](#)

Reaction Search

Match any fragment of the reaction name



Name, abbreviation, PDBID 

[About the project](#)

The project is aimed at solving the problem of integrating an array of experimental and theoretical data on the intermediates of chemical reactions in proteins, estimation of the quality of data, and filling databases with new results.

[Search by name](#)

- Reaction name
- Abbreviation
- PDBID

Благодарности:

- соавторам доклада

(Григоренко Б.Л., Луцкекина С.В., Поляков И.В., Хренова М.Г.)

- коллегам

(Книга А.Е., Кулакова А.М., Миронов В.А., Московский А.А.,
Цирельсон В.Г. (РХТУ))

- Центру коллективного пользования

сверхвысокопроизводительными вычислительными
ресурсами МГУ

- РФФИ

Спасибо за внимание!