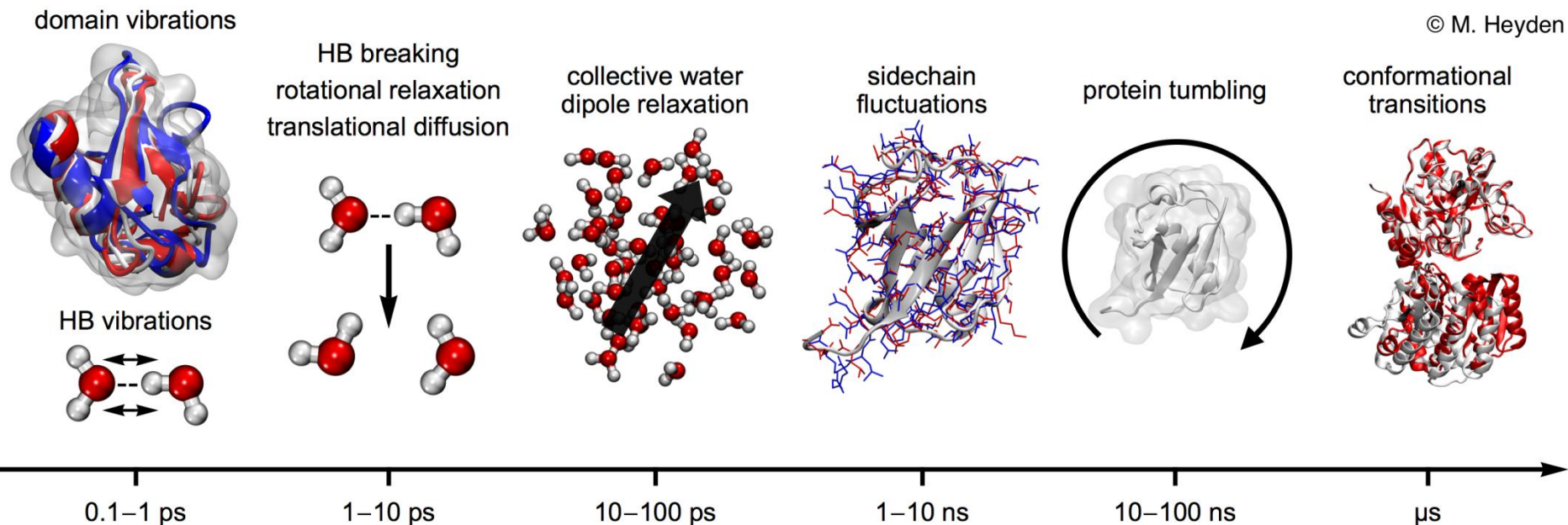


Молекулярная динамика в силовом поле FF14SB в воде TIP4P-Ew, и в силовом поле FF15IPQ в воде SPC/Eb: сравнительный анализ на GPU и CPU

Д.А. Суплатов, Я.А. Шарапова, Н.Н. Попова,
К.Е. Копылов, Вл.В. Воеводин, В.К. Швядас

НИИ ФХБ имени А.Н.Белозерского,
Факультет биоинженерии и биоинформатики,
Факультет вычислительной математики и кибернетики,
НИВЦ МГУ имени М.В. Ломоносова

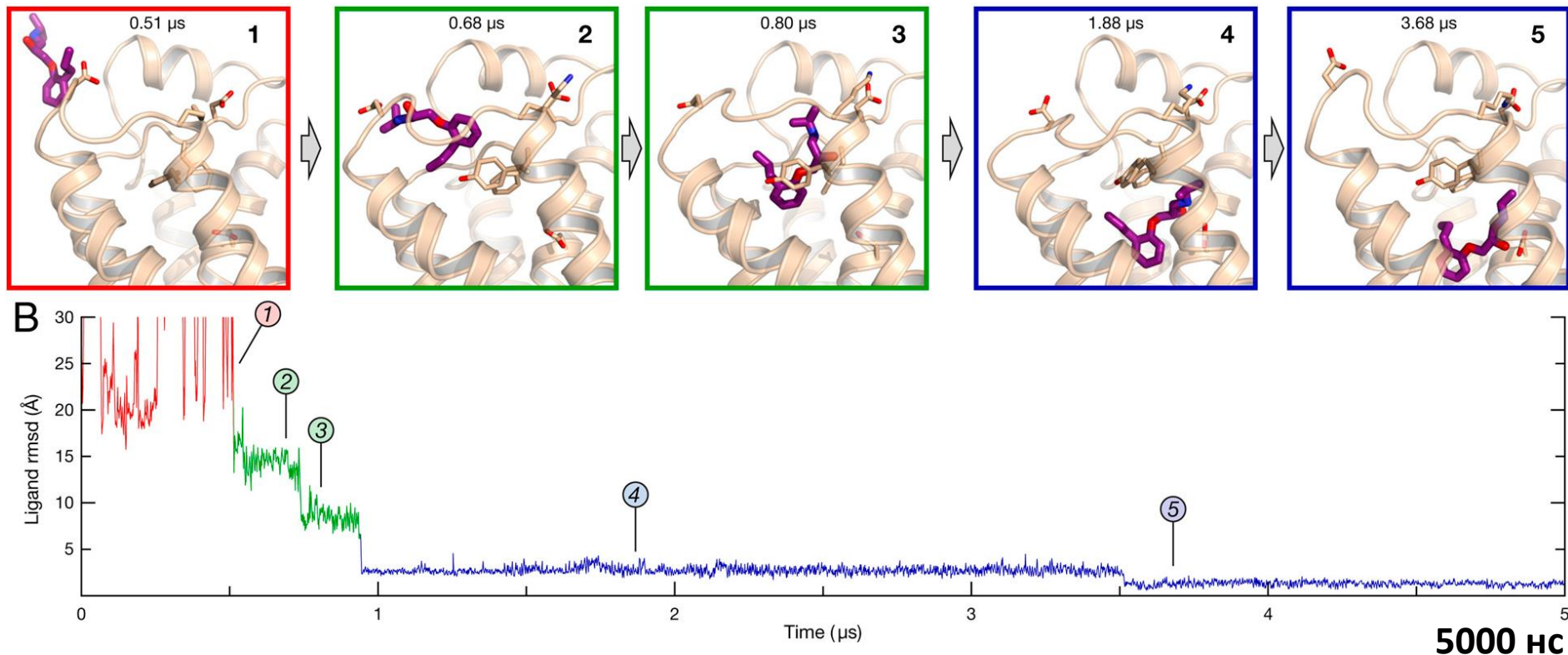
Молекулярная динамика: метод для изучения структуры белка в движении



<https://isearch.asu.edu/profile/3160407>

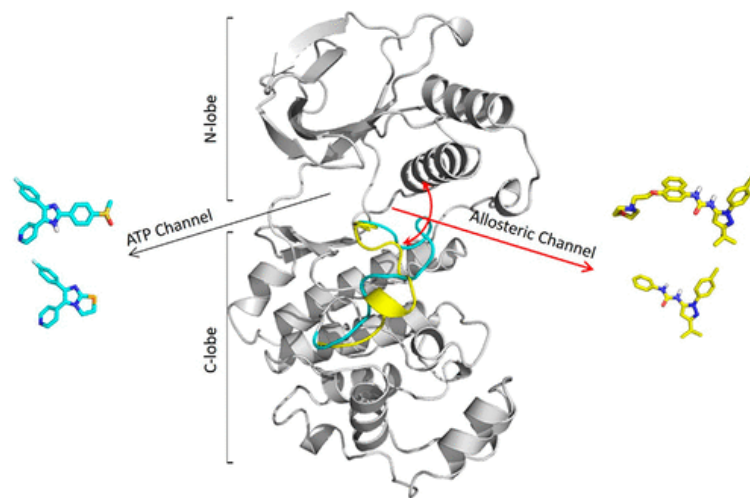
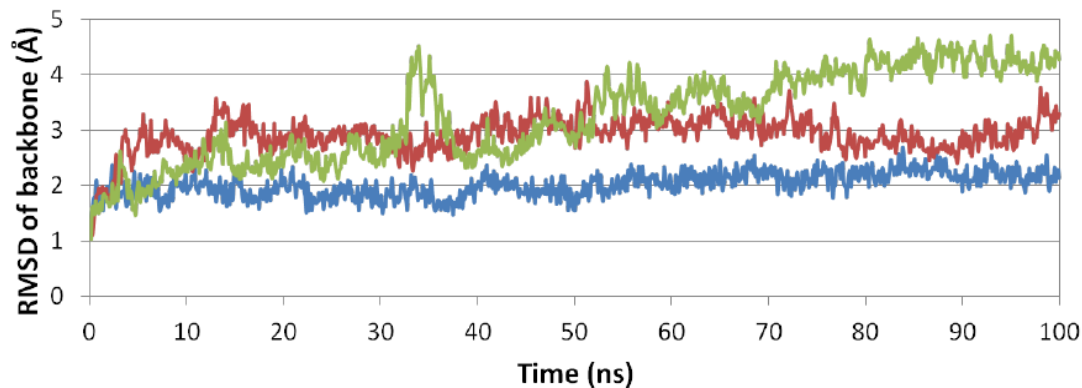
Для изучения более сложных конформационных изменений в структуре белка необходимо рассчитывать более длинные траектории молекулярной динамики

Примеры работы с использованием молекулярной динамики (2011 год)



Dror R. O. et al. Pathway and mechanism of drug binding to G-protein-coupled receptors //Proceedings of the National Academy of Sciences. – 2011. – Т. 108. – №. 32. – С. 13118-13123.

Примеры работы с использованием молекулярной динамики (2018 год)



You W., Chang C. A. Role of Molecular Interactions and Protein Rearrangement in the Dissociation Kinetics of p38 α MAP Kinase Type-I/II/III Inhibitors //Journal of chemical information and modeling . – **2018**. – T. 58. – №. 5. – С. 968–981.

Длина траекторий молекулярной динамики

2011	2018
5 000 нс	100 нс
?	?

Длина траекторий молекулярной динамики

2011	2018
5 000 нс	100 нс
Особое оборудование	Длина траектории в повседневной лабораторной практике



Специализированные суперкомпьютеры Anton/Anton-2 для молекулярной динамики



(a)



(b)



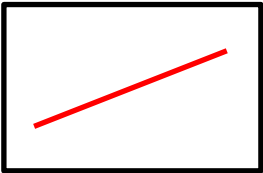
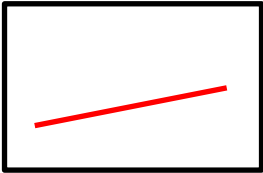
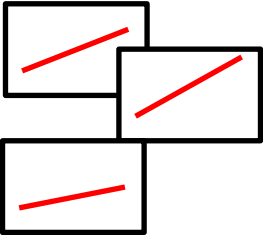
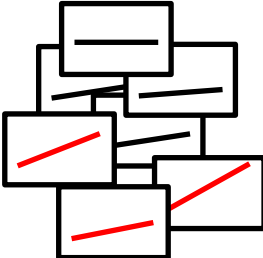


(c)

An Anton 2 ASIC. (b) An Anton 2 node board, with one ASIC underneath a phase-change heat sink. (c) A 512-node Anton 2 machine, with

- Продукт частной биотехнологической компании «DE Shaw Research»;
- Результат глубокой переработки и ко-дизайна программного обеспечения и аппаратных средств для его исполнения;
- Предназначен для исполнения только специального программного обеспечения;
- Очень быстрое, но редкое (штучное) оборудование, доступное только в США


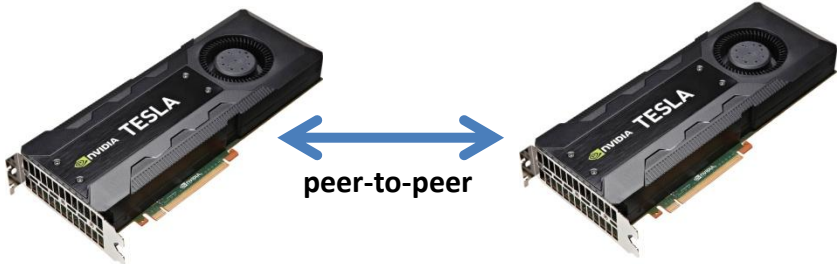
Shaw D. E. et al. Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer //Proceedings of the international conference for high performance computing, networking, storage and analysis. – IEEE Press, 2014. – С. 41-53.

Объем вычислений молекулярной динамики в повседневной лабораторной практике

До 2000	2000 – 2008	2008 – 2016	2016 – 2017	2018
< 1 нс	10 нс	10 – 100 нс	100 нс	?
				
CPU	CPU	CPU Чебышев Ломоносов-1	Ломоносов-2 (Tesla K40) 	

Постановка задачи

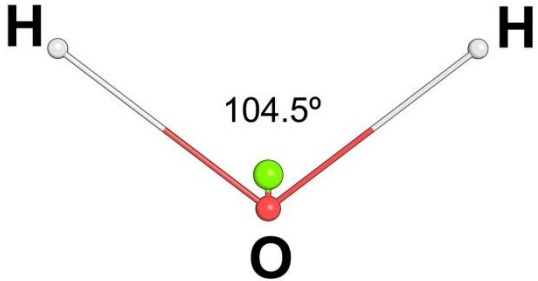
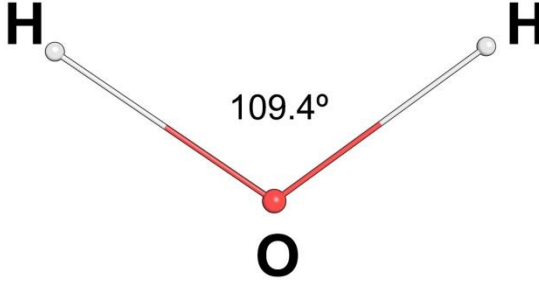
Доступное оборудование

Ломоносов-2 «Compute» (2016)	Ломоносов-2 «Pascal» (март 2018)
<p data-bbox="112 601 683 729">Intel Xeon E5-2697 v3 1 × Tesla K40</p> 	<p data-bbox="981 601 1534 729">Intel Xeon Gold 6126 2 × Tesla P100</p> 

Постановка задачи

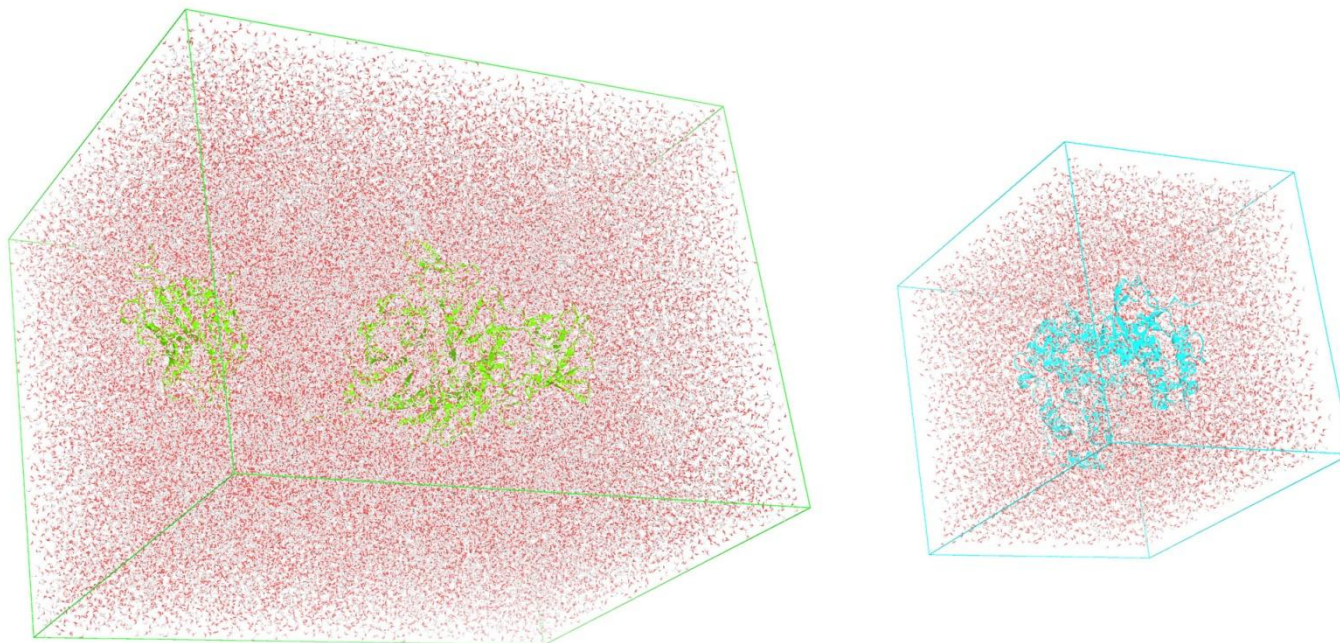
Силовые поля и модели воды

AMBER FF14SB	AMBER FF15IPQ
«Классическое» силовое поле, основное рекомендуемое силовое поле для молекулярной динамики белков в AMBER	Новое «улучшенное» силовое поле, предназначенное для длительных молекулярно-динамических симуляций [Ahmed & Papaleo, 2018; Koes & Vries, 2017; Debiec et al., 2016]

TIP4P-Ew	SPC/Eb
	

Постановка задачи

Макромолекулярные биологические системы

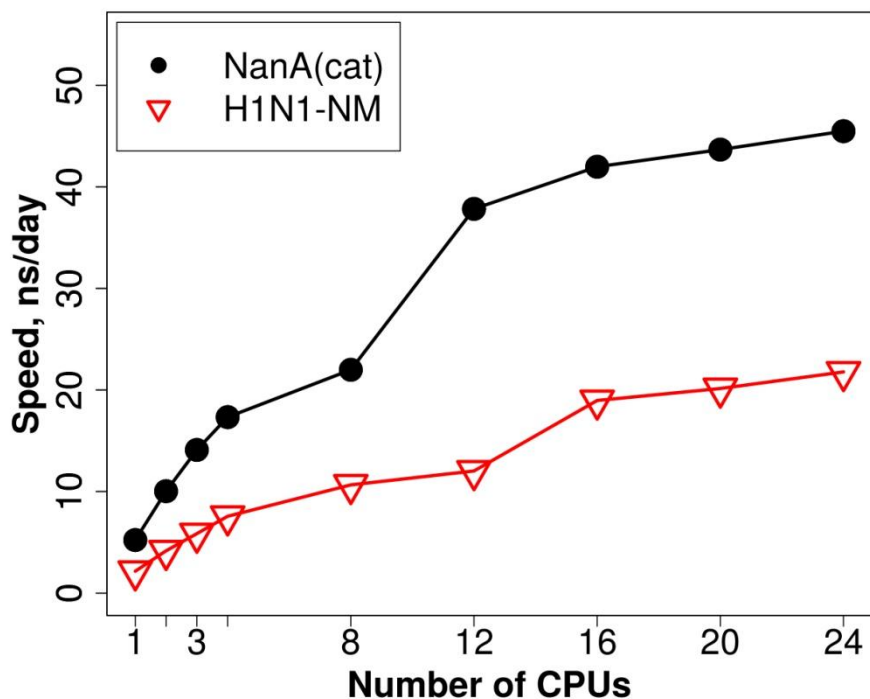


Использовались системы от **65 000 до 350 000 атомов** на основе пяти белков разных размеров, представляющих интерес как мишеней для дизайна новых лекарств от заболеваний человека:

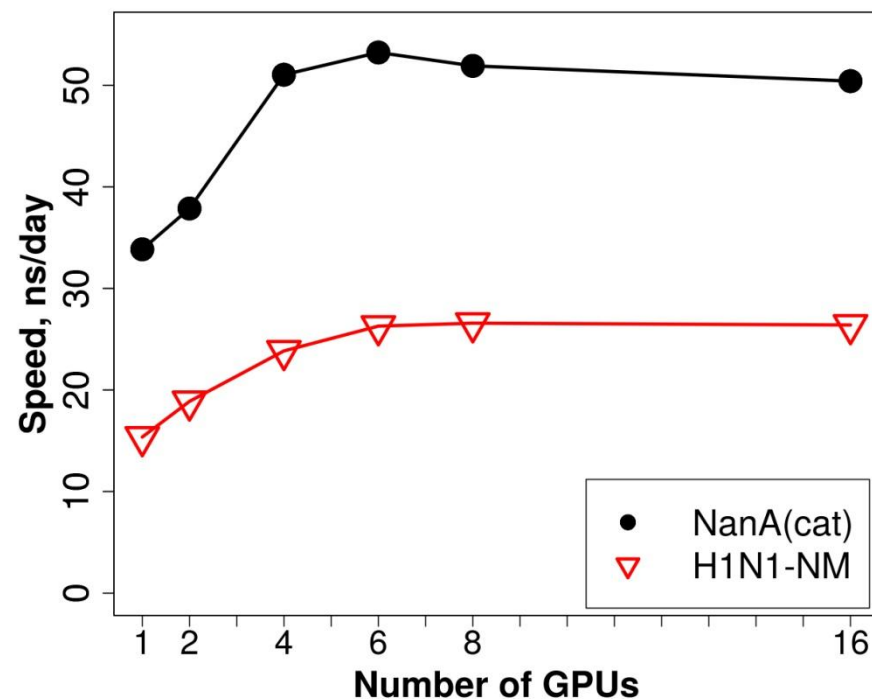
- **MAPK** — митоген-активируемая протеинкиназа человека;
- **NanC** — нейраминидаза C из *Streptococcus pneumoniae*;
- **LDH** — лактатдегидрогеназа из *Staphylococcus aureus*;
- **H1N1-NM** — нейраминидаза вируса гриппа H1N1;
- **NanA** — нейраминидаза A из *Streptococcus pneumoniae*.

Молекулярная динамика на CPU vs GPU

Intel Xeon Gold 6126



Tesla K40

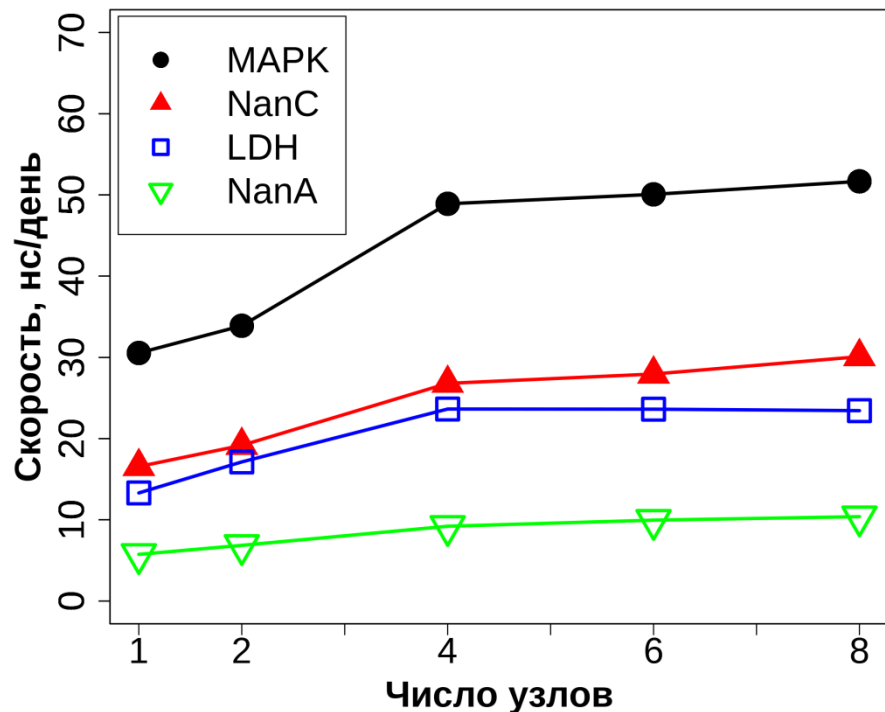


Ускорение МД на CPU возможно, но не эффективно

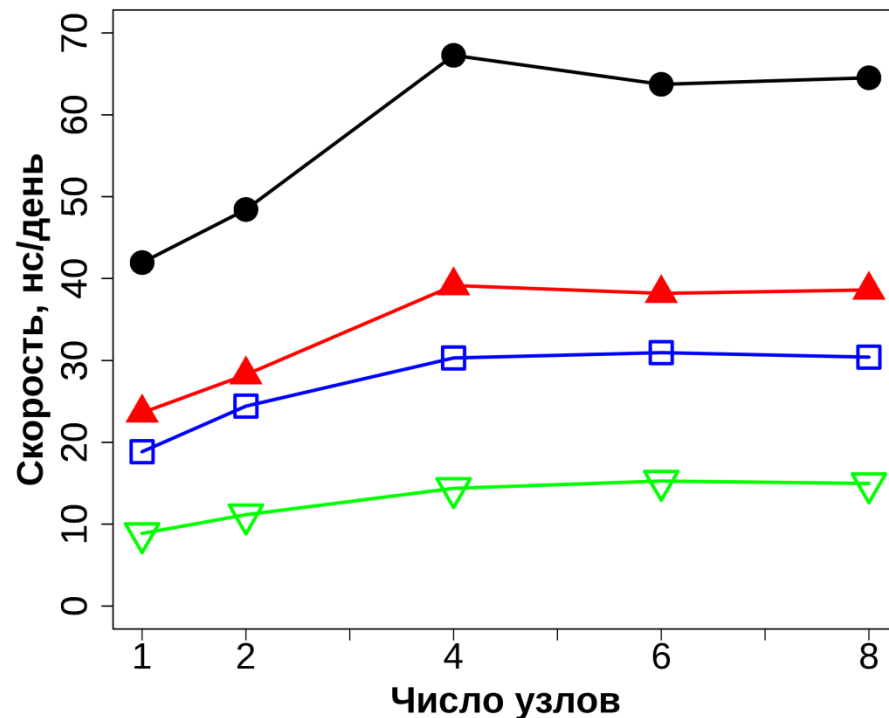
Молекулярная динамика на Tesla K40 GPU

биологических систем разных размеров и плотности

FF14SB + TIP4P-Ew



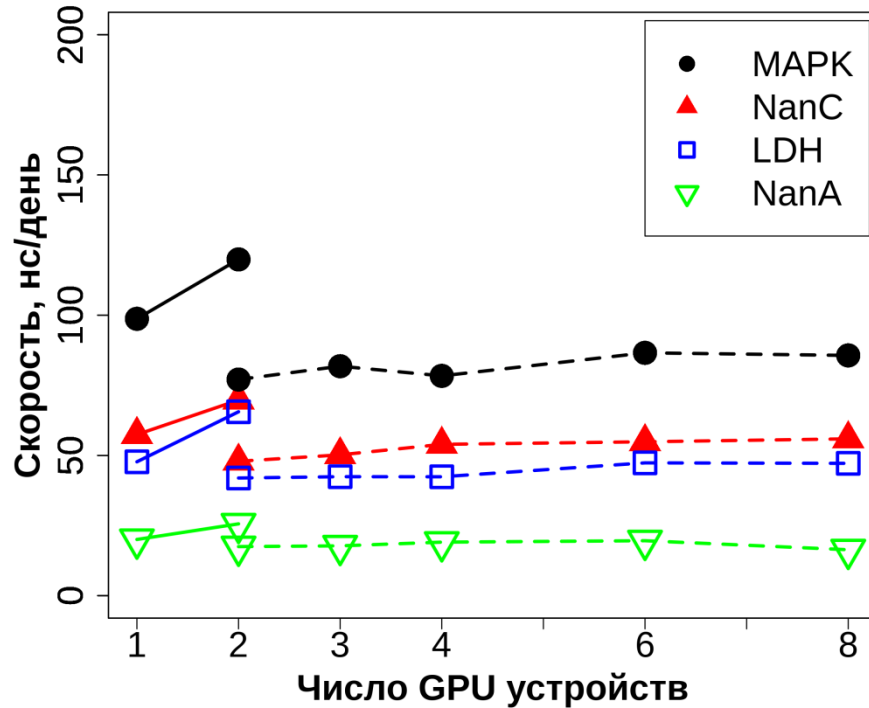
FF15IPQ + SPC/Eb



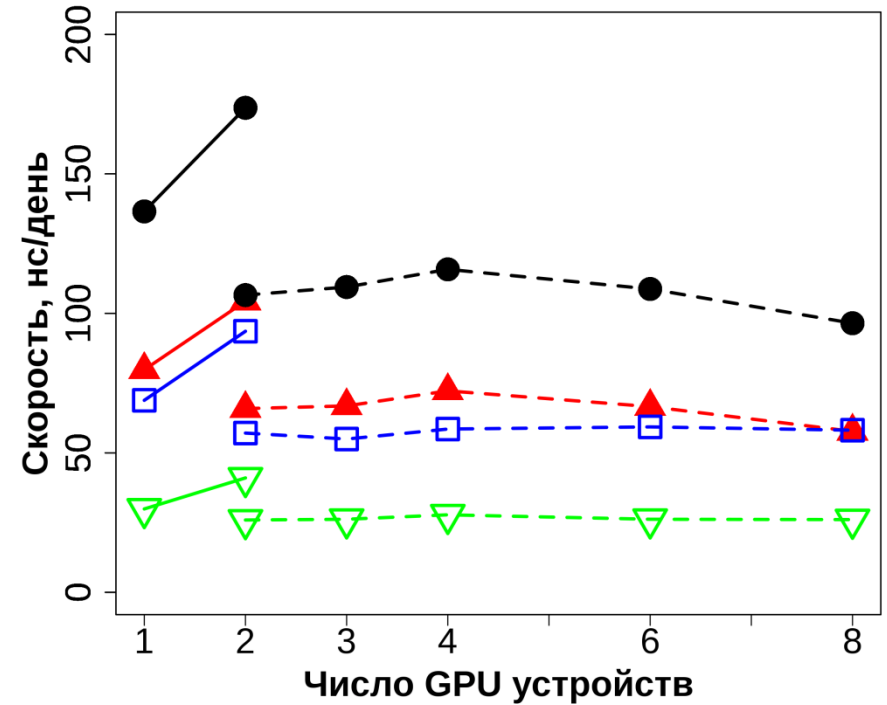
- Уменьшение количества атомов в ячейке в **×1,30** раза при одинаковой плотности ускоряет расчет МД в **×1.30** раза;
- Уменьшение количества атомов в ячейке в **×1,30** раза в результате использования 3-центральной модели растворителя (т.е. с уменьшением плотности) ускоряет расчет МД в **×1.60** раза;
- Использование 2-4 узлов с Tesla K40 ускоряет расчет по сравнению с одним узлом.

Молекулярная динамика на Tesla P100 GPU биологических систем разных размеров и плотности

FF14SB + TIP4P-Ew



FF15IPQ + SPC/Eb



Ускорение на двух Tesla P100 одного узла в режиме peer-to-peer по сравнению с одной Tesla P100 составило $\times 1,21$ – $\times 1,37$, однако производительность на двух и более узлах во всех случаях была существенно ниже производительности на одной видеокарте

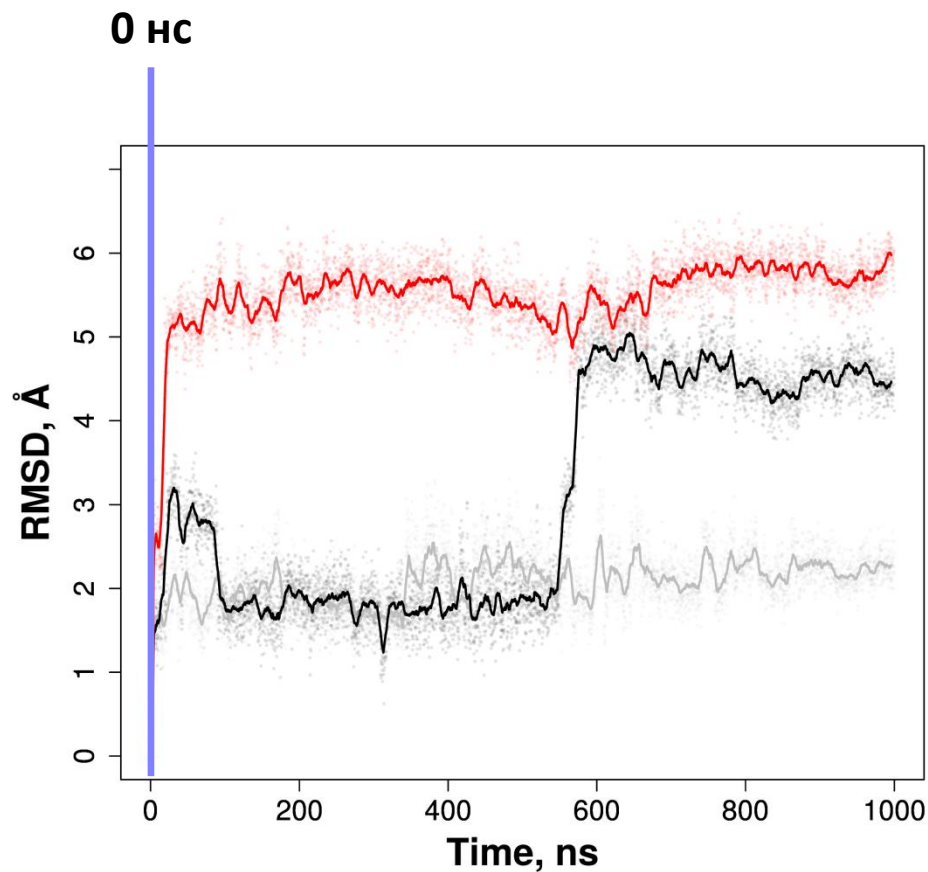
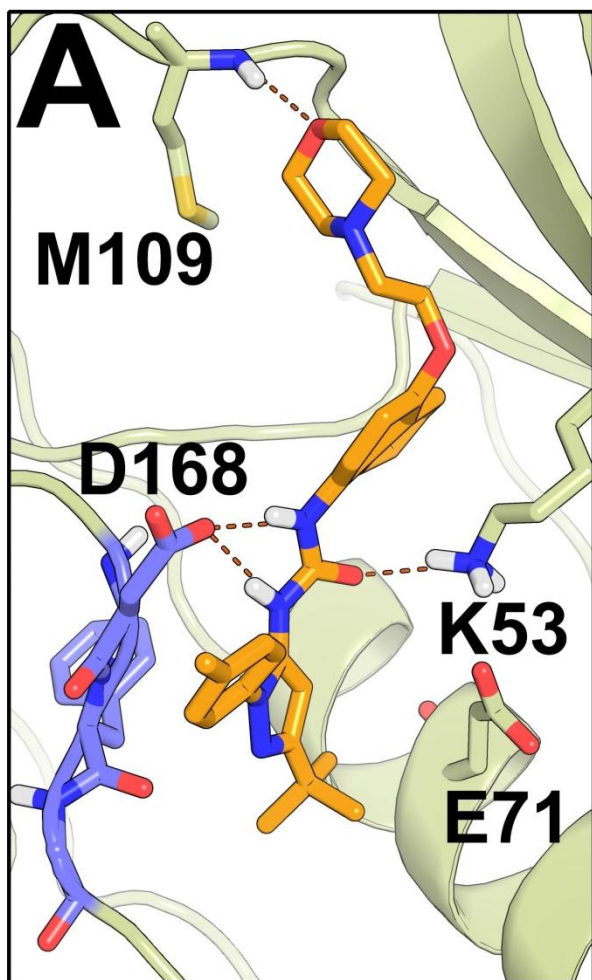
Молекулярная динамика на Pascal GPU биологических систем разных размеров и плотности

Скорость расчет МД биологических систем наиболее распространенных размеров (65 000 – 150 000 атомов) на двух видеокартах Tesla P100 **одного узла** в режиме peer-to-peer составляет 73 – 150 нс/день, т.е. **микросекундная шкала может быть достигнута за 1-2 недели.**

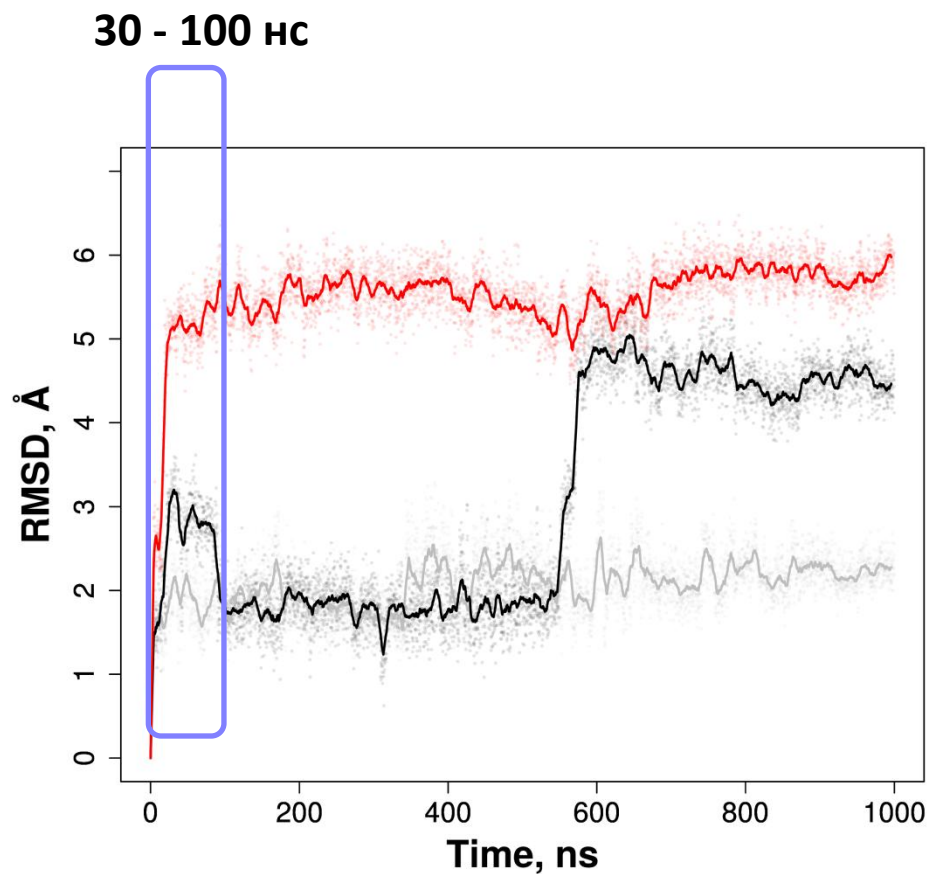
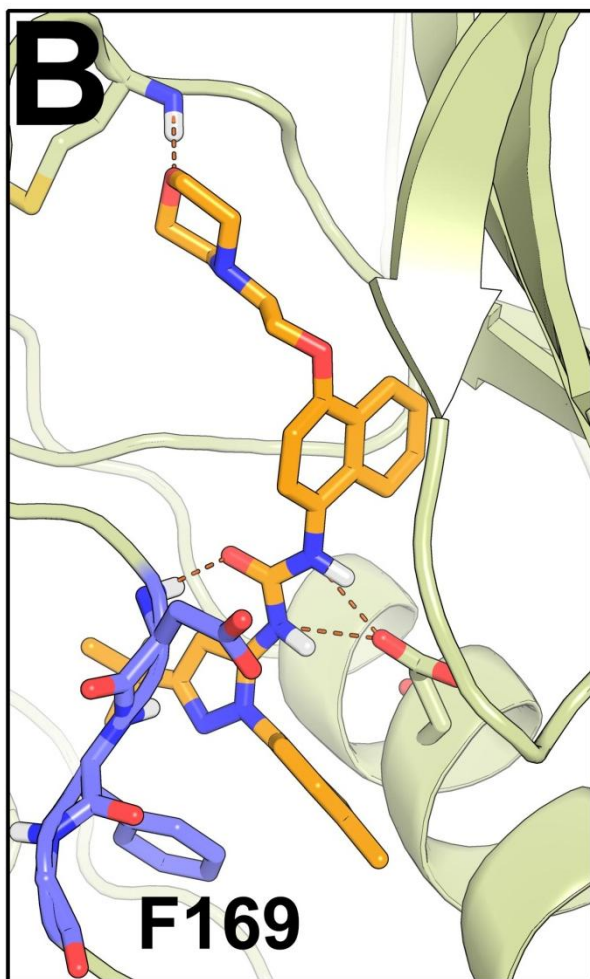
Это позволяет:

- считать существенно **более длинные траектории;**
- считать **большее количество независимых траекторий** (запущенных из разных начальных состояний или с разными начальными условиями), что критически важно, поскольку позволяет собрать больше данных для последующего анализа

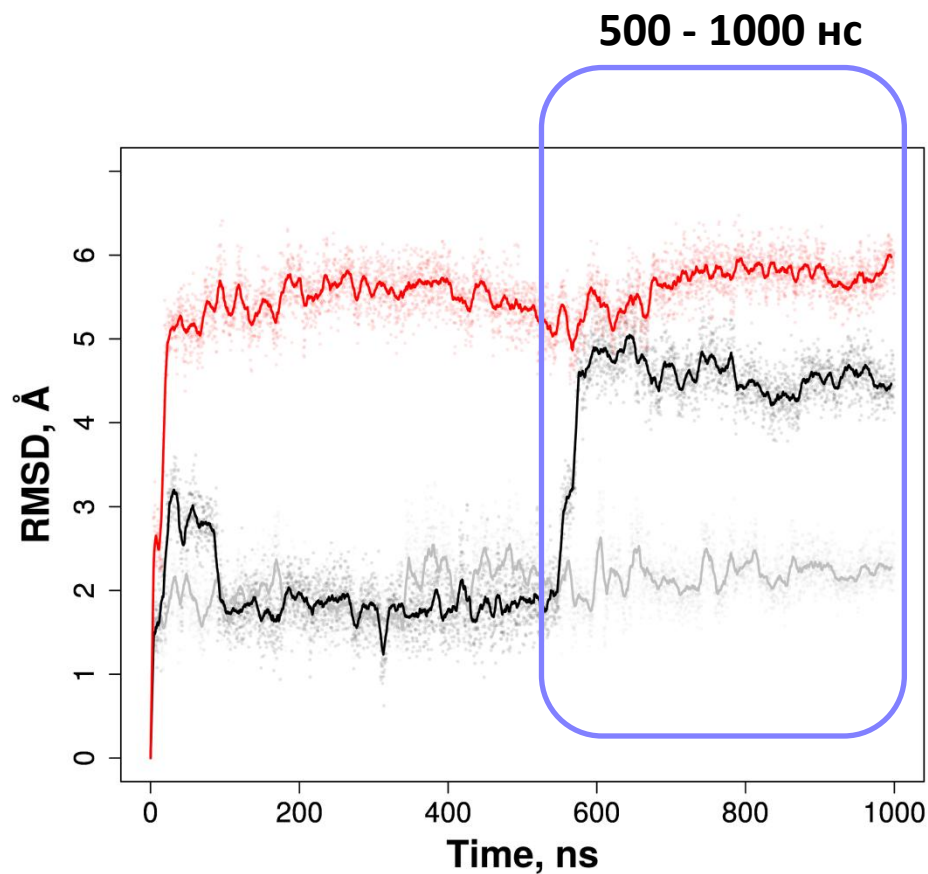
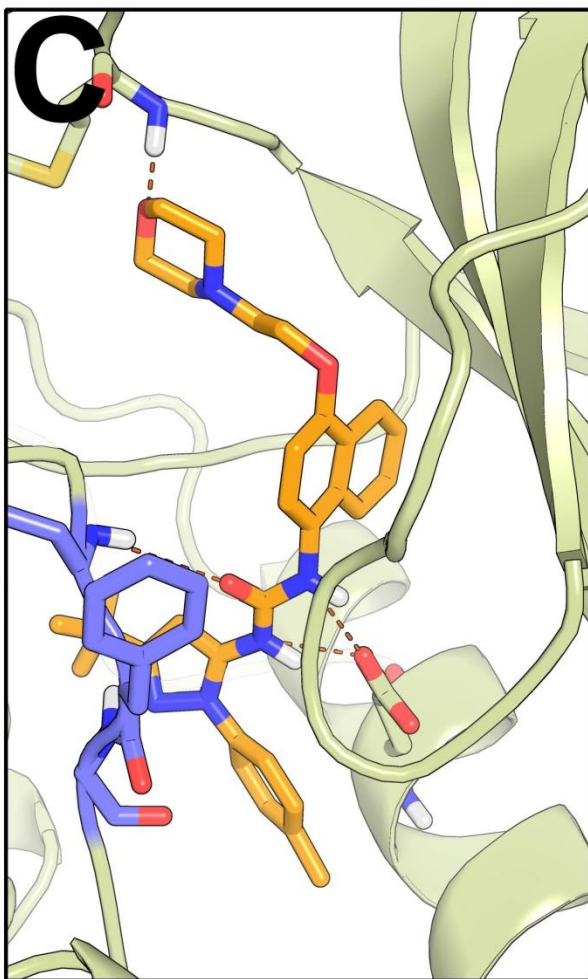
Симуляция структурной подвижности комплекса р38α MAPK человека с ингибитором дорамапимодом



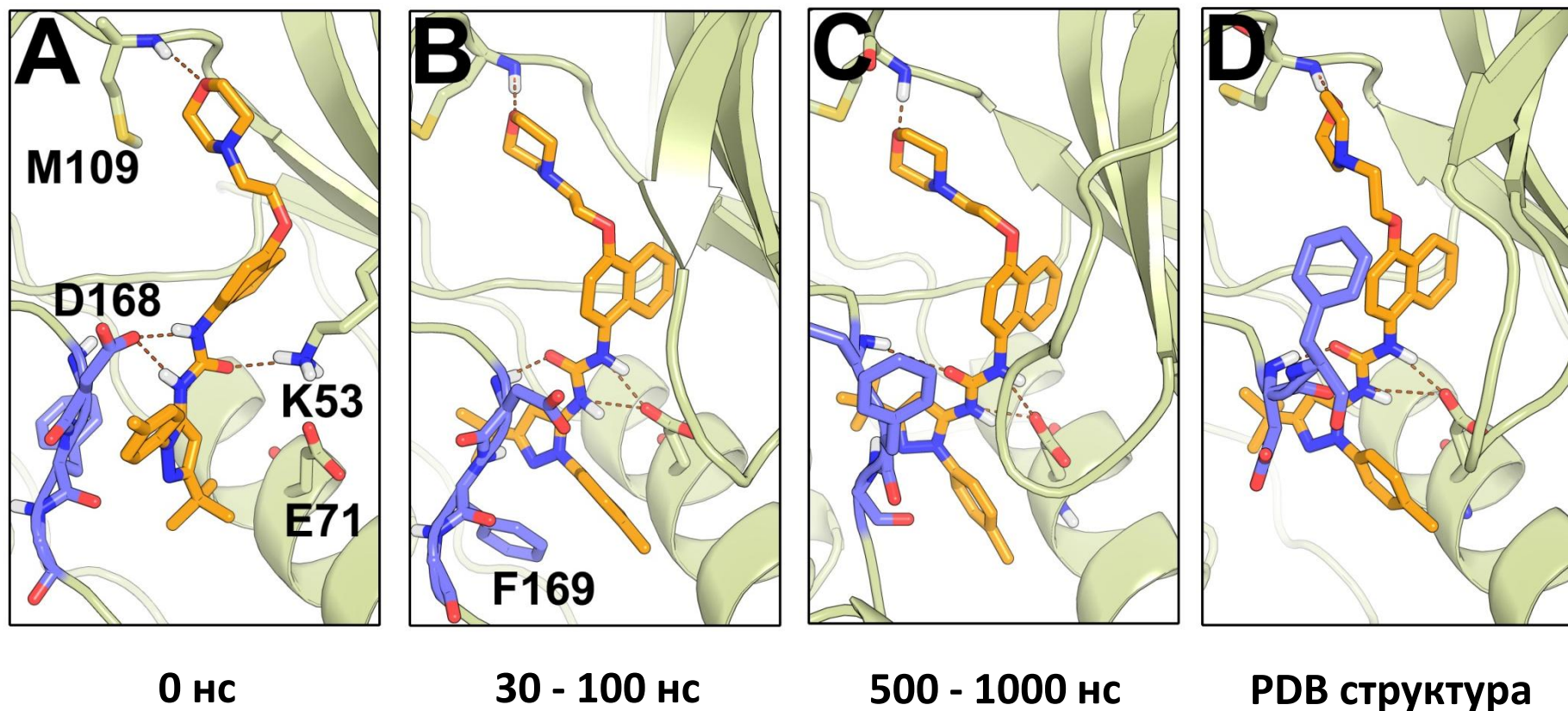
Симуляция структурной подвижности комплекса р38α МАРК человека с ингибитором дорамапимодом



Симуляция структурной подвижности комплекса р38α МАРК человека с ингибитором дорамапимодом

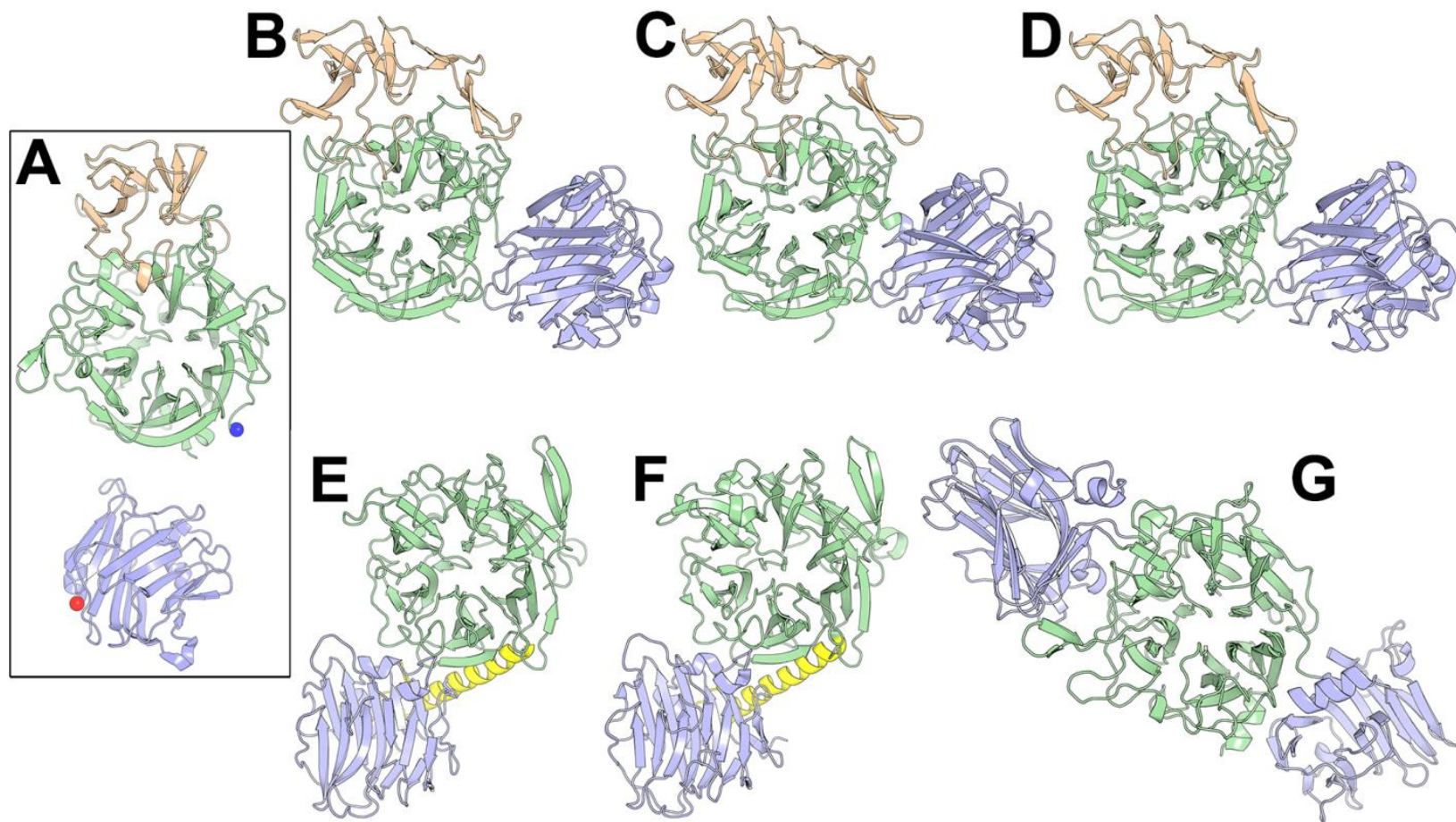


Длительная симуляция структурной подвижности p38a MAPK человека позволила изучить последовательные стадии механизма связывания дорамапимода



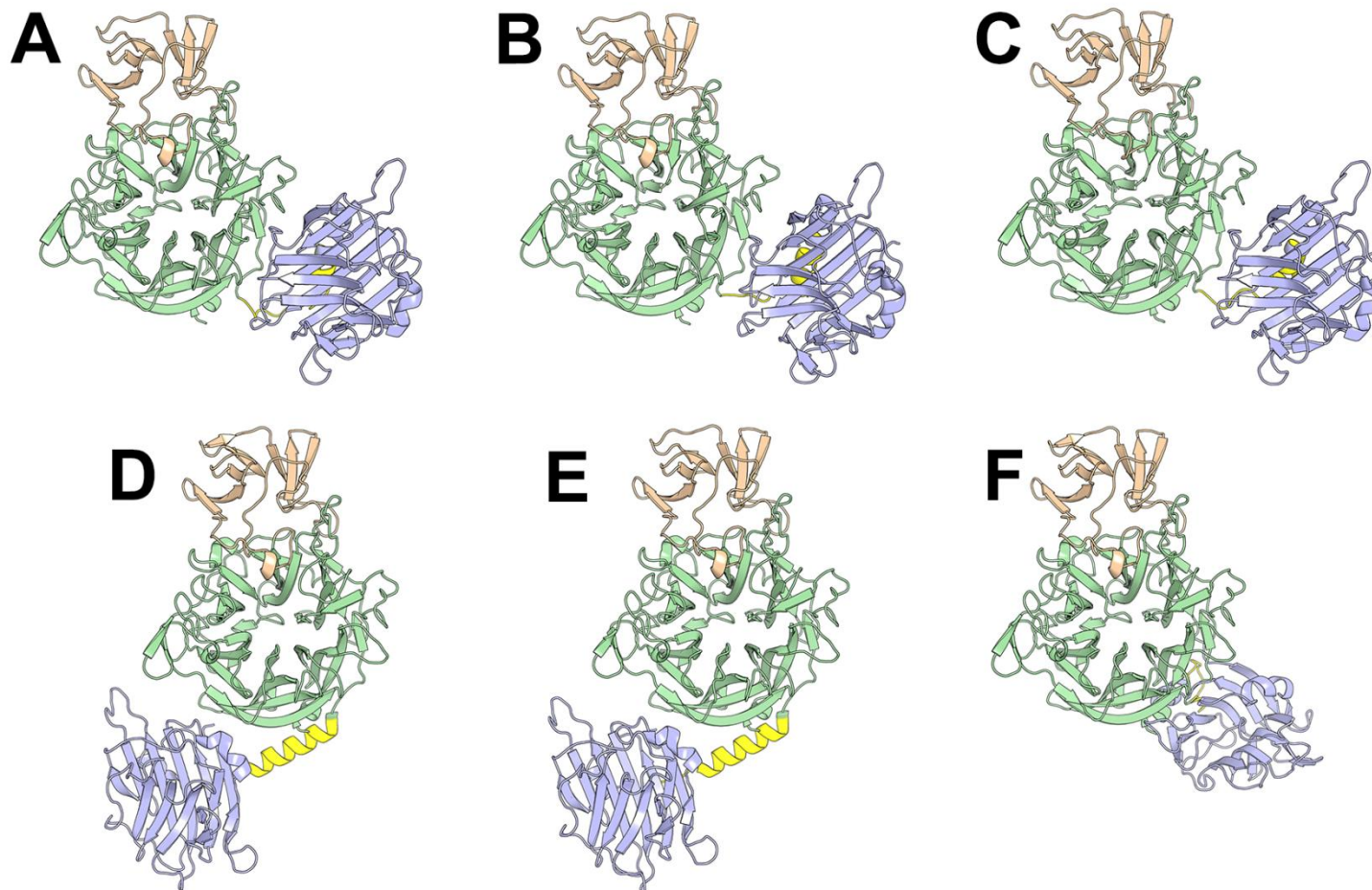
Suplatov, D. A., Kopylov, K. E., Sharapova Y.A., Švedas V. Human p38 α Mitogen-Activated Protein Kinase in the Asp168-Phe169-Gly170-in (DFG-in) state can bind allosteric inhibitor Doramapimod // *J. Biomol. Struct. Dyn.* – 2018. DOI:10.1080/07391102.2018.1475260

Симуляция структурной подвижности для изучения доменной организации бактериальных нейраминдаз



Sharapova Y., Suplatov D., Švedas V. Neuraminidase A from *Streptococcus Pneumoniae* Has a Modular Organization of Catalytic and Lectin Domains Separated by a Flexible Linker. *The FEBS Journal*. 2018. vol. 285, no. 13. pp. 2428–2445. DOI: 10.1111/febs.14486.

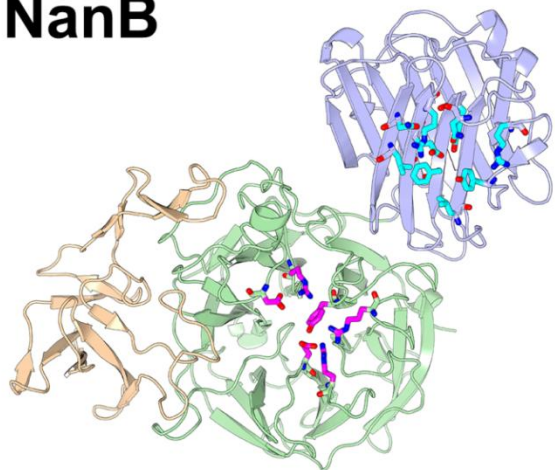
Симуляция структурной подвижности для изучения доменной организации бактериальных нейраминдаз



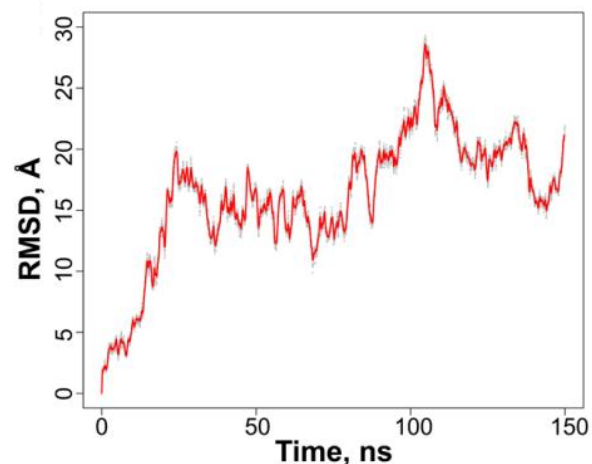
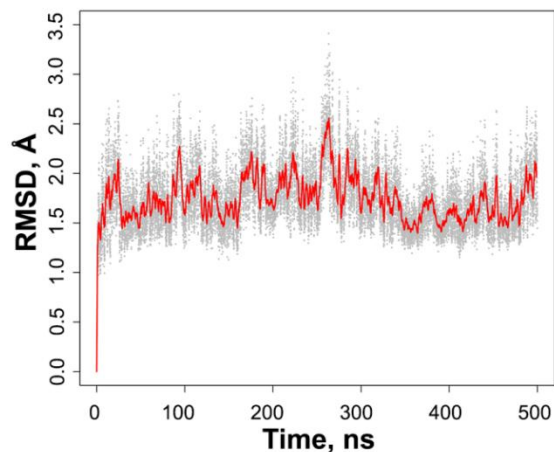
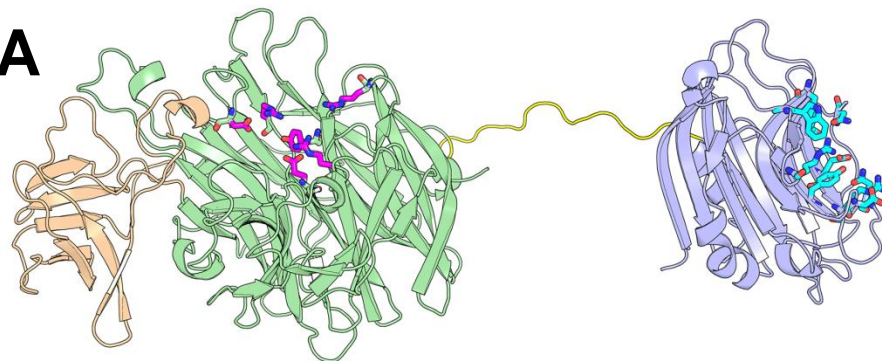
Sharapova Y., Suplatov D., Švedas V. Neuraminidase A from *Streptococcus Pneumoniae* Has a Modular Organization of Catalytic and Lectin Domains Separated by a Flexible Linker. *The FEBS Journal*. 2018. vol. 285, no. 13. pp. 2428–2445. DOI: 10.1111/febs.14486.

Симуляция структурной подвижности для изучения доменной организации бактериальных нейраминдаз

NanB

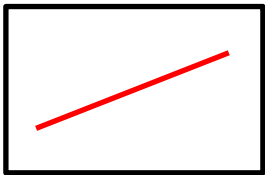
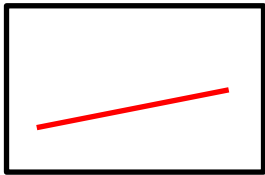
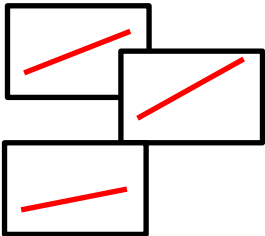
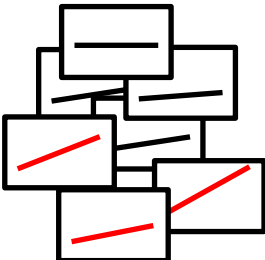
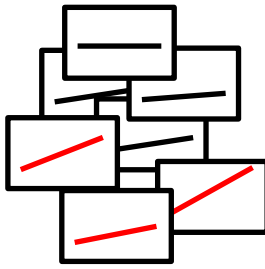




NanA

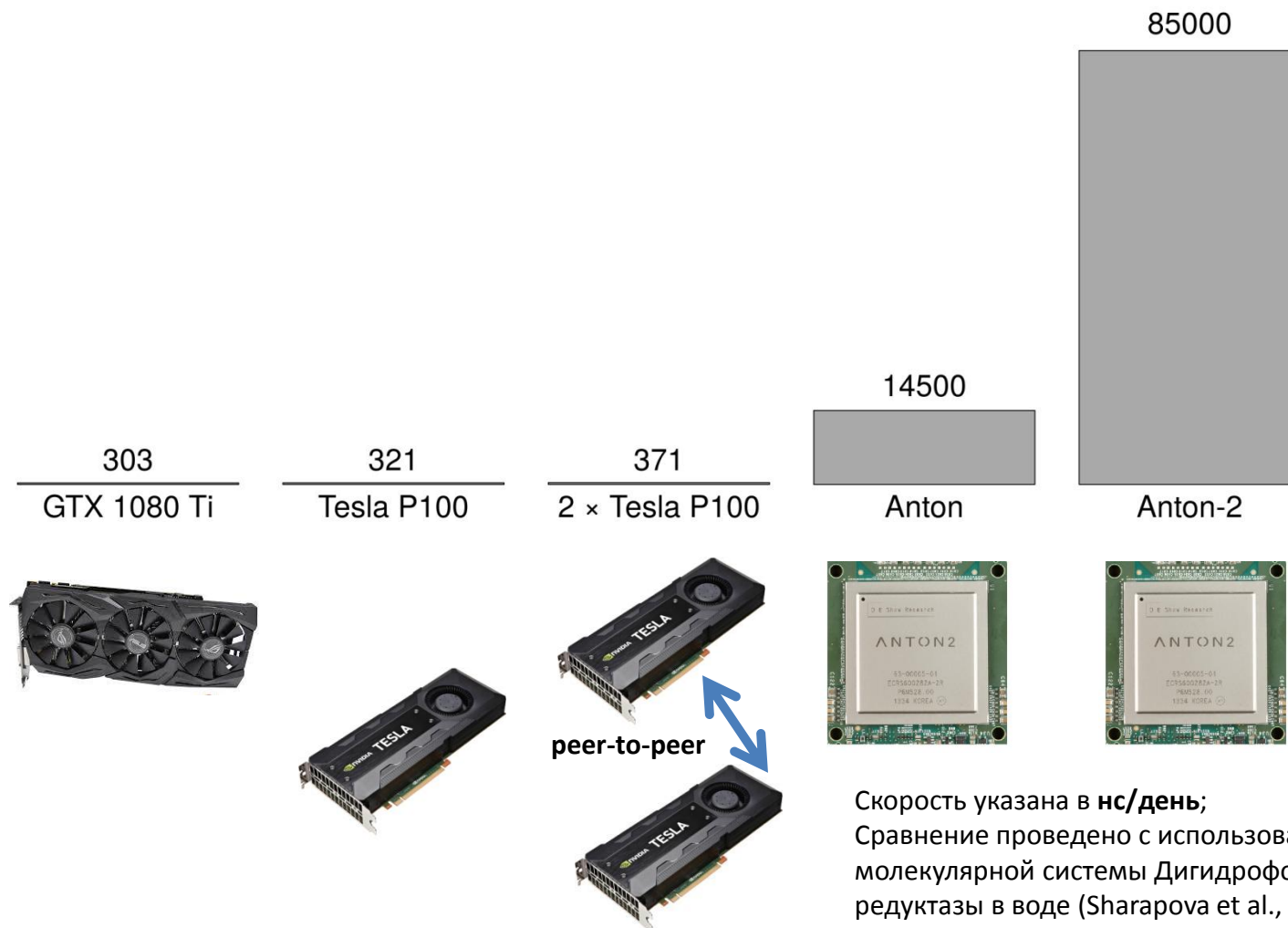


Sharapova Y., Suplatov D., Švedas V. Neuraminidase A from *Streptococcus Pneumoniae* Has a Modular Organization of Catalytic and Lectin Domains Separated by a Flexible Linker. *The FEBS Journal*. 2018. vol. 285, no. 13. pp. 2428–2445. DOI: 10.1111/febs.14486.

Объем вычислений молекулярной динамики в повседневной лабораторной практике

До 2000	2000 – 2008	2008 – 2014	2014-2017	2018
< 1 нс	10 нс	10 – 100 нс	100 нс	1000 нс
				
		Чебышев Ломоносов-1	Ломоносов-2 	Ломоносов-2 

Сравнение скорости молекулярной динамики на Pascal GPU и Anton (2008) /Anton-2 (2014)



Выводы

- Новое поколение GPU ускорителей Pascal, в том числе Tesla P100, установленные в новом разделе Ломоносов-2, позволяют достичь микросекундной шкалы при использовании молекулярной динамики в повседневной практике;
- Уменьшение количества атомов в ячейке на 25-31% в результате использования 3-центральной модели растворителя ускоряет расчет МД до 63% и ухудшает масштабируемость до 11%;
- Результаты, полученные в разных силовых полях и моделях воды, могут качественно отличаться, что говорит о необходимости совместного использования разных силовых полей при изучении биологических систем;
- Дальнейшее развитие доступных специализированных архитектур необходимо для достижения миллисекундной шкалы при симуляции больших белков в повседневной практике.

Благодарности

Работа выполнена

- при поддержке грантов
РФФИ №17-07-00751 и РНФ №15-14-00069
- с использованием оборудования Центра коллективного пользования сверхвысокопроизводительными вычислительными ресурсами МГУ имени М.В. Ломоносова

Публикации

Sharapova Y., Suplatov D., Švedas V. Neuraminidase A from *Streptococcus pneumoniae* has a modular organization of Catalytic and Lectin Domains separated by a flexible linker // *The FEBS journal*. – 2018. – no. 285 – pp. 2428–2445; <https://doi.org/10.1111/febs.14486>

Suplatov, D. A., Kopylov, K. E., Sharapova Y.A., Švedas V. Human p38 α Mitogen-Activated Protein Kinase in the Asp168-Phe169-Gly170-in (DFG-in) state can bind allosteric inhibitor Doramapimod // *J. Biomol. Struct. Dyn.* – 2018. <https://doi.org/10.1080/07391102.2018.1475260>

Sharapova Y., Suplatov D., Švedas V. Simulating the long-timescale structural behavior of bacterial and influenza Neuraminidases with different HPC resources // *Supercomputing Frontiers and Innovations*. — 2018. — Принято в печать

Суплатов Д.А., Шарапова Я.А., Попова Н.Н., Копылов К.Е., Воеводин Вл.В., Швядас В.К. Молекулярная динамика в силовом поле FF14SB в воде TIP4P-Ew, и в силовом поле FF15IPQ в воде SPC/Eb: сравнительный анализ на GPU и CPU // *Вестник ЮУрГУ. Серия: Вычислительная математика и информатика*. 2018 — Принято в печать