

УДК 541.64:547.458

ИСТОРИЯ И ПЕРСПЕКТИВЫ АТОМИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ ПОЛИСАХАРИДОВ

© 2023 г. В. И. Дещеня^{a,b,*}, Н. Д. Кондратюк^{a,b,c}

^aМосковский физико-технический институт (национальный исследовательский университет)
141701 Московская обл., Долгопрудный, Институтский пер., 9, Россия

^bОбъединенный институт высоких температур Российской академии наук
125412 Москва, ул. Ижорская, 13, стр.2, Россия

^cНациональный исследовательский университет Высшая школа экономики
101000 Москва, ул. Мясницкая, 20, Россия

*e-mail: deshchenia.vi@phystech.edu

Поступила в редакцию 17.04.2023 г.

После доработки 09.06.2023 г.

Принята к публикации 13.07.2023 г.

Представлены история и перспективы развития потенциалов межатомного взаимодействия для моделирования молекулярных систем и, в частности, полисахаридов. Рассмотрены популярные семейства потенциалов, такие как CHARMM, GROMOS, AMBER и OPLS. Обозначены проблемы применимости данных моделей для мономеров сахаров: плохая воспроизводимость экспериментальных свойств и гиперагрегация при малых концентрациях. Также рассмотрены векторы развития данной области: модификация невалентных взаимодействий в потенциале и использование поляризуемых потенциалов. Продемонстрированы примеры применения атомистического моделирования полисахаридов в актуальных фундаментальных и промышленных задачах на примере целлюлозы. Приведены важные вычислительные работы, внесшие существенный вклад в понимание структуры и процессов, происходящих в кристалле целлюлозы. Кроме того, дан обзор доступных на сегодня масштабов времен и характерных размеров систем для моделирования в пакетах GROMACS, LAMMPS, OpenMM и AMBER для модели раствора сахарозы.

DOI: 10.31857/S2308114723700231, EDN: HOXFTS

ВВЕДЕНИЕ

Материалы на основе полисахаридов широко применяются в промышленности для различных задач. Так, декстран и фиколл являются эффективными криопротекторами [1–6], а мембраны на основе целлюлозы служат для опреснения и очистки воды [7, 8]. Соединения и комплексы на основе полисахаридов используются в биомедицинских приложениях [9, 10] и робототехнике [11]. В последнее десятилетие научное сообщество проявляет интерес к изучению углеводов, вызванный в значительной степени признанием их ключевой роли во многих биохимических и технологических процессах, а также их потенциала для создания новых материалов. Например, мембраны на основе целлюлозы являются перспективными материалами для проточных окислительно-восстановительных (редокс) батарей [12, 13], а композиты, содержащие хитозан, имеют потенциал для использования в регенеративной медицине [14]. Кроме того, нанокompозиты на основе полисахаридов перспективны благодаря высоким физико-механическим характеристикам, биосов-

местимости, биоразлагаемости и низкой стоимости [15].

Современные методы молекулярного моделирования могут выступать альтернативой экспериментам по поиску соединений с необходимыми свойствами, но для этого нужна верификация моделей межатомного взаимодействия [16–21]. На данный момент вычислительных работ, посвященных углеводам, немного. Причина заключается в сложности разработки соответствующих моделей для углеводов разной молекулярной массы, широком наборе функциональных групп, крайней конформационной неоднородности и ограниченности экспериментальных данных [22, 23]. Однако появление первых квантово-механических расчетов для гексапираноз сделало возможным создание и апробирование потенциалов межатомного взаимодействия для углеводов [24–27].

Разработка надежных теоретических моделей углеводов необходима, например, для лучшего понимания механизмов, определяющих их биологические функции, и описания теплофизических свойств растворов, требующихся в технологиче-

ских приложениях. В данной работе приведены базовые идеи метода молекулярной динамики, который служит мощным инструментом для исследования различных физических и химических систем, история развития наиболее распространенных семейств силовых полей, таких как CHARMM, GROMOS, AMBER и OPLS, а также история создания и текущее состояние потенциалов межатомного взаимодействия для моделирования углеводных систем. В конце обзорной статьи приведены современные примеры применения данных методов, а также сведения о развитии вычислительной инфраструктуры.

МЕТОД КЛАССИЧЕСКОЙ МОЛЕКУЛЯРНОЙ ДИНАМИКИ

Метод молекулярной динамики является мощным инструментом для исследования физических, химических и биологических систем. Он позволяет рассчитывать траектории каждой частицы и исследовать систему на атомарном уровне. Временной масштаб исследуемых траекторий при этом существенно зависит от вычислительных ресурсов.

Первое упоминание метода молекулярной динамики (МД) в литературе относится к 1950-м годам [28], когда происходил значительный рост вычислительных мощностей. Тогда же был создан первый электронный компьютер Mathematical Analyzer Numerical Integrator and Automatic Computer Model I (MANIAC I) мощностью 10 кфлопс (10 тысяч операций в секунду). Первые МД-работы исследовали системы из 100–1000 атомов [29, 30] и использовали достаточно простые модели взаимодействия атомов между собой. Современное состояние вычислительной техники позволяет рассматривать системы вплоть до сотен миллионов атомов [31] с более сложными моделями взаимодействия, что дает возможность моделировать полноатомную динамику целых органелл. А в 2021 году с участием разработчиков из NVIDIA и LAMMPS ученым удалось рассчитать миллиард атомов углерода на суперкомпьютере Summit [32]. Однако такие расчеты доступны только малому числу научных групп, имеющих доступ к лидирующим суперкомпьютерам. Рутинные расчеты, доступные большинству научных групп, описывают системы размером порядка 100 тысяч атомов на наносекундных временах.

Отдельно стоит выделить линейку специализированных для молекулярно-динамических расчетов суперкомпьютеров Anton, созданных D.E. Shaw Research [33–36]. Специализация архитектуры суперкомпьютеров позволила достигнуть существенной производительности. Так, скорость расчета систем, содержащих сотню миллионов атомов, превосходит наносекунду в день, что более чем на порядок выше по сравне-

нию с лучшими неспециализированными для МД суперкомпьютерами.

Суть метода молекулярной динамики заключается в численном интегрировании для каждой из частиц системы уравнений Ньютона

$$m_i \frac{d^2 r_i}{dt^2} = F_i(r_1, \dots, r_N), \quad (1)$$

где m_i – масса i -й частицы, r_i – ее координата, F_i – сила, действующая на нее. Важно заметить, что сила, действующая на конкретную частицу, зависит от положения других N частиц системы и определяется потенциалом межатомного взаимодействия $U(r_1, \dots, r_N)$:

$$F_i(r_1, \dots, r_N) = -\frac{\partial U(r_1, \dots, r_N)}{\partial r_i} \quad (2)$$

Параметризация потенциала $U(r_1, \dots, r_N)$ определяет достоверность модели и точность прогнозов свойств. В следующем разделе приводится история развития потенциалов (силовых полей) для моделирования молекулярных систем и, в частности, углеводов.

Общая схема расчетов при этом имеет вид, представленный на рис. 1. Ключевыми моментами являются формула для расчета энергии системы, определяющей силы, действующие на атомы, и схема численного интегрирования. Путем повторения шагов 2–4 схемы можно получить траекторию движения каждого атома $\{r_i(t), v_i(t), F_i(t)\}_{i=1, t=0}^{i=N, t=T}$. Знание траекторий атомов позволяет рассчитывать макроскопические параметры системы и значения интересующих физических величин (например, вязкость раствора или коэффициент диффузии иона в мембране).

Наиболее простой является схема интегрирования Эйлера, однако она имеет целый ряд критичных недостатков. В частности, она необратима по времени и не сохраняет фазовый объем при эволюции системы. Более сложной, но существенно более точной является схема Верле и эквивалентные ей алгоритмы, такие как скоростная схема Верле [37] и алгоритм Leap Frog [38]. Наиболее часто используется схема Верле:

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{F_i(t)}{2m_i} \Delta t^2, \quad (3)$$

$$v_i(t + \Delta t) = v_i(t) + \frac{F_i(t + \Delta t) + F_i(t)}{2m_i} \Delta t. \quad (4)$$

ИСТОРИЯ РАЗВИТИЯ ПОТЕНЦИАЛОВ МЕЖАТОМНОГО ВЗАИМОДЕЙСТВИЯ

Общий вид потенциала

Взаимодействия между атомами системы молекул в нереакционной МД могут быть разделены



Рис. 1. Общая схема расчета методом классической молекулярной динамики. Цветные рисунки можно посмотреть в электронной версии.

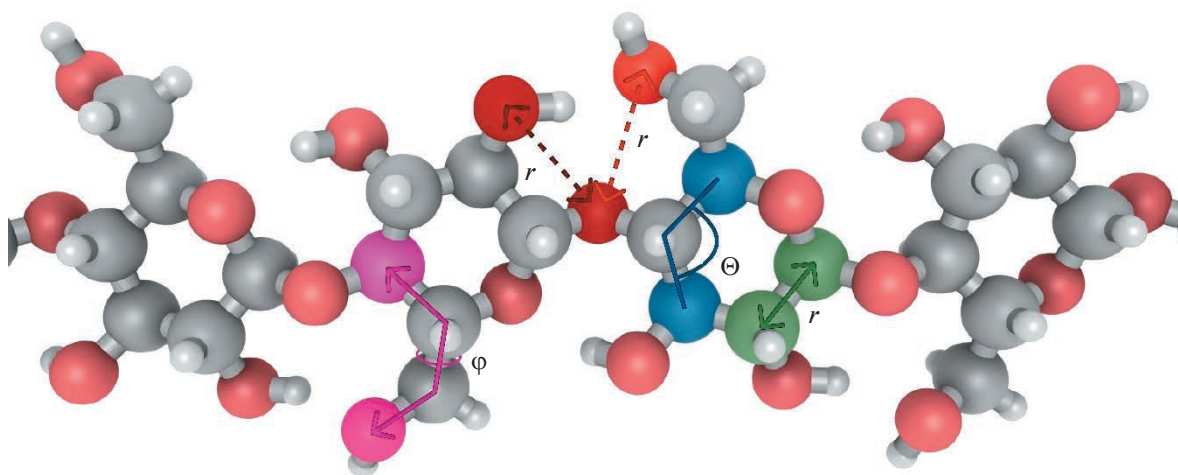


Рис. 2. Молекула целлюлозы с основными взаимодействиями, представленными в большинстве силовых полей: колебания ковалентных связей (зеленый), колебания углов между связями (синий), колебания двугранных углов (фиолетовый), невалентные взаимодействия 1–4 (темно-красный) и 1–5 (ярко-красный). Цветами отмечены пары атомов, связанные соответствующим взаимодействием. Визуализация молекулы выполнена в программе iRASPА [225].

на две группы: внутримолекулярные (взаимодействия между атомами, принадлежащими одной молекуле) и межмолекулярные (между атомами разных молекул). Первая группа включает валентные (колебания длины связей, углов между связями и углов двугранных вращений) и невалентные (электростатическое E_C и ван-дер-ваальсово E_{vdw}) взаимодействия. Вторая группа – только электростатическое E_C и ван-дер-ваальсо-

во E_{vdw} взаимодействия. На рис. 2 на примере молекулы целлюлозы показаны основные виды взаимодействий.

Ван-дер-ваальсово взаимодействие между атомами обычно описывается потенциалом Леннарда-Джонса

$$E_{vdw} = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6], \quad (5)$$

где r – расстояние между центрами частиц, а ϵ и σ – параметры взаимодействия: ϵ – глубина по-

тенциальной ямы, σ – модельный радиус частицы. В случае, когда взаимодействуют частицы разных типов (у которых отличаются значения ϵ и σ), для расчета энергии взаимодействия используются параметры, полученные с помощью правил смешения. Наиболее часто используют правила смешения Лоренца–Бертло [39, 40]

$$\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}, \quad \sigma_{ij} = (\sigma_i + \sigma_j)/2 \quad (6)$$

(ϵ_i и σ_i – параметры i -й частицы, а ϵ_j и σ_j – j -й).

Электростатическое взаимодействие задается законом Кулона

$$E_C = \frac{q_i q_j}{4\pi\epsilon_0 r}, \quad (7)$$

где r – расстояние между взаимодействующими частицами, q_i и q_j – их заряды, а ϵ_0 – диэлектрическая проницаемость.

Ковалентные связи между атомами и углы между ними довольно часто описываются гармоническим потенциалом

$$E_{bond} = E_b(r - r_0)^2, \quad E_{angle} = E_\Theta(\Theta - \Theta_0)^2 \quad (8)$$

Такое описание используется в силовых полях семейств AMBER [41], CHARMM [42, 43], OPLS [44], а также GROMOS87 [45, 46].

Для более точного задания взаимодействий иногда нужно добавлять в модель перекрестные слагаемые, отвечающие, например, за взаимосвязь угла и длины связей. Силовые поля, учитывающие эти поправки, показывают лучшие результаты по воспроизведению экспериментальных уравнений состояния и транспортных коэффициентов для, например, углеводородов [47–49]. Примерами таких полей являются UFF [50], COMPASS [51] и MMFF94 [52]. Так, в потенциале MMFF94 [52] энергия связи имеет негармонический вид

$$E_{bond} = E_{b2}(r - r_0)^2 + E_{b3}(r - r_0)^3 + E_{b4}(r - r_0)^4, \quad (9)$$

$$E_{angle} = E_{a2}(\Theta - \Theta_0)^2 + E_{a3}(\Theta - \Theta_0)^3$$

Кроме того, потенциал учитывает взаимосвязь между изменением длины связи и угла:

$$E_{bond-angle} = (E_{ba}^{ij}(r^{ij} - r_0^{ij}) + E_{ba}^{kj}(r^{kj} - r_0^{kj}))(\Theta^{ijk} - \Theta_0^{ijk}) \quad (10)$$

Добавление перекрестных слагаемых, как и рассмотрение дополнительных, более высоких порядков колебаний ковалентных связей и углов, приводят к значительному увеличению количества параметров.

Помимо ковалентных связей и углов между ними в функциональной форме силовых полей принимаются во внимание и другие валентные взаимодействия, например, потенциал колеба-

ния двугранных углов. Важно отметить, что невалентные взаимодействия между атомами, связанными ковалентно (1–2 взаимодействия) или разделенными двумя ковалентными связями (1–3 взаимодействия), не учитываются. Невалентные взаимодействия 1–4 часто масштабируются, поскольку часть взаимодействия уже учтена в потенциале колебаний двугранного угла. Так, в силовом поле ff14sb из семейства AMBER электростатические 1–4 взаимодействия уменьшены в 1.2 раза, а ван-дер-ваальсовы – в 2 раза [53]. В силовом поле OPLS оба типа невалентных взаимодействий уменьшены в 2 раза [54], а в GLYCAM06 они учитываются в полном объеме [55]. Такое отличие делает силовые поля несовместимыми друг с другом. Так, последние версии силового поля GLYCAM, которое ранее входило в семейство AMBER, стали несовместимы с потенциалами AMBER.

Существуют и более сложные потенциалы, которые учитывают эффекты, необходимые для более точного описания соединений в органической химии, например, поляризации и стереоэлектронных эффектов. Примерами силовых полей этого класса являются Atomic Multipole Optimized Energetics for Biomolecular Simulation (АМОЕВА) и DRUDE [56–58].

Выбор потенциала для моделирования основывается на рассматриваемой системе и на том, какие свойства являются приоритетными при расчете. Для конформационного анализа органических молекул было показано [59] превосходство потенциалов MMFF94 [52], MM2 [60] и MM3 [61], а также поляризуемого потенциала АМОЕВА [57]. Плотности жидкостей малых органических молекул лучше всего предсказываются [62] потенциалами TraPPE-UA [63–67] и CHARMM22 [68]. А потенциалы OPLS [69–71] и GROMOS-2016H66 [72] оказываются точнее при расчете свободных энергий сольватации [73].

Для моделирования углеводов в основном используются силовые поля без перекрестных слагаемых в потенциале валентных взаимодействий, например, поля семейств CHARMM, GROMOS, AMBER, OPLS, которые подробно рассмотрены в следующей главе статьи. Однако разработка поляризуемых силовых полей также активно ведется [74–77].

Модель CHARMM

Семейство силовых полей Chemistry at Harvard Macromolecular Mechanics (CHARMM) [42, 43] разрабатывается с 1983 года научной группой профессора Мартина Карплуса из гарвардского университета. Его вклад был высоко оценен научным сообществом, и в 2013 году он вместе с коллегами получил нобелевскую премию по химии за

“компьютерное моделирование химических систем”. Наиболее важными силовыми полями семейства CHARMM являются CHARMM22/CMAP (2003) [68, 78, 79], параметризованное для моделирования белков, CHARMM27 (2000) [80] для нуклеиновых кислот (РНК и ДНК) и CHARMM36 для липидов [81]. Последнее было расширено для моделирования нуклеиновых кислот [82, 83] в 2011–2012 годах и белков [84] в 2016 году. Кроме того, пользуется популярностью силовое поле CHARMM General Force Field (CGenFF) для малых органических молекул, разрабатываемое с 2009 года [85, 86].

Уже в 1988 году была опубликована версия силового поля с функциональной формой CHARMM для моделирования пираноз [87]. Данная версия имела ряд недостатков, например, давала неверные конформации гидроксиметильных групп [88]. Последующие адаптации CHARMM на класс углеводов изменяли энергетическую функцию потенциала, добавляя в нее слагаемые, соответствующие водородным связям, и меняя форму потенциала колебания двугранных углов [89–93]. Это привело к отсутствию совместности с другими полями семейства и невозможности моделирования систем, включающих другие, отличные от углеводов, соединения.

Начиная с 2008 года идет разработка силового поля CHARMM для углеводов, совместимого с другими силовыми полями семейства. На данный момент оно охватывает основные классы моносахаридов гексопиранозы [94], альдопентофуранозы и фруктофуранозы [95], но также подобраны параметры для гликозидных связей [94, 96], необходимые для изучения дисахаридов и полисахаридов.

В силовых полях данного семейства энергия системы рассчитывается по общей формуле

$$\begin{aligned}
 E = & \sum_{bonds} K_b (r - r_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \\
 & + \sum_{dihedrals} K_\varphi (1 + \cos(n\varphi - \delta)) + \\
 & + \sum_{improper\ dihedrals} K_\phi (\phi - \phi_0)^2 + \\
 & + \sum_{Urey-Bradley} K_{UB} (r_{1,3} - r_{1,3,0})^2 + \\
 & + \sum_{nonbonded} (E_C + E_{LJ})
 \end{aligned} \quad (11)$$

Здесь r – расстояние между ковалентно-связанными атомами; θ – угол между ковалентными связями; φ – двугранный угол, проходящий через ковалентную связь; n – параметр, определяющий периодичность потенциала вращения; δ – фазовый сдвиг; ϕ – двугранный угол между плоскостями, образованными тройками атомов; $r_{1,3}$ – расстояние между атомами, которые ковалентно связаны с одним атомом; r_0 , θ_0 , ϕ_0 , $r_{1,3,0}$ – соответ-

ствующие равновесные значения; K_b , K_θ , K_φ , K_ϕ , K_{UB} – константы энергии для соответствующих взаимодействий; E_C , E_{LJ} – энергии электростатического и лэннард-джонсоновского взаимодействий между атомами.

В общей формуле первые два слагаемых описывают ковалентные связи и углы между ними в гармонической форме, третье слагаемое (dihedrals) – энергия вращения вокруг валентной связи, четвертое (improper dihedrals) – энергия отклонения атома от плоскости, образованной тремя его соседями (это слагаемое поддерживает структуру молекулы плоской), а последнее слагаемое – стандартная форма описания невалентных взаимодействий.

Пятое слагаемое формулы (11), потенциал Юри–Брэдли, характеризует взаимодействия между двумя атомами, соединенными с третьим, расположенным между ними. При уменьшении угла между соседними связями изменяется длина связи, обусловленной взаимодействием между внешними атомами связанного триплета. Потенциал Юри–Брэдли позволяет добиться лучшего согласия с колебательными спектрами, но обычно несуществен, особенно в случае большой конформационной выборки.

Важным силовым полем семейства является поляризуемое поле CHARMM Drude [97], работающее для белков [98], нуклеиновых кислот [99] и некоторых углеводов, таких как гексопиранозы [74], альдопентофуранозы [75] и полисахариды, состоящие из пираноз и фураноз [77].

Модель GROMOS

Силовые поля GROMOS изначально создавались для одноименного программного пакета GRONingen MOlecular Simulation (GROMOS), который разрабатывался с 1978 года в Гронингском университете в Нидерландах научной группой профессора Вильфреда ван Ганстерена. В 1990 году он перешел работать в Швейцарскую высшую техническую школу Цюриха (ETH Zurich), где продолжает заниматься разработкой уже совместно с научными группами профессоров Филиппа Хюненбергера и Серейны Риникер.

Силовые поля этого семейства не являются полноатомными, в них алифатические группы CH_n ($n = 1, 2, 3, 4$) представлены едиными частицами, называемыми объединенными атомами. Функциональная форма силового поля была пересмотрена в 1996 году и сейчас имеет следующий вид [100]:

$$\begin{aligned}
 E = & \sum_{bonds} K_b (r^2 - r_0^2)^2 + \sum_{angles} K_\theta (\cos\theta - \cos\theta_0)^2 + \\
 & + \sum_{dihedrals} K_\phi (1 + \cos\delta \cos(n\phi)) + \\
 & + \sum_{improper\ dihedrals} K_\phi (\phi - \phi_0)^2 + \sum_{nonbonded} (E_C + E_{LJ})
 \end{aligned} \quad (12)$$

Первая версия силового поля 26C1 была создана в 1981 году [45, 46] и включала всего 26 типов атомов. После нескольких доработок в 1987 году было выпущено силовое поле GROMOS87 (37D4), которое осталось неизменным на протяжении 10 лет. Наконец, в 1996 году была выпущена версия GROMOS96 (43A1), в которой были уточнены леннард-дженсоновские параметры для алифатических объединенных атомов углерода, а также пересмотрена функциональная форма силового поля для повышения производительности [100]. Начиная с этой версии силовые поля делятся на две группы. А-версия силового поля – это базовое силовое поле, предназначенное для молекул в растворе или в кристаллической форме. В-версия является производной от А-версии. Она пренебрегает эффектом диэлектрического экранирования окружающей среды и предназначена для моделирования молекул в вакууме.

В версии 45A4 от 2003 года добавлены параметры моделирования углеводов на основе гексопиранозы в явном растворителе [101]. Набор параметров был проверен для небольшого ряда моно- и дисахаридов в растворе путем сравнения результатов моделирования с экспериментальными данными рентгеновской кристаллографии и спектроскопии ЯМР. Тогда проверка была ограничена α - и β -D-глюкозой, α - и β -D-галактозой, а также дисахаридами трегалозой, мальтозой и целлобиозой. Одновременно были разработаны наборы параметров 53A5 и 53A6, где была выполнена полная репараметризация параметров невалентных взаимодействий для моделирования жидкостей малых молекул (53A5) и растворов молекулярных систем в воде и неполярных растворителях (53A6) [102].

Силовое поле GROMOS 45A4 протестировано во многих независимых исследованиях моно-, ди- и полисахаридов (включая их производные) [103–105]. Большинство расчетов показало хорошее согласие с имеющимися экспериментальными данными, однако был обнаружен ряд расхождений по воспроизведению конформаций углеводов [106–108]. Это указывало на необходимость повторной оптимизации для лучшего воспроизведения углеводов. Наконец, в 2010 году версия силового поля 45A4 была репараметризована для углеводов на основе гексопиранозы и опубликована под названием 56A_{CARBO} [109]. В 2016 была выпущена ее доработанная версия 56A_{CARBO_R}, а позднее были опубликованы версии для произ-

водных хитозана [110] и фураноз [111]. Однако область применения данных силовых полей крайне ограничена, поскольку у них отсутствует совместимость с общими силовыми полями для биомолекул. Напротив, модификация 53A6_{GLYC} расширяет потенциал 53A6 для моделирования гексопираноз и имеет полную совместимость с силовыми полями 53A6 и 54A7 [112].

Отдельно стоит упомянуть популярное крупнозернистое (coarse-grained) силовое поле MARTINI [113]. В отличие от полноатомных силовых полей в нем группа атомов рассматривается как одна частица. Это позволяет достичь значительно большей производительности расчетов. Например, в силовом поле MARTINI 3 [114] в качестве таких частиц выступают набор из 4 молекул воды (сразу 12 атомов), ацетат (11 атомов) и *n*-бутан (14 атомов). Первая версия силового поля MARTINI была расширена на класс углеводов в 2009 году [115]. Для параметризации использовалось полноатомное поле GROMOS 45A4. Версия силового поля MARTINI 3 [114], вышедшая в 2021 году, была расширена на класс углеводов несколькими группами уже в 2022 году [116, 117].

Модель AMBER

Семейство силовых полей Assisted Model Building with Energy Refinement (AMBER), как и CHARMM, включает несколько потенциалов, параметризованных для различных классов химических соединений. Оно разрабатывается научной группой профессора Питера Кольмана из калифорнийского университета в Сан-Франциско с 1984 года [41].

Функциональная форма потенциала AMBER имеет вид

$$\begin{aligned}
 E = & \sum_{bonds} K_b (r - r_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \\
 & + \sum_{dihedrals} K_\phi (1 + \cos(n\phi - \delta)) + \\
 & + \sum_{nonbonded} (E_C + E_{LJ})
 \end{aligned} \quad (13)$$

В ней отсутствует слагаемое, описывающее энергию выхода атома из плоскости (improper dihedral).

В AMBER входят силовое поле, параметризованное для моделирования белков ff19SB [118], опубликованное в 2020 году как улучшение более ранних версий (ff14SB [53], ff99SB [119], ff94 [120]), а также силовые поля для моделирования нуклеиновых кислот bsc0 [121, 122] и bsc1 [123] и липидов lipids21 [124]. Кроме того, в семейство включено обобщенное силовое поле General AMBER Force Field (GAFF) [125] для малых органических молекул.

Потенциал GLYCAM из семейства AMBER для моделирования углеводов разрабатывается научной группой профессора Роберта Вудса из университета Джорджии. Первая версия GLYCAM93 [126] была опубликована в 1995 году. После этого выходило несколько обновлений, в которых вносились улучшения и уточнения [127–129]. В 2008 году вышло обновление GLYCAM06 [55], в котором силовое поле было полностью перестроено. В отличие от предыдущих версий, разработанных с намерением ввести минимально возможное количество параметров для углеводов в AMBER, новая версия полностью устраняет зависимость от силового поля и параметров AMBER. Кроме того, GLYCAM06 теперь может быть расширен на все классы молекул. Последующие улучшения маркировались суффиксом. Последняя версия GLYCAM06j, опубликованная в 2014 году, имеет совместимость с ff14SB (версией потенциала AMBER для белков).

Кроме группы Вуда адаптацией силового поля AMBER для моделирования углеводов занимаются и другие ученые. Еще в 1990 году первая версия AMBER была дополнена для моделирования олигосахаридов [130, 131]. В 2000 году под названием AMB99C вышла версия силового поля ff94, расширенная на класс углеводов [132], которая была доработана в 2009 году [133].

Модель OPLS

Силовое поле Optimized Potentials for Liquid Simulations (OPLS) разрабатывается группой профессора Уильяма Йоргенсена из Йельского университета с 1978 года, когда были опубликованы параметры (Transferable Intermolecular Potential Functions, TIP5) [134–137] для моделирования жидкого фтористого водорода, аммиака и метанола. В цитируемых работах впервые было проведено молекулярно-динамическое моделирование всех указанных жидкостей. Невалентные взаимодействия, которые в значительной степени определяют такие важные свойства, как плотность и энтальпия испарения жидкостей, были ключевой проблемой в то время. Эти ранние модели (теперь называемые OPLS-UA) рассматривали атомы водорода, связанные с алифатическим углеродом, как одну частицу.

Силовое поле изначально создавалось для моделирования жидкостей, содержащих органические молекулы. Первые расчеты белков были сделаны в 1985 году [69], когда к научной группе присоединился Джулиан Тирадо-Ривес, возглавивший разработку потенциала для пептидов и белков. Последующая работа над более чем 100 органическими жидкостями послужила основой для модели OPLS для органических систем, белков и нуклеиновых кислот [44]. Типы атомов и параметры валентных взаимодействий тогда брались из си-

лового поля AMBER 1984 года, а получившееся силовое поле было названо AMBER/OPLS. Как и в случае с другими потенциалами, позже была разработана полноатомная версия (OPLS-All Atom), где атомы водорода явно присутствовали в модели. Заряды атомов, как и параметры потенциала Леннарда-Джонса, все так же получали из моделирования чистых жидкостей [70]. Параметры для двугранных углов были разработаны по энергетическим профилям Hartree-Fock/6-31G* [138]. Параметры для ковалентных связей и углов между ними были стандартизированы, но в основном взяты из полноатомного силового поля AMBER 1986 года. Основными исключениями были параметры для алканов, взятые из CHARMM (параметры из статьи 1994 года [139], которые позже послужили основой для CHARMM22/CMAP [78, 79]), поскольку они приводили к значительному улучшению как структуры, так и энергетики этих соединений.

Общая форма для энергии выглядит следующим образом:

$$E = \sum_{bonds} K_b (r - r_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} [K_\phi (1 + \cos(\phi + f_1)) + K_{\phi_2} (1 - \cos(2\phi + f_2)) + K_{\phi_3} (1 + \cos(3\phi + f_3))] + \sum_{nonbonded} (E_C + E_{LJ}) \quad (14)$$

Первая попытка моделирования углеводов была сделана в 1997 году [54]. Обычно при разработке потенциалов для углеводов ученые концентрировались на оптимизации параметров колебания двугранных углов и воспроизведении структуры молекулы, а параметры невалентных взаимодействий брались из значений для простых спиртов и эфиров. Основная же идея силовых полей OPLS изначально состояла в точном воспроизведении термодинамических свойств чистых органических жидкостей, а следовательно, в аккуратном расчете параметров невалентных взаимодействий, от которых в большей степени зависят объемные свойства жидкости. Параметры валентных взаимодействий, в том числе параметры энергий колебания двугранных углов, могут быть оптимизированы для воспроизведения молекулярных структур и конформационных энергий из экспериментов и расчетов ab initio. В отличие от других силовых полей для углеводов параметры в OPLS были получены по результатам ab initio расчетов для полных гексопираноз. В 2002 году силовое поле было улучшено посредством ослабления электростатических взаимодействий 1–5 (взаимодействия между атомами, разделенными 4 ковалентными связями) и некоторых взаимодействий 1–6 (взаимодействия между атомами, разделенными 5 ковалентными связями), а также

репараметризации параметров вращения [140]. В 2017 году было показано, что комбинация OPLS с коррекцией парциальных зарядов $1.14 \cdot \text{CM1A}$ на атомах и коррекцией локализации заряда (LBCC) позволяет улучшить точность описания конденсированных сред [141, 142].

В 2022 году авторами статьи была исследована применимость новой версии силового поля для водных растворов сахарозы и продемонстрировано хорошее соответствие расчетных и экспериментальных данных по транспортным коэффициентам и плотности [143, 144].

ПОТЕНЦИАЛЫ ДЛЯ УГЛЕВОДОВ

Создание первых потенциалов

В предыдущем разделе была рассмотрена история развития наиболее популярных семейств силовых полей, каждое из которых включает параметры для моделирования углеводов. Потенциалы для углеводов начали создаваться во второй половине 1990-х годов, когда были опубликованы параметры для семейств AMBER (GLYCAM, 1995 год [126]) и OPLS (1997 год [54]). В то время рос интерес к исследованиям взаимодействия углеводов с белками. Основными методами изучения олигосахаридов (углеводов, содержащих от 2 до 10 моносахаридных остатков) и гликопротеинов (белки, ковалентно соединенные с гетероолигосахаридами) были экспериментальные методы, такие как ЯМР и флуоресцентная спектроскопия. Хотя ЯМР широко применяется к олигосахаридам, определить данным методом конформации гликозидных связей в этих системах крайне трудно.

В качестве альтернативного метода исследования рассматривался метод молекулярной динамики. Однако существовавшие на тот момент силовые поля (в том числе первые версии полей CHARMM, AMBER, GROMOS, OPLS) для исследования олигосахаридов не подходили. Одним из ограничений, присущих всем существовавшим на тот момент параметризациям углеводов, было отсутствие точных энергетических профилей вращения атомов вокруг ковалентных связей. Тем не менее тогда уже начали появляться первые результаты *ab initio* расчетов для гексопираноз, которые могли обеспечить хорошую основу для разработки и тестирования силовых полей. Это позволило начать параметризацию моделей для углеводов и их последующие уточнения, затронутые в предыдущем разделе статьи.

Недостатки современных потенциалов

В 2012 году в обзоре потенциалов для углеводов [145] группа Вудса указала на некоторые проблемы, связанные с разработкой точных и пере-

носимых параметров для сахаров. Например, распределение парциальных зарядов различных стереоизомеров одного и того же моносахарида может значительно отличаться. Несмотря на обнаруженные сложности, было показано, что ряд современных силовых полей для сахаров довольно точно отражает важные особенности конформационного поведения простых углеводов. Это значит, что современные силовые поля для углеводов могут работать хорошо в плане воспроизведения внутримолекулярных конформаций.

Кроме того, некоторые исследования показывают, что конфигурации и свойства углеводов могут быть воспроизведены не так хорошо. Например, при моделировании целлюлозы с помощью силового поля GLYCAM06 может происходить нефизическое перекручивание волокон, а для растворов β -D-глюкозы была показана нефизическая гиперагрегация молекул сахара [146, 147]. Последний эффект более явно показан в исследовании с расчетами осмотического давления [147], а также обнаружен в растворах сахарозы авторами обзора (рис. 3). Кроме того, термодинамические и транспортные свойства концентрированных растворов находятся в плохом согласии с экспериментом. Так, в 2015 году была проведена проверка силового поля $56A_{\text{CARBO}}$ для водных растворов глюкозы при температуре 313 K [148]. Было показано, что $56A_{\text{CARBO}}$, как и OPLS [54], существенно превышает плотность высококонцентрированных растворов (на 4% при концентрации 54%). Вязкость же при концентрации в 54% превышалась более, чем в 2 раза (для сравнения, OPLS превышал ее в 3 раза). Данные несоответствия приводят к необходимости дальнейшей разработки потенциалов для описания углеводов.

Способы улучшения потенциалов

Нефизичная гиперагрегация молекул в водном растворе в расчетах характерна не только для углеводов, но и, например, для поверхностно-активных веществ [149]. Такой эффект является одним из артефактов силовых полей, обнаруженных при увеличении временных масштабов молекулярно-динамического моделирования. Он объясняется тем, что параметризация силовых полей проводилась в первую очередь на стабильных молекулах и на коротких расчетах. Это привело к завышению констант невалентных взаимодействий для удержания структуры, что и послужило причиной гиперагрегации водорастворимых молекул.

Научная группа Аксментьева занимается систематическим уточнением леннард-джонсоновских параметров взаимодействия, сопоставляя рассчитанное осмотическое давление растворов с экспериментальными данными [150–154]. Такое

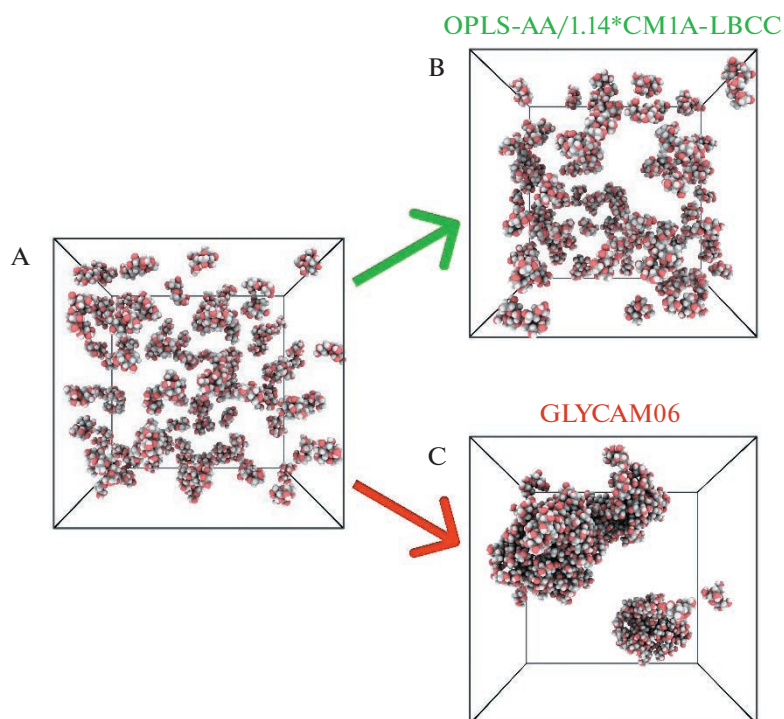


Рис. 3. Расчетная ячейка водного раствора сахарозы, включающая 100 молекул сахарозы и 17100 молекул воды (на рисунке не показаны), что соответствует содержанию сахара в растворе 10 мас.%. Ячейка А отвечает начальной конфигурации, ячейки В и С – состояниям системы через 15 нс эволюции в NPT-ансамбле в силовых полях OPLS-AA/1.14*CM1A-LBCC [141, 142] (ячейка В) и GLYCAM06 [55] (ячейка С). Расчеты проведены авторами обзора. В ячейке С наблюдается 2 крупных кластера сахарозы, не свойственные для данной области фазовой диаграммы.

уточнение позволяет привести результаты моделирования к хорошему согласию с экспериментальными данными и избежать гиперагрегации молекул в растворе.

Аналогичное уточнение леннард-джонсоновских взаимодействий для углеводов проводится группой Элкока [147, 155]. Оптимальные параметры подбираются из согласования с экспериментальными значениями осмотического коэффициента, который определяется как отношение осмотического давления к таковому для идеального раствора. Он интерпретируется таким образом: осмотический коэффициент выше единицы означает сильное отталкивающее взаимодействие растворенных веществ, а осмотический коэффициент ниже единицы указывает на сильное притягивающее взаимодействие между растворенными молекулами. В работе [147] показано, что силовые поля GLYCAM06 и CHARMM36 занижают осмотический коэффициент, что является следствием гиперагрегации молекул. Кроме того, авторы обнаружили, что новые параметры могут улучшить описание олиго- и полисахаридов в водном растворе на примере декстрана. Был рассчитан радиус инерции неразветвленного декстрана, состоящего из 55 глюкозных мономеров, для оригинального и уточненного набора па-

раметров. Показано, что силовое поле GLYCAM06 приводит к занижению размера молекулы почти в 2.5 раза, в то время как уточненный набор параметров занижает его менее, чем на 5%.

Другой подход к улучшению силовых полей связан с пересмотром электростатических взаимодействий. Он основан на использовании различных моделей для оценки атомных парциальных зарядов или на рассмотрении поляризуемых силовых полей. Например, было показано, что уменьшение парциальных зарядов на 20–30% позволяет улучшить воспроизведение экспериментальных данных по диффузии и вязкости ионных жидкостей [156, 157]. Физически такое масштабирование можно объяснить смещением электронной плотности от аниона к катиону, что приводит к уменьшению эффективных зарядов ионов. Сравнение методов масштабирования зарядов и использования поляризуемых силовых полей проведено, например, в обзоре [158]. Показано, что использование поляризуемых силовых полей приводит к лучшим результатам, но требует существенно больших вычислительных ресурсов.

Таким образом, величина зарядов оказывает прямое влияние на свойства, а масштабирование атомных зарядов является простым и малозатратным способом повышения точности описания

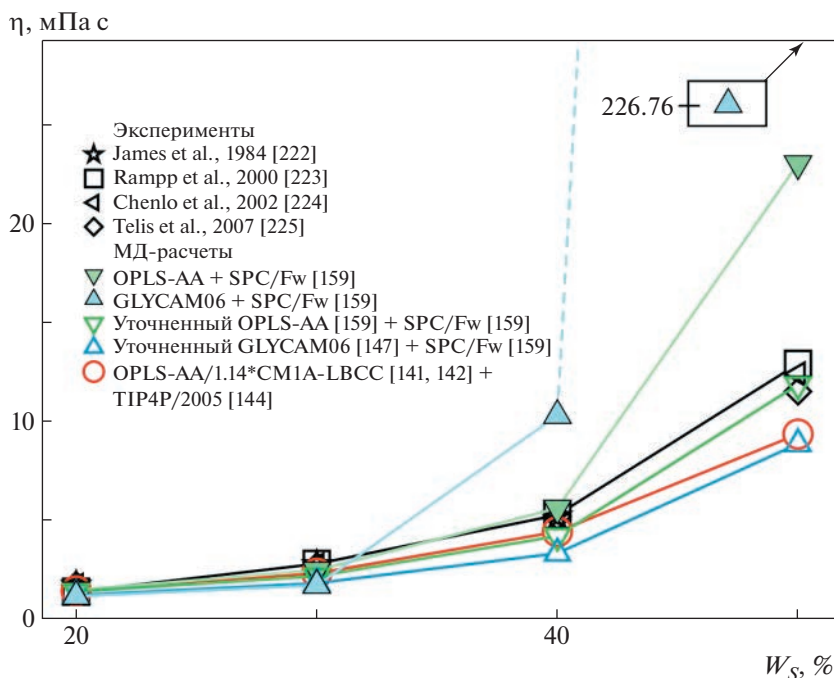


Рис. 4. Концентрационная зависимость вязкости водного раствора сахарозы при температуре 293 К. Черными символами показаны экспериментальные значения [226–229], сплошными треугольниками и перевернутыми треугольниками – рассчитанные в работе [159] значения для силовых полей OPLS и GLYCAM06, светлыми треугольниками и перевернутыми треугольниками – рассчитанные значения для уточненных силовых полей OPLS [159] и GLYCAM06 [147], а кругами – рассчитанные в работе [144] значения для силового поля OPLS-AA/1.14*CM1A-LBCC [141, 142].

исследуемой системы, особенно транспортных свойств (таких, как вязкость и коэффициент диффузии). Этот метод применяется и для водных растворов углеводов [148, 159]. Так, в работе 2015 года показано, что домножение зарядов на коэффициент 0.8 (подобранный перебором) приводит к хорошему согласованию с экспериментальными данными по вязкости и плотности водных растворов глюкозы [148]. При концентрации глюкозы 70% отклонение коэффициента вязкости от экспериментального значения составляет 35% (против 650% в моделировании с исходными атомными зарядами глюкозы силового поля GROMOS56A_{CARBO}). Однако такое улучшение было достигнуто только там, где потенциал GROMOS 56A_{CARBO} не смог предсказать значения плотности и вязкости. Таким образом, для достижения высокой точности в широком диапазоне концентраций может потребоваться использование поляризуемых силовых полей.

Первый подход улучшает описание термодинамических свойств, таких как осмотическое давление. Однако изменение леннард-джонсоновских энергетических параметров оказывает меньшее влияние на транспортные свойства по сравнению с изменением зарядов атомов. В то же время в работе [148] не были исследованы термодинамические свойства. Это послужило основанием для более широкого исследования влияния

параметров невалентных взаимодействий на свойства растворов, опубликованного в 2018 году [159]. В работе было продемонстрировано, что масштабирование зарядов недостаточно для воспроизведения одновременно термодинамических и транспортных свойств концентрированных водных растворов сахаров. Масштабирование же только дисперсионных взаимодействий приводит к хорошему согласию с экспериментальными данными по транспортным, а также по термодинамическим свойствам. Наилучшие результаты достигаются при одновременном масштабировании зарядов и леннард-джонсоновских энергетических параметров. В работе показано, что домножение дисперсионных взаимодействий в силовом поле OPLS на 0.8, а зарядов атомов на 0.95 приводит к лучшему согласию с экспериментальными результатами по ряду свойств.

Описанные выше уточнения силовых полей демонстрируют существенно лучшее согласие результатов молекулярно-динамического моделирования углеводов с экспериментальными данными в сравнении с предыдущими версиями силовых полей. Кроме того, авторы обзора исследовали применимость последней версии силового поля OPLS [141, 142] для моделирования водных растворов сахарозы при концентрациях от 10 до 50% и показали, что оно работает хуже, чем уточненные версии полей OPLS и GLYCAM06 (рис. 4).

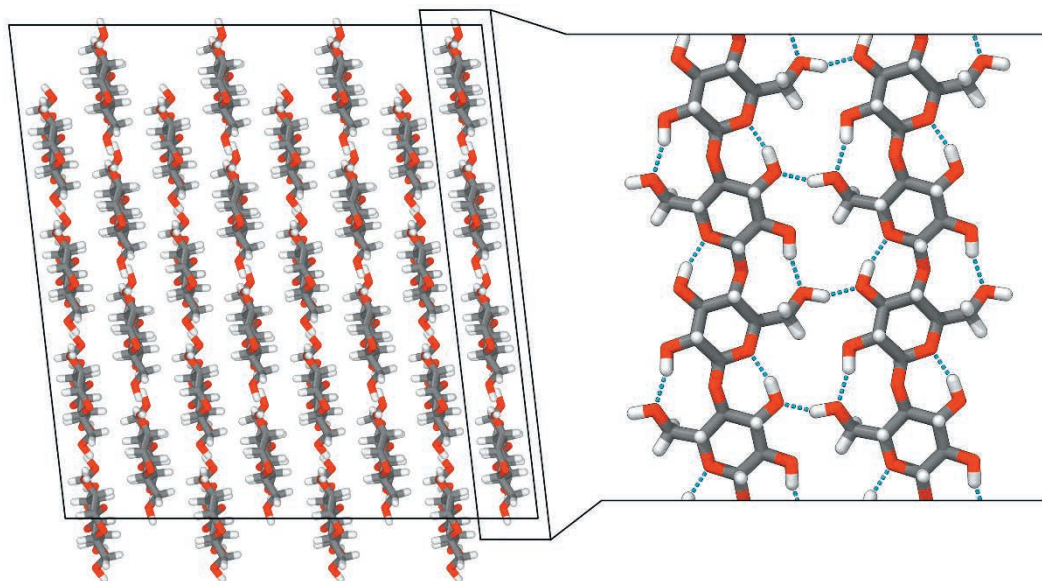


Рис. 5. 3D-модель кристалла целлюлозы I β , полученная в программе Cellulose-Builder [230], (слева) и две соседние целлюлозные цепочки в этой кристаллической модификации с отмеченными водородными связями (справа).

МОДЕЛИРОВАНИЕ ЦЕЛЛЮЛОЗЫ

Целлюлоза является не только самым широко распространенным полисахаридом, но и самым распространенным биополимером. Она имеет несколько полиморфных модификаций, обнаруженных экспериментальными методами [15, 160, 161]. Целлюлоза I является нативной формой и присутствует в двух модификациях I α [162] и I β [163] (рис. 5) в клеточных стенках растений. Целлюлоза II является более стабильной модификацией из-за антипараллельного расположения молекул. Целлюлоза III может быть получена обратимым образом из целлюлозы I (III $_I$) [164] и целлюлозы II (III $_II$) [165]. Из них же могут быть получены целлюлоза IV $_I$ и IV $_II$ [166, 167], которые были обнаружены в клеточных стенках растений [168]. Все эти полиморфные модификации определяются по конформационным различиям, различным устройствам упаковки и разным наборам водородных связей.

В последнее десятилетие наблюдается рост интереса к целлюлозе, а именно к ее нановолокнам и нанокристаллам [169–171]. Они обладают уникальными механическими и химическими свойствами, низкой стоимостью и биоразлагаемостью, а композиты на их основе имеют твердость, сравнимую с твердостью алюминиевых сплавов [172]. Наноцеллюлоза имеет приложения в пищевой промышленности [173, 174] и в медицине [175–177]. Кроме того, химические модификации целлюлозы являются перспективными материалами для создания ион-селективных мембран [13, 12].

Молекулярно-динамические исследования внесли большой вклад в изучение целлюлозы.

Еще в 1995 году А.Р. Heiner с коллегами провели моделирование двух модификаций нативной целлюлозы I α и I β [178]. Они показали более высокую стабильность целлюлозы I β , что подтверждается экспериментальными данными. В работе это обстоятельство объясняется различным набором водородных связей и разным поворотом остатков глюкозы в кристаллических модификациях.

В 2014 году Wu с коллегами исследовали растяжение кристалла целлюлозы I β в трех направлениях при различных скоростях деформации [179]. Это позволило им рассчитать модули упругости и коэффициенты Пуассона. Кроме того, они с помощью реактивного силового поля [180] промоделировали разрыв ковалентных связей при растяжении и вычислили предельные напряжения вдоль трех осей. В данном исследовании впервые был вычислен полный набор предельных свойств целлюлозы I β , что можно рассматривать как первый шаг к пониманию деформации целлюлозных нанокристаллов.

Ту же модификацию исследовали М. Bergenstråhle с сотрудниками [181]. Они показали реконструкцию кристалла целлюлозы I β при нагревании выше 450 К, которая сопровождается изменением структуры водородных связей и ориентацией гидроксиметильных групп. Наблюдения хорошо согласуются с экспериментальными данными и могут улучшить понимание пути перехода между низкотемпературной и высокотемпературной структурами.

Н. Miyamoto с сотрудниками в своих работах исследуют механизм образования регенерированной целлюлозы, получаемой из растворов

[182–185]. Так, изучая структурную реорганизацию двух различных молекулярных листов, полученных из кристалла целлюлозы II, они показали, что молекулярный лист, образованный ван-дер-ваальсовыми силами, оказывается стабильным в воде, а лист, образованный водородными связями, перестраивается [182]. Это свидетельствует о том, что первая структура может быть исходной структурой кристалла, формирующегося в водном растворе. Кроме того, авторы обнаружили, что в бензольном растворе наблюдается обратная ситуация и молекулярный лист, образованный водородными связями, является стабильным.

Для описания целлюлозы на большем масштабе используются крупнозернистые силовые поля, такие как MARTINI [113, 115]. В 2015 году был разработан набор параметров специально для моделирования кристаллической целлюлозы [186]. Показано, что модель хорошо описывает механические свойства целлюлозы I β , а также способна воспроизводить переход к целлюлозе III β . Каждый мономер целлюлозы в данной модели описывается всего тремя частицами, что существенно повышает скорость расчетов по сравнению с полноатомными силовыми полями. Кроме того, авторы утверждают, что способ параметризации силового поля MARTINI может быть использован и для других кристаллических биополимеров.

Помимо поведения целлюлозы, ученых интересуют и взаимодействия целлюлозы с другими веществами, например, растворение в ионных жидкостях. Liu с соавторами [187] проанализировали механизм растворения олигомеров целлюлозы в растворе 1-этил-3-метилимидазолия ацетата. При параметризации ионной жидкости использовалось силовое поле GAFF из семейства AMBER, а для целлюлозы – версия GLYCAM 1995 года (GLYCAM93 [126]), поскольку более современные версии не имеют совместимости с семейством AMBER.

Важным направлением является изучение взаимодействия целлюлозы с клеточными мембранами, которым занимается научная группа А.А. Гуртовенко [188–191]. Такие исследования необходимы для биомедицинских приложений целлюлозы, например для тканевой инженерии и костной имплантации. Кроме того, они позволяют определить механизм взаимодействия раневых повязок с кожей и усовершенствовать их для более быстрого заживления ран. Для этих целей, например, было проведено моделирование взаимодействия кристалла целлюлозы с модельными мембранами, состав которых соответствует составу липидного матрикса рогового слоя [191]. В работе были выявлены два механизма взаимодействия целлюлозы с кожей и показано, что механизм зависит от доли протонированных жирных кислот в бислое рогового слоя. В рассматриваемых

работах используются силовые поля семейства CHARMM: CHARMM35 [94, 192] для целлюлозы и CHARMM36 [81] для липидов.

ПРОИЗВОДИТЕЛЬНОСТЬ ВЫЧИСЛЕНИЙ МОДЕЛЕЙ ПОЛИСАХАРИДОВ

Разработка высоко оптимизированных программных пакетов для молекулярного моделирования, обеспечивающих сверхдлинные расчеты траекторий атомов и экстремальные размеры систем, является одним из важных векторов развития методов высокопроизводительных вычислений [33, 35, 36, 193]. Вскоре после внедрения технологии Nvidia CUDA в 2007 году появились гибридные алгоритмы МД, показавшие свою многообещающую производительность. В настоящее время архитектуры с графическими ускорителями (GPU) обеспечивают наиболее эффективный и доступный способ проведения исследований в области МД [194–196] и делают возможным выполнение различных прикладных исследований [47, 197–200].

Появление суперкомпьютеров с распределенной памятью стимулировало разработку параллельных кодов для молекулярного моделирования. Среди прочего, Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [201], GROningen MACHine for Chemical Simulations (GROMACS) [202] и AMBER [203] представляют собой три программных пакета, которые обладают широким функционалом и повсеместно используются в настоящее время. С момента появления возможности проведения вычислений на графических ускорителях все пакеты были дополнены возможностями расчета на них [204–209]. Новые пакеты для молекулярной динамики, такие как HOOMD [210, 211] и OpenMM [212, 213] используют алгоритмы, ориентированные на GPU, а центральный процессор (CPU) нужен только для обмена данными.

Перенос алгоритмов МД на гибридные архитектуры – трудоемкая задача, требующая больших усилий. Недавно такой опыт был описан для систем высокопроизводительных вычислений на базе процессоров Sunway [214–217]. Другой пример – портирование алгоритмов на архитектуру Adapteva EpiPhany [218]. Приведенные примеры уникальны, поскольку данные архитектуры запрограммированы без установленной программной экосистемы.

На настоящий момент лучшая программная экосистема для вычислений на GPU основана на технологии Nvidia CUDA. Она имеет развитую поддержку драйверов и среды выполнения, отличные инструменты отладки и профилирования, подробную документацию и примеры. Практически каждый алгоритм, обладающий вы-

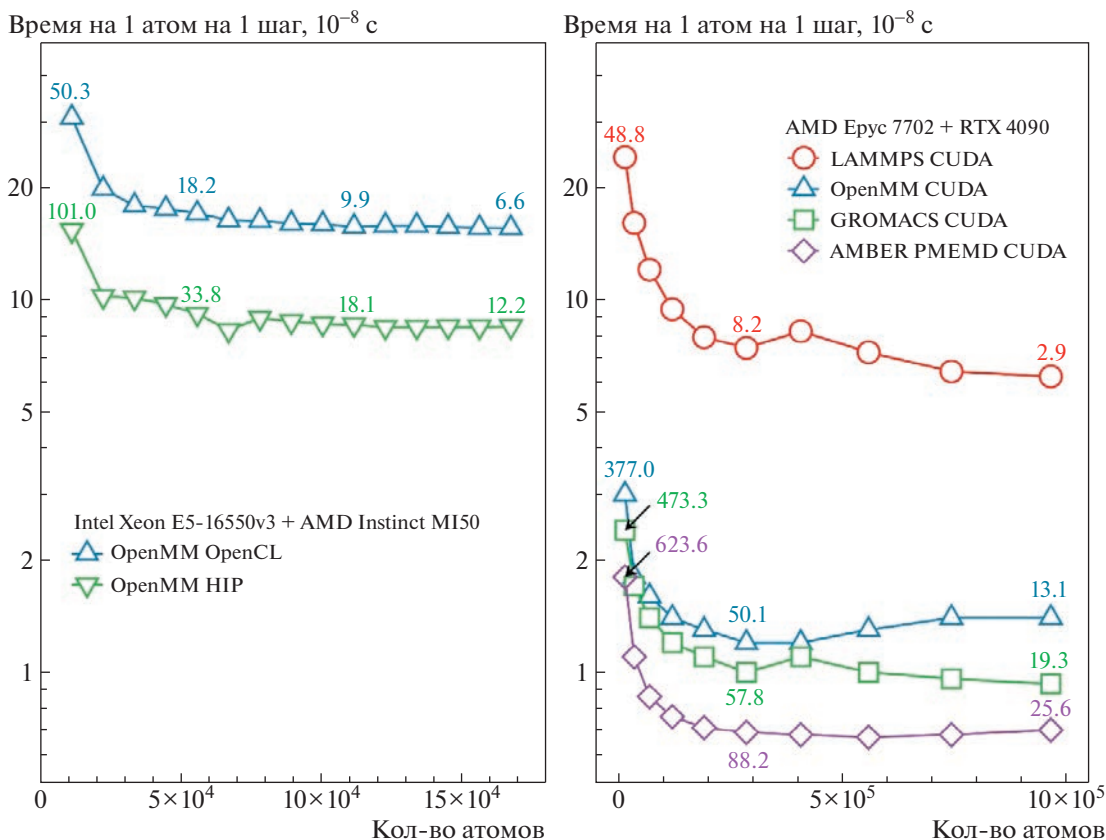


Рис. 6. Время расчета одного МД-шага для водного раствора сахарозы в расчете на один атом для разных конфигураций узлов. Числами рядом с точками показана скорость расчета всей системы в наносекундах в день. В левой части рисунка приведены результаты для OpenMM с использованием OpenCL и HIP. Применение технологии HIP обеспечивает ускорение расчетов в 2 раза. В правой части рисунка приведены результаты для четырех разных пакетов: LAMMPS, OpenMM, GROMACS и AMBER. Для всех расчетов шаг по времени составлял 2 фемтосекунды.

сокой вычислительной сложностью и возможностью распараллеливания, был реализован на CUDA. К сожалению, CUDA – технология с закрытым исходным кодом, что ограничивает переносимость. Стандарт OpenCL призван стать кросс-платформенной альтернативой CUDA, но такая переносимость достигается за счет гораздо более сложной настройки производительности для конкретного типа графического процессора.

Напротив, технология AMD HIP предназначена для разработки программного обеспечения портативных графических процессоров по аналогии с CUDA как для новых архитектур графических процессоров AMD, так и для графических процессоров Nvidia (без ущерба для производительности). HIP является частью платформы AMD Radeon Open Compute (ROCm) с открытым исходным кодом, анонсированной в 2015 году компанией AMD.

Авторами обзора был проведен анализ производительности расчета водного раствора сахарозы при разных размерах систем в пакетах

LAMMPS, GROMACS, AMBER, а также OpenMM с технологиями OpenCL, HIP и CUDA.

Для тестов были взяты силовое поле OPLS-AA/1.14*CM1A-LBCC и модель воды TIP4P/2005 [219]. Моделирование выполняли в каноническом ансамбле с термостатом Нозе–Гувера [220–222] с шагом по времени 2 фемтосекунды. В случае AMBER использовали термостат Андерсена [223]. Применяли периодические граничные условия.

Тесты проводили на 1 узле кластера Десмос в Объединенном институте высоких температур РАН: CPU Intel Xeon E5-1650v3, RAM 32 GB; GPU AMD Instinct MI50. Также для тестов была задействована 1 рабочая станция в Московском физико-техническом институте (Soft cluster): CPU AMD Epyc 7702, RAM 256 GB; GPU NVIDIA RTX 4090. Результаты представлены на рис. 6. По горизонтальной оси отложено количество атомов в модели включая атомы воды, по вертикальной – требуемое время расчета на 1 атом на 1 шаг интегрирования. Выход времени расчета на плато соответствует оптимальному режиму эксплуатации

узла. Числами рядом с точками указаны достижимые производительности в нс/день.

Для узла Десмоса (левая часть рис. 6) переход от технологии OpenCL к HIP в OpenMM для карт AMD Instinct MI50 дает ускорение в 2 раза. Результаты тестов производительности на Soft cluster изображены на правой части рис. 6. Мы подтверждаем то, что производительности OpenMM и GROMACS превосходят LAMMPS [224]. В то же время LAMMPS обладает гораздо большей универсальностью, позволяет использовать огромное количество методов расчета свойств и моделей взаимодействия, включая потенциалы машинного обучения. Реализация pmemd.cuda, доступная в рамках академической лицензии, лидирует по производительности.

Для расчета свойств растворов требуются траектории длиной порядка 100 нс [144]. Исследования конформационной динамики молекул сахарозы и взаимодействия сахаров с белками требуют времен расчета порядка 1000 нс (1 мкс). Таким образом, на 1 узле гибридного суперкомпьютера можно за сутки предсказать свойства раствора с заданной концентрацией.

ЗАКЛЮЧЕНИЕ

В настоящей работе рассмотрены основные семейства потенциалов для атомистического моделирования углеводов: CHARMM, GROMOS, AMBER, OPLS. Изложен их исторический путь от самых первых потенциалов, которые позволяли описывать конформации гексопираноз, до современных потенциалов, с хорошей точностью воспроизводящих как термодинамические, так и динамические свойства растворов ди- и полисахаридов.

На текущий момент для полисахаридов можно выделить следующие потенциалы: GLYCAM06, CHARMM35 и модификацию OPLS-AA/1.14*CM1A-LBCC. Первые два подходят для гликопротеинов, но плохо описывают водные растворы сахаров. При их использовании наблюдается отклонение транспортных свойств от эксперимента и гипер-агрегация молекул сахара. Существует методика уточнения потенциалов типа GLYCAM06 и CHARMM35, в которой производится репараметризация невалентных взаимодействий. Также с задачами моделирования углеводов отлично справляется потенциал OPLS-AA/1.14*CM1A-LBCC, в котором для каждой молекулы сахара осуществляется параметризация парциальных зарядов.

Развитие данной области было бы невозможно без роста суперкомпьютерных мощностей. Благодаря развитию архитектур с графическими ускорителями и соответствующих программных пакетов (LAMMPS, GROMACS, OpenMM, AMBER и т.п.) становятся доступными расчеты в десятки

микросекунд и модели до триллионов атомов. Такие масштабы моделей позволяют решать задачи по прогнозированию свойств новых соединений и исследованию корреляции структура–свойство.

Важно отметить, что в последнее десятилетие к оптимизации параметров для потенциалов стали присоединяться различные научные группы, не связанные напрямую с разработчиками крупных семейств потенциалов. Доступность методик и участие большого числа специалистов по всему миру существенно ускорят получение надежных параметров, подходящих для точного предсказания свойств большого числа углеводных соединений.

Авторские расчеты, приведенные в данной работе, выполнены на суперкомпьютере Московского физико-технического института и на кластере Лаборатории многомасштабного моделирования в физике мягкой материи Московского физико-технического института Soft cluster, а также на суперкомпьютере Десмос Объединенного института высоких температур РАН.

Настоящая работа выполнена в рамках Программы стратегического академического лидерства “Приоритет-2030” (соглашение 075-02-2021-1316 от 30.09.2021).

СПИСОК ЛИТЕРАТУРЫ

1. *Venturoli D. and Rippe B.*, American Journal of Physiology-Renal physiology. 2005. V. 288. P. F605.
2. *Yuan Y., Yang Y., Tian Y., Park J., Dai A., Roberts R.M., Liu Y., and Han X.*, Scientific Reports. 2016. V. 6. P. 34476.
3. *Zhao G., Liu X., Zhu K., and He X.*, Advanced Healthcare Materials. 2017. V. 6. P. 1700988.
4. *Mao Y., Han X., and Zhang Y.*, International Journal of Heat and Mass Transfer. 2018. V. 127. P. 319.
5. *Shahid S., Hasan I., Ahmad F., Hassan M.I., and Islam A.*, Biomolecules. 2019. V. 9. P. 477.
6. *Gore M., Narvekar A., Bhagwat A., Jain R., and Dandekar P.*, Journal of Materials Chemistry B. 2022. V. 10. P. 143.
7. *Vatanpour V., Pasaoglu M.E., Barzegar H., Teber O.O., Kaya R., Bastug M., Khataee A., and Koyuncu I.*, Chemosphere. 2022. V. 295. P. 133914.
8. *Li S., Wang X., Guo Y., Hu J., Lin S., Tu Y., Chen L., Ni Y., and Huang L.*, Journal of Cleaner Production. 2022. V. 333. P. 130171.
9. *Spiridonov V.V., Panova I.G., Sybachin A.V., Kuznetsov V.V., Afanasov M.I., Alekhina Y.A., Melik-Nubarov N.S., and Yaroslavov A.A.*, Polymer Science, Series A. 2019. V. 61. P. 296.
10. *Zashikhina N.N., Yudin D.V., Tarasenko I.I., Osipova O.M., and Korzhikova-Vlakh E.G.*, Polymer Science, Series A. 2020. V. 62. P. 43.
11. *Shibaev A.V., Doroganov A.P., Larin D.E., Smirnova M.E., Cherkaev G.V., Kabaeva N.M., Kitaeva D.K., Buyanov-*

- skaya A.G., and Philippova O.E., *Polymer Science, Series A*. 2021. V. 63. P. 24.
12. Lander S., Vagin M., Gueskine V., Erlandsson J., Boisard Y., Korhonen L., Berggren M., Wågberg L., and Crispin X., *Advanced Energy and Sustainability Research*. 2022. V. 3. P. 2200016.
 13. Yang H., Edberg J., Gueskine V., Vagin M., Say M.G., Erlandsson J., Wågberg L., Engquist I., and Berggren M., *Carbohydrate Polymers*. 2022. V. 278. P. 118938.
 14. Demina T.S., Akopova T.A., and Zelenetsky A.N., *Polymer Science, Series C*. 2021. V. 63. P. 219.
 15. Bogdanova O.I. and Chvalun S.N., *Polymer Science, Series A*. 2016. V. 58. P. 629.
 16. Ewen J., Gattinoni C., Thakkar F., Morgan N., Spikes H., and Dini D., *Materials*. 2016. V. 9. P. 651.
 17. Glova A.D., Volgin I.V., Nazarychev V.M., Larin S.V., Lyulin S.V., and Gurtovenko A.A., *RSC Advances*. 2019. V. 9. P. 38834.
 18. Orekhov N., Ostroumova G., and Stegailov V., *Carbon*. 2020. V. 170. P. 606.
 19. Nazarychev V.M., Glova A.D., Volgin I.V., Larin S.V., Lyulin A.V., Lyulin S.V., and Gurtovenko A.A., *International Journal of Heat and Mass Transfer*. 2021. V. 165. P. 120639.
 20. Bakulin I., Kondratyuk N., Lankin A., and Norman G., *The Journal of Chemical Physics*. 2021. V. 155. P. 154501. <https://doi.org/10.1063/5.0059337>
 21. Vaganova M., Nesterova I., Kanygin Y., Kazennov A., and Khlyupin A., *Chemical Engineering Science*. 2022. V. 250. P. 117383.
 22. Mackerell A.D., *Journal of Computational Chemistry*. 2004. V. 25. P. 1584.
 23. Perez S. and Makshakova O., *Chemical Reviews*. 2022. V. 122. P. 15914.
 24. Salzner U. and Schleyer P.v.R., *The Journal of Organic Chemistry*. 1994. V. 59. P. 2138.
 25. Glennon T.M., Zheng Y.-J., Le Grand S.M., Shutzberg B.A., and Merz K.M., *Journal of Computational Chemistry*. 1994. V. 15. P. 1019.
 26. Barrows S.E., Dulles F.J., Cramer C.J., French A.D., and Truhlar D.G., *Carbohydrate Research*. 1995. V. 276. P. 219.
 27. Brown J.W. and Wladkowski B.D., *Journal of the American Chemical Society*. 1996. V. 118. P. 1190.
 28. Alder B.J. and Wainwright T.E., *The Journal of Chemical Physics*. 1957. V. 27. P. 1208.
 29. Gibson J.B., Goland A.N., Milgram M., and Vineyard G.H., *Phys. Rev.* 1960. V. 120. P. 1229.
 30. Rahman A., *Phys. Rev.* 1964. V. 136. P. A405.
 31. Singharoy A., Maffeo C., Delgado-Magnero K.H., Swainsbury D.J., Sener M., Kleinekathöfer U., Vant J.W., Nguyen J., Hitchcock A., Isralewitz B., Teo I., Chandler D.E., Stone J.E., Phillips J.C., Pogorelov T.V., Mal-lus M.I., Chipot C., Luthey-Schulten Z., Teleman D.P., Hunter C.N., Tajkhorshid E., Aksimentiev A., and Schulten K., *Cell*. 2019. V. 179. P. 1098.
 32. Nguyen-Cong K., Willman J.T., Moore S.G., Belonoshko A.B., Gayatri R., Weinberg E., Wood M.A., Thompson A.P., and Oleynik I.I., in *Proceedings of the international conference for high performance computing, networking, storage and analysis, SC '21 (2021)*.
 33. Shaw D.E., Deneroff M.M., Dror R.O., Kuskin J.S., Larson R.H., Salmon J.K., Young C., Batson B., Bowers K.J., Chao J.C., Eastwood M.P., Gagliardo J., Grossman J.P., Ho C.R., Ierardi D.J., Kolossv'ary I., Klepeis J.L., Layman T., McLeavey C., Moraes M.A., Mueller R., Priest E.C., Shan Y., Spengler J., Theobald M., Towles B., and Wang S.C., *Communications of the ACM*. 2008. V. 51. P. 91.
 34. Shaw D.E., Dror R.O., Salmon J.K., Grossman J.P., Mackenzie K.M., Bank J.A., Young C., Deneroff M.M., Batson B., Bowers K.J., Chow E., Eastwood M.P., Ierardi D.J., Klepeis J.L., Kuskin J.S., Larson R.H., Lindorff-Larsen K., Maragakis P., Moraes M.A., Piana S., Shan Y., and Towles B., in *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis (Nov. 2009)*. P. 1–11.
 35. Shaw D.E., Grossman J., Bank J.A., Batson B., Butts J.A., Chao J.C., Deneroff M.M., Dror R.O., Even A., Fenton C.H., Forte A., Gagliardo J., Gill G., Greskamp B., Ho C.R., Ierardi D.J., Iserovich L., Kuskin J.S., Larson R.H., Layman T., Lee L.-S., Lerer A.K., Li C., Killebrew D., Mackenzie K.M., Mok S.Y.-H., Moraes M.A., Mueller R., Nociolo L.J., Peticolas J.L., Quan T., Ramot D., Salmon J.K., Scarpazza D.P., Schafer U.B., Siddique N., Snyder C.W., Spengler J., Tang P.T.P., Theobald M., Toma H., Towles B., Vitale B., Wang S.C., and Young C., in *SC14: International Conference for High Performance Computing, Networking, Storage and Analysis (Nov. 2014)*. P. 41–53.
 36. Shaw D.E., Adams P.J., Azaria A., Bank J.A., Batson B., Bell A., Bergdorf M., Bhatt J., Butts J.A., Correia T., Dirks R.M., Dror R.O., Eastwood M.P., Edwards B., Even A., Feldmann P., Fenn M., Fenton C.H., Forte A., Gagliardo J., Gill G., Gorlatova M., Greskamp B., Grossman J., Gullingsrud J., Harper A., Hasenplough W., Heily M., Heshmat B.C., Hunt J., Ierardi D.J., Iserovich L., Jackson B.L., Johnson N.P., Kirk M.M., Klepeis J.L., Kuskin J.S., Mackenzie K.M., Mader R.J., McGowen R., McLaughlin A., Moraes M.A., Nasr M.H., Nociolo L.J., O'Donnell L., Parker A., Peticolas J.L., Pocina G., Pre-descu C., Quan T., Salmon J.K., Schwink C., Shim K.S., Siddique N., Spengler J., Szalay T., Tabladillo R., Tartler R., Taube A.G., Theobald M., Towles B., Vick W., Wang S.C., Wazłowski M., Weingarten M.J., Williams J.M., and Yuh K.A., in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (Nov. 2021)*. P. 1–11.
 37. Swope W.C., Andersen H.C., Berens P.H., and Wilson K.R., *The Journal of Chemical Physics*. 1982. V. 76. P. 637.
 38. Hockney R. and Eastwood J., *Computer Simulation Using Particles*, 0th ed. (CRC Press, Mar. 2021).
 39. Lorentz H.A., *Ann. Phys.* 1881. V. 248. P. 127.
 40. Berthelot D., *Comptes rendus hebdomadaires des s'eances de l'Acad'emie des Sciences*. 1898. V. 126. P. 1703.
 41. Weiner S.J., Kollman P.A., Case D.A., Singh U.C., Ghio C., Alagona G., Profeta S., and Weiner P., *J. Am. Chem. Soc.* 1984. V. 106. P. 765.
 42. Brooks B.R., Bruccoleri R.E., Olafson B.D., States D.J., Swaminathan S., and M. Karplus, *J. Comput. Chem.* 1983. V. 4. P. 187.
 43. Zhu X., Lopes P.E.M., and MacKerell A.D., *WIREs Comput Mol Sci*. 2012. V. 2. P. 167.
 44. Jorgensen W.L. and Tirado-Rives J., *Journal of the American Chemical Society*. 1988. V. 110. P. 1657.

45. *Van Gunsteren W.F. and Karplus M.*, *Macromolecules*. 1982. V. 15. P. 1528.
46. *Hermans J., Berendsen H.J.C., Van Gunsteren W.F., and Postma J.P.M.*, *Biopolymers*. 1984. V. 23. P. 1513.
47. *Kondratyuk N.D. and Pisarev V.V.*, *Fluid Phase Equilibria*. 2019. V. 498. P. 151.
48. *Kondratyuk N.D.*, *Journal of Physics: Conference Series* 1385, 012048 (2019).
49. *Kondratyuk N.D., Pisarev V.V., and Ewen J.P.* *The Journal of Chemical Physics*. 2020. V. 153. P. 154502, (2020).
<https://doi.org/10.1063/5.0028393>
50. *Rappe A.K., Casewit C.J., Colwell K.S., Goddard W.A., and Skiff W.M.*, *Journal of the American Chemical Society*. 1992. V. 114. P. 10024.
51. *Sun H.*, *The Journal of Physical Chemistry B*. 1998. V. 102. P. 7338.
52. *Halgren T.A.*, *Journal of Computational Chemistry*. 1999. V. 20. P. 730.
53. *Maier J.A., Martinez C., Kasavajhala K., Wickstrom L., Hauser K.E., and Simmerling C.* *Journal of Chemical Theory and Computation*. 2015. V. 11. P. 3696.
54. *Damm W., Frontera A., Tirado-Rives J., and Jorgensen W.L.*, *Journal of Computational Chemistry*. 1997. V. 18. P. 1955.
55. *Kirschner K.N., Yongye A.B., Tschampel S.M., Gonz'alez-Outeirin'õ J., Daniels C.R., Foley B.L., and Woods R.J.*, *Journal of Computational Chemistry*. 2008. V. 29. P. 622.
56. *Shi Y., Xia Z., Zhang J., Best R., Wu C., Ponder J.W., and Ren P.*, *Journal of Chemical Theory and Computation*. 2013. V. 9. P. 4046.
57. *Ponder J.W., Wu C., Ren P., Pande V.S., Chodera J.D., Schnieders M.J., Haque I., Mobley D.L., Lambrecht D.S., DiStasio R.A., HeadGordon M., Clark G.N.I., Johnson M.E., and Head-Gordon T.*, *The Journal of Physical Chemistry B*. 2010. V. 114. P. 2549.
58. *Lemkul J.A., Huang J., Roux B., and MacKerell A.D.*, *Chemical Reviews*. 2016. V. 116. P. 4983.
59. *Lewis-Atwell T., Townsend P.A., and Grayson M.N.*, *Tetrahedron*. 2021. V. 79. P. 131865.
60. *Allinger N.L.*, *Journal of the American Chemical Society*. 1977. V. 99. P. 8127.
61. *Allinger N.L., Yuh Y.H., and Lii J.H.*, *Journal of the American Chemical Society*. 1989. V. 111. P. 8551.
62. *Martin M.G.*, *Fluid Phase Equilibria*. 2006. V. 248. P. 50.
63. *Martin M.G. and Siepmann J.I.*, *The Journal of Physical Chemistry B*. 1998. V. 102. P. 2569.
64. *Martin M.G. and Siepmann J.I.*, *The Journal of Physical Chemistry B*. 1999. V. 103. P. 4508.
65. *Wick C.D., Martin M.G., and Siepmann J.I.*, *The Journal of Physical Chemistry B*. 2000. V. 104. P. 8008.
66. *Chen B., Potoff J.J., and Siepmann J.I.*, *The Journal of Physical Chemistry B*. 2001. V. 105. P. 3093.
67. *Lubna N., Kamath G., Potoff J.J., Rai N., and Siepmann J.I.*, *The Journal of Physical Chemistry B*. 2005. V. 109. P. 24100.
68. *MacKerell A.D., Bashford D., Bellott M., Dunbrack R.L., Evanseck J.D., Field M.J., Fischer S., Gao J., Guo H., Ha S., Joseph-McCarthy D., Kuchnir L., Kuczera K., Lau F.T.K., Mattos C., Michnick S., Ngo T., Nguyen D.T., Prodhom B., Reiher W.E., Roux B., Schlenkrich M., Smith J.C., Stote R., Straub J., Watanabe M., Wi'orkiewicz-Kuczera J., Yin D., and Karplus M.*, *J. Phys. Chem. B*. 1998. V. 102. P. 3586.
69. *Jorgensen W.L. and Swenson C.J.*, *Journal of the American Chemical Society*. 1985. V. 107. P. 569.
70. *Jorgensen W.L., Maxwell D.S., and Tirado-Rives J.*, *Journal of the American Chemical Society*. 1996. V. 118. P. 11225.
71. *Kaminski G.A., Friesner R.A., Tirado-Rives J., and Jorgensen W.L.*, *The Journal of Physical Chemistry B*. 2001. V. 105. P. 6474.
72. *Horta B.A.C., Merz P.T., Fuchs P.F.J., Dolenc J., Riniker S., and Hu'nenberger P.H.*, *Journal of Chemical Theory and Computation*. 2016. V. 12. P. 3825.
73. *Kashefolgheta S., Wang S., Acree W.E., and Hünenberger P.H.*, *Physical Chemistry Chemical Physics*. 2021. V. 23. P. 13055.
74. *Patel D.S., He X., and MacKerell A.D.*, *The Journal of Physical Chemistry B*. 2015. V. 119. P. 637.
75. *Jana M. and MacKerell A.D.*, *The Journal of Physical Chemistry B*. 205. V. 119. P. 7846.
76. *Yang M., Aytenfisu A.H., and MacKerell A.D.*, *Carbohydrate Research*. 2018. V. 457. P. 41.
77. *Aytenfisu A.H., Yang M., and MacKerell A.D.*, *Journal of Chemical Theory and Computation*. 2018. V. 14. P. 3132.
78. *Mackerell A.D., Feig M., and Brooks C.L.*, *Journal of Computational Chemistry*. 2004. V. 25. P. 1400.
79. *MacKerell A.D., Feig M., and Brooks C.L.*, *Journal of the American Chemical Society*. 2004. V. 126. P. 698.
80. *Foloppe N. and MacKerell Jr.A.D.*, *J. Comput. Chem*. 2000. V. 21. P. 86.
81. *Klauda J.B., Venable R.M., Freites J.A., O'Connor J.W., Tobias D.J., Mondragon-Ramirez C., Vorobyov I., MacKerell A.D., and Pastor R.W.*, *J. Phys. Chem. B*. 2010. V. 114. P. 7830.
82. *Denning E.J., Priyakumar U.D., Nilsson L., and Mackerell A.D.*, *J. Comput. Chem*. 2011. V. 32. P. 1929.
83. *Hart K., Foloppe N., Baker C.M., Denning E.J., Nilsson L., and MacKerell A.D.*, *J. Chem. Theory Comput*. 2012. V. 8. P. 348.
84. *Huang J., Rauscher S., Nawrocki G., Ran T., Feig M., de Groot B.L., Grubmüller H., and MacKerell A.D.*, *Nat Methods*. 2017. V. 14. P. 71.
85. *Vanommeslaeghe K., Hatcher E., Acharya C., Kundu S., Zhong S., Shim J., Darian E., Guvench O., Lopes P., Vorobyov I., and Mackerell A.D.*, *J. Comput. Chem., NA*. 2009.
86. *Soteras Gutierrez I., Lin F.-Y., Vanommeslaeghe K., Lemkul J.A., Armacost K.A., Brooks C.L., and MacKerell A.D.*, *Bioorganic & Medicinal Chemistry*. 2016. V. 24. P. 4812.
87. *Ha S.N., Giammona A., Field M., and Brady J.W.*, *Carbohydrate Research*. 1988. V. 180. P. 207.
88. *Brady J.W.*, *Journal of the American Chemical Society*. 1989. V. 111. P. 5155.
89. *Grootenhuis P.D.J. and Haasnoot C.A.G.*, *Molecular Simulation*. 1993. V. 10. P. 75.
90. *Kouwizjer M.L.C.E. and Grootenhuis P.D.J.*, *The Journal of Physical Chemistry*. 1995. V. 99. P. 13426.
91. *Reiling S., Schlenkrich M., and Brickmann J.*, *Journal of Computational Chemistry*. 1996. V. 17. P. 450.

92. Palma R., Zuccato P., Himmel M.E., Liang G., and Brady J.W., "Molecular Mechanics Studies of Cellulases", in *Glycosyl Hydrolases for Biomass Conversion*, V. 769, ACS Symposium Series (American Chemical Society, Oct. 2000). P. 112–130.
93. Kuttel M., Brady J.W., and Naidoo K.J., *Journal of Computational Chemistry*. 2002. V. 23. P. 1236.
94. Guvench O., Greene S.N., Kamath G., Brady J.W., Venable R.M., Pastor R.W., and Mackerell A.D., *J. Comput. Chem.* 2008. V. 29. P. 2543.
95. Hatcher E., Guvench O., and MacKerell A.D., *J. Phys. Chem. B*. 2009. V. 113. P. 12466.
96. Raman E.P., Guvench O., and MacKerell A.D., *J. Phys. Chem. B*. 2010. V. 114. P. 12981.
97. Vanommeslaeghe K. and MacKerell A., *Biochimica et Biophysica Acta (BBA) – General Subjects*. 2015. V. 1850. P. 861.
98. Huang J., Lopes P.E.M., Roux B., and MacKerell A.D., *The Journal of Physical Chemistry Letters*. 2014. V. 5. P. 3144.
99. Baker C.M., Anisimov V.M., and MacKerell A.D., *The Journal of Physical Chemistry B*. 2011. V. 115. P. 580.
100. Daura X., Mark A.E., and Van Gunsteren W.F., *Journal of Computational Chemistry*. 1998. V. 19. P. 535.
101. Lins R.D. and Hünenberger P.H., *J. Comput. Chem.* 2005. V. 26. P. 1400.
102. Oostenbrink C., Villa A., Mark A.E., and Van Gunsteren W.F., *Journal of Computational Chemistry*. 2004. V. 25. P. 1656.
103. Pereira C.S., Kony D., Baron R., Müller M., van Gunsteren W.F., and Hünenberger P.H., *Biophysical Journal*. 2006. V. 90. P. 4337.
104. Franca E.F., Lins R.D., Freitas L.C.G., and Straatsma T.P., *Journal of Chemical Theory and Computation*. 2008. V. 4. P. 2141.
105. Bergensträhle M., Thormann E., Nordgren N., and Berglund L.A., *Langmuir*. 2009. V. 25. P. 4635.
106. Kräutler V., Müller M., and Hünenberger P.H., *Carbohydrate Research*. 2007. V. 342. P. 2097.
107. Autieri E., Sega M., Pederiva F., and Guella G., *The Journal of Chemical Physics*. 2010. V. 133. P. 095104.
108. Hansen H.S. and Hünenberger P.H., *Journal of Computational Chemistry*. 2010. V. 31. P. 1.
109. Hansen H.S. and Hünenberger P.H., *J. Comput. Chem.* 2011. V. 32. P. 998.
110. Naumov V.S. and Ignatov S.K., *Journal of Molecular Modeling*. 2017. V. 23. P. 244.
111. Nester K., Gaweda K., and Plazinski W., *J. Chem. Theory Comput.* 2019.
112. Pol-Fachin L., Rusu V.H., Verli H., and Lins R.D., *Journal of Chemical Theory and Computation*. 2012. V. 8. P. 4681.
113. Marrink S.J., Risselada H.J., Yefimov S., Tieleman D.P., and de Vries A.H., *J. Phys. Chem. B*. 2007. V. 111. P. 7812.
114. Souza P.C.T., Alessandri R., Barnoud J., Thallmair S., Faustino I., Grünewald F., Patmanidis I., Abdizadeh H., Bruininks B.M.H., Wassenaar T.A., Kroon P.C., Melcr J., Nieto V., Corradi V., Khan H.M., Domariski J., Javanainen M., Martinez-Seara H., Reuter N., Best R.B., I. Vattulainen, Monticelli L., Periole X., Tieleman D.P., de Vries A.H., and Marrink S.J., *Nature Methods*. 2021. V. 18. P. 382.
115. López C.A., Rzepiela A.J., de Vries A.H., Dijkhuizen L., Hünenberger P.H., and Marrink S.J., *J. Chem. Theory Comput.* 2009. V. 5. P. 3195.
116. Lutsyk V., Wolski P., and Plazinski W., *Journal of Chemical Theory and Computation*. 2022. V. 18. P. 5089.
117. Gruñewald F., Punt M.H., Jefferys E.E., Vainikka P.A., König M., Virtanen V., Meyer T.A., Pezeshkian W., Gormley A.J., Karonen M., Sansom M.S.P., Souza P.C.T., and Marrink S.J., *Journal of Chemical Theory and Computation*. 2022. V. 18. P. 7555.
118. Tian C., Kasavajhala K., Belfon K.A.A., Raguet L., Huang H., Miguez A.N., Bickel J., Wang Y., Pincaj J., Wu Q., and Simmerling C., *J. Chem. Theory Comput.* 2020. V. 16. P. 528.
119. Hornak V., Abel R., Okur A., Strockbine B., Roitberg A., and Simmerling C., *Proteins: Structure, Function, and Bioinformatics*. 2006. V. 65. P. 712.
120. Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., and Kollman P.A., *Journal of the American Chemical Society*. 1995. V. 117. P. 5179.
121. Pérez A., Marchán I., Svozil D., Sponer J., Cheatham T.E., Laughton C.A., and Orozco M., *Biophysical Journal*. 2007. V. 92. P. 3817.
122. Zgarbová M., Sponer J., Otyepka M., Cheatham T.E., Galindo-Murillo R., and Jurečka P., *J. Chem. Theory Comput.* 2015. V. 11. P. 5723.
123. Ivani I., Dans P.D., Noy A., Pérez A., Faustino I., Hospital A., Walther J., Andrio P., Gonñ R., Balaceanu A., Portella G., Battistini F., Gelpi J.L., González C., Vendruscolo M., Laughton C.A., Harris S.A., Case D.A., and Orozco M., *Nature Methods*. 2016. V. 13. P. 55.
124. Dickson C.J., Walker R.C., and Gould I.R., *J. Chem. Theory Comput.* 2022. V. 18. P. 1726.
125. He X., Man V.H., Yang W., Lee T.-S., and Wang J., *J. Chem. Phys.* 2020. V. 153. P. 114502.
126. Woods R.J., Dwek R.A., Edge C.J., and Fraser-Reid B., *The Journal of Physical Chemistry*. 1995. V. 99. P. 3832.
127. Kirschner K.N. and Woods R.J., *Proceedings of the National Academy of Sciences*. 2001. V. 98. P. 10541.
128. Basma M., Sundara S., Çalgan D., Vernali T., and Woods R.J., *Journal of Computational Chemistry*. 2001. V. 22. P. 1125.
129. Tschampel S.M., Kennerty M.R., and Woods R.J., *Journal of Chemical Theory and Computation*. 2007. V. 3. P. 1721.
130. Homans S.W., *Biochemistry*. 1990. V. 29. P. 9110.
131. Edge C.J., Singh U.C., Bazzo R., Taylor G.L., Dwek R.A., and Rademacher T.W., *Biochemistry*. 1990. V. 29. P. 1971.
132. Momany F. and Willett J., *Carbohydrate Research*. 2000. V. 326. P. 194.
133. Momany F.A., Willett J., and Schnupf U., *Carbohydrate Polymers*. 2009. V. 78. P. 978.
134. Jorgensen W.L., *Journal of the American Chemical Society*. 1978. P. 8.
135. Jorgensen W.L., *The Journal of Chemical Physics*. 1979. V. 71. P. 5034.
136. Jorgensen W.L. and Ibrahim M., *Journal of the American Chemical Society*. 1980. V. 102. P. 3309.
137. Jorgensen W.L., *Journal of the American Chemical Society*. 1981. V. 103. P. 335.

138. Maxwell D.S., Tirado-Rives J., and Jorgensen W.L., *J. Comput. Chem.* 1995. V. 16. P. 984.
139. Dunbrack R.L. and Karplus M., *Nat Struct Mol Biol.* 1994. V. 1. P. 334.
140. Kony D., Damm W., Stoll S., and Van Gunsteren W.F., *Journal of Computational Chemistry.* 2002. V. 23. P. 1416.
141. Dodda L.S., Vilseck J.Z., Tirado-Rives J., and Jorgensen W.L., *The Journal of Physical Chemistry B* 121, 3864 (2017).
142. Dodda L.S., Cabeza de Vaca I., Tirado-Rives J., and Jorgensen W.L., *Nucleic Acids Research.* 2017. V. 45. P. W331.
143. Deshchenya V.I., Kondratyuk N.D., Lankin A.V., and Norman G.E., *Russian Journal of Physical Chemistry A.* 2022. V. 96. P. 556.
144. Deshchenya V., Kondratyuk N., Lankin A., and Norman G., *Journal of Molecular Liquids.* 2022. V. 367. P. 120456.
145. Foley B.L., Tessier M.B., and Woods R.J., *WIREs Comput Mol Sci.* 2012. V. 2. P. 652.
146. Sauter J. and Grafmüller A., *Journal of Chemical Theory and Computation.* 2015. V. 11. P. 1765.
147. Lay W.K., Miller M.S., and Elcock A.H., *Journal of Chemical Theory and Computation.* 2016. V. 12. P. 1401.
148. Batista M.L.S., Pérez-Sánchez G., Gomes J.R.B., Coutinho J.A.P., and Maginn E.J., *The Journal of Physical Chemistry B.* 2015. V. 119. P. 15310.
149. Kopanichuk I., Scerbacova A., Ivanova A., Cheremisin A., and Vishnyakov A., *Journal of Molecular Liquids.* 2022. V. 360. P. 119525.
150. Yoo J. and Aksimentiev A., *Physical Chemistry Chemical Physics.* 2018. V. 20. P. 8432.
151. Yoo J. and Aksimentiev A., *The Journal of Physical Chemistry Letters.* 2012. V. 3. P. 45.
152. Yoo J. and Aksimentiev A., *Journal of Chemical Theory and Computation.* 2016. V. 12. P. 430.
153. Yoo J. and Aksimentiev A., *The Journal of Physical Chemistry Letters.* 2016. V. 7. P. 3812.
154. Yoo J., Wilson J., and Aksimentiev A., *Biopolymers.* 2016. V. 105. P. 752.
155. Lay W.K., Miller M.S., and Elcock A.H., *Journal of Chemical Theory and Computation.* 2017. V. 13. P. 1874.
156. Bhargava B.L. and Balasubramanian S., *The Journal of Chemical Physics.* 2007. V. 127. P. 114510.
157. Youngs T.G.A. and Hardacre C., *Chem Phys Chem.* 2008. V. 9. P. 1548.
158. Schröder C., *Phys. Chem. Chem. Phys.* 2012. V. 14. P. 3089.
159. Jamali S.H., Westen T.V., Moulton O.A., and Vlucht T.J.H., *Journal of Chemical Theory and Computation.* 2018. V. 14. P. 6690.
160. Kovalenko V.I., *Russian Chemical Reviews.* 2010. V. 79. P. 231.
161. Nishiyama Y., *Journal of Wood Science.* 2009. V. 55. P. 241.
162. Nishiyama Y., Sugiyama J., Chanzy H., and Langan P., *Journal of the American Chemical Society.* 2003. V. 125. P. 14300.
163. Nishiyama Y., Langan P., and Chanzy H., *Journal of the American Chemical Society.* 2002. V. 124. P. 9074.
164. Wada M., Chanzy H., Nishiyama Y., and Langan P., *Macromolecules.* 2004. V. 37. P. 8548.
165. Hori R. and Wada M., *Cellulose.* 2006. V. 13. P. 281.
166. Wada M., Heux L., and Sugiyama J., *Biomacromolecules.* 2004. V. 5. P. 1385.
167. Newman R.H., *Cellulose.* 2008. V. 15. P. 769.
168. Kulshreshtha A.K., *The Journal of the Textile Institute.* 1979. V. 70. P. 13.
169. Dufresne A., *Materials Today.* 2013. V. 16. P. 220.
170. Vanderfleet O.M. and Cranston E.D., *Nature Reviews Materials.* 2020. V. 6. P. 124.
171. Dhali K., Ghasemlou M., Daver F., Cass P., and Adhikari B., *Science of the Total Environment.* 2021. V. 775. P. 145871.
172. Rao A., Divoux T., Owens C.E., and Hart A.J., *Cellulose.* 2022. V. 29. P. 2387.
173. Mu R., Hong X., Ni Y., Li Y., Pang J., Wang Q., Xiao J., and Zheng Y., *Trends in Food Science & Technology.* 2019. V. 93. P. 136.
174. Kalashnikova I., Bizot H., Cathala B., and Capron I., *Langmuir.* 2011. V. 27. P. 7471.
175. Li R., Liu Y., Seidi F., Deng C., Liang F., and Xiao H., *Advanced Materials Interfaces.* 2022. V. 9. P. 2101293.
176. Lugoloobi I., Maniriho H., Jia L., Namulinda T., Shi X., and Zhao Y., *Journal of Controlled Release.* 2021. V. 336. P. 207.
177. Aziz T., Ullah A., Fan H., Ullah R., Haq F., Khan F.U., Iqbal M., and Wei J., *Journal of Polymers and the Environment.* 2021. V. 29. P. 2062.
178. Heiner A.P., Sugiyama J., and Teleman O., *Carbohydrate Research.* 1995. V. 273. P. 207.
179. Wu X., Moon R. J., and Martini A., *Cellulose.* 2014. V. 21. P. 2233.
180. van Duin A.C.T., Dasgupta S., Lorant F., and Goddard W.A., *The Journal of Physical Chemistry A.* 2001. V. 105. P. 9396.
181. Bergensträhle M., Berglund L.A., and Mazeau K., *The Journal of Physical Chemistry B.* 2007. V. 111. P. 9138.
182. Miyamoto H., Umemura M., Aoyagi T., Yamane C., Ueda K., and Takahashi K., *Carbohydrate Research.* 2009. V. 344. P. 1085.
183. Miyamoto H., Ago M., Yamane C., Seguchi M., Ueda K., and Okajima K., *Carbohydrate Research.* 2011. V. 346. P. 807.
184. Yang G., Miyamoto H., Yamane C., and Okajima K., *Polymer Journal.* 2007. V. 39. P. 34.
185. Yamane C., Miyamoto H., Hayakawa D., and Ueda K., *Carbohydrate Research.* 2013. V. 379. P. 30.
186. López C.A., Bellesia G., Redondo A., Langan P., Chundawat S.P.S., Dale B.E., Marrink S.J., and Gnana-karan S., *The Journal of Physical Chemistry B.* 2015. V. 119. P. 465.
187. Liu H., Sale K.L., Holmes B.M., Simmons B.A., and Singh S., *The Journal of Physical Chemistry B.* 2010. V. 114. P. 4293.
188. Kostrikskii A.Y., Tolmachev D.A., Lukasheva N.V., and Gurtovenko A.A., *Langmuir.* 2017. V. 33. P. 12793.
189. Gurtovenko A.A., Mukhamadiarov E.I., Kostrikskii A.Y., and Karttunen M., *The Journal of Physical Chemistry B.* 2018. V. 122. P. 9973.
190. Gurtovenko A.A. and Karttunen M., *Langmuir.* 2019. V. 35. P. 13753.

191. Gurtovenko A.A. and Karttunen M., *Soft Matter*. 2021. V. 17. P. 6507.
192. Guvench O., Hatcher E., Venable R.M., Pastor R.W., and MacKerell A.D., *J. Chem. Theory Comput.* 2009. V. 5. P. 2353.
193. Tchipev N., Seckler S., Heinen M., Vrabec J., Gratl F., Horsch M., Bernreuther M., Glass C.W., Niethammer C., Hammer N., Krischok B., Resch M., Kranzlmüller D., Hasse H., Bungartz H.-J., and Neumann P., *Int. J. High Perf. Comp. Applications* 0, 1094342018819741. (2019).
194. Kutzner C., Páll S., Fechner M., Esztermann A., de Groot B.L., and Grubmüller H., *Journal of computational chemistry*. 2015. V. 36. P. 1990.
195. Kutzner C., Páll S., Fechner M., Esztermann A., de Groot B.L., and Grubmüller H., *Journal of Computational Chemistry*. 2019. V. 40. P. 2418.
196. Stegailov V., Dlinnova E., Ismagilov T., Khalilov M., Kondratyuk N., Makagon D., Semenov A., Simonov A., Smirnov G., and Timofeev A., *The International Journal of High Performance Computing Applications*. 2019. V. 33. P. 507.
197. Pisarev V. and Kondratyuk N., *Fluid Phase Equilibria*. 2019. V. 501. P. 112273.
198. Smirnov G.S. and Stegailov V.V., *Journal of Physics: Condensed Matter*. 2019. V. 31. P. 235704.
199. Antropov A. and Stegailov V., *Journal of Nuclear Materials*. 2020. V. 533. P. 152110.
200. Kolotinskii D. and Timofeev A., *Computer Physics Communications*. 2023. V. 288. P. 108746.
201. Plimpton S., *J. Comput. Phys.* 1995. V. 117. P. 1.
202. Berendsen H., van der Spoel D., and van Drunen R., *Computer Physics Communications*. 1995. V. 91. P. 43.
203. Salomon-Ferrer R., Case D.A., and Walker R.C., *WIREs Computational Molecular Science*. 2013. V. 3. P. 198.
204. Brown W.M., Wang P., Plimpton S.J., and Tharrington A.N., *Computer Physics Communications*. 2011. V. 182. P. 898.
205. Brown W.M., Kohlmeyer A., Plimpton S.J., and Tharrington A.N., *Computer Physics Communications*. 2012. V. 183. P. 449.
206. Brown W.M. and Yamada M., *Computer Physics Communications*. 2013. V. 184. P. 2785.
207. Abraham M.J., Murtola T., Schulz R., Páll S., Smith J.C., Hess B., and Lindahl E., *SoftwareX*. 2015. V. 1–2. P. 19.
208. Götz A.W., Williamson M.J., Xu D., Poole D., Le Grand S., and Walker R.C., *Journal of Chemical Theory and Computation*. 2012. V. 8. P. 1542.
209. Salomon-Ferrer R., Götz A.W., Poole D., Le Grand S., and Walker R.C., *Journal of Chemical Theory and Computation*. 2013. V. 9. P. 3878.
210. Anderson J.A., Lorenz C.D., and Travesset A., *Journal of Computational Physics*. 2008. V. 227. P. 5342.
211. Glaser J., Nguyen T.D., Anderson J.A., Lui P., Spiga F., Millan J.A., Morse D.C., and Glotzer S.C., *Computer Physics Communications*. 2015. V. 192. P. 97.
212. Eastman P., Friedrichs M.S., Chodera J.D., Radmer R.J., Bruns C.M., Ku J.P., Beauchamp K.A., Lane T.J., Wang L.-P., Shukla D., Tye T., Houston M., Stich T., Klein C., Shirts M. R., and Pande V.S., *Journal of Chemical Theory and Computation*. 2013. V. 9. P. 461.
213. Eastman P., Swails J., Chodera J.D., McGibbon R.T., Zhao Y., Beauchamp K.A., Wang L.-P., Simmonett A.C., Harrigan M.P., Stern C.D., Wiewiora R.P., Brooks B.R., and Pande V.S., *PLOS Comput. Biology*. 2017. V. 13. P. 1.
214. Dong W., Kang L., Quan Z., Li K., Li K., Hao Z., and Xie X., in 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (hpcc/smartcity/dss) (Dec. 2016). P. 443–450.
215. Dong W., Li K., Kang L., Quan Z., and Li K., *Concur. and Comput.: Practice and Experience*. 2018. V. 30. P. e4468.
216. Yu Y., An H., Chen J., Liang W., Xu Q., and Chen Y., in Algorithms and architectures for parallel processing, edited by S. Ibrahim, K.-K. R. Choo, Z. Yan, and W. Pedrycz. 2017. P. 18–32.
217. Duan X., Gao P., Zhang T., Zhang M., Liu W., Zhang W., Xue W., Fu H., Gan L., Chen D., Meng X., and Yang G., in Sc18: international conference for high performance computing, networking, storage and analysis. Nov. 2018. P. 148–159.
218. Nikolskii V. and Stegailov V., *Lobachevskii Journal of Mathematics*. 2018. V. 39. P. 1228.
219. Abascal J.L. and Vega C., *J. Chem. Phys.* 2005. V. 123. P. 234505.
220. Nos'e S., *Mol. Phys.* 1984. V. 52. P. 255.
221. Hoover W.G., *Phys. Rev. A*. 1985. V. 31. P. 1695.
222. Shinoda W., Shiga M., and Mikami M., *Phys. Rev. B*. 2004. V. 69. P. 134103.
223. Andrea T.A., Swope W.C., and Andersen H.C., *The Journal of Chemical Physics*. 1983. V. 79. P. 4576.
224. Kondratyuk N., Nikolskiy V., Pavlov D., and Stegailov V., *The International Journal of High Performance Computing Applications*. 2021. V. 35. P. 312.
225. Dubbeldam D., Calero S., and Vlugt T.J., *Molecular Simulation*. 2018. V. 44. P. 653.
226. James C.J., Mulcahy D.E., and Steel B.J., *Journal of Physics D: Applied Physics*. 1984. V. 17. P. 225.
227. Rampp M., Buttersack C., and Lüdemann H.-D., *Carbohydrate Research*. 2000. V. 328. P. 561.
228. Chenlo F., Moreira R., Pereira G., and Ampudia A., *Journal of Food Engineering*. 2002. V. 54. P. 347.
229. Telis V., Telis-Romero J., Mazzotti H., and Gabas A., *International Journal of Food Properties*. 2007. V. 10. P. 185.
230. Gomes T.C.F. and Skaf M.S., *Journal of Computational Chemistry*. 2012. V. 33. P. 1338.