

Методы молекулярного и интегративного моделирования в структурной биологии

А.К. Шайтан, alex@intbio.org

Уделим внимание обсуждению и определению ключевых понятий, внесенных в название данной главы, а именно, “структурная биология”, “молекулярное моделирование” и “интегративное моделирование”, в том числе в историческом контексте.

Журнал Nature дает следующее определение структурной биологии: “Структурная биология – это изучение молекулярных структур и динамики биологических макромолекул, в особенности белков и нуклеиновых кислот, а также изучение того, как изменения в их структуре влияют на функцию” [1]. Данное определение, тем не менее, не дает ответа на вопрос о том, что мы понимаем под структурой биологической макромолекулы. В биологии активно используется подразделение уровней структурной организации биомолекул на первичный (химический) уровень, и последующие уровни (вторичный, третичный, четвертичный), которые описывают укладку в пространстве биомолекул. Первые технологии секвенирования — чтения последовательности мономеров в биополимерах появились в 1950-1970-ых годах. К 1951 году при участии Фредерика Сенгера была разработана технология определения первичной последовательности белков [2], к 1964 Робертом Холли определена первичная последовательность молекулы транспортной РНК [3], к 1973 году Аланом Максамом и Вальтером Гилбертом была определена последовательность ДНК *lac* оператора [4], а к 1977 году ими [5], а также Ф. Сенгером [6] были развиты рутинные методы секвенирования ДНК, которые используются до сих пор. Среди работ отечественных ученых важный вклад внесла группа академика Александра Александровича Баева – к 1967 году была впервые определена первичная структура валиновой транспортной РНК пекарских дрожжей [7]. Несмотря на успехи определения первичной структуры биополимеров, именно, пространственная структурная и динамическая организация биомолекул является ключевой для понимания механизмов их функционирования (что описывается термином “mechanistic insight”, часто встречающимся в англоязычной литературе). В силу исторических и

технических причин структура биомакромолекул обычно воспринимается как некоторая трехмерная модель, отображающая укладку биомакромолекул (белков, нуклеиновых кислот, липидов и т.д.) в пространстве. В некотором роде такой подход продолжает парадигму принятую в химии для описания структуры химических соединений в виде взаимного расположения атомов и связей между ними. На Рис. 0.1а приведены примеры белковых структур в руках основоположников белковой кристаллографии Макса Перутца и Джона Кендрию, впервые разрешивших структуру миоглобина в 1958 году [8]. На Рис. 0.1б модели белков в 1980-ых годах в руках основоположника отечественной белковой кристаллографии академика Бориса Константиновича Вайнштейна [9]. В компьютерную эпоху модели биомакромолекул принято представлять в виде файлов (например, в формате PDB) с описанием пространственных координат отдельных атомов. Следует однако отметить, что такое представление моделей позволяет описать лишь некоторые статичные конфигурации атомов. Насколько такой подход позволяет действительно представить и описать происходящие процессы в микромире? В реальном мире молекулы не являются некоторыми статическими структурами, а находятся в постоянном динамическом изменении. Для описания поведения газов, жидкостей, мягкой материи (soft matter), разупорядоченных и слабоупорядоченных систем в физике были разработаны формализмы термодинамики и статистической физики (см. Рис. 0.2). В этих подходах, приходится отказаться от описания молекулярных процессов в виде статических структур в пользу некоторых средних характеристик, ансамблей состояний, вероятностного описания различных микросостояний согласно статистике Гиббса-Больцмана. В микромире наличие, например, как у глобулярных белков более менее детерминированных пространственных структур является скорее счастливым исключением. В то же время при попытке описания многих биологических систем парадигма представления структуры в виде даже ограниченного набора статических структур испытывает значительные трудности. С одной стороны здесь можно упомянуть неструктурированные белки (англ. intrinsically disordered proteins), у которых отсутствует какая-то одна статическая структура. С другой стороны подобные проблемы характерны при попытке описания многих комплексных биологических образований, например, структуры хроматина или даже структуры нуклеосомы. В этой связи высказывается точка зрения, что понятие структуры биомолекулы следует трактовать более широко, в идеальном случае как ансамбль возможных конформаций дан-

ной биомолекулы с указанием статистических весов различных конформаций, а также их зависимости от внешних условий [10]. Такое понимание структуры биомакромолекулярных комплексов наиболее близко связано с пониманием их функционирования и, по-видимому, будет являться основой для познания функционирования живых систем. Развивая данную концепцию, стоит упомянуть о том, что большинство биологических объектов на молекулярном уровне являются неравновесными системами, в которых происходят процессы превращения энергии. Для описания таких систем не достаточно подходов только равновесной статистической физики, а требуется развитие подходов описания, так называемой, активной материи (active matter) [11].

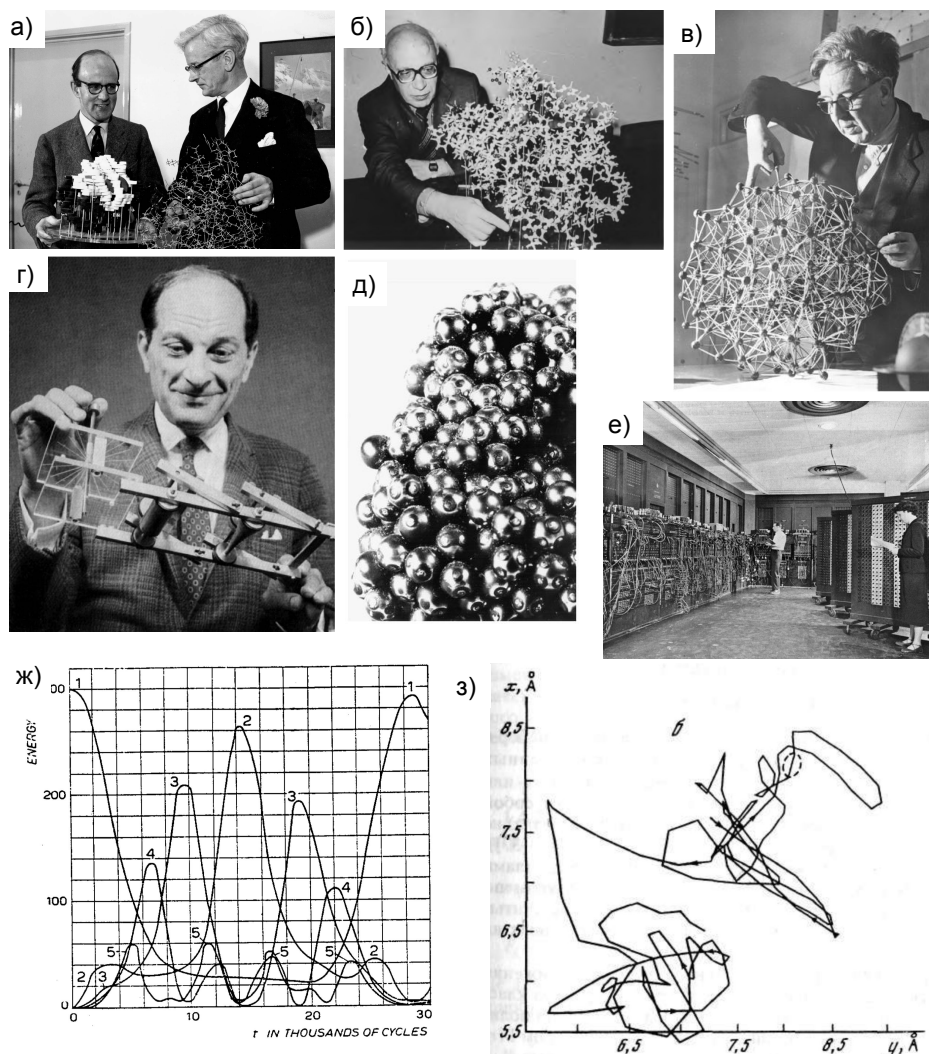


Рисунок 0.1 — Иллюстрации к истории структурной биологии и моделирования. а) М. Перуц (слева) и Д. Кендрю (справа) со структурными моделями миоглобина. б) Б.К. Вайнштейн со структурой белка [9], в) Дж. Д. Бернал с моделью укладки атомов в жидкости [12], г) С. Улам с тележкой Монте-Карло Энрико Ферми [13], д) результаты физического моделирования жидкости в виде стальных шаров в работах Д.Д. Бернала [12], е) первый программируемый компьютер общего назначения ENIAC [14], ж) изменение полной энергии со временем для первых пяти мод вибрирующей струны в задаче Ферми-Пасты-Улама-Цингу, первая работа по исследованию статистических свойств механических систем методом молекулярной динамики, 1955 г. [15], (з) Траектории движения атомов в жидкости, рассчитанные методом МД из кандидатской диссертации А.Г. Гривцова, начало 1970-ых годов [16].

Сам по себе термин *молекулярное моделирование* (*molecular modeling*) является достаточно широким понятием, значение которого постепенно расширяется и видоизменяется с развитием различных новых методов. В одних

случаях упор может делаться непосредственно на задачу создания модели (например, пространственной модели белка), в других случаях на анализ и исследование свойств молекул на основе физических моделей поведения и взаимодействия атомов. К первому типу задач можно отнести традиционные задачи реконструкция структуры биомакромолекул по экспериментальным данным. Данный процесс всегда включает в себя шаг создания модели, которая удовлетворяет этим данным. На заре структурной биологии реконструкция биомолекул по результатам кристаллографических экспериментов долгое время сводилась к построению пространственных моделей из проволоки, дерева, других материалов (Рис. 0.1аб). Позднее данный процесс был значительно автоматизирован с применением компьютерных технологий, разработаны соответствующие программные пакеты по автоматизации построения моделей и их визуализации (например, CCP4 [17], O, Phenix [18], Coot [19], PyMol [20], Chimera [21]). Построение моделей молекул в виде пространственного расположения атомов, удовлетворяющего набору экспериментальную данных, - одно из возможных значений термина компьютерное молекулярное моделирование. В то же время со времен возникновения молекулярно-кинетической теории (основы которой были в том числе заложены М.В. Ломоносовым [22]) ученые стремились к описанию и предсказанию свойств молекулярных систем на основе учета взаимодействий между отдельными атомами, описываемых законами физики. В до-компьютерную эпоху такие предсказания можно было делать либо на основе теоретических расчетов, либо на основе физических механических моделей. Так, например, в 1950-60-ых годах для молекулярного моделирования жидкостей Джон Десмонд Бернал активно использовал механические модели в виде металлических и пластиковых шариков [23](Рис. 0.1вд)¹. С развитием компьютерных технологий появились возможности численного моделирования поведения молекулярных систем на основе их физических моделей. Следует отметить, что в английском языке для описания такого рода деятельности активно используется термин *molecular simulations* (см. например [26]), который на русский язык обычно переводится также “молекулярное моделирование”, однако слово “simulations” имеет важные отличия в своей смысловой нагрузке. Под этим термином понимается не столько создание моделей, сколько изу-

¹Любопытным также является факт, что тот же Дж. Д. Бернал является одним из основателей области биомолекулярной кристаллографии, под его руководством работали классики структурной биологии Макс Перутц, Дороти Ходжкин, Розалинд Франклин [24], а также в честь него назван пакет по молекулярному моделированию DESMOND [25].

чение и анализ теоретических моделей с помощью вычислительных методов, так называемые “вычислительные эксперименты”. В области инженерного моделирования в русском языке для обозначения термина “simulation modeling” используется термин *имитационное моделирование*, однако он не вошел в обиход в среде исследователей, занимающихся молекулярным моделированием. Подход к изучению различных систем на основе проведения вычислительных экспериментов на данный момент многими воспринимается как одна из парадигм научного познания наравне с экспериментальным и теоретическими подходами [27] и получил название *in silico* подхода. В случае изучения свойств вещества данный подход предполагает наличие определенной физической модели исследуемого объекта, которая описывает взаимодействия между его составными частями, и использует различные вычислительные алгоритмы, чтобы исследовать свойства данной модели вещества. Если для описания физической модели вещества используются различные приближения, основанные на законах квантовой механики, то такие подходы зачатую называют *ab initio* подходами (от лат. “из первых принципов” - подразумевая, что для моделирования необходимо знать лишь базовые физические константы). Для моделирования многих задач и веществ такой формализм является слишком вычислительно затратным, поэтому исторически область “molecular simulations” развивалась в первую очередь на основе формализма классической механики (частицы вещества представлялись в виде материальных точек, взаимодействующих по законам классической механики).

Зарождение области вычислительных экспериментов в молекулярном моделировании происходило в середине-конце 1940-ых годов во время работы над созданием ядерного и термоядерного оружия в США и связано в первую очередь с именами Станислава Улама, Джона фон Неймана, Николаса Метрополиса, а также с вводом в строй в 1945 году первого электронного цифрового вычислителя общего назначения ENIAC (Рис. 0.1e). С. Улам и Д. фон Нейман, работая над проблемой транспорта нейтронов в различных типах ядерных систем, предложили в 1947 году моделировать диффузию нейтронов путем многократного численного расчета траекторий движения отдельных нейтронов с учетом случайности их взаимодействия с окружающим веществом. Таким образом, решение задачи о вычислении средних характеристик данного процесса сводилось к статистической выборке (statistical sampling), а не к полному численному решению уравнений, которые описывали этот процесс. Благода-

ря наличию компьютера ENIAC стало возможным написать компьютерную программу для соответствующих расчетов, которая использовала в том числе генератор псевдослучайных чисел для моделирования случайных событий и соответствующих статистических выборок [28]. Подходы основанные на статистическом моделировании с использованием случайных величин были названы ими **методами Монте-Карло**, по предложению **Николаса Метрополиса**, а первая статья по этому методу опубликована в 1949 году [29]. Любопытным является факт, что физик Энрико Ферми еще до формулировки метода Монте Карло Уламом и фон Нейманом использовал метод статистических выборок для своих задач, проводя расчеты вручную в предрассветные часы, когда страдал бессонницей [30]. В 1947 году, когда постановка расчетов на компьютере **ENIAC** была временно невозможной из-за его перевозки, Ферми придумал аналоговый компьютер, который можно было использовать для моделирования транспорта нейтронов. Данный компьютер представлял собой тележку с карандашом, которую нужно было передвигать по бумажному чертежу и периодически изменять ее направление на основе случайных чисел, таким образом моделируя передвижения нейтронов и их столкновения с атомами среды. Данное устройство получило название **тележки Монте-Карло или FERMIAC** (см Рис. 0.1г).

Следующей важной вехой в развитии методов Монте Карло в молекулярном моделировании и статистической механике явилась разработка **Метрополисом**, супругами Розенблют и Теллер в 1953 году метода Монте Карло на основе марковских цепей (Markov Chain Monte Carlo, MCMC) для изучения уравнений состояния вещества состоящего из индивидуальных взаимодействующих молекул [31]. В этой работе авторы провели исследование двумерного газа из жестких сфер на компьютере MANIAC. Инновацией данного метода стало вычисление статистической выборки состояний согласно распределению Больцмана в многомерном пространстве через построение марковской цепи состояний. Случайные переходы между состояниями принимались или отклонялись согласно **критерию Метрополиса-Хастингса** таким образом, что **лимитирующее распределение состояний в данной марковской цепи соответствовало Больцмановскому распределению состояний по энергиям**. На идеях данной работы основано большинство подходов по моделированию молекулярных систем методами Монте Карло и в наши дни [32].

1950-ые годы также ознаменовались первыми работам в области применения метода молекулярной динамики - численного решения уравнений движения

классической механики. Впервые Ферми, Паста, Улам и Цингу исследовали колебания струны с нелинейной упругостью и показали наличие необычной периодичности возникающей в нелинейных системах [15]. Идея моделирования динамики молекулярных систем была сформулирована Олдером и Вайнрайтом в 1957 году – с помощью компьютера UNIVAC исследовались фазовые переходы в системе твердых сфер [33]. К первым попыткам применить методы молекулярной динамики к реалистичным системам можно отнести работу Гибсона и др. 1960 года по моделированию радиационного поражения меди [34] и работу Анисура Рамана по моделированию жидкого аргона 1964 года [35]. В нашей стране первые шаги по развитию методов молекулярной динамики связаны с деятельностью А.А. Гривцова, Э.Э. Шноля, Н.К. Балабаева из Института прикладной математики (ИПМ, Москва) начиная с 1967 года [16]. С тех пор методы молекулярной динамики стали активно применяться для моделирования газов, жидкостей [36], а позднее стали активно применяться и для моделирования биомолекулярных систем. Развитие применения методов молекулярного моделирования к биомолекулярным системам связаны с деятельностью групп Мартина Карплуса, Арье Варшеля, Майкла Левитта, которые получили в 2013 году нобелевскую премию за “разработку мультимасштабных моделей сложных химических систем”, а также групп Питера Коллмана, Хермана Берендсена, Клауса Шультена, Дэвида Кейса, Эрика Линдаля, Алекса МакКерела и др. Работы группа М. Карплуса легли в основу набора силовых полей CHARMM, а также одноименной программы для моделирования [37]. На основе подходов, силовых полей и программного кода CHARMM были развиты многие родственные программы, например, программа для параллельных расчетов молекулярной динамики NAMD (в группе К. Шультена) [38], программа для моделирования на основе данных ЯМР XPLOR [39], программа для моделирования по гомологии MODELLER [40]. Работы группы Питера Коллмана положили основу программному пакету AMBER и одноименной группе силовых полей [41]. Работы группы Хермана Берендсена в Голландии заложили основы пакета по моделированию GROMACS [42] и силового поля GROMOS [43]. В нашей стране развитие программных продуктов в области молекулярной динамики связано в том числе с разработкой программы PUMA и ее версий под руководством Н.К. Балабаева [44].

В первые десятилетия XXI века область молекулярного моделирования продолжала активно развиваться и совершенствоваться, в том числе и в мето-

дологическом плане. Стоит отметить адаптацию программного обеспечения для использования **массивно-параллельных архитектур графических процессоров**, что позволило уверенно проводить **расчеты МД** в микросекундном диапазоне не только на суперкомпьютерах, но и отдельных рабочих станциях [45]. Усилиями нескольких групп создавались специализированные компьютерные платформы для расчета молекулярной динамики, в частности архитектуры MD GRAPE в Японии [46] и компьютеры ANTON компании D.E. Shaw Research [47]. Последние позволяют проводить вычисления в миллисекундном диапазоне времен моделируемых систем.

Серьезный прогресс был достигнут и в совершенствовании методов повышения эффективности статистических выборок (enhanced sampling methods) в ходе расчетов молекулярной динамики. К таким методам можно отнести методы **метадинамики (metadynamics)** [48], **адаптивной смещающей силы (adaptive biasing force)** [49], **адаптивно смещаемой молекулярной динамики (adaptively biased molecular dynamics)** [50], **динамики с обменом репликами (replica exchange molecular dynamic/parallel tempering)** [51], **управляемой молекулярной динамики (steered molecular dynamics)** [52]. Многие из этих методов получили активное распространение благодаря созданию **кросс-платформенного дополнения к пакетам молекулярной динамики PLUMED** [53]. Были в том числе развиты теоретические основы, связывающие характеристики расчета неравновесных молекулярных процессов с равновесными параметрами моделируемых системы. К таковым можно отнести равенство Джарзинского, связывающее работу в ходе неравновесного процесса над системой с разницей свободной энергии между конечным и начальным состояниями [54] и **флуктуационную теорему Крукса (Crooks fluctuation theorem)** [55]. Любопытно, что открытие и формулировка данных достаточно лаконичных и простых соотношений неравновесной статистической физики произошла только на рубеже XX и XXI веков, хотя явные предпосылки для их формулировки существовали со времен Больцмана. Отдельно стоит отметить развитие и успехи различных эмпирических подходов к моделированию структуры и дизайну белков. В отличие от методов, исходящих в первую очередь из физических представлений о структуре, подвижности и взаимодействии атомов в биомолекулах (physics-based methods), такие подходы активно используют различные эмпирические скоринговые функции для оценки вероятности той или иной конформации белка, которые основываются на статистическом анализе большого количества уже

известных пространственных структур белков, а также эволюционном анализе похожести различных аминокислотных мотивов в белках. Перебор конформационного пространства полипептидной цепи также зачастую ограничивают набором конформаций фрагментов присутствующих в уже известных структурах белков. К таким методам можно отнести подходы основанные на пакете ROSETTA [56], AWSEM-MD [57] и др.

Выше мы обсудили в историческом контексте различные взгляды на задачи и методы молекулярного моделирования. В одном случае речь идет о задачах поиска оптимальной пространственной модели молекулы на основе большого количества экспериментальных данных (например, реконструкции структуры белка по набору рефлексов картины дифракции его кристаллов), в другом случае упор делается на имитационное моделирование (molecular simulations) систем на основе физических моделей взаимодействий между атомами, которые иногда называют также *ab initio* подходами (напр. *ab initio protein folding*² и т.д.). Разные аспекты данных подходов просуммированы на рисунке 0.2, они отличаются по наличию, качеству и информационному содержанию используемых экспериментальных данных (Рис. 0.2в), по используемым математическим подходам и физическим теориям (Рис. 0.2б) и по самому представлению моделей (Рис. 0.2а) - в одном случае речь обычно идет о некоторой статической структуре или наборе структур, в другом о статистических ансамблях различных конформаций.

В то же время в структурной биологии существует большое количество задач, которые не решаются в рамках одного из двух вышеописанных подходов. С одной стороны для многих систем, экспериментальные данные могут присутствовать лишь в ограниченном количестве и зачастую иметь опосредованное отношение к деталям молекулярной структуры (напр. данные спектроскопических экспериментов, где оценивается взаимодействие введенных в структуру меток, данные по химической доступности и реакционной способности различных групп, данные малоуглового рассеяния и т.д.). С другой стороны, решение структурных задач напрямую *in silico* методами физического моделирования на основе взаимодействия частиц невозможно для большинства систем в силу чрезвычайной вычислительной сложности и зачастую качества имеющихся моделей взаимодействия атомов в биомакромолекулах. Отдельную проблему

²Не стоит путать с пониманием термина *ab initio* в квантовой химии, где под этим понимается моделирование на основе решения уравнения Шредингера.

представляют задачи описания структуры биомакромолекул, у которых нет явной статической структуры, а она представлена ансамблем конформаций. К примерам таких систем можно отнести внутренне разупорядоченные белки (англ. *intrinsically disordered proteins*), а также большинство биомакромолекулярных комплексов (вплоть до структуры хроматина на масштабах клеточного ядра) – по мере роста размеров комплексов, они все реже формируют какую-то одну статическую конформацию, а их функционирование завязано сложную внутри- и межмолекулярную динамику. Для описания таких молекул динамический формализм используемый в имитационном молекулярном моделировании (*molecular simulations*) подходит значительно лучше, однако зачастую решение таких задач из первых принципов опять сводится к вышеописанным затруднениям.

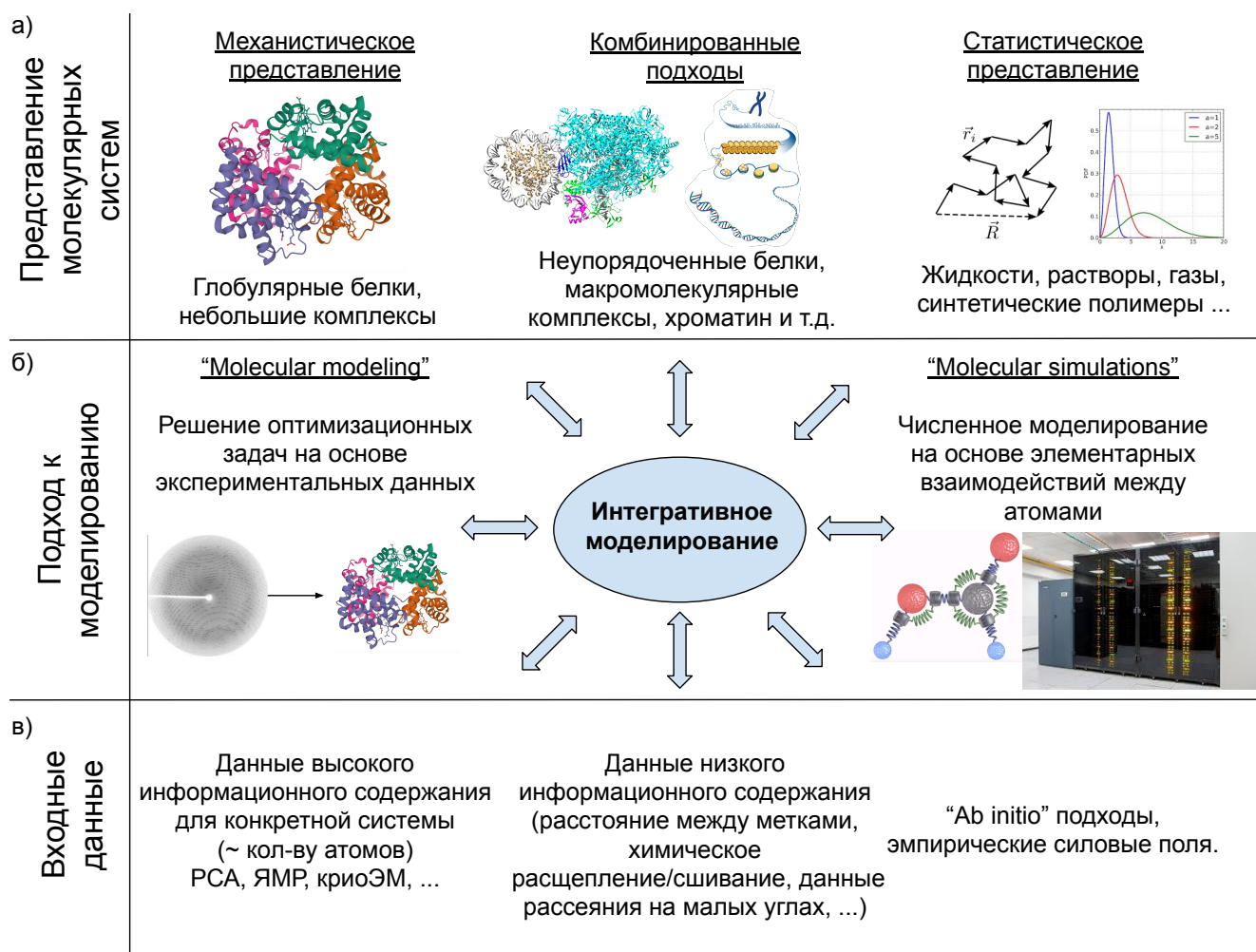


Рисунок 0.2 — Диаграмма, суммирующая различные подходы к описанию, моделированию и изучению структуры и динамики биомакромолекул.

Таким образом, актуальным является разработка гибридных методов и подходов по моделированию биомакромолекулярных комплексов, которые

сочетают возможности построения структурных моделей как на основе экспериментальных данных (в том числе данных с низким информационным содержанием), так и на основе моделирования атом-атомных взаимодействий, а также учитывают структурные динамические особенности многих биомакромолекулярных систем. Такие подходы в последнее время в литературе получили название *интегративного моделирования* (от англ. integrative modeling, см. напр. [58]). Одной из задач интегративного моделирования является в том числе интеграция различного рода экспериментальных данных при построении моделей биомакромолекулярных систем. На Рис. 0.3 представлены различные типы экспериментальных данных, которые могут быть использованы в интегративном моделировании. К ним могут относиться структуры биомакромолекул или отдельных их частей полученные традиционными методами структурной биологии с высоким разрешением, а также набор данных низкого информационного содержания (когда данных недостаточно для реконструкции структуры традиционными методами), полученных различными биофизическими, биохимическими и другими методами. К последним могут относиться данные полученные методами электронной микроскопии низкого разрешения, данные малоуглового рассеяния рентгеновских лучей или нейтронов, данные о взаимодействии введенных в структуру меток, получаемые методами ЯМР, ЭПР, спектроскопии Ферстеровского резонансного переноса энергии (FRET spectroscopy), данные по экспонированности в растворитель различных групп биомакромолекул, получаемые на основе экспериментов по дейтериевому обмену, данные экспериментов по химическому сшиванию или разрушения биомакромолекул, данные получаемые с применением методов геномики (например, методов ChIP-seq, Hi-C и др.). Актуальной является и задача обратного рода – имея некоторую структурно-динамическую модель биомакромолекулярной системы, рассчитать параметры ее отклика в различных типах экспериментов (см. Рис. 0.3). Обсуждаемые подходы также важны при дизайне и конструировании новых биомакромолекулярных комплексов с полезными функциями.

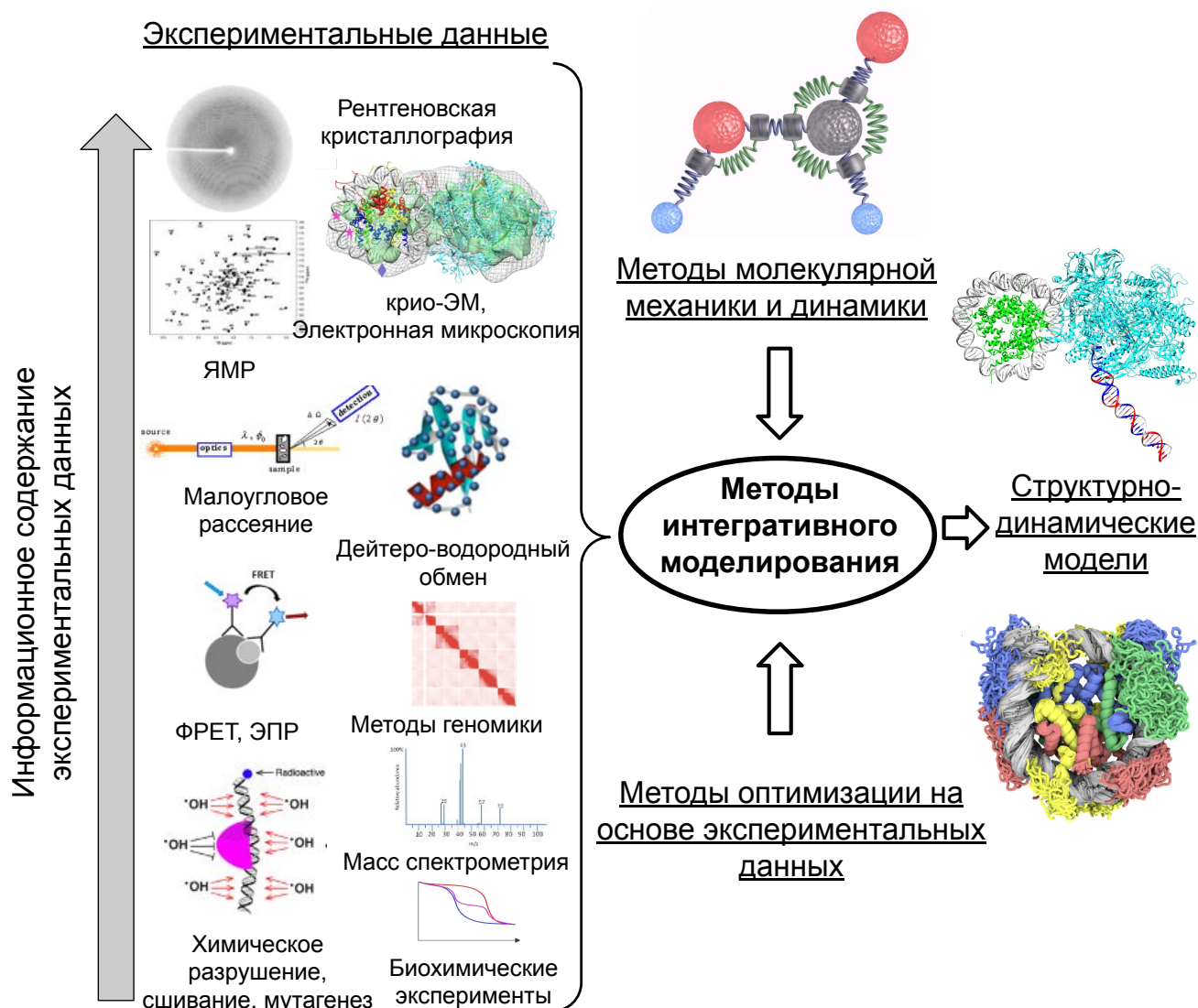


Рисунок 0.3 — Идея методов интегративного моделирования.

Список литературы

1. Structural Biology | Nature / <https://www.nature.com/subjects/structural-biology>. URL: <https://www.nature.com/subjects/structural-biology> (дата обр. 22.06.2019).
2. *Sanger, F.* The Amino-Acid Sequence in the Phenylalanyl Chain of Insulin. 1. The Identification of Lower Peptides from Partial Hydrolysates / F. Sanger, H. Tuppy // *Biochemical Journal*. — 1951. — Sept. 1. — Vol. 49, no. 4. — P. 463—481. — DOI: [10.1042/bj0490463](https://doi.org/10.1042/bj0490463).
3. Structure of a Ribonucleic Acid / R. W. Holley, J. Apgar, G. A. Everett, J. T. Madison, M. Marquisee, S. H. Merrill, J. R. Penswick, A. Zamir // *Science*. — 1965. — Mar. 19. — Vol. 147, no. 3664. — P. 1462—1465. — DOI: [10.1126/science.147.3664.1462](https://doi.org/10.1126/science.147.3664.1462).
4. *Gilbert, W.* The Nucleotide Sequence of the Lac Operator / W. Gilbert, A. Maxam // *Proceedings of the National Academy of Sciences*. — 1973. — Dec. 1. — Vol. 70, no. 12. — P. 3581—3584. — DOI: [10.1073/pnas.70.12.3581](https://doi.org/10.1073/pnas.70.12.3581).
5. *Maxam, A. M.* A New Method for Sequencing DNA / A. M. Maxam, W. Gilbert // *Proceedings of the National Academy of Sciences*. — 1977. — Feb. 1. — Vol. 74, no. 2. — P. 560—564. — DOI: [10.1073/pnas.74.2.560](https://doi.org/10.1073/pnas.74.2.560).
6. *Sanger, F.* DNA Sequencing with Chain-Terminating Inhibitors / F. Sanger, S. Nicklen, A. R. Coulson // *Proceedings of the National Academy of Sciences*. — 1977. — Dec. 1. — Vol. 74, no. 12. — P. 5463—5467. — DOI: [10.1073/pnas.74.12.5463](https://doi.org/10.1073/pnas.74.12.5463).
7. Первичная Структура Валиновой Транспортной РНК 1 Пекарских Дрожжей / А. Баев, Т. Венкстерн, А. Мирзабеков, А. Крутилина, Л. Ли, В. Аксельрод // *Молекулярная биология*. — 1967. — Т. 1, № 5. — С. 745—766.
8. A Three-Dimensional Model of the Myoglobin Molecule Obtained by x-Ray Analysis. / К. Јс, В. G, D. Hm, P. Rg, W. H, P. Dc // *Nature*. — 1958. — Mar. 1. — Vol. 181, no. 4610. — P. 662—666. — DOI: [10.1038/181662a0](https://doi.org/10.1038/181662a0).
9. Кристаллография и жизнь / под ред. М. V. Kovalchuk. — Moskva : Fizmatlit, 2012. — 376 с. — ISBN 978-5-9221-1403-5.

10. *Mannige, R. V.* Dynamic New World: Refining Our View of Protein Structure, Function and Evolution / R. V. Mannige // *Proteomes*. — 2014. — 7 марта. — Т. 2, № 1. — С. 128—153. — DOI: [10.3390/proteomes2010128](https://doi.org/10.3390/proteomes2010128).
11. *Needleman, D.* Active Matter at the Interface between Materials Science and Cell Biology / D. Needleman, Z. Dogic // *Nature Reviews Materials*. — 2017. — July 20. — Vol. 2, issue 9, no. 9. — P. 1—14. — DOI: [10.1038/natrevmats.2017.48](https://doi.org/10.1038/natrevmats.2017.48).
12. *Finney, J. L.* Bernal's Road to Random Packing and the Structure of Liquids / J. L. Finney // *Philosophical Magazine*. — 2013. — 5 нояб. — Т. 93, № 31—33. — С. 3940—3969. — DOI: [10.1080/14786435.2013.770179](https://doi.org/10.1080/14786435.2013.770179).
13. *Gregg, C.* MNCP Software Quality: Then and Now / C. Gregg. — URL: <https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-00-2532>.
14. English: ENIAC (Electronic Numerical Integrator And Computer) in Philadelphia, Pennsylvania. Glen Beck (Background) and Betty Snyder (Foreground) Program the ENIAC in Building 328 at the Ballistic Research Laboratory (BRL). [https:// Commons.Wikimedia.Org / Wiki/ File:Eniac.Jpg](https://commons.wikimedia.org/wiki/File:Eniac.jpg). — [1947].
15. *Studies of the Nonlinear Problems* / E. Fermi, P. Pasta, S. Ulam, M. Tsingou ; Los Alamos Scientific Lab., N. Mex. — 05/01/1955. — LA—1940. — DOI: [10.2172/4376203](https://doi.org/10.2172/4376203).
16. *Гривцов, А.Г.* Методика Численных Экспериментов и Динамика Микрогетерогенных Систем / Гривцов, А.Г. // *Метод Молекулярной Динамики в Физической Химии*. — Москва : Наука, 1996. — С. 16. — ISBN 5-02-001907-0.
17. Overview of the CCP4 Suite and Current Developments / M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson // *Acta Crystallographica Section D: Biological Crystallography*. — 2011. — Apr. 1. — Vol. 67, issue 4, no. 4. — P. 235—242. — DOI: [10.1107/S0907444910045749](https://doi.org/10.1107/S0907444910045749).

18. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix / D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams // *Acta Crystallographica Section D: Structural Biology*. — 2019. — Oct. 1. — Vol. 75, issue 10, no. 10. — P. 861–877. — DOI: [10.1107/S2059798319011471](https://doi.org/10.1107/S2059798319011471).
19. *Emsley, P.* Coot: Model-Building Tools for Molecular Graphics / P. Emsley, K. Cowtan // *Acta Crystallographica Section D: Biological Crystallography*. — 2004. — Dec. 1. — Vol. 60, issue 12, ARRAY(0xa66f734). — P. 2126–2132. — DOI: [10.1107/S0907444904019158](https://doi.org/10.1107/S0907444904019158).
20. *Schrodinger.* The PyMOL Molecular Graphics System, Version 1.8 / Schrodinger. — 2015.
21. UCSF Chimera—a Visualization System for Exploratory Research and Analysis / E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin // *Journal of Computational Chemistry*. — 2004. — Oct. — Vol. 25, no. 13. — P. 1605–1612. — DOI: [10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084).
22. *Ломоносов, М.* Полное Собрание Сочинений: Ломоносов Михаил Васильевич — Алфавитный Каталог — Электронная Библиотека Руниверс / М. Ломоносов. — 1950—1983. — URL: <https://runivers.ru/lib/book6882/> (дата обр. 22.09.2020).
23. *Bernal, J.* The Bakerian Lecture, 1962 The Structure of Liquids / J. Bernal // *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*. — 1964. — 28 июля. — Т. 280, № 1382. — С. 299–322. — DOI: [10.1098/rspa.1964.0147](https://doi.org/10.1098/rspa.1964.0147).
24. *Breathnach, C. S.* Desmond Bernal and His Role in the Biological Exploitation of X-Ray Crystallography / C. S. Breathnach // *Journal of Medical Biography*. — 1995. — Nov. — Vol. 3, no. 4. — P. 197–200. — DOI: [10.1177/096777209500300403](https://doi.org/10.1177/096777209500300403).

25. *Kevin, J.* Schrödinger Desmond High-Performance Molecular Dynamics Simulations / J. Kevin ; Schrödinger. — 2020. — URL: <https://www.schrodinger.com/desmond>.
26. *Frenkel, D.* Understanding molecular simulation / D. Frenkel, B. Smit. — Academic Press, 2002.
27. *Hey, T.* The Fourth Paradigm: Data-Intensive Scientific Discovery / T. Hey, S. Tansley, K. Tolle. — Microsoft Research, 10.2009. — ISBN 978-0-9825442-0-4.
28. *Eckhardt, R.* And the MONTE CARLO METHOD / R. Eckhardt // Los Alamos Science Special Issue. — 1987. — P. 131—141.
29. *Metropolis, N.* The Monte Carlo Method / N. Metropolis, S. Ulam // Journal of the American Statistical Association. — 1949. — Sept. — Vol. 44, no. 247. — P. 335—341. — DOI: [10.1080/01621459.1949.10483310](https://doi.org/10.1080/01621459.1949.10483310).
30. *Metropolis, N.* The Beginnig of the Monte Carlo Method / N. Metropolis // Los Alamos Science Special Issue. — 1987. — C. 125—130.
31. Equation of State Calculations by Fast Computing Machines / N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller // The Journal of Chemical Physics. — 1953. — 1 июня. — Т. 21, № 6. — С. 1087—1092. — DOI: [10.1063/1.1699114](https://doi.org/10.1063/1.1699114).
32. *Иванов, В.* Методы Компьютерного Моделирования Для Исследования Полимеров и Биополимеров / В. Иванов. — 2009. — URL: <http://www.chem.msu.ru/rus/books/polymer/welcome.html> (дата обр. 18.09.2020).
33. *Alder, B. J.* Phase Transition for a Hard Sphere System / B. J. Alder, T. E. Wainwright // The Journal of Chemical Physics. — 1957. — 1 нояб. — Т. 27, № 5. — С. 1208—1209. — DOI: [10.1063/1.1743957](https://doi.org/10.1063/1.1743957).
34. Dynamics of Radiation Damage / J. B. Gibson, A. N. Goland, M. Milgram, G. H. Vineyard // Physical Review. — 1960. — 15 нояб. — Т. 120, № 4. — С. 1229—1253. — DOI: [10.1103/PhysRev.120.1229](https://doi.org/10.1103/PhysRev.120.1229).
35. *Rahman, A.* Correlations in the Motion of Atoms in Liquid Argon / A. Rahman // Physical Review. — 1964. — 19 окт. — Т. 136, 2A. — A405—A411. — DOI: [10.1103/PhysRev.136.A405](https://doi.org/10.1103/PhysRev.136.A405).

36. *Allen, M. P.* Computer Simulation of Liquids / M. P. Allen, D. J. Tildesley. — New York, NY, USA : Clarendon Press, 1989. — ISBN 0-19-855645-4.
37. *Karplus, M.* Molecular Dynamics of Biological Macromolecules: A Brief History and Perspective / M. Karplus // Biopolymers. — 2003. — Mar. — Vol. 68, no. 3. — P. 350—358. — DOI: [10.1002/bip.10266](https://doi.org/10.1002/bip.10266).
38. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD / J. C. Phillips, D. J. Hardy, J. D. C. Maia, J. E. Stone, J. V. Ribeiro, R. C. Bernardi, R. Buch, G. Fiorin, J. Hénin, W. Jiang, R. McGreevy, M. C. R. Melo, B. K. Radak, R. D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V. Kalé, K. Schulten, C. Chipot, E. Tajkhorshid // The Journal of Chemical Physics. — 2020. — 28 июля. — T. 153, № 4. — C. 044130. — DOI: [10.1063/5.0014475](https://doi.org/10.1063/5.0014475).
39. The Xplor-NIH NMR Molecular Structure Determination Package / C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore // Journal of Magnetic Resonance (San Diego, Calif.: 1997). — 2003. — Jan. — Vol. 160, no. 1. — P. 65—73. — DOI: [10.1016/s1090-7807\(02\)00014-9](https://doi.org/10.1016/s1090-7807(02)00014-9).
40. *Sali, A.* Comparative Protein Modelling by Satisfaction of Spatial Restraints / A. Sali, T. L. Blundell // J Mol Biol. — 1993. — Dec. 5. — Vol. 234, no. 3. — P. 779—815. — DOI: [10.1006/jmbi.1993.1626](https://doi.org/10.1006/jmbi.1993.1626).
41. AMBER 17 / D. A. Case, D. S. Cerutti, T. E. C. III, T. Darden, R. Duke, T. J. Giese, H. Gohlke, A. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, D. York, P. Kollman. — University of California, San Francisco.
42. GROMACS 2020 Manual / Lindahl, Abraham, Hess, V. D. Spoel. — 2020. — 1 янв. — DOI: [10.5281/ZENODO.3562512](https://doi.org/10.5281/ZENODO.3562512).
43. An Improved Nucleic Acid Parameter Set for the GROMOS Force Field / T. A. Soares, P. H. Hünenberger, M. A. Kastenholtz, V. Kräutler, T. Lenz, R. D. Lins, C. Oostenbrink, W. F. van Gunsteren // Journal of Computational Chemistry. — 2005. — May. — Vol. 26, no. 7. — P. 725—737. — DOI: [10.1002/jcc.20193](https://doi.org/10.1002/jcc.20193).

44. *Likhachev, I.* Parallelism of different levels in the program of molecular dynamics simulation PUMA-CUDA / I. Likhachev, N. Balabaev // (Mathematical Biology and Bioinformatics). — Pushchino, Moscow Region, Russia, 07.11.2018. — e44. — DOI: [10.17537/icmbb18.53](https://doi.org/10.17537/icmbb18.53).
45. NVIDIA GPU and Amber 16 Configurations / NVIDIA. — URL: <https://www.nvidia.com/es-la/data-center/gpu-accelerated-applications/amber/> (дата обр. 30.09.2020).
46. MDGRAPE-4: A Special-Purpose Computer System for Molecular Dynamics Simulations / I. Ohmura, G. Morimoto, Y. Ohno, A. Hasegawa, M. Taiji // Philosophical transactions. Series A, Mathematical, physical, and engineering sciences. — 2014. — 6 авг. — Т. 372, № 2021. — DOI: [10.1098/rsta.2013.0387](https://doi.org/10.1098/rsta.2013.0387).
47. Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer / D. E. Shaw, J. P. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. Dror, A. Even, C. H. Fenton, A. Forte, J. Gagliardo, G. Gill, B. Greskamp, C. R. Ho, D. J. Ierardi, L. Iserovich, J. S. Kuskin, R. H. Larson, T. Layman, L. Lee, A. K. Lerer, C. Li, D. Killebrew, K. M. Mackenzie, S. Y. Mok, M. A. Moraes, R. Mueller, L. J. Nociolo, J. L. Peticolas, T. Quan, D. Ramot, J. K. Salmon, D. P. Scarpazza, U. B. Schafer, N. Siddique, C. W. Snyder, J. Spengler, P. T. P. Tang, M. Theobald, H. Toma, B. Towles, B. Vitale, S. C. Wang, C. Young // SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis). — 11.2014. — С. 41–53. — DOI: [10.1109/SC.2014.9](https://doi.org/10.1109/SC.2014.9).
48. *Laio, A.* Escaping Free-Energy Minima / A. Laio, M. Parrinello // Proceedings of the National Academy of Sciences. — 2002. — Oct. 1. — Vol. 99, no. 20. — P. 12562–12566. — DOI: [10.1073/pnas.202427399](https://doi.org/10.1073/pnas.202427399).
49. Smoothed Biasing Forces Yield Unbiased Free Energies with the Extended-System Adaptive Biasing Force Method / A. Lesage, T. Lelièvre, G. Stoltz, J. Hénin // The Journal of Physical Chemistry B. — 2017. — 20 авг. — Т. 121, № 15. — С. 3676–3685. — DOI: [10.1021/acs.jpcb.6b10055](https://doi.org/10.1021/acs.jpcb.6b10055).

50. *Marchi, M.* Adiabatic Bias Molecular Dynamics: A Method to Navigate the Conformational Space of Complex Molecular Systems / M. Marchi, P. Ballone // The Journal of Chemical Physics. — 1999. — 11 февр. — Т. 110, № 8. — С. 3697—3702. — DOI: [10.1063/1.478259](https://doi.org/10.1063/1.478259).
51. *Sugita, Y.* Replica-Exchange Molecular Dynamics Method for Protein Folding / Y. Sugita, Y. Okamoto // Chemical Physics Letters. — 1999. — Nov. 26. — Vol. 314, no. 1. — P. 141—151. — DOI: [10.1016/S0009-2614\(99\)01123-9](https://doi.org/10.1016/S0009-2614(99)01123-9).
52. Неравновесная Молекулярная Динамика Наноструктур, Включая Биологические / К. В. Шайтан, Е. В. Турлей, Д. Н. Голик, К. Б. Терешкина, О. В. Левцова, И. В. Федик, А. К. Шайтан, М. П. Кирпичников // Химическая физика. — 2006. — Т. 25, № 9. — С. 31—48. — IF WoS 0.58.
53. Promoting Transparency and Reproducibility in Enhanced Molecular Simulations / M. Bonomi, G. Bussi, C. Camilloni, G. A. Tribello, P. Banáš, A. Barducci, M. Bernetti, P. G. Bolhuis, S. Bottaro, D. Branduardi, R. Capelli, P. Carloni, M. Ceriotti, A. Cesari, H. Chen, W. Chen, F. Colizzi, S. De, M. De La Pierre, D. Donadio, V. Drobot, B. Ensing, A. L. Ferguson, M. Filizola, J. S. Fraser, H. Fu, P. Gasparotto, F. L. Gervasio, F. Giberti, A. Gil-Ley, T. Giorgino, G. T. Heller, G. M. Hocky, M. Iannuzzi, M. Invernizzi, K. E. Jelfs, A. Jussupow, E. Kirilin, A. Laio, V. Limongelli, K. Lindorff-Larsen, T. Löhner, F. Marinelli, L. Martin-Samos, M. Masetti, R. Meyer, A. Michaelides, C. Molteni, T. Morishita, M. Nava, C. Paissoni, E. Papaleo, M. Parrinello, J. Pfaendtner, P. Piaggi, G. Piccini, A. Pietropaolo, F. Pietrucci, S. Pipolo, D. Provasi, D. Quigley, P. Raiteri, S. Raniolo, J. Rydzewski, M. Salvalaglio, G. C. Sosso, V. Spiwok, J. Šponer, D. W. H. Swenson, P. Tiwary, O. Valsson, M. Vendruscolo, G. A. Voth, A. White, The PLUMED consortium // Nature Methods. — 2019. — Aug. — Vol. 16, issue 8, no. 8. — P. 670—673. — DOI: [10.1038/s41592-019-0506-8](https://doi.org/10.1038/s41592-019-0506-8).
54. *Jarzynski, C.* Nonequilibrium Equality for Free Energy Differences / C. Jarzynski // Physical Review Letters. — 1997. — 7 апр. — Т. 78, № 14. — С. 2690—2693. — DOI: [10.1103/PhysRevLett.78.2690](https://doi.org/10.1103/PhysRevLett.78.2690).
55. *Crooks, G. E.* Entropy Production Fluctuation Theorem and the Nonequilibrium Work Relation for Free Energy Differences / G. E. Crooks // Physical Review E. — 1999. — 1 сент. — Т. 60, № 3. — С. 2721—2726. — DOI: [10.1103/PhysRevE.60.2721](https://doi.org/10.1103/PhysRevE.60.2721).

56. Macromolecular Modeling and Design in Rosetta: Recent Methods and Frameworks / J. K. Leman [et al.] // Nature Methods. — 2020. — July. — Vol. 17, issue 7, no. 7. — P. 665—680. — DOI: [10.1038/s41592-020-0848-2](https://doi.org/10.1038/s41592-020-0848-2).
57. *Wu, H.* AWSEM-IDP: A Coarse-Grained Force Field for Intrinsically Disordered Proteins / H. Wu, P. G. Wolynes, G. A. Papoian // The Journal of Physical Chemistry. B. — 2018. — Dec. 13. — Vol. 122, no. 49. — P. 11115—11125. — DOI: [10.1021/acs.jpcc.8b05791](https://doi.org/10.1021/acs.jpcc.8b05791).
58. *Braitbard, M.* Integrative Structure Modeling: Overview and Assessment / M. Braitbard, D. Schneidman-Duhovny, N. Kalisman // Annual Review of Biochemistry. — 2019. — T. 88, № 1. — C. 113—135. — DOI: [10.1146/annurev-biochem-013118-111429](https://doi.org/10.1146/annurev-biochem-013118-111429).