

На правах рукописи

Шайтан Алексей Константинович

**Интегративное моделирование структуры и динамики
биомакромолекулярных комплексов**

Специальность 03.01.09 —
«Математическая биология, биоинформатика»
доктора физико-математических наук

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
доктора физико-математических наук

Работа выполнена на кафедре биоинженерии биологического факультета
МГУ имени М.В.Ломоносова.

Научный д-р бiol. наук, академик РАН
Киричников Михаил Петрович

консультант:

Официальные оппоненты:
Разин Сергей Владимирович,
доктор биологических наук, член-корреспондент РАН,
профессор, заведующий кафедрой молекулярной
биологии биологического факультета МГУ имени
М.В.Ломоносова, главный научный сотрудник, заве-
дующий лабораторией структурно-функциональной
организации хромосом Института биологии гена РАН

Ефремов Роман Гербертович,
доктор физико-математических наук, профессор, глав-
ный научный сотрудник, заведующий лабораторией
моделирования биомолекулярных систем Института био-
органической химии им. академиков М.М. Шемякина и
Ю.А. Овчинникова РАН

Бриллиантов Николай Васильевич,
доктор физико-математических наук,
директор центра по научным и инженерным вычисли-
тельным технологиям для задач с большими массивами
данных, профессор Сколковского института науки и тех-
нологий

Зашита состоится 15 апреля 2021 г. в 14 часов на заседании диссер-
тационного совета МГУ.03.02 при Московском государственном универси-
тете имени М.В.Ломоносова по адресу: 119234, Москва, Ленинские горы,
МГУ, дом 1, стр. 12, Биологический факультет, ауд. М1.

E-mail: maristra@yandex.ru

Диссертация находится на хранении в отделе диссертаций науч-
ной библиотеки МГУ имени М.В. Ломоносова (Ломоносовский просп.,
д. 27). Со сведениями о регистрации участия в защите в удаленном
интерактивном режиме и с диссертацией в электронном виде можно ознаком-
иться, перейти на страницу диссертационного совета по ссылкам: https://istrina.msu.ru/dissertation_councils/councils/32241428/ <https://www.msu.ru/science/dis Sov-msu.html>.

Автореферат разослан — — — — — 2021 года.

Ученый секретарь
диссертационного совета МГУ.03.02,
д-р бiol. наук

Страховская Марина Глебовна

Общая характеристика работы

Актуальность темы.

Понимание механизмов функционирования живых систем на молекулярном и супрамолекулярном уровнях является одной из ключевых задач биологии XXI века. Такое понимание важно не только с фундаментальной точки зрения, но и имеет прямые практические приложения в области медицины, здравоохранения, бионженерии и биотехнологии. Особенностью биологических молекул является формирование ими сложных пространственных структур, которые могут перстриаваться в ходе различных функциональных процессов. Исследование *структурно-динамических* моделей биомакромолекул является одним из ключевых подходов к изучению механизмов функционирования живой материи. Методы компьютерного молекулярного моделирования являются неотъемлемой частью этого подхода. Среди методов построения моделей важное место занимает методы, использующие экспериментальные данные рентгеноструктурного анализа (PCA), ядерно-магнитного резонанса (ЯМР), криоэлектронной микроскопии (криоЭМ). В настоящее время они сочетаются с методами молекулярной динамики (МД), Монте-Карло, позволяющих проводить *компьютерные эксперименты*, имитационное моделирование молекулярных систем. В этой области существуют и ограничения, которые затрудняют исследования биологический систем значительных размеров таких, как большинство биомакромолекулярных комплексов. Сложность интеграции и самого получения данных РСА, ЯМР, криоЭМ возрастает по мере увеличения размера и конформационной подвижности исследуемой системы. По состоянию на 2020 год 92% структур депонированные в базе РДВ имеют молекулярный вес менее 75 кДа. Возможности исследования больших биомакромолекулярных комплексов (от 100 кДа и более) методами молекулярной динамики ограничиваются вычислительными возможностями и качеством параметризации атом-атомных взаимодействий (так называемых силовых полей). Несмотря на то, что благодаря прогрессу компьютерных технологий, методом МД стал возможен расчет систем, состоящих из сотен тысяч и миллионов атомов на временах десятков микросекунд (а с использованием специализированных суперкомпьютеров и на миллисекундных временах), временные масштабы многих физиологически важных процессов, включая важные моды динамики многих биомакромолекулярных комплексов, находятся пока за пределами возможностей моделирования в субсекундном и секундном диапазоне времен. С ростом размера биомолекулярной системы неизбежны задания атом-атомных взаимодействий суммируются и сильнее отражаются на качестве воспроизведения ее динамики, в то время как количество динамических состояний увеличивается, а разница свободной энергии между ними уменьшается.

Шайтан Алексей Константинович

Интегративное моделирование структуры и динамики биомакромолекулярных комплексов

Автореф. дис. на соискание ученой степени докт. физ.-мат. наук
Подписано в печать _____. Заказ № _____

Формат 60×90/16. Авт. печ. л. 2. Тираж 100 экз.

Типография _____

Преодоление вышеописанных ограничений с целью построения и анализа структурно-динамических моделей биомакромолекулярных комплексов является важной задачей. На современном этапе одним из основных путей решения данной задачи является разработка и совершенствование подходов, которые используют для построения моделей как расчет физических взаимодействий между атомами, так и различные экспериментальные данные. К таким экспериментальным данным могут относиться как данные классических методов структурной биологии (данные РСА, ЯМР, крио-ЭМ), так и данные спектральных методов, различных видов микроскопии, биохимического анализа.

Степень разработанности темы исследований.

Методы молекулярного компьютерного моделирования зародились в 1940-ых годах с появлением первых вычислительных систем, однако активное их применение к биологическим молекулам стало возможным в 1970-ых – 1980-ых годах. Активно стали развиваться как подходы к моделированию биомолекул методом МД, так и оптимизационные методы моделирования структуры на основе экспериментальных данных (главным образом РСА и ЯМР). Методы МД основаны на представлении молекулярной системы в виде набора атомов, взаимодействующих согласно законам классической механики Ньютона, и численном решении уравнений движения. На сегодняшний день разработана обширная методическая база для моделирования белков, нуклеиновых кислот, липидов молекулярной динамики, разработаны наборы функциональных форм и параметров, описывающих взаимодействия между атомами (например, семейства силовых полей CHARMM, AMBER, OPLS и др.), алгоритмы численного интегрирования уравнений движения, поддержания температуры, давления, использования параллельных суперкомпьютерных технологий, созданы программные комплексы для расчетов (например, AMBER, CHARMM, GROMACS, NAMD, PUMA и др.). В то же время уровень развития методов МД, как с точки зрения доступных времен моделирования, так и с точки зрения качества силовых полей, не позволяет использовать их полностью независимым образом в отрыве от экспериментальных данных. Стандартной практикой для расчетов методом МД является использование в качестве стартовой конформации биомакромолекулярной системы структурной модели, полученной на основе данных РСА, ЯМР, крио-ЭМ.

Важный вклад в построение структурных моделей биомакромолекулярных систем вносят комбинированные вычислительные подходы, которые, кроме данных РСА, ЯМР, крио-ЭМ, могут использовать различные экспериментальные данные, получаемые в ходе биофизических, биохимических, функциональных, спектроскопических и других экспериментов. Такого рода эксперименты могут предоставить информацию о расстояниях между введенными метками в белке или ДНК (например, методы FRET, ЭПР), характере укладки белковой цепи (например,

7. Cole, H. A. The Centromeric Nucleosome of Budding Yeast Is Perfectly Positioned and Covers the Entire Centromere / H. A. Cole, B. H. Howard, D. J. Clark // Proceedings of the National Academy of Sciences. – 2011. – Aug. 2. – Vol. 108, no. 31. – P. 12687–12692. – DOI: [10.1073/pnas.1104978108](https://doi.org/10.1073/pnas.1104978108).
8. Mechanism of Transcription through a Nucleosome by RNA Polymerase II / O. I. Kulaeva, F.-K. Hsieh, H.-W. Chang, D. S. Luse, V. M. Sturdivsky // Biochimica Et Biophysica Acta. – 2013. – Jan. – Vol. 1829, no. 1. – P. 76–83. – DOI: [10.1016/j.bbagen.2012.08.015](https://doi.org/10.1016/j.bbagen.2012.08.015).
9. Free Energy Profiles of Amino Acid Side Chain Analogs near Water-Vapor Interface Obtained via MD Simulations / A. K. Shaytan, V. A. Ivanov, K. V. Shaitan, A. R. Khokhlov // Journal of Computational Chemistry. – 2010. – Jan. 15. – Vol. 31, no. 1. – P. 204–216. – DOI: [10.1002/jcc.21267](https://doi.org/10.1002/jcc.21267). – IF WoS 3.28.

Зарегистрированные патенты и программы

1. Заявка 2580006 Рос. федерация, МПК G 06 F 19/100. Способ скрининга потенциальных противоопухолевых препаратов ингибиторов FACT [Текст] / В. М. Студитский, О. И. Студитская, А. К. Шайтан (Российская Федерация). – № 2013132806/10 ; заявл. 16.07.2013 ; опубл. 27.01.2015, Бюл. № 3 ; приоритет 16.07.2013 (Рос. Федерации). – 10 с.
2. Свидетельство о гос. регистрации программы для ЭВМ. Программный комплекс реконструкции пространственной структуры белков и комплексов на основе карт электронной плотности низкого разрешения [Текст] / Д. Л. Щуров, А. К. Шайтан, Г. А. Армееев, Д. А. Турченков, В. Н. Блинов, М. П. Кирничников, К. В. Шайтан. – № 2013614397 ; заявл. 13.05.2013 ; опубл. 17.07.2013 ; заявл. 13.05.2013 ; заявл. 2013614397 (Рос. Федерации).

ИК-, КД-спектроскопия, дифракция рентгеновских лучей на фибрillах), реакционной доступности химических групп (методы фурье-принтинга, химического сплавления) и др. Такой подход, основанный на интеграции различных экспериментальных данных при построении структурно-динамических моделей, в литературе называется *интегративным моделированием* [1].¹ Программы, используемые для реконструкции структур по данным экспериментов ЯМР, вероятно, первыми начали позволять использовать некоторые типы данных из других экспериментов, например, данные малоуглового рентгеновского и нейтронного рассеяния [2]. Другим примером является использование в методах МД дополнительных членов потенциальной энергии, которые зависят от экспериментально полученных карт электронной плотности низкого разрешения методами электронной микроскопии или PCA. Такой подход позволяет деформировать начальные структуры биомолекул для прицелия им экспериментально наблюдаемой формы [3]. Разрабатываются программные подходы, предстаивающие возможность одновременно использовать данные различных биофизических и биохимических экспериментов (например, карты электронной плотности и расстояния между группами атомов) [4]. В целом, подходы интегративного моделирования могут различаться по (i) уровню представления модели (например, атомистическое или огрубленное представление модели), (ii) способу оценки соответствия модели экспериментальным данным (использование различных силовых полей или оценочных функций), (iii) методам поиска структуры, соответствующей экспериментальным данным, включая использование различных алгоритмов оптимизации, в том числе минимизации энергии, расчетов методом Монте-Карло, (iv) заданию типа варьируемых параметров (координат атомов, групп атомов, обобщенные координаты реакции). В настоящее время универсальных подходов интегративного моделирования не существует, такие подходы требуют специальных методов и программных реализаций под конкретный тип моделируемых систем (например, комплексы белок-ДНК, амилоидоподобные фибрillы, мембранные системы и т.д.) и конкретный набор экспериментальных данных.

Целью данной работы является разработка интегративных подходов к построению и методов анализа структурно-динамических моделей больших ДНК-белковых комплексов и амилоидоподобных фибрill.

В работе поставлены и решены следующие основные задачи:

1. Разработать подходы к интегративному моделированию ДНК-белковых комплексов методом молекулярной динамики и проанализировать динамику нуклеосом с атомистическим уровнем детализации в микросекундном временном диапазоне.

¹Ссылки на список литературы приведены цифрами, ссылки на список статей автора – цифрами со звездочкой.

2. На основе различных экспериментальных данных разработать интегративные подходы к моделированию и верификации моделей комплексов ДНК и белков.
3. Для интегративного моделирования биомакромолекулярных комплексов разработать методы анализа экспериментальных данных по расщеплению ДНК гидроксильными радикалами. Применить данные методы для изучения организаций ДНК в нуклеосомах.
4. Разработать подходы и методы построения моделей пуклеосом с белками хроматина. Применить данные подходы для изучения комплексов нуклеосом с белками Н1, СЕНР-С, комплексом белков FACT.
5. Разработать и применить интегративные подходы по построению атомистических моделей амилоидоподобных фибрill на основе экспериментальных данных спектроскопии и микроскопии.

Объект и предмет исследований. Предмет исследования состоит в разработке новых системных подходов и алгоритмов построения структурно-динамических моделей биомакромолекулярных комплексов на основе методов молекулярного моделирования с использованием различных экспериментальных данных. Объектами исследования являлись ДНК-белковые комплексы, нуклеосомы, комплексы нуклеосом с белками хроматина, амилоидоподобные фибрillы, состоящие из пептидов и конъюгатов пептидов с синтетическими полимерами.

Научная новизна

Предложен оригинальный интегративный подход к созданию моделей ДНК-белковых комплексов с использованием атомистического моделирования белокарбонатных молекул методами молекулярной динамики, отрубленного молекулярного моделирования ДНК, с учетом свойств симметрии ДНК-белковых комплексов и экспериментальных данных по футпринтингу ДНК, распределению электронной плотности в комплексах, спектроскопических данных о расстояниях между мечеными нуклеотидами.

Впервые на атомистическом уровне в микросекундном временном диапазоне рассчитаны траектории молекулярной динамики нуклеосом и определены микроконформационные состояния ДНК в процессах “дыхания” и отворачивания ДНК от октамера гистонов.

Установлена конформация нуклеосом и их комплексов с рядом белков хроматина (белком СЕНР-С, РНК полимеразами, белками комплекса FACT), установлены структуры ряда амилоидоподобных фибрill.

Методами мульти尺度ного молекулярного моделирования с использованием данных ИК-, КД-спектроскопии, рентгеновской дифракции, атомно-силовой микроскопии установлены структуры ряда амилоидоподобных фибрill.

Теоретическая и практическая значимость

- *28. Voltage-Gated Ion Channel Modulation by Lipids: Insights from Molecular Dynamics Simulations / M. A. Kasimova, M. Tarek, A. K. Shaytan, K. V. Shaitan, L. Delemotte // Biochimica et Biophysica Acta (BBA) - Biomembranes. – 2014. – May. – Vol. 1838, no. 5. – P. 1322–1331. – DOI: [10.1016/j.bbamem.2014.01.024](https://doi.org/10.1016/j.bbamem.2014.01.024). – IF WoS 3.4 - (1.2/0.1).
- *29. *Nishi, H.* Physicochemical Mechanisms of Protein Regulation by Phosphorylation / H. Nishi, A. Shlytan, A. R. Panchenko // Frontiers in Genetics. – 2014. – Aug. 7. – Vol. 5. – DOI: [10.3389/fgene.2014.00270](https://doi.org/10.3389/fgene.2014.00270). – IF WoS 3.789 - (1.2/0.4).
- *30. Genome Packaging in EL and Lin68, Two Giant phiKZ-like Bacteriophages of *P. aeruginosa* / O. Sokolova, O. Shaburova, E. Pechnikova, A. Shaytan, S. Krylov, N. Kiselev, V. Krylov // Virology. – 2014. – Nov. – Vol. 468–470. – P. 472–478. – DOI: [10.1016/j.virol.2014.09.022](https://doi.org/10.1016/j.virol.2014.09.022). – IF WoS 2.464 - (0.8/0.1).
- *31. Peptide Nanofibrils Boost Retroviral Gene Transfer and Provide a Rapid Means for Concentrating Viruses / M. Yolamanova, C. Meier, A. K. Shaytan, V. Vas, C. W. Bertoncini, F. Arnold, O. Zirraf, [et al.] // Nature Nanotechnology. – 2013. – Feb. – Vol. 8, no. 2. – P. 130–136. – DOI: [10.1038/nnano.2012.248](https://doi.org/10.1038/nnano.2012.248). – IF WoS 31.538 - (0.8/0.2).
- *32. Influence of Interionic Interactions on Functional State and Blocker Binding of Voltage-Gated Potassium Channels / K. V. Shlytan, O. S. Sokolova, A. K. Shaytan, M. A. Kasimova, V. N. Novoseletskei, M. P. Kirpichnikov // Moscow University Biological Sciences Bulletin. – 2013. – Mar. – Vol. 68, no. 1. – P. 8–14. – DOI: [10.3103/S0096392513010057](https://doi.org/10.3103/S0096392513010057). – IF RIN 0.76 - (0.8/0.1).
- *33. *Orekhov, P. S.* Calculation of Spectral Shifts of the Mutants of Bacteriorhodopsin by QM/MM Methods / P. S. Orekhov, A. K. Shaytan, K. V. Shlytan // Biophysics. – 2012. – Mar. – Vol. 57, no. 2. – P. 144–152. – DOI: [10.1134/S0006350912020170](https://doi.org/10.1134/S0006350912020170). – IF SJR 0.226 - (1.0/0.2).
- *34. Self-Assembling Nanofibers from Thiophene-Peptide Diblock Oligomers: A Combined Experimental and Computer Simulations Study / A. K. Shaytan, E.-K. Schillinger, P. G. Khalatur, E. Mena-Osteritz, J. Hentschel, H. G. Börner, P. Bäuerle, A. R. Khokhlov // ACS Nano. – 2011. – Sept. 27. – Vol. 5, no. 9. – P. 6894–6909. – DOI: [10.1021/nm2011943](https://doi.org/10.1021/nm2011943). – IF WoS 13.7 - (1.8/1.0).
- *35. Self-Organizing Bioinspired Oligothiophene-Oligopeptide Hybrids / A. K. Shaytan, E.-K. Schillinger, E. Mena-Osteritz, S. Schmid, P. G. Khalatur, P. Bäuerle, A. R. Khokhlov // Beilstein Journal of Nanotechnology. – 2011. – Sept. 5. – Vol. 2. – P. 525–544. – DOI: [10.3762/bjnano.2.57](https://doi.org/10.3762/bjnano.2.57). – IF WoS 2.44 - (2.3/2.0).

- *21. Armeev, G. A. Molecular Dynamics Study of the Ionic Environment and Electrical Characteristics of Nucleosomes / G. A. Armeev, K. V. Shaitan, A. K. Shaitan // Moscow University Biological Sciences Bulletin. — 2015. — Oct. — Vol. 70, no. 4. — P. 173–176. — DOI: [10.3103/S0096392515040033](https://doi.org/10.3103/S0096392515040033). — IF RINC 0.76 - (0,5/0,2).
- *22. Direct Prediction of Residual Dipolar Couplings of Small Molecules in a Stretched Gel by Stochastic Molecular Dynamics Simulations: Direct Prediction of Residual Dipolar Couplings by Stochastic MD Simulations / A. O. Frank, J. C. Freudenberger, A. K. Shaytan, H. Kessler, B. Luy // Magnetic Resonance in Chemistry. — 2015. — Mar. — Vol. 53, no. 3. — P. 213–217. — DOI: [10.1002/mrc.4181](https://doi.org/10.1002/mrc.4181). — IF WoS 1.731 - (0,6/0,1).
- *23. Structural Analysis of Nucleosomal Barrier to Transcription / D. A. Gaykalova, O. I. Kulaeva, O. Volokh, A. K. Shaytan, F.-K. Hsieh, M. P. Kirpichnikov, O. S. Sokolova, V. M. Studitsky // Proceedings of the National Academy of Sciences. — 2015. — Oct. 27. — Vol. 112, no. 43. — E5787–E5795. — DOI: [10.1073/pnas.1508371112](https://doi.org/10.1073/pnas.1508371112). — IF WoS 9.412 - (1,0/0,3).
- *24. Structural Perspectives on the Evolutionary Expansion of Unique Protein-Protein Binding Sites / A. Gonçalves, A. K. Shaytan, B. A. Shoemaker, A. R. Panchenko // Biophysical Journal. — 2015. — Sept. — Vol. 109, no. 6. — P. 1295–1306. — DOI: [10.1016/j.bpj.2015.06.056](https://doi.org/10.1016/j.bpj.2015.06.056). — IF WoS 3.665 - (1,4/0,4).
- *25. Shaytan, A. K. Nucleosome Adaptability Conferred by Sequence and Structural Variations in Histone H2A-H2B Dimers / A. K. Shaytan, D. Landsman, A. R. Panchenko // Current Opinion in Structural Biology. — 2015. — June. — Vol. 32. — P. 48–57. — DOI: [10.1016/j.sbi.2015.02.004](https://doi.org/10.1016/j.sbi.2015.02.004). — IF WoS 6.908 - (1,2/1,0).
- *26. Comparative Computational Study of Interaction of C60-Fullerene and Tris-Malonyl-C60-Fullerene Isomers with Lipid Bilayer: Relation to Their Antioxidant Effect / M. E. Bozdaganyan, P. S. Orekhov, A. K. Shaytan, K. V. Shaitan // PLoS ONE / ed. by C. M. Soares. — 2014. — July 14. — Vol. 9, no. 7. — e102487. — DOI: [10.1371/journal.pone.0102487](https://doi.org/10.1371/journal.pone.0102487). — IF WoS 2.740 - (0,9/0,1).
- *27. Analysis of the Mechanism of Nucleosome Survival during Transcription / H.-W. Chang, O. I. Kulaeva, A. K. Shaytan, M. Kibarov, K. Kuznedelov, K. V. Severinov, M. P. Kirpichnikov, D. J. Clark, V. M. Studitsky // Nucleic Acids Research. — 2014. — Feb. — Vol. 42, no. 3. — P. 1619–1627. — DOI: [10.1093/nar/gkt1120](https://doi.org/10.1093/nar/gkt1120). — IF WoS 11.501 - (1,0/0,2).

Разработанные в данной работе методы и подходы позволяют решать ряд актуальных научных и практических задач, связанных с изучением структурно-динамических характеристик биомакромолекулярных комплексов. Результаты работы применимы для изучения упаковки ДНК в хроматине на нуклеосомному /супрануклеосомному уровнях и упаковки пептидов в амилоидоподобных фибриллах. Структурно-динамические модели нуклеосом и их комплексов могут быть использованы в разработке лекарственных препаратов, направленных на модуляцию эпигенетических механизмов в клетке, например, для поиска ингибиторов связывания белков хроматина с нуклеосомами или ингибиторов посттрансляционных модификаций гистонов. Полученные модели амилоидоподобных фибрилл служат основой для дизайна функциональных самособирающихся фильтров (электропроводящих фибрилл, фибрилл, усиливающих вирусную трансдукцию).

Методология и методы исследования.

Разработанные в работе комплексный подход к изучению структуры и динамики биомакромолекулярных комплексов основан на комбинации различных методов атомистической молекулярной динамики, огрубленного моделирования, оптимизации геометрии молекулярной системы при заданных ограничениях на экспериментально определенные параметры системы, отбора конформаций молекулярной системы, соответствующих определенным экспериментальным параметрам, оцифровки экспериментальных данных, с целью определения численных значений параметров, используемых при моделировании.

Методы атомистической молекулярной динамики позволяют проследить за характером временной эволюции, оценить оптимальность стартовой структуры молекулярной системы, изучить ее конформационную релаксацию, охарактеризовать ансамбль динамических состояний стартовой структуры. В результате МД расчетов становится возможным оценить различные макропараметры системы, связанные с набором динамических состояний системы, и сравнить их значения с данными с экспериментальными данными. В случае нуклеотидов таким параметром, например, является доступность нуклеотидов для атаки и расщепления гидроксильными радикалами, которую можно сравнить с результатами экспериментов по гидроксильному футипринтигу. В случае амилоидоподобных фибрилл – шаг спирали фибриллы, который можно сравнить с данными атомно-силовой микроскопии, а структурные факторы – с данными рентгеновской дифракции.

Используемые нами методы огрубленного моделирования основаны на представлении ДНК в виде последовательности Уотсон-Криковских пар оснований, где каждая пара представляется в виде плоского абстрактного элемента. Взаимное расположение соседних вдоль по цепи элементов определяется шестью параметрами Rise, Shift, Slide, Twist, Tilt, Roll (три

оптимизают за поступательные, а три за вращательные степени свободы).

Функция потенциальной энергии задается в гармоническом динуклеотидном приближении, то есть зависит от квадратов смещений этих параметров вдоль цепи ДНК. Такое представление ДНК позволяет достаточно быстро проводить оптимизацию ее геометрии путем поиска минимума энергии. Задание внешних экспериментальных ограничений, например, ограничений на расстояния между определенными нуклеотидами, измеренные методом Ферстеровского резонансного переноса энергии (FRET), позволяет рассчитывать модели, удовлетворяющие экспериментальным данным.

Важным методологическим подходом в нашем исследовании является также метод (автоматизированного или полу-автоматизированного) перебора возможной структурной организации биомакромолекулярного комплекса среди большого набора детерминированных вариантов и оценка его соответствия экспериментальным данным, а также оценка его относительной энергетической выгодности. Таким образом решались, в частности, задачи по поиску типов укладки бета-листов в амилоидоподобных фибриллах. Например, в данных ИК- и КД-спектров проявление пиков на определенной длине волн можно связать с наличием в структуре бета-листов с параллельной или антипараллельной укладкой бета-нитей, ламелльные рентгеновской дифракции с определенным расстоянием между бета-листами. По этим параметрам можно проводить отбор конформаций генерируемых молекулярных систем, соответствующих определенной стартовой структуре и ее эволюции. Аналогично, при построении моделей нуклеосом с определенным положением ДНК, можно использовать данные гидроксильного футпринтинга ДНК для отбора моделей с правильным положением ДНК на нуклеосоме.

При моделировании методами молекулярной динамики использовались программные пакеты NAMD, LAMMPS и Gromacs различных версий, силовые поля AMBER, CHARMM, PCFF с различными моделями воды и параметрами ионов. Большинство процедур обработки данных, автоматизации и огрубленного моделирования производились с применением языка Python, библиотек NumPy, SciPy, MDAnalysis и других. В ходе работы был разработан ряд программных библиотек для обработки экспериментальных данных (HYDORID, HYDROID_GUI), а также алгоритмы и копиейеры для моделирования нуклеосом, комплексов нуклеосом с белками хроматина, амилоидоподобных фибрилл.

Основные положения, выносимые на защиту:

1. Для построения структурно-динамических моделей сложных биомакромолекулярных комплексов концептуально обосновано применение новых интерактивных подходов на основе сочетания методов атомистического и огрубленного молекулярного моделирования, методов учета разнородных экспериментальных данных рентгеноструктурного анализа, атомно-силовой и электронной

*14. HistoneDB 2.0: A Histone Database with Variants—an Integrated Resource to Explore Histones and Their Variants / E. J. Drazen, A. K. Shaytan, L. Marín-Ramírez, P. B. Talbert, D. Landsman, A. R. Panchenko // Database. — 2016. — Vol. 2016. — baw014. — DOI: [10.1093/database/baw014](https://doi.org/10.1093/database/baw014). — IF WoS 2.593 - (1,2/0,6).

*15. Structure and Functions of Linker Histones / A. V. Lyubitelev, D. V. Nikitin, A. K. Shaytan, V. M. Studitsky, M. P. Kirpichnikov // Biochemistry (Moscow). — 2016. — Mar. — Vol. 81, no. 3. — P. 213–223. — DOI: [10.1134/S0006297916030032](https://doi.org/10.1134/S0006297916030032). — IF WoS 1.978 - (1,3/0,1).

*16. Shaitan, K. V. The Dynamics of Irreversible Evaporation of a Water-Protein Droplet and the Problem of Structural and Dynamic Experiments with Single Molecules / K. V. Shaitan, G. A. Armeev, A. K. Shaytan // Biophysics. — 2016. — Mar. — Vol. 61, no. 2. — P. 177–184. — DOI: [10.1134/S0006350916020172](https://doi.org/10.1134/S0006350916020172). — IF SJR 0.226 - (0,9/0,1).

*17. Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale: Atomistic Insights into Nucleosome Functions / A. K. Shaytan, G. A. Armeev, A. Goncharenco, V. B. Zhurkin, D. Landsman, A. R. Panchenko // Journal of Molecular Biology. — 2016. — Jan. — Vol. 428, no. 1. — P. 221–237. — DOI: [10.1016/j.jmb.2015.12.004](https://doi.org/10.1016/j.jmb.2015.12.004). — IF WoS 5.04 - (2,0/1,8).

*18. Trajectories of Microsecond Molecular Dynamics Simulations of Nucleosomes and Nucleosome Core Particles / A. K. Shaytan, G. A. Armeev, A. Goncharenco, V. B. Zhurkin, D. Landsman, A. R. Panchenko // Data in Brief. — 2016. — June. — Vol. 7. — P. 1678–1681. — DOI: [10.1016/j.dib.2016.04.073](https://doi.org/10.1016/j.dib.2016.04.073). — IF WoS 0.97 - (0,5/0,5).

*19. Large-Scale ATP-Independent Nucleosome Unfolding by a Histone Chaperone / M. E. Valieva, G. A. Armeev, K. S. Kudryashova, N. S. Gerasimova, A. K. Shaytan, O. I. Kulaeva, L. L. McCullough, [et al.] // Nature Structural & Molecular Biology. — 2016. — Dec. — Vol. 23, no. 12. — P. 1111–1116. — DOI: [10.1038/nsmb.3321](https://doi.org/10.1038/nsmb.3321). — IF WoS 11.98 - (0,9/0,1).

*20. Armeev, G. A. Conformational Flexibility of Nucleosomes: A Molecular Dynamics Study / G. A. Armeev, K. V. Shaitan, A. K. Shaytan // Moscow University Biological Sciences Bulletin. — 2015. — July. — Vol. 70, no. 3. — P. 147–151. — DOI: [10.3103/S0096392515030025](https://doi.org/10.3103/S0096392515030025). — IF RINC 0.76 - (0,6/0,3).

- *7. Hydroxyl-Radical Footprinting Combined with Molecular Modeling Identifies Unique Features of DNA Conformation and Nucleosome Positioning / A. K. Shaytan, H. Xiao, G. A. Armeev, C. Wu, D. Landsman, A. R. Panchenko // Nucleic Acids Research. — 2017. — Sept. 19. — Vol. 45, no. 16. — P. 9229–9243. — DOI: [10.1093/nar/gkx616](https://doi.org/10.1093/nar/gkx616). — IF WoS 11.501 - (1,7/1,5).
- *8. Gribkova, A. K. Investigation of Histone-DNA Binding Energy as a Function of DNA Unwrapping from Nucleosome Using Molecular Modeling / A. K. Gribkova, G. A. Armeev, A. K. Shaytan // Moscow University Biological Sciences Bulletin. — 2017. — July. — Vol. 72, no. 3. — P. 142–145. — DOI: [10.3103/S009639251703004X](https://doi.org/10.3103/S009639251703004X). — IF RINN 0.76 - (0,5/0,2).
- *9. MS_HistoneDB, a Manually Curated Resource for Proteomic Analysis of Human and Mouse Histones / S. El Kenmani, A. Adrait, A. K. Shaytan, S. Khochbin, C. Bruley, A. R. Panchenko, D. Landsman, D. Pfleiger, J. Govin // Epigenetics & Chromatin. — 2017. — Dec. — Vol. 10, no. 1. — DOI: [10.1186/s13072-016-0109-x](https://doi.org/10.1186/s13072-016-0109-x). — IF WoS 5.333 - (2,1/0,5).
- *10. Dual Active Site in the Endolytic Transglycosylase Gp144 of Bacteriophage phiKZ / O. V. Chertkov, G. A. Armeev, I. V. Uporov, S. A. Legotsky, N. N. Sykilinda, A. K. Shaytan, N. L. Klyuchko, K. A. Miroshnikov // Acta Naturae. — 2017. — Vol. 9, no. 1. — P. 7. — DOI: [10.32607/20758251-2017-9-1-81-87](https://doi.org/10.32607/20758251-2017-9-1-81-87). — IF WoS 1.62 - (0,8/0,1).
- *11. Modeling of the Structure of Protein-DNA Complexes Using the Data from FRET and Footprinting Experiments / G. A. Armeev, T. K. Gorkovets, D. A. Efimova, K. V. Shaitan, A. K. Shaytan // Moscow University Biological Sciences Bulletin. — 2016. — Jan. — Vol. 71, no. 1. — P. 29–33. — DOI: [10.3103/S0096392516010016](https://doi.org/10.3103/S0096392516010016). — IF RINN 0.76 - (0,6/0,2).
- *12. Armeev, G. A. Nucleosome Structure Relaxation during DNA Unwrapping: Molecular Dynamics Simulation Study / G. A. Armeev, K. V. Shaitan, A. K. Shaytan // Moscow University Biological Sciences Bulletin. — 2016. — July. — Vol. 71, no. 3. — P. 141–144. — DOI: [10.3103/S0096392516030020](https://doi.org/10.3103/S0096392516030020). — IF RINN 0.76 - (0,5/0,2).
- *13. Genomic Profiling of Multiple Sequentially Acquired Tumor Metastatic Sites from an “Exceptional Responder” Lung Adenocarcinoma Patient Reveals Extensive Genomic Heterogeneity and Novel Somatic Variants Driving Treatment Response / R. Biswas, S. Gao, C. M. Cultraro, T. K. Maity, A. Venugopalan, Z. Abdullaev, A. K. Shaytan, [et al.] // Molecular Case Studies. — 2016. — Nov. — Vol. 2, no. 6. — a001263. — DOI: [10.1101/mcs.a001263](https://doi.org/10.1101/mcs.a001263). — IF WoS 1.750 - (3,1/0,5).

микроскопии, футпринтинга ДНК, ИК-, КД-спектроскопии, измерений расстояний между флуоресцентными метками на основе эффекта Ферстеровского резонансного переноса энергии.

2. С использованием разработанного интегративного подхода возможно создание атомистических моделей нуклеосом и комплексов нуклеосом с белками хроматина, при этом учет симметрии белковых комплексов позволяет значительно повысить точность построения молекулярных моделей на основе данных футпринтинга ДНК.

3. Интегративное моделирование позволяет воспроизвести на атомистическом уровне функциональную динамику нуклеосом, важную с точки зрения эпигенетической регуляции функционирования генома, включая крупномасштабные конформационные перестройки структуры ДНК-белковых комплексов (углы входа-выхода ДНК в нуклеосоме, диффузия гистоновых хвостов вдоль ДНК), а также позволяет обнаружить новые моды динамичской подвижности, связанные с изменением конформации ДНК, перестройкой взаимодействий гистоновых хвостов, деформацией глобулярных доменов гистонов.
4. Разработанный подход интегративного моделирования применяется для построения молекулярных моделей амилоидоподобных фибрill, реконструкции укладки пептидов в фибрillах и установления связи между морфологией фибрill и межмолекулярной укладкой пептидов.

Достоверность полученных результатов обеспечивается их публикацией в рецензируемых журналах международного уровня с высокими импакт-факторами. Результаты находятся в соответствии с результатами, полученными другими авторами. Материалы докторской работы доказывались и обсуждались на научных семинарах в МГУ, ряде международных университетов и исследовательских центров (включая Национальные Институты Здоровья, СПА, Университет г. Ульма, Германия, Университет Джонса Хопкинса, СПА, Классический Университет Центрального Китая, Китай и др.). Основные результаты работы до-кладывались на следующих научных конференциях: Biophysical Society 64th Annual Meeting Сан-Диего, СПА, 15–19 февраля 2020; The 44th FEBS Congress, Краков, Польша, 6–11 июля 2019; Multiscale Modeling of Chromatin: Bridging Experiment with Theory, Les Houches, Франция, 31 марта – 5 апреля 2019; Keystone Symposia on Molecular and Cellular Biology “Genomic instability and DNA repair”, Санта Фе, СПА, 2–7 апреля 2017; Biophysical Society 60th Annual Meeting, Los Angeles, СПА, 27 февраля – 2 марта 2016; 27 Международная конференция “Математика. Компьютер. Образование” г. Дубна, Государственный Университет “Дубна”, Россия, 27 января – 1 февраля 2020; Russian International Conference on Cryo-Electron

Microscopy 2019, Москва, Россия, 2–5 июня 2019; Russian International Conference on Cryo-Electron Microscopy 2019, Москва, Россия, 2–5 июня 2019; “Ломоносовские чтения – 2019” Секция “Биология”, Москва, МГУ, Россия, 15–25 апреля 2019; XXVI международная конференция “Математика. Компьютер. Образование”, Пущино, Россия, 28 января – 2 февраля 2019 и др.

Личный вклад. Основные идеи, методология и результаты исследований, изложенные в диссертации, получены автором лично. Вклад автора в совместных экспериментально-теоретических работах заключается в применении разработанных методов, моделировании и интерпретации экспериментальных результатов, полученных коллегами.

Публикации. Основные результаты по теме диссертации изложены в 35 статьях в рецензируемых научных изданиях, индексируемых в базах данных Web of Science, Scopus, RSCI. Зарегистрированы 1 патент и 1 программа для ЭВМ.

Благодарности

Работы, описанные в диссертации были поддержаны рядом российских и международных грантов и стипендий, в том числе, грантами РНФ 18–74–0006, 14-24-00031, 19-74-30003, РФФИ 20-34-70039, 12-04-31942, стипендий сотрудничества России-СПА в области биомедицинских наук, стипендий Национальной медицинской библиотеки СПА, Немецким научно-исследовательским обществом, проект SFB 569 A11. Работа выполнена с использованием оборудования Центра коллективного пользования сверхвысокопроизводительными ресурсами МГУ имени М.В. Ломоносова, кластера BioWulf Национальных институтов здоровья (США), кластера Aldan университета г. Ульм (Германия).

Автор выражает благодарность своим научным руководителям и консультантам, под руководством которых автору посчастливилось работать, М.П.Кирличникову, А.Р.Хохлову, А.Р. Пантенко, Д. Ландсману, П.Г. Халатту, В.А. Иванову, всем соавторам своих научных работ и коллегам за плодотворное сотрудничество, в особенности, Г.А. Армееvu, Б.Б. Журкину, В.М. Студитскому, Х.-В. Чанг, Д.А. Гайкаевой, К.Ву, Х. Жкао, А. Гогчаренко, И. Драйзену, Е.-К. Шиллингер, О.С. Соколовой, А.В. Феофанову, Н.В. Малюченко, М. Вальевой, А. Любителеву и многим другим, коллектикам кафедры физики полимеров и кристаллов физического факультета МГУ, кафедры биоинженерии биологического факультета МГУ, Национального Центра Биотехнологической Информации Национальных Институтов Здоровья за продуктивную рабочую атмосферу и обсуждение работы.

Автор выражает благодарность своей семье за поддержку, без которой написание этой работы не было бы возможным, и А.Д. Шайтану за помощь с версткой текста.

Основные публикации Шайтана Алексея Константиновича по теме диссертации в рецензируемых научных изданиях, индексируемых в базах данных Web of Science, Scopus, RSCI, рекомендованных для защиты в диссертационном совете МГУ по специальности 03.01.09 – «Математическая биология, биинформатика»²

- *1. Histone Octamer Structure Is Altered Early in ISW2 ATP-Dependent Nucleosome Remodeling / A. Hada, S. K. Hota, J. Luo, Y.-c. Lin, S. Kale, A. K. Shaytan, S. K. Bhardwaj, [et al.] // Cell Reports. – 2019. – July 2. – Vol. 28, no. 1. – 282–294.e6. – DOI: 10.1016/j.celrep.2019.05.106. – IF WoS 7.7 – (2.3/0.3).
- *2. The Effect of Oncomutations and Posttranslational Modifications of Histone H1 on Chromatosome Structure and Stability / M. V. Bass, G. A. Armeev, K. V. Shaitan, A. K. Shaytan // Moscow University Biological Sciences Bulletin. – 2019. – July. – Vol. 74, no. 3. – P. 121–126. – DOI: 10.3103/S0096392519030015. – IF RINC 0.76 – (0.7/0.3).
- *3. Linking Chromatin Composition and Structural Dynamics at the Nucleosome Level / G. A. Armeev, A. K. Gribkova, I. Pospelova, G. A. Komarova, A. K. Shaytan // Current Opinion in Structural Biology. – 2019. – June. – Vol. 56. – P. 46–55. – DOI: 10.1016/j.sbi.2018.11.006. – IF WoS 6.908 – (1.2/0.6).
- *4. Structural Interpretation of DNA-Protein Hydroxyl-Radical Footprinting Experiments with High Resolution Using HYDROID / A. K. Shaytan, H. Xiao, G. A. Armeev, D. A. Gaykalova, G. A. Komarova, C. Wu, V. M. Studitsky, D. Landsman, A. R. Panchenko // Nature Protocols. – 2018. – Nov. – Vol. 13, no. 11. – P. 2535–2556. – DOI: 10.1038/s41596-018-0048-z. – IF WoS 11.334 – (2.5/2.2).
- *5. Joint Effect of Histone H1 Amino Acid Sequence and DNA Nucleotide Sequence on the Structure of Chromatosomes: Analysis by Molecular Modeling Methods / T. K. Gorkovets, G. A. Armeev, K. V. Shaitan, A. K. Shaytan // Moscow University Biological Sciences Bulletin. – 2018. – Apr. – Vol. 73, no. 2. – P. 82–87. – DOI: 10.3103/S0096392518020025. – IF RINC 0.76 – (0.7/0.3).
- *6. Molecular Basis of CENP-C Association with the CENP-A Nucleosome at Yeast Centromeres / H. Xiao, F. Wang, J. Wisniewski, A. K. Shaytan, R. Ghirlando, P. C. Fitzgerald, Y. Huang, [et al.] // Genes & Development. – 2017. – Oct. 1. – Vol. 31, no. 19. – P. 1958–1972. – DOI: 10.1101/gad.304782.117. – IF WoS 9.527 – (1.7/0.3).

²В скобках приведен объем публикации в печатных листах и вклад автора в печатных листах.

от гистонового октамера (10 мкм), установлены флукуационные структурные механизмы образования дефектов кручения ДНК, пе- реключения конформаций гистоновых хвостов.

3. Обработка экспериментальных данных по расщеплению ДНК гидроксильными радикалами (футпринтинга ДНК) позволила вычислить вероятность расщепления ДНК для каждого нуклеотида. Показано, что профиля расщепления ДНК гидроксильными радикалами в нуклеосоме мало зависят от последовательности ДНК, и определяются в основном позиционированием ДНК на нуклеосоме. Предложен алгоритм точного определения положения ДНК в нуклеосоме по данным футпринтинга для двух цепей ДНК на основании положения оси псевдосимметрии.

4. С помощью методов интегративного моделирования были усташновлены структуры и параметры взаимодействий комплексов нуклеосом с РНК полимеразами, белком СЕNP-С, гистоном Н1, белками комплекса FACT. Установлено, что в положении +49 после входа в нуклеосому РНК полимераза II может формировать компактный комплекс с нуклеосомой, в котором контакты гистонов с ДНК сохраняются по обе стороны активного центра. В модели центромерной нуклеосомы дрожжей определено положение ДНК и установлено, что белок СЕNP-С взаимодействует с нуклеосомой в районе 20 нуклеотидов от центра симметрии нуклеосомы. Предложены модели конформации линкерных сегментов ДНК при связывании гистона Н1. Установлены амплитуды конформационной подвижности ДНК в нуклеосомах при связывании с комплексом FACT.

5. Разработанные подходы построения моделей амилоидоподобных фибрилл позволили изучить связь взаимного расположения пептидов в фибрillлярных структурах с крупномасштабной морфологией амилоидоподобных фибрилл и таким образом установить структурную организацию филаментов на основе диплок олигомеров кватерцифена и пептида ($Thr - Val$)₃, а также на основе фрагмента белка gp120. Установлено, что рассмотренные амилоидоподобные фибриллы могут образовываться на основе двух взаимодействующих бета-листов, которые формируют либо плоскую ленту, либо левозакрученную ленту с периодом 24-30 нм.

Объем, структура и логика работы. Диссертация состоит из введения, пяти глав и заключения. **Первая глава** посвящена обсуждению общих методических вопросов молекулярного моделирования, систематизируются области применения различных методов, излагаются основы методов молекулярной динамики, огрубленного моделирования, концептуальные подходы интегративного моделирования. **Вторая глава** посвящена применению методов суперкомпьютерной молекулярной динамики для изучения нуклеосом. На примере нуклеосомы МД, делаются выводы о возможностях и ограничениях методов МД, делаются выводы о необходимости применения интегративных подходов к моделированию, учитывающих дополнительные данные. В **третьей главе** показывается, как, используя данные биохимических экспериментов по расщеплению ДНК гидроксильными радикалами, можно проводить моделирование положения ДНК на нуклеосоме и определять ее точную позицию. В **четвертой главе** излагаются подходы по интегративному моделированию более сложных систем – комплексов нуклеосом с белками хроматина. Для решения данной задачи дополнительно используются методы огрубленного моделирования ДНК, данные экспериментов электронной микроскопии и мономолекулярного FRET. **Пятая глава** посвящена разработке и применению методов интегративного моделирования для решения более сложного класса задач – определения структуры амилоидоподобных фибрилл. Особенностью таких систем является полиморфизм локальной структурной организации и необходимость учета ее взаимосвязи с крупномасштабной морфологией комплекса. Полный объём диссертации составляет 488 страниц, включая 106 рисунков и 9 таблиц. Список литературы содержит 677 наименований.

Содержание работы

Во **введении** обосновывается актуальность исследований, проводимых в рамках данной диссертационной работы, формулируется цель, ставится задачи работы, сформулированы научная новизна и практическая значимость представляемой работы, дана общая характеристика работы. **Первая глава (Методы молекулярного и интегративного моделирования в структурной биологии)** посвящена анализу литературы по предмету исследования, систематизации понятийного и терминологического аппарата, изложению принципиальных основ методов и подходов молекулярного и интегративного моделирования. Данная глава является обзорной, основана на анализе и обобщении результатов моделирования различных биологических систем.

В простейшем случае в биологии под моделью биомолекулы понимают пространственное расположение атомов в ее структуре. Однако многие биомакромолекулы обладают множеством конформаций, описание

пекорых систем даже в виде ограниченного набора конформаций за-труднительно. Методы молекулярного моделирования (*molecular modeling, model building*) уже достаточно давно применяются для решения опти-мизационных задач на основе экспериментальных данных в структурной биологии, например, данных РСА или ЯМР. Такие подходы обычно позволяют получать статические модели беломолекул при наличии экспериментальных данных достаточного качества и количества. Другим подходом является имитационное моделирование (*molecular simulations*), разновидность компьютерных экспериментов, в которых расчет свойств молекулярной системы происходит, исходя из моделирования физиче-ских взаимодействий между атомами. Наиболее известным подходом является метод молекулярной динамики, который открывает возмож-ность описания динамики, конформационной подвижности молекулярных систем. Однако у существующих подходов к моделированию беломак-ромолекулярных систем есть ряд существенных ограничений. К ним относятся сложности получения качественных экспериментальных дан-ных методами РСА, ЯМР, криоЭМ для больших молекулярных систем, ограничения по временам моделирования в методах МД из-за вычис-лительной сложности расчетов и ограничения, связанные с качеством силовых полей. Возможным решением данной проблемы является раз-витие методов интегративного моделирования. Методы интегративного моделирования позволяют объединить разные подходы к моделированию беломолекул с учетом разного рода экспериментальных данных. Более по-дробно идея методов интегративного моделирования проиллюстрирована на рисунке 1. Такие данные могут предоставить информацию о форме комплекса (различные виды микроскопии, малоугловое рассеяние), рас-стояниях между атомными группами (FRET, ЭПР, ЯМР), вероятностях взаимодействия с химическими агентами (футпринтинг, дейтеро-водород-ный обмен), контактах различных групп и доменов (химическое сплавление, масс-спектрометрия, методы 3D геномики), наличия определенных паттер-нов структурной организации (ИК-, КД-спектроскопия, дифракционные эксперименты) и т.д. Задача интегративного моделирования состоит в под-боре оптимальных структурных или структурно-динамических моделей молекулярных систем, которые с одной стороны удовлетворяют имею-щемуся набору экспериментальных данных, а с другой не противоречат известным или расчетным параметрам межмолекулярных и межатомных взаимодействий. В задачах интегративного моделирования может исполь-зоваться два принципиальных класса вычислительных процедур. Первый класс предполагает добавление в физическую модель системы (например, функцию потенциальной энергии) дополнительных эмпириче-ских членов, способствующих тому, чтобы состояния, удовлетворяющие экспериментальным значениям измеренных макропараметров, станови-лись более энергетически выгодными. В главе обсуждается пример такого

динамика ДНК в процессе реакции, шумовой фон. Однако, в силу на-личия оси симметрии второго порядка в нуклеосоме такие эффекты будут симметричны для двух цепей ДНК и не влияют на алгоритм по-иска положения в ДНК, относительно которого профили симметричны. Аналогичные соображения можно рекомендовать использовать при ин-тегративном моделировании других ДНК белковых систем, обладающих симметрией. К таким системам относятся многие транскрипционные фак-торы, формирующие гомодимеры.

Заключение

В настоящей работе автором развиты новые системные подходы для построения структурно-динамических моделей сложных беломакромолеку-лярных комплексов. Разработанные интегративные методы молекулярного моделирования основаны на сочетании (интеграции) физических моде-лей взаимодействия молекул с информацией получаемой из различных источников экспериментальных данных. В ходе исследования развиты подходы, (i) позволяющие проводить анализ динамики атомистических структур в микросекундном временном диапазоне, (ii) учитывать раз-личные экспериментальные данные при построении моделей (например, данные экспериментов по футпринтингу ДНК, данные электронной мик-роскопии, данные FRET-микроскопии, данные рентгеновской дифракции на фибрillах, атомно-силовой микроскопии и т.д.), (iii) сочетать атоми-стическое и отрубленное представление молекулярных систем (в частности при моделировании ДНК-содержащих комплексов).

Разработанные автором подходы и методы рекомендуется применять для установления структуры и динамической подвижности беломакромоле-кулярных комплексов, размер и свойства которых затрудняют применение стандартных методов структурной биологии. Практическое применение данных подходов продемонстрировано в диссертации на примере ряда био-макромолекулярных систем, в том числе нуклеосом, комплексов нуклеосом с белками хроматина, амилоидоподобных фибрилл.

Выводы диссертационной работы

1. Разработан интегративный подход и программные решения для мультиаспектного моделирования комплексов ДНК и белков, в котором используются различные данные экспериментов по ДНК футпринтингу, электронной микроскопии, FRET-микроскопии и комбинированное представление ДНК в атомистическом и в ли-нуклеотидном приближении.
2. В микросекундном временном диапазоне определены параметры крупномасштабных функциональных конформационных измене-ний структуры нуклеосом: вычислен ансамбль конформаций линкерной ДНК, установлено влияние электростатического оттал-кивания на геометрию сегментов линкерной ДНК, определено характерное время диссоциации концов нуклеосомальной ДНК

структуре и значения которых может сильно повлиять неточность силовых полей, необходимо дополнительно привлекать экспериментальные данные. Вместе с тем, моделирование конформационно устойчивых структурных элементов можно производить на основании данных силовых полей. Примером служат полученные результаты моделирования комплексов нуклеосом с РНК полимеразой или моделирования геометрии линкерной ДНК при связывании гистона H1. Предполагалось, что свободная ДНК в целом должна находиться в В-конформации двойной спирали, согласно параметрам, заложенным в силовом поле, описывающем структуру ДНК. В то же время более тонкие параметры, такие как небольшой изгиб ДНК, уровень ее отворота от нуклеосомы определяются ограничениями, накладываемыми согласно дополнительными экспериментальными данными. В целом аналогичной логике мы следовали при интегративном моделировании амилоподобных фибрill. Были выделены параметры, которые сложно воспроизвести только с помощью моделирования, а именно, параметры укладки пептидов внутри фибрillы, и параметры, которые воспроизводятся моделированием с хордой точностью, а именно, связь морфологии фибрill с межмолекулярной укладкой пептидов. Таким образом, экспериментальные данные использовались частично при выборе укладок, а также для отбора фибрill на уровне их крупномасштабной морфологии.

Таким образом общий принцип разработки интегративных подходов различных биомакромолекулярных комплексов, который можно рекомендовать, состоит в следующем. В первую очередь с учетом известных экспериментальных и литературных данных проводится анализ различных уровней организаций моделируемой системы, выделяются уровни организации, моделирование которых эффективно может быть проведено на основе имеющихся силовых полей, а также уровня организации, при моделировании которых необходимо привлекать дополнительные экспериментальные данные. На следующем этапе необходимо использовать различные интегративные методы моделирования, которые позволяют одновременно учитывать как различные экспериментальные данные, так и данные основанные на представлениях о внутримолекулярной структуре и межмолекулярных взаимодействиях.

В интегративном моделировании весьма полезным походом является использование соображений структурной симметрии. В некоторых случаях они позволяют кардинальным образом упростить задачу, повысить точность моделирования и уменьшить влияние возможных систематических ошибок в экспериментальных данных. Именно этот подход был продемонстрирован при использовании данных фундаментального моделирования нуклеосом. На качество профиля фундаментального моделирования ДНК могут оказывать влияние такие факторы, как наличие мицерных альтернативных продуктов химического расщепления, локальная

подхода для моделирования конформации ДНК по данным FRET. Экспериментально определенные расстояния между мечеными нуклеотидами используются для добавления к физической модели ДНК эмпирического гармонического потенциала, способствующего уменьшению/увеличению расстояния между этими нуклеотидами до экспериментальных значений. Соответственно проводится минимизация энергии системы. Второй класс процедур предполагает автоматизированный или полу-автоматизированных перебор различных физически допустимых вариантов структурной организации молекулярной системы и нахождение той структурной организации, которая удовлетворяет экспериментальным данным. В процессе оценки физической допустимости конкретной модели и ее соответствия экспериментальным данным могут использоваться методы молекулярной динамики. Главы 3-5 более подробно раскрывают данные подходы на примере конкретных систем и наборов экспериментальных данных.

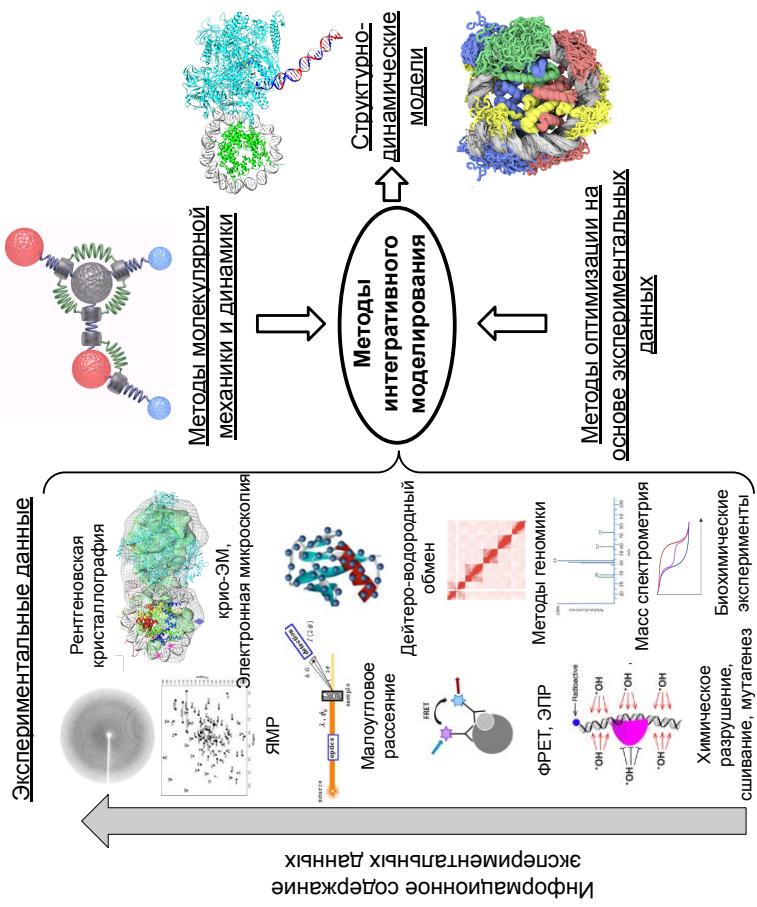


Рис. 1 — Идея методов интегративного моделирования.

В диссертационной работе метод молекулярной динамики является одним из важных компонентов разрабатываемых нами интегративных подходов. Понимание его возможностей и границ применимости необходимо для рационального выбора возможных стратегий моделирования. В связи с этим **вторая глава** (*Применение методов молекулярной динамики для изучения нуклеосом*) иллюстрирует современные возможности и ограничения метода суперкомпьютерной атомистической молекулярной динамики, направленного на получение и анализ структурно-динамических моделей беломакромолекулярных комплексов.

Нуклеосомы являются примером беломакромолекулярного комплекса существенных размеров (около 200 кДа) и играют важную роль в функционировании генома. Они состоят из примерно 200 пар оснований ДНК, организованных гистоновыми белками (Рис. 2а). Центральные 145–147 пар оснований ДНК образуют коровую часть нуклеосомы, плотно наматываются на октамер гистонов в виде ~1,7 витка левой суперспирали [^{*14}; ^{*25}; ⁵]. Нуклеосомы претерпевают множество важных структурных и динамических перестроек во время всех ключевых процессов транскрипции, репликации, reparации ДНК и т.д. [^{*3}]. Способность нуклеосом претерпевать определенные типы конформационных переходов играет важную роль при взаимодействии с белками, включая ремоделирование комплексов, факторы транскрипции, шапероны [^{*19}], РНК-полимеразы [^{*23}] и т.д.

На рисунке 3 проиллюстрированы предполагаемые динамические режимы нуклеосом и влияние на них различных факторов:

последовательности ДНК и гистонов, пост-трансляционных модификаций аминокислот гистонов, взаимодействия с белками. Динамика нуклеосом может происходить спонтанно при данной температуре и нести функционально направленный характер за счет воздействия специальных АТФ-зависимых комплексов. Белковый состав и последовательность ДНК в нуклеосомах влияют на характер динамических процессов в нуклеосомах и таким образом модулируют различные функциональные процессы в геноме. Так показано существование тонких режимов конформационной динамики нуклеосом [^{*3}]. Например, кручение ДНК внутри нуклеосом обеспечивает путь для АТФ-зависимого ремоделирования нуклеосом [6]. Речь идет о существовании особых режимов динамики нуклеосом, в которых важны атомистические детали структурных изменений. Метод молекулярной динамики является мощным инструментом, который может раскрыть природу этих динамических режимов.

Для этого нам были разработаны подходы и алгоритмы моделирования нуклеосом методами молекулярной динамики в полноатомном приближении в микросекундном временном диапазоне с использованием суперкомпьютерных технологий. Разработаны алгоритмы подготовки систем для моделирования и анализа динамического поведения нуклеосом

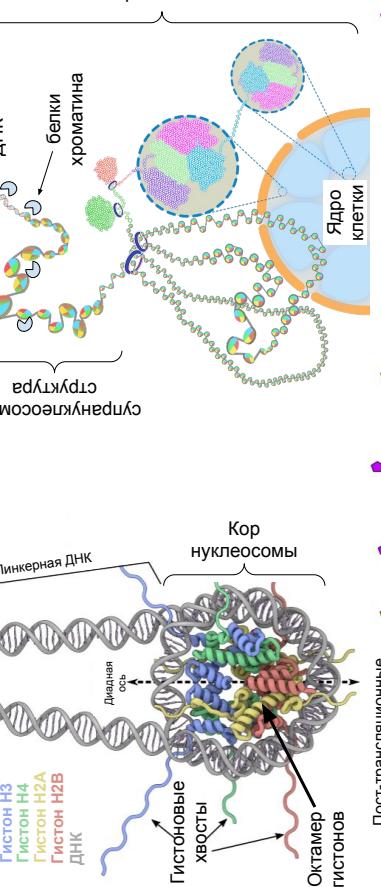
методами МД моделирования, необходимо оценить с одной стороны, в каких временных диапазонах МД модель адекватно описывает систему, а с другой стороны, в каких временных диапазонах возможно компьютерное воспроизведение функционально интересных конформационных перестроек. Результаты нашего моделирования на примере нуклеосом показывают, что современные методы МД позволяют создавать структурно-динамические модели ДНК-белковых комплексов без как-либо видимых артефактов по крайней мере на временах в десятки микросекунд. Свободная ДНК сохраняет конформацию, свойственную В-форме ДНК, ионное окружение является стабильным, глобулярные домены являются стабильными, а неупорядоченные домены сохраняют повышенную конформационную подвижность. В тоже время проведенный нами анализ динамики показал, что на временном диапазоне в пределах одной микросекунды наблюдаются такие процессы, как флуктуации формы свободных участков ДНК, конденсация неупорядоченных белковых доменов по поверхность ДНК, появления локальных конформационных возмущений в участках ДНК, связанных с белком. На временных масштабах порядка десятков микросекунд можно уже изучать диссоциацию/реассоциацию ДНК от белка в отдельных сайтах связывания, частичную диффузию неупорядоченных доменов вдоль поверхности ДНК. В то же время, более крупномасштабные перестройки ДНК-белковых комплексов пока находятся за пределами возможностей МД. Так, вероятно, процессы комплексообразования, например, связывания транскрипционных факторов и поиск ими сайтов на последовательности ДНК или процессы образования гетероструктур ДНК-РНК (например, при связывании комплексов системы CRISPR-Cas9) находятся за пределами возможностей МД. Аналогичная ситуация возникает и при моделировании самосборки амилоидоподобных фибрилл. Несмотря на высокую точность задания силовых полей белков по сравнению с силовыми полями нукleinовых кислот, ситуация здесь усложняется повышенными требованиями к точности определения энергии для различных конформаций белков. Известно, что амилоидоподобные фибриллы обладают конформационным полиморфизмом и в экспериментальных работах небольшое изменение условий (состав растворителя, температура) и, соответственно, энергетического баланса может приводить к формированию фибрилл различного типа.

Именно для решения такого рода задач, мы обратились к развитию и разработке методов интегративного моделирования. Базовая идея разработанных методов интегративного моделирования состоит в объединении информации, заложенной в силовых полях с информацией о взаимном расположении элементов структуры, которую мы получаем из дополнительных экспериментальных методов. Общая логика построения интегративных подходов может быть сформулирована следующим образом. При моделировании структурных элементов и параметров системы, на

методы моделирования для больших комплексов, для изучения свойств которых недостаточно стандартных.

Одним из наиболее широкораспространенных и развитых методов молекулярного моделирования является метод молекулярной динамики (МД). Физические модели молекул в методе МД основаны на рассмотрении взаимодействий между атомами в приближении классической механики Ньютона. Согласно этому для биомолекул по некоторым правилам задается эмпирическая функция зависимости потенциальной энергии системы от координат атомов (так называемое силовое поле). Теоретически, при идеальной точности задания силового поля и неограниченных вычислительных возможностях метод МД может работать в режиме "вычислительного микроскопа", который воспроизводит реальные процессы структурообразования и взаимодействия биологических молекул. Однако на практике мы имеем дело лишь с некоторым приближенным описанием взаимодействий, построенным по аддитивному принципу. Кроме того мы ограничены во временах моделирования, что делает возможным исследование молекулярных систем лишь в ограниченной области конформационного пространства вблизи их стартовых конформаций. Современное развитие методов МД обычно представляет собой повторяющийся двухстадийный процесс, в котором вслед за совершенствованием вычислительных возможностей следует совершенствование силовых полей. Характерным примером здесь является совершение модели ДНК. Так, силовые поля класса AMBER были изначально параметризованы для воспроизведения экспериментальной геометрии ДНК в наносекундном диапазоне, однако, в середине 2000-ых, оказалось, что при временах моделирования свыше 10 нс изменения торсионных углов остива ДНК приводят к неправильной геометрии ДНК, и поэтому силовое поле было усовершенствовано. Были изменены параметры силового поля, отвечающие за энергетические термы вращения торсионных углов в ДНК, что привело к сохранению правильной В-формы ДНК на доступных временах моделирования. В середине 2010-ых ситуация повторилась, когда стандартные времена моделирования увеличились до микросекундного масштаба, и силовое поле было вновь усовершенствовано. Аналогичные примеры можно привести и для силовых полей белков, воды и ионов. Кроме правильного воспроизведения равновесных параметров молекулярных систем, отдельным вопросом является качество воспроизведения их динамики. Здесь с точки зрения силового поля важным является воспроизведение правильного энергетического баланса (заселенности) между различными конформациями. Мы показали [9], что на больших временах моделирования начинают проявляться энтропийные эффекты, которые в случае белков приводят к излишней стабилизации структурных моделей. На практике при исследовании динамических процессов в биомакромолекулярных комплексах

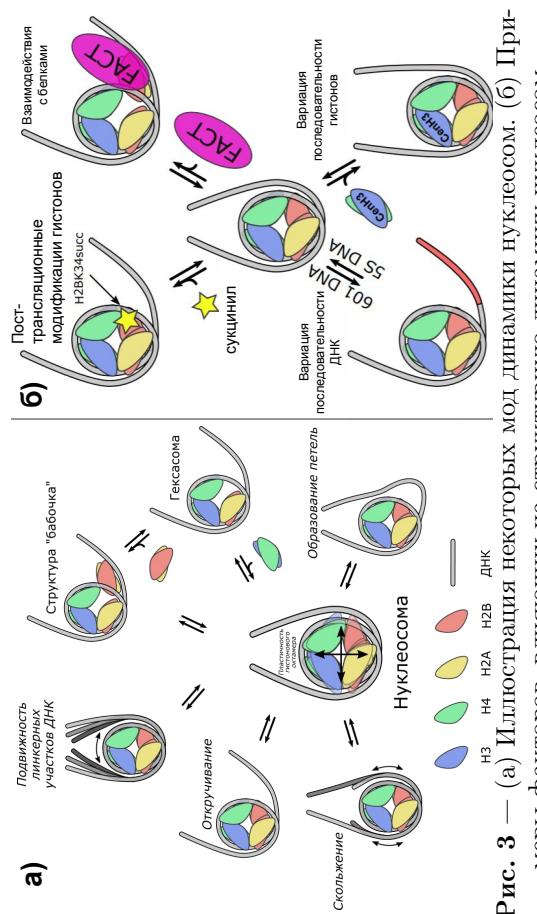
а)



б)

Линкерная ДНК
Белковые хвосты
Октаамер гистонов
Гистоновая фольга
Гистоновые хвосты

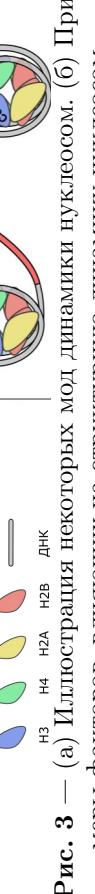
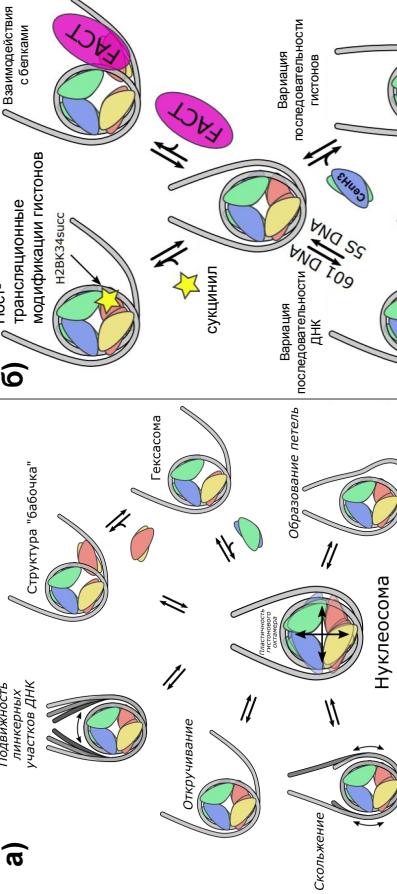
Рис. 2 — (а) Структура нуклеосомы. **(б)** Схема типичной структуры гистонового белка и расположение пост-трансляционных модификаций. **(в)** Место нуклеосом в организации хроматина.



в)

в явном растворителе. В частности, разработаны алгоритмы задания нуклеосомной системы координат на основе определения оси суперспиралей ДНК и оси псевдосимметрии нуклеосом. **(б)** Примеры факторов, влияющих на структурную динамику нуклеосом.

а)



б)

(а) Иллюстрация некоторых моделей динамики нуклеосом. **(б)** Примеры факторов, влияющих на структурную динамику нуклеосом.

определение различных геометрических параметров молекулярной системы таких, как геометрия ДНК, углы вращения ДНК, расположение остина полипептидных цепей.

В разделе 2.2 приводятся результаты равновесного микросекундного моделирования молекулярной динамики нуклеосом, включающих линкерные сегменты ДНК и полноразмерные гистоны (рисунок 2), в ящике растворителя в атомистическом приближении. Было выявлено, что линкерные участки ДНК изменяют свое взаимное расположение из-за электростатического отталкивания, что влияет на углы входа/выхода ДНК из нуклеосом и следовательно на структуру супрануклеосомной организации хроматина. Мы показали, что определенные конформации гистонового хвоста способствуют выпрямлению ДНК близи участков ее входа/выхода из нуклеосомы, что приводит к образованию дефектов кручения внутри ДНК.

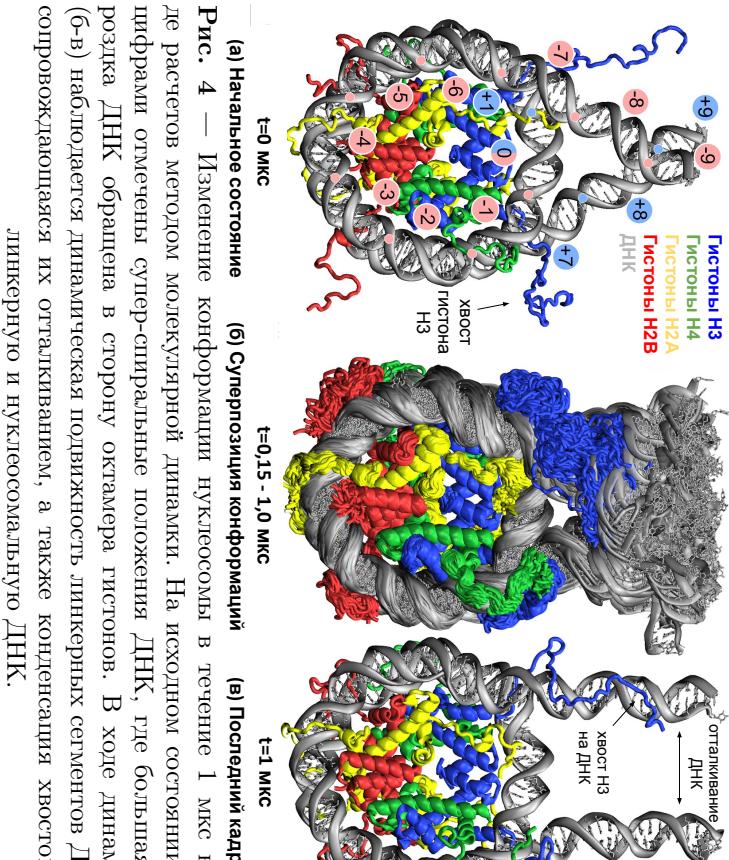


Рис. 4 — Изменение конформации нуклеосомы в течение 1 мкс в ходе расчетов методом молекулярной динамики. На исходном состоянии (а) цифрами отмечены супер-спиральные положения ДНК, где большая бороздка ДНК обращена в сторону октамера гистонов. В ходе динамики (б-в) наблюдается динамическая подвижность линкерных сегментов ДНК, сопровождающаяся их отталкиванием, а также конденсация хвостов на линкерную и нуклеосомальную ДНК.

Показано, что если в начальной структуре гистоновые хвосты экспонированы в растворитель, то в ходе динамики их конденсация на ДНК происходит в течение нескольких десятков наносекунд. Далее начинается

фибрillы, позволяющие реконструировать укладку пептидов и морфологию фибрill на основе соединения экспериментальных данных ИК- и КД-спектроскопии, рентгеновской дифракции, электронной и атомно-силовой микроскопии.

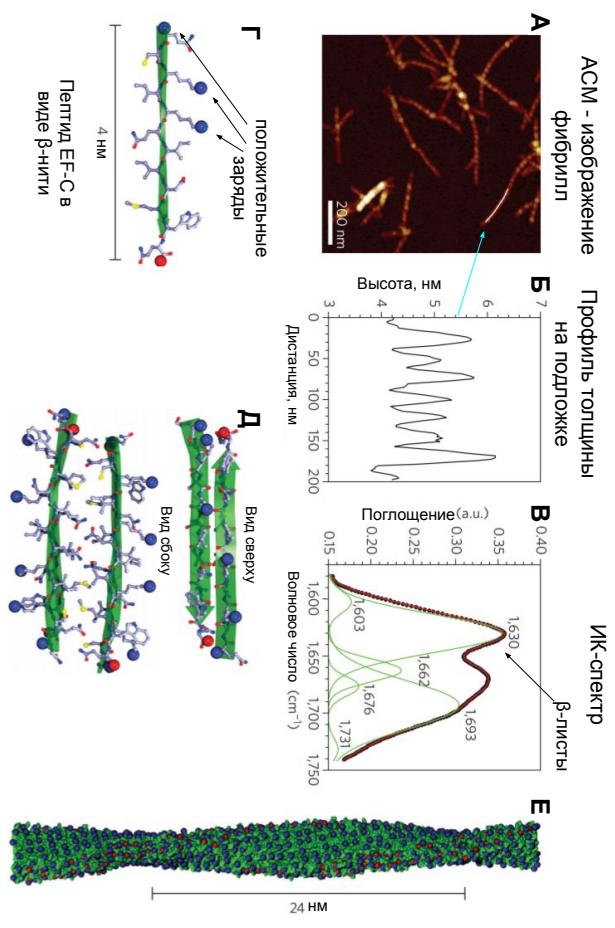


Рис. 11 — Структурная характеристика и молекулярное моделирование фибрill EF-C. А - АСМ-изображение фибрill EF-C. Б, профиль по указанной линии в А. В, ИКФС-спектр, Г - Молекулярная модель пептида EF-C. Д, вид сверху и сбоку элементарной единицы фибрillы, состоящей из четырех β -нитей, собранных в столкнувшись из двух антипараллельных β -листов. Е, Уточненная молекулярная модель фибрillы с шагом спирали 28 нм. Atомы C, серый; N, синий; O, красный; S, желтый.

В заключении приведено обсуждение результатов работы, рекомендации по их использованию и выводы, которые сформулированы ниже.

Обсуждение результатов

Построение и анализ структурных моделей биомакромолекулярных комплексов имеет ключевое значение в изучении структурной организации биологических систем. В реализации этого подхода принципиальную роль играют методы компьютерного молекулярного моделирования. В данной работе рассмотрены возможности методов молекулярного моделирования в построении и анализе больших ДНК-белковых комплексов и амилодподобных фибрill. В результате были разработаны новые интегративные

диффузия хвостов вдоль поверхности ДНК. При этом хвосты могут принимать конформационно ограниченные позиции из-за вставки лизинов и аргининов в малые бороздки ДНК. Ряд аминокислот в гистоновых хвостах, являющиеся сайтами важных эпигенетических пост-трансляционных модификаций (см. рисунок 2б) и сайтами связывания белков хроматина (например, Н3K9, Н3K27, Н4K16). На рисунке 4в видно, что многие из этих сайтов находятся в тесном контакте с ДНК, что ограничивает их доступность для пост-трансляционной модификации и взаимодействий с белками хроматина. В то же время, такие взаимодействия способствуют возникновению кооперативных эффектов при связывании нескольких белков с соседними сайтами на гистоновых хвостах. Полученные результаты находятся в согласии с рядом экспериментов по химическому сплаванию гистонов с ДНК, анализу влияния хвостов на взаимодействия белков хроматина с нуклеосомами. Рисунок 4 суммирует наблюдавшуюся конформационную подвижность нуклеосом.

Результаты моделирования коровых частиц нуклеосом методом молекулярной динамики приведены в разделе 2.3. Разработанные нами оригинальные алгоритмы моделирования позволили получить динамические траектории нуклеосомного корда с полноразмерными гистоновыми хвостами при физиологической температуре и ионной силе на временах более 10 микросекунд. С помощью специально разработанных алгоритмов анализа траекторий (проекция координат в системе отсчета нуклеосом, расчет относительного скручивания ДНК, анализ стабильных контактов и т. д.) мы выявили новые режимы динамики и пластичности нуклеосом. Мы показали возможность реконфигурации и разворачивания концов ДНК в нуклеосомах, опосредованных гистоновыми хвостами Н3 и H2A. Выявленные режимы отворачивания ДНК представлены на рисунке 5а. В ходе моделирования были выявлены ключевые взаимодействия ДНК-белок с гистонами, которые являются барьерами для откручивания ДНК от октамера гистонов (рисунок 5б). В ходе моделирования также была показана возможность локального спонтанного изменения степени закрученностии ДНК (образование/релаксация так называемых дефектов кручения) и сдвиг ориентационного положения для части коровой ДНК. Данные наблюдения указывают на возможность транслокации ДНК вдоль нуклеосомы путем “червячного” передвижения дефектов кручения вдоль ДНК (inchworm mechanism). При этом ДНК подобно винту прокручивается вдоль поверхности октамера гистонов. В ходе моделирования было показано наличие динамической пластичности ядра гистонов, способствующей подвижности ДНК. В результате сравнительного моделирования коровых частиц нуклеосом с хвостами и без хвостов гистонов было выявлено, что гистоновые хвосты замедляют как время отворота, так и время приворота ДНК, а также при отвороте ДНК поддерживают ее волнистый исходной траектории.

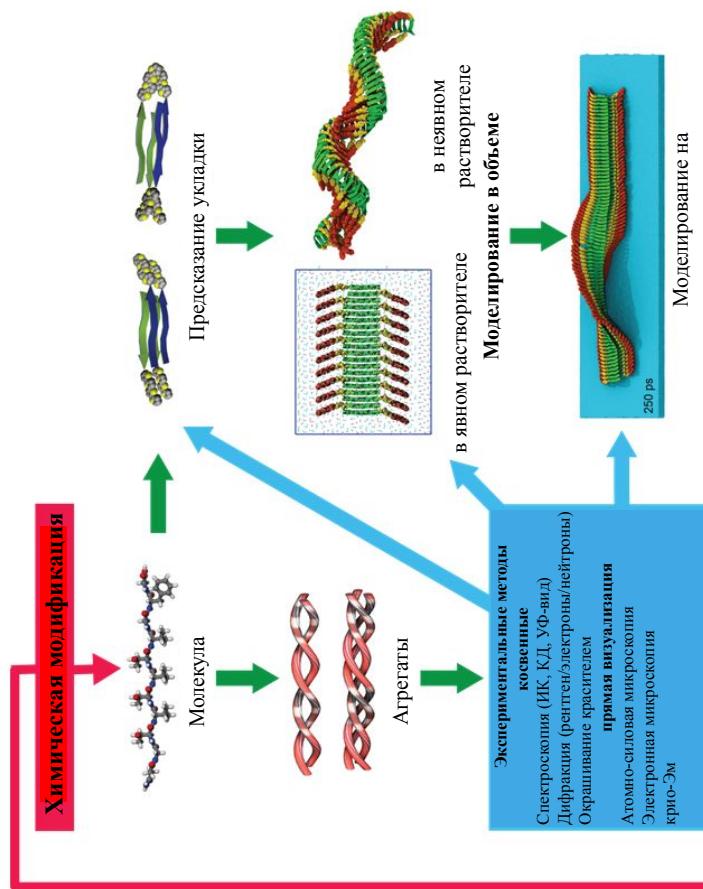


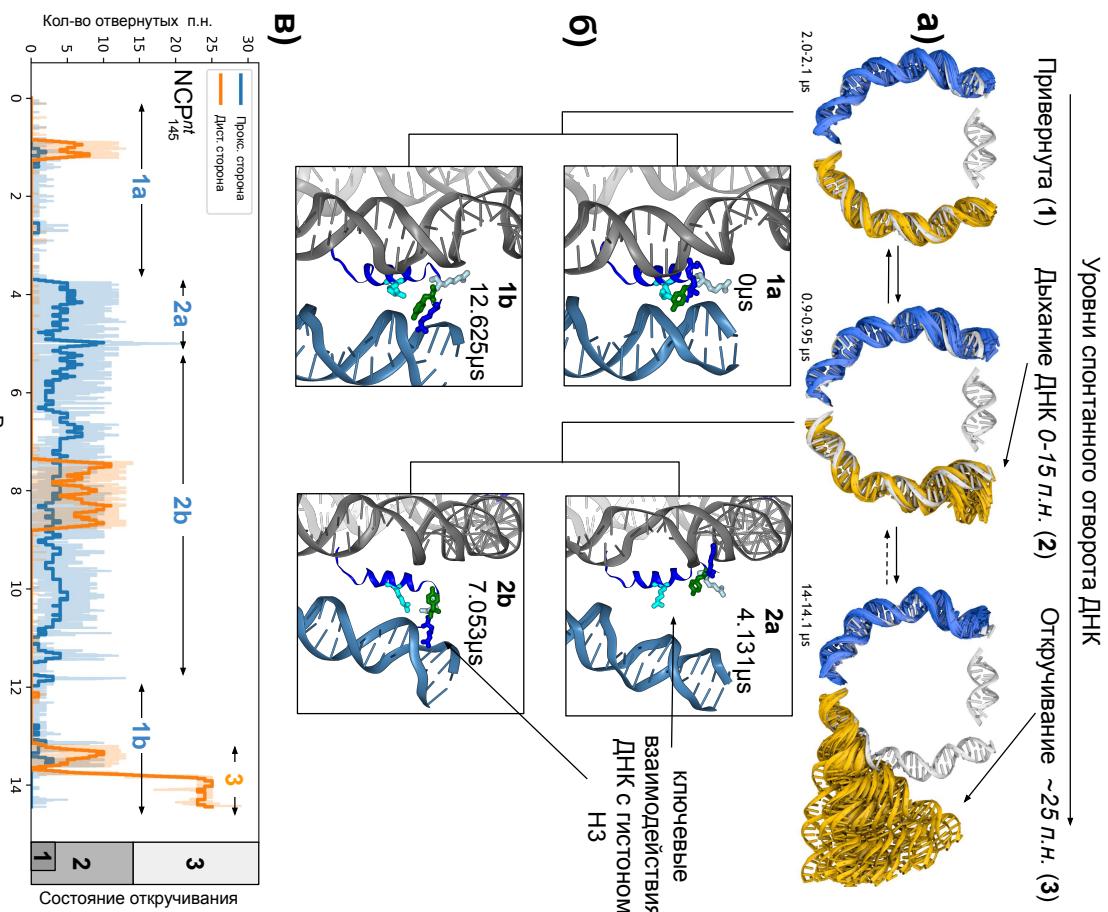
Рис. 10 — Схематическое изображение комбинированной экспериментальной/вычислительной методологии для изучения амилоидоподобных фибрилл.

На основе данных атомно-силовой микроскопии, кругового дихроизма, инфракрасной спектроскопии, порошковой дифракции рентгеновских лучей с применением разработанного нами интегративного подхода была построена молекулярная модель фибрилл EF-C (рис. 11г – е) [31]. Установлено, что боковые цепи лизина образуют гидрофильную поверхность с высокой плотностью катионных зарядов при физиологическом pH (рис. 11е), таким образом усиливая проникновение генетического материала вируса в клетку.

Материалы данной главы основаны на следующих ключевых статьях [31; *34; *35].

Выходы главы 5

Разработаны подходы интегративного моделирования амилоидоподобных



комплексов, образующихся при самосборке олигопептидов. Задача заключается в определении формы комплекса и его внутренней организации. В этом подходе экспериментальные данные используются уже на этапе конструирования возможных межмолекулярных укладок пептидов и на этапе оценки крупномасштабной морфологии фибрill.

В начале главы дается подробный обзор литературы о свойствах и характеристиках амилоидных и амилоидоподобных фибрилл (в том числе конъюгированных с синтетическими молекулами), типах известной или предполагаемой их структурной организации.

Раздел 5.2 посвящен работам по интегративному моделированию структуры амилоидоподобных фибрилл из либок олигомеров тетрапефен-(Thi-Va)₂-пептид. Была разработана методология, позволяющая на основе ограниченного набора данных (ИК-, КД-спектроскопия, рентгеновская дифракция, данные атомно-силовой микроскопии, просвечивающей электронной микроскопии и т.д.) реконструировать наиболее вероятные атомистические модели и укладки молекул в фибриллах. Схема разработанной стратегии изображена на рисунке 10. Известно, что в основе амилоидных фибрилл лежит укладка пептидов в β -листы. Данные спектроскопии могут помочь отличить наличие параллельных анти-параллельных β -листов. Данные дифракции оценить периодичность укладки β -нитей в листе и β -листов относительно друг друга. Варианты укладки можно определить на основе расчета энталпии образования при моделировании укладок методом МД в растворителе и использовать для построения начальных моделей фибрилл. С использованием алгоритмов диссипативной динамики частиц можно определить крупномасштабную морфологию фибрилл (включая период кручения и персистентную длину) и сопоставить с данными микроскопии. На основе данного подхода было проведено комплексное исследование различных вариантов молекулярной организации фибрилл на основе олигомеров тиоферн-пептид, определены наиболее вероятные модели, наблюдавшихся в эксперименте фибрилл, в том числе на графите. Практической мотивацией создания такого рода фибрилл, являются проводящие свойства тиофиена, что может приводить к формированию проводящих фибрилл “нанопроводов”.

Рис. 5 — Динамика отворота ДНК в нуклеосоме с укороченными гистоновыми хвостами. (а) Наблюдаемые динамические состояния **1-3**, между которыми система переходит в ходе динамики. (б) Для состояний **1** и **2** наблюдаются полостояния, связанные с реориентацией белкового сегмента в районе НЗ оН спирали (область НЗ-“застежки”). (в) Количество пар нуклеотидов отвернутых от октамера гистонов как функция времени моделирования. Критерий отворота ДНК — смещение центра пары оснований

полимеразы назад [[*23](#)]. Модель данного комплекса была также построена нами и представлена на рисунке [9Б](#).

Раздел 4.3 посвящен построению моделей разворачивания нуклеосом при их взаимодействии с комплексом гистоновых шаперонов FACT по данным экспериментов мономолекулярного FRET. В данном случае по данным FRET были оценены расстояния между положениями ДНК, в которых были введены флюоресцентные метки. К опубличенной физической модели ДНК (представленной в динуклеотидном приближении) добавлялись эмпирический гармонический потенциал, призванный свести расстояние между метками к наблюдаемому. В результате оптимальные энергии были построены модели, в которых показаны возможные моды разворачивания ДНК и гистонов, удовлетворяющие экспериментальным данным [[*19](#)].

В этой главе также изложены результаты анализа связывания центромерных нуклеосом некарнских дрожжей с белком внутренней кинетохоры Mif2 по данным гидроксильного футпринтинга. В главе 3 рассмотрена модель центромерной нуклеосомы и точно определено положение ДНК последовательности в ней. При анализе комплекса на профиля доступности ДНК к расщеплению оказывало влияние взаимодействие ДНК не только с гистонами, но и с белком Mif2. Путем сравнения экспериментальных данных, обработанных с помощью разработанной программы HYDROID, было показано, что димер белка Mif2 взаимодействует с одной из сторон нуклеосомы в районе 5-35 пар нуклеотидов от центра нуклеосомы [[*6](#)]. В главе также приводятся результаты по анализу комплексов нуклеосом с гистоном H1. Обсуждается результаты моделирования конформации линкерной ДНК при связывании гистона H1 по данным FRET, а также энергетические предпочтения изгиба линкерной ДНК при различных модах связывания гистона H1 [[*2; *5; *11; *19; *23](#)].

Материалы главы основаны на следующих статьях [[*2; *5; *6; *11; *19; *23](#)].

Выходы главы 4

Автором продемонстрировано, что методы интегративного моделирования способны решать важный класс задач при изучении супрануклеосомной структуры хроматина, исследования которой затруднительны классическими методами структурной биологии, а именно создавать модели комплексов нуклеосом с белками хроматина на основе использования экспериментальных данных электронной микроскопии низкого разрешения, данных FRET-микроскопии, данных футпринтинга ДНК. Получены структурные и динамические модели различных комплексов, возникающих при взаимодействии нуклеосом с РНК-полимеразами, гистоном H1, комплексом FACT.

В **пятой главе** (*Интегративное моделирование амилоподобных фибринилов*) рассмотрены подходы к определению структуры фибриллярных

Показано, что область гистона H3, находящаяся между двумя супервитками ДНК (см. среднюю часть Рис. [5](#)), играет важную роль в стабилизации концов ДНК, а также влияет на стабильность ДНК вблизи диады нуклеосомы.

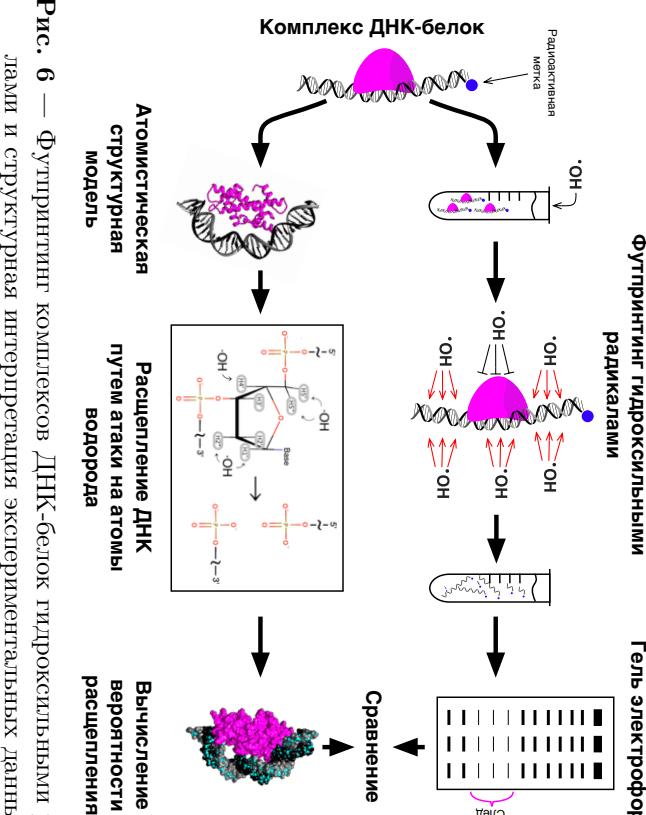
Результаты, приведенные в данной главе, основаны на следующих ключевых статьях [[*1—*3; *8; *9; *12; *14; *17; *18; *20; *21; *25](#)].

Выходы главы 2

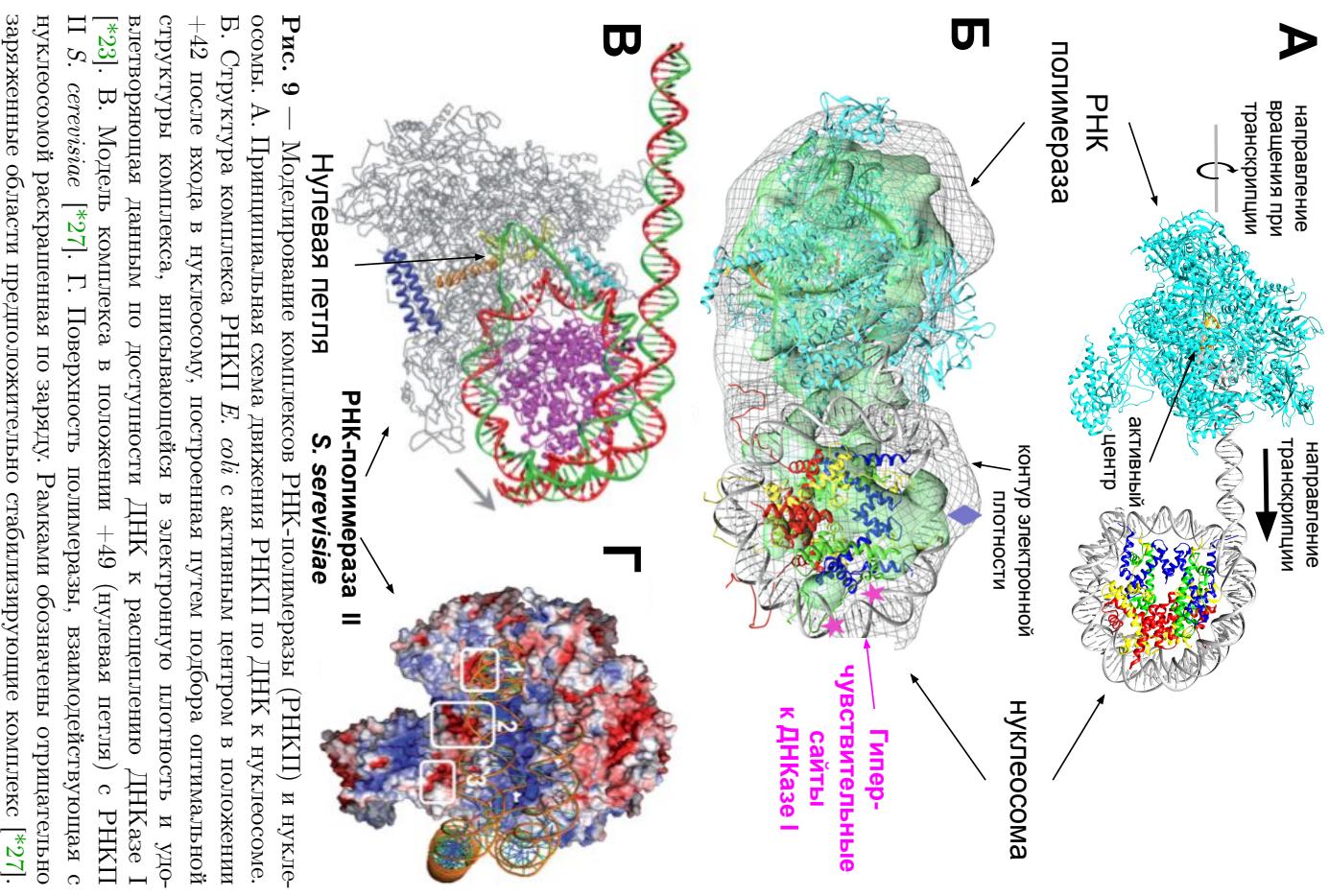
- Электростатическое отталкивание двух сегментов линкерной ДНК является одним из факторов, которые влияют на конформацию ДНК и углы входа-выхода ДНК из ядра нуклеосомы.
- В процессе спонтанного ступенчатого откручивания ДНК в нуклеосоме последовательно теряются стабильные взаимодействия ДНК с гистонами. Конформация ДНК подтверждена значительным флукутирующим в наносекундном диапазоне. В переходе нуклеосом из кристаллоподобного состояния в состояние с отвернутой ДНК ключевую роль играют взаимодействия остатков гистона H3 между двумя супервитками ДНК.
- Формирование дефектов кручения ДНК в нуклеосоме и конформационные перестройки внутри глобулярных частей гистонового октамера, происходящие на микросекундных временах, указывают на возможность скольжения нуклеосом вдоль ДНК по винтовому механизму и транслокации ДНК вдоль октамера.

Процессы динамики нуклеосом, изученные нами методом МД, ограничены по времени в силу текущих вычислительных возможностей. Крупномасштабные процессы, например, скольжения нуклеосом вдоль ДНК, сборка и разборка нуклеосом, полная диссоциации/реассоциации ДНК пока находятся за пределами моделируемых временных интервалов. Примером задачи не решаемой методами МД является выбор нуклеосомой наиболее энергетически выгодного положения на большом участке ДНК. Для решения такой задачи могут применяться интегративные подходы, изложенные в **третьей главе** (*Интегративное моделирование с применением данных экспериментов по футпринтингу ДНК*). В главе развиты подходы по моделированию ДНК-белковых комплексов, основанные на использовании экспериментальных данных по расщеплению ДНК (футпринтингу) гидроксильными радикалами. Гидроксильные радикалы обладают высокой реакционной способностью, небольшим размером, не являются заряженными. При взаимодействии с ДНК они атакуют атомы водорода дезоксирибозы, что приводит к расщеплению ДНК с потерей атакованного нуклеотида. При наличии белка на ДНК, белок стерически экранирует ДНК от расщепления. Это свойство

можно использовать для анализа структуры ДНК-белковых комплексов. Общая схема данного подхода представлена на рисунке 6.



Комплексы ДНК-белок создаются с использованием радиоактивно меченной ДНК и подвергаются расщеплению гидроксильными радикалами преимущественно на участках, не занятых белком. Репертуар расщепленных фрагментов ДНК анализируется с помощью денатурирующего гель-электрофореза. По интенсивности полос в геле можно судить о вероятности расщепления ДНК в том или ином месте последовательности. Атомистические структурные модели могут использоваться для прогнозирования профилей футпринтинга, поскольку частота расщепления зависит от доступности атомов водорода дезоксирибозы. Сравнение предсказанных и экспериментальных результатов может быть использовано для уточнения или создания структурных моделей. Несмотря на кажущуюся простоту данной концепции, ее комплексное решение представляет собой нетривиальную задачу, требует подбора алгоритмов анализа данных, их оптимизацию и программную реализацию. Нами такие алгоритмы были реализованы в виде программного комплекса HYDROID [**4] (<https://github.com/pcbi/HYDROID>). Ключевой задачей решаемой пакетом HYDROID является количественная оценка частоты расщепления



электронная микроскопия, FRET-микроскопия). Главу предваряет введение с описанием современных представлений о структуре хроматина и некоторых подходах к его моделированию.

Раздел 4.2 посвящен моделированию элонгационных комплексов РНК-полимеразы при транскрипции через нуклеосому на основе данных электронной микроскопии низкого разрешения и данных фурье-принтинга ДНК. Продвижение РНК-полимеразы вдоль ДНК является сложным биохимическим процессом (Рис. 9А), моделирование которого на данный момент лежит за пределами возможностей молекуллярной динамики и квантовой химии, поэтому является хорошей задачей для применения интегративных методов. Согласно экспериментальным данным при транскрипции РНК-полимеразой II через нуклеосому наблюдается ряд необычных фактов [8]. Нуклеосомы зачастую “выживают” при транскрипции, то есть сохраняют свое положение на ДНК, что очень важно для сохранения разметки генома в виде эпигенетических пост-трансляционных модификаций гистонов. Предполагается, что выживание нуклеосом становится возможным благодаря образованию интегральной “нулевой” петли в районе положения +49 после входа полимеразы в нуклеосому. В этом интермедиате ДНК за полимеразой восстанавливает контакты с октамером гистонов, таким образом, поддерживает его стабильность. Структуры полимераз и нуклеосомы известны по отдельности, однако каким образом образуется их комплекс при прохождении полимеразы через различные положения на нуклеосоме оставалось неизвестным. С использованием представления геометрии ДНК в виде параметров взаимного расположения пар оснований нами был разработан подход по генерации ансамбля комплексов с различным положением полимеразы, а также возможность варьировать уровень отворота ДНК от нуклеосомы. Данный подход, в том числе позволил связать ориентационное и трансляционное положение полимеразы и нуклеосомы в комплексе. Далее проводился анализ моделей на предмет оптимальной корреляции пространственного распределения атомов с экспериментальной плотностью и отсутствия стерических перекрываний атомов в комплексе. В результате анализа выбиралась оптимальная модель.

Модель комплекса нулевой петли, построенная нами в совместных экспериментально-теоретических работах, представлена на рисунке 9В. Детальный анализ данных моделей позволил выдвинуть гипотезу о том, что ряд отрицательно заряженных областей РНК-полимеразы II способствует стабильности данного комплекса (рис. 9Г). Данные предположения подтверждаются сравнением моделей и экспериментальных данных для ряда бактериальных РНК-полимераз [27]. Поскольку комплекс нулевой петли требует времени для образования, то пауза, возникающая при транскрипции, может приводить к откату (backtracking) активного центра

ДНК в каждом положении на основе анализа изображений гель электрофореза, а также теоретическая количественная оценка этого параметра на основе атомистических структурных моделей.

Процесс деконволюции изображения дорожек геля в частоты расщепления ДНК представлен на рисунке 7 и предполагает аппроксимацию профиля дорожки в виде суммы кривых Гаусса или Лоренца, описывающих вклад каждой полосы на геле. Математическая процедура данной аппроксимации была реализована в виде многопараметрической оптимизации аналитической функции (являющейся суммой вкладов отдельных пиков) по методу наименьших квадратов на основе алгоритма Левенberга — Марквардта. Данная процедура не всегда приводит к физически обоснованному решению и может быть неустойчивая.

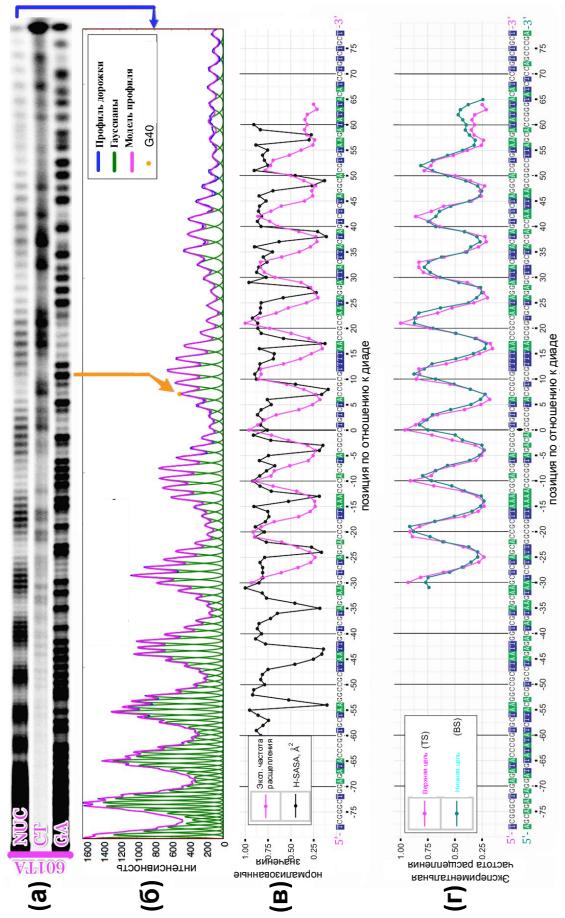


Рис. 7 — Количественная оценка экспериментов по гидроксильному фурнитингу и сравнение с теоретическими профилиями, полученными на основе атомной структуры известной нуклесомы (601TA-NUC). (А) Изображение полиакриламидного геля с продуктами реакции для верхней цепи ДНК. (Б) Профиль интенсивности дорожек геля и его деконволюция в индивидуальные интенсивности полос путем подбора функций Гаусса. (В) суперпозиция количественных экспериментальных частот расщепления ДНК и профилей H-SASA, рассчитанных на основе атомистической структуры. (Г) Суперпозиция экспериментальных частот расщепления ДНК для верхней и нижней цепей ДНК.

Поэтому были разработаны дополнительные методы регуляризации результатов, накладывающие ограничения на положения и ширину отдельных Гауссианов /Лоренцианов.

Для предсказания вероятности расщепления ДНК по атомистическим структурам были разработаны методы расчета поверхности атомов водорода дезоксирибозы доступной растворителю (H-SASA). Данный параметр чувствителен к небольшим вариациям в геометрии ДНК, параметрам выбираемым для длин атом-атомных связей и размерам атомов. Были проведены наборы расчетов методом молекулярной динамики для выяснения влияния динамических эффектов на профили H-SASA и вычисления наиболее точных профилей H-SASA. В результате были предложены пропедуры (в том числе способы добавления атомов водорода к тяжелотонким структурам) и наборы параметров для оценки профилей H-SASA по кристаллическим структурам.

На основании разработанных подходов были изучены данные по расщеплению ДНК в нуклеосомах. В результате моделирования было предсказано, что профили расщепления ДНК слабо зависят от последовательности ДНК и аналогичны в различных нуклеосомах. Было показано, что наличие в нуклеосоме оси псевдосимметрии второго порядка, находит свое отражение в идентичности профилей расщепления ДНК для верхней и нижней цепей ДНК (рис. 7Г). Таким образом, для каждого нуклеотида верхней цепи ДНК существует нуклеотид нижней цепи с аналогичной вероятностью расщепления. Эти нуклеотиды находятся на одинаковом расстоянии от диады нуклеосомы, но по разные ее стороны. На основании данного наблюдения был сформулирован и апробирован метод определения положения ДНК на нуклеосоме на основе подбора положения диады, которое обеспечивает выполнение вышеуказанного правила. Данный подход позволяет с точностью до одного нуклеотида определить позицию последовательности ДНК на нуклеосомном коре. С помощью данного подхода было определено ранее неизвестное положение ДНК в *in vitro* модели центромерной нуклеосомы третьей хромосомы *S. cerevisiae* (CEN3-NUC) и построена ее структурная модель (рис. 8). Обсуждается значение данной модели для понимания функционирования центромерных нуклеосом пекарских дрожжей, нуклеосомная ДНК которых содержит функционально важные последовательности аденинов (A-тракты).

Материалы данной главы основаны на следующих статьях [**4; **7; *11].

Выходы главы 3

Данные гидроксильного футпринтинга ДНК могут быть эффективно использованы для построения структурных моделей ДНК-белковых комплексов. Для ДНК-белковых комплексов, обладающих осью симметрии второго порядка, сравнение профилей расщепления ДНК для двух цепей

позволяет дополнительно повысить точность определения положения ДНК относительно белка.

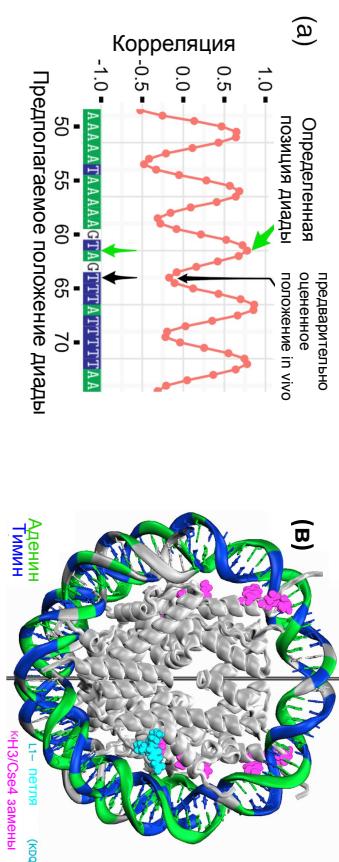


Рис. 8 — Применение алгоритма идентификации диады к нуклеосоме CEN3-NUC с неизвестным положением диады. (А) коэффициент корреляции Пирсона между экспериментальными профилями расщепления верхней (TS) и нижней (BS) цепей ДНК как функция предполагаемого положения диады; зеленая стрелка – определенное положение, черная стрелка ранее предполагавшееся по *in vivo* данным [7]. (Б) Наложение профилей TS и BS. (В) Полученная модель нуклеосомы (диада установлена в положение 61). Нити ДНК окраинены в соответствии с их последовательностью, выделяя АТ-участки: А – зеленый, Т – синий, Г или С – серый.