ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ УЧРЕЖДЕНИЕ «ФЕДЕРАЛЬНЫЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР «КРИСТАЛЛОГРАФИЯ И ФОТОНИКА» РОССИЙСКОЙ АКАДЕМИИ НАУК»

ИНСТИТУТ КРИСТАЛЛОГРАФИИ ИМ. А.В. ШУБНИКОВА РАН

На правах рукописи

КОНАРЕВ ПЕТР ВАЛЕРЬЕВИЧ

РАЗВИТИЕ И ПРИМЕНЕНИЕ МЕТОДОВ АНАЛИЗА ДАННЫХ МАЛОУГЛОВОГО РЕНТГЕНОВСКОГО РАССЕЯНИЯ МНОГОКОМПОНЕНТНЫМИ БИОЛОГИЧЕСКИМИ СИСТЕМАМИ

Диссертация

на соискание ученой степени доктора физико-математических наук по специальности 1.3.8 - физика конденсированного состояния

Научный консультант: Доктор химических наук Волков Владимир Владимирович

ОГЛАВЛЕНИЕ

Вве	дение	5
Гла	ва 1. Малоугловое рентгеновское и нейтронное рассеяние	
(M:	УРР/МУРН) как эффективный наноразмерный структурный	
мет	од исследования	16
1.1.	Теоретические основы метода малоуглового рассеяния	17
1.2.	Эксперимент: синхротронные/нейтронные станции и лабораторные	
p	ентгеновские установки	23
1.3.	Оценка структурных параметров	29
1.4.	Методы ab initio восстановления формы макромолекул	39
1.5.	Методы молекулярной тектоники для моделирования комплексов	
б	иомакромолекул	51
1.6.	Восстановление неизвестных частей структуры и учет гибкости	
M	иакромолекул	56
1.7.	Оценка эффективности ab initio методов восстановления формы	
M	акромолекулярных комплексов с неоднородной электронной	
П	лотностью по данным МУРР/МУРН	70
Гла	ва 2. Апостериорная оценка оптимального углового диапазона	
эксг	периментальных данных МУРР/МУРН с использованием теоремы	
Кот	ельникова-Шеннона	95
2.1	Автоматизированный алгоритм оценки оптимального углового диапазона	
данн	ных МУРР/МУРН	97
2.2 (Оценка работы алгоритма на теоретических и экспериментальных данных	
МУ	РР/МУРН	110
Гла	ва 3. Количественный анализ состава многокомпонентных	
полі	идисперсных систем по данным МУРР/МУРН	119
3.1 (Оценка числа компонентов и нахождение объемных долей компонентов	
белк	ковой смеси с известными атомными структурами	120

3.2 Приближение многокомпонентной системы простыми геометрическими	
телами с учетом полидисперсности и межчастичного взаимодействия	123
3.3 Применение разработанных методов для анализа олигомерных смесей и	
многокомпонентных самоорганизующихся систем	130
3.3.1 Смесь активной и неактивной форм фермента	130
3.3.2 Олигомерные смеси, находящиеся в динамическом равновесии	140
3.3.3 Самоорганизующиеся многокомпонентные системы	156
Глава 4. Количественный анализ состава липидных и липид-белковых	
смесей по данным МУРР	176
4.1 Анализ полидисперсных систем липидных везикул: восстановление	
электронной плотности липидного бислоя, распределения везикул по	
размерам и мультислойной ламеллярной организации везикул	177
4.2 Применение разработанных методов для анализа эволюции липидных вез	икул
в процессе экструзии	184
4.3 Применение разработанных методов для анализа взаимодействия липидни	ЫХ
везикул с белками	194
Глава 5. Прямое определение формы промежуточного компонента в	
эволюционирующих многокомпонентных системах по данным МУРР	205
5.1 Метод ab initio определения формы промежуточного компонента в	
эволюционирующей системе	207
5.2 Проверка эффективности работы алгоритма на теоретических	
данных МУРР	210
5.3 Применение метода для ряда трехкомпонентных	
эволюционирующих систем и двухкомпонентных олигомерных смесей	216
5.3.1 Процесс образования фибрилл белка инсулина	216
5.3.2. Множественные состояния само-сборки фермента люмазинсинтазы	219
5.3.3. Концентрационно-зависимая олигомеризация белка фактора роста	
нервов	222
5.3.4. Общие замечания по работе алгоритм	225

1 лава 6. Метод эволюционного факторного анализа для восстановлени	RI
индивидуальных профилей рассеяния компонентов белковыми смесям	ии по
данным МУРР с использованием хроматографической колонки	230
6.1 Метод эволюционного факторного анализа (ЭФА) для анализа	
набора данных МУРР в сочетании с гель-хроматографией	232
6.2 Возможности и ограничения метода ЭФА на примерах	
теоретических данных МУРР	238
6.3 Применение метода ЭФА для экспериментальных	
наборов данных МУРР с использованием гель-фильтрации	
и ионно-обменной хроматографии	253
6.4 Автоматический и интерактивный режимы работы метода ЭФА	260
Глава 7. Разработка программного обеспечения для интерактивного	
моделирования малоугловых данных с помощью	
графических приложений	265
7.1 Первичная обработка экспериментальных данных в	
интерактивном режиме	266
7.2 Интерактивное моделирование четвертичной структуры	
белковых комплексов	273
7.3 Суперпозиция атомных моделей с использованием разложения на	
сферические гармоники	282
7.4 Анализ структурного состава полидисперсных систем	296
7.5 Оценка структурных параметров частично упорядоченных систем	302
Выводы	306
Приложение – Программное обеспечение для анализа	
данных малоуглового рассеяния	309
Литература	
Список работ автора	313
Список цитируемой литературы	347

ВВЕДЕНИЕ

Актуальность

В настоящее время большую роль в практических разработках функциональных материалов и лекарственных препаратов играют структурные методы исследования вещества, среди которых особо следует отметить рентгеновскую дифракцию и малоугловое рентгеновское рассеяние (МУРР). Метод МУРР позволяет изучать неупорядоченные образцы различной природы с размерами частиц от 1 до 200 нм и более. Это и биологические системы (растворы белков и их комплексов, пептиды и нуклеиновые кислоты), полимеры, композитные наноматериалы, коллоидные растворы, микроэмульсии, минералы и металлические наночастицы. Благодаря доступности источников рентгеновского излучения высокой яркости и развитию методов анализа и интерпретации данных последние десятилетия, метод МУРР стал стандартным инструментом исследования структур для биологов и биохимиков. Размер макромолекул, которые можно анализировать, варьируется от небольших пептидов или фрагментов нуклеиновых кислот до больших комплексов с молекулярным весом более мегадальтона или даже целых вирусов. Важным преимуществом метода МУРР ПО сравнению с другими структурными методами является его неразрушающая природа и относительная простота подготовки образцов. В большинстве случаев необходимо просто отфильтровать предварительно очищенный раствор белка для удаления крупных частиц (агрегатов) и загрузить его в измерительный капилляр или ячейку. Таким образом, биологические образцы можно изучать при физиологических условиях, или близких к естественным, а также изучать изменения структуры в зависимости от физико-химических условий среды (рН раствора, температура, концентрация образца, состав буферного раствора, добавление соли или лиганда и т.д.). Поэтому, исследования биологических макромолекул с помощью метода МУРР представляют особую важность для диагностики и лечения болезней в медицине и фармацевтике, а также

создания новых функциональных биоматериалов, используемых в промышленности.

Однако, тот факт, что подготовка образца для измерений методом МУРР относительно проста по сравнению с другими структурными методами (такими как рентгеновская кристаллография, ядерный магнитный резонанс или криоэлектронная микроскопия), не означает отсутствие каких-либо критериев, которым должен соответствовать образец. Одним из требований является монодисперсность образца, когда в растворе находятся только одинаковые макромолекулы. Для таких систем разработаны надежные ab initio методы определения пространственной формы однородных частиц с разрешением 1-2 нм, основывающиеся исключительно на использовании экспериментальных данных МУРР. В то же время, для макромолекулярных комплексов с неоднородной электронной плотностью (например, комплексов белок/нуклеиновая кислота) требуется дополнительная информация, которую можно получить либо путем проведения измерений для неполных (парциальных) суб-комплексов, либо с помощью метода вариации контраста с использованием малоуглового рассеяния нейтронов (МУРН). В настоящей работе будет уделено особое внимание возможностям и ограничениям итерационных алгоритмов восстановления электронной плотности (формы частиц) для неоднородных макромолекулярных комплексов.

Проведение измерений МУРР в растворе ограничивает пространственное разрешение моделей, которые можно восстановить *ab initio*, т.е. без какой-либо дополнительной информации. Случайная ориентация частиц в растворе приводит к изотропному рассеянию, это позволяет проводить азимутальное усреднение регистрируемого изображения на двумерном детекторе, но одновременно приводит к потере существенной части информации. Поэтому восстановление 3D-модели по одномерной кривой малоуглового рассеяния в общем случае является неоднозначной процедурой. В этой связи особую важность приобретает оценка оптимального углового диапазона экспериментальных данных МУРР, в настоящей

работе был разработан и предложен новый автоматизированный алгоритм определения оптимального углового диапазона регистрации с использованием теоремы Котельникова-Шеннона для разбавленных монодисперсных систем, позволяющий добиться более надежной оценки интегральных структурных параметров частиц (таких как гидратированный объем частицы и молекулярная масса), что, в свою очередь, улучшает эффективность работы алгоритмов восстановления 3D формы частицы.

До недавнего времени анализ полидисперсных многокомпонентных систем по данным МУРР был существенным образом ограничен и позволял оценивать лишь общие структурные характеристики, усредненные по ансамблю частиц. Однако, последние достижения в разработке программ анализа данных, в том числе предложенные и реализованные автором и приведенные в настоящей работе, позволили гораздо более осмысленно исследовать сложные полидисперсные системы, представляющие собой белковые или липидные многокомпонентные смеси. При наличии известных кривых рассеяния от каждого макромолекулярного компонента, можно определять их объемные доли в общей смеси. Если структуру компонентов можно аппроксимировать простыми геометрическими телами, то для ряда систем, таких как микроэмульсии или растворы металлических наночастиц, помимо определения объемных долей компонентов появилась возможность функции распределения частиц по размерам оценивать характер межчастичного взаимодействия в случае концентрированных систем. Для липидных многокомпонентных систем ПОМИМО определения функции распределения везикул по размерам, проводят восстановление электронной липидного плотности бислоя внутри везикулы оценку параметров мультислойной ламеллярной организации везикул. Все последние достижения в области интерпретации данных рассеяния были реализованы автором в виде ряда компьютерных программ, в том числе были предложены и новые подходы. Созданное автором программное обеспечение предоставляет исследователям уникальную возможность одновременного структурного

моделирования большого числа компонентов в смесях, которая была недоступна в ранее предложенных известных схемах. Разработанные методы были апробированы на большом ряде олигомерных смесей и самоорганизующихся систем.

Следует отметить, что прямой анализ формы частиц для многокомпонентных смесей сильно затруднен и практически невозможен, так как кривые рассеяния всегда представляют собой линейные комбинации интенсивностей от различных компонентов смеси. Однако, благодаря разработанному автором новому подходу формы промежуточных прямого определения эволюционирующих системах, в ряде случаев это стало возможным. Системы, для которых известны начальное и конечное состояния, но которые также содержат неизвестное промежуточное состояние, теперь могут быть непосредственно Несколько наборов проанализированы c помощью ЭТОГО метода. экспериментальных данных МУРР эволюционирующей смеси, записанные во времени, анализируются совместно с целью получения наилучшего общего соответствия всем имеющимся данным МУРР. Изучение процесса фибрилляции подхода, позволило определить инсулина помощью данного промежуточного состояния инсулина, а также объемные доли его мономеров, промежуточных продукта и зрелых фибрилл в каждый момент времени. Фактически, в данном случае можно говорить о восстановлении формы компонента в эволюционирующей системе, который напрямую измерить в эксперименте не представляется возможным.

Для успешного анализа данных МУРР агрегированные макромолекулы должны быть удалены из раствора, чтобы избежать их влияния на кривую рассеяния, так как метод МУРР очень чувствителен к размеру частиц: увеличение радиуса частицы в 10 раз увеличивает интенсивность сигнала в нулевом угле сразу в 106 раз. Если раньше общей практикой было отфильтровывать или осаждать агрегаты путем центрифугирования, то в настоящее время стало популярным проводить измерения МУРР в сочетании с использованием гель-хроматографической ячейки

(SEC-SAXS, Size-Exclusion Chromatography-Small-Angle X-ray Scattering). В такой установке раствор образца, выходящий из хроматографической колонки (элюент) сразу направляется в камеру образца, так что сигнал МУРР собирается непосредственно после разделения компонентов смеси за счет различного времени их прохождения через хроматографическую колонку. Эта установка позволяет разделять отдельные компоненты в образце и регистрировать данные МУРР от изолированных фракций, что чрезвычайно важно для последующей интерпретации данных, анализа и структурного моделирования. Однако, в случае частично перекрывающихся пиков выхода (элюирования) смесей компонентов, задача разложения этих данных на отдельные фракции нетривиальна и потенциально неоднозначна. В данной работе предложен новый алгоритм для восстановления профилей рассеяния и концентрации компонентов на основе эволюционного факторного анализа (ЭФА) по данным МУРР, исследованы его возможности и Эффективность ограничения. алгоритма продемонстрирована смоделированных и экспериментальных наборов данных SEC-SAXS. В частности, с помощью данного метода удалось описать необычную олигомерную смесь белка пируват альдолазы, состоящую из гексамеров и октамеров, которую невозможно было бы промоделировать стандартными способами.

Хотя в ряде случаев удается создать полностью автоматизированные алгоритмы, требующие вмешательства пользователя при анализе данных МУРР, оптимальные решения не всегда могут быть найдены в автоматическом режиме, поэтому разработка программного обеспечения для интерактивной обработки данных МУРР с помощью графических приложений остается чрезвычайно актуальной задачей. Автором диссертационной работы были разработаны и реализованы в виде программ графические приложения, позволяющие проводить полноценное интерактивное моделирование данных МУРР, начиная с первичной обработки и заканчивая моделированием четвертичной структуры белковых комплексов, проведением структурного анализа многокомпонентных смесей и оценкой структурных параметров для частично

упорядоченных систем. Разработанные программы составляют значительную часть свободно распространяемого программного пакета ATSAS (https://www.embl-hamburg.de/biosaxs/software.html), который широко используется в научном сообществе при анализе данных МУРР от биологических систем.

Цели и задачи работы

Целью настоящей работы было развитие методов структурного анализа многокомпонентных биологических систем по данным рентгеновского и нейтронного малоуглового рассеяния. Важной частью работы была разработка алгоритма оценки углового диапазона экспериментальных данных малоуглового рассеяния, содержащего полезную информацию, и применение предложенных методов к ряду практических задач исследований состава и структуры многокомпонентных систем биологической и неорганической природы. Кроме того, разрабатываемые методы должны были учитывать, как наличие полидисперсности частиц по форме и размерам, так и неоднородность электронной плотности исследуемых частиц.

Научная новизна

Развитые методы позволяют анализировать состав и структуру многокомпонентных систем как в условиях отсутствия априорных сведений об изучаемой системе, так и с привлечением дополнительной информации. Компьютерное моделирование позволило определить возможности и ограничения использования разработанных методов интерпретации данных малоуглового рассеяния сложными биологическими системами.

Разработанные автором методы анализа данных малоуглового рассеяния позволили провести детальное структурное исследование ряда биологических многокомпонентных смесей, в частности определить форму неизвестного промежуточного компонента в процессе образования фибрилл, оценить распределение липидных везикул по размерам и их мультислоевую ламмеллярную организацию, восстановить профили рассеяния отдельных компонентов белковых

смесей, которые не удается полностью разделить с помощью хроматографических или других методов. Методы применены для количественной оценки состава ряда биологических систем по данным малоуглового рассеяния, причем в некоторых случаях другими способами получить ценную структурную информацию не представляется возможным. Новизна полученных результатов подтверждается высокими индексами цитирования работ с участием автора.

Научное и практическое значение работы

Научное значение работы определяется практическое актуальностью поставленных задач и возможностью практического использования разработанных методов данных малоуглового рассеяния многокомпонентными анализа биологическими системами. Разработанные методы существенно расширяют область применимости метода МУРР/МУРН. Эффективность методов продемонстрирована на примере решения широкого круга задач структурного анализа сложных полидисперсных и неоднородных объектов.

Среди основных применений развитых методов, следует отметить следующие:

- 1) структурный анализ белковых и липидных смесей по данным малоуглового рассеяния;
- 2) восстановление профилей рассеяния отдельных компонентов в неразделяемых белковых смесях по набору данных малоуглового рассеяния с использованием гель-хроматографической колонки;
- 3) структурная характеризация равновесных состояний и протекания динамических процессов в эволюционирующих системах по данным малоуглового рассеяния.

Предложенные в работе методы анализа белковых и липидных смесей и алгоритмы определения формы биомакромолекул реализованы в ряде компьютерных программ, вошедших в пакет ATSAS, который использовался на момент написания работы более чем в 2000 лабораторий в мире. Данный пакет находится в свободном доступе для академических пользователей на сайте URL: https://www.embl-hamburg.de/biosaxs/software.html.

Степень достоверности результатов

Достоверность представленных в диссертации результатов анализа малоуглового рассеяния многокомпонентных биологических OT систем согласуется с данными, полученными, в частности, методом электронной микроскопии. Восстановленные формы частиц биомакромолекул сопоставлялись со структурами, найденными другими авторами методами рентгеноструктурного анализа или ядерного магнитного резонанса, если такая информация имелась в наличии. С целью подтверждения надежности работы созданного программного обеспечения автором было проведено масштабное моделирование возможных объектов исследований.

Основные положения, выносимые на защиту

- 1. Метод количественной оценки углового диапазона экспериментальных данных малоуглового рассеяния, содержащего полезную информацию, на основе теоремы Котельникова-Шеннона, в случае монодисперсных разбавленных растворов.
- 2. Методы поиска объемных долей компонентов и распределений частиц по размерам в приближении простых геометрических тел или с использованием атомных моделей по данным малоуглового рассеяния в многокомпонентных смесях.
- 3. Методы поиска профилей электронной плотности липидного бислоя, параметров мультислоевой организации и распределений липидных везикул по размерам по данным малоуглового рассеяния в многокомпонентных липидных и липидбелковых смесях.
- 4. Метод прямого восстановления формы частицы неизвестного промежуточного состояния по данным малоуглового рассеяния в динамической трехкомпонентной системе, включая оценку объемных долей компонентов в процессе ее эволюции.
- 5. Эволюционный факторный анализ наборов данных малоуглового рентгеновского рассеяния, полученных использованием гельхроматографической колонки, ДЛЯ оценки количества компонентов И

восстановления индивидуальных профилей рассеяния от компонентов в белковых смесях.

6. Многофункциональная графическая платформа для анализа данных малоуглового рассеяния от полидисперсных и частично упорядоченных систем.

Личный вклад автора

Вклад автора состоит в разработке алгоритмов, методов и программ для анализа экспериментальных данных малоуглового рассеяния, приведенных в данной работе. Ряд программ, созданных автором, включен в широкоиспользуемый в научном сообществе пакет программного обеспечения ATSAS, предназначенный для анализа данных малоуглового рассеяния растворами биомакромолекул и наночастиц. Автор принимал активное участие в проведении экспериментов по малоугловому рентгеновскому рассеянию на синхротронных и лабораторных станциях, проводил обработку результатов экспериментов и интерпретацию полученных данных.

Апробация работы

Результаты работы доложены и обсуждены на II, III, IV, и V Национальных конференциях по применению Рентгеновского, Синхротронного излучений, Нейтронов и Электронов для исследования материалов — РСНЭ (Москва, 1999, 2001, 2003, 2005), VII Национальной конференции Рентгеновское, Синхротронное излучения, Нейтроны и Электроны для исследования наносистем и материалов. Нано-Био-Инфо-Когнитивные технологии - РСНЭ-НБИК (Москва, 2009), XI, XII, XIII, XIV, XVI, XVII и XVIII Международных конференциях по малоугловому рассеянию (Нью-Йорк, США, 1999, Венеция, Италия, 2002; Киото, Япония, 2006; Оксфорд, Англия, 2009; Берлин, Германия 2015, Траверс-Сити, США, 2018, Кампинас, Бразилия, 2022), 9-й Международной конференции по дифракции волокон и некристаллической дифракции (Шеффилд, Англия, 2000), III Международной конференции «Химия высокоорганизованных соединений и научные принципы нанотехнологии» (Санкт-Петербург, 2001), Европейской конференции "Тенденции в энзимологии" (Комо, Италия, 2006), Сателлитной

конференции Международного союза кристаллографов по Рентгеновским и нейтронным методам в исследовании наноструктур (Хиого, Япония, 2008), Международном Химическом конгрессе сообщества тихоокеанского бассейна (Гонолулу, США, 2010), XX, XXI и XXIV Конгрессах Международного союза кристаллографов (Флоренция, Италия, 2005, Осака, Япония, 2008, Хайдарабад, Индия, 2017), IX, XI и XIII Международных конференциях по Биологии и синхротронному излучению (Манчестер, Англия, 2007, Гамбург, Германия, 2013, Шанхай, Китай, 2019), Первом Российском кристаллографическом конгрессе (Москва, Россия, 2016), II, III, IV, V, VI, VII, VIII и IX Международной конференции "Лазерные и плазменные исследования и технологии" (ЛаПлаз) (Москва, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023), VII, VIII, IX и X Российских Национальных Симпозиумах "Белки и пептиды" (Новосибирск, 2015, Москва, 2017, Дагомыс, 2019, Сочи, 2021), 51-й, 52-й, 53-й и 54-й Зимних Школах по физике конденсированного состояния (ФКС) (Санкт-Петербург, 2017, 2018, 2019, 2020), Российской конференции по использованию рассеяния нейтронов в конденсированных средах РНИКС-2018 (Санкт-Петербург, 2018), VI конференции по малоугловому рассеянию и рефлектометрии (Гатчина, 2023), XIII и XIV Международных конференциях "Проблемы сольватации и комплексообразования в растворах" (Суздаль, 2018, Иваново 2021), VIII и IX Международных научных семинарах и VI и VII международных молодежных научных школах-семинарах "Современные методы анализа дифракционных данных и актуальные проблемы" рентгеновской оптики (Великий Новгород, 2016, Москва, Санкт-Петербург, 2020), Hamburg P12 Virtual User Meeting (Гамбург, Германия, 2021), **EMBL** Международной конференции "Synchrotron and Free electron laser Radiation: generation and application" ("SFR-2022"), (Новосибирск, 2022).

Публикации по теме диссертации

По теме диссертации опубликованы 117 статей в отечественных и международных рецензируемых журналах, рекомендуемых ВАК РФ, и тезисы 92 докладов на конференциях.

Структура и объем работы

Диссертация состоит из введения, 7 глав, основных выводов, приложения с указанием программного обеспечения, разработанного автором для анализа данных малоуглового рассеяния, списка цитируемой литературы и списка публикаций по теме диссертации. Работа изложена на 371 странице, включая 99 рисунков и 23 таблицы. Список цитируемой литературы содержит 240 библиографических ссылок. Список работ автора содержит 117 публикаций, обозначенные в тексте диссертации буквой А.

Автор с огромной благодарностью посвящает эту работу своим научным наставникам доктору физико-математических наук Д. И. Свергуну и доктору химических наук В. В. Волкову.

Автор также выражает глубокую признательность д.ф.-м.н., профессору В.Е. Асадчикову, д.х.н. Э.В.Штыковой, д.ф.-м.н. Ю.В. Писаревскому, к.ф-м.н. В.Р.Самыгиной, к.ф.-м.н. М.В.Петухову, к.ф.-м.н. Л.А.Дадиновой, д.ф.-м.н. О.В. Батищеву, к.ф.-м.н. Л.В.Кордюковой, к.ф.-м.н. М.А. Марченковой, к.ф.-м.н. А.Е. Крюковой, к.ф.-м.н. К.Б. Ильиной, Г.С. Петерсу за сотрудничество и плодотворные дискуссии, а также всем своим коллегам из Института кристаллографии и особенно лаборатории рефлектометрии и малоуглового рассеяния за помощь, сотрудничество и конструктивное, теплое участие на разных стадиях работы.

Глава 1

Малоугловое рентгеновское и нейтронное рассеяние (МУРР/МУРН) как эффективный структурный метод исследования наноразмерных систем

Метод малоуглового рентгеновского рассеяния (МУРР) является универсальным методом исследования структурных характеристик наноразмерных неоднородностей частично упорядоченных или разупорядоченных систем [1-4]. Метод МУРР имеет долгую историю развития, начиная с первой половины XX века после открытия рентгеновских лучей. Однако, если рентгеновские лучи использовались уже с 1912 года для проведения экспериментов по дифракции кристаллов (Фридрих, Книппинг, Лауэ), то должно было пройти несколько десятков лет, прежде чем стало ясно, что они также могут быть полезны и для экспериментов по малоугловому рассеянию от растворов частиц. До 1938 года рентгеновские камеры не были пригодны для регистрации областей вблизи первичного пучка (в области малых углов) из-за паразитного рассеяния. Французский физик Андре Гинье, который сконструировал первую рентгеновскую камеру с низким рассеивающим фоном и в дальнейшем использовал ее для изучения неоднородных образцов (сплавов металлов), содержащих зерна с размером частиц от 10 до 100 нм [5], положил тем самым начало методу МУРР в области структурного анализа. Эти исследования привели к открытию взаимосвязи между формой кривой интенсивности рассеяния рентгеновских лучей и размером частиц. Вскоре ученые поняли, что в дополнение к информации о размере и форме частицы частиц, профили рассеяния также содержат информацию о внутренней структуре.

В 1960-70-х годах рентгеновская кристаллография биологических макромолекул способствовала росту интереса и к другим структурным методам. Малоугловое рентгеновское рассеяние постепенно становилось популярным методом для определения размера и формы наноразмерных образований в образцах, которые было трудно закристаллизовать. Одним из основных препятствий для расширения экспериментов МУРР была сложность анализа данных. В большинстве случаев

анализ данных МУРР в то время сводился к определению радиуса инерции и объема белковой молекулы. Значительные успехи в структурных работах с использованием малоуглового рассеяния произошли только в 1970-х годах, когда стали доступны источники синхротронного рентгеновского излучения и появились возможности для проведения малоугловых экспериментов с использованием нейтронов (МУРН) методом вариации контраста за счет изменения состава растворителя (H_20/D_20) [6, 7]. Исследователи объединили информацию, полученную из экспериментов МУРР и МУРН, чтобы определить, как молекулы ДНК сворачиваются вокруг ядра нуклеосомы [8] за 20 лет до того, как была получена кристаллическая структура высокого разрешения [9]. Эти же методы были использованы для картирования положения белков и субъединиц РНК в рибосоме. Полная карта прокариотической 30S субъединицы [10] и частичная карта 50S субъединицы [11] были составлены на 10 лет раньше, чем была получена первая кристаллическая структура этой субъединицы [12]. Тем не менее, анализ данных был настолько сложной процедурой, что чаще всего для того, чтобы сделать значимые выводы из полученных результатов, требовалось привлечение специалистов из нескольких областей науки.

Более мощные компьютеры и прогресс в развитии программного обеспечения на рубеже 2000х годов сделали восстановление трехмерных форм частиц по данным малоуглового рассеяния рутинным и относительно простым процессом [13]. В настоящее время биологическое малоугловое рассеяние является быстро развивающейся областью исследований. Простота и гибкость проведения экспериментов МУРР/МУРН, в сочетании с последними достижениями в области интерпретации результатов, с каждым годом привлекает все больше исследователей и расширяет возможности этих методов.

1.1 Теоретические основы метода малоуглового рассеяния

Хотя физические механизмы упругого малоуглового рассеяния рентгеновских лучей и нейтронов (МУРР/МУРН) веществом принципиально различны, они могут

быть описаны одним и тем же математическим формализмом. Поэтому теоретические основы рассеяния обоими методами будут представлены единым образом, но с указанием на имеющиеся различия между двумя типами излучения. Явление рентгеновского или нейтронного рассеяния возникает вследствие взаимодействия падающего пучка фотонов с электронами атомов образца или нейтронов с атомными ядрами, соответственно. При облучении образца монохроматическим пучком рентгеновских фотонов с длиной излучения λ или нейтронов с длиной волны де Бройля λ , т.е. с волновым вектором $k_0=|k_0|=2\pi/\lambda$, фотоны взаимодействуют с электронами (нейтроны - с ядрами) внутри объекта, и последние становятся источниками рассеянных фотонов/нейтронов. В данной работе мы будем рассматривать только эффекты упругого рассеяния, когда энергия рассеянного излучения остается равной энергии падающего. При таком упругом рассеянии длина волны λ рассеянных фотонов/нейтронов равна длине волны падающего пучка, а модуль волнового вектора рассеянной волны $k_1 = |k_1|$ равен k_0 . Амплитуда волны, рассеянной каждым атомом/ядром, описывается его длиной рассеяния, f (название отражает тот факт, что эта характеристика имеет единицы длины). Поскольку рентгеновские лучи взаимодействуют с электронами атомов образца, то атомная длина рассеяния пропорциональна количеству имеющихся в нем электронов. Взаимодействие нейтронов с ядерным потенциалом описывается длиной ядерного рассеяния, которая не имеет систематической зависимости от атомного номера, а скорее чувствительна к изотопному содержанию. Особого внимания заслуживает случай водорода и дейтерия. Если для рентгеновских лучей эти два изотопа неразличимы (оба имеют по одному электрону), то при рассеянии нейтронов между атомами Н и D наблюдается большая разница: протоны рассеивают преимущественно некогерентно. Это различие дает основание для использования селективного дейтерирования и измерений в смесях вода/тяжелая вода для анализа внутренней структуры сложных частиц. Рассмотренные выше представляют собой соображения упрощенную картину взаимодействия рентгеновских лучей и нейтронов с веществом.

Рассмотрение процессов рассеяния (подобно дифракционным данным) включает в себя преобразование из прямого пространства (координаты r объекта) в обратное пространство, т.е. координаты векторов рассеяния $s=(s,\Omega)=k_I-k_\theta$ (Рис. 1.1). Здесь переданный момент импульса (или модуль вектора рассеяния) определяется как $s=4\pi \sin(\theta)/\lambda$, где 2θ - угол рассеяния, а ось телесного угла Ω - определяет направление вектора рассеяния. Заметим, что для обозначения вектора рассеяния используется как термин "s" (от слова scattering), так и термины "q" или "Q" (например, в нейтронном научном сообществе), но во всех случаях они имеют одно и то же определение. С математической точки зрения это представляет собой преобразование Фурье

$$A(s) = \Im[\rho(r)] = \int \rho(r) \exp(isr) dr$$
(1.1)

где $\rho(\mathbf{r})$ - распределение рассеивающей плотности, A(s) - амплитуда рассеяния, а интегрирование производится по облученному объему образца.

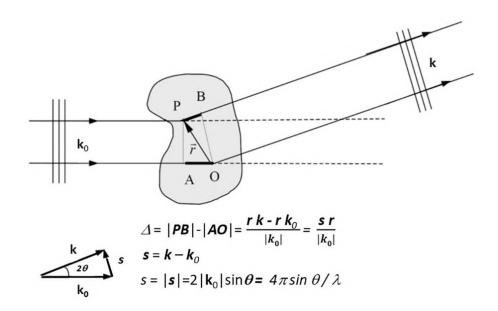


Рис. 1.1 Расчет разности хода лучей Δ от рассеивателей в точках О и Р исследуемой частицы в случае упругого рассеяния. На рисунке k и k_{θ} - волновые векторы в направлениях рассеянного и падающего пучков, соответственно, точка Р смещена от О на вектор r. 2θ обозначает угол рассеяния, а s - вектор рассеяния.

Примем во внимание свойства преобразования Фурье (1.1), которое зависит от произведения \mathbf{sr} , т.е. чем больше размер объекта \mathbf{r} в прямом пространстве, тем

короче будет соответствующий вектор рассеяния *s* в обратном пространстве. Разрешение данных МУРР/МУРН в прямом пространстве можно формально оценить как $d=2\pi/s$, что является аналогом хорошо известного уравнения Брэгга в кристаллографии. Таким образом, рассеяние в области малых углов, т.е. при малых значениях *s*, дает информацию о больших расстояниях (намного больше длины волны), т.е. об общей структуре частицы, с низким (обычно 2-3 нм) разрешением. Для описания рассеяния от совокупности атомов распределение плотности длины рассеяния $\rho(\mathbf{r})$ может быть представлено как общая длина рассеяния электронов/ядер на единицу объема. Если рассмотреть рассеяние от системы, содержащей отдельные частицы, и предположить, что пространство (матрица), в которой содержится образец (например, растворитель в белковых растворах) имеет постоянную плотность рассеяния ρ_s , то разница между амплитудой рассеяния от отдельной частицы и амплитудой рассеяния матрицей (растворителем) в том же объеме определяется преобразованием Фурье от избыточной плотности длины рассеяния (или контраста) $\Delta \rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_s$

$$A(s) = \Im[\Delta \rho(r)] = \int_{V} \Delta \rho(r) \exp(isr) dr$$
 (1.2)

где интегрирование производится по объему частицы V, который представляет собой когерентно рассеивающий объем (положения атомов коррелированы между собой). Экспериментально измеряется интенсивность рассеяния, т.е. число фотонов/нейтронов, рассеянных в данном направлении s, которое пропорционально квадрату амплитуды $I(s) = [A(s) A(s)^*]$.

Если рассматривать ансамбль одинаковых частиц, то полное рассеяние будет зависеть от пространственного распределения этих частиц, и можно рассмотреть два основных предельных случая. В случае идеального монокристалла все частицы в образце имеют определенные коррелированные ориентации и расположены регулярным образом в пространстве, так что должны суммироваться амплитуды рассеяния от отдельных частиц с учетом направления векторов рассеяния. В результате общая интенсивность рассеяния перераспределяется вдоль определенных направлений, определяемых обратной решеткой, и измеренная

дискретная трехмерная функция $I(s_{hkl})$ соответствует распределению плотности в одной элементарной ячейке кристалла. Если же частицы распределены случайным образом, а их положения и ориентации некоррелированы, то суммируются уже **интенсивности** рассеяния, так как фазы рассеянных от разных частиц волн носят случайный характер. Соответственно, интенсивность от всего ансамбля такой системы является непрерывной изотропной функцией, пропорциональной рассеянию от одной частицы, усредненному по всем ориентациям в обратном пространстве $I(s) = \langle I(s) \rangle_{\Omega}$. В этом случае интенсивность рассеяния преобразуется к виду

$$I(s) = \int_{r=0}^{\infty} \langle \Delta \rho(\mathbf{r}) \otimes \Delta \rho(\mathbf{r}) \rangle_{\omega} \frac{\sin(sr)}{sr} d\mathbf{r}$$
(1.3)

где \otimes - означает операцию свертки, ω - телесный угол в прямом пространстве. Разбавленные растворы монодисперсных невзаимодействующих биологических макромолекул часто соответствуют этому второму предельному случаю. Основной структурной задачей в этом случае является восстановление структуры частицы (т.е. ее контраста $\Delta \rho(r)$) на низком разрешении по данным рассеяния. Интенсивность рассеяния от таких систем пропорциональна квадрату контраста $(\Delta \rho)^2$, где $\Delta \rho$ - средняя по объему плотность избыточной длины рассеяния $\Delta \rho(r)$. Плотности рассеяния рентгеновского и нейтронного излучения и точки совпадения контрастов биологических макромолекул (matching points) в растворах H_2O/D_2O приведены в Таблице 1.1.

Таблица 1.1 Плотности рентгеновского и нейтронного рассеяния биологических макромолекул [14].

	Рентгеновские лучи		Нейтроны		
Компонент	ρ , 10^{10}	Точка	ho в H ₂ O,	ρв D ₂ O,	Точка
	cm ⁻²	совпаден	10^{10}cm^{-2}	10^{10}cm^{-2}	совпадения
		ия			контраста
		контраста			% D ₂ O
H ₂ O	9.42	-	-0.6	-	-
D ₂ O	9.42	-	6.4	-	-
Липиды	8.46	-	0.3	-6.0	≈ 15%
Белки	11.8	65%	1.8	3.1	pprox 40%
		сахарозы			
D-белки	11.8	65%	6.6	8.0	-
		сахарозы			
Нуклеиновые	15.5	-	3.7	4.8	≈ 70%
кислоты					
D- Нуклеиновые	15.5	-	6.6	7.7	-
кислоты					

Для рентгеновских лучей плотность длины рассеяния часто выражается в терминах электронной плотности (количество электронов на $Å^3$, 1 электрон/ $Å^3 = 2.82 \times 10^{11}$ cm). В случае нейтронов увеличение плотности длины рассеяния одного и того же компонента в D_2O по сравнению с H_2O обусловлено H/D-замещением обмениваемых протонов.

На практике часто приходится иметь дело с неидеальными системами, когда частицы отличаются друг от друга по размеру и/или форме, а в случаях концентрированных растворов еще и межчастичными взаимодействиями, которыми нельзя пренебречь. В частности, полидисперсные системы, такие как мицеллы, микроэмульсии, блок-сополимеры или металлические наночастицы,

часто состоят из частиц с похожей формой, которые отличаются только размером. Такие системы удобно описывать функцией распределения по объему $D_V(R) = N(R)V(R)$, где N(R) - количество частиц с характерным размером R, а V(R) - объем частиц этого размера. Предположим, что контраст одинаков для всех частиц, тогда интенсивность рассеяния может быть представлена в следующем виде

$$I(s) = T(s)(\Delta \rho)^2 \int_{R_{min}}^{R_{max}} D_V(R)V(R)i_0(sR)dR$$
(1.4)

где $i_0(sR)$ - нормированная интенсивность рассеяния частицы ($i_0(0) = 1$) (формфактор), а R_{min} и R_{max} - минимальный и максимальный размеры частицы, соответственно, T(s) — структурный фактор, описывающий межчастичную интерференцию. Одновременное восстановление форм-фактора частицы и функции распределения частиц по размерам невозможно в силу неоднозначности решения такой задачи, поэтому в большинстве практических случаев при анализе такого рода полидисперсных систем форма частиц аппроксимируется простыми геометрическими телами (сферой, эллипсоидом, цилиндром) и проводится поиск функции распределения частиц по размерам.

1.2 Эксперимент: синхротронные/нейтронные станции и лабораторные рентгеновские установки

Общая схема малоугловых измерений сама по себе достаточно простая (Рис. 1.2), разработка лабораторного прибора НО рентгеновского ИЛИ синхротронной/нейтронной станции является сложной технической задачей, поскольку требует решения вопроса по уменьшению паразитного рассеяния вблизи Малоугловой первичного пучка. эксперимент обычно выполняется использованием геометрии пропускания, когда падающий монохроматический луч проходит через образец перпендикулярно поверхности образца. Зарегистрированный сигнал представляет собой суммарный сигнал от всех контрастно-взвешенных вкладов атомов в облучаемом объеме образца.

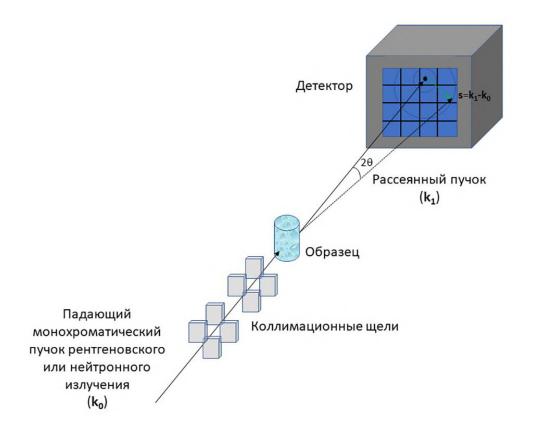


Рис. 1.2 Общая схема малоуглового эксперимента, где k_1 и k_0 — волновые вектора падающего и рассеянного пучка, 2θ - угол рассеяния, $s=k_1-k_0$ — вектор рассеяния (показан зеленым цветом), черный круг на детекторе вблизи первичного пучка представляет собой защитную заслонку детектора (beamstop).

Существует два основных типа лабораторных малоугловых приборов: дифрактометры с точечным источником и коллимационные камеры Кратки. В первом случае рентгеновское излучение подается на образец в виде почти параллельного (слабо расходящегося) пучка, геометрия которого формируется с помощью рентгеновской оптики, что обеспечивает минимальное влияние размера падающего пучка на форму кривой рассеяния. Во втором случае, рентгеновский пучок подается на образец в виде линии, которая может быть описана как набор бесконечных точечных источников, расположенных параллельно. В результате этого волны рассеяния, исходящие от любой точки вдоль этой линии, смешиваются ("замазываются") со слегка смещенными волнами рассеяния от соседних точек. Таким образом, "замазанный" профиль I(s) в конфигурации камеры Кратки определяется как составом образца, так и геометрией пучка, и для преобразования данных в идеальный (свободный от эффектов "замазывания") профиль *I(s)*

требуется введение коллимационных поправок [15-17]. Типичным примером камеры Кратки является лабораторный малоугловой дифрактометр АМУР-К (Институт кристаллографии РАН, Москва) с позиционно-чувствительным линейным газовым детектором [18] (Рис. 1.3), на котором был проведен ряд экспериментов, упоминающихся в данной диссертации.



Рис. 1.3 Лабораторный малоугловой дифрактометр АМУР-К (ИК РАН, Москва).

В качестве источников рентгеновского излучения в основном используются герметичные рентгеновские трубки с медными, молибденовыми или серебряными анодами (Си K_{α} , 1.54 Å; Мо K_{α} , 0.71 Å; Ag K_{α} , 0.59 Å) [19, 20]. Более высокий рентгеновский поток можно получить, используя либо источник рентгеновского излучения на жидком аноде [21, 22], где мишенью является поток расплавленного металла (например, галлия), либо трубку, с вращающимся анодов, где твердая мишень анода вращается для снижения локальной тепловой нагрузки. Последние модели лабораторных установок (среди которых можно выделить приборы, производимые компаниями Bruker, Rigaku, Anton Paar, Xenocs) позволили получить потоки фотонов до 10^9 фотонов в секунду на площади в несколько сотен квадратных микрометров, что позволяет измерять разбавленные растворы полимеров и биомакромолекул в течение нескольких минут и проводить исследования с временным разрешением в секундном диапазоне.

Синхротронные источники предоставляют уникальные возможности для проведения экспериментов МУРР, которые невозможно осуществить с помощью лабораторных приборов. Высокая яркость синхротронного источника в сочетании с настройкой размера пучка позволяет достичь потока фотонов на образце на несколько порядков больше (10¹²-10¹⁴ фотонов/с) по сравнению с лабораторными приборами. Это открывает новые возможности для проведения высокопоточного скрининга образцов на низких концентрациях (для одного измерения обычно требуется около 10-20 микролитров), использования микрофлюидных систем доставки образцов и исследований с высоким временным разрешением (до 2-3 миллисекунд).

Типичная синхротронная малоугловая станция может одновременно измерять угловой диапазон, на котором вектор рассеяния s меняется на два порядка (примерно от 0.05 до 5 нм⁻¹). Этого можно достичь либо путем параллельного использования двух детекторов, либо перестройкой длины волны излучения (обычно в диапазоне $4\div 20$ кэВ), поскольку модуль вектора s обратно пропорционален длине волны λ .

Значительное количество синхротронных станций, частично или полностью предназначенных для проведения малоугловых экспериментов на биологических образцах, были построены за последние два десятилетия, здесь приведен далеко неполный список таких станций: X33 и P12 (Doris-III и Petra-III, DESY, Hamburg, Germany) [A1-A2], [23], SAXS/WAXS (Australian Synchrotron Melbourne, Australia) [24], SAXS/D (SSRL, SLAC, USA), SIBYLS (ALS, Berkeley, USA) [25], ID02 и ВМ29 (ESRF, Grenoble, France) [26], SWING (Synchrotron SOLEIL Saint-Aubin, France) [27], cSAXS (SLS, Villigen, Switzerland), G1 (CHESS, Cornell University, USA), 4C PAL SAXS II (POSTECH, Pohang, South Korea) [28], 18-ID (Argonne National Laboratory, USA) [29], BL45XU (SPring-8, Japan) [30], B21 (Diamond, Oxford, UK) [31], TP25A (Taiwan Light Source (TLS), Taiwan), BL19U2 (SSRF, Shanghai, China) [32], "БиоМУР" (Курчатовский институт, Москва) [A3-A4] и др. Особо стоит отметить станцию X33 [A1-A2], которая была одной из первых малоугловых

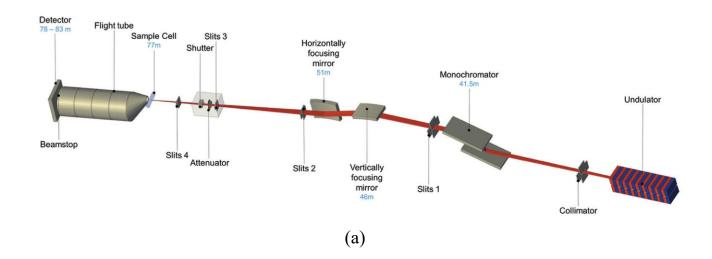
установок с внедренной роботизированной сменой образцов и автоматизированной системой сбора и анализа данных. После завершения работы синхротронного накопительного кольца DORIS-III, часть оборудования станции X33 была использована при строительстве станции "БиоМУР" в Курчатовском институте [A3-A4].

Схема установки малоугловой станции P12 синхротронного кольца Petra-III (Гамбург, Германия), на которой была получена значительная часть данных МУРР, рассматриваемых в данной диссертации, показана на Рис. 1.4 [23]. Оптическая часть станции включает в себя перестраиваемый двойной монохроматор и фокусирующее зеркало для подавления высших гармоник ($\lambda/2, \lambda/3, ...$). Паразитное рассеяние вблизи первичного пучка уменьшается за счет использования нескольких пар коллимационных щелей (наиболее эффективными из них являются нерассеивающие коллимационные щели, на краях которых крепятся монокристаллы кремния). Кроме того, вся линия станции (на длине от источника до детектора 83 м), включая "пролетную" трубу (с регулируемым расстоянием образец-детектор 1.6, 2.1, 3.1, 4.1 или 6.1 м), находится под вакуумом.

Синхротронные источники, обладая высокой плотностью потока квантов, имеют связанный с этим недостаток, заключающийся в высокой вероятности радиационного повреждения образца во время съемки данных рассеяния. Этот недостаток преодолевают с помощью покадровой съемки с небольшой длительностью одного кадра. Программа обработки данных, сравнивая формы последовательно получаемых кривых интенсивности рассеяния, определяет начало такого повреждения, и усредняет только временные кадры, снятые до этого момента.

Нейтронные источники бывают двух типов: ядерные реакторы, которые обеспечивают непрерывный поток нейтронов с различной кинетической энергией, и импульсные ускорительные источники, где используется столкновение пучка протонов с вольфрамовой или ртутной мишенью для генерации нейтронов. Однако, поскольку потоки интенсивности нейтронов на единицу площади пучка

значительно ниже, чем в рентгеновских методах, а размер пучка обычно больше, то все это требует более длительного времени экспозиции и большего объема/концентрации образца.



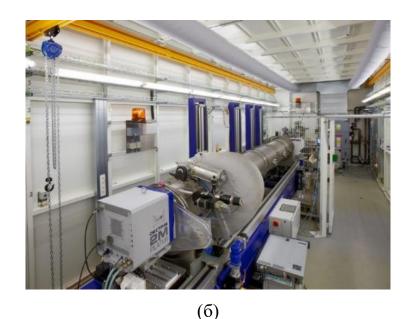


Рис. 1.4 (а) Схема основных элементов малоугловой синхротронной станции P12 (Гамбург, Германия). Расстояния основных компонентов от источника-ондулятора указаны синим цветом. (б) Вид конечной части станции P12 с перестраиваемой по длине вакуумной трубой и детектором Pilatus 2M [23].

Тем не менее, нейтронные малоугловые исследования обладают рядом дополнительных возможностей, в частности большой глубиной проникновения через исследуемые материалы, отсутствием механизмов повреждения материалов

за время измерений и возможностью варьирования контраста с помощью частичного дейтерирования образцов. В ряде случаев совместное использование МУРР/МУРН позволяет извлекать детальную информацию о структуре многокомпонентных неоднородных биологических систем.

1.3 Оценка основных структурных параметров

В данных МУРР/МУРН существенная часть информации теряется вследствие ориентационного усреднения. Тем не менее, и из таких данных можно извлечь достаточно ценной информации с помощью довольно простых методов, которые будут рассмотрены в данном разделе.

Как было показано в разделе 1.1, в случае изотропного рассеяния монодисперсной системой невзаимодействующих частиц интенсивность ее рассеяния определяется уравнением (1.3), где можно ввести следующие обозначения

$$I(s) = 4\pi \int_{r=0}^{\infty} r^2 \gamma(r) \frac{\sin(sr)}{sr} d\mathbf{r}$$
 (1.5)

$$\gamma(r) = \langle \Delta \rho(r) \otimes \Delta \rho(r) \rangle_{\omega} = \langle \int \Delta \rho(r) \Delta \rho(r+u) \, du \rangle_{\omega}$$
 (1.6)

Функция $\gamma(r)$ — это сферически усредненная автокорреляционная (характеристическая) функция избыточной плотности (контраста) рассеяния, которая очевидно, обращается в нуль для расстояний, превышающих максимальный диаметр частицы D_{max} . На практике, часто используется функция $p(r) = r^2 \gamma$ (r), соответствующая распределению количеств расстояний между элементами объема внутри частицы, взвешенными на соответствующие значения контрастов. Эта функция распределения расстояний p(r) вычисляется с помощью обратного преобразования Фурье

$$p(r) = \frac{1}{2\pi^2} \int_{s=0}^{\infty} sr \, I(s) \sin(sr) ds$$
 (1.7)

Таким образом, поведение интенсивности рассеяния при очень малых $(s \to 0)$ и очень больших $(s \to \infty)$ значениях вектора рассеяния s напрямую связано с общими геометрическими параметрами частиц.

Андре Гинье в своей классической работе в 1939 году [5] вывел соотношение для зависимости кривой рассеяния в области малых углов ($s \to 0$)

$$I(s) \approx I(0)e^{-\frac{s^2R_g^2}{3}}$$
 (1.8)

которое легко получить, воспользовавшись разложением в ряд Маклорена для функции $sin(sr)/sr = 1 - (sr)^2/3! + O(s^4)$ в уравнении (1.5). Интенсивность рассеяния в нулевой угол I(0) определяется как

$$I(0) = \int_{V} \int_{V} \Delta \rho(\mathbf{r}) \Delta \rho(\mathbf{r}') d\mathbf{r} d\mathbf{r}' = (\Delta \rho)^{2} V^{2}$$
(1.9)

а радиус инерции частицы R_g (среднеквадратичное расстояние от центра масс частицы) может быть рассчитан с помощью функции распределения по расстояниям p(r) как

$$R_g^2 = \frac{\int_0^{Dmax} r^2 p(r) dr}{2 \int_0^{Dmax} p(r) dr}$$
 (1.10)

Поскольку в последнем случае при оценке R_g используется вся кривая рассеяния, а не ограниченный угловой диапазон в области малых углов, полученные значения R_g будут более надежны и менее чувствительны к остаточным эффектам межчастичного взаимодействия или наличия низкого уровня агрегации частиц, чем значения, полученные на основе приближения Гинье. Различие оценок R_g в (1.8) и (1.10) может свидетельствовать либо об агрегации частиц, либо о межчастичной интерференции.

В результате мы получаем, что для идеальных монодисперсных систем график Гинье (ln(I(s)) в зависимости от s^2) должен аппроксимироваться линейной функцией, пересечение которой с вертикальной осью дает оценку для I(0), по тангенсу наклона кривой можно определить радиус инерции R_g . Линейность графика Гинье можно рассматривать как проверку монодисперсности образца, тогда как наличие отклонений от линейной зависимости может указывать на полидисперсность системы или наличие межчастичных взаимодействий (притягивающего или отталкивающего характера) (Рис. 1.5). При этом всегда

следует помнить, что приближение Гинье справедливо только для очень малых углов, а именно в диапазоне $s < 1.3/R_g$.

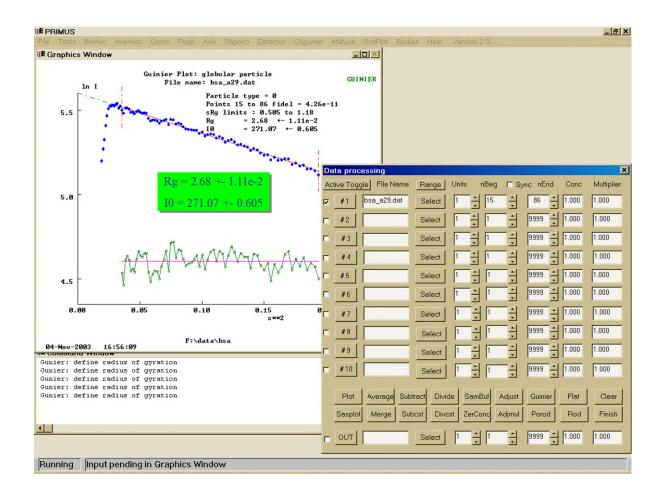


Рис. 1.5 График Гинье ln(I(s)) от s^2 , представленный на графическом интерфейсе программы Primus [A5] из пакета ATSAS [A6-A10]. Синим цветом показаны экспериментальные данные МУРР, красная кривая — наилучшее приближение Гинье на выбранном угловом диапазоне, ниже зеленым цветом показана кривая остатков между экспериментальными данными и приближением Гинье. Спад кривой в области самых малых углов обусловлен наличием защитной заслонки детектора.

Интенсивность рассеяния в нулевой угол I(0) пропорциональна квадрату молекулярной массы (MM), а также числу частиц в облученном объеме образца. Оценку молекулярной массы исследуемого белка можно проводить с

использованием стандартного белка (например, лизоцима, бычьего альбумина или изомеразы) согласно следующему соотношению

$$MM = \frac{I(0) * c_{\text{стандарт}} * MM_{\text{стандарт}}}{c * I(0)_{\text{стандарт}}}$$
(1.11)

где c и $c_{cman\partial apm}$ — концентрации исследуемого и стандартного белка, соответственно.

Приближение Гинье (уравнение (1.8)) справедливо для частиц произвольной формы. Отметим, что в пределе очень вытянутых и очень сплюснутых частиц справедливы модифицированные соотношения Гинье, позволяющие оценивать радиус инерции поперечного сечения R_c стержнеобразных частиц по графику зависимости sI(s) от s^2 , а радиус инерции толщины R_t дискообразных частиц вычисляется из графика зависимости $s^2I(s)$ от s^2 согласно

$$sI(s) \approx I_c(0)e^{-\frac{s^2R_c^2}{2}}, \qquad s^2I(s) \approx I_t(0)e^{-s^2R_t^2}$$
 (1.12)

В работе Порода [33] была проанализирована асимптотическая зависимость интенсивности в области больших углов ($s \to \infty$) для однородных частиц. После интегрирования уравнения (1.5) по частям, приняв во внимание, что $\gamma(r)=0$ npu $r \ge D_{max}$, можно получить следующее соотношение для интенсивности при $s \to \infty$

$$I(s) = -\frac{8\pi}{s^4} \gamma'(0) + \frac{Asin(sD_{max})}{s^3} + \frac{Bcos(sD_{max})}{s^4} + \frac{1}{s^4} \sum_i c_i \cos(sR_i) + O(s^{-5}) \quad (1.13)$$

где $A = 4\pi D \gamma' (D_{max})$, $B = 4\pi \left(2\gamma' (D_{max}) + D\gamma'' (D_{max})\right)$, c_i — постоянные коэффициенты, а R_i — точки разрыва функции $\left(r\gamma(r)\right)^{\prime\prime\prime}$ на интервале $[0,D_{max}]$. В уравнении (1.13) все члены за исключением первого слагаемого являются осциллирующими тригонометрическими функциями, поэтому интенсивность будет убывать на больших углах как s^{-4} , эта асимптотическая зависимость получила название закона Порода. Для однородных частиц также справедливо соотношение $\gamma'(0) = -(\Delta\rho)^2 S$, где S — площадь поверхности частицы. Если ввести величину Q (так называемый инвариант Порода)

$$Q = \int_{s=0}^{\infty} s^2 I(s) ds = 2\pi^2 \int_{V} (\Delta \rho)^2 dr$$
 (1.14)

Последняя часть равенства в (1.14) следствие теоремы Парсеваля для интегралов в прямом и обратном пространствах [34]. Если принять во внимание уравнение (1.9), то исключенный объем частицы (объем Порода) V_P и площадь поверхности S можно оценить с помощью инварианта Порода

$$V_P = 2\pi^2 \frac{I(0)}{Q}, \qquad \frac{S}{V_P} = \frac{\pi}{Q} \lim_{S \to \infty} (S^4 I(S))$$
 (1.15)

Для реальных макромолекул электронная плотность не является полностью однородной, что приводит к появлению дополнительной константы в асимптотике Порода $s^4I(s)\approx B's^4+A'$. Тем не менее, вычитание соответствующей константы B' из кривой I(s) в большинстве случаев дает разумное приближение рассеяния от соответствующего однородного тела. Рассчитанный объем частицы, V_P позволяет получить альтернативную оценку молекулярной массы MM, не зависящую от погрешностей в определении концентрации образца, и для глобулярных белков имеется эмпирическая оценка, что V_P (в нм³) в 1,5÷1.7 раза больше MM (выраженного в кДа) [A8].

Все вышеперечисленные структурные параметры (R_g , D_{max} , MM, V_p) могут быть рассчитаны и для полидисперсных систем, однако следует учитывать, что, в отличие от монодисперсных систем, экспериментально полученные значения этих параметров будут относиться не к отдельным частицам, а являться усредненными оценками по ансамблю частиц.

Функция распределения по расстояниям p(r) содержит ту же информацию, что и интенсивность рассеяния I(s), но позволяет на интуитивном уровне делать выводы о форме частицы по самому виду кривой p(r) (Рис. 1.6). Например, глобулярные частицы имеют колоколообразные профили p(r) с максимумом вблизи $D_{max}/2$, для вытянутых частиц характерны асимметричные распределения p(r) с максимумом на малых расстояниях, соответствующим радиусу их поперечного сечения. Плоские частицы обладают довольно широким максимумом, смещенным к расстояниям меньшим, чем $D_{max}/2$. В то же время, максимум, смещенный в сторону расстояний, превышающих $D_{max}/2$, обычно указывает на наличие полых частиц.

Наконец, многодоменные частицы или макромолекулярные комплексы часто обладают профилями с несколькими максимумами и осцилляциями между ними, соответствующими внутри- и межсубъединичным расстояниям (Рис. 1.6).

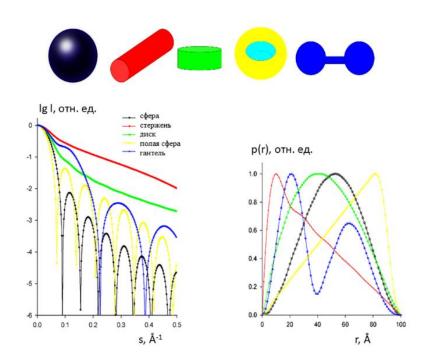


Рис. 1.6 Интенсивности малоуглового рассеяния (слева) и функции распределения по расстояниям p(r) (справа) от различных геометрических тел с одинаковым максимальным размером.

Даже для простых геометрических тел существует лишь несколько случаев, когда функции I(s) и/или p(r) могут быть выражены аналитически. Наиболее известными являются выражения для однородной сферы радиуса R: $I(s) = A(s)^2$, $A(s) = (4\pi R^3/3)[\sin(x) - x\cos(x)]/x^3$ где x = sR, и $p(r) = (4\pi R^3/3)r^2(1 - 3t/4 + t^3/16)$, где t = r/R. Полуаналитические (численные) уравнения для интенсивности рассеяния эллипсоидами, цилиндрами и параллелепипедами были получены Миттельбахом и Породом в 1960-х годах [35, 36], позже также были опубликованы аналитические формулы для функции p(r) для некоторых форм (например, куба [37] и тетраэдра [38]).

Вычисление p(r) является нетривиальной процедурой, поскольку нам доступен только ограниченный угловой диапазон I(s) (от s_{min} до s_{max}), и прямое проведение

преобразования Фурье будет приводить к сильным искажениям из-за эффектов обрыва ряда. Решением этой проблемы является косвенное преобразование Фурье, впервые предложенное О. Глаттером [39].

В этом подходе p(r) параметризуется в виде линейной комбинации ортогональных функций в диапазоне $[0, D_{max}]$,

$$p(r) = \sum_{k=1}^{K} c_k \, \varphi_k(r) \tag{1.16}$$

причем D_{max} в этом случае является задаваемым параметром, а коэффициенты c_k , описывающие p(r), определяются так, чтобы получить гладкую функцию, при этом одновременно обеспечивающую наилучшее соответствие с экспериментальными данными и минимизацию функционала

$$\Phi(\alpha) = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} [I(s_i) - \sum_{k=1}^{K} c_k \, \mathcal{Y}_k(s_i)]^2 + \alpha \int_0^{D_{max}} \left[\frac{dp(r)}{dr} \right]^2 dr$$
 (1.17)

Здесь $\Psi_{\kappa}(s)$ являются Фурье-образами функций $\varphi_{\kappa}(s)$ (в случае данных МУРН или использовании камеры Кратки необходимо также введение коллимационных поправок для учета "замазывания" сигнала). Регуляризационный параметр α используется для баланса между точностью соответствия данным (первое слагаемое в уравнении (1.17)) и гладкостью функции p(r) (второе слагаемое). Априорная оценка D_{max} , может быть уточнена позднее путем итерационных расчетов p(r) с различными значениями D_{max} . Степень гладкости p(r) должна быть соответствующим образом учтена для получения стабильных решений, но главное, что такой подход не требует экстраполяции данных за пределы экспериментально измеренного интервала и позволяет учитывать инструментальное размытие Существуют различные реализации сигнала. метода косвенного преобразования в компьютерных программах GIFT [39], GNOM [40], IRENA [41] и др.

Основной проблемой при использовании метода косвенного преобразования является выбор правильного значения регуляризационного параметра α . При слишком малых значениях α решения будут неустойчивы к экспериментальным ошибкам интенсивностей, в то время как при слишком больших значениях α

решения будут иметь систематические отклонения от экспериментальных данных (Рис. 1.7). В программе GNOM [40] было предложено использовать набор критериев ('perceptual criteria'), который оценивает качество решения (по шкале от 0 до 100%), и помогает пользователю программы сделать правильный выбор параметра α .

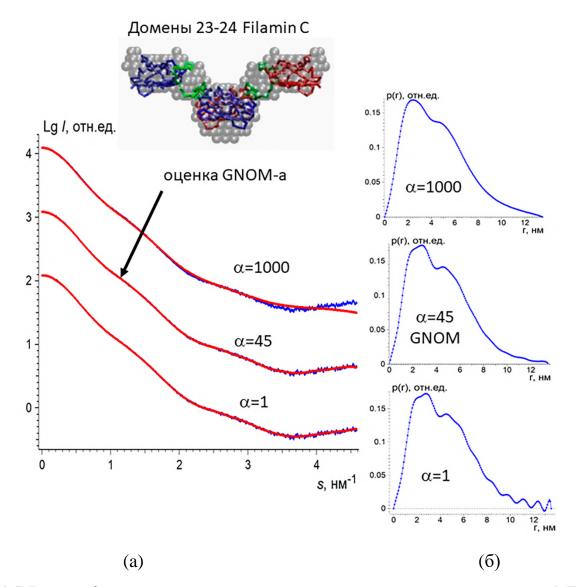


Рис. 1.7 Расчет функции распределения по расстояниям p(r) по данным МУРР с помощью программы GNOM. Экспериментальная кривая МУРР димерной структуры доменов 23-24 Filamin C [A11] (слева, синяя кривая). Влияние изменения параметра регуляризации α на качество приближения экспериментальных данных (а) и на степень гладкости функции p(r) (б). Если значение параметра α слишком велико, рассчитанный профиль p(r) хотя и гладкий/стабильный, но рассчитанная от него кривая рассеяния не соответствует экспериментальным данным, если же параметр α слишком мал, то происходит неоправданное приближение низкочастотной шумовой составляющей в данных,

приводящее к нефизичным осцилляциям в функции p(r). Программа GNOM использует набор критериев для автоматического выбора оптимального значения параметра α (α =45).

Для автоматического расчета функции p(r) и определения ее максимального размера D_{max} была разработана автоматизированная версия программы GNOM (под названием DATGNOM) [A7]. Для этого в программе DATGNOM автоматически выполняется несколько запусков программы GNOM с целью поиска оптимальных D_{max} и p(r). Значения D_{max} в диапазоне от $2R_g$ до $4R_g$ сканируются с шагом $0.1*R_g$. В дополнение к стандартным критериям программы GNOM здесь учитывается нормированное значение $p'(D_{max})$, гарантирующее, что функция p(r) плавно подходит к нулю при D_{max} .

преобразования Фурье Методы косвенного можно применять ДЛЯ полидисперсных систем, в которых частицы имеют схожую форму и отличаются только размером (см. уравнение 1.4). В большинстве практических случаев предполагается, что форма частиц известна (в частности, для изотропных систем частицы часто считаются сферическими), и уравнение (1.4) используется для поиска функции объемного распределения частиц по размерам $D_V(R)$. Это можно сделать с помощью описанного выше метода косвенного преобразования (функция $D_{V}(R)$ раскладывается на ортогональные функции, аналогично функции p(r) в уравнении (1.16), на интервале [R_{min} , R_{max}]). В программе GNOM также имеется возможность расчета функций распределения частиц по размерам как для сферических частиц, так и для частиц произвольной формы, задаваемой пользователем. В Главе 3 будут рассмотрены альтернативные подходы к восстановлению функций распределения частиц по размерам, основанные на параметрическом задании функции $D_V(R)$ набором аналитических функций или ее прямого поиска в виде гистограммы.

По форме кривой МУРР/МУРН можно оценить степень сворачивания белковых макромолекул и возможное наличие гибких участков в структуре белка. Традиционно для этого используется график Кратки ($s^2I(s)$ как функция s)

Интенсивность рассеяния от однородного тела с небольшой анизометрией спадает под большими углами согласно закону Порода как s^{-4} , что приводит к колоколообразному графику Кратки с хорошо выраженным максимумом [42]. С другой стороны, идеальная гауссова цепь (например, в случае полимеров) имеет асимптотику s^{-2} и поэтому график Кратки выходит на плато при больших значениях s. В целом, эксперименты по химическому или термическому разворачиванию белковых цепей, наблюдаемые с помощью МУРР, являются хорошими примерами того, как уменьшение компактности белка отражается на графиках Кратки [43, 44]. Для того чтобы сравнить состояния сворачивания различных белков, удобно пользоваться безразмерным графиком Кратки (Рис. 1.8), где по вертикальной оси графика откладываются значения $(sR_g)^2I(s)/I(0)$, а по горизонтальной оси безразмерная величина sR_g .

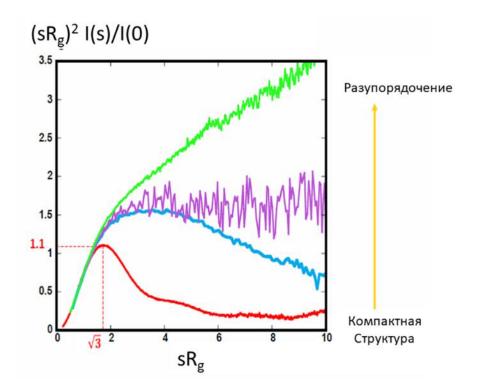


Рис. 1.8 Безразмерный график Кратки. Для глобулярных белков характерно наличие колоколообразного пика при sR_g равным $\sqrt{3}$. По мере развертывания белков или увеличения их гибкости график Кратки выходит на плато в области больших углов или даже имеет возрастающую зависимость.

В работе [45] было показано, что компактные глобулярные белки на таком графике дают колоколообразную кривую в области малых углов с максимумом при sR_g

равным √3. Отклонения от этого поведения указывают на гибкость или разупорядоченность частиц.

1.4 Методы ab initio восстановления формы макромолекул

Ab initio реконструкция трехмерных (3D) структур по данным изотропного малоуглового рассеяния возможна, строго говоря, только для монодисперсных невзаимодействующих ансамблей частиц и с достаточно низким разрешением (1-2 нм). При таком разрешении поиск обычно ограничивается однородными моделями (т.е. определением формы частицы). В 1960-е годы моделирование формы частиц проводилось методом проб и ошибок путем вычисления кривых рассеяния от различных форм, построенных из элементов конечного объема (например, шариков), и сравнения их с экспериментальными данными. Эти методы проб и ошибок были вытеснены в 1970-х годах, благодаря использованию сферических гармоник при расчете малоугловой кривой рассеяния [46, 47]. Начиная с конца 1990-х годов, в связи со значительным увеличением вычислительной мощности компьютеров, основное "бремя" построения моделей было передано последним. Идея определения формы частицы методом конечных элементов заключается в том, чтобы начать поиск с некоторой стартовой (обычно случайной) конфигурации модели и модифицировать ее небольшими шагами, чтобы в итоге построить компактную модель, обеспечивающую хорошее соответствие малоугловому эксперименту. На практике применяются различные стохастические алгоритмы минимизации (например, генетический алгоритм или имитация отжига), и различные стратегии, использующие ограниченные и неограниченные объемы поиска, которые реализованы в компьютерных программах, таких как SASHA [48], DALAI_GA [49], SAXS3D [50], SASMODEL [51], DAMMIN [52], GASBOR [53], GA_STRUCT [54], DAMMIF [55], DENFERT [56] и др.

В настоящее время *ab initio* определение 3D формы частиц стало стандартным инструментом анализа данных МУРР/МУРН. Однако, из-за потери существенной доли информации вследствие сферического усреднения различных ориентаций

частиц в растворе, однозначное восстановление формы частицы принципиально невозможно, и необходимо следить за тем, чтобы не переинтерпретировать данные МУРР/МУРН. Методы *ab initio*, направленные на реконструкцию трехмерной формы из одномерной кривой малоуглового рассеяния, не могут дать однозначного решения, и при многократном запуске программ получаются несколько разные модели. Это связано как с математической неоднозначностью определения трехмерной функции формы частиц по одномерным данным рассеяния, так и с неустойчивостью решения плохо обусловленной решаемой обратной задачи. Тем не менее произвол в решениях ограничен тем, что все они должны иметь одинаковые структурные инварианты (рассмотренные выше параметры, математически однозначно определяемые из функции рассеяния): радиус инерции, максимальный размер, объем, площадь поверхности.

Сравнение находимых решений позволяет выделить наиболее устойчивые особенности из ряда моделей, которые одинаково хорошо соответствуют данным, но имеют различия в форме. Программы SUPCOMB [57], SUPALM [A12] и DAMAVER [58] позволяют выбрать наиболее вероятную/представительную модель из полученного ансамбля моделей и создать "консенсусную" усредненную модель. Кроме того, необходимо всегда помнить, что кривые малоуглового рассеяния инвариантны по отношению к зеркальному отражению координат атомов, поэтому при анализе всегда следует рассматривать энантиоморфный вариант имеющейся модели (в частности, программы суперпозиции и усреднения SUPCOMB/SUPALM/DAMAVER моделей допускают учет такого рода энантиоморфизма).

Программа *ab initio* определения формы DAMMIN [52] представляет частицу как набор большого числа ($N\sim10^3$ - 10^4) плотно упакованных шариков внутри искомого объема (например, сферы с диаметром D_{max} , равным максимальному размеру частицы). Размеры шариков r_0 должны быть намного меньше максимального диаметра частицы ($r_0 << D_{max}$). Каждый шарик может принадлежать либо частице (тогда его амплитуда рассеяния принимается равной 1), либо растворителю (в этом

случае амплитуда рассеяния равна 0), и форма частицы, таким образом. описывается бинарной строкой X длины M (где M общее число шариков внутри объема поиска) (Рис. 1.9).

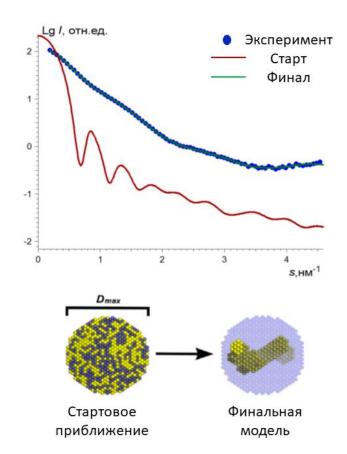


Рис. 1.9 Процедура *ab initio* моделирования с помощью программы DAMMIN [52]. Начиная со случайной конфигурации модели внутри сферического объема поиска, проводится минимизация методом имитации отжига до тех пор, пока не будет получена конечная модель, не только наилучшим образом приближающая экспериментальные данные, но и представляющая собой компактную и связанную модель. Внизу изображен сферический объем поиска с шариками, принадлежащими частице (желтый) или растворителю (синий).

Интенсивность рассеяния от такой модели можно быстро рассчитать с использованием разложения по сферическим гармоникам [46, 47]

$$I(s) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} |A_{lm}(s)|^2$$
 (1.18)

причем парциальные амплитуды рассеяния $A_{lm}(s)$ малых объемных элементов (шариков) определяются как

$$A_{lm}(s) = i^{l} \sqrt{2/\pi} \sum_{j} j_{l}(sr_{j}) Y_{lm}^{*}(\omega_{j})$$
(1.19)

где суммирование по j ведется по всем шарикам с полярными координатами $(r_j \ \omega_j)$, принадлежащим частице, $j_l(sr)$ - сферические функции Бесселя, а $Y_{lm}(\omega)$ — функции сферических гармоник.

Начиная с некоторого состояния системы (произвольной бинарной строки X), применяется поиск методом Монте-Карло с помощью случайного изменения принадлежности шариков для нахождения конфигурации, которая соответствует экспериментальным данным МУРР/МУРН. Такой поиск в ограниченном объеме был впервые предложен [49], который использовал генетический алгоритм (программа DALAI_GA). В программе DAMMIN используется альтернативный метод поиска глобального минимума — алгоритм имитации отжига (Simulated Annealing, SA) [59] для поиска компактной модели, которая соответствует кривой рассеяния "от формы частицы", полученной после вычитания фона атомного рассеяния как соответствующей константы из экспериментальных данных. Это вычитание проводится с целью приведения спадания интенсивности в области больших углов в максимальное соответствие с законом Порода для однородных частиц (см. уравнение (1.13)). В минимизируемый функционал f(X) помимо невязки χ^2 между рассчитанной и экспериментальной кривыми рассеяния

$$\chi^{2} = \frac{1}{N-1} \sum_{j=1}^{N} \left[\frac{I(s_{j}) - I_{\exp}(s_{j})}{\sigma(s_{j})} \right]^{2}, \qquad f(X) = \chi^{2} + P(X)$$
(1.20)

также входят штрафные слагаемые, P(X), требующие, чтобы шарики находились в контакте, а модель была компактной. Это условие гарантирует получение физически осмысленных структур низкого разрешения.

Следует также отметить, что программа DAMMIN позволяет учитывать различные типы симметрии частиц. Симметрия вводится как параметр модели, т.е. начальная генерация, рандомизация и каждая модификация модели во время процедуры SA выполняются для групп симметричных шариков одновременно. Наличие симметрии дает правила отбора сферических гармоник, которые используются для

расчета рассеяния от шариковой модели в программе DAMMIN. Эти условия позволяют не только получить более адекватные модели симметричных частиц, но и значительно ускорить расчеты.

В настоящее время программа DAMMIN поддерживает точечные группы Pn до n=19-кратной оси симметрии вращения, группы Pn2 до n=12, кубические группы P23 и P432, а также икосаэдрическую симметрию. Все эти группы либо совместимы с плотной гексагональной упаковкой шариков, либо могут быть реализованы путем генерации шариков на соответствующих слоях цилиндрических оболочек. Икосаэдрическая симметрия требует более сложной начальной генерации, которая подробно описана в работе [A13]. Возможность использования икосаэдрической симметрии особенно полезна при анализе вирусных частиц.

Если имеется априорная информация о форме частицы, ее также можно принять во внимание в процессе *ab initio* моделирования. В частности, когда известна общая анизометричность частицы, требование вытянутости или сплюснутости ее формы может быть добавлено в качестве мягкого ограничения при проведении минимизации. Кроме того, программа DAMMIN может проводить поиск и по несферической области (могут использоваться объемы эллипсоида, цилиндра, полого цилиндра или параллелепипеда). Для существенно анизометричных частиц такой поиск может улучшить стабильность решений и сэкономить время вычислений.

Как и в других методах типа Монте-Карло, при запуске программы DAMMIN, начиная с различных начальных приближений, могут быть получены различные модели. Для проверки результатов моделирования программу часто запускают несколько раз и результаты отдельных реконструкций сравнивают друг с другом и усредняют с помощью программ SUPCOMB/SUPALM и DAMAVER.

Хотя программа DAMMIN работает достаточно быстро, обычно это все же занимает от нескольких минут до нескольких часов на современных компьютерах. Для дальнейшего повышения скорости и производительности программа

DAMMIN была переписана, ее новая реализация получила название DAMMIF [55], где F означает быстрый. DAMMIF отличается от DAMMIN несколькими важными ограниченный объем поиска модели был неограниченный, который может увеличиваться по мере необходимости во время процедуры минимизации. Это улучшение помогает избежать артефактных граничных эффектов, которые могут возникнуть при использовании объема поиска со слегка недооцененным значением D_{max} ; (ii) перед расчетом амплитуд рассеяния модели, которые являются несвязными, немедленно отбрасываются, и расчет выполняется только для моделей, расположение шариков которых представлено связным графом; (ііі) необходимые амплитуды рассеяния заранее рассчитываются для каждого шарика, вносящего вклад в общее рассеяние. Все эти меры ускоряют процедуру моделирования в 25-40 раз по сравнению с оригинальным DAMMIN. Несмотря общую эффективность работы алгоритма, однозначное на восстановление формы частиц по данным МУРР, как было отмечено выше, не всегда оказывается возможным вследствие потери существенной информации, происходящей при сферическом усреднении по ориентациям частиц. В работе [60] была предложена количественная оценка (мера) возможной неоднозначности решений и показано, что она существенно выше для плоских частиц по сравнению с глобулярными и вытянутыми частицами. Поэтому для поиска частиц сложной формы, например, спиралевидных структур, необходимо использовать разные тактики в процессе оптимизации. В работе [61] была предложена новая схема попеременного варьирования относительного вклада членов целевой функции в формы (невязки процессе поиска частиц между рассчитанной экспериментальной кривыми рассеяния, штрафов за рыхлость и разрыв структуры, а также за отклонение среднего числа контактов шариков от заданного значения), позволившая существенно улучшить качество восстановления сложных форм частиц, в частности, спиралевидных.

Разрешение шариковых моделей при определении формы частицы ограничено предположением о равномерном распределении электронной плотности внутри

частицы, следовательно кривые рассеяния могут быть приближены только в ограниченном угловом диапазоне (обычно до значений векторов рассеяния *s* около 2.5 нм^{-1} , т.е. с разрешением 2.5 нм, где разрешение d формально определяется как $d=2\pi/s$). Другой подход для *ab initio* определения формы белков по данным МУРР использует тот факт, что белки представляют собой полипептидные цепи, состоящие из аминокислотных остатков, разделенных между собой примерно на 0.38 нм (типичное расстояние между соседними атомами C_{α} в первичной аминокислотной последовательности). Поэтому, до разрешения около 0.5 нм структуру белка можно рассматривать как совокупность идентичных фиктивных остатков (Dummy Residues - DR), с центром на позициях C_{α} атомов. 3D модель белка может быть построена по данным МУРР путем нахождения цепочечного пространственного расположения фиктивных остатков, которое наилучшим образом соответствует экспериментальным данным МУРР в области малых и средних углов. Использование положений C_{α} атомов позволяет накладывать ограничения на пространственное расположение виртуальных остатков. В дополнение к условию нахождения соседних (по номеру аминокислотной последовательности) остатков на расстоянии 0.38 нм друг от друга локальные взаимодействия близлежащих аминокислотных остатков приводят к характерному распределению ближайших соседей, которое можно использовать в качестве дополнительного ограничения при поиске модели. Количество виртуальных остатков должно быть равно количеству остатков в последовательности белка. Программа GASBOR [53] использует метод имитации отжига для построения белковоподобной цепочечной DR-модели внутри сферического объема, определенного аналогично программе DAMMIN. Такое DR-моделирование способно лучше учитывать внутреннюю структуру белковой макромолекулы, приближать данные МУРР в области углов, соответствующих более высокому разрешению (вплоть до s < 10 нм⁻¹), и в целом дает более детальные модели, чем те, которые дает алгоритм определения формы в программе DAMMIN. Помимо оригинальной версии GASBOR, приближающей данные МУРР в обратном

пространстве, существует версия программы для прямого пространства, которая приближает функцию распределения расстояний p(r). Программа GASBOR также способна учитывать симметрию и анизометрию частиц.

Ав initio анализ белков и белковых комплексов с помощью МУРР/МУРН очень полезно дополнять информацией, полученной другими методами. Например, кристаллические и ЯМР-структуры высокого разрешения могут быть пристыкованы к малоугловым моделям низкого разрешения, а электронные карты, полученные с помощью электронной микроскопии (ЭМ), могут быть использованы в качестве начальных объемов поиска при ab initio моделировании данных МУРР/МУРН [52].

Поскольку всегда имеется внутренняя неоднозначность *ab initio* восстановлений 3D формы частиц по данным МУРР/МУРН, то апостериорная оценка изменчивости моделей также имеет важное значение. Программа SASRES использует критерий корреляции в обратном Фурье пространстве (широко применяемый в электронной микроскопии [62]) для оценки изменчивости между несколькими *ab initio* реконструкциями и оценки разрешения получаемой 3D формы частицы [63]. Этот объективный критерий корреляции позволяет избежать чрезмерной интерпретации моделей, полученных на основе данных малоуглового рассеяния, и рассматривать только те особенности, которые важны на данном уровне разрешения.

Таким образом, *ab initio* 3D модели могут быть быстро восстановлены по данным МУРР/МУРН в автоматическом режиме, соответствующие программы были включены в пакеты сбора и обработки данных на многих синхротронных станциях. В частности, на станции X33 [A1-A2] была реализована одна из первых автоматизированных систем анализа данных МУРР [A7]. Позднее были созданы их более совершенные варианты [64, 65], и в настоящее время на многих станциях МУРР пользователь может получать предварительную информацию о результатах синхротронного эксперимента МУРР фактически в режиме реального времени.

В настоящее время применение *ab initio* методов стало рутинной процедурой для структурных биологов при определении трехмерных форм белков и белковых комплексов на низком разрешении, что можно легко найти во многих публикациях [66-69]. Приведем несколько характерных примеров их использования.

В работе [А14] методом МУРР исследовалась структура белка филамина А, который сшивает актиновые нити в ортогональные сети в цитоплазме и участвует в закреплении мембранных белков для актинового цитоскелета. Этот белок состоит из N-концевого актин-связывающего домена и 24 иммуноглобулиноподобных доменов (IgFLN) (Рис. 1.10). Филамин А взаимодействует с многочисленными белками, включая важные мембранные рецепторы, и большая часть таких взаимодействий затрагивает С-концевые домены IgFLN 16-24. По данным МУРР были восстановлены ab initio 3D формы ряда тандемных доменных фрагментов филамина A: IgFLN 12-13, IgFLN 16-17, IgFLN18-19, IgFLN20-21, IgFLN22-23 (Рис. 1.10). Было обнаружено, что структура тандемных доменов IgFLNa16-17, IgFLNa18-19 и IgFLN 20-21 значительно более компактна, чем можно было бы ожидать для двух независимых доменов, и указывает на специфичные домендоменные взаимодействия. В отличие от них фрагменты IgFLN12-13, IgFLN22-23 имеют более удлиненную форму, что характерно при малой поверхности взаимодействия между тандемными доменами. Впоследствии, независимым образом были получены структуры высокого разрешения тандемных доменов филамина А методом ЯМР, которые идеально вписались в 3D формы белков, подтвердив надежность работы ab initio алгоритма.

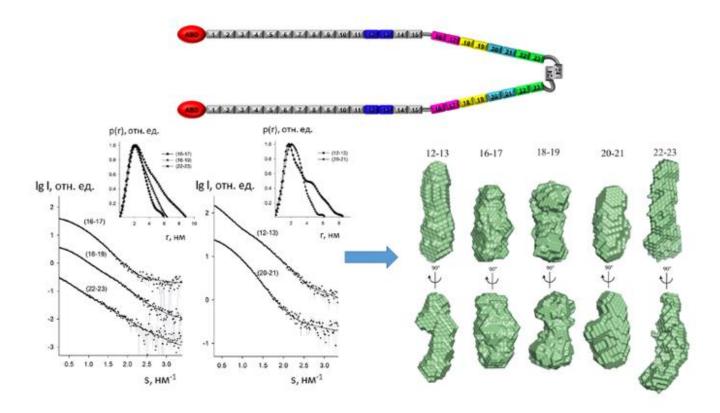


Рис. 1.10 Доменная организация филамина A (сверху), экспериментальные данные МУРР (слева) и восстановленные *ab initio* 3D формы тандемных доменов филамина A (справа).

Выживаемость холерных бактерий в окружающей среде усиливается благодаря белку GbpA фактора колонизации, который облегчает колонизацию молекул вибриона путем медиации бактериального присоединения к хитиновым поверхностям. Исследование структуры методом МУРР показало, что белок GbpA состоит из четырех доменов и имеет стержнеобразную форму с частичным поворотом в одном из доменов [A15] (Рис. 1.11). Полученная структура полноразмерного белка GbpA показывает, что в ней не существует прямых междоменных взаимодействий, кроме полипептида главной цепи, связывающего домены вместе. Кроме того, поверхность каждого домена может быть полностью подвержена воздействию внешней среды. Белок обладает удлиненной структурой в растворе со значительной гибкостью между каждым доменом. Эти результаты позволили предположить механизм молекулярных взаимодействий белка GbpA и бактериальной клетки, что чрезвычайно важно при разработке вакцин.

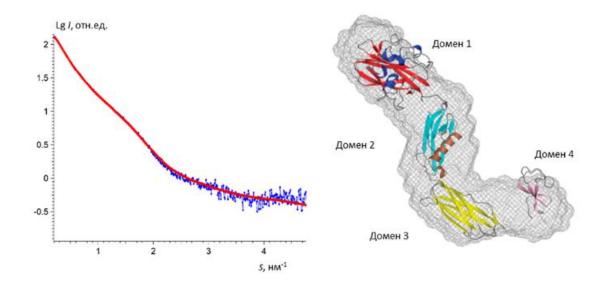


Рис. 1.11 Фактор колонизации GbpA холерного вибриона, экспериментальные данные МУРР (слева) и восстановленная *ab initio* 3D форма полноразмерного белка GbpA (справа). Известные структуры доменов высокого разрешения хорошо вписываются в *ab initio* форму.

В некоторых случаях *ab initio* методы позволяют получить не только форму молекул, но и определить структуру сольватационного слоя вокруг нее. В работе [A16] исследовалась конформация и сольватация олигомеров нативных таннинов (полифенольных соединений, содержащихся в коже растений и защищающих их от проникновения бактерий) в водно-спиртовом растворе. Размеры олигомеров и малоуглового сольватационного вклад слоя интенсивность рассеяния изменяются немонотонным образом с содержанием этанола, что является проявлением амфипатической природы молекул танина (Рис. 1.12а). Известно, что молекулы белков покрыты гидратной оболочкой толщиной около 3 Å, электронная плотность которой на 20% выше, чем у растворителя в объеме. Однако, геометрия сольватации в присутствии этанола оказывается более сложной. Минимальный исключенный объем белка достигается при 20% этанола, что соответствует случаю, когда плотность сольватационного слоя максимально близка плотности растворителя. Полифенолы взаимодействуют с молекулами растворителя по двум через гидрофобные взаимодействия ароматических механизмам:

посредством водородных связей с гидрофильными ОН-группами фенолов. Гидрофобные и гидрофильные участки на молекулах полифенолов располагаются на участках, отстоящих друг от друга на 3 Å и более, что сравнимо с размером молекул этанола и воды. Таким образом, возникает сложная микрогетерогенность вокруг молекул полифенола в смешанном растворителе. Чтобы более детально изучить гетерогенную сольватацию таннинов, был использован метол дифференциальной модели при помощи анализа различий между формами таннинов, восстановленными для различных составов растворителей. С помощью этого метода были найдены сольватационные слои вокруг молекул таннина в воде этаноле (Рис. 1.126). Модели показывают, что сольватация имеет неоднородности в масштабе 1 нм, что соответствует длине одного таннинового мономера плотноупакованного димера. Это подтверждает, или ЧТО неоднородности сольватационного слоя соответствуют размерам, определяемым расположением полифенольных единиц в полимере.

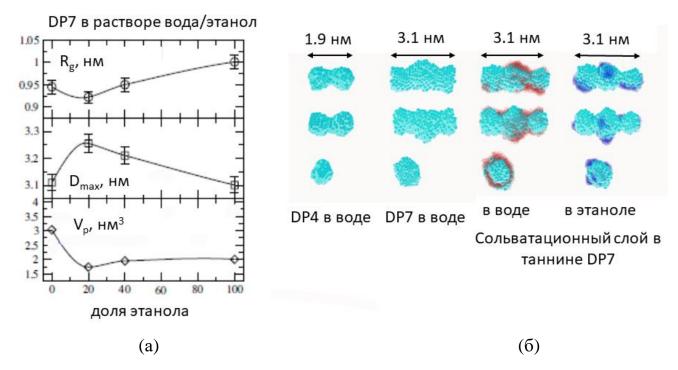


Рис. 1.12 (а) Радиус инерции R_g , максимальный размер частиц D_{max} и исключенный объем Порода V_p являются немонотонными функциями объемной доли этанола в растворе. (б) Эффект преференционной сольватации для нативного танина DP7 (средняя степень полимеризации N=6.3). Сольватационный слой показан красным и синим цветом.

1.5 Методы молекулярной тектоники для моделирования комплексов биомакромолекул

Если имеется структура высокого разрешения или гомологичная модель макромолекулы или комплекса, то от нее можно рассчитать теоретическую кривую рассеяния и сравнить с экспериментальными данными МУРР/МУРН. Этот расчет должен учитывать, вообще говоря, три фактора: атомное рассеяние частицей в вакууме, исключенный объем (объем, занимаемый биомолекулой в растворе и недоступный для растворителя) и рассеяние, возникающее от гидратной оболочки, образующейся вокруг молекулы в растворе. Получение хорошего согласия с экспериментом свидетельствует о том, что общая структура модели высокого разрешения действительно совпадает со структурой макромолекулы в растворе. Такой подход часто используется для проверки предсказанных моделей, идентификации биологически активных конформаций и кристаллографических олигомеров в растворе, а также для анализа возможных различий четвертичной структуры белка в кристалле и растворе. Для расчета теоретических кривых рассеяния на основе атомных моделей ранее был разработан ряд программ, такие как CRYSOL [70], FoXS [71], AXES [72], AquaSAXS [73], HyPred [74] и др. Алгоритмы программ в основе своей схожи, но имеют разные реализации для представления вкладов от молекул воды, в том числе, с использованием методов молекулярной динамики.

Программы CRYSOL для данных МУРР [70] и CRYSON [75] для измерений МУРН широко используются для расчета профилей рассеяния от атомных моделей макромолекулярных структур. Эти программы рассчитывают рассеяние от атомной модели частицы в растворе как

$$I(s) = \langle |A(s)|^2 \rangle_{\Omega} = \langle |A_a(s) - \rho_s A_s(s)|^2 \rangle_{\Omega}$$
(1.21)

где $A_a(s)$ - амплитуда рассеяния (рентгеновского или нейтронного) от частицы в вакууме, $A_s(s)$ и $A_b(s)$ - соответственно амплитуды рассеяния от исключенного объема и гидратной оболочки, обе с единичной плотностью, а <...> обозначает сферическое среднее в обратном пространстве (Рис. 1.13).

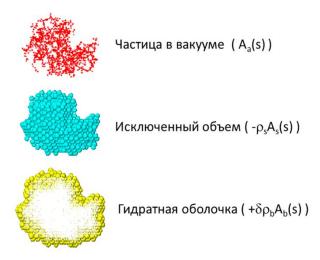


Рис. 1.13 Схематичное представление основных вкладов в амплитуду рассеяния при расчете интенсивности МУРР/МУРН от атомной модели высокого разрешения.

Уравнение (1.21) учитывает, что плотность связанного с частицей растворителя ρ_b может отличаться от плотности основного объема растворителя ρ_s . Это приводит к оболочки ненулевому контрасту гидратной $\delta \rho_h$ Программы $=\rho_b$ - ρ_s . CRYSOL/CRYSON приближают экспериментальную кривую рассеяния с учетом атомных координат модели, оптимизируя исключенный объем частицы и контраст гидратного слоя, окружающего частицу в растворе, чтобы минимизировать невязку χ^2 (см. уравнение (1.20)). При отсутствии экспериментальных данных они могут теоретическую картину рассеяния, используя предсказать параметры умолчанию или заданные пользователем. В частности, предполагается, что плотность рассеяния гидратной оболочкой на 10% больше, чем плотность рассеяния в объеме растворителя, это предположение было подтверждено экспериментально в комбинированном исследовании МУРР/МУРН [75]. Три члена в уравнении (1.21) вычисляются с помощью разложения через сферические гармоники, аналогичного уравнениям (1.18-1.19). Кроме того, парциальные амплитуды растворе $A_{lm}(s)$, вычисленные программами частицы В CRYSOL/CRYSON, могут быть использованы для быстрого расчета рассеяния от комплексов методом молекулярной тектоники.

Для более точного предсказания данных широкоугольного рассеяния (Wide-Angle X-ray Scattering - WAXS), где потенциально может быть извлечена информация как о третичной, так и о вторичной структуре в частности, используются программы SoftWAXS [76] и WAXSIS [77]. Эти подходы также включают альтернативные методы для аппроксимации гидратной оболочки, в том числе добавление в явном виде молекул воды в растворителе с плотностью, соответствующей плотности растворителя в объеме образца.

Собрать полную структуру макромолекулярного комплекса можно путем стыковки отдельных его компонентов внутрь *ab initio* формы, полученной по данным МУРР/МУРН [78]. Однако, поскольку разрешение *ab initio* моделей является низким, то более надежным способом является моделирование сборки таких комплексов путем прямого приближения данных МУРР/МУРН. Ряд интерактивных и автоматизированных подходов для определения положения и ориентации субъединиц в макромолекулярных комплексах был разработан с использованием малоугловых данных [79-82], [А17]. На основе известных структур (или гомологичных моделей) субъединиц, теоретическое рассеяние комплекса может быть быстро рассчитано с помощью использования разложения на сферические гармоники, которое было рассмотрено в разделе 1.4. Этот формализм, лежащий в основе программ CRYSOL и CRYSON, применяется для интерактивного моделирования в программе MASSHA [А17] (более подробно она будет рассмотрена в Главе 7) или автоматизированного моделирования в программах SASREF [81], FoxsDock и MODELLER [83].

Программа SASREF - это комплексная автоматизированная программа моделирования методом молекулярной тектоники (или rigid body modelling - моделирование жесткими телами) по данным МУРР/МУРН, позволяющая одновременно приближать несколько кривых рассеяния (например, когда измеряется несколько неполных конструкций белка или комплекса, а также для данных МУРН с вариацией контраста). При моделировании имеется возможность использования симметрии, взаимных ориентационных ограничений субъединиц

(например, полученных методом ЯМР), ограничений на контакты между аминокислотными остатками (например, по результатам мутагенеза экспериментов по химическому сшиванию связей) и меж-субъединичным расстояниям (например, по данным FTIR и FRET). Начиная с произвольного расположения субъединиц, программа SASREF [81] проводит итеративную серию случайных движений и поворотов субъединиц комплекса (или доменов белка), используя метод имитации отжига для поиска наилучшего соответствия рассчитанного от модели рассеяния к экспериментальным данным МУРР/МУРН. Целевая функция, которую необходимо минимизировать, имеет форму уравнения (1.20), в которую входит невязка χ^2 (возможно, рассчитанная по нескольким кривым рассеяния) и штрафной член P(X). Последний может включать ограничения, полученные из других методов, таких как упомянутые выше, но при этом он всегда включает в себя и ограничения, гарантирующие, что получаемые модели взаимосвязаны и не имеют пространственных пересечений основных полипептидных цепей (пересечения боковых цепей при этом не учитываются). В качестве примера эффективного использования метода моделирования жесткими телами рассмотрим исследование методами МУРР и ЯМР тройного белкового комплекса IscS/IscU/CyaY, проведенное в работе [A18]. В этот комплекс входит белок фратаксин IcsS, ответственный за развитие нейродегенеративной болезни атаксии Фридриха и участвующий в сборке Fe-S кластера, фермент десульфураза (IscU) и каркасный (scaffold) белок-ингибитор (CyaY). Фратаксин действует либо как шаперон, который доставляет железо в комплекс IscS/IscU [84], либо как акцептор [85], который за счет образования крупных железосодержащих ассоциатов предотвращает излишки железа в митохондриях, сохраняя при этом биологическую доступность железа для доставки к ключевым акцепторам. В работе [А18] было проведено структурное исследование бинарных и тройных IscS/IscU/CyaY комплексов методом МУРР. С помощью программы SASREF была определена стехиометрия и построена модель тройного комплекса путем

одновременного приближения данных МУРР от двух бинарных (IscS/IscU и IscS/CyaY) и полного тройного комплекса (IscS/IscU/CyaY) (Puc. 1.14).

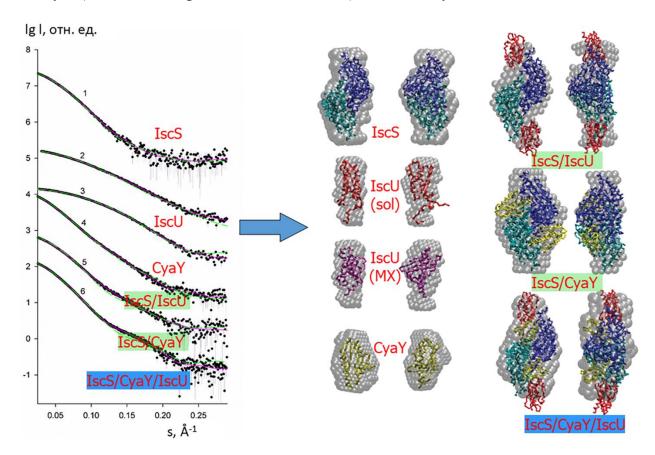


Рис. 1.14 Экспериментальные кривые МУРР от IscS/IscU/CyaY белков/белковых комплексов и приближения, полученные методом ab initio и молекулярной тектоники (малиновые и зеленые кривые, соответственно) (слева). Восстановленные ab initio 3D формы и модели белковых комплексов (справа).

Было установлено, что белок IscS существует в растворе в виде димера, тогда как два других белка (IscU и CyaY) остаются в мономерном состоянии. Оценка значений исключенных объемов по данным МУРР указывает на следующую стехиометрию бинарных (2:2) и тройного (2:2:2) комплексов. Форма бинарного комплекса IscS/CyaY является достаточно глобулярной, это позволяет предположить, что белок CyaY расположен вблизи полости у границы раздела между двумя мономерами IscS. Комплекс IscS/IscU имеет более удлиненную структуру, где молекулы IscU расположены на периферии димера IscS. Полученная методом МУРР модель тройного комплекса (IscS/IscU/CyaY) была затем подтверждена методом ЯМР путем измерения спектрального возмущения ¹⁵N-

меченного CyaY, титрованного IscS и далее IscU до молярного соотношения 1:1:1. Поверхность взаимодействия белка IscS с IscU и CyaY была подтверждена с помощью использования мутагенеза, для чего методом ЯМР исследовались образцы IscS с мутациями остатков, влияющих на взаимодействие с CyaY и IscU. На основании полученной модели тройного комплекса был сделан важный вывод о механизме сборки и регуляции Fe-S кластера.

1.6. Восстановление неизвестных частей структуры и учет гибкости макромолекул

В атомных моделях белков высокого разрешения, полученных с помощью рентгеноструктурного анализа или методом ЯМР, часто не хватает некоторых частей полипептидных цепей. Типичным примерами являются многодоменные белки с линкерами (короткими связующими звеньями между доменами). Такие белки можно разделять биохимическими методами на отдельные домены для облегчения кристаллизации, однако неупорядоченные области (например, N- или С-концы аминокислотной последовательности белка), могут быть не видны даже при высоком разрешении из-за их гибкости. Тем не менее, эти области, отсутствующие в моделях высокого разрешения, вносят свой вклад в малоугловое рассеяние макромолекулы в растворе.

Ранее были разработаны подходы, позволяющие достраивать отсутствующие части структуры и восстанавливать их приблизительные конфигурации по данным МУРР/МУРН [81, 86]. В ходе такого гибридного моделирования, сочетающего в себе *ab initio* подход и метод молекулярной тектоники, на место неизвестных областей структуры добавляются фиктивные атомные остатки (Dummy Residues - DR) так, чтобы можно было адекватно рассчитать интенсивность рассеяния от всей молекулы. В программе BUNCH [81] многодоменный белок может содержать жесткие домены с известной структурой, соединенные цепочками DR, играющими роль недостающих гибких линкеров. Метод имитации отжига (SA) используется для поиска оптимального расположения доменов и вероятной конфигурации

линкеров DR, соответствующей данным МУРР/МУРН, при этом конформационное пространство исследуется путем случайного вращения частей модели вокруг случайно выбранных DR остатков.

Для описания более сложных систем, например, комплексов, состоящих из нескольких субъединиц, в которых у одного или нескольких компонентов отсутствуют фрагменты их структуры заметной длины, была разработана программе CORAL [A8]. CORAL, аналогично программе SASREF, перемещает и вращает атомные модели отдельных доменов, принадлежащих к нескольким компонентам комплекса. Разница заключается в том, что эти перестановки не являются полностью случайными: расстояния между N- и C-концевыми частями соседних доменов, принадлежащих одной белковой цепи, ограничиваются. Это достигается путем использования библиотеки случайных линкеров, состоящих из DR остатков, которая содержит различные конфигурации линкеров длиной от 5 до 100 аминокислот. Когда домен в процессе моделирования перемещается, его новая позиция исследуется путем запроса данной библиотеки. Тогда в структуру добавляется линкер соответствующей длины для соединения этого домена с предыдущим доменом, если же подходящий линкер не может быть найден, то такое движение отвергается. N- и С-концевые части субъединиц, если они отсутствуют, также могут быть выбраны случайным образом из библиотеки линкеров, но они не ограничивают движение соответствующего домена. Кроме того, в программе CORAL имеется возможность согласованного движения группы доменов, когда выбранные домены сохраняют свое взаимное расположение при изменении их положения и ориентации по отношению к остальным. Эта бывает особенно полезно, когда в одной системе существует несколько известных интерфейсов олигомеризации.

Интересным примером достраивания неизвестных фрагментов белковой цепи по данным МУРР является исследование белка PSH, обладающего мощными радиопротекторными свойствами [A19]. Белок PSH был создан генно-инженерным путем и включает в себя человеческий пероксиредоксин (Prx6) и Mn-содержащей

супероксиддисмутазы из Escherichia (MnSOD) (Рис. 1.15). Этот белок бифункционален, так как одновременно обладает супероксиддисмутазной и пероксидазной активностью. Такой искусственно созданный фермент является мощным антиоксидантом, который способен вдвойне эффективнее бороться с разного рода свободными радикалами, возникающими при окислительном стрессе (который развивается, например, при наличии резаных и ожоговых ран, радиационных поражений, трансплантаций органов и др.).

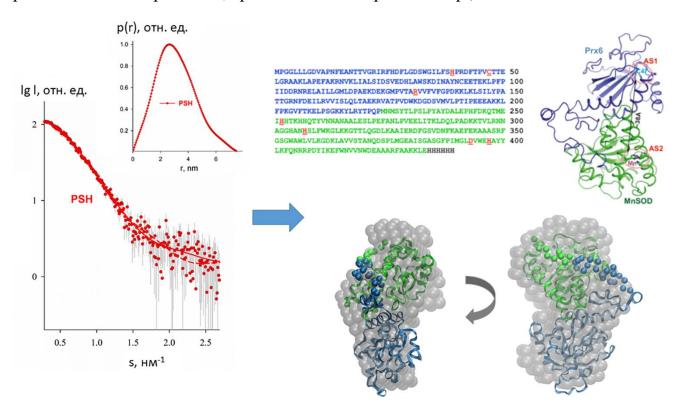


Рис. 1.15 Экспериментальная кривая МУРР гибридного белка-фермента PSH и приближения, полученные ab initio методом и молекулярной тектоникой с достраиванием неизвестного фрагмента аминокислотной последовательности (172-234) (сплошная и пунктирная красные кривые, соответственно). На вставке изображена функция распределения ПО расстояниям p(r). Ha полной аминокислотной последовательности синим цветом показан пероксиредоксин, зеленым цветом Мп-содержащая супероксиддисмутаза, красным цветом – активные аминокислотные остатки. Достроенный по данным МУРР фрагмент (172-234) показан синими сферами.

Структуры индивидуальных белков PSH были известны, но было важно понять, что с ними произошло, когда их «сшили» в одно целое (Рис. 1.15). Анализ данных МУРР полученной новой структуры позволил восстановить модель комплекса с

программы CORAL (при моделировании были использованы помощью предсказательные структуры Prx6 и MnSOD с достройкой неизвестного фрагмента) и определить, что активные центры ферментов (область белка, где происходит «обезвреживание» свободного радикала) находятся на расстоянии 36 Å друг от друга и с одной стороны не мешают работе друг друга, а с другой - то, что они находятся в одной связке, позволяет им эффективно воздействовать на свободные радикалы разного типа (Рис. 1.15). Параллельно исследованиям структуры методом МУРР были проведены успешные клинические испытания радиопротекторных свойств белка на животных. В этих экспериментах белок PSH успешно защитил подопытных мышей от летальных и сублетальных доз радиации. Фактор снижения дозы для него составил 1.33 [А19]. На основе полученных данных в будущем планируется создать еще более совершенные версии белка PSH. заслуживает способность Отдельного рассмотрения метода МУРР/МУРН оценивать подвижность и гибкость частиц в растворе. Гибкие системы, строго говоря, являются полидисперсными системами, в которых содержатся различные В белковая ЭТОМ случае, каждая отдельная ТИПЫ присутствующая в исследуемом образце, дает аддитивный вклад в общий профиль рассеяния. Для системы невзаимодействующих частиц, содержащей K различных МУРР/МУРН компонентов смеси, интенсивность является линейной В комбинацией их вкладов:

$$I(s) = \sum_{k=1}^{K} \nu_k I_k(s)$$
 (1.22)

где v_k и $I_k(s)$ - объемная доля и интенсивность рассеяния от k-го компонента, соответственно.

Полидисперсность, обусловленная присутствием более чем одного типа гибкостью, требует макромолекул, связанного cИХ специального не моделирования, которое будет рассмотрено подробно в Главе 3. Например, программа OLIGOMER [A5] может быть использована для анализа смеси макромолекул и выяснения индивидуальных объемных долей компонентов, если известны их профили рассеяния, а для оценки общего числа компонентов часто применяется сингулярное разложение.

Наличие конформационной полидисперсности можно ожидать для большого круга объектов, например, это могут быть частично разупорядоченные белки, белки с метастабильными компонентами, обладающие гибкими связующими звеньями, белки Маловероятно, развернутые И т.д. ЧТО такие системы будут кристаллизоваться, и даже если это произойдет, то их гибкие компоненты не обязательно будут иметь специфические конформации. Одним из важных преимуществ метода МУРР/МУРН является возможность изучения таких систем в нативном состоянии, хотя для этого требуются специальные подходы к моделированию.

Гибкие системы объединяет то, что они обладают множеством различных конформаций в растворе; таким образом, имеется астрономическое количество возможных сосуществующих конформаций. Разработка реалистичных ансамблевых моделей разупорядоченных состояний белков является важным предметом исследований на протяжении многих лет [87, 88]. Для анализа данных МУРР/МУРН было разработано несколько подходов, позволяющих оценить подвижность/гибкость белков: метод оптимизации ансамбля (программа ЕОМ) [89, 90], поиск минимального ансамбля (программа MES) [91], программы BSS-SAXS [92], EROS [93], ENSEMBLE [94], ASTEROIDS [95], MultiFoXS [71]. Эти методы основаны на общей стратегии, содержащей два последовательных шага: (і) генерацию большого количества статических случайных моделей различных возможных конформаций данного белка, и (ii) выбор подмножества конформаций, которые вместе описывают экспериментальный профиль рассеяния с помощью уравнения (1.22). Реализации программ и методы оптимизации для выбора ансамбля могут отличаться для разных подходов. В некоторых случаях, анализ ансамблей на основе данных МУРР сочетается с анализом экспериментальных ЯМР данных (как в программах ENSEMBLE и EROS).

Рассмотрим более подробно подход ЕОМ, который исторически первым был применен для анализа данных МУРР/МУРН от подвижных систем. Программа ЕОМ [89, 90] использует большой случайный пул возможных конформаций, доступное конформационное охватывающий пространство, которого выбирается подмножество конформаций, которые, как смесь, приближают экспериментальные данные МУРР/МУРН. ЕОМ состоит из двух основных RanCh (Random Chain) программа создает начальный программ: рандомизированных конформаций и программа GAJOE (Genetic Algorithm Judging Optimization of Ensembles) выбирает оптимизированный под-ансамбль. Последний состоит из 10-50 конформаций и выбирается генетическим алгоритмом из большого набора моделей (обычно десять тысяч различных конформаций, кривые рассеяния от каждой конформации, предварительно вычисляются с помощью CRYSOL/CRYSON). программ Данные ансамбля рассеяния каждого рассчитываются путем суммирования индивидуальных картин рассеяния каждой конформации, согласно уравнению (1.22). Под-ансамбль с наилучшим согласием (наименьшей невязкой χ^2) к экспериментальным данным отбирается для дальнейшего анализа. Конечно, ансамбль из 10-50 структур слишком мал, чтобы объяснить конформационное поведение гибкого белка в растворе, но этих субпопуляций обычно достаточно для описания глобальных свойств белка с точки зрения размера и формы (т.е. R_g , D_{max} и анизометрии). Общая схема работы алгоритма ЕОМ представлена на Рис. 1.16.

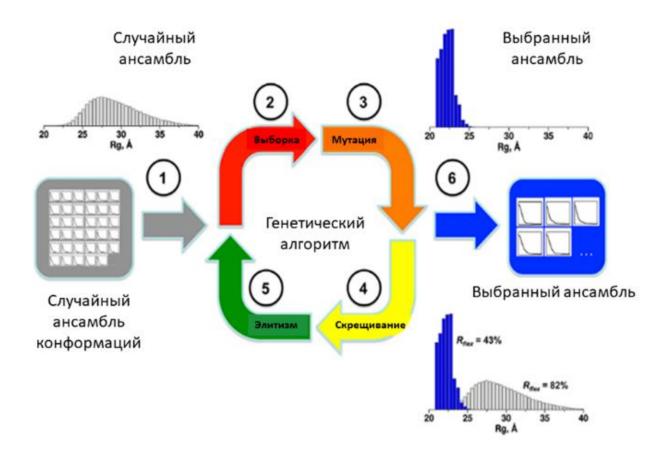


Рис. 1.16 Схема работы генетического алгоритма в методе оптимизации ансамбля (ЕОМ). Конформационное пространство системы исследуется путем генерации случайного пула с кривыми МУРР/МУРН, рассчитанными для каждой конформации (1). Подмножества этих кривых случайным образом выбираются для формирования "хромосом" (2), и эти хромосомы подвергаются серии этапов мутации (3) и скрещивания (4). Этап мутации позволяет менять местами/заменять кривые в пределах текущего набора хромосом, а также кривые, случайно выбранные из пула. На этапе скрещивания случайным образом выбираются две хромосомы из текущего набора и меняются местами минимум две кривые для создания новой хромосомы. Состав хромосом, которые обеспечивают наилучшее соответствие экспериментальным данным рассеяния, оценивается на этапе элитизма (5) и передается следующему поколению в итерационном процессе генетического алгоритма. Процесс повторяется до стадии (6), где параметры усредняются и строятся окончательные распределения размеров (например, $R_{\rm g}$, D_{max}) и энтропийной метрики (R_{flex}) отобранных (оптимизированных) ансамблей, которые сравниваются с параметрами исходного пула.

Начальный ансамбль (пул) может состоять не только из случайных моделей одноцепочечных неструктурированных белков. Программа ЕОМ позволяет следовать различным сценариям с учетом имеющейся априорной информации.

Например, сгенерированные модели могут включать упорядоченные области, домены или даже нуклеиновые кислоты. Для олигомерных структур могут быть сгенерированы асимметричные или симметричные мультимеры. Возможность объединения нескольких независимых пулов еще больше расширяет диапазон применимости программы ЕОМ. Например, в работе по исследованию полноразмерной митохондриальной глутаминазы С [96], ансамбли, состоящие из гибких димеров, тримеров и октамеров были объединены, что позволило объяснить, как неупорядоченные области белка влияют на его ферментативную активность и олигомерное состояние в растворе.

Если распределение R_g моделей в оптимизированном ансамбле такое же широкое, как и в начальном случайном пуле, белок, вероятно, является гибким; получение же узкого пика распределения R_g свидетельствует о том, что система является жесткой [97]. Более того, положение максимума распределения выбранного ансамбля по отношению к пулу дает представление о степени компактности или вытянутости структуры. Программа ЕОМ предлагает также количественную оценку степени гибкости макромолекулы, параметр R_{flex} , характеризующий информационную энтропию системы. Этот параметр изменяется в пределах от 0 до 100%, и позволяет количественно оценить разницу между гибкой и жесткой системами. В этом случае R_{flex} для выбранного (оптимизированного) распределения R_g ансамбля численно сравнивается со значением R_{flex} для распределения случайного пула, последнее представляет собой своего рода "эталон гибкости" исследуемой системы [90].

Использование метода ЕОМ оказалось очень эффективным при анализе гибкости белков и белковых комплексов в растворе, что подтверждается большим количеством публикаций. Так, был выполнен целый ряд работ по исследованиям белков семейства нейротрофинов [A20-A28], далее будут приведены результаты некоторых из этих работ.

В работе [A21] рассматривался белок семейства нейротрофинов ProNGF (Nerve Growth Factor), который участвует в поддержании и росте определенных

популяций нейронов, как в центральной, так и в периферической нервной системе, и выполняет свою биологическую роль, связываясь с различными рецепторами (антителами). ProNGF состоит из компактной димерной части NGF (26 кДа) и 2*113 остатков гибкого Pro-домена (N-концевой части). Кристаллическая структура гомодимера NGF известна, тогда как о Pro-домене белка ProNGF структурная информация отсутствует. Анализ данных МУРР с помощью метода оптимизированного ансамбля показал, что компактная форма с обоими Nконцевыми доменами в компактной конфигурации должна быть преобладающим типом частиц в растворе, тогда как меньшая часть частиц proNGF имеет один Nконцевой домен в развернутой конформации (Рис. 1.17). Следует, однако, отметить, что термины "компактный" и "развернутый", используемые для по существу разупорядоченного пропептида, просто соотносят две конформации друг с другом, так как даже "компактная" конформация является гораздо более развернутой по сравнению с конформацией идеально свернутого глобулярного белка. Таким образом, было показано, что пропептидная область proNGF является подвижной и приводящей к различным конформациям белка, являясь по существу разупорядоченным доменом. Это свойство белка ProNGF представляет собой функцию, способствуя модуляции взаимодействия биологическую фактора роста нервов с мембраносвязанными рецепторами.

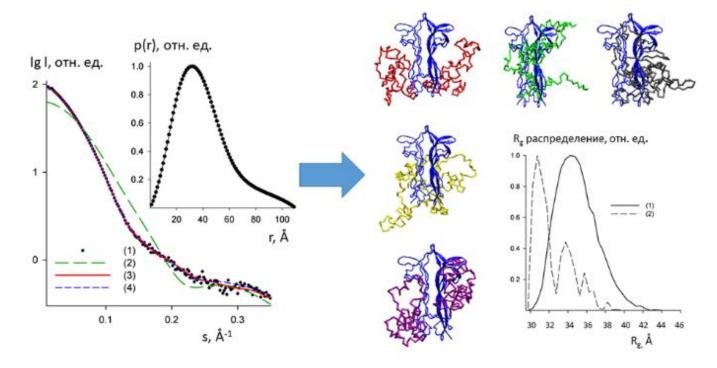


Рис. 1.17 Экспериментальные данные МУРР белка фактора роста нервов ProNGF (кривая 1), и наилучшие приближения, полученные программами DAMMIN (кривая 3) и ЕОМ (кривая 4), а также рассчитанная кривая от кристаллической модели димера белка NGF (известная С-концевая часть структуры ProNGF). С помощью программы ЕОМ были достроены неизвестные фрагменты Pro-домена белка ProNGF и получены типичные конформации. Также показаны распределения R_g начального случайного (кривая 1) и оптимизированного (кривая 2) ансамблей. Одна из цепей Pro-домена достаточно компактна, тогда как другая имеет более развернутую структуру.

Комбинированный эволюционный и структурный анализ семейства пронейротрофинов с помощью методов МУРР, дифференциальной сканирующей флуорометрии и биоинформатики был проведен в работе [A26]. В это семейство нейротрофинов (NTs) входят фактор роста нервов (NGF), нейротрофический фактор мозга (BDNF), нейротрофин 3 (NT3) и нейротрофин 4 (NT4). Данные белки обладают важными функциями для развития периферической и центральной нервной системы.

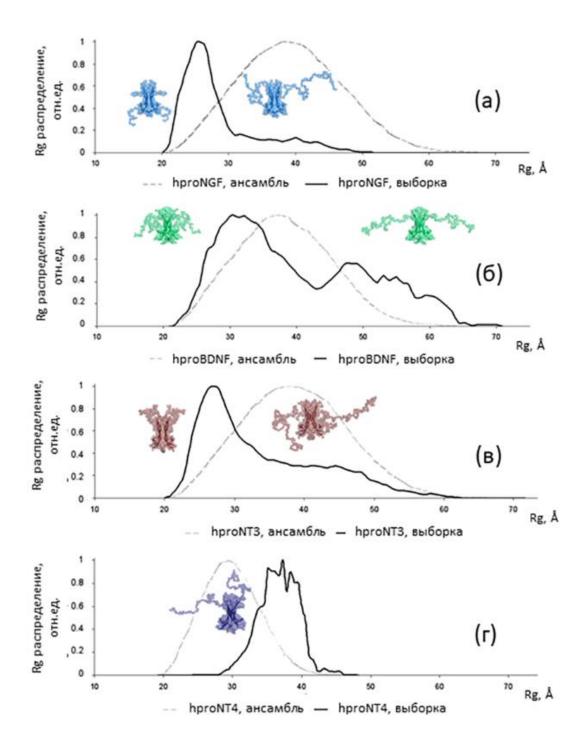


Рис. 1.18 Распределения по радиусам инерций R_g , полученные по данным МУРР для белков hproNGF (a), hproBDNF (б), hproNT3 (в) и hproNT4 (г). Распределения для начальных случайных и оптимизированных ансамблей показаны серыми пунктирными и черными сплошными линиями, соответственно. Рядом с распределениями показаны типичные конформации (полупрозрачные поверхности): компактные конформации слева и расширенные конформации справа. Димеры зрелых нейротрофинов показаны в центральной части структур (cartoons). Видно, hproNT4 обладает меньшей конформационной ЧТО гетерогенностью по сравнению с другими членами семейства нейротрофинов.

В результате анализа с помощью программы ЕОМ были найдены типичные конформации белков семейства нейротрофинов, вычислены их распределения по радиусам инерций R_g , и дана оценка их конформационной гетерогенности. Как видно из Рис. 1.18. белок hproNT4 обладает меньшей конформационной гетерогенностью по сравнению с другими членами семейства нейротрофинов. Полученные результаты помогли **ПОНЯТЬ** молекулярные механизмы, регулирующие биологические функции частично разупорядоченных нейротрофинов.

В работе [А27] был проведен анализ структуры белка нейротрофического фактора мозга человека proBDNF и его варианта с точечной мутацией V66M. Точечная мутация V66M в гене нейротрофического фактора мозга человека тесно связана с отклонениями в эпизодической памяти, уменьшением объема гиппокампа (части головного мозга, ответственного за эмоции и память) и усилением риска развития депрессии и тревожных расстройств у людей. Однако оставалось неясным, приводит ли точечная мутация V66M к структурным изменениям белка proBDNF. С помощью метода МУРР было показано, что, хотя наиболее вероятная конформация модифицированного белка более генетически оказывается компактной по сравнению с его природной формой, наличие мутации приводит к существенному увеличению гибкости молекул в растворе (Рис. 1.19).

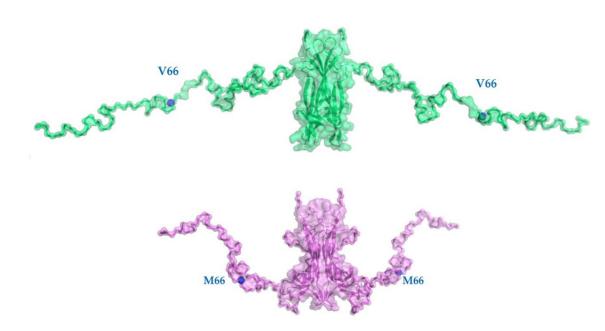


Рис. 1.19 Типичные модели нейротрофического фактора мозга человека hproBDNF: природная форма (зеленый цвет) и мутантый вариант V66M (фиолетовый цвет), полученные программой CORAL по данным МУРР. Положения аминокислотных остатков V66 (природная форма) и M66 (мутантная форма) изображены синими точками.

Проведенное структурное исследование позволило прояснить механизмы функционирования белка proBDNF и его патологического варианта, что важно для развития методов лечении психоневрологических расстройств.

Еще одним объектом, подходящим для исследования методом ЕОМ, является атаксин-3 – белок, участвующий в регуляции процессов внутриклеточной белков. преимущественно взаимодействует деградации Он других полиубиквитиновыми цепями, прикрепленными к белкам, доставляемым в убиквитин- и протеасомную систему. Атаксин-3 состоит из N-концевого Josephin домена и гибкой С-концевой части. который содержит два или три убиквитинвзаимодействующих мотива (UIMs) и полиглутаминовый тракт (polyQ), который, когда расширяется сверх порогового значения, приводит к агрегации и неправильному сворачиванию белка и вызывает спиноцеребеллярную атаксию типа 3. Структура домена Josephin с высоким разрешением была получена ранее, но структурная и динамическая гетерогенность атаксина-3 до сих пор препятствовала структурному описанию полноразмерного белка. В работе [А29]

были исследованы методом МУРР нерасширенные (нормальный, polyQ13) и расширенные (потенциально-патогенный, polyQ54) варианты атаксина-3 с точки зрения конформационных ансамблей, которые образуются этими белками в растворе (Рис. 1.20).

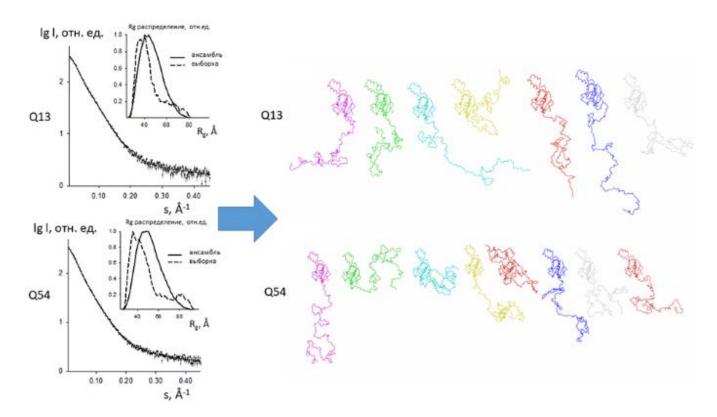


Рис. 1.20 Экспериментальные данные МУРР (точки) и кривые рассеяния, рассчитанные с помощью программы ЕОМ (сплошные линии) для нерасширенного и расширенного (потенциально-патогенного) вариантов белка Ataxin-3 (Q13 и Q54, соответственно). На вставке показаны распределения R_g для начальных случайных и оптимизированных ансамблей. Справа представлены типичные конформации белков, выбранные программой ЕОМ.

Из Рис. 1.20 видно, что как атаксин-3(Q13), так и атаксин-3(Q54) образуют широкий спектр конформаций в растворе с преобладанием удлиненных конформаций С-концевой части хвоста. Расширение полиглутаминового тракта (polyQ) выше патологического порога (что соответствует варианту Q54) не нарушает конформационный ансамбль атаксина-3. Эти результаты являются важной вехой на пути к детальному пониманию роли, которую конформационная пластичность атаксина-3 может играть в развитии заболеваний.

1.7. Оценка эффективности ab initio методов восстановления формы макромолекулярных комплексов с неоднородной электронной плотностью по данным МУРР/МУРН

Для систем случайно ориентированных одинаковых и невзаимодействующих между собой частиц (монодисперсные системы, разбавленные растворы биологических метода МУРР/МУРН макромолекул) с помощью онжом восстановить форму частиц с низким разрешением 1-2 нм ab initio, т.е. без какихлибо предварительных предположений о структуре частицы. При таком разрешении поиск обычно ограничивается однородными моделями. Из-за потери информации при сферическом усреднении однозначная реконструкция 3D формы частицы принципиально невозможна, и необходимо следить за тем, чтобы не переинтерпретировать данные МУРР/МУРН. Если ранее моделирование формы проводилось методом проб и ошибок путем вычисления профилей рассеяния от простых геометрических тел и сравнения их с экспериментальными данными, то позднее был разработан ряд алгоритмов, использующих различные представления формы - угловую огибающую функцию [46, 47, 98], шариковые модели [49, 50, 52, 54, 99] или, для белков, псевдо-аминокислотные остатки [53]. Детали этих алгоритмов были рассмотрены нами ранее в разделе 1.4 настоящей Главы. В указанных работах было показано, что при определенных обстоятельствах из данных МУРР/МУРН можно извлечь уникальную трехмерную форму частицы (с точностью до энантиоморфной конфигурации структуры, которая всегда дает одну и ту же кривую рассеяния; эта неоднозначность справедлива и для всех ab initio методов).

Отметим, во всех упомянутых подходах электронная плотность частиц предполагается однородной, что для белковых макромолекул или нуклеиновых кислот является вполне оправданным приближением. Но для ряда систем, например, комплексов "белок-нуклеиновая кислота" или "белок-липидная везикула", электронные плотности отдельных частей комплекса будут сильно

различаться между собой и здесь уже необходимо явным образом учитывать неоднородности в электронной плотности исследуемых частиц.

В настоящем разделе мы рассмотрим два достаточно общих *ab initio* метода, позволяющих проводить восстановление формы частиц с неоднородной электронной плотностью по данным МУРР/МУРН, подробно исследуем их возможности и ограничения и приведем примеры их практического использования. Рассмотрим частицу, состоящую из K компонентов с сильно различающимися электронными плотностями (например, для бинарного комплекса типа "белок-ДНК" или "белок-РНК" K=2, и этими компонентами будут белковые молекулы и нуклеиновые кислоты). Объем, достаточно большой, чтобы вместить эту частицу (например, сфера радиуса $R=D_{max}/2$), заполняется упакованными маленькими сферами радиуса $r_0 << R$ (называемыми шариками или "фиктивными атомами"). Каждому шарику присваивается индекс X_i , указывающий на компонент ("фазу"), которому она принадлежит (X_i варьируется от 0 (растворитель) до K). Поскольку шариков положения фиксированы, структура такой модели полностью фазовой "принадлежности" (конфигурации) X с вектором описывается размерностью $M \approx (R/r_0)^3$. Если шарики k-й фазы имеют контраст $\Delta \rho_k$, то интенсивность рассеяния составляет

$$I(s) = \left\langle \sum_{k=1}^{K} \Delta \rho_k A_k^2(\mathbf{s}) \right\rangle_{\Omega} = \sum_{k=1}^{K} \Delta \rho_k I_k(s) + 2 \sum_{n>k} \Delta \rho_k \Delta \rho_n I_{kn}(s)$$
 (1.23)

где $A_k(s)$ и $I_k(s)$ - амплитуда и интенсивность рассеяния, соответственно, от объема, занимаемого k-й фазой, а $I_{kn}(s)$ — перекрестные члены. Чтобы быстро рассчитать рассеяние от такой модели, используются сферические гармоники. Каждая трехмерная амплитуда рассеяния от отдельной фазы представляется в виде ряда

$$A(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm}(\mathbf{s}) Y_{lm}(\Omega) , \qquad (1.24)$$

Здесь сферические гармоники - это зависящие от угла функции $Y_{lm}(\Omega)$, определенные на поверхности единичной сферы, а $A_{lm}(s)$ — радиально-зависимые функции. Это представление, впервые введенное в методе МУРР в работах [46, 47], чрезвычайно полезно для описания изотропного рассеяния и будет широко

использоваться в дальнейшем. Самое главное, что использование сферических гармоник позволяет получить замкнутое выражение для сферически усредненной интенсивности, которая для многофазной модели записывается как

$$I(s) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left\{ \sum_{k=1}^{K} \left[\Delta \rho_k A_{lm}^{(k)}(s) \right]^2 + 2 \sum_{n>k} \Delta \rho_k A_{lm}^{(k)}(s) \Delta \rho_n \left[A_{lm}^{(n)}(s) \right]^* \right\}$$
(1.25)

где парциальные амплитуды

$$A_{\rm lm}^{(k)}(s) = i^{l} \sqrt{2/\pi} \ v_a \ \sum_{j} j_l({\rm sr}_j) Y_{\rm lm}^*(\omega_j)$$
 (1.26)

здесь сумма проходит по шарикам k-й фазы, $\mathbf{r}_j = (r_j \ \omega_j)$ - полярные координаты, $v_a = (4\pi r_0^3/3)/0.74$ - вытесненный объем, приходящийся на один шарик.

Если имеется набор $N_C \ge 1$ кривых рассеяния с вариацией контраста $I^{(j)}_{exp}(s)$, j=1,... N_C , то следует искать конфигурацию X, приближающую несколько кривых одновременно, т.е. минимизирующую общую невязку

$$\chi_{ov}^{2}(X) = \sum_{j=1}^{N_{c}} \chi_{j}^{2}$$
 (1.27)

куда входят индивидуальные невязки между экспериментальными и расчетными кривыми для каждой реализации контраста.

Модели шариков содержат тысячи элементов, и может быть найдено множество конфигураций, сопоставимых с экспериментальными данными. Для ограничения решений добавляется штрафной член P(X), обеспечивающий компактность и связность отдельных компонентов в полученной модели. Минимизируемая целевая функция принимает вид $f(X) = \chi^2 + \alpha P(X)$, где $\alpha > 0$ - штрафной вес, обеспечивающий надлежащий учет ограничения.

В случае большого количества переменных минимизация может быть выполнена только с помощью поиска типа Монте-Карло. Одним из наиболее подходящих методов является метод имитации отжига (SA – Simulated Annealing, [59]). Идея этого метода заключается в том, чтобы, начав поиск со случайного вектора X, произвести случайные модификации этого вектора X, как правило, переходя к конфигурациям, которые уменьшают f(X), но иногда допуская и те, которые увеличивают f(X). Вероятность принятия этого последнего типа модификации

уменьшается в процессе минимизации (система "охлаждается"). Вначале, когда "температура отжига" высокая, модификации происходят почти случайным образом, тогда как в конце процесса достигается конфигурация, соответствующая минимуму целевой функции.

На базе рассмотренного алгоритма, была создана программа MONSA [100], являющаяся многофазной версией программы DAMMIN [52]. В программе MONSA номер "фазы" шарика меняется при каждой модификации структуры, однако амплитуды, входящие в уравнения (3.1-3.2) только обновляются, но не полностью пересчитываются. Этот факт, вместе с использованием сферических гармоник, значительно ускоряет вычисления и позволяет выполнять процедуру метода имитации отжига, требующую миллионов оценок целевых функций, за разумное время (в зависимости от задачи и компьютера, от нескольких часов до нескольких дней).

Приведем несколько практических примеров использования программы MONSA для анализа данных МУРР/МУРН. Стоит отметить, что эксперименты МУРР на белковом комплексе, хотя и способны представить его общую форму, но не могут позволить различить молекулы, принадлежащие разным частям комплекса, что затрудняет определение их интерфейса связывания. Чтобы разрешить этот вопрос, часто используется вариация контраста образцов в экспериментах МУРН. Так, в работе [101] проводилось моделирование комплекса рецептора Меt с белком InlB321. Общая структура комплекса и способ связывания InlB321 к Met были получены из данных МУРН *ab initio*, т.е. без какой-либо информации о структуре. Для этого был взят набор данных, содержащий десять кривых МУРН, шесть от комплекса полноразмерного белка Меt с полностью (100%) дейтерированным белком InlB321, измеренных в растворах с 0, 35, 50, 60, 81 и 100% D₂O, и четыре кривые от комплекса полноразмерного Меt с частично (50%) дейтерированным InlB321 при 0, 50, 81 и 100% D_2O . В этом случае комплекс Met-InlB321 можно рассматривать как двухкомпонентную систему для ab initio реконструкции формы с помощью программы MONSA, которая была измерена при различных условиях

контраста. Действительно, специфическое дейтерирование InlB321 делает его контраст существенно отличным от контраста Меt, что позволяет использовать MONSA (фаза 1 = Met, фаза 2 = InlB321). В результате такого моделирования удалось получить воспроизводимые *ab initio* модели, точно соответствующие одновременно всем десяти кривым рассеяния (Рис. 1.21а). Более того, восстановленная форма, отображающая части Меt и InlB321 в их комплексе (рис. 1.21б), позволяет предположить, что InlB321 присоединяется к Меt в его серединной части. *Ab initio* модель также хорошо совпадает с моделью, полученной методом молекулярной тектоники программой SASREF с использованием моделей доменов Меt высокого разрешения [101]. Этот пример показывает, что метод МУРН может предоставить ценную информацию об общей форме и внутренней структуре сложных частиц, а также, что его комбинация с методами высокого разрешения еще больше повышает уровень структурной интерпретации данных малоуглового рассеяния.

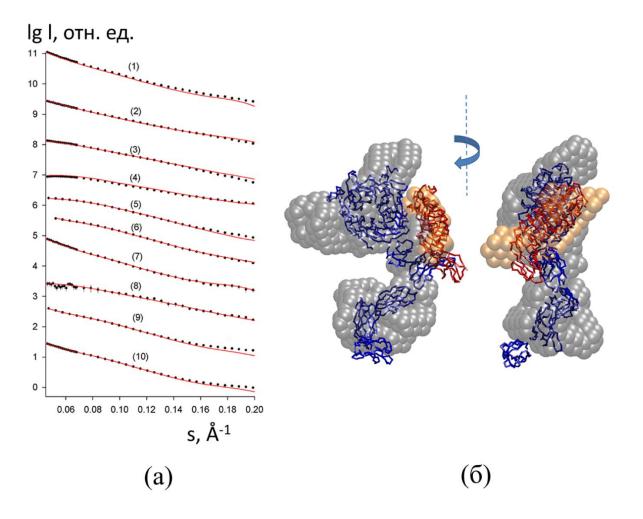
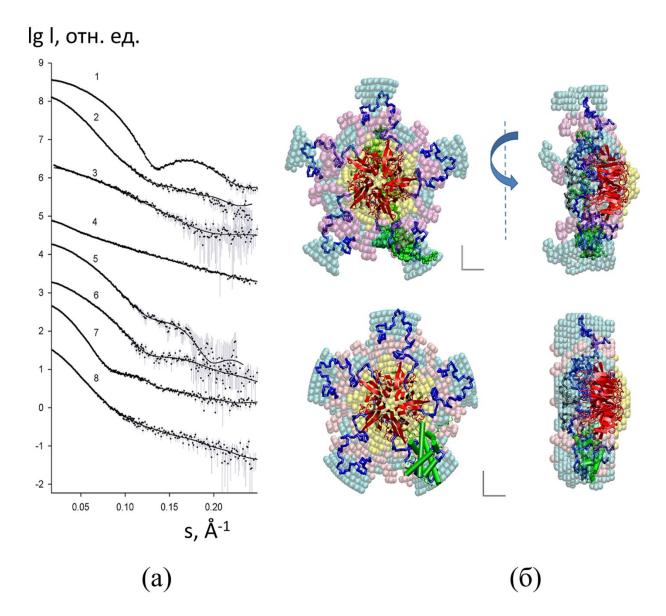


Рис. 1.21 (а) Метод вариации контраста в данных МУРН от комплекса InlB321/Меt. Экспериментальные данные обозначены точками; приближения с помощью программы MONSA представлены в виде красных сплошных линий. Кривые (1-6) соответствуют 100% дейтерированному InlB321 в комплексе с Меt, измеренному при 0, 35, 50, 60, 81 и 100% D2O, соответственно Кривые (7-10) соответствуют 50% дейтерированному InlB321 в комплексе с Меt, измеренному при 0, 50, 81 и 100% D2O, соответственно. (б) *Ab initio* модель комплекса Met/Inlb, полученная с помощью программы MONSA (серые полупрозрачные сферы соответствуют Меt, оранжевые полупрозрачные сферы изображают InlB321), совмещенная с молекулярной моделью комплекса, построенной с помощью программы SASREF [101]. Белок Меt показан в виде синей C_{α} -цепи (сверху: домен Sema, снизу: Ig домены), молекула InlB321 - в виде красной C_{α} -цепи.

Стоит отметить, что программа MONSA может быть применена не только для комплексов частиц с неоднородной электронной плотностью (типа "белокнуклеиновая кислота"), но и для восстановления формы многодоменных белков или белковых комплексов, когда имеется набор кривых как от полноразмерного белка (комплекса), так и от его отдельных доменов (частей комплекса). В этом

случае, можно обойтись только экспериментом МУРР без проведения вариации контраста и использования нейтронного рассеяния. В работе [А30] был исследован полноразмерный белок нуклеоплазмин (NP) и его комплексы с белками-гистонами Н2АН2В и Н5 (участвующими в регуляции процессов транскрипции, репликации и репарации). NP представляет собой гомопентамерный белок, причем каждый мономер состоит из ядра (120 N-концевых остатков) и хвостового (80 C-концевых остатков) домена [102]. Этот белок участвует в процессах ремоделирования хроматина и способствует сборке нуклеосом, модулируя взаимодействие между гистонами и ДНК. В результате моделирования с помощью программы MONSA были восстановлены трехфазные формы комплексов белков NP/H5 и NP/H2AH2B (Рис. 1.22) путем одновременного приближения трёх экспериментальных кривых рассеяния (от NP ядра, полноразмерного NP и комплекса NP/гистон). Общая структура NP не претерпевает существенных изменений при образовании комплекса с белками-гистонами, центральный домен ядра NP имеет компактную структуру, тогда как хвостовой NP домен простирается за пределы периферии центрального домена. Молекулы белков-гистонов взаимодействуют как с ядром NP, так и с хвостовым доменом NP. В то время как молекулы H2AH2B имеют компактную форму, молекулы Н5 выглядят более удлиненными. Основное различие между NP/H5 и NP/H2AH2B моделями удобно выразить в терминах межфазных областей, которые можно приблизительно оценить по числу атомовшариков, принадлежащих гистонной "фазе" и одновременно контактирующей с "фазой" нуклеоплазмина в моделях, полученных программой MONSA. Отношение интерфейсных областей (NP ядро/гистон) к (NP хвост/гистон) составляет 0.5÷0.8 для гистона H5 и 1.8÷2.3 для гистона H2AH2B. Это означает, что гистон H5 в значительно большей степени взаимодействует с хвостами NP, тогда как H2AH2B имеет более сильное связывание с поверхностью ядра NP. Таким образом, нами было показано, что компактный домен ядра NP и динамические гибкие хвосты NP формируют поверхность распознавания гистонов, причем хвосты NP особенно важны для взаимодействий с линкерными гистонами Н5, которые являются более

энергетически выгодными по сравнению с нуклеосомными гистонами H2AH2B. Это различие в распознавании гистонов может обеспечивать эффективный механизм регулирования динамического гистонового обмена.



МУРР Рис. 1.22 Данные NP/H5 NP/H2A-H2B. (a) OT комплексов Экспериментальные данные показаны точками, кривые рассеяния от *ab initio* моделей, рассчитанных программой MONSA, в виде непрерывных линий. NP ядро: кривая 1; полноразмерный NP: кривая 2; H2A-H2B димер: кривая 3; H5: кривая 4; комплекс NP-ядро/H2A-H2B: кривая 5; комплекс NP-ядро/H5: кривая 6; комплекс NP/H2A-H2B: кривая 7; и комплекс NP/H5: кривая 8. (б) Ab initio модели комплексов NP/H5 (сверху) и NP/H2A-H2B (снизу), полученные с помощью программы MONSA. Ядро и хвостовая часть NP представлены в виде желтой и фиолетовой сфер, а гистоны Н5 и Н2А-Н2В - в виде голубых шариков, соответственно. Обе ab initio модели наложены на структуру NP, восстановленную

программой BUNCH, где домен NP ядра (PDB ID: 1К5J) показан красным цветом. N-концевой сегмент и C-концевой хвост, добавленные программой BUNCH, показаны серым и синими C_{α} -цепями, соответственно. Зеленые структуры указывают на возможное расположение мономерного H5 и гетеродимерного H2A-H2B гистонов. Для наглядности представлена только одна модель гистона H5 и H2A-H2B, соответственно. Правые виды моделей повернуты против часовой стрелки вокруг вертикальной оси. Масштаб составляет 2 нм.

Еще одним интересным примером практического использования программы MONSA является определение структуры трехкомпонентного комплекса белков PEX5p/PEX14/PTS1 по данным МУРР [A31]. Белок-рецептор Pex5p распознает белки пероксисомного матрикса, которые содержат сигнальный пероксисомный белок PTS1. После стыковки Pex5p с PTS1 на мембране эти белки переносятся через мембрану, взаимодействуя с белком пероксисомной мембраны Рех14р. Исследование методом МУРР и статического светорассеяния структуры тройного предположения комплекса позволило сделать 0 Эксперименты по титрованию указали на образование трехкомпонентных комплексов PEX5p/PEX14/PTS1 со стехиометрией (1:6:1). Полноразмерный белок Рех5р содержит 639 аминокислотных остатков и содержится в растворе в мономерной форме, причем его N-концевая часть является удлиненной и частично разупорядоченной. Восстановление формы тройного комплекса программой MONSA выполнялось одновременного приближения путем ИТКП экспериментальных кривых МУРР (одна кривая от полноразмерного белка Рех5р, две кривые от двойных комплексов Pex5p/Pex14p и Pex5p/PTS1 и две кривые от тройных комплексов Pex5p/Pex14p/PTS1 и Pex5p(C-концевая часть)/Pex14p/PTS1). Каждый шарик модели мог принадлежать либо растворителю (индекс = 0) или одной из частей комплекса (index=1 соответствовал Pex5p (N-концевая часть, остатки 1-316), index=2 Pex5p C-концевой части (остатки 317-639), index=3 белку PTS1 и index=4 мембранным белкам Pex14p). Ожидаемые объемные фракции четырех фаз, соответствующие их теоретическим молекулярным массам, использовались в качестве ограничения при проведении минимизационной процедуры имитации отжига. На Рис. 1.23 представлена типичная модель тройного

комплекса, полученная программой MONSA. Как видно из Рис. 1.23, форма N-концевой части Pex5p в присутствии Pex14p и PTS1 остается удлиненной, даже когда молекулы Pex14p плотно связываются с N-концевой областью Pex5p. При этом C-терминальный домен Pex5p имеет компактную глобулярную форму и связан с PTS1 белком. Молекулярное моделирование с использованием известных структур высокого разрешения хорошо согласуется с *ab initio* моделью (Puc. 1.24). Таким образом, Pex5p, содержащий развернутый N-концевой домен, может связываться с другими белками или мембраной на относительно больших расстояниях. Поскольку Pex5p имеет несколько сайтов связывания в относительно широкой области, это повышает эффективность стыковки белков и их налипания на мембрану. Также можно предположить, что неструктурированная N-концевая часть Pex5p и ее удлиненная форма могут быть необходимы для взаимодействия комплекса Pex5p/Pex14p/PTS1 с другими субстратами, такими как липиды и мембранные белки, для образования функциональных мультипротеиновых комплексов.

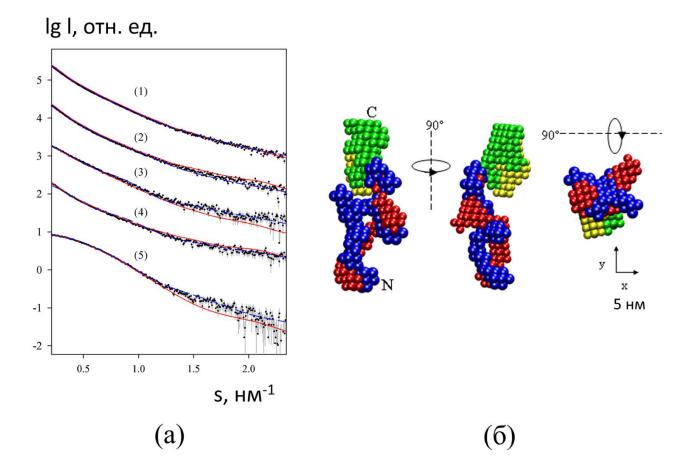


Рис. 1.23 (а) Данные МУРР от комплекса Pex5p/Pex14p/PTS1. Кривая 1, тройной комплекс Pex5p/Pex14p/PTS1 со стехиометрией (1:6:1). Кривые 2 и 3 представляют двойные комплексы Pex5p/Pex14p (1:6) и Pex5p/PTS1 (1:1), соответственно. Кривая 4 соответствует белку Pex5p, а кривая 5 — комплексу C-концевой части Pex5p(C)/Pex14p/PTS1 (соотношение 1:1:1). Приближения, полученные с помощью программ MONSA и SASREF, показаны сплошными (красной) и пунктирными (синими) линиями. Кривые смещены вдоль вертикальной оси для лучшей визуализации. (б) Аb initio модель комплекса Pex5p/Pex14p/PTS1, полученная программой MONSA. C-концевая часть Pex5p (PDB ID: 1FCH) показана в виде зеленых шариков. N-концевая часть Pex5p - в виде синих шариков. Пять белков Pex14p связанные с N-концевой частью Pex5p, показаны красными шариками, шестой Pex14p и белок PTS1 показаны желтыми шариками. Правый и нижний виды повернуты против часовой стрелки на 90° вокруг осей у и х, соответственно.

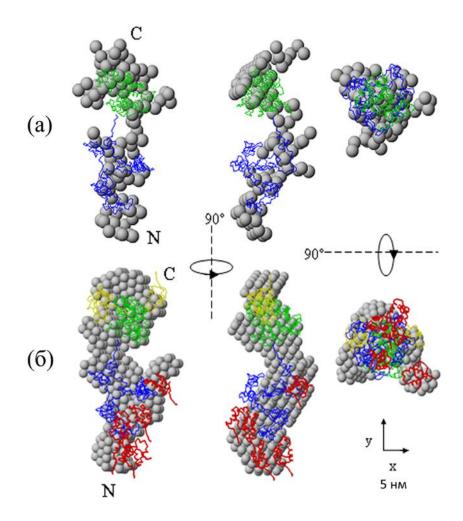


Рис. 1.24 Модели полноразмерного белка Pex5p (a) и комплекса Pex5p/Pex14p/PTS1 (стехиометрия 1:6:1) (б) по данным МУРР, полученные программами DAMMIN (серые сферы) и SASREF (C_{α} -цепи показаны в соответствии с цветовой схемой на Puc. 1.23.

Продемонстрированные возможности *ab initio* моделирования формы частиц с помощью представления структуры в виде многофазной шариковой модели говорят о большом потенциале данного подхода, но в то же время требуют в качестве необходимого условия моделирования наличия нескольких кривых рассеяния, полученных либо с помощью вариации контраста в нейтронных экспериментах, либо путем измерения помимо полноразмерных белков (или их комплексов) и неполных конструкций белков (или отдельных частей комплексов). Только в этом случае, информативность данных МУРР/МУРН будет достаточной для получения однозначных и воспроизводимых решений при *ab initio* восстановлении формы частиц с неоднородной электронной плотностью.

Однако, недавно был предложен новый итеративный алгоритм *ab initio* восстановления электронной плотности и формы частиц по данным МУРР [103], который как утверждал его разработчик, позволяет по одной кривой рассеяния проводить восстановление даже для систем с неоднородной электронной плотностью (например, комплексов типа "белок-нуклеиновая кислота"). Кратко воспроизведем его основные принципы.

В рассеянию, эксперименте ПО малоугловому регистрируется только интенсивность рассеяния, а вся информация о фазе теряется. Кроме того, в случае рассеяния от раствора частиц сферическое усреднение интенсивности происходит по мере того, как частицы меняют ориентацию относительно падающего пучка, что приводит к потере существенной доли информации. Таким образом, в отличие от итерационных алгоритмов решения "фазовой проблемы" (когда известны только амплитуды преобразования Фурье объекта, а фазы неизвестны, что типично для рентгеновской дифракции, томографии, электронной микроскопии), в случае данных МУРР необходимо решить 'проблему нахождения структурного фактора' частицы, когда и 3D амплитуды и фазы объекта неизвестны. Для того, чтобы рассчитать карту электронной плотности по данным малоуглового рассеяния от раствора частиц, кубический объем прямого пространства, содержащий частицу, перенесем на сетку, содержащую N × N × N объемных пикселей. Затем это пространство заполним электронной плотностью путем присвоения каждой точке сетки случайного значения. Амплитуды рассеяния, соответствующие этой случайной карте электронной плотности, рассчитываются с помощью дискретного, быстрого преобразования Фурье (Fast Fourier Transformation - FFT). Затем трехмерные интенсивности вычисляются как амплитуды рассеяния, умноженные на их комплексно-сопряженные величины. Чтобы наложить ограничение на экспериментальную 1D интенсивность рассеяния, 3D интенсивность рассеяния сводится к 1D профилю путем сферического усреднения. Каждой точке сетки обратного пространства присваивается 'бин' (ячейка) в соответствии с модулем его вектора рассеяния s, в результате чего получается серия концентрических sоболочек. Одномерное сферическое среднее значение интенсивности рассчитывается путем усреднения интенсивностей всех точек сетки, соответствующих одной s-оболочке (Рис. 1.25).

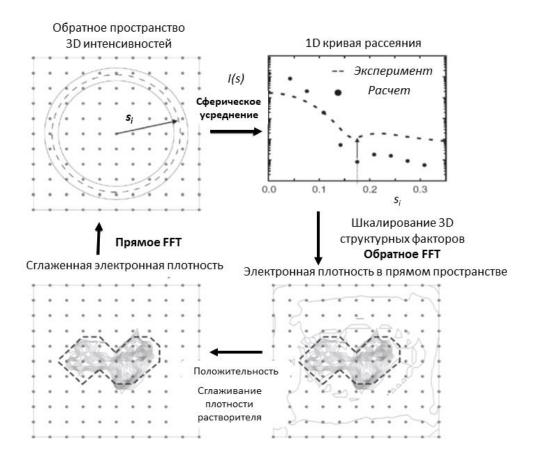


Рис. 1.25 Итеративный алгоритм восстановления электронной плотности частицы с использованием данных МУРР.

Затем экспериментальные 1D кривые рассеяния совмещаются с рассчитанными интенсивностями путем шкалирующих коэффициентов, рассчитываемых для всех точек сетки в *s*-оболочке как квадратный корень отношения экспериментальной 1D интенсивности к рассчитанной 1D интенсивности (Рис. 1.25). По сути, это есть операция шкалирования амплитуд при сохранении фаз. Далее рассчитывается обратное преобразование Фурье коэффициентов трехмерной структуры для получения новой карты электронной плотности. Полученная карта электронной плотности имеет соответствующий профиль рассеяния, который точно соответствует профилю рассеяния экспериментальных данных. Однако эта карта

электронной плотности вряд ли отражает фактическую электронную плотность объекта, поскольку сферическое усреднение в обратном пространстве приводит к потере значительного количества информации, необходимого для однозначного определения электронной плотности частицы. Вследствие этого, существует множество функций электронной плотности, которые дают одинаковые одномерные кривые рассеяния. К счастью, многие из этих функций электронной плотности не согласуются с известными свойствами реальных частиц; например, электронная плотность является положительной и вещественной величиной, а объемная плотность растворителя должна быть константной величиной. Эти ограничения накладываются итеративно, определяя так называемую область "поддержки", т.е. область пространства, где электронная плотность положительна внутри и равна нулю снаружи этой области [104]. Область "поддержки" определяется с помощью алгоритма сжатия [105], где электронная плотность свертывается с гауссовой функцией, а область "поддержки" определяется как точки сетки, чья электронная плотность превышает минимальный порог. Цикл наложения ограничений в прямом и обратном пространствах итеративно повторяется до тех пор, пока расчетный профиль рассеяния не сойдется максимально возможно с экспериментальными данными МУРР. Данный алгоритм был реализован в виде программы DENSS (DENsity from Solution Scattering), находящейся в свободном доступе на сайте https://github.com/tdgrant1/denss.

О возможностях и ограничениях рассмотренного выше алгоритма пойдет речь ниже. Нами было проведено систематическое и детальное исследование работы программы DENSS на ряде многокомпонентных систем с неоднородной электронной плотностью и показано, что этот подход не способен реконструировать внутреннюю структуру таких систем, поэтому пользователям данного алгоритма надо быть осторожными при интерпретации результатов программы DENSS [A32].

В работе [103] рассмотренный выше итеративный алгоритм был проверен на ряде различных биологических макромолекул, для которых экспериментальные данные МУРР были взяты из базы данных SASBDB. Результаты восстановлений приведены на Рис. 1.26, все эти образцы представляют собой белковые молекулы разного размера и формы. Во всех случаях рассчитанные карты электронной плотности близко соответствовали известным атомным структурам, полученным с помощью рентгеновской кристаллографии (Рис. 1.26), что говорит о достаточной эффективности алгоритма, сравнимой с ранее развитыми подходами в программах DAMMIN, DAMMIF, GASBOR и др.

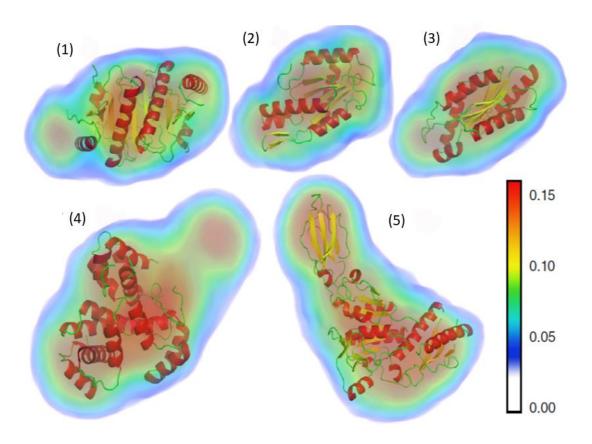


Рис. 1.26 Реконструкции электронной плотности белковых макромолекул по данным МУРР с помощью итеративного алгоритма программой DENSS [103]. Электронные плотности показаны в виде объемов, окрашенных в соответствии с плотностью (цветовая полоса указывает значения электронной плотности в е/ų). Рентгеновские кристаллические структуры показаны в формате Cartoon. Обозначения для белков: (1) HIT гидролаза (PDB ID: 3I24), (2) EAL/GGDEF домен (PDB ID: 3ICL), (3) Дигуанилатциклаза (PDB ID: 3IGN), (4) MmoQ регулятор (PDB ID: 3LJX), (5) белок Nmul A1745 (PDB ID: 3LMF).

Однако представленный дополнительный пример для комплекса белка эндофилина A1-BAR с двумя липидными мицеллами арахидонил-CoA (код SASDAX5 из базы данных SASBDB), который должен был продемонстрировать работу алгоритма для восстановления формы частиц с неоднородной электронной плотностью (Рис. 1.27а), при более внимательном рассмотрении оказался неподходящим примером многокомпонентной системы, так как электронный контраст головных групп липидов намного выше, чем у белковой части. Сигнал от двух мицелл СоА доминирует по интенсивности по сравнению с сигналом от белка эндофилина (в пятнадцать раз выше при рассеянии в нулевой угол), и поэтому модель хорошо аппроксимируется формой полой гантели. Действительно, в работе [А32] нами была рассчитана теоретическая кривая МУРР от структуры SASDAX5 с помощью программы CRYSOL, и оказалось, что ab initio формы, восстановленные по этой кривой рассеяния в стандартном однородном приближении с помощью программы DAMMIN, хорошо воспроизводят полую гантель (Рис. 1.276). Таким образом, было показано, что мицеллы CoA не являются подходящим примером многокомпонентной системы.

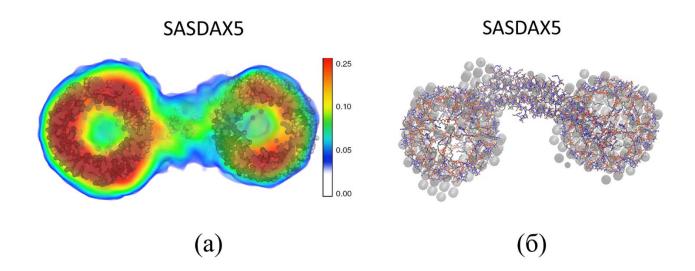


Рис. 1.27 (а) Реконструкция трехмерной карты электронной плотности A1 BAR домена белка эндофилина A Bar1 в комплексе с двумя арахидонил-CoA липидными мицеллами (SASBDB ID: SASDAX5). по данным МУРР (SASBDB ID: SASDAX5) с помощью итеративного алгоритма программой DENSS. Карты электронной плотности представлены в виде объемов, окрашенных в соответствии с плотностью

(цветовая полоса указывает на электронную плотность значения в $e/Å^3$). (б) Ab *initio* реконструкция формы данного комплекса с помощью программы DAMMIN. Реконструированная модель была получена на основе кривой рассеяния, рассчитанной от атомной модели (SASDAX5) программой CRYSOL. Для демонстрации полого вида восстановленной формы показан срез *ab initio* модели. Атомная модель комплекса наложена на *ab initio* модель с помощью программы SUPALM.

Важно заметить, что естественным и практически важным тестом способности реконструировать многокомпонентные системы является использование данных рассеяния от комплексов типа "белок-нуклеиновая кислота". В водных растворах рентгеновский контраст ДНК (~0,22 e/Å³) примерно в два раза выше, чем у белков $(\sim 0,10 \text{ e/Å}^3)$, т.е. такие комплексы являются классическими двухкомпонентными частицами. Для проверки работы алгоритма программы DENSS нами были взяты несколько моделей "белок-ДНК" из белковой базы данных (PDB, www.wwpdb.org) и рассчитаны теоретические кривые МУРР по этим моделям с помощью программы CRYSOL. Далее по этим данным мы восстановили 3D электронные плотности с помощью программы DENSS и наложили полученные электронные плотности на атомные модели с помощью программы SUPALM (Рис. 1.28a). В то время как общие формы частиц напоминают кристаллические структуры, ни одна из восстановленных карт не имеет более высокой плотности в месте расположения молекул ДНК. Вместо этого, независимо от положения ДНК, части с более высокой плотностью на усредненных картах всегда расположены вблизи их геометрических центров.

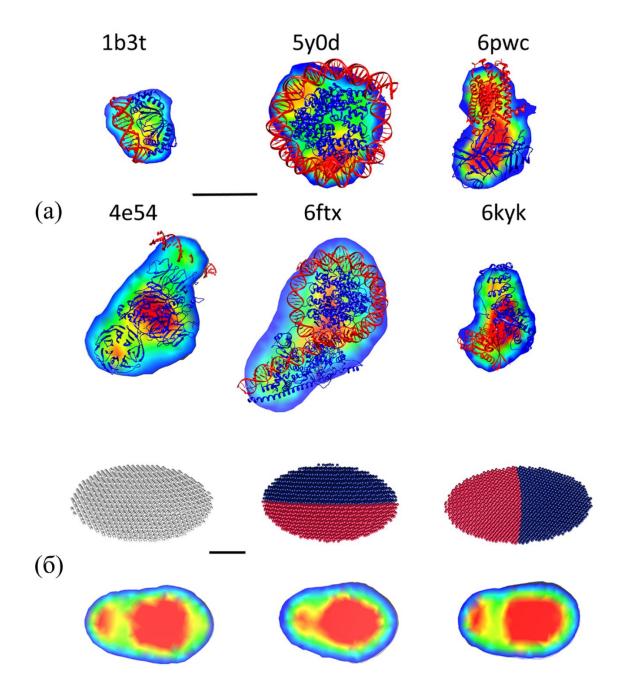


Рис. 1.28. Реконструкции электронной плотности программой DENSS по теоретическим данным МУРР, рассчитанные до номинального разрешения 12 Å. Усреднение проводилось по 100 реконструкциям с помощью скрипта 'superdenss' [А32]. Двумерные срезы окрашены от красного (более высокая плотность, относительное значение 20) до синего (более низкая плотность, значение 5), масштаб - 50 Å. (а) Комплексы белок-ДНК (PDB ID: 1B3T, 4E54, 5Y0D и 6FTX) и две модели белков с (PDB ID: 6PWC (А и R цепи) и 6KYK (А и D цепи)) с искусственно удвоенным контрастом одного домена (R и D цепи, соответственно). Кристаллические структуры представлены в Сагtооп представлении (белок показан синим цветом; ДНК и удвоенные домены — красным цветом). (б) Эллипсоиды

вращения: верхний ряд, модельные частицы (слева - однородный эллипсоид; в середине и справа - неоднородные эллипсоиды с плотностями 1 (синий) и 2 (красный)). Нижний ряд, восстановленные программой DENSS электронные плотности.

Следует отметить, что попытки дальнейшего увеличения контраста ДНК, например, путем искусственного "удвоения" ДНК для структуры 5Y0D (Рис. 1.28а), привели бы к получению модели, напоминающей полый диск, где рассеяние от комплекса полностью определяется сигналом от его ДНК части. Подобно мицеллам СоА, упомянутым ранее, такая полая форма восстанавливается как с помощью программы DENSS, так и в стандартном однородном приближении программой DAMMIN.

Чтобы избежать спекулятивных предположений о различиях в упаковке, компактности и т.д. между белками и ДНК, нами также было рассчитано малоугловое рассеяние от нескольких двухдоменных моделей белков, где один домен дважды повторялся в структуре, т.е. имел вдвое больший контраст по сравнению с другим белком. Два типичных восстановления программой DENSS с использованием таких данных представлены на Рис. 1.28a (PDB ID: 6PWC и 6KYK), и снова видно, что повышенные плотности находятся ближе к центрам частиц, а не к расположению повторяющихся доменов.

Далее нами была проверена способность программы DENSS выявлять еще более четкие неоднородности. В качестве тестового примера было взято рассеяние от трех частиц, имеющих одинаковую простую форму (эллипсоиды вращения с соотношением полуосей 1:2, Рис. 1.28б). В одном случае эллипсоид был однородным. В другом случае он был разделен на две половины с разными плотностями (равными 1.0 и 2.0, соответственно), разделенные вдоль длинной оси эллипсоида. В третьем случае разделение частей с разной плотностью происходило вдоль короткой оси эллипсоида. Теоретическое рассеяние от таких шариковых моделей было рассчитано программы MONSA, \mathbf{c} помощью карты, восстановленные программой DENSS по этим кривым, представлены на Рис. 1.286, нижний ряд. Примечательно, что эти карты показывают схожие распределения

плотности, однако при этом не имеют ничего общего с истинными неоднородностями. Во всех случаях более высокие плотности находятся в центре, и результаты аналогичные тем, что представлены на Рис. 1.28а.

При анализе упомянутых выше данных мы обнаружили еще один поразительный факт, а именно, что рассеяние, рассчитанное по усредненным картам (по результатам 100 запусков DENSS), не соответствует кривой рассеяния от исходных моделей. Стоит отметить, что в статье по алгоритму DENSS [103] действительно не было представлено никаких приближений, рассчитанных по результирующим усредненным моделям электронной плотности. Очевидно, что все индивидуальные реконструкции программы DENSS (не имеющие физического смысла вследствие своей разрывности) хорошо соответствуют данным МУРР, однако усредненные модели им уже не соответствуют, и выявленные нами расхождения с данными МУРР являются существенными, для всех случаев, которые представлены на Рис. 1.29 и 1.30.

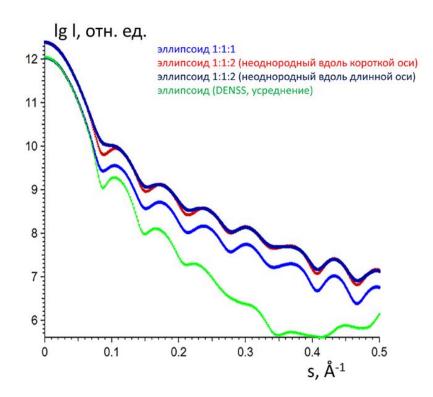


Рис. 1.29 Теоретические кривые МУРР от трех эллипсоидов (изображенных на Рис. 1.28б) (1 - однородный (синий); 2 - неоднородный вдоль короткой оси (красный); 3 - неоднородный вдоль длинной оси (синий)), рассчитанные MONSA. Кривая

МУРР, рассчитанная по восстановленной программой DENSS электронной плотности (с усреднением по 100 запускам программы), показана зеленым цветом.

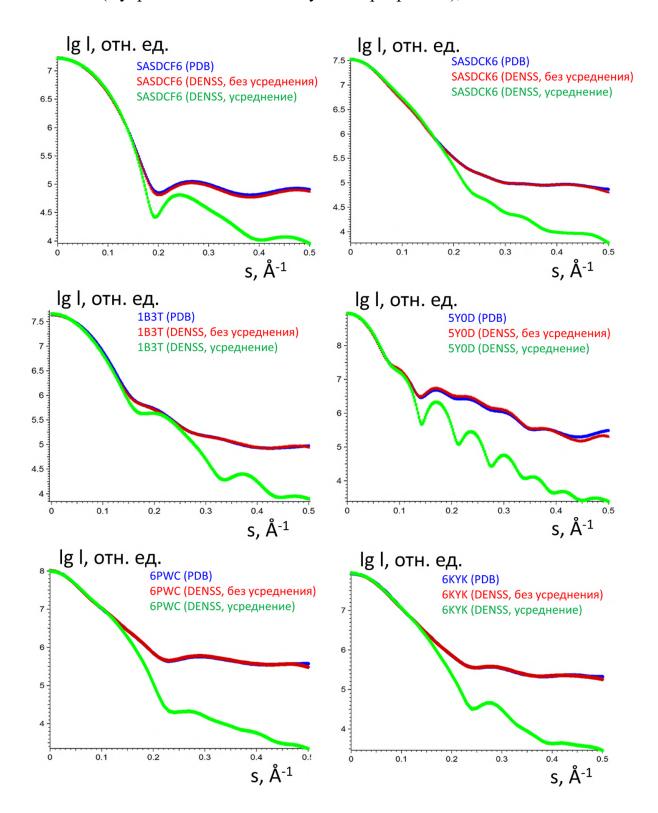


Рис. 1.30 Теоретические кривые МУРР, рассчитанные от атомных моделей (синий) и кривые, рассчитанные от восстановленных карт электронной плотности программой DENSS: красный - индивидуальное восстановление, зеленый -

усредненная модель. Верхний ряд: два примера из оригинальной статьи Т.D. Grant [103] (SASBDB ID: SASDCF6 и SASDCK6; усреднение проводилось по 20 реконструкциям DENSS-а); средний ряд: комплексы белок-ДНК (PDB ID: 1В3Т и 5Y0D); нижний ряд: комплексы белков с PDB ID: 6PWC (цепи А и удвоенная R) и 6KYK (цепи А и удвоенная D), усреднения проведены по 100 реконструкциям DENSS-а). Рассеяние от электронных карт плотности рассчитывалось с помощью программы EM2DAM из пакета ATSAS.

Мы считаем, что такие неоднозначные результаты связаны с ключевой процедурой, используемой в методе DENSS, а именно, с усреднением нескольких слабо ограниченных 3D плотностей прямого пространства. Эта процедура обычно дает повышенную плотность в центральной части усредняемой карты, одновременно приводя к несоответствию данным в обратном пространстве. Эти эффекты похожи на низкочастотную фильтрацию плотности частиц с учетом не только когерентно рассеивающего объема самой частицы, но и некогерентной части, обусловленной окружающим растворителем (например, с помощью программ типа EMAN). Даже для правильных структур такая фильтрация всегда приводит к искусственно повышенной плотности в центре частицы, что приводит к последующему несоответствию данным МУРР.

В этой необходимо связи. отметить, что усреднение ПО нескольким восстановленным формам частиц является стандартной процедурой при ab initio анализе данных МУРР и действительно необходимо для фильтрации выбросов (случайных единичных восстановлений, сильно отличающихся от основной массы), оценки неоднозначности восстановления и построения согласующихся моделей. При этом, усреднение не является обязательным этапом постобработки для таких методов, как DAMMIN/DAMMIF/MONSA, где все модели физически обоснованы. Однако, метод DENSS в огромной степени полагается на процедуру усреднения, и повышенная плотность вблизи центра карты является следствием этого факта.

Таким образом нами было проведено систематическое исследование эффективности работы итеративного алгоритма, реализованного в программе DENSS, при восстановлении формы частиц с неоднородной электронной

плотностью и показано, что данный алгоритм не позволяет корректно восстанавливать внутреннюю структуры таких систем, причем систематически повышенная плотность в центральной части является артефактом в подавляющем числе случаев, а кривые рассеяния, рассчитанные от усредненных моделей, имеют существенные расхождения с данными МУРР.

Поэтому, можно заключить, что для однородных систем можно использовать как стандартные алгоритмы в представлении структур плотноупакованными шариками (DAMMIN, DAMMIF), так и недавно предложенный итеративный алгоритм DENSS. Все указанные алгоритмы позволяют приблизительно одинаково хорошо и качественно восстанавливать форму частиц. Однако, в случае частиц с неоднородной электронной плотностью, восстановления программой DENSS могут приводить к серьезным артефактам, и поэтому более надежно в этом случае использовать многофазный *ab initio* подход, реализованный в программе MONSA. В последнем случае при анализе данных МУРР/МУРН нужно использовать не единичную кривую рассеяния, а набор кривых (используя метод вариации контраста или проведение измерений для неполных конструкций белковых комплексов).

Заключение к Главе 1

В настоящей главе были кратко рассмотрены теоретические основы метода МУРР/МУРН, типы экспериментальных установок и принципы проведения измерений, а также имеющиеся к настоящему времени подходы анализа и интерпретации данных. Было показано, что метод МУРР/МУРН позволяет восстанавливать трехмерные структуры низкого разрешения, используя *ab initio* моделирование, молекулярную тектонику и гибридные подходы, а также оценивать олигомерное состояние белков и белковых комплексов, и кроме того является мощным инструментом для количественного анализа гибких систем.

Также был проведен краткий обзор основных *ab initio* методов восстановления 3D формы частиц для систем с неоднородной электронной плотностью по данным МУРР/МУРН, которые включают в себя многофазное моделирование в виде частицы плотноупакованными представления структуры шариками большого набора использованием максимально кривых рассеяния полноразмерной частицы и ее доступных фрагментов, а также итеративный алгоритм восстановления 3D электронной плотности по данным МУРР, для которого достаточно использования одной кривой рассеяния. Были детально рассмотрены возможности и ограничения данных подходов на ряде практических задач.

Глава 2

Апостериорная оценка оптимального углового диапазона экспериментальных данных МУРР/МУРН с использованием теоремы Котельникова-Шеннона

Важной характеристикой метода МУРР является контраст, который физически возникает из-за разницы в электронной плотности между образцом и буферным раствором. Поэтому для малоуглового эксперимента необходимо проводить раздельные измерения образца и буфера. При этом надо иметь ввиду, что при каждом измерении все, что находится на пути рентгеновского пучка, вносит свой вклад в рассеяние, регистрируемое на детекторе. Это включает в себя, в частности, рассеяние от окошек ячейки для образца и самих исследуемых макромолекул. Большинство биологических образцов содержат не только макромолекулы, но и стабилизирующие вещества в буфере. Поэтому измеряют не только сам образец (макромолекулы в матричном растворе), но и сам матричный раствор, и чтобы получить чистый сигнал от образца, последний вычитается из первого. В эксперименте МУРР контраст определяется как разница в электронной плотности, а она у воды (334 е/нм³) не намного ниже, чем у белков (420 е/нм³). Нуклеиновые кислоты имеют более высокую электронную плотность - около 550 е/нм³. Чтобы контраст не оказался слишком маленьким, по возможности избегают богатых электронами веществ, таких как сульфаты, фосфаты и ионы тяжелых металлов или, по крайней мере, используют их в очень низких концентрациях. Метод вариации контраста также широко используется в экспериментах МУРН, где исследуется разность плотности рассеяния между растворителями Н2О и D2O, а также специально дейтерированными макромолекулами.

Таким образом, для биологических систем контрастность частиц в водных растворах довольно мала, и полезный сигнал может быть слабым по сравнению с фоновым [106]. Это приводит к низкому отношению сигнал/шум в данных МУРР, особенно в области больших углов рассеяния. В связи с этим возникает вопрос, как можно определить полезный угловой диапазон экспериментальных данных

малоуглового рассеяния, который следует выбрать для последующей интерпретации и построения моделей. Ранее обычной практикой являлось использование только той части кривой рассеяния, где отношение сигнал/шум превышает определенный порог [107], но выбор такого порога остается довольно субъективной процедурой. Кроме того, выбор углового диапазона, базирующийся только на отношении сигнал/шум, не учитывает шаг угловой сетки данных МУРР/МУРН, от которого зависит информативность данных.

Проблема оценки полезного углового диапазона данных актуальна и для других дифракционных методов, например, для рентгеновской кристаллографии. Принятые критерии качества и точности данных включают как отношение сигнал/шум интенсивности в координационной сфере с наивысшим разрешением ($\langle I/\sigma(I)\rangle$), так и функцию разброса эквивалентных отражений (R_{merge}) [108]. Однако, в анализе данных МУРР до недавнего времени не существовало такого рода согласованных критериев, а с учетом последних изменений в направлении обеспечения публичного доступа к экспериментальным данным МУРР и соответствующим моделям [109, 110], отсутствие объективного метода оценки полезного углового диапазона в наборе данных МУРР представляется серьезным упущением.

В свете изложенных выше причин, нами был разработан новый подход, использующий теорему Котельникова-Шеннона [111], для определения полезного углового диапазона в экспериментальных данных рассеяния от разбавленной монодисперсной системы. Ниже будут приведены основные детали данного алгоритма и показана эффективность его работы на ряде теоретических и экспериментальных данных МУРР/МУРН.

2.1 Автоматизированный алгоритм оценки оптимального углового диапазона данных МУРР/МУРН

Интенсивность рассеяния I(s) от разбавленного раствора одинаковых частиц (например, монодисперсного раствора макромолекул) связана с функцией распределения расстояний p(r) в прямом пространстве как

$$I(s) = 4\pi \int_{0}^{D_{\text{max}}} p(r) \frac{\sin sr}{sr} dr$$
 (2.1)

где $s=4\pi \, sin(\theta)/\lambda,\, 2\theta$ - угол рассеяния, а λ - длина волны излучения. Уравнение (2.1) учитывает тот факт, что функция p(r) определена на ограниченном интервале и равна нулю для всех $r>D_{max}$ (где D_{max} - максимальный размер частицы). Если I(s) известна, то p(r) может быть вычислена с помощью обратного преобразования

$$p(r) = \frac{r}{2\pi^2} \int_0^\infty sI(s)\sin srds \tag{2.2}$$

Из уравнений (2.1) и (2.2) легко заметить, что функции sI(s) и p(r)/r являются синус-преобразованиями Фурье друг для друга. Далее p(r) удобно представить в виде ряда Фурье по синусам:

$$p(r) = \frac{r}{2\pi^2} \sum_{n=1}^{\infty} a_n \sin(\pi r n / D_{\text{max}})$$
 (2.3)

где n - целое число. Подстановка уравнения (2.3) в (2.1) дает интерполяционную формулу Шеннона [111]:

$$U(s) = sI(s) = \sum_{n=1}^{\infty} s_n a_n \left[\frac{\sin D_{\max}(s - s_n)}{D_{\max}(s - s_n)} - \frac{\sin D_{\max}(s + s_n)}{D_{\max}(s + s_n)} \right]$$
(2.4)

где $s_n = \pi n/D_{max}$ — так называемые положения шенноновских каналов.

Уравнение (2.4) содержит, вообще говоря, бесконечное число шенноновских каналов. Однако для экспериментальных данных МУРР, измеренных в ограниченном диапазоне векторов рассеяния ($s < s_{max}$), вклад каналов за пределами этого диапазона (т.е. с индексами $n > s_{max}D_{max}/\pi$) в приближение интенсивности в этом диапазоне будет относительно мал. Поэтому число шенноновских каналов в

измеренном диапазоне $N_S = \mathrm{s}_{\mathrm{max}} D_{\mathrm{max}}/\pi$ было предложено использовать в качестве оценки информационного содержания в данных рассеяния [112, 113]. Также были предложены методы расчета функции p(r) [114] и оценки соответствия экспериментальным данным [115, 116] на основе представления Шеннона.

Хотя большие значения N_S обычно указывают на большее информационное содержание, очевидно, что сама по себе эта величина не может дать окончательную оценку из-за того, что в данном случае никак не учитывается отношение сигнал/шум. Кроме того, данные МУРР обычно измеряются на плотной угловой сетке, т.е. с угловым приращением Δs , значительно меньшим, чем расстояние между шенноновскими каналами π/D_{max} . Таким образом, количество информации в данных МУРР должно быть связано как с уровнем экспериментальных ошибок интенсивности, так и со степенью плотности угловой сетки данных МУРР.

Когда индекс суммирования в уравнениях (2.3) и (2.4) ограничен целым числом (M), соответствующие усеченные шенноновские приближения обозначим как $p_M(r)$ и $U_M(s)$, соответственно. Для набора экспериментальных данных МУРР/МУРН можно построить его усеченное приближение $U_M(s)$ с использованием M шенноновских каналов путем минимизации невязки

$$\chi^{2}(M) = \sum_{i=1}^{N} \frac{1}{2s_{i}^{2} \sigma_{i}^{2}} [s_{i} I(s_{i}) - U_{M}(s_{i})]^{2}$$
(2.5)

где индекс суммирования i проходит по N экспериментальным точкам, а σ_i^2 - стандартное отклонение для измеренной интенсивности в точке s_i . Наилучшее решение методом наименьших квадратов должно удовлетворять условию $\delta \chi^2/\delta a_m = 0$, что приводит к системе нормальных уравнений:

$$\sum_{i=1}^{N} \frac{1}{s_{i}^{2} \sigma_{i}^{2}} \left[s_{i} I(s_{i}) B(m,i) \right] = \sum_{i=1}^{N} \sum_{n=1}^{M} \frac{a_{n}}{s_{i}^{2} \sigma_{i}^{2}} \left[B(m,i) B(n,i) \right]$$
(2.6)

где
$$B(n,i) = \left[\frac{\sin D_{\max}\left(s_i - s_n\right)}{D_{\max}\left(s_i - s_n\right)} - \frac{\sin D_{\max}\left(s_i + s_n\right)}{D_{\max}\left(s_i + s_n\right)}\right].$$

Для экспериментов МУРР/МУРН от растворов макромолекул экспериментальные данные $I(s_i)$ представляют собой разницу между рассеянием от растворенного

вещества и чистого растворителя и могут иметь отрицательные значения из-за экспериментальных ошибок. Эти отрицательные значения должны войти в уравнения (2.5-2.6), однако поскольку рассчитанная теоретическая интенсивность МУРР $U_M(s_i)$ всегда должна быть неотрицательной, то система уравнений (2.6) может быть решена стандартными методами при ограничении на неотрицательность значений коэффициентов a_n [117].

Усеченное приближение шенноновское дает возможность оценить информативность и полезный угловой диапазон в наборе экспериментальных данных MУРР/MУРН. Действительно, при слишком малом M эта аппроксимация Шеннона не будет иметь достаточного количества членов (слагаемых) для соответствия экспериментальным данным. С ростом М приближение будет улучшаться, но на определенном этапе будет наблюдаться плато, когда значения a_n , определенные из системы уравнений (2.6) уже не будут существенно улучшать невязку, определяемую по экспериментальным данным. Поэтому должно существовать оптимальное (эффективное) значение шенноновских каналов M_S , отражающее информативность данных, а полезный угловой диапазон данного набора экспериментальных данных МУРР/МУРН будет определяться как $\pi M_S/D_{max}$. Заметим, что M_S не обязательно должно совпасть с N_S , и далее будет представлена процедура надежного автоматизированного определения эффективного числа шенноновских каналов M_S .

Но прежде покажем, как на усеченное шенноновское приближение влияют шум и плотность точек угловой сетки данных МУРР, для этого нами был смоделирован ряд кривых малоуглового рассеяния от различных геометрических тел (Таблица 2.1). Данные были рассчитаны при фиксированном значении момента переноса импульса (вектора рассеяния) s_{max} =4 нм⁻¹ и содержали различное число шенноновских каналов для разных тел из-за их различных размеров. Плотная сетка с угловым шагом Δ s=0.0025 нм⁻¹ использовалась для имитации типичного набора данных МУРР синхротронного рентгеновского излучения, а более разреженная сетка с Δ s=0.042 нм⁻¹ (имеющая примерно в 17 раз меньше точек в том же угловом

диапазоне) имитировала данные МУРН. Для каждой точки интенсивности добавлялся случайный гауссовский шум, относительная ошибка моделируемого шума варьировалась от 1% до 400% для различных наборов данных.

Для каждого смоделированного набора данных вычислялись усеченные шенноновские приближения $U_M(s)$ для различных чисел М в соответствии с уравнениями (2.4-2.6), и качество приближения оценивалось по коэффициенту R_M между идеальной теоретической кривой без шума $I_{ref}(s)$ и соответствующим приближением Шеннона $U_M(s)/s$ по следующей формуле:

$$R_{M} = \sum_{i=1}^{N} s_{i} \left(U_{M}(s_{i}) / s_{i} - I_{ref}(s_{i}) \right) / \sum_{i=1}^{N} s_{i} I_{ref}(s_{i})$$
(2.7)

Смоделированные наборы данных и лучшие шенноновские приближения (соответствующие минимальным *R*-факторам) показаны на Рис. Оптимальное число шенноновских каналов M_B , обеспечивающее наилучшее согласие с идеальной кривой, зависит как от уровня шума, так и от углового шага данных МУРР/МУРН (Таблица 2.1). Следует также отметить, что качество усеченного шенноновского приближения зависит от анизометрии объекта. Для очень анизометричных частиц высокие уровни шума (100% шума на Рис. 2.1а; 20% и 100% шума на Рис. 2.1б) приводят к значительным осцилляциям для шенноновских приближений. Тем не менее, все приближения на Рис. 2.1, даже с осцилляциями, обеспечивают наилучшее согласие с идеальной кривой по сравнению с шенноновскими приближениями с другими значениями числа М, и поэтому являются наилучшими приближениями, полученными в рамках интерполяционной формулы Шеннона.

Как видно из Таблицы 2.1А, для данных с плотной угловой сеткой (типичной для данных МУРР) и хорошей точностью измерений (при наличии от 1 до 5% шума), для получения наилучшего шенноновского приближения иногда требуется больше шенноновских каналов M_B , чем их формально содержится в малоугловой кривой рассеяния N_S , это указывает на то, что количество информации в данных позволяет проводить экстраполяцию за пределы доступного углового диапазона. Эта

возможность отражает хорошо известное свойство аналитических функций, измеренных на плотной сетке значений аргумента функции (учтем, что интенсивность рассеяния, являющаяся преобразованием Фурье функции p(r), отличной от нуля на конечном диапазоне значений r от 0 до D_{max} , также является аналитической функцией согласно теореме Винера-Пейли-Шварца [118]). Этот получения "сверхразрешения" эффект используется, например, ДЛЯ реконструкции оптических изображений [119], но его можно наблюдать только для очень точных данных. Очевидно, что M_B уменьшается с увеличением уровня добавляемого шума, но интересно и несколько неожиданно, что для данных МУРР на плотной угловой сетке, даже при очень высоком (100% и выше) уровне шума, M_B все еще может быть практически равен N_S (если принять во внимание, что неопределенность в определении M_B составляет ± 1). Другими словами, данные МУРР, измеренные на плотной угловой сетке, даже если они выглядят очень шумными (например, Рис. 2.1в, кривая 4), все равно содержат полезную информацию об идеальной кривой рассеяния во всем своем угловом диапазоне. Напротив, для данных, смоделированных на более разреженной угловой сетке (типичной для МУРН эксперимента), M_B начинает уменьшаться уже при уровне шума 20-50% (Таблица 2.16), что указывает на недостаточное количество информации для определения имеющихся шенноновских каналов N_S в случае разреженных зашумленных данных.

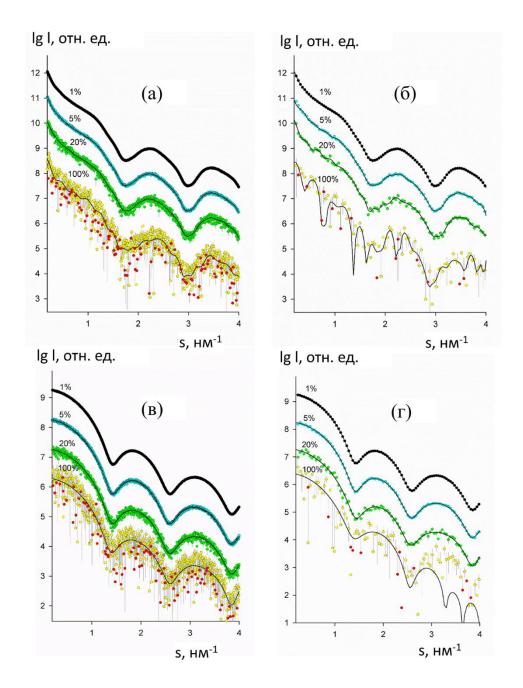


Рис. 2.1 Теоретические кривые рассеяния от сплюснутого эллипсоида (полуоси 1, 15 и 15 нм, панели (а) и (б)) и куба с ребром 5 нм (панели (в) и (г)). Кривые 1-4 соответствуют добавленному гауссовскому шуму с относительным уровнем 1, 5, 20 и 100 % (точки с линейками ошибок). Наилучшие усеченные шенноновские приближения показаны сплошными линиями. Последующие кривые сдвинуты на один логарифмический порядок для лучшей визуализации. Интенсивность отображается в логарифмическом масштабе. Для зашумленных теоретических данных, где значения могут стать отрицательными из-за шума, логарифмы модуля интенсивности показаны в виде красных точек. Панели (а), (в) рассчитаны на угловой сетке типичной для рентгеновского эксперимента, панели (б), (г) – для нейтронных данных.

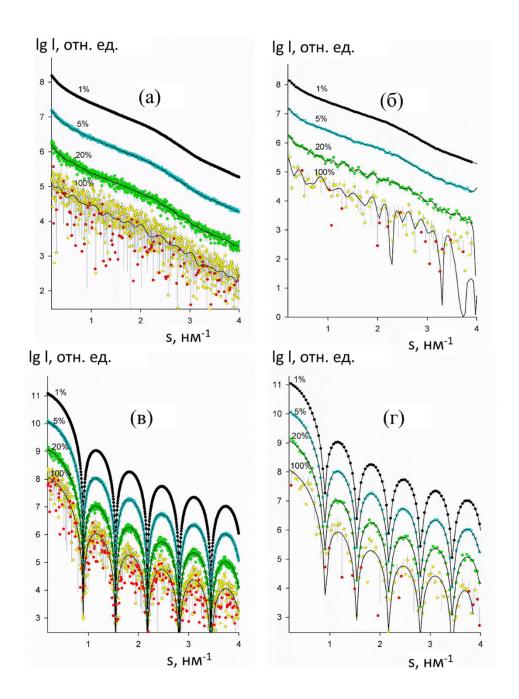


Рис. 2.2 Теоретические кривые рассеяния от вытянутого эллипсоида (полуоси 1, 1 и 15 нм, панели (а) и (б)) и от сферы (радиус 5 нм) (панели (в) и (г)). Обозначения и цветовая схема такие же, как на Рис. 2.1.

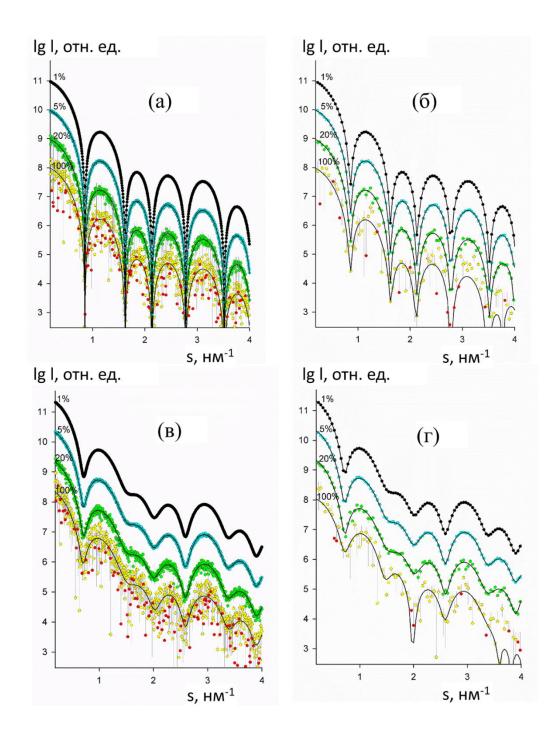


Рис. 2.3 Теоретические кривые рассеяния от полой сферы (внутренний радиус 2,5 нм, внешний радиус 5 нм, панели (а) и (б)) и от полого цилиндра (внутренний радиус 2,5 нм, внешний радиус 5 нм, высота 10 нм) (панели (в) и (г)). Обозначения и цветовая схема такие же, как на Рис. 2.1.

Таблица 2.1 Результаты тестирования метода на теоретических наборах данных рассеяния, рассчитаных от геометрических тел. Кривые малоуглового рассеяния рассчитаны на угловом диапазоне до s_{max} =4 нм⁻¹, к ним добавлен шум с относительным уровнем от 1% до 400%. Столбцы, расположенные справа от номинального числа шенноновских каналов N_S = s_{max} D_{max} / π показывают оптимальное число шенноновских каналов M_B , которое обеспечивает наилучшее согласие с идеальной (незашумленной) кривой. Максимальный уровень шума, при котором полезная информация все еще присутствует на всем угловом диапазоне кривой выделен жирным шрифтом. Таблица 2.1а соответствует рентгеновским данным, Таблица 2.1б - данным нейтронного рассеяния.

Таблица 2.1а (данные МУРР, плотная угловая сетка)

Тип тела	D_{max}	N_S	M_B							
	(нм)		1%	5%	20%	50	100	200	400	
						%	%	%	%	
Сплюснутый	30	38	41	39	39	38	38	38	25	
эллипсоид										
(полуоси										
15, 15, 1 нм)										
Вытянутый	30	38	39	39	39	38	38	37	23	
эллипсоид										
(полуоси										
1, 1, 15 нм)										
Полая сфера	10	13	14	14	13	13	13	12	11	
(Rin 2.5 нм,										
Rout 5 нм)										
Полый цилиндр	14	18	19	19	18	18	18	16	14	
(Rin 2.5 нм,										
Rout 5 нм,										
Н 10 нм)										
Куб (сторона	8.6	11	12	12	12	11	11	10	9	
5 нм)										
Сфера (радиус	10	13	15	14	14	14	14	14	13	
5 нм)										

Таблица 2.16 (данные МУРН, разреженная угловая сетка)

Тип тела	D_{max}	N_S	M_B						
	(нм)		1%	5%	20%	50	100	200	400
						%	%	%	%
Сплюснутый	30	38	38	38	38	38	38	37	25
эллипсоид									
(полуоси									
15, 15, 1 нм)									
Вытянутый	30	38	38	38	38	37	36	34	15
эллипсоид									
(полуоси									
1, 1, 15 нм)									
Полая сфера	10	13	13	13	13	12	11	11	10
(Rin 2.5 нм,									
Rout 5 HM)									
Полый цилиндр	14	18	18	18	18	17	16	16	14
(Rin 2.5 нм,									
Rout 5 нм,									
Н 10 нм)									
Куб (сторона	8.6	11	11	11	11	10	10	9	8
5 нм)									
Сфера (радиус	10	13	13	13	13	13	13	11	10
5 нм)									

В реальном эксперименте МУРР/МУРН идеальная кривая рассеяния и, следовательно, число M_R . конечно, недоступны, и эффективное число каналов M_S должно шенноновских определяться только экспериментальных данных. Моделирование на теоретических наборах данных, приведенное выше, позволило нам определить количественные критерии для выбора M_S . Такой выбор может быть осуществлен путем мониторинга невязки χ^2 между экспериментальными данными и шенноновским приближением $U_M(s)$ как функции M, поскольку плохо определенные шенноновские каналы уже не будут значительно улучшать невязку. Такая процедура, например, используется для определения числа независимых компонентов при сингулярном разложении набора данных [120], однако, формализация условия "несущественности" компонента не является тривиальной, и результаты такого разложения

оказываются не всегда точными. К счастью, надежную оценку M_S можно получить, комбинируя критерии качества приближения для обратного прямого пространства. Действительно, каждое шенноновское приближение $U_M(s)$, выраженное набором коэффициентов a_n , соответствует распределению расстояний в прямом пространстве $p_M(r)$ согласно уравнению (2.3). Увеличение числа M добавляет дополнительные члены к $p_M(r)$, осциллирующие со все более высокой частотой n/D_{max} . Можно ожидать, что ненадежно определенные шенноновские каналы a_n не дадут ничего, кроме нарастающих осцилляций в функции $p_M(r)$, и это можно уловить с помощью меры интегральной производной функции $\Omega(p)$:

$$\Omega(p) = \int_{0}^{D_{\text{max}}} \left[\frac{dp_M(r)}{dr} \right]^2 dr \tag{2.8}$$

Следовательно, качество шенноновского приближения может быть оценено с помощью комбинированного критерия

$$f(M) = \chi^2(M) + \alpha \Omega(p_M) \tag{2.9}$$

где коэффициент α обеспечивает правильное масштабирование двух метрик χ^2 и $\Omega(p)$ (см. ниже). Таким образом, процедура определения оптимального числа шенноновских каналов M_S может быть сформулирована следующим образом:

- 1. Для данного набора экспериментальных данных оценить максимальный размер частиц D_{max} (это делается, например, с помощью программ DATRG и DATGNOM [A7].
- 2. Рассчитать номинальное число шенноновских каналов как $N_S = s_{max} D_{max} / \pi$ и задать диапазон поиска наилучшего шенноновского приближения по числу M: $[M_{min}, M_{max}]$. На практике удобно использовать $M_{min} = \max(3, 0.2*N_S)$ и $M_{max} = 1.25*N_S$.
- 3. Для всех M_{min} </br/>M</br/> M_{max} вычислить коэффициенты шенноновского приближения a_n (n=1,...M), решив систему уравнений (2.6) с помощью процедуры линейных наименьших квадратов с ограничением на неотрицательность решений a_n [117].

- 4. Для каждого шенноновского приближения вычислить невязку $\chi^2(M)$ и меру интегральной производной $\Omega(p_M)$.
- 5. Оценить масштабный коэффициент α как отношение между χ^2 (M_{max}) и $\Omega(p(M_{min}))$.
- 6. Определить оптимальное значение M_S , соответствующее минимуму целевой функции f(M), согласно ее определению из уравнения (2.9).

Типичный пример шенноновских приближений с использованием различного числа каналов и соответствующих функций p(r) показан на Рис. 2.4 для случая вытянутого эллипсоида.

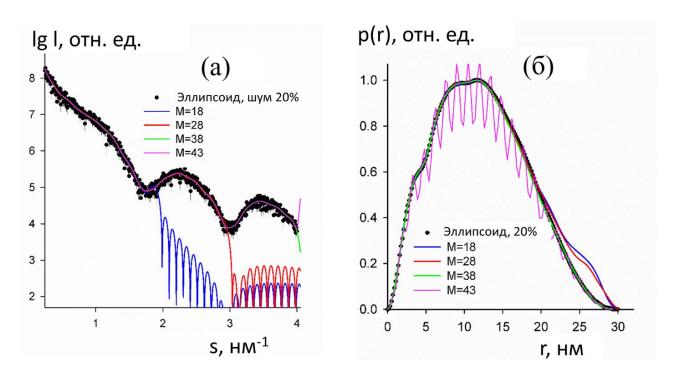


Рис. 2.4 Усеченное шенноновское приближение для теоретической кривой рассеяния от сплюснутого эллипсоида (полуоси 1, 15 и 15 нм) с добавлением шума на уровне 20%. Панель (а), кривая рассеяния с шумом (точки с доверительным интервалом ошибок) и шенноновские приближения, полученные при M=18 (синяя кривая), 28 (красная кривая), 38 (зеленая кривая) и 43 (малиновая кривая). Панель (б), функция распределения по расстояниям p(r), рассчитанная по данным рассеяния без шума (кривая 1) и функции $p_M(r)$, вычисленные по соответствующим шенноновским приближениям (кривые 2-5). Цветовая схема та же, что и на панели (а).

Значения χ^2 ожидаемо уменьшаются с увеличением числа шенноновских каналов M (Рис. 2.5, кривая 1), достигая плато при приближении к M_B (что для данного примера совпадает с N_S). Интегральная производная $\Omega(p_M)$ вначале лишь немного увеличивается с ростом M, но начинает резко расти, когда M превышает M_B (Рис. 2.5, кривая 2). Такое поведение еще раз подтверждает тот факт, что шенноновские каналы за пределами диапазона их достоверного определения уже существенно не улучшают невязку для интерполированной кривой, но в то же время приводят к сильным осцилляциям в функции p(r) (что хорошо видно на Рис. 2.4В). Таким образом, в целевой функции f(M) при малых значениях M доминирует слагаемое невязки χ^2 , тогда как при больших значениях M основной вклад уже вносит интегральная производная $\Omega(p_M)$ из-за резкого роста осцилляций функции p(r) (Рис. 2.5, кривая 3). Это приводит к характерному U-образному профилю целевой функции f(M) и позволяет относительно легко определить значение M_S , соответствующее минимуму целевой функции f(M).

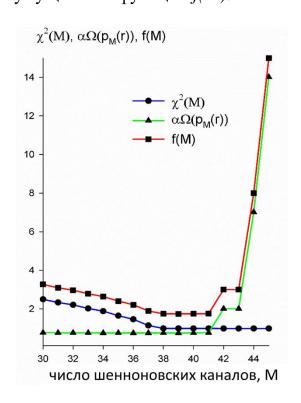


Рис. 2.5 Невязка шенноновского приближения $\chi^2(M)$ и интегральная гладкость $\Omega(p_M(r))$ в зависимости от числа шенноновских каналов M для теоретических данных, изображенных на Рис. 2.4. Кривая 1 (кружки, синяя линия) показывает

невязку $\chi^2(M)$, кривая 2 (треугольники, зеленая линия) - шкалированную интегральную гладкость $\alpha\Omega(p_M(r))$. Целевая функция f(M) соответствует кривой 3, (квадраты, красная линия).

2.2 Оценка работы алгоритма на теоретических и экспериментальных данных МУРР/МУРН

Описанный в предыдущем разделе алгоритм был реализован в виде компьютерной программы SHANUM [A33] Для проверки эффективности работы алгоритма программа SHANUM была протестирована на теоретических наборах кривых малоуглового рассеяния (Таблица 2.1), и с помощью нее были определены значения M_S , совпавшие со значениями M_B с точностью до одного шенноновского канала для всех случаев (Таблица 2.1). Данные тестовые расчеты показали, что предложенный алгоритм позволяет надежно определять эффективное число шенноновских каналов M_S по набору данных и, следовательно, оценивать полезный угловой диапазон данных МУРР/МУРН (согласно $s=\pi M_S/D_{max}$).

После проверки на теоретических данных, метод был применен к ряду экспериментальных наборов рентгеновских и нейтронных малоугловых данных, собранных в различных угловых диапазонах от растворов макромолекул, содержащих частицы различных размеров при различных концентрациях. Ниже представлены результаты, иллюстрирующие эффективность работы метода по определению полезного углового диапазона данных МУРР/МУРН. Данные рентгеновского синхротронного малоуглового рассеяния были получены в рамках X33 **EMBL** совместных пользовательских проектов на станции накопительного кольца DORIS-III (DESY, Гамбург). На Рис. 2.6а представлены данные МУРР от белкового комплекса импортин α/β с молекулярной массой (*MM*) 160 кДа и максимальным размером D_{max} =19 нм [A34]. Из-за низкой концентрации белка (0.5 мг/мл) данные МУРР сильно зашумлены в области больших углов. Несмотря на то, что измеренный диапазон векторов рассеяния (до s_{max} =6 нм⁻¹) номинально содержит N_S =36 шенноновских каналов, программа SHANUM дает

оценку для M_s =9, указывая на то, что высокоугловые данные за пределами s=1.5 нм⁻¹ не содержат полезной информации. В то же время, для данных МУРР от комплекса ДНК-метилтрансферазы SsoII (MM=45 кДа, D_{max} =11 нм) [A35], хотя кривая рассеяния представленная на Рис. 2.66 кажется довольно шумной, начиная с s=2 нм⁻¹, алгоритм показывает, что данные содержат полезную информацию вплоть до 4 нм⁻¹. Наконец, данные МУРР по ферменту люмазинсинтазе LSAQ-IDEA [A13], образующему икосаэдрические капсулы в растворе (с MM=2 МДа и D_{max} =33 нм), имеют хорошее соотношение сигнал/шум во всем угловом диапазоне, изображенном на Рис. 2.6в, и алгоритм ожидаемо выдал оценку, что весь угловой диапазон данных с 20 шенноновскими каналами содержит полезную информацию. Интересно, что оценки, сделанные программой SHANUM хорошо коррелируют с диапазонами данных, которые фактически использовались для анализа данных в вышеупомянутых публикациях.

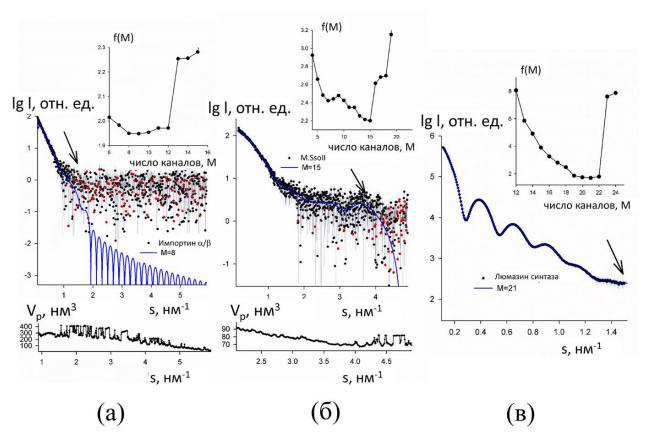


Рис. 2.6 Экспериментальные данные МУРР и шенноновские приближения, соответствующие оптимальному числу шенноновских каналов, определенному программой SHANUM. Панель (а), белковый комплекс импортин α/β [A34]; панель

(б), белок M.SSoII [A35]; панель (в), белок LSAQ-IDEA [A13]. Верхняя граница полезного углового диапазона обозначена стрелкой, а целевая функция f(M) показана на вставке. На вставках внизу панелей (а) и (б) показана зависимость исключенного объема частицы (Породовский объем) от диапазона экспериментальных данных, использованных для расчета инварианта Порода.

Также интересно проверить, применим ли разработанный метод к данным широкоуглового рентгеновского рассеяния (WAXS: Wide Angle X-ray Scattering). Кривые WAXS предоставляют информацию с более высоким разрешением и обычно содержат большее количество шенноновских каналов по сравнению с данными МУРР. Работа алгоритма была проверена на данных WAXS от концентрированного раствора миоглобина (28 мг/мл) (данные депонированы в базу данных SASBDB, www.sasbdb.org, код SASDAK2 [110]) и разбавленного раствора цитохрома С (2 мг/мл) (данные получены на станции X33, неопубликованные данные). Если для первого случая весь имеющийся угловой диапазон данных WAXS был признан содержащим полезную информацию, то для второго случая только половина этого диапазона была признана информативной (Рис. 2.7).

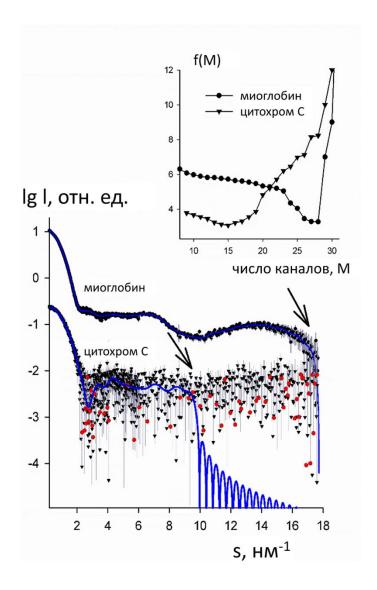


Рис. 2.7 Экспериментальные широкоугловые рентгеновские данные и шенноновские приближения, соответствующие оптимальному числу каналов, определенному программой SHANUM. Панель (а), концентрированный раствор белка миоглобина; панель (б), разбавленный раствор белка цитохрома С. Верхняя граница полезного углового диапазона обозначена стрелкой, а целевая функция f(M) показана на вставке.

Для полноты картины проиллюстрируем использование алгоритма еще на нескольких опубликованных наборах данных МУРН. На Рис. 2.8а представлены данные МУРН от тиоредоксин-редуктазы, димерного белка с MM=68кДа и D_{max} =11 нм, полученные на станции D22 в Институте Лауэ Ланжевена, Гренобль, Франция [75]. Два набора данных, собранных в растворах H_2O и в D_2O в одном и том же угловом диапазоне (до s_{max} =5.2 нм⁻¹), номинально оба набора содержат N_S =17 шенноновских каналов. Однако данные в растворе H_2O более шумные из-за

меньшего контраста и некогерентного фона, поэтому программа SHANUM дает оценку в 14 эффективных шенноновских каналов для H_2O и 16 шенноновских каналов для данных в D_2O .

Следующий пример показывает, предложенный что нами подход ограничивается применением только для биологических макромолекул в водных растворах. Данные МУРН на Рис. 2.8б были измерены на линии KWS-2 (нейтронный реактор FRM-II, Технический Университет, Мюнхен, Германия) от гибридных золотых наночастиц, стабилизированных додеканэтиол-гексанэтиолом (С12-С6) и растворенных в дейтерированном хлороформе [121]. Кривые (1) и (2) относятся К гибридным частицам co специально дейтерированными гексанэтиолом, соответственно. Растворы композитных додеканэтиолом и наночастиц близки к монодисперсным с диаметром частиц 8 нм, о чем свидетельствуют формы кривых рассеяния, а также дополнительные исследования [121]. Программа SHANUM дает довольно реалистичные результаты, предполагая, что большая часть кривой 1 информативна, в то время как последняя треть более шумной кривой 2 не несет никакой полезной информации. Так как химически синтезированные наночастицы неизбежно обладают определенной степенью полидисперсности, представленный пример указывает на применимость алгоритма не только для небиологических, но и для умеренно полидисперсных систем.

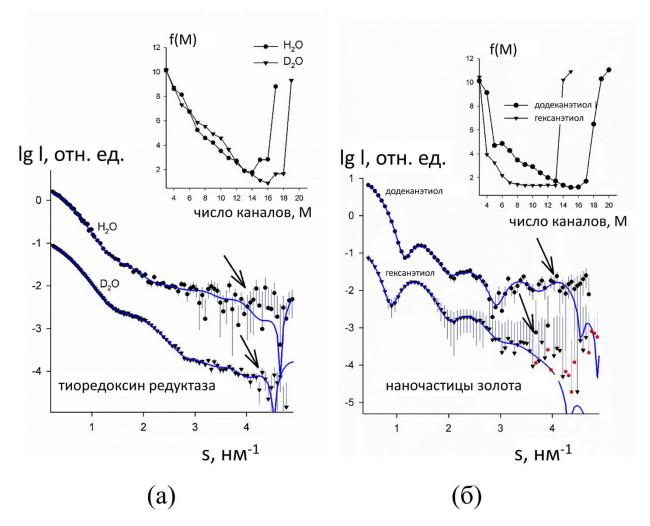


Рис. 2.8 Экспериментальные данные МУРН и шенноновские приближения, соответствующие оптимальному числу каналов, определенному программой SHANUM. Панель (а), данные от растворов белка тиоредоксин-редуктазы; кривые (1) и (2) измерены в H_2O и D_2O , соответственно [75]. Панель (б), кривые рассеяния от композитных наночастиц золота в дейтерированном хлороформе, со специально дейтерированным додеканэтиолом (кривая 1) и гексанэтиолом (кривая 2) [121]. Верхняя граница полезного углового диапазона обозначена стрелкой, а целевая функция f(M) показана на вставке.

Следует отметить, что программа SHANUM требует только задания максимального размера частицы D_{max} для определения полезного углового диапазона. По умолчанию для оценки D_{max} используются программы DATRG и DATGNOM [A7], однако, если это значение известно *a priori* (например, при анализе данных от белка с известной структурой), оно может быть напрямую задано пользователем. Важно отметить, что формализм Шеннона (уравнения (2.4-2.6)) справедлив не только для максимального размера D_{max} , но и для любого

значения $D > D_{max}$. Это делает всю процедуру еще более надежной, позволяя безопасно использовать несколько завышенный максимальный размер, а также работать со слегка полидисперсными системами (см. пример с наночастицами, представленный выше). В тестах и практических расчетах, использование завышенных на 5-10% значений давало практически те же полезные угловые диапазоны данных МУРР/МУРН.

рентгеновской кристаллографии полезный угловой диапазон оцениваемый по $< I/\sigma(I) >$ и R_{merge} , определяет набор отражений, использующийся при уточнении моделей, и следовательно, напрямую определяет разрешение моделей. В методе МУРР/МУРН отбрасывание данных в области больших углах не повлияет на точность оценки структурных параметров, например, радиуса инерции, определяемого по данным МУРР/МУРН в области малых углах с помощью аппроксимации Гинье [5]. Очевидно, что удаление неинформативной данных должно улучшить результаты анализа косвенных преобразований, процедур восстановления формы частиц, использующих данные WAXS (например, с помощью программы GASBOR [53]), а также расчет общих параметров частиц, таких как Породовский объем частицы V_p . Последний представляет собой объем исключенной ИЗ растворителя рассчитывается по формуле [33]:

$$V_p = \frac{2\pi^2 I(0)}{Q}, \quad Q = \int_0^\infty s^2 I(s) ds$$
 (2.10)

На практике инвариант Порода Q рассчитывается на конечном угловом диапазоне $[0, s_m]$, и затем для компенсации недостающих данных от s_m до бесконечности применяются дополнительные поправки (например, как реализовано в модуле POROD программы PRIMUS [A5]). На нижней панели Рис. 2.6А представлена оценка объема Порода для белкового комплекса импортин α/β как функции верхнего предела интегрирования s_m . Если учесть эмпирическое соотношение между объемом и молекулярной массой частицы V_p [нм³] $\approx 1,7$ -1,8*MM [кДа] [A7], то ожидаемый объем Порода для комплекса импортин α/β составляет около 280

нм³. Объем, непосредственно вычисленный модулем POROD в программе PRIMUS, дает стабильные значения с небольшими колебаниями, находясь внутри полезного углового диапазона данных МУРР, оцененного программой SHANUM (т.е. до s_m около 1.3 нм⁻¹), и начинает резко колебаться, как только при его расчетах учитываются данные МУРР в области больших углов. Аналогично, для ДНК-метилтрансферазы SsoII, V_p показывает разумные значения около 75 нм³, когда s_m остается в пределах полезного углового диапазона данных, и проявляет резкие колебания за пределами этого диапазона (нижняя панель Рис. 2.66). Заметим, что на практике приведенное выше эмпирическое соотношение используется в обратном направлении, и V_p считается одним из способов оценки MM, не требующим проведения абсолютной калибровки данных МУРР. Эти примеры иллюстрируют важность удаления части данных, не содержащих полезной информации, для предотвращения потенциальных проблем и неточностей при оценке основных структурных параметров частиц.

Особо следует подчеркнуть, что предложенный алгоритм не должен служить в качестве фильтра низких частот для обеспечения снижения шума путем приближения экспериментальных данных. Как видно из Рис. 2.1, при высоком уровне шума шенноновские приближения могут иметь заметные осцилляции, особенно это характерно для анизометричных частиц. Кроме того, усеченные шенноновские приближения неизбежно приводят к эффектам обрыва при проведении Фурье преобразования из-за отсутствия высокочастотных гармоник (в нефизичные частности, может иметь отрицательные осциллирующие вокруг нуля для значений аргументов s, превышающих $\pi M/D_{max}$). Поэтому разработанный подход следует использовать как средство оценки информативности данных МУРР/МУРН, а не как инструмент сглаживания зашумленных данных.

Заключение к Главе 2

До настоящего времени не существовало надежной процедуры для оценки полезного углового диапазона экспериментальных данных МУРР/МУРН. Основными проблемами оценки на основе отношения сигнал/шум являются отсутствие объективности при выборе порога и тот факт, что не учитывается шаг угловой сетки данных МУРР/МУРН. Разработанный автором метод преодолевает обе проблемы и предлагает объективную процедуру определения полезного углового диапазона. Процедура реализована в виде программы SHANUM, входящей в пакет ATSAS. Программа может быть использована для выбора части экспериментальных данных МУРР/МУРН, содержащей полезную информацию, в автоматизированных системах обработки данных, в процедурах моделирования и при депонировании кривых рассеяния в имеющиеся базы данных.

Глава 3

Количественный анализ состава многокомпонентных полидисперсных систем по данным МУРР/МУРН

Как было показано в Главе 1 методы *ab initio* восстановления 3D формы частиц надежно работают для данных малоуглового рассеяния от растворов белков (или других макромолекул), которые принимают одну конформацию в растворе (монодисперсные системы). Однако во многих случаях возникают очень интересные биохимические вопросы относительно поведения ансамбля частиц в растворе в условиях, когда они принимают несколько конформаций. В частности, ансамбли, состоящие из смесей мономеров, димеров и олигомеров более высокого порядка, или смеси различных конформаций белка представляют собой важные практические приложения метода МУРР/МУРН.

Белки - это сложные молекулярные машины, функции которых часто зависят от конформационных "крышка", изменений, например, когда специальная имеющаяся в структуре белка (шаперона) закрывается, чтобы захватить субстрат сайте фермента [122]. Эксперименты активном ПО рентгеновской на кристаллографии позволяют получать подробные "снимки" белковых структур на атомном разрешении, что безусловно способствует пониманию высоком взаимосвязи структуры И функции белка [123]. Однако, как правило кристаллические структуры дают информацию о единственной конформации фактически являясь 'единственным снимком', выбранным белка конформационного ансамбля структур, которые белок может/должен принимать во время своей работы [124]. Следовательно, конформационные состояния, которые являются биологически важными, например, долгоживущие промежуточные состояния в процессе конформационного перехода, трудно идентифицировать с помощью кристаллографии.

В отличие от этого, методы МУРР/МУРН предоставляют информацию о структуре белка, усредненную по ансамблю имеющихся структур при любом заданном состоянии раствора, что в принципе позволяет идентифицировать изменения в

конформации белка вследствие связывания лиганда [125] или изменения его окислительно-восстановительного состояния [126]. В частности, если белок существует в различных конформационных состояниях, то данные МУРР/МУРН будут представлять собой взвешенную комбинацию интенсивности рассеяния от каждого отдельного состояния, где веса определяются объемной долей соответствующего состояния. Следовательно, эксперименты МУРР/МУРН, которые проводят при варьировании таких физико-химических параметров, как температура или концентрация белка, для изучения изменения относительного количества этих конформаций, дают возможность получать новые сведения о функции белка. Целью таких исследований является определение относительной концентрации и/или структуры каждого из присутствующих компонентов.

В настоящей Главе будут детально рассмотрены основные подходы, разработанные автором, позволяющие проводить количественный анализ состава многокомпонентных полидисперсных систем по данным МУРР/МУРН. Они могут быть использованы как для анализа биологических систем (сложных белковых смесей), так и для анализа состава микроэмульсий и растворов наночастиц. Для каждого из развитых методов будут представлены практические примеры их использования.

3.1 Оценка числа компонентов и нахождение объемных долей компонентов белковой смеси с известными атомными структурами

Существует несколько подходов для количественного анализа данных малоуглового рассеяния от систем, содержащих отдельные компоненты, картина рассеяния которых представляет собой линейную комбинацию, согласно уравнению (1.22).

Когда ни количество, ни интенсивность компонентов неизвестны, модельнонезависимый анализ нескольких наборов кривых рассеяния, полученных при различных состояниях системы, проводится с помощью разложения по сингулярным значениям (SVD, Singular Value Decomposition [120]). Матрица данных $\mathbf{A} = \{A_{ik}\} = \{I^{(k)}(s_i)\}, (i=1,...N, k=1,...K, где N- число экспериментальных$ точек в кривой рассеяния, K - номер кривой рассеяния в наборе данных) представляется как $A = U*S*V^T$, где матрица S диагональная, а столбцы ортогональных матриц U и V - собственные векторы матриц $A*A^T$ и A^T*A , соответственно. Матрица U дает набор так называемых левых сингулярных векторов, т.е. ортонормированных базисных кривых $U^{(k)}(s_i)$, которые охватывают диапазон матрицы А, тогда как диагональ S содержит их соответствующие сингулярные значения в порядке убывания (чем больше сингулярное значение, тем значимее вектор). Физически, количество значимых сингулярных векторов (неслучайных кривых co значимыми сингулярными значениями) минимальное количество независимых кривых, необходимых для представления всего набора данных их линейными комбинациями (например, для смесей, см. ниже). Данный алгоритм был реализован в программе SVDPLOT, которую можно также вызывать из меню программы PRIMUS [A5]. Программа SVDPLOT позволяет проводить SVD разложение активных наборов данных МУРР/МУРН (кривых рассеяния), содержащихся в инструментальной панели программы PRIMUS, и получать сингулярные вектора (фиктивные кривые рассеяния, обеспечивающие ортогональный базис для линейного представления всего набора данных) и связанные с ними сингулярные значения (веса компонентов в этом базисе). Количество неслучайных сингулярных векторов co значимыми сингулярными значениями (которое оценивается с помощью непараметрического теста Вальда-Вольфовица на случайность осцилляций [127]) дает минимальное количество независимых кривых, необходимых для представления всего набора данных, т.е. количество значимых компонентов в смеси (например, количество промежуточных продуктов реакции в развивающемся динамическом процессе). Таким образом, в ряде случаев с помощью SVD оценки удается достаточно точно установить количество независимых компонентов (олигомерных состояний) в исследуемой системе, что помогает обосновать и выбрать соответствующее описание системы при наличии информации об атомных структурах компонентов.

Если число компонентов и их интенсивности рассеяния известны, то для анализа данных МУРР/МУРН можно использовать программу OLIGOMER [A5], в которой реализован линейный алгоритм наименьших квадратов с ограничением на неотрицательность решений [117] для нахождения объемных компонентов смеси, минимизирующих невязку между экспериментальными и Программа **OLIGOMER** рассчитанными кривыми рассеяния. позволяет обрабатывать сразу несколько наборов экспериментальных данных МУРР/МУРН относительно одного и того же набора функций $I_k(s)$, что особенно полезно, например, для быстрой оценки состава олигомерных равновесных смесей белков с известной структурой высокого разрешения при титровании раствора.

В качестве входных данных программа OLIGOMER запрашивает файл с формфакторами, содержащий набор кривых рассеяния от известных структур компонентов $I_k(s)$. Его можно составить вручную, рассчитав с помощью программы CRYSOL кривые рассеяния от имеющихся PDB структур, и затем скопировав полученные кривые рассеяния в один файл. Однако ручная подготовка такого файла может быть громоздкой процедурой, особенно если необходимо объединить данные из разнородных моделей. Для облегчения этой задачи была разработана программа FFMAKER [A7]. С помощью нее можно автоматически объединить интенсивности рассеяния, набранные из различных источников: (а) теоретические интенсивности рассеяния от PDB структур высокого разрешения, рассчитанных программой CRYSOL [70], (б) экспериментальные кривые рассеяния и (в) регуляризованные кривые рассеяния, полученные с помощью программы GNOM [40]. FFMAKER может быть запущен в интерактивном или автоматическом режиме, причем для последнего наиболее важные параметры могут быть заданы в качестве аргументов в командной строке. Программа FFMAKER также предоставляет возможность автоматической оценки нескольких интенсивностей рассеяния от ансамбля структуры, полученной методом ЯМР и хранящегося в одном PDB файле.

Дополнительной опцией ('-compar') в программе OLIGOMER является возможность автоматического перебора всех неполных комбинаций от имеющегося набора кривых рассеяния компонентов $I_k(s)$. В качестве результата выдается сравнительная таблица, где отсутствующие (не присутствующие) компоненты обозначены как 'N/A', причем последовательность комбинаций отсортирована по увеличению значения невязки между экспериментальными данными и рассчитанными приближениями. Эти результаты удобно использовать для определения наиболее значимых компонентов, присутствующих в смеси.

Если соотношение между долями компонентов известно и должно сохраняться (например, при диссоциации олигомера или белкового комплекса на отдельные фракции), то его можно сохранить в ходе вычислений программы OLIGOMER и найти наилучшее соответствие, которое все еще удовлетворяет желаемому соотношению долей. В целом, приближение может стать хуже по сравнению с результатом вычислений, проведенных без ограничений на соотношение долей компонентов, однако оно обеспечит физически обоснованные решения. Ограничения могут быть заданы как через соотношения объемных долей компонентов, так и в терминах соотношения для числа молекул между разными компонентами.

Таким образом, программа OLIGOMER позволяет анализировать количественный состав смеси при наличии информации о структуре компонентов, примеры ее практического применения будут рассмотрены в следующих разделах.

3.2 Приближение многокомпонентной системы простыми геометрическими телами с учетом полидисперсности и межчастичного взаимодействия

Для более сложных смесей различных типов частиц, когда рассеивающие частицы имеют определенную форму, но отличаются по размеру и полидисперсности, эти характеристики могут быть параметризованы и восстановлены вместе с их объемными долями с помощью нелинейной процедуры минимизации.

Рассмотрим систему, состоящую из K невзаимодействующих компонентов, и опишем полидисперсность k-го компонента функцией распределения по размерам $N_k(R)$, соответствующей количеству частиц с характерным размером R. Интенсивность рассеяния от k-го компонента является изотропной функцией, зависящей от момента передачи импульса (вектора рассеяния) $s=(4\pi/\lambda)sin\theta$, (λ - длина волны, 2θ - угол рассеяния), которая может быть записана как

$$I_k(s) = \int_0^\infty N_k(R) \ [V_k(R) \ \Delta \rho_k(R)]^2 i_{0k}(s, R) \ dR, \tag{3.1}$$

где $\Delta \rho_k(R)$, $V_k(R)$ и $i_{0k}(s, R)$ обозначают контраст, объем и нормированную интенсивность рассеяния (форм-фактор) частицы размером R (эти функции определяются формой и внутренней структурой частиц, а $i_{0k}(0, R) = 1$). Общий объем компонента задается следующим образом

$$v_k = \int_0^\infty V_k(R) \ N_k(R) \ dR \tag{3.2}$$

так что уравнение (3.1) может быть переписано как

$$I_k(s) = V_k \frac{\int_0^\infty D_k(R) V_k(R) \left[\Delta \rho_k(R) \right]^2 i_{0k}(s,R) dR}{\int_0^\infty D_k(R) dR}$$
(3.3)

где $D_k(R) = N_k(R)V_k(R)$ обозначает функцию распределения по объему, которая, как видно из уравнения (3.3), без потери общности может быть нормирована, $\int\limits_0^\infty D_k(R)dR = 1$. Тогда интенсивность рассеяния от смеси записывается как линейная

комбинация парциальных интенсивностей

$$I(s) = \sum_{k=1}^{K} v_k I_{k0}(s)$$
 (3.4)

где

$$I_{k0}(s) = \int_0^\infty D_k(R) \ V_k(R) \ [\Delta \rho_k(R)]^2 i_{0k}(s, R) \ dR$$
 (3.5)

Функции распределения по объему частиц $D_k(R)$ удобно представлять нормированным распределением Шульца [128] с использованием двух параметров, среднего значения R_{0k} и дисперсии ΔR_k :

$$D_k(R) = G(R, R_{0k}, \Delta R_k) = \left(\frac{z+1}{R_{0k}}\right)^{z+1} \frac{R^z}{\Gamma(z+1)} \exp\left[-\frac{(z+1)R}{R_{0k}}\right]$$
(3.6)

где $z=(R_{0k}/\Delta R_k)^2-1$.

В уравнениях (3.1) и (3.5) эффектами межчастичной интерференции мы пренебрегали, что справедливо только при низких концентрациях частиц. Поскольку в общем случае не существует аналитического выражения для рассеяния от смеси взаимодействующих частиц различной формы, необходимо сделать упрощающие предположения. В дальнейшем нами будут рассматриваться только взаимодействия между частицами, принадлежащими одному компоненту, а парциальные интенсивности будут приниматься в виде

$$J_{k0}(s) = I_{k0}(s) \ T_k(s), \tag{3.7}$$

где $T_k(s)$ структурный фактор, описывающий эффекты межчастичного взаимодействия для k-го компонента.

Для сферически симметричных потенциалов взаимодействия структурный фактор может быть вычислен аналитически в приближении Перкуса-Йевика [129] для потенциала притягивающихся жестких сфер, введенный Бакстером [130]:

$$T_k(s, R_k^{sh}, \eta_k, \tau_k) = \left[A^2(s, R_k^{sh}, \eta_k, \tau_k) + B^2(s, R_k^{sh}, \eta_k, \tau_k) \right]^{-1}, \tag{3.8}$$

где R^{hs} - радиус взаимодействия жестких сфер, η - объемная доля жестких сфер и τ - параметр притяжения, определяющий величину притягивающего потенциала на поверхности сфер (выражения для коэффициентов A и B можно найти в работе [130]).

Рассмотренный выше алгоритм был реализован в программе MIXTURE [A5] для проведения количественного анализа смесей частиц простой геометрической формы на основе данных МУРР/МУРН. Алгоритм позволяет моделировать смеси, содержащие до десяти различных компонентов (т.е. различных типов частиц), каждый описывается объемной долей, форм-фактором, ИЗ которых контрастностью, полидисперсностью и, для сферических частиц, потенциалом межчастичного взаимодействия. В программе можно использовать сплошные или полые сферические частицы или круговые цилиндры, эллипсоиды и гантели, для которых форм-фактор определяется их геометрическими параметрами (например, радиусами сферической оболочки R_k). Каждый компонент может соответствовать монодисперсному или полидисперсному состоянию, причем во втором случае размерная полидисперсность $D_k(R)$ описывается двухпараметрическим мономодальным распределением Гаусса или Шульца [128] (см. уравнение (3.6)), характеризующимся средним размером R_{0k} и дисперсией ΔR_k . Межчастичные взаимодействия могут быть учтены в пределах каждого компонента, состоящего из сферически симметричных частиц. Структурный фактор $T_k(s)$ рассчитывается в приближении Перкуса-Йевика с использованием потенциала притягивающихся жестких сфер [130], описываемого двумя параметрами, радиусом взаимодействия жесткой сферы R^{hs} и параметром притяжения сфер τ .

Нелинейная минимизация в программе MIXTURE использует программный комплекс OPTIS, включающий более десятка методов оптимизации (см. В.В.Волков, 2013, Докторская диссертация). По умолчанию программа MIXTURE использует квази-ньютоновскую процедуру минимизации по схеме Бройдена-Флетчера-Гольдфарба-Шанно с простыми ограничениями на переменные [131], что позволяет накладывать физически обоснованные ограничения на параметры. В качестве входных данных программе MIXTURE требуется командный файл, содержащий спецификации модели и начальные значения параметров, т.е. количество компонентов, тип каждого компонента, его относительную объемную долю, размерные параметры, средний радиус сферы, полидисперсность, тип функции распределения (Гаусс или Шульц) и т.д. Для всех параметров модели необходимо также задавать их верхние и нижние возможные пределы. Пример формата командного файла для системы полидисперсных сфер и цилиндров представлен в Таблице 4.1. Каждый приближаемый параметр (отмечен *) имеет следующий формат: начальное значение, нижняя граничная граница, верхняя граничная граница. Если три значения равны, то значение параметра становится фиксированным.

Таблица 4.1 Пример командного файла задания в программе MIXTURE для двухкомпонентной смеси сферических и цилиндрических частиц.

Test command file	Comment Line 1 (Строка комментария 1)
Spheres+Cylinders	Comment Line 2 (Строка комментария 2)
EXPERIMENT	Data fitting mode (Мода работы программы)
2	Number of components (Общее число компонент)
0.20	Total volume fraction (Объемная доля образца)
SPHERE	Type of the first component (Тип компонента 1)
0.25, 0.0, 1.0	*sphere volume fraction (Объемная доля сфер)
0.0, 0.0, 0.0	*inner shell radius (Внутренний радиус сферы)
0.0, 0.0, 0.0	*inner shell contrast (Контраст ядра сферы)
48.0, 40.0, 60.0	*outer shell radius (Внешний радиус сферы)
1.0, 1.0, 1.0	*outer shell contrast (Контраст сферической оболочки)
5.0, 0.1, 25.0	*sphere polydispersity (Полидисперсность сфер)
75.0, 55.0, 95.0	*hard sphere radius (Радиус притягивающихся жестких
73.0, 33.0, 93.0	сфер)
2	polydispersity (1=Gauss, 2=Schultz) (Тип распределения)
0.0, 0.0, 0.0	*sticky parameter (0=hard sphere) (Параметр
0.0, 0.0, 0.0	притягивания сфер)
CYLINDER	Type of the second component (Тип компонента 2)
0.25, 0.0, 1.0	* cylinder volume fraction (Объемная доля цилиндров)
0.0, 0.0, 0.0	* inner shell radius (Внутренний радиус цилиндра)
0.0, 0.0, 0.0	* inner shell contrast (Контраст ядра цилиндра)
40.0, 30.0, 60.0	* outer shell radius (Внешний радиус цилиндра)
1.0, 1.0, 1.0	* outer shell contrast (Контраст цилиндрической
1.0, 1.0, 1.0	оболочки)
3.0, 2.1, 10.1	* radius polydispersity (Полидисперсность цилиндра по
, ,	радиусу)
2	polydispersity (1=Gauss, 2=Schultz) (Тип распределения)
300.0	cylinder length (Длина цилиндра)
2	1=OTOKO data, 2=ASCII data (Тип данных)
test1.dat	experimental data file (Имя экспериментального файла)
1.0	angular scale multiplier (шкалирующий коэффициент
1.0	угловой оси)

Программа MIXTURE может запускаться автономно, либо с помощью всплывающего меню программы PRIMUS [A5], запрашивающего имена командного файла задания и файла экспериментальных данных. После проведения

нелинейной минимизации рассчитанное наилучшее приближение к экспериментальным данным МУРР/МУРН, а также найденные парциальные интенсивности рассеяния от компонентов и соответствующие структурные факторы сохраняются в отдельных файлах. Найденные (оптимизированные) параметры каждого компонента записываются в log-файл после каждого запуска, чтобы позволяет сохранять историю запусков программы MIXTURE при различных стартовых условиях.

Прототип MIXTURE был разработан программы при исследовании вода-АОТ-масло (где микроэмульсий вида AOT трехкомпонентных поверхностно-активное вещество - бис(2-этилгексил) сульфосукцинат натрия) и применен для количественного описания перехода сферических частиц в цилиндрические ассоциаты в этих классических микроэмульсионных системах [АЗ6]. Было установлено, что система представляет собой смесь трех компонентов (1) сферических капель воды, умеренно полидисперсных, средний радиус которых зависит от молярного соотношения вода/АОТ, (2) длинных цилиндрических агрегатов, (3) обратных мицелл АОТ (пример рассеяния от отдельных компонентов и от всей смеси показан на Рис. 4.1).

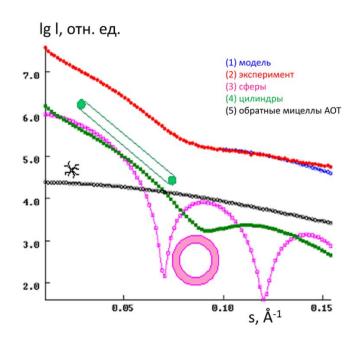


Рис. 4.1 Пример приближения данных МУРР от микроэмульсии вида вода-АОТ-масло трехкомпонентной моделью. Экспериментальные данные МУРР (кривая 2)

и интенсивность, рассчитанная от модели (кривая 1), являющаяся линейной комбинацией вкладов от сферических капель (3, малиновая кривая), длинных цилиндров (4, зеленая кривая) и обращенных мицелл (5, черная кривая). Формы этих трех компонентов схематично изображены на рисунке.

Алгоритмы восстановления функций распределения частиц по размерам для простых геометрических тел, заданных в параметрическом виде, реализованы в ряде программ, таких как SASFIT [132], SASVIEW (http://www.sasview.org), SCATTER [133], IRENA [41], X+ [134], но из них только MIXTURE обладает возможностями одновременного приближения несколькими типами полидисперсных частиц.

Помимо использования косвенного преобразования Фурье с регуляризацией по Тихонову (программы GNOM [40], GIFT [39]) и параметрического представления функции распределения, можно также проводить прямой поиск функции $D_V(R)$ в виде непараметрической гистограммы. Такие алгоритмы реализованы в программах McSAS [135] и VOLDIS [136].

Каждый из этих алгоритмов обладает своими преимуществами и недостатками, и использование только одного из них не всегда позволяет получать устойчивые решения, свободные от артефактов. В цикле работ [А37-А42] на примере ряда МУРР было теоретических И экспериментальных данных проведено систематическое исследование устойчивости решений в зависимости от начальных стартовых приближений модели, уровня шума и углового диапазона данных, и на основании полученных результатов предложена новая схема совместного использования алгоритмов, позволившая существенно повысить надежность восстановления функции распределения частиц по размерам [А43-А44]. Работу схемы можно разбить на три этапа:

- 1) расчет программой VOLDIS гистограммы распределения и оценка среднего радиуса частиц;
- 2) расчет программой GNOM функции распределения по размерам, используя максимальный размер частиц R_{max} , найденный программой VOLDIS;

3) уточнение параметров распределения с помощью программ POLYMIX/MIXTURE (программа POLYMIX имеет то же ядро целевой функции, как и программа MIXTURE, но использует модифицированный алгоритм Левенберга-Марквардта), используя в качестве стартового приближения распределения, максимально близкие к решениям, полученным на шаге (1) и (2), и комбинируя локальные и глобальные минимизационные методы при поиске распределения.

3.3 Применение разработанных методов для анализа олигомерных смесей, многокомпонентных самоорганизующихся систем и липидных везикул

3.3.1 Смесь активной и неактивной форм фермента

Относительное движение белковых доменов, допускаемое структурной гибкостью белка, часто играет ведущую роль в механизме действия различных ферментов. Закрытие домена над активным сайтом фермента может создать оптимальную среду для катализа, надлежащим образом ориентирует реакционные группы субстрата и/или препятствует побочным реакциям. Однако структурные принципы, регулирующие движение домена, не всегда понятны. Типичными примерами такого механизма, например, являются различные группы киназ, катализирующие перенос фосфогруппы с одного субстрата на другой [137].

Метод МУРР является наиболее подходящим методом для тестирования таких изменений конформации белка, происходящих в растворе, особенно если они сопровождаются изменениями в форме молекулы, которые действительно происходят при закрытии домена. Если известны атомные структуры открытой и закрытой форм фермента, то их можно использовать, чтобы оценить по данным МУРР какая из конформаций белка содержится в растворе, либо с помощью программы OLIGOMER определить относительные объемные доли этих форм, присутствующих в растворе.

Целый ряд работ был посвящен исследованиям структуры фермента 3-фосфоглицерат-киназы (сокращенно PGK) с помощью метода МУРР [A45-A48].

Этот фермент состоит из двух структурных доменов приблизительно равного размера (около 22 кДа) [138]. Каждый домен связывает один из двух субстратов: N-концевой домен связывает 3-фосфоглицерат (3-PG) и 1,3-бисфосфоглицерат (1,3-BPG), в то время как С-концевой домен связывает нуклеотидные субстраты, такие как MgATP или MgADP [139, 140]. PGK является важным ферментом для большинства живых организмов, он катализирует фосфо-перенос из 1,3-BPG в MgADP и продуцирует 3-PG и MgATP во время гликолиза. Следует также отметить, что PGK играет важную роль в фосфорилировании аналогов L-нуклеозидов, которые входят в состав лекарственных средств против рака и вирусных инфекций [141].

В указанных выше работах было проведено систематическое исследование двойных и тройных фермент-субстратных комплексов в растворе с помощью метода МУРР. Скрининг известных кристаллографических структур фермента (Рис. 4.4) относительно экспериментальных данных малоуглового рассеяния показал необходимость присутствия обоих субстратов для закрытия домена. На Рис. 4.5а представлены наилучшие приближения данных МУРР различными кристаллическими моделями, на Рис. 4.5б показаны типичные представители открытой (РGК без субстрата) и закрытой (тройной комплекс PGK*3-PG*MgATP) форм фермента.

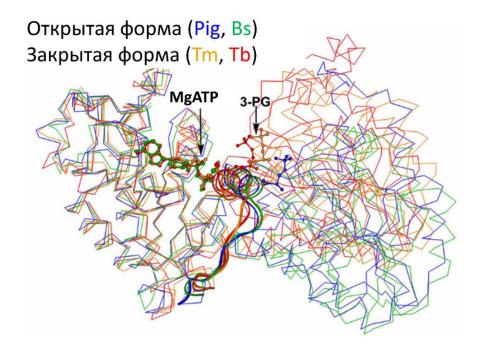


Рис. 4.4 Известные кристаллические структуры фермента PGK без субстрата (синяя и зеленая C_{α} цепи) (открытая форма) и тройного комплекса PGK*3-PG*MgATP (оранжевая и красная C_{α} цепи) (закрытая форма).

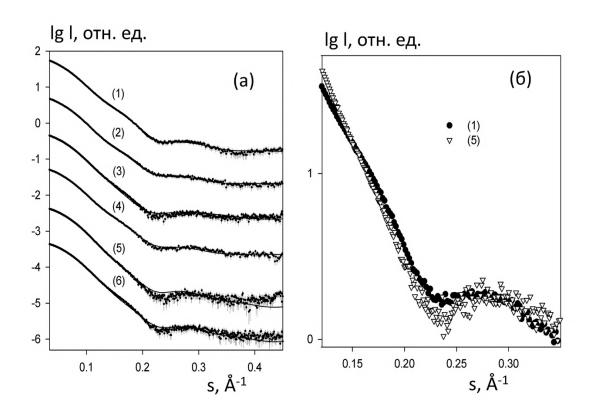


Рис. 4.5 (а) Кривые рассеяния МУРР фермента PGK. Эксперименты проводились в отсутствие субстратов (кривая 1), а также в присутствии 10 mM 3-PG (кривая 2), 10 mM ATP+12 mM $MgCl_2$ (кривая 3), 10 mM ADP+12 mM $MgCl_2$ (кривая 4), 10 mM 3-PG, 10 mM ATP+12 mM $MgCl_2$ (кривая 5) и 10 mM 3-PG, 10 mM 10 mM

MgCl₂ (кривая 6). Наилучшие приближения кристаллическими моделями были рассчитаны программой CRYSOL (сплошные линии). Последовательные кривые смещены вниз на одну логарифмическую единицу для лучшей визуализации. (б) Кривые МУРР от открытой (РGK без субстрата) (кривая 1) и закрытой (тройной комплекс PGK*3-PG*MgATP) форм фермента (кривая 5).

Следует отметить, что существуют примеры закрытия домена даже среди термостабильных ферментов, которые, как известно, обладают менее гибкой структурой. Таким примером 3-изопропилмалатдегидрогеназа является (сокращенно IPMDH) из термостабильной бактерии Thermus thermophilus (Tt). **IPMDH** относится классу бета-гидроксикислотных К окислительных декарбоксилаз и катализирует окисление и декарбоксилирование (2R,3S) -3изопропилмалата (IPM) до 2-оксо-4-метил-пентаноата в присутствии NAD и двухвалентного катиона (Mn^{2+}) или Mg^{2+} в биосинтетическом пути лейцина бактерий. Недавно опубликованные кристаллические структуры различных субстратных комплексов димерного Tt IPMDH выявили относительное движение двух доменов внутри каждой субъединицы и позволили идентифицировать два наиболее важных участка в междоменной области, с помощью которых происходит закрытие домена [142].

Ранее были получены кристаллографические модели аро Тt IPMDH дикого типа (PDB ID: 2Y3Z) (открытая форма) и его комплекса с субстратами Mn²⁺, IPM и NADH (PDB ID: 4F7I) (закрытая форма). Они представляют собой димеры с общей молекулярной массой около 74 кДа, при этом каждый мономер состоит из двух примерно равных по размеру доменов. Данные модели были использованы при моделировании данных МУРР, их теоретические кривые рассеяния от моделей были рассчитаны программой CRYSOL. Кривая от открытой формы фермента Тt IPMDH содержит характерное "плечо" в средней области углов, тогда как у закрытой формы фермента оно отсутствует (аналогично ферменту PGK, Рис. 4.5б). В работе [A49] было проведено систематическое исследование всех возможных комбинаций комплексов фермент-лиганд для Тt IPMDH. Экспериментальные кривые рассеяния различных комплексов Тt IPMDH в присутствии и в отсутствии

ионов металлов Mn^{2+} или Mg^{2+} показаны на Рис. 4.6а и 4.6б, соответственно. В Таблице 4.3 приведены вычисленные значения радиусов инерции R_g , а также степень закрытия домена фермента Tt IPMDH, выраженная в процентах от открытого и закрытого состояний, оцениваемых с помощью программы OLIGOMER. Как видно из Таблицы 4.3, данные MУPP свидетельствуют, по крайней мере, о частичном закрытии домена фермента в Tt IPMDH в комплексе с Mn*IPM и в стабильном (нефункционирующем) комплексе Mn*IPM*NADH. Однако интересно, что конформационное равновесие функционирующего комплекса Mn*IPM*NAD смещено в сторону открытой формы фермента. Кроме того, данные MVPP косвенно указывают на существенную роль иона металла в комплексообразовании с IPM: закрытие домена происходит исключительно при связывании металла с IPM, тогда как наличие не содержащего металл IPM не является достаточным условием для закрытия домена.

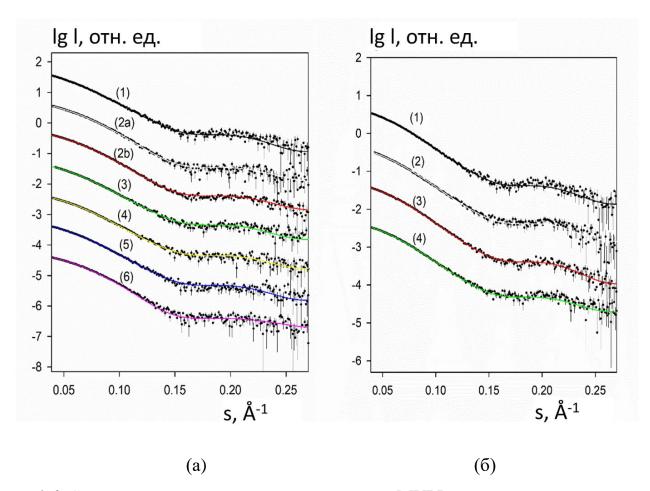


Таблица 4.3 Сравнение экспериментальных данных МУРР фермента IPMDH без субстрата и в присутствии субстратов с кривыми рассеяния, рассчитанными от кристаллографических моделей с помощью программ CRYSOL и OLIGOMER.

Эксперимент МУРР				Невязка <i>х</i> ²	Объемные доли открытой/закрытой форм фермента		
Фермент IPMDH + субстраты	R_g (эксперимент), Å $^{ m a}$		Закрытая форма	Открытая форма	Смесь открытой и закрытой форм	$V_{closed},\%$	Vopen, %
	GNOM	Гинье					
без субстрата	28.6±0.2	28.7±0.3	2.21	1.08	1.08	0	100
Mn	28.5±0.2	28.6±0.3	2.29	1.15	1.15	0	100
Mn+IPM	27.5±0.2	27.6±0.3	1.28	1.36	1.09	55±5	45±5
Mg+IPM	27.6±0.2	27.7±0.3	1.41	1.45	1.07	50±5	50±5
Mn+NAD	28.2±0.2	28.3±0.3	1.80	1.09	1.04	10±5	90±5
Mn+NADH	28.1±0.2	28.2±0.3	1.74	1.06	1.01	12±5	88±5
Mn+IPM+ NAD	28.1±0.2	28.2±0.3	1.65	1.06	1.03	10±5	90±5
Mn+IPM+ NADH	27.5±0.2	27.6±0.3	1.47	1.78	1.08	55±5	45±5
EDTA	28.4±0.2	28.4±0.3	2.10	1.07	1.07	0	100
EDTA+IPM	28.4±0.2	28.4±0.3	2.16	1.03	1.03	0	100
EDTA+IPM+ NAD	28.3±0.2	28.4±0.3	2.18	1.05	1.05	0	100

Наилучшие приближения с минимальными значениями невязки χ^2 выделены жирным шрифтом.

В ряде случаев мутации аминокислотных остатков, относящихся к активному сайту фермента, также могут влиять на его функциональные свойства. Так в работе [A50] роль активного остатка сайта, E270 фермента Тt IPMDH была исследована посредством проведения сайт-направленного мутагенеза. Образец-мутант E270A проявляет в значительной степени пониженную ($\sim 1\%$) каталитическую активность и незначительную степень активации при добавлении ионов K^+ по сравнению с ферментом дикого типа. Эксперименты МУРР и флуоресцентно-резонансного переноса энергии (FRET) действительно выявили существенную роль остатка E270

в стабилизации активной (закрытой) конформации фермента, в этом случае даже добавление ионов К⁺ не способствует закрытию домена фермента Тt IPMDH. Результаты анализа данных МУРР, приведенные в Таблице 4.4, ясно указывают на сохранение образцом-мутантом E270A открытой конформации даже в четвертичном комплексе фермент-лиганд Mn-IPM-NADH-IPMDH.

Таблица 4.4. Сравнение экспериментальных данных МУРР фермента IPMDH в присутствии лиганда Mn-IPM-NADH в нативном состоянии и с наличием мутации E270A (в присутствии и отсутствии ионов K^+) с кривыми рассеяния, рассчитанными от кристаллографических моделей открытой и закрытой формы фермента и их смеси.

Эксперимент МУРР			Невязка χ^2			Объемные доли открытой/закрытой форм фермента	
Фермент Tt - IPMDH в присутствии лиганда Mn^{2+} -		Закрытая форма	Открытая форма	Смесь открытой и закрытой форм	$V_{closed},\%$	$V_{open,}$ %	
IPM-NADH	GNOM	Гинье					
wild-type (no K ⁺)	27.4±0.2	27.4±0.3	1.18	1.19	1.09	53±3	47±3
wild-type (50 mM K ⁺)	27.2±0.2	27.3±0.3	1.17	1.20	1.11	58±3	42±3
E270A mutant (no K ⁺)	28.3±0.2	28.4±0.3	2.03	1.15	1.15	0	100
E270A mutant (1 M K ⁺)	28.4±0.2	28.5±0.3	1.71	1.08	1.08	0	100
R_g (теорет.), Å			26.67	28.45			
ММ, кДа			74.3	73.8			

Значения R_g вычисляли двумя альтернативными способами, используя программу GNOM и приближение Гинье, соответственно. Кристаллическая структура открытой формы Tt IPMDH (PDB ID: 4F7I) и закрытой формы комплекса Tt IPMDH-Mn-IPM-NADH (PDB ID: 2Y3Z). Минимальные значения невязки χ^2 (выделены жирным шрифтом) указывают на наилучшие приближения данных МУРР. Небольшие изменения в молекулярной массе MM обусловлены различным количеством остатков, разрешенных в кристаллических структурах.

В работе [А51] были продолжены исследования влияния мутаций аминокислотных остатков активного участка сайта фермента Тt IPMDH на закрытие его домена. Методом МУРР была протестирована роль ключевых аминокислотных остатков активного сайта K185, Y139, D217, D241, D245 и N102. Радикальное уменьшение каталитической активности (0.06% по сравнению с активностью фермента дикого типа) наблюдалось для K185A и D241A мутантов. Другие мутанты сохраняют некоторую (1-13%) каталитическую активность фермента дикого типа. При этом рН-зависимость активности фермента дикого типа (рН = 7.4) смещается в сторону более высоких значений (рН = 8.4 и 8.5 соответственно) для мутантов K185A и D241A. Для других мутантов наблюдались меньшие изменения. Измерения МУРР показали полное отсутствие закрытия домена в случае D217A и D241A мутантов, в то время как для других мутантов можно было обнаружить только частичное закрытие домена (Таблица 4.5). Это убедительно показывает, что те же самые остатки, которые важны для катализа, также необходимы для индуцирования закрытия домена фермента Тt IPMDH.

Таблица 4.5. Сравнение экспериментальных данных МУРР фермента IPMDH в присутствии лиганда Mn-IPM-NADH в нативном состоянии и с наличием точечных мутаций K185A, Y139A, D217A, D241A, D245A, N102A, E270A с кривыми рассеяния, рассчитанными от кристаллографических моделей открытой и закрытой формы фермента и их смеси. Остальные обозначения такие же, как в Таблице 4.4.

Эксперимент МУРР						Объемные доли	
			Невязка	χ^2	открытой/закрытой		
						форм фермента	
Фермент <i>Тt</i> -			Закрытая	Открыта	Смесь		
IPMDH в			форма	я форма	открытой		
	R_g	, Á ^a			И	$V_{closed,}$ %	$V_{open,}$ %
присутствии					закрытой		
лиганда Mn ²⁺ -					форм		
IPM-NADH	GNOM	Гинье					
Дикий тип	27.4±0.2	27.4±0.3	1.18	1.19	1.09	53±3	47±3
К185А мутант	28.1±0.2	28.2±0.3	2.26	1.19	1.15	15±3	85±3
Ү139А мутант	27.6±0.2	27.7±0.3	1.53	1.28	1.22	46±3	54±3
D217A мутант	28.4±0.2	28.5±0.3	2.41	1.22	1.22	0	100
D241A мутант	28.4±0.2	28.5±0.3	2.49	1.26	1.26	0	100
D245A мутант	27.9±0.2	28.0±0.3	2.16	1.41	1.32	32±3	68±3
N102A мутант	28.2±0.2	28.3±0.3	2.32	1.36	1.16	22±3	78±3
Е270А мутант	28.4±0.2	28.5±0.3	1.71	1.08	1.07	5±3	95±5
R_g (теорет.), Å			26.67	28.45			
ММ, кДа			74.3	73.8			

Значения R_g вычисляли двумя альтернативными способами, используя программу GNOM и приближение Гинье, соответственно. Кристаллическая структура открытой формы Tt IPMDH (PDB ID: 4F7I) и закрытой формы комплекса Tt IPMDH-Mn-IPM-NADH (PDB ID: 2Y3Z). Минимальные значения невязки χ^2 (выделены жирным шрифтом) указывают на наилучшие приближения данных МУРР. Небольшие изменения в молекулярной массе MM обусловлены различным количеством остатков, разрешенных в кристаллических структурах.

Таким образом, приведенные примеры исследований активной (закрытой) и неактивной (открытой) форм ферментов РСК и Тt IPMDH с помощью метода МУРР демонстрируют его высокую эффективность в получении детальной информации о равновесном состоянии различных комплексов фермент-лиганд и позволяют объяснить роль аминокислотных остатков активного сайта в конформационных переходах фермента.

3.3.2 Олигомерные смеси, находящиеся в динамическом равновесии

Как уже было сказано, для равновесных белковых смесей анализ малоугловых данных сводится к нахождению объемных долей компонентов при условии наличия информации об их структуре. Такие системы образуются, например, при частичной диссоциации олигомеров в растворе. Так, в работе [А52] был исследован белок трикорн протеаза, который является основным компонентом при протеолизе (процессе расщепления белков до пептидов), осуществляемого протеасомой. Было показано, что гексамер этого белка, который образуется в кристалле, частично диссоциирует в растворе на мономеры и димеры (Рис. 4.7а). По экспериментальной кривой рассеяния от трикорн протеазы (концентрация 2 мг/мл) (Рис. 4.76, кривая 1) можно оценить средний радиус инерции R_g =(5.9±0.2) нм и максимальный размер частиц D_{max} =(19 \pm 1) нм. Эти значения сопоставимы со значениями, рассчитанными от кристаллической структуры гексамерного трикорна (5.95 нм и 19 нм, соответственно). Однако теоретическая кривая, рассчитанная от гексамерной модели, не соответствует экспериментальным данным (невязка χ^2 =7.15, Рис. 4.76, кривая 2). Кроме того, рассчитанная по инварианту Порода молекулярная масса раствора, (380±40) кДа, значительно ниже, чем теоретическая масса гексамера (720 кДа), что также указывает на частичную диссоциацию белка в растворе. Действительно, OLIGOMER приблизить программа позволила хорошо экспериментальные данные МУРР (невязка χ^2 =1.08, Рис. 4.76, кривая 3) и оценить объемные доли мономеров (53%), димеров (14%) и гексамеров (33%) в растворе белка трикорн протеазы.

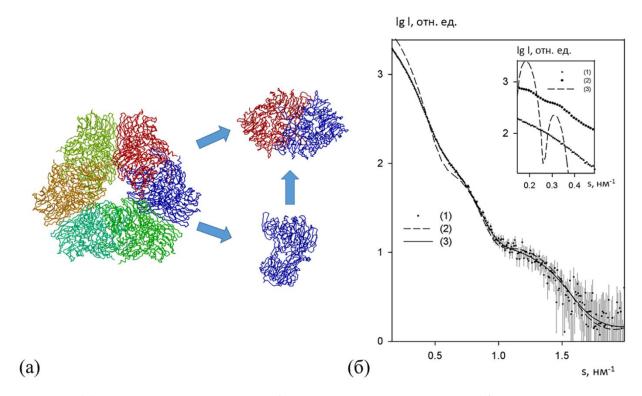


Рис. 4.7 (а) Схема диссоциации белка трикорн протеазы (б) Экспериментальные данные МУРР от раствора белка трикорн протеазы (2 мг/мл) (кривая 1), профиль рассеяния от гексамерной кристаллической модели трикорна (кривая 2), (3) наилучшее приближение смесью мономеров (объемная доля 53%), димеров (объемная доля 14%) и гексамеров (объемная доля 33%), полученное программой OLIGOMER. На вставке изображены кривые рассеяния в области самых малых углов от раствора белка трикорна с концентрацией 4.0 и 10.0 мг/мл (кривые 1 и 2, соответственно) и рассчитанная кривая рассеяния от сферической капсулы с диаметром D = 50 нм и толщиной стенок 7 нм, показывающая осцилляции на тех же угловых положениях, что и на кривых 2,3.

Помимо этого, данные МУРР от трикорн протеазы имеют ярко выраженную и очень специфическую зависимость от концентрации растворенного вещества. В области самых малых углов кривые рассеяния белка, измеренные при более высоких концентрациях 4 и 10 мг/мл, содержат заметные осцилляции, отсутствующие для раствора с низкой концентрацией. Такое поведение наблюдалось в нескольких повторяющихся экспериментах (вставка на Рис. 4.7, кривые 1 и 2 соответственно). Такие осцилляции можно объяснить наличием очень больших изометрических частиц, как показано на вставке Рис. 4.7, кривая 3, вычисленной от полой сферической частицы с наружным и внутренним

диаметрами 50.0 и 36.0 нм соответственно. Интересно, что размер этой полой частицы хорошо соответствует общему размеру икосаэдрических супермолекулярных частиц, которые ранее наблюдались на крио-ЭМ [143]. Поэтому наши результаты показывают, что мономеры, димеры, гексамеры и гигантские сборки белка трикорн протеазы находятся в растворе в равновесии, зависящем от концентрации белка.

С другой стороны, олигомерные смеси могут образовываться и за счет обратного процесса - ассоциации мономеров в олигомеры более высоких порядков. Одним из простейших примеров таких систем является смесь мономер-димер, находящаяся в динамическом равновесии.

В работе [А53] методом МУРР были исследованы мономер-димерные смеси H_c домена-рецептора белка столбнячного нейротоксина TeNT. Молекулярная масса мономера H_c домена составляет около 50 кДа. С помощью гель-хроматографии были выделены мономерные, димерные и промежуточные (полидисперсные) фракции H_c домена в диапазоне концентраций $2.5 \div 18.0$ мг/мл. Экспериментальные кривые МУРР от различных фракций H_c домена показаны на Рис. 4.8а, а общие структурные параметры, вычисленные на основе данных МУРР, представлены в Таблице 4.6. Экспериментальная молекулярная масса (MM) мономерной фракции близка к теоретическому значению 53.9 кДа. Оценки радиуса инерции R_g и максимального размера частицы D_{max} также близки к значениям, вычисленным от кристаллографической модели мономера (PDB ID: 1FV2). Кроме того, кривая рассеяния, рассчитанная на основе данной модели, дает относительно хорошее соответствие экспериментальным данным с $\chi^2 = 1.97$ (Рис. 4.8а). Все это указывает на то, что мономерная (по гель-фильтрации) фракция H_c домена действительно в значительной степени соответствует мономерным состояниям белка в растворе.

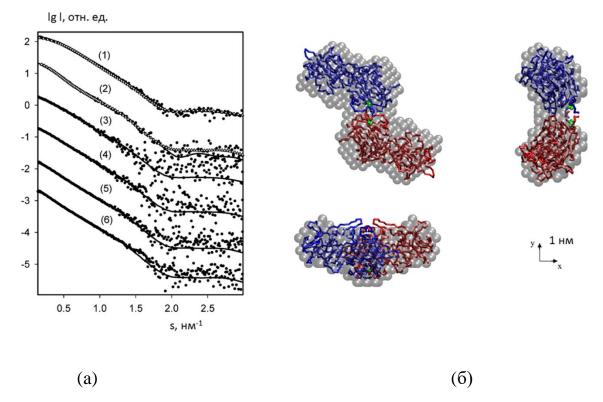


Рис. 4.8 (а) Экспериментальные данные МУРР от Н_с домена белка столбнячного нейротоксина TeNT (кружки - экспериментальные данные, сплошные линии рассчитанные приближения). Для мономерной фракции Н_с домена (кривая 1) кристаллографической приближение модели показано открытыми треугольниками, приближение от смеси мономеров и димеров - как сплошная линия. Для димерной фракции Н_с домена (кривая 2) сплошная линия приближение от ab initio модели, открытые треугольники - это приближение от димерной модели, полученной методом молекулярной тектоники. Кривые (3-6) представляют собой полидисперсные фракции Н_с домена для концентраций белка 2.5, 5.0, 10.0 и 18.0 мг/мл, соответственно, сплошные линии – наилучшие приближения от смеси мономеров и димеров, полученные программой OLIGOMER. Последовательные кривые смещены вниз на одну логарифмическую единицу для ясности. (б) Ab initio модель димера H_c домена, полученная программой DAMMIN (серые полупрозрачные сферы), наложена на атомную модель, восстановленную программой SASREF. Мономеры представлены синей и красной С_α цепями; остатки Суѕ689 в обоих мономерах, обозначены зелеными сферами. Вид сверху слева расположен вдоль оси симметрии второго порядка; правый и нижний виды повернуты против часовой стрелки на 90° вокруг вертикальной и горизонтальной осей, соответственно.

Таблица 4.6. Общие структурные параметры H_c домена белка столбнячного нейротоксина TeNT, рассчитанные по данным МУРР. R_g - радиус инерции, D_{max} - максимальный размер частиц, V_{Porod} - исключенный объем частицы, MM_{exp} - молекулярная масса, v_{mon} и v v_{dim} - объемные доли мономеров и димеров, соответственно, полученные программой OLIGOMER. Для полидисперсных образцов H_c домена показаны невязки χ^2 от приближений смесью мономер-димер с помощью программы OLIGOMER; для мономерной фракции H_c домена - от кристаллической модели и смеси мономер/димер; для димерной фракции H_c домена - от ab initio модели и димерной модели, полученной программой SASREF.

H _C домен	R_g , нм	$D_{max},$ HM	V_{Porod} , HM 3	<i>MM_{exp}</i> , кДа	v _{mon} ,	v_{dim} ,%	χ^2
Мономер	2.90±0.05	13±1	76±7	55 ±5	90.1	9.9	1.04 / 1.01
Димер	4.10±0.05	9.5±1.0	175±10	80 ±10	0.0	100	1.97 / 1.59
Полидисперсная							
фракция							
(2.5 мг/мл)	3.40±0.10	13.0±1	96±10	62 ±10	64.2	35.8	1.06
Полидисперсная							
фракция							
(5.0 мг/мл)	3.80 ± 0.10	13.0±1	116±10	68 ±10	42.6	57.4	1.28
Полидисперсная							
фракция							
(10.0 мг/мл)	3.90±0.05	13.0±1	140±10	75 ±10	20.5	79.5	1.89
Полидисперсная							
фракция							
(18.0 мг/мл)	4.00±0.05	13.0±1	155±10	77 ±10	14.1	85.9	2.05

Экспериментальная молекулярная масса MM образца димерной фракции H_c домена в Таблице 4.6 почти в два раза больше, чем у мономерной фракции, что указывает на образование димеров в растворе. Это также подтверждается оценкой для

исключенного гидратного объема образца (175 \pm 10) нм 3 , поскольку для глобулярных белков гидратированный объем в нм³ должен быть в численном выражении примерно в 1.5÷1.7 раза больше молекулярной массы, выраженной в кДа [A8]. Значения R_g и D_{max} в Таблице 4.6 значительно превышают значения для мономерной фракции Нс домена и указывают на удлиненную структуру частиц. В совокупности все эти данные показывают, что в растворе присутствуют димеры. Ab initio форму димера с низким разрешением восстановили с использованием программы DAMMIN в присутствии оси симметрии второго порядка. Типичная модель на Рис. 4.8В имеет полный размер приблизительно 4 x 8 x 13 нм³ и хорошо приближает экспериментальные данные (невязка $\chi^2 = 1.04$) (кривая 2, Рис. 4.8a). В полученной ab initio модели можно ясно различить две мономерные части, образующие димер Н_с домена с общей зигзагообразной конфигурацией. Более детальная модель димера была получена методом молекулярной тектоники с использованием двух моделей мономеров с высоким разрешением. Несколько серий запусков программы SASREF дали два типа моделей, совместимых с ab initio моделью, но с различным интерфейсом димеризации (один через С-концевую область, другой - через N-конец домена). Затем нами была учтена информация направленного мутагенеза, согласно которой димеризация должна опосредована дисульфидной связью через аминокислотный остаток Cys689. Близость остатка Cys689 к димерному интерфейсу была включена в моделирование в качестве ограничения, и в этих условиях программа SASREF позволила однозначно восстановить модель димера H_c домена, показанную на Рис. 4.8б. Для полидисперсных фракций H_c домена экспериментальные кривые рассеяния сильно различаются между собой (Рис. 4.8а). Как видно из Таблицы 4.6, общие структурные параметры находятся в диапазоне значений между мономерной и димерной фракциями Н_с домена, причем радиусы инерции увеличиваются с ростом концентрации белка. Поэтому логично предположить, что полидисперсные фракции Н_С домена представляют собой смеси мономерных и димерных состояний. Чтобы количественно описать равновесие мономер-димер, была

использована программа OLIGOMER. В результате было найдено, что объемная доля мономеров, обеспечивающая наилучшее приближение экспериментальным данным МУРР для полидисперсного образца, составляет 65% при 2.5 мг/мл, но неуклонно снижается с увеличением концентрации (Таблица 4.6) до 14% при 18 мг/мл. Интересно, что содержание 10% димеров несколько улучшило приближение для данных МУРР от мономерной фракции H_c домена ($\chi^2 = 1.59$), что указывает на образца. начальную стадию димеризации ЭТОГО Полученные данные свидетельствуют о важности учета олигомерных состояний белка столбнячного нейротоксина TeNT, которые могут образовываться в процессе интоксикации живых организмов.

В общем случае, образование белкового комплекса представляет собой, по меньшей мере, двухстадийный процесс, в котором динамический комплекс, образованный за счет случайных столкновений («встреч»), предшествует стабильному финальному комплексу. Предполагается, что «динамическое» состояние комплекса существует недолгое время и быстро переходит в стабильное конечное состояние. Однако для ряда слабо связанных белковых комплексов такое состояние, образованное за счет случайного соударения ("встречи") двух белков, может стать доминирующей (основной) формой. Это подтверждается данными ЯМР, когда для таких комплексов средние возмущения химического сдвига, наблюдаемые комплексообразовании, при оказываются очень аминокислотные остатки одного белка, вовлеченные в процесс образования комплекса, распределены по значительной части поверхности второго белка [144, 145].

В работе [A54] была исследована динамическая природа образования комплекса адренодоксина (Adx) с митохондриальным цитохромом с (Сс) с помощью методов МУРР и ЯМР. Адренодоксин (Adx) - белок с молекулярной массой 14.4 кДа, относящийся к семейству 2Fe-2S ферродоксинов. Он участвует в биосинтезе стероидных гормонов, действуя в качестве электронного челнока между NADPH-зависимой адренодоксин редуктазой и несколькими цитохромами Р450 [146]. *In*

vitro перенос электронов от адренодоксин редуктазы до Adx часто контролируют с помощью быстрого последующего переноса электронов от Adx до митохондриального цитохрома с (Сс). Молекулярная масса цитохрома с составляет около 12.5 кДа.

В данном исследовании помимо нативного комплекса был также сконструирован комплекс CL (Cross-Linked) с искусственно внедренной межмолекулярной дисульфидной связью для использования в качестве контрольного комплекса, обладающего «жесткой» структурой. В этом случае белки сшивали, используя дисульфидный мостик между аминокислотными остатками цистеина цитохрома C (V28C) и неполной формой адренодоксина Adx (остатки 4-108) (L80C) [147].

Метод МУРР был использован для анализа олигомерного состава комплексов Cc/Adx в растворе. В буфере, содержащем 10мМ Hepes при рН 7.4, общие структурные параметры CL комплекса (Таблица 4.7) указывают на то, что он остается связанным и гетеродимерным при всех исследовавшихся концентрациях образца (от 2.4 до 24 мг/мл) независимо от присутствия хлорида натрия в растворе. Ab initio форма комплекса, восстановленная по данным МУРР (Рис. 4.9a, кривая 1 и Рис. 4.10), хорошо соответствует модели, восстановленной методом ЯМР [А54]. В то же время, нативный комплекс Cc/Adx демонстрирует совершенно другое поведение, сильно зависящее от концентрации образца и ионной силы раствора в 10 мМ буфере Hepes. Так, экспериментальная кривая при 200 мМ NaCl может быть хорошо приближена рассеянием от смеси отдельных белков Сс и Adx (Рис. 4.9a, кривая 2), что подтверждает электростатический характер нативного комплекса, который диссоциирует при более высоких концентрациях NaCl. Снижение концентрации соли способствует образованию гетеродимерных комплексов, но также наблюдаются и более высокие олигомеры (Таблица 4.7). В растворах, не содержащих хлорид натрия, при низкой концентрации белка (с = 2.4 мг/мл) структурные параметры близки к параметрам гетеродимерного комплекса (Таблица 4.7). Максимальный размер комплекса растет с концентрацией белка (Таблица 4.7), и при 24 мг/мл в растворе присутствуют только тетрамеры, полученные ab initio форма и атомная модель, составленная из двух димеров, пристыкованных друг к другу бок о бок (модели получены программами DAMMIN и SASREF, соответственно) хорошо приближают экспериментальные данные (Рис. 4.9а, кривая 3)). Для получения количественных оценок олигомерного равновесия экспериментальные данные МУРР приближались смесью различных олигомеров Сс и Adx с помощью программы OLIGOMER. Эти олигомеры включали в себя мономеры, гетеродимеры, тетрамеры, а также гипотетические тримеры, созданные путем удаления одного мономера из найденной тетрамерной модели, и именно такие смеси действительно обеспечили хорошее соответствие расчетных кривых экспериментальным данным при различных концентрациях белков (Рис. 4.96, кривые (1-4)). Интересно, что присутствие значительной доли тримеров было выявлено для всех концентраций нативного белка, и попытки приблизить данные без учета тримеров значительно ухудшали качество приближений (Таблица 4.7). Таким образом, было установлено, что образование тетрамеров происходит именно через тримеры, а не посредством ассоциации двух гетеродимеров. Такой механизм ожидался бы для случая специфической димеризации и тетрамеризации, тогда как фактически наблюдаемая олигомеризация предполагает образование стохастического «столкновительного» комплекса. Этот результат дополнительно подтверждается тем фактом, что искусственно «сшитый» гетеродимерный комплекс CL не образует тетрамеров даже при высоких концентрациях белков. Повидимому, гетеродимеры CL не обладают подвижностью в растворе, необходимой для эффективного формирования междоменных связей через столкновительный механизм.

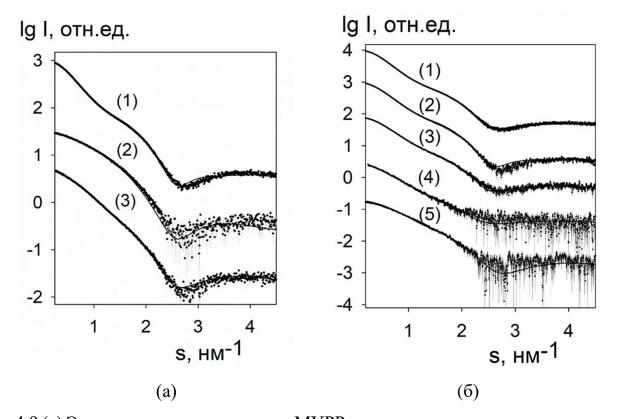


Рис. 4.9 (а) Экспериментальные данные МУРР от нативных и искусственно сшитых дисульфидной связью (CL) комплексов белков цитохром С - адренодоксин (кружки: эксперимент, сплошные линии: рассчитанные приближения). Кривая (1) - CL комплекс с приближением от модели гетеродимера комплекса (в соотношении 1:1), полученного программой SASREF; кривая (2) нативный комплекс в буфере Hepes 200 мМ NaCl и приближение от эквимолярной смеси цитохрома С и адренодоксина; кривая (3) - нативный комплекс с концентрацией 24 мг/мл и приближение от модели гетеротетрамера комплекса (в соотношении 2:2), полученного программой SASREF. На кривых (1) и (3) показаны только приближения от моделей молекулярной тектоники, приближения от ab initio моделей практически с ними совпадают. Кроме того, кривая рассеяния от модели CL-комплекса, восстановленной по ЯМР данным, графически неотличима от рассеяния моделью, полученной методом молекулярной тектоникой по данным МУРР (кривая 1). (б) Данные МУРР от нативного комплекса Cc/Adx в растворах, не содержащих хлорид натрия. Кривые (1) - (4) соответствуют концентрациям комплекса 24, 12, 6.0 и 2.4 мг/мл в буфере 10 мМ Нерез, рН 7.4, соответственно, кривая (5) соответствует концентрации комплекса 2.4 мг/мл в 20 мМ буфере фосфата калия рН 7.4). Экспериментальные данные изображены в виде кружков с интервалами ошибок; наилучшие приближения программы OLIGOMER показаны сплошными линиями, они соответствуют смесям мономеров, гетеродимеров, тримеров и тетрамеров с объемными долями, указанными в Таблице 4.7 (для кривых (1-4)). Для кривой (5) наилучшим приближением является гетеродимерный комплекс (в соотношении 1:1). Последовательные кривые смещены вниз на один логарифмический порядок для наглядности.

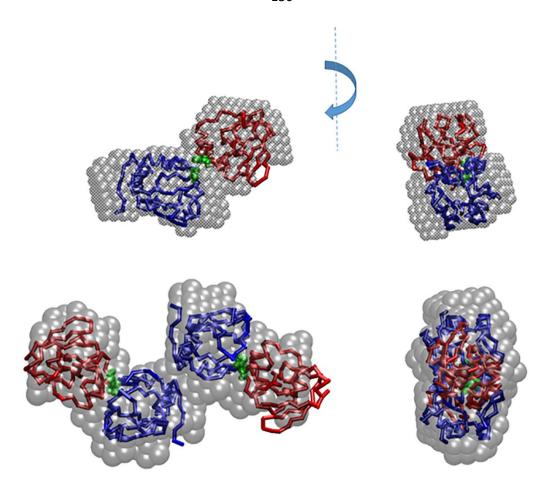


Рис. 4.10 Верхний ряд: *ab initio* модель "сшитого" СL комплекса белков цитохром С - адренодоксин, полученная программой DAMMIN (серые полупрозрачные сферы), наложенная на атомную модель гетеродимера, восстановленную программой SASREF. Модели цитохрома С и адренодоксина представлены в виде синих и красных C_{α} цепей, соответственно. Остатки Val28 цитохрома С и Leu80 адренодоксина указаны зелеными сферами. Правый вид повернут на 90° по часовой стрелки вокруг вертикальной оси. Нижний ряд: *ab initio* модель (полученная программой DAMMIN) нативного комплекса белков цитохром С - адренодоксин при высокой концентрации (24 мг/мл) в буфере, не содержащим хлорид натрия, совмещенная с атомной моделью гетеротетрамера, найденной методом молекулярной тектоники (программа SASREF). Представления и цветовые коды моделей такие же, как на верхнем ряду рисунка.

Таблица 4.7 Структурные параметры нативного комплекса цитохром C – адренодоксин по данным МУРР.

Нати	вный комп	лекс, буфе	p 10 mM H	epes,	CL	Cc/Adx
		pH 7.4			комплекс	(ЯМР)
c, мг/мл	24	12	6	2.4	3-12	-
R_g , нм	2.83±0.07	2.83±0.07	2.65±0.05	2.44±0.07	2.14±0.05	2.04
D_{max} , нм	9.0±0.5	9.0±0.5	9.0±0.5	8.0±0.5	8.0±0.5	7.0
V_p , HM^3	63±6	52±5	43±5	35±4	42±5	45
<i>ММ</i> , кДа	44±5	42±5	35±4	25±4	22±3	24.3
χ^2_{MDT}	2.45	1.72	1.56	1.25	-	-
χ^2_{MTT}	1.84	1.47	1.41	1.17	-	-
χ^2_{MDTT}	1.84	1.43	1.35	1.15	1.45	-
$V_{\scriptscriptstyle MOH,}\%$	0	0	6±5	24±5	0	-
$V_{\scriptscriptstyle{\partial \mathit{UM}}}$ %	0	8±5	25±5	24±5	100	-
$V_{mpu\scriptscriptstyle M}$,%	48±5	47±5	54±5	52±5	0	-
$V_{memp,}\%$	52±5	45±5	15±5	0	0	-

Обозначения: R_g , радиус инерции; D_{max} , максимальный размер частицы, V_p , исключенный объем гидратированной частицы, MM, молекулярная масса, χ^2 обозначает невязку наилучших приближений к экспериментальным данным от модельных смесей: мономеры/димеры/тетрамеры χ^2_{MDT} , мономеры/тримеры/тетрамеры χ^2_{MDT} , и мономеры/димеры/тримеры/тетрамеры χ^2_{MDTT} , V_{MOH} , $V_{\partial uM}$, V_{mpuM} и V_{memp} —соответствующие объемные доли олигомерных компонентов смеси. В крайнем правом столбце приведены значения структурных параметров, вычисленные на основе ЯМР модели, полученной для искусственно "сшитого" комплекса.

Таким образом, метод МУРР позволил оценить степень олигомеризации белкового комплекса адренодоксин-цитохром с, оказавшуюся различной для искусственно

"сшитого" и нативного комплекса. Результаты анализа данных МУРР показывают, что этот комплекс является полностью динамическим и образуется исключительно вследствие случайно-«столкновительного» характера взаимодействий белков. Данные ЯМР также указывают на динамический характер образования комплекса, подтверждая, что белок Сс должен "охватывать» большую площадь поверхности белка Adx, чтобы уменьшить межмолекулярные химические сдвиги до незначительных значений, согласуется большим интерфейсом ЧТО \mathbf{c} взаимодействия белков, определяемым возмущениями химического сдвига [А54]. Олигомерная смесь раствора макромолекул может иметь и более сложный состав, определяемый свойствами димеризационных интерфейсов соответствующих белков, а также наличием точечных мутаций в данных областях. Одним из примеров таких сложных систем является белок Ataxin-1, ответственный за наследственную нейродегенеративную болезнь атаксия типа 1, связанную с агрегацией белка и его неправильным сворачиванием. В работе [А55] методом МУРР были исследованы структурные последствия двух типов мутаций АХНдомена белка Ataxin-1, каждый из которых предназначен для нарушения специфических межмолекулярных интерфейсов. Первая мутация, А567G влияет на остаток, расположенный на границе раздела димеризации. Во втором наборе были введены одиночные и двойные аланиновые мутации I580 и L588, которые могут влиять на равновесие в смеси мономер-димер и димер-тетрамер. Молекулярная масса мономера данного белка составляет около 14 кДа.

Эксперименты МУРР были проведены для нативного белка wt-АХН и двух его точечных мутаций I580A-АХН и А567G-АХН в диапазоне концентраций белка от 0.04 до 1.35 мМ. Анализ данных МУРР показал, что нативный белок wt-АХН находится в динамическом равновесии между мономерными, димерными и тетрамерными состояниями с существенной долей мономерных частиц. Для I580A-АХН и А567G-АХН наблюдалось заметное увеличение доли мономерных частиц в соответствии с ролью обеих этих мутаций в дестабилизации димерного интерфейса. Наилучшие приближения, полученные программой OLIGOMER

показаны на Рис. 4.11, количественная оценка олигомерного состава АХН домена белка Ataxin-1 приведена в Таблице 4.8.

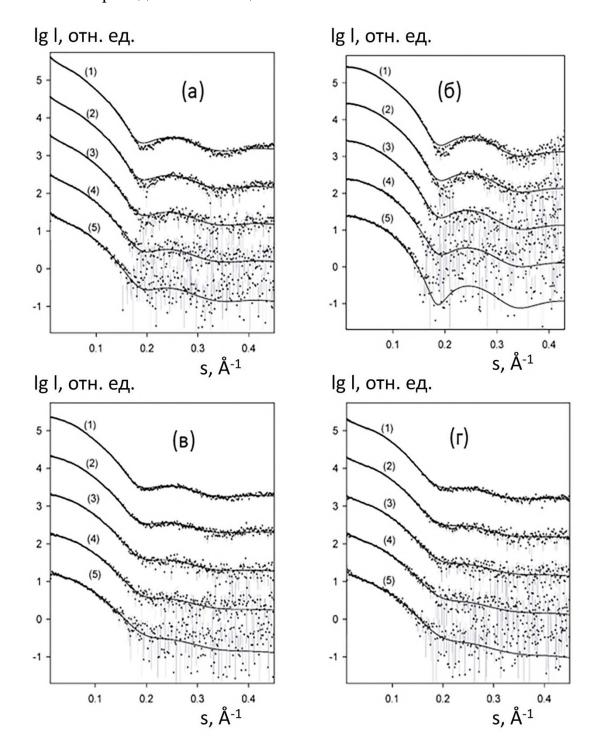


Рис. 4.11 Данные МУРР для нативного wt-АХН домена белка Ataxin-1 и его мутантых вариантов. (а) wt-АХН, измеренный после хранения в холодильнике и последующего размораживания (б) свежеприготовленный wt-АХН без замораживания (в) мутантный вариант A567G-AXH, (г) мутантный вариант I580A-AXH. Кривые 1-5 соответствуют концентрациям белка 1.35, 0.67, 0.33, 0.17 и 0.08

мМ в 20 мМ Tris-HCl pH 7.0, 2 мМ TCEP, соответственно. Экспериментальные данные показаны кружками, сплошные линии — наилучшие приближения для смесей мономеров, димеров и димеров, рассчитанных программой OLIGOMER. Последовательные кривые смещены вниз на одну логарифмическую единицу для ясности.

Таблица 4.8 Результаты приближения данных МУРР для оценки содержания тетрамеров, димеров и мономеров в нативном домене wt-AXH белка Ataxin-1 и его мутантных вариантов A567G-AXH и I580A-AXH при различных концентрациях образца, полученные программой OLIGOMER. Значения в скобках соответствуют свежеприготовленной партии образца wt-AXH.

образец	концентрация	мономеры	димеры	тетрамеры	χ^2
	(mM)	%	%	%	,,
wt-AXH	1.35	21 (28)	31 (55)	48 (17)	1.78 (1.76)
	0.67	27 (32)	33 (55)	40 (13)	1.53 (1.38)
	0.33	34 (30)	34 (60)	32 (7)	1.35 (1.23)
	0.17	41 (27)	34 (73)	25 (0)	0.91 (1.15)
	0.08	45 (25)	36 (75)	19 (0)	0.92 (1.33)
	0.04	47 (30)	36 (70)	17 (0)	0.99 (1.12)
A567G-AXH	1.35	37	42	21	1.53
	0.67	46	40	14	1.14
	0.33	56	33	11	1.03
	0.17	66	26	8	1.00
	0.08	75	16	9	0.98
	0.04	81	11	8	0.98
I580A-AXH	1.35	53	42	5	1.83
	0.67	60	33	7	1.26
	0.33	68	24	8	1.03
	0.17	76	15	9	0.95
	0.08	80	10	10	0.92
	0.04	80	12	8	0.91

Таким образом, если в кристалле АХН домен белка Ataxin-1 (который модулирует процесс агрегации белка Ataxin-1) содержится в виде тетрамера, то в растворе он образует сложную комбинацию из мономеров, димеров, тетрамеров и (возможно) олигомеров более высокого порядка. Такое поведение представляется необычным примером белка «хамелеона», мульти-конформационные состояния которого влияют на агрегацию белка и, таким образом, на протекание болезни.

Исследованиям олигомерного состава белков в условиях благоприятных для кристаллизации посвящен цикл работ с использованием методов МУРР и МУРН [A56-A67]. Согласно трехстадийной модели кристаллизации белков предшествует формирование промежуточной фазы кластеров-прекурсоров [148, 149]. Такие кластеры представляют собой трехмерные фрагменты кристаллической структуры, которые были экспериментально обнаружены МУРР/МУРН для ряда белков. С использованием методом имеющихся кристаллографических структур и программы OLIGOMER было показано, что для тетрагонального лизоцима кластеры-прекурсоры представляют собой октамеры [А56-А58, А66]; для термолизина - гексамеры [А63]; для аминотрансферазы додекамеры [А65], а для протеиназы К - димеры [А60]. Кроме того, полученные результаты были подтверждены независимым способом c помощью моделирования молекулярной динамикой [Абб, Аб7]. Олигомерный состав белков зависит от многих факторов, таких как температура, концентрация белка, концентрация и тип осадителя, состав буферного раствора, тип симметрии, образующегося кристалла. Объемная доля кластеров-прекурсоров растет с понижением температуры, а также при повышении концентрации белка и/или осадителя. Исследования в этой области активно продолжаются, упомянутый выше подход был недавно успешно применен и для анализа небиологических систем – раствора дигидрофосфата калия (KDP) [A68].

Представленные примеры не являются полным обзором всевозможных случаев различных структурных состояний белковых олигомерных смесей, находящихся в динамическом равновесии, но при этом наглядно демонстрируют большие возможности метода МУРР/МУРН и разработанных методов анализа для проведения количественной оценки олигомерного состава таких сложных систем.

3.3.3 Самоорганизующиеся многокомпонентные системы

Способность белков принимать различные четвертичные структуры является существенной и важной для многих биологических процессов, таких как передача сигнала, регуляция клеточного цикла и ферментативный катализ. Среди факторов, которые вызывают изменение четвертичной структуры, можно указать следующие - температурные сдвиги, изменение ионной силы и рН раствора, связывание лигандов. В некоторых случаях белковая инженерия, использующая мутации и геномные вставки, приводит к изменению структуры белка *in vitro* [150]. самособирающиеся Макромолекулярные структуры икосаэдрической \mathbf{c} симметрией имеют преимущество, заключающееся в том, что они состоят из сравнительно малого числа уникальных строительных блоков и при этом теоретически способны образовывать большое число стабильных состояний. Многие вирусные самосборки имеют оптимизированные межсубъединичные икосаэдрической контакты, согласно геометрическим принципам квазиэквивалентности. Каспар и Клюг [151] обнаружили, что квази-эквивалент икосаэдрической капсулы может быть сконструирован путем введения пентамеров в поверхностную решетку гексамеров, таким образом генерируя изогнутую оболочку с икосаэдрической симметрией. Число триангуляции (Т) указывает количество субъединиц в икосаэдрической асимметричной элементарной ячейке. Число субъединиц во всей икосаэдрической частице, таким образом, оказывается равным 60Т [152].

Фермент люмазинсинтаза (Lumazine synthase, сокращенно LS) катализирует образование 6,7-диметил- 8-рибитиллюмазина (6,7-dimethyl-8-ribityllumazine) на предпоследней стадии биосинтеза рибофлавина (витамина В₂) (Рис. 4.12) [153]. В организмах Saccharomyces cerevisiae, Schizosaccharomyces pombe, Magnaporthe grisea и Mycobacterium tuberculosis фермент формирует пентамеры, в то время как в Bacillus subtilis, Escherichia coli, Spinacia oleracea и Aquifex aeolicus, образуются икосаэдрические капсулы диаметром около 16 нм [154]. Икосаэдрические LS капсулы создаются 60 субъединицами, размещенными в 12 пентамерах, которые

составляют икосаэдрическую капсулу с числом триангуляции T=1 и с общей молекулярной массой около 960 кДа. Активные сайты LS расположены на интерфейсах двух соседних субъединиц внутри пентамера. В икосаэдральных LS активный участок расположен на внутренней поверхности капсулы вблизи оси симметрии 2-го порядка икосаэдра, которая относится к двум соседним пентамерным узлам.

Рис. 4.12 Синтез 6,7-диметил- 8-рибитиллюмазина (6,7-dimethyl-8-ribityllumazine) с помощью фермента Lymazine synthase.

В работе [A13] распределение частиц икосаэдрических капсул Lumazine synthase из *Bacillus subtilis* (LSBS) и *Aquifex aeolicus* (LSAQ) для нативных и мутантных образцов было исследовано с помощью синхротронного малоуглового рентгеновского рассеяния (МУРР) и крио-электронной микроскопии при различных значениях рН и в присутствии различных буферов.

Данные крио-ЭМ различных образцов нативных и модифицированных LS показали, что частицы LS образуют икосаэдрические капсулы с диаметрами в трех различных диапазонах: 16-20 нм (тип I), 27-30 нм (тип II) и 30-33 нм (тип III). Капсулы типа I состоят из 60 субъединиц с совершенной икосаэдрической симметрией (T = 1), тогда как диаметры капсул типа II и типа III хорошо

соотносятся с числами триангуляции T = 3 и T = 4, и состоят из 180 и 240субъединиц. соответственно. В фосфатном буфере рН 8.0, LSBS образует исключительно частицы типа I (Рис. 4.13a, Таблица 4.9). Анализ LSBS в фосфатных буферах при значениях рН между 7.0 и 9.0 показывает, что рН, по-видимому, не оказывает влияния на сборку капсул. В буфере Tris гидрохлорида pH = 8.0, 54% идентифицированных частиц нативного LSBS имеют сферически всех симметричную капсульную форму, другие частицы имеют неправильную форму и агрегированы, деформированы или неполны (Рис. 4.13б). Капсуло-образные частицы появляются в диапазонах диаметров, типичных для типа I (29%) и типа ІІ/ІІІ (25%). При рН 9.0 только 49% всех обнаруженных частиц имеют правильную капсульноподобную форму (Рис. 4.13в). Остальные частицы имеют некруглую форму, что указывает на деформацию или неполную капсульную оболочку. Симметрично образовавшиеся капсулы относятся к категории II/III типа. В боратном буфере pH = 7.0 LSBS образует как капсулы типа I (80%), так и капсулы типа ІІ/ІІІ (20%) (Рис. 4.13г). В фосфатном буфере при рН 7.0, 89% мутантных частиц LSBS Arg127Thr (R127T) (остаток Arg127 находится вблизи активного центра фермента) имеют круглую форму, в то время как незначительный процент (11%) имеет неправильную форму. При этом 75% частиц имеют круглую форму с диаметром, соответствующим капсулам типа ІІ/ІІІ, и 14% частиц соответствуют капсулам типа I (Рис. 4.13д). Остальные частицы неправильной формы имеют размеры, сравнимые либо с типом I, либо с типом II/III. Было обнаружено, что мутант LSAQ-IDEA формирует исключительно капсулы типа II/III как в Трис гидрохлоридном буфере pH = 8.5 (Рис. 4.13e), так и в фосфатном буфере pH = 7.0(Рис. 4.13ж).

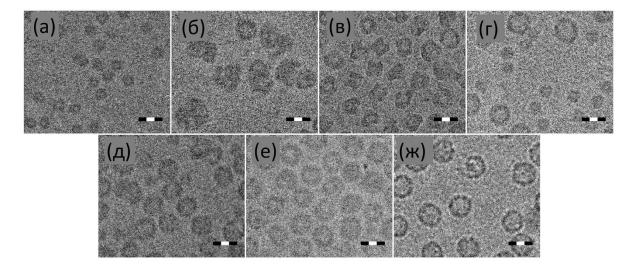


Рис. 4.13 Контрастные увеличенные крио-ЭМ изображения (а) LSBS фосфатный буфер, pH 8.0 (б) LSBS в Tris/HCl, pH 8.0 (в) LSBS в Tris/HCl, pH 9.0 (Γ) LSBS в боратном буфере, pH 7.0 (д) мутант LSBS R127T в фосфатном буфере, pH 7.0 (е) мутант LSAQ-IDEA в Tris/HCl, pH 8.5, (ж) мутант LSAQ-IDEA в фосфатном буфере, pH 7.0. Шкала соответствует 30 нм.

Таблица 4.9 Оценка популяций частиц по данным крио-ЭМ.

Кр	рио-ЭМ		Капсулы Тип I (16-20 нм)	Капсулы Тип II/III (27-33 нм)	Нерегулярные частицы
Образец	Буфер	pН	Доля [%]	Доля [%]	Доля [%]
LSBS WT	Phosphate	8.0	100	-	-
LSBS WT	Tris/HCl	8.0	29	25	46 ^a
LSBS WT	Tris/HCl	9.0	-	49	51
LSBS WT	Borate	7.0	80	20	-
LSBS R127T	Phosphate	7.0	14	75	11
LSAQ IDEA	Tris/HCl	8.5	-	100	-

^а большая часть нерегулярных частиц, наблюдаемых в этом образце, представляет собой агрегаты неповрежденных капсул различных размеров.

Эксперименты МУРР проводились с растворами LSBS в фосфатных, Tris гидрохлоридных и боратных буферах и с растворами LSAQ в фосфатных и Tris гидрохлоридных буферах. Данные МУРР от белковых растворов и соответствующих буферов были собраны в диапазоне значений рН между 6.0 и 10.0. На Рис. 4.14 показаны экспериментальные кривые рассеяния для ферментов LSBS и LSAQ, соответственно. Интегральные структурные параметры LS частиц (эффективный радиус инерции, R_g , молекулярная масса MM, а также максимальный размер частиц D_{max}) приведены в Таблице 4.10.

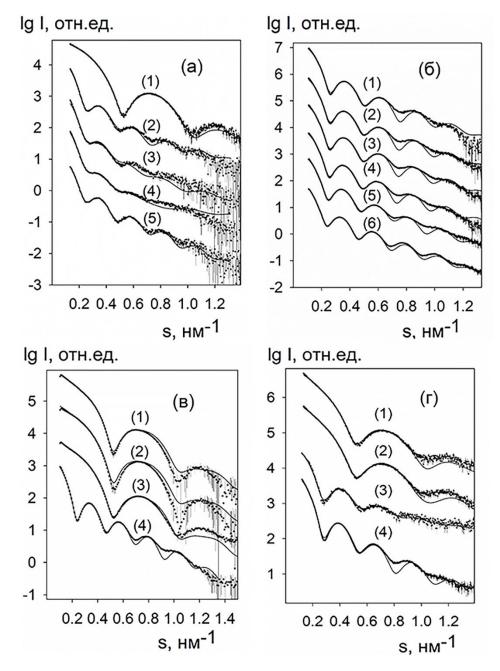


Рис. 4.14 Экспериментальные данные МУРР от фермента LSBS дикого типа, LSBS R127T мутанта, LSAQ дикого типа и LSAQ-IDEA мутанта (кружки) и наилучшие приближения (сплошные линии), полученные программой MIXTURE. (а) кривая рассеяния LSBS дикого типа (1) в фосфатном буфере при рН 6.0 и в Tris гидрохлоридном буфере (2-5) при рН 7.0, 7.6, 8.4 и 9.0, соответственно; (б) кривая рассеяния мутанта LSBS R127T (1) в фосфатном буфере при рН 6.0 и кривые (2-6) Трис гидрохлоридном буфере при рН 7.0, 7.5, 8.0, 8.5 и 9.0, соответственно; (в) кривые рассеяния LSBS дикого типа (1-4) в боратном буфере при рН 7.0, 8.0, 9.0 и 10.0, соответственно; (г) кривые рассеяния LSAQ дикого типа (1) в фосфатном буфере при рН 6.0; LSAQ дикого типа (2) в Трис гидрохлоридном буфере при рН 7.0; мутанта LSAQ-IDEA (3) в фосфатном буфере при рН 6.0; мутанта LSAQ-IDEA (4) в Трис гидрохлоридном буфере при рН 7.0.

Таблица 4.10. Структурные параметры смесей фермента люмазинсинтазы, вычисленные по данным МУРР. D_{max} , M_e , R_{ge} представляют собой максимальный размер, средние по ансамблю молекулярную массу и радиус инерции частицы.

Образец	Буфер	рН	D_{max}	M_e	R_{ge}
		1	НМ	кДа	НМ
LSBS, дикий тип	Phosphate,	6.0	30.0	595	6.9
, , , , , , , , , , , , , , , , , , ,	0.1M	6.5	30.0	658	7.1
		7.0	30.0	607	7.3
		7.5	30.0	510	6.9
		8.0	30.0	581	7.1
	Tris-HCl, 0.1M	7.0	32.0	856	12.8
	,	7.6	32.0	962	12.9
		8.4	32.0	936	12.8
		9.0	32.0	854	12.7
	Borate, 0.1M	7.0	16.0	1637	6.8
	,	8.0	16.0	1096	6.9
		9.0	16.0	814	6.8
		10.0	31.0	1608	12.7
LSBS, R127 мутант	Phosphate,	6.0	31.0	785	12.5
•	0.1M	6.5	31.0	730	12.4
		7.0	31.0	683	12.6
		7.5	31.0	805	12.5
		8.0	31.0	752	12.4
	Tris-HCl, 0.1M	7.0	31.0	854	12.5
		7.5	31.0	661	12.7
		8.0	31.0	715	12.9
		8.5	31.0	644	13.1
		9.0	31.0	623	13.3
LSAQ, дикий тип	Phosphate	6.0	32.0	806	10.0
		6.5	32.0	720	9.9
		7.0	32.0	687	9.8
		7.5	32.0	816	10.1
		8.0	32.0	855	10.0
	Tris-HCl	7.0	32.0	965	10.2
		7.5	32.0	732	10.3
		8.0	32.0	772	9.8
		8.5	32.0	820	10.0
		9.0	32.0	678	8.8
LSAQ-IDEA, мутант	Phosphate,	6.0	28.0	621	11.2
	0.1M	7.0	28.0	617	11.2
		8.0	28.0	888	11.1
	Tris-HCl, 0.1M	7.0	28.0	683	11.3
		7.5	28.0	949	11.2
		8.0	28.0	1182	11.2
		8.5	28.0	1386	11.3
		9.0	28.0	1670	11.3

Функции распределения по расстояниям p(r), вычисленные по экспериментальным данным МУРР с помощью программы GNOM, показаны на Рис. 4.15. Для смесей, содержащих множество различных частиц (компонентов), экспериментальная кривая рассеяния представляет собой линейную комбинацию рассеяния от отдельных компонентов, взвешенных по их относительным концентрациям (т.е. их объемным долям в растворе). Эффективные значения R_g и MM также являются средними по совокупности, взвешенной на объемные доли частиц, а p(r) функции смесей являются взвешенными суперпозициями распределения расстояний отдельных компонентов, тогда как D_{max} соответствует максимальному расстоянию самой большой частицы в смеси.

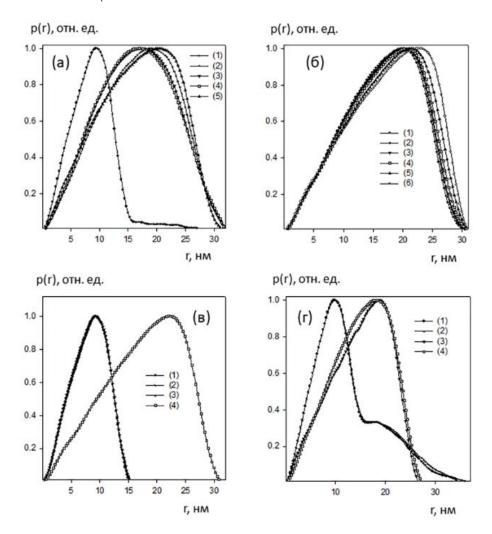


Рис. 4.15 Функции распределения p(r) ферментов LSBS (a) и LSAQ (б) дикого типа и мутантов LSBS R127T (в) и LSAQ-IDEA (г) Обозначения кривых такие же, как на Рис. 4.14. Все функции p(r) нормированы по максимуму на единицу.

Для частиц, имеющих форму капсул максимальный размер эквивалентен наружному диаметру капсулы, а форма функции p(r) позволяет сделать качественные выводы о составе смесей в различных буферах при различных значениях рН. Таким образом, LSBS дикого типа образует небольшие капсулы (тип I) в фосфатном буфере (Рис. 4.15а, кривая 1), в то время, как в Трис-буфере формируются большие капсулы (тип III) (Рис. 4.15а, кривые 2-5). В боратном буфере наблюдается переход от капсул типа I (рН от 7.0 до 9.0) к капсулам типа III при рН 10.0 (Рис. 4.15в). Растворы LSAQ дикого типа (Рис. 4.15г, кривые 1-2) содержат только небольшие капсулы (тип I). В то же время большие капсулы преобладают в растворах мутантов LSBS R127T (Рис. 4.15б) и LSAQ-IDEA (Рис. 4.15г, кривые 3-4) при всех исследованных условиях.

Функция p(r) от LSBS в фосфатном буфере практически не зависит от pH. Экспериментальные кривые рассеяния указывают на присутствие капсул LSBS с внешним диаметром приблизительно 16 нм, который соответствует размеру исосаэдрической капсулы T=1, наблюдаемой в кристалле и небольшом количестве больших частиц (приводящих к небольшому вкладу в межчастичные расстояния в диапазоне от 16 до 33 нм на Puc. 4.15а, кривая 1). В противоположность этому функции распределения по расстояниям p(r) LSBS в Tris гидрохлориде (Puc. 4.15а, кривые 2-5) указывают на наличие больших капсул с внешним диаметром около 32 нм. Интересно, что кривые рассеяния при pH 7.0 и 9.0 имеют выраженные максимумы, предполагающие довольно изометрические формы частиц, в то время как при pH 7.6 и 8.4 максимумы кажутся замазанными, что указывает на возможную полидисперсность или деформацию частиц. Таким образом, можно сделать вывод, что структурный переход LSBS происходит в Tris буфере при pH около 8.0.

Функция p(r) мутанта LSBS R127T в фосфатном буфере (Рис. 4.15б, кривая 1) указывает на образование капсул с диаметром около 30 нм с небольшой тенденцией к образованию более крупных частиц при увеличении рН. В противоположность этому функции распределения по расстояниям p(r) мутанта

LSBS R127T в буферах Tris гидрохлорида (Рис. 4.156, кривые 2-6) значительно меняются в зависимости от рН. При рН 7.0 функция p(r) аналогична LSBS R127T в фосфатном буфере (соответствующей частицам типа III), но при рН 7.5 и 8.0 появляются более крупные частицы диаметром около 32-33 нм (что соответствует более крупным частицам типа III). В боратном буфере (Рис. 4.15в) LSBS дикого типа имеется еще более выраженная зависимость образования капсул от рН. При значениях рН (7.0, 8.0 и 9.0) в растворе присутствуют главным образом мелкие частицы (тип I) (Рис. 4.15в, кривые 1-3), тогда как при рН 10.0 образуются в основном крупные частицы (тип III) (Рис. 4.15в, кривая 4).

Малоугловые кривые рассеяния и функции распределения по расстояниям от LSAQ дикого типа в фосфатных и Tris гидрохлоридных буферах очень похожи друг на друга и практически не имеют зависимости от рН (Рис. 4.15г, кривые 1-2). Профиль p(r) соответствует малым икосаэдрическим частицам типа I, хотя в растворе присутствует и небольшое количество более крупных частиц. Довольно интересные результаты были получены для мутанта LSAQ-IDEA, где эти четыре вставки остатков активного сайта фермента препятствуют формированию капсул типа I (так как при такой вставке IDEA два соседних пентамера в модели капсулы типа I будут сильно пространственно перекрываться друг с другом) и вместо этого приводят к образованию пентамеров [155]. Исследования каталитической активности мутанта LSAQ-IDEA выявили менее 0.5% активности по отношению к ферменту дикого типа. Однако довольно неожиданно было обнаружено, что вместо пентамеров мутант LSAQ-IDEA образует довольно крупные частицы как в фосфатном, так и в Tris гидрохлоридном буфере. Кривые рассеяния в двух буферах сходны друг с другом и мало зависят от рН (Рис. 4.14г, кривые 3-4). Рассчитанные функции p(r) (Рис. 4.15г, кривые 3-4) указывают на наличие полых капсул с D_{max} около 27.5 нм (капсулы типа II), который существенно больше T1 (тип I), но меньше диаметров наиболее крупных частиц (30-32 нм, тип III), образованного R127T мутантом LSBS.

Вышеприведенный анализ функций p(r) показывает, что как растворы LSBS, так и растворы LSAQ-IDEA являются смесями частиц различного размера. Кривые рассеяния, показывающие резкие минимумы, указывают на то, что частицы являются довольно изометричными, тогда как профили функций p(r) в случаях, когда доминирует один тип частиц (например, кривая 1 на Рис. 4.15а, кривая 4 на Рис. 4.15 в,г), имеют асимметричный вид с более резким спаданием при приближении к D_{max} , что характерно для полых частиц. Для дальнейшей оценки формы частиц и проверки того, являются ЛИ эти частицы в идеале икозаэдрическими, была использована программа DAMMIN, учитывающая икозаэдрическую симметрию, с помощью которой были восстановлены формы частиц LSBS в боратном буфере при pH 7.0 и pH 10.0, где преимущественно содержатся капсиды типа I и типа III, соответственно. Две восстановленные формы частиц (Рис. 4.16 а,в) качественно похожи на модели капсул типа I и типа III высокого разрешения, однако кривые рассеяния от этих моделей имеют систематические отклонения от экспериментальных данных (Рис. 4.17). Минимумы на теоретических кривых рассеяния от икосаэдрических структур выглядят резче, чем в экспериментальных данных. Это указывает на то, что данные модели слишком изометричные. Для подтверждения этого предположения модели икосаэдра на (Рис. 4.16 а,в) были дополнительно уточнены путем исключения ограничений, накладываемых симметрией. Полученные модели, представленные на Рис. 4.16 б,г, уже хорошо соответствуют экспериментальным данным МУРР (Рис. 4.17) и выглядят как полые икосаэдрические частицы, у которых отсутствуют отдельные грани, сходные по размеру с пентамерным строительным блоком капсулы. Аналогичные результаты были получены для данных МУРР от других образцов, что позволяет предположить, что самосборки фермента LS могут содержать значительные популяции неполных икосаэдрических частиц с отдельными гранями, отсутствующими в их структуре. Интересно, что частицы типа I (Рис. 4.16б) выглядят более изометричными (возможно, из-за меньшего количества диссоциированных блоков), чем частицы типа III (Рис. 4.16г).

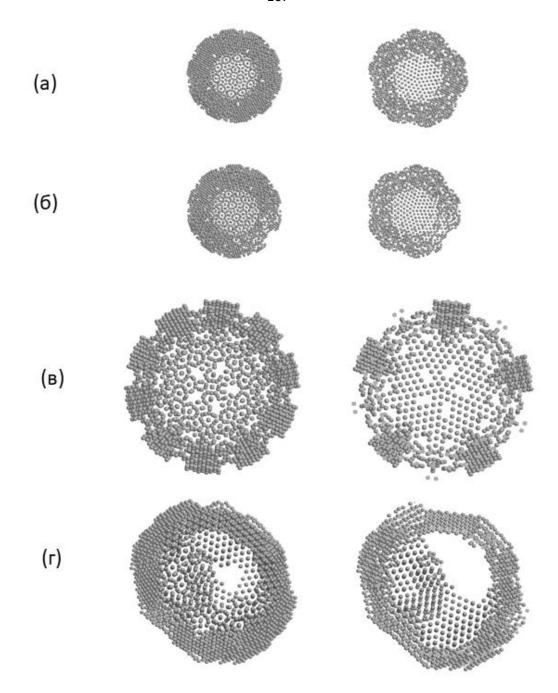


Рис. 4.16 Типичные *ab initio* модели нативного LSBS в боратном буфере при рН 7.0 (а,б) и рН 10.0 (в,г) восстановленные программой DAMMIN. Модели (а) и (б) были получены с ограничениями на икосаэдрическую симметрию, тогда как модели (в) и (г) имеют Р1 симметрию. Модели отображались с помощью графической программы RASMOL [156] в режимах space filling (левая сторона) и slab mode (разрез) (правая сторона).

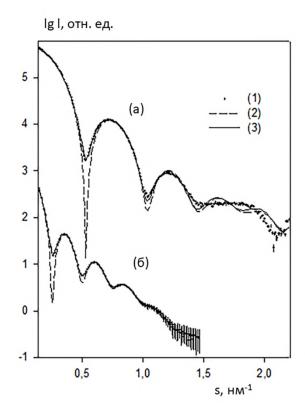


Рис. 4.17 Экспериментальные данные МУРР от LSBS дикого типа в боратном буфере (кружки) и наилучшие приближения, полученные при *ab initio* моделировании программой DAMMIN с ограничениями на икосаэдрическую симметрию (пунктирные кривые) и без ограничений на симметрию (сплошные кривые). (а) соответствует рН буфера 7.0, (б) соответствует рН буфера 10.0.

Приближения отдельными типами частиц, или полными икосаэдрическими капсулами, или неполными с (одной или несколькими) отсутствующими гранями не дала разумного согласия с экспериментальными данными. Поэтому было сделано предположение, что капсулы типа I и типа II или III могут сосуществовать в растворе, другими словами кривые рассеяния должны быть приближены смесью, содержащей несколько компонентов (Рис. 4.18).

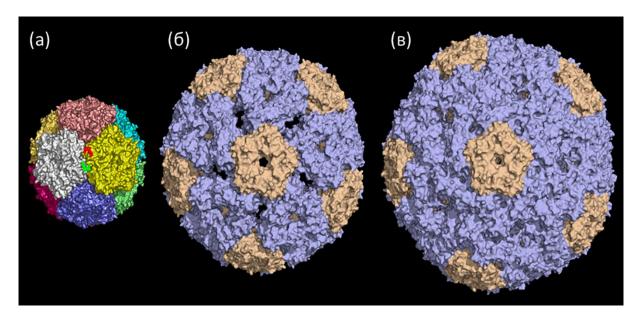


Рис. 4.18 (а) Кристаллическая структура капсул LSAQ дикого типа (капсулы типа I, число триангуляции T=1) с диаметром 16 нм; (б) модельная икосаэдрическая структура с диаметром 29 нм (капсулы типа II, T=3); (в) модельная икосаэдрическая структура с диаметром 32 нм (капсулы типа III, T=4)

С учетом наличия неполных капсул, выявленных при *ab initio* моделировании, было установлено, что исследуемая система может быть адекватно описана пятью типами компонентов: компоненты (1-2) - полные и неполные капсулы типа II или типа III; компоненты (3-4) - полные и неполные капсулы типа I; компонент (5) - малые структуры (грани) капсул типа I. Неполные капсулы типа I и типа II или III не имели одной и или двух граней. Для учёта возможной вариации размеров капсул, а также перехода от капсул типа I к капсулам типа II или III была введена полидисперсность полных и неполных капсул. Для этого кривые рассеяния от капсул типа I и типа II/III представлялись как "основные" форм-факторы частиц. Далее, рассеяние от частиц типа I и типа II/III с диаметром d может быть представлено путем шкалирования этих форм-факторов в угловом диапазоне как $I(s,d) = (d/d_0)^3 I(s*d_0/d)$, где d_0 - номинальный средний диаметр капсулы. Этот параметр составлял d_0 = 16 нм для частиц типа I и 29 нм для крупных капсул, которые могли быть либо типа II (для частиц d < 28 нм), либо типа III (для частиц

d > 30 нм). Используя такой подход, соответствующее гауссовское распределение было добавлено компонентов размерам В программу MIXTURE-M MIXTURE), (модифицированную версию программы которая позволила приблизить экспериментальные данные линейной комбинацией рассеяния от пяти компонентов и найти их объемные доли и полидисперсность. Средние радиусы компонентов 1-4 могли изменяться вблизи соответствующих значений d_0 , в то время как их полидисперсность (ширина Гауссова распределения) была ограничена не более чем 0.5 нм для полных капсул, но могла достигать 3 нм для неполных капсул типа I и типа II/III.

Следует отметить, что выбор малых подструктур капсул (соответствующих размерам пентамеров) в качестве дополнительного компонента был сделан в некоторой степени произвольно, но результаты не сильно зависят от выбора конкретного форм-фактора, представляющего малые частицы в растворе, пока эти частицы остаются намного меньших размеров по сравнению с капсулами. Кроме того, был проверен альтернативный подход, предполагающий, что отклонения от икосаэдрической симметрии обусловлены не отсутствием граней, а деформациями капсул. Было обнаружено, что аффинные деформации идеальных икосаэдрических частиц до степени в 15-20% требуются для замазывания минимумов в рассчитанных данных МУРР, но даже эти очень большие деформации дают приближения, которые оказываются намного хуже, чем приближения от моделей с отсутствующими гранями. Кроме того, *ab initio* восстановленные формы также указывают скорее на отсутствующие грани, а не на деформированные частицы, поэтому гипотеза о наличии деформаций оказывается менее совместима с экспериментальными данными МУРР.

Результаты моделирования приведены в Таблицах 4.11-4.15, наилучшие приближения экспериментальных данных МУРР представлены на Рис. 4.14. В целом, предложенная пятикомпонентная модель обеспечивает достаточно хорошее соответствие ко всему имеющемуся набору экспериментальных данных МУРР от LS фермента.

Таблица 4.11 Структурные параметры смесей фермента LSBS дикого типа в фосфатном и Tris гидрохлоратном буфере, полученные программой MIXTURE. Обозначения: χ^2 — невязка к данным МУРР, ν — объемные доли, R — средние радиусы, dR — полидисперсность капсул.

Образец	pН	χ^2	Маль	Малые капсулы			шие капс	улы	Свободные			
			((T=1)			(T=3,4)			пентамерные грани		
			П	олные	;		полные					
			v	R,	dR,	v	R,	dR,	v	R,	dR,	
				HM	HM		HM	HM		HM	HM	
LSBS	6.0	1.08	0.912	7.83	0.4	0.022	13.47	0.5	0.065	3.0	0.2	
WT	6.5	1.10	0.907	7.85	0.4	0.019	13.42	0.5	0.074	3.0	0.2	
(Phospha	7.0	1.02	0.887	7.86	0.3	0.032	13.43	0.5	0.081	3.0	0.3	
te)	7.5	1.30	0.863	7.89	0.2	0.015	13.48	0.2	0.122	3.0	0.2	
	8.0	1.43	0.849	7.90	0.3	0.014	13.84	0.1	0.135	3.0	0.3	
LSBS	7.0	1.51	0.161	9.00	0.3	0.762	15.11	0.5	0.077	3.0	0.3	
WT	7.6	2.57	0.305	9.00	0.5	0.633	14.52	0.5	0.062	3.0	0.3	
(Tris)	8.4	2.27	0.455	9.00	0.4	0.454	14.67	0.5	0.091	3.0	0.2	
	9.0	1.09	0.104	9.00	0.3	0.573	15.52	0.5	0.323	3.0	0.2	

Таблица 4.12 Структурные параметры смесей фермента LSBS дикого типа в боратном буфере, полученные программой MIXTURE.

Образец	рН	χ^2	Малые капсулы (T=1)		Большие капсулы (T=3,4)							
			I	юлны	e	полные]	неполные		
			v	R,	dR,	v	R,	dR,	v	R,	dR,	
				HM	HM		HM	HM		HM	HM	
LSBS	7.0	1.28	0.97	7.69	0.50	0.02	15.72	0.5	0.01	14.77	0.27	
WT	8.0	2.47	0.96	7.72	0.50	0.02	15.64	0.5	0.01	14.65	0.24	
(Phosphate)	9.0	2.81	0.98	7.54	0.50	0.01	15.95	0.5	0.01	14.72	0.15	
	10.0	3.72	0.02	9.00	0.10	0.62	15.92	0.5	0.36	12.94	0.99	
	7.0	1.28	0.97	7.69	0.50	0.02	15.72	0.5	0.01	14.77	0.27	

Таблица 4.13 Структурные параметры смесей фермента LSBS R127T мутанта в фосфатном и Tris гидрохлоратном буфере, полученные программой MIXTURE.

Образец	pН	χ^2		Малые капсулы (T=1)			Большие капсулы (T=3,4)						
				ПОЛНЫ	ie		полные			неполн	ые		
			v	R,	dR,	v	R,	dR,	v	R,	dR,		
				HM	HM		HM	HM		HM	HM		
LSBS	6.0	5.41	0.04	9.00	1.27	0.80	14.70	0.5	0.16	11.70	0.10		
R127T	6.5	6.71	0.03	9.00	1.34	0.79	14.73	0.5	0.18	11.73	0.10		
(Phosphate)	7.0	7.25	0.03	9.00	1.19	0.79	14.76	0.5	0.18	11.75	0.10		
	7.5	2.98	0.03	9.00	1.52	0.75	14.83	0.5	0.22	11.81	0.10		
	8.0	2.78	0.04	9.00	1.47	0.75	14.92	0.5	0.21	11.98	0.10		
LSBS	7.0	4.62	0.05	9.00	1.33	0.75	14.87	0.5	0.20	11.87	0.10		
R127T	7.5	4.91	0.04	9.00	1.25	0.74	14.96	0.5	0.22	11.93	0.10		
(Tris)	8.0	4.60	0.04	9.00	1.13	0.70	15.22	0.5	0.26	12.24	0.10		
	8.5	3.30	0.03	9.00	1.23	0.54	15.76	0.5	0.43	12.77	0.10		
	9.0	5.66	0.03	9.00	1.19	0.47	16.27	0.5	0.50	13.24	0.10		

Таблица 4.14 Структурные параметры смесей фермента LSAQ дикого типа в фосфатном и Tris гидрохлоратном буфере, полученные программой MIXTURE.

Образец	pН	χ^2	Маль	Малые капсулы			Большие капсулы			Свободные		
				(T=1)			(T=3,4)			пентамерные грани		
			П	юлные			полные					
			v	R,	dR,	v	R,	dR,	v	R,	dR,	
				HM	HM		HM	HM		HM	HM	
LSAQ	6.0	2.37	0.62	7.95	0.5	0.21	13.11	0.5	0.17	3.0	0.1	
WT	6.5	2.06	0.65	7.92	0.5	0.21	13.15	0.5	0.14	3.0	0.1	
(Phosphate)	7.0	2.05	0.62	7.93	0.5	0.23	13.23	0.5	0.15	3.0	0.1	
	7.5	2.08	0.64	7.94	0.5	0.22	13.12	0.5	0.14	3.0	0.1	
	8.0	1.96	0.61	7.98	0.5	0.23	13.34	0.5	0.16	3.0	0.1	
LSAQ	7.0	2.47	0.63	7.98	0.5	0.19	14.50	0.5	0.18	3.0	0.1	
WT	7.5	1.70	0.69	7.95	0.5	0.21	14.56	0.5	0.10	3.0	0.1	
(Tris)	8.0	1.99	0.69	7.94	0.5	0.18	14.58	0.5	0.13	3.0	0.1	
	8.5	1.86	0.68	7.93	0.5	0.20	14.52	0.5	0.12	3.0	0.1	
	9.0	2.12	0.67	7.95	0.5	0.17	14.51	0.5	0.16	3.0	0.1	

Таблица 4.15 Структурные параметры смесей фермента LSAQ-IDEA мутанта в фосфатном и Tris гидрохлоратном буфере, полученные программой MIXTURE.

Образец	pН	χ^2	Маль	Малые капсулы			шие капс	улы	Свободные			
				(T=1)			(T=3,4)			пентамерные грани		
			не	еполны	e		полные					
			v	R,	dR,	v	R,	dR,	v	R,	dR,	
				HM	HM		HM	HM		HM	HM	
LSAQ-	6.0	2.38	0.01	9.00	0.1	0.67	13.41	0.5	0.32	3.0	0.1	
IDEA	7.0	2.09	0.02	9.00	0.1	0.81	13.29	0.5	0.17	3.0	0.1	
(Phosphate)	8.0	2.21	0.01	9.00	0.1	0.77	13.55	0.5	0.22	3.0	0.1	
LSAQ-	7.0	3.75	0.05	9.00	0.1	0.95	13.73	0.5	-	-	-	
IDEA	7.5	2.47	0.04	9.00	0.1	0.96	13.73	0.5	-	-	-	
(Tris)	8.0	2.61	0.05	9.00	0.1	0.95	13.74	0.5	-	-	-	
	8.5	2.13	0.03	9.00	0.1	0.97	13.76	0.5	-	-	-	
	9.0	2.17	0.04	9.00	0.1	0.96	13.82	0.5	_	_	-	

Для LSBS дикого типа в фосфатном буфере были найдены три компонента: полные капсулы типа I (с объемной долей от 85 до 90%), свободные «диссоциированные» грани капсул (объемные доли от 6 до 13%) и полные капсулы типа II (объемные доли от 1 до 3%). Размеры капсул типа I и типа II близки к теоретическим значениям 16 нм и 27 нм, соответственно. Напротив, для LSBS дикого типа в Trisбуфере присутствует значительная объемная доля крупных капсул III типа (от 45 до 76%) с диаметром в диапазоне от 29 нм до 31 нм. Также наблюдается сильная зависимость от рН, при рН 7.0 и 9.0 доминирующие фракции состоят из крупных капсул III типа, в то время как при рН 7.6 и 8.4 обнаружено значительное количество (от 30 до 45%) малых капсул типа I. Замазанные максимумы на кривых рассеяния при рН 7.6 и 8.4 не очень хорошо приближаются данной моделью, что может указывать, например, на присутствие деформированных частиц в растворе во время этого структурного перехода.

Мутант LSBS R127T образует три различных компонента в фосфатном и Tris буферах, которые соответствуют форм-факторам полных капсул типа III (75-80% объемная доля), неполных капсул типа III (16-21% объемная доля) и неполных капсул типа I (4% объемная доля). Относительно высокая степень полидисперсности (от 1 нм до 2 нм) наблюдается для неполных структур. Следует подчеркнуть, что для мутанта LSBS R127T в фосфатном буфере объемные фракции и размер компонентов не зависят от рН, тогда как для мутанта LSBS R127T в Tris буфере наблюдается сильная зависимость от рН. Средний размер компонента типа III увеличивается с рН (от 30 нм при рН 7.0 до 33 нм при рН 9.0). Это можно объяснить структурным перестроением икосаэдрических капсул типа III между рН 7.0 и 9.0.

Результаты для LSBS в боратном буфере также сильно зависят от рН. При значениях рН 7.0, 8.0 и 9.0 образуются главным образом капсулы типа I (объемная доля 97%) с диаметром около 16 нм, тогда как при рН 10.0 основная фракция (62% полных и 36% неполных капсул) относится к капсулам типа III с диаметрами около 32 нм. Это означает, что процесс образования капсул может подстраиваться

(управляться) рН буферного раствора, при этом ансамбли различных типов капсул находятся в динамическом равновесии друг с другом.

Для LSAQ дикого типа в фосфатном и Tris-буфере были обнаружены три компонента: 10-18% свободных "диссоциированных» граней, 61-70% полных капсул типа I диаметром 16.0 нм и 20-22% полных капсул типа II диаметром 26-27 нм. Вставка четырех аминокислотных остатков в мутантном типе LSAQ-IDEA приводит к существенным изменениям профиля рассеяния (Рис. 4.14г). Компонент с капсулами типа І практически исчезает (было найдено лишь до 5% этих частиц), вместо них образуются в основном капсулы типа II диаметром 27 нм и небольшое количество малых свободных граней капсул. Эти результаты показывают, что мутация IDEA предотвращает образование капсул типа I, но не капсул типа II. Таким образом, в работе [A13] было показано, что фермент LS подобно некоторым вирусным белкам может самособираться в капсулы различного размера и количества субъединиц вследствие мутаций или изменений химических и физических условий среды. Причем эта изменчивость форм капсул, как показывают данные МУРР и крио-ЭМ, коррелирует с наличием определенных остатков вблизи активного центра фермента LS и, таким образом, с его ферментативной (каталитической) функцией.

Заключение к Главе 3

В Главе 3 был рассмотрен ряд подходов, позволяющих проводить количественный анализ сложных многокомпонентных и полидисперсных смесей по данным МУРР/МУРН. С помощью разработанных методов можно определять число независимых компонентов, оценивать их объемные доли при наличии информации об атомных структурах компонентов, определять функции распределения частиц по размерам и параметры межчастичного взаимодействия при аппроксимации формы компонентов простыми геометрическими телами. Широкий спектр рассмотренных практических приложений данных подходов подчеркивает большую востребованность метода МУРР/МУРН для исследования таких систем и высокий потенциал использования разработанных подходов.

Глава 4

Количественный анализ состава липидных и липид-белковых смесей по данным МУРР

МУРР/МУРН Методы широко используются ДЛЯ изучения растворов макромолекул, а также частиц, которые встроены в липидные носители и системы доставки лекарств, такие как поверхностно-активные мицеллы, фосфолипидные везикулы и двуслойные нанодиски [157-161]. Для адекватного описания рассеяния от таких систем необходимо учитывать, как форм-фактор (общая форма частицы), так и отражения Брэгга дальнего порядка, возникающие из-за упорядоченной организации бислоев, что является нетривиальной задачей. Нами разработан общий подход вычисления и анализа данных МУРР от липидных смесей, который позволяет восстановить электронную плотность липидного бислоя и одновременно везикул размерам. оценить распределение ПО Алгоритм реализован компьютерных программах BILMIX [A69] и LIPMIX [A70]. Данный подход может быть полезен для анализа различных типов систем на основе липидов, например, для характеристики взаимодействий между молекулами лекарственного средства (мишени) и потенциальными системами доставки.

Фосфолипиды играют важную биологическую роль в клетках. Они являются амфифильными молекулами и самоорганизуются в водных растворах в слоистые структуры. *In vivo* фосфолипиды образуют двойные слои в биологических мембранах, служащие каркасом как для гидрофильных, так и для гидрофобных областей макромолекул и разделения клеточных компонентов. *In vitro* фосфолипиды образуют упорядоченные комплексы, такие как везикулы и липосомы, которые имеют однослойную (SUV – Single Unilamellar Vesicles) или многослойную структуру (MLV - Multi-Lamellar Vesicles).

Системы доставки лекарства на основе фосфолипидов, таких как липосомы, часто используются для доставки лекарственного средства-мишени. Липосомы преимущественно используются в качестве носителей для гидрофильных молекул, которые инкапсулированы в водный внутренний объем липосомы, который

ограничен липидным бислоем. Гидрофобные лекарственные средства могут быть включены в гидрофобную часть (липидный бислой). Анализ распределения электронной плотности и морфологии липидных бислоев может дать представление о распределении и включении целевых лекарственных средств в липосомы [159].

4.1 Анализ полидисперсных систем липидных везикул: восстановление электронной плотности липидного бислоя, распределения везикул по размерам и мультислойной ламеллярной организации везикул

Существуют различные подходы к решению проблемы моделирования данных малоуглового рассеяния от модельных липидных мембран [162-167]. В большинстве случаев профили плотности длины рассеяния симметричных липидных бислоев представляются в виде ступенчатых функций [163] или гауссианов [162]. При этом липидные бислои могут иметь значительные изгибные колебания энтропийного происхождения. В мультиламеллярных везикулах (MLV) эти эффекты приводят к затуханию корреляционной функции и характерным формам брэгговских пиков [168]. Упругость мембраны оценивается по анализу формы брэгговских пиков с помощью модифицированной теории Кайле (МСТ) [169, 170]. Степень колебаний определяется параметром Кайле, который является функцией модуля изгиба бислоя и объемного модуля межслоевого сжатия. Для такого анализа рентгеновское рассеяние имеет преимущество перед нейтронными исследованиями из-за более высокого разрешения данных МУРР по сравнению с тем, которое обычно достигается с помощью нейтронов.

Для оценки состава и структуры растворов липидных смесей нами была разработана программа LIPMIX [A70]. В ней используется тот факт, что данные МУРР, полученные от симметричных липидных везикул, могут быть хорошо аппроксимированы произведением форм-фактора тонкой сферической оболочки (определяющей размер везикул) и форм-фактора плоского липидного бислоя (содержащего информацию о плотности электронов в липидном бислое) [171, 172].

Это так называемое приближение разделенного форм-фактора является действительным, когда размер везикул намного больше, чем толщина липидного бислоя. Упорядоченное расположение бислоев внутри многослойной везикулы может быть учтено введением межслоевого структурного фактора [166].

В общем случае, интенсивность разбавленной полидисперсной липидной смеси может быть представлена следующим образом:

$$I(s) = \sum_{k=1}^{N} v_k I_k(s)$$
 (4.1)

где $s=4\pi sin\theta/\lambda$, 2θ - угол рассеяния, N - число липидных везикул с различными плотностями бислоя, v_k и $I_k(s)$ - соответствующие объемные доли и парциальные интенсивности рассеяния от везикул. В свою очередь, каждая парциальная интенсивность может быть выражена как:

$$I_k(s) \cong \frac{1}{s^2} |\int F_{TS}(s, r)_k D_V(r)_k dr|^2 \times |F_{FB}(s)_k|^2 \times \sum_{i=1}^M w_i T_i^{FB}(s)$$
 (4.2)

где $D_V(r)_k$ - объемное распределение везикул, $F_{TS}(s,r)_k$ - форм-фактор тонкой сферической оболочки с радиусом г, $F_{FB}(s)_k$ - форм-фактор плоского липидного бислоя k-го компонента в смеси, M - общее количество частиц с различной многослойной организацией, $T_i^{FB}(s)$ - межслоевой структурный фактор равномерно расположенных плоских бислоев, а w_i - коэффициент заселенности частиц с заданным числом упорядоченных липидных бислоев. Объемное распределение $D_V(r)$ обычно задается мономодальным распределением Шульца со средним радиусом R и шириной σ .

Форм-фактор липидного бислоя F_{FB} (s) задается преобразованием Фурье профиля электронной плотности липидного бислоя, которое может быть аппроксимировано пятью гауссовыми функциями, аналогично работе [173]

$$\rho(z) = \sum_{i=1}^{2} A_i \left[\exp\left(\frac{-(z - z_{Hi})^2}{2\sigma_{Hi}^2}\right) + \exp\left(\frac{-(z + z_{Hi})^2}{2\sigma_{Hi}^2}\right) \right] - \rho_r \exp\left(\frac{-z^2}{2\sigma_C^2}\right)$$
(4.3)

где первые четыре гауссовых слагаемых с шириной σ_{Hi} и центром Z_{Hi} (i = 1,2) представляют собой гидрофильные фосфолипидные полярные головные группы и предназначены для моделирования как симметричных, так и асимметричных профилей плотности. Пятая гауссовская функция с шириной σ_C в центре бислоя

учитывает гидрофобные углеводородные цепи, а ρ_r - отношение электронной плотности углеводородных цепей к электронной плотности полярных головных групп липидной молекулы.

Межслоевой структурный фактор $T^{FB}(s)$ из L равномерно разнесенных плоских бислоев конечного размера, приводящий к появлению брэгговских пиков, рассчитывается согласно модифицированной теории Кайле [169, 170]:

$$T^{FB}(s) = L + 2\sum_{k=1}^{L-1} (L-k)\cos(ksd)\exp\left\{-(\frac{d}{2\pi})^2 s^2 \eta[\gamma + \ln(\pi k)]\right\}$$
(4.4)

где L - общее число упорядоченных плоских бислоев в везикуле, d – межслоевое расстояние и η - параметр Кайле, который является мерой изгиба бислоя, γ - постоянная Эйлера.

Уравнения (4.1-4.4) позволяют параметризовать рассеяние от полидисперсной смеси частиц MLV с использованием нескольких структурных параметров (Таблица 4.2). Программа LIPMIX оптимизирует эти параметры для соответствия экспериментальным квази-ньютоновской заданным данным помощью процедуры минимизации по схеме Бройдена-Флетчера-Гольдфарба-Шанно [131]. Параметры начального приближения оцениваются по экспериментальным данным, а на переменные накладываются простые ограничения, отражающие физически обоснованные значения параметров системы. Так, начальный средний радиус везикулы R_k получается из радиуса инерции R_g , рассчитанного по приближению Гинье [5], а полидисперсность размера везикулы может изменяться в интервале $[0.02*R_g, 0.5*R_g]$. Для параметров электронной плотности липидного бислоя центры полярных головных групп фосфолипидов Z_{Hi} можно варьировать в интервале $[1.5 \div 2.5]$ нм, а их ширина может изменяться в пределах $[0.1 \div 0.3]$ нм, что соответствует общей толщине бислоя порядка 4.5÷6.5 нм. Программа LIPMIX может быть запущена автоматически с помощью командного файла, содержащего спецификации модели и начальные значения параметров, а также их верхние и нижние пределы.

Таблица 4.2 Структурные параметры липидных смесей, содержащих везикулы различного размера и с различными профилями электронной плотности липидного бислоя.

k-й липидный компонент смеси (k = 1 ÷ N)											
Объемная доля {v _k }	Средний размер везикулы и ее полидисперсностьу $\{R_k, \ \sigma_k\}$	Профиль электронной плотности липидного бислоя $\{A_i, z_{Hi}, \sigma_{Hi}, \rho_r, \sigma_c\}$ $(i=1,2)$	Мультиламеллярная организация везикулы SUV - $\{w_i, L_i\}, L_i$ =1 MLV - $\{w_i, L_i\}, L_i$ \geq 2 $(i=1 \div M)$ Параметр Кайле η								

Для проверки эффективности и надежности работы алгоритма нами были выбраны три случая липидных смесей: (1) только частицы SUV, (2) SUV и MLV с количеством упорядоченных слоев до 3, (3) SUV и MLV с количеством упорядоченных слоев до 7. Средний радиус частиц был равен 50 нм, степень полидисперсности - 10%, электронная плотность имела симметричный профиль и толщину бислоя 6.5 нм, а параметр Кайле для MLV составлял 0.10. Относительная ошибка Пуассона в 5% была добавлена к сгенеренным теоретическим данным (показанным на Рис. 4.2 в виде кружков с интервалами ошибок); соответствующее распределение размеров везикул и профиль электронной плотности бислоя показаны на вставке Рис. 4.2.

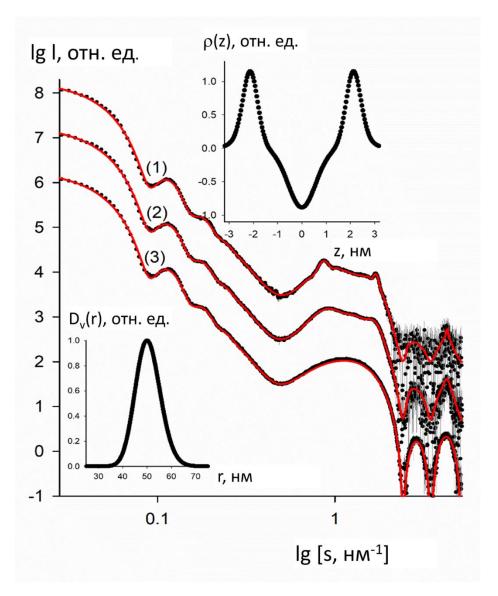


Рис. 4.2 Теоретические кривые МУРР от смесей липидов (кружки с интервалом ошибок) и приближения, полученные программой LIPMIX (красные сплошные кривые). Кривая 1 соответствует смеси SUV и MLV (до 7 упорядоченных слоев), кривая 2 - смеси SUV и MLV (до 3 упорядоченных слоев), кривая 3 — однослойным частицам SUV. Заданные распределение размеров везикул Dv(r) и профиль электронной плотности липидного бислоя $\rho(z)$ показаны на вставках.

Нами были протестированы различные начальные приближения (с относительными отклонениями до 40% от истинных значений параметров), и во всех случаях алгоритм сходился к истинному решению. Наилучшие приближения, полученные с помощью программы LIPMIX, показаны на Рис. 4.2 красными сплошными линиями. Полученные результаты подтверждают, что программа LIPMIX позволяет надежно восстанавливать структурные параметры липидных

смесей. Практические примеры применения программы LIPMIX (в частности для анализа состава везикул DMPC/DPPC во время процесса экструзии) будут приведены в разделе 4.4.5.

Еще одна разработанная нами программа BILMIX [A69], предназначена для анализа смесей только однослойных везикул, однако в отличие от программы LIPMIX, обладает дополнительными возможностями, во-первых, программа BILMIX позволяет моделировать везикулы как сферической, так и эллипсоидной формы, во-вторых, в ней можно задавать асимметричную электронную плотность бислоя, которая может возникать при взаимодействии липосомы с белком или другим связывающим партнером.

В этом случае форм-фактор липидного бислоя $F_{FB}(s)$ задается преобразованием Фурье профиля электронной плотности липидного бислоя, которое содержит четыре гауссовых функции

$$\rho(z) = \exp\frac{-(z - z_{H1})^2}{2\sigma_{H1}^2} + \exp\frac{-(z + z_{H1})^2}{2\sigma_{H1}^2} - \rho_r \exp\frac{-z^2}{2\sigma_C^2} + A_2 \exp\frac{-(z + z_{H2})^2}{2\sigma_{H2}^2}, \tag{4.5}$$

Первые три слагаемых в уравнении (4.5) аналогичны тем, которые содержатся в уравнении (4.3). Последний член в уравнении (4.5) соответствует общему случаю асимметрии профиля электронной плотности, возникающему, например, в случае взаимодействия липосомы с белком или другим связывающим партнером. На Рис. 4.3 эта схема представлена графически.

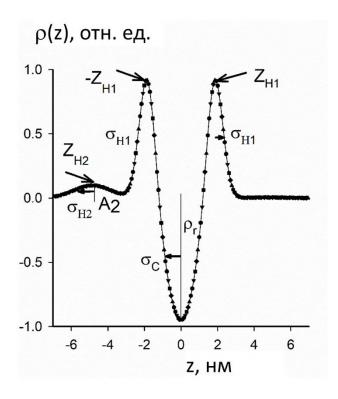


Рис. 4.3 Параметрическое описание профиля электронной плотности липидного бислоя, используемого в программе ВІLМІХ. Левая часть профиля состоит из двух гауссовских функций, где первый гауссиан с центром при z_{H1} относится к полярной головной группе липидной молекулы, а второй гауссиан с центром при z_{H2} представляет связанную с липосомой частицу (например, молекулу белка). Центральная часть профиля имеет отрицательный контраст, соответствующий углеводородным цепям липидной молекулы. Правая часть профиля описывается одним гауссианом, соответствующим головной липидной группе с центром при z_{H1} . ρ_r - нормированный контраст углеводородных цепей по отношению к средней электронной плотности растворителя. Ширины гауссовских функций для полярной головной группы липида, связанной частицы (молекулы белка) и углеводородных цепей липидной молекулы составляют σ_{H1} , σ_{H2} и σ_C , соответственно.

Обе программы (LIPMIX и BILMIX) после завершения своей работы создают 3 файла с расширениями ".fit", ".den", ".vr", где имена файлов совпадают с именем файла экспериментальных данных, а также log-файл с историей запусков. Файл с расширением ".fit" содержит наилучшие приближения к экспериментальным данным МУРР, файл с расширением ".den" — восстановленный профиль электронной плотности липидного бислоя везикулы, и наконец файл ".vr" содержит информацию о распределении везикул по размерам. Программы LIPMIX

и BILMIX входят в состав пакета ATSAS, находящемся в свободном доступе для академических пользователей (https://www.embl-hamburg.de/biosaxs/software.html).

4.2 Применение разработанных методов для анализа эволюции липидных везикул в процессе экструзии

Как уже было отмечено в разделе 4.3, метод МУРР широко используется для изучения структуры и морфологии фосфолипидных везикул (липосом) и их взаимодействия с белковыми макромолекулами. Поскольку липосомы часто используются для доставки лекарственных средств, то по изменению распределения электронной плотности липидного бислоя можно судить о характере внедрения лекарственных средств в бислой липосомы.

Эффективность подхода количественного описания состава липидных смесей по данным МУРР, описанного в разделе 4.3, была проверена на примере эволюционирующих везикул DMPC/DPPC во время процесса их экструзии [A70]. Общая схема организации липидных везикул в растворе показана на Рис. 4.19. Полный процесс экструзии состоял из 25 проходов через экструдер. Такая повторяющаяся процедура обычно должна приводить к образованию достаточно узкодисперсных популяций везикул. При проведении экспериментов были использованы поликарбонатные мембранные фильтры с различным размером пор (30, 50, 100 и 200 нм). На Рис. 4.20 представлены данные МУРР (кружки), собранные от полностью экструдированных везикул DMPC/DPPC, полученных после 25 проходов через фильтры поликарбонатных мембран. С помощью LIPMIX получить хорошие приближения программы удалось экспериментальным данным МУРР для всех размеров фильтров в полном угловом диапазоне (Рис. 4.20, сплошные линии).

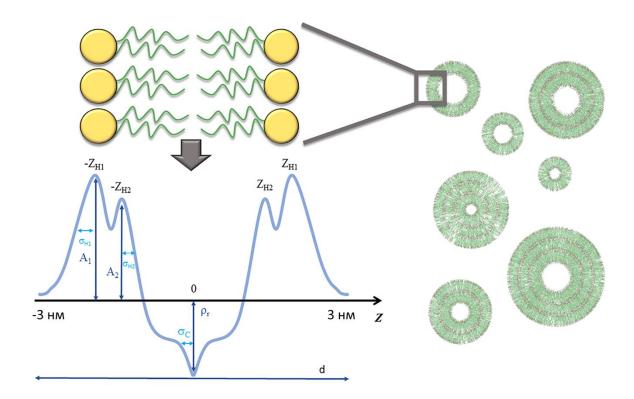


Рис. 4.19 Общая организация смеси однослойных (SUV) и многослойных (MLV) везикул в растворе (справа). Параметры модели липидного бислоя, используемые для представления электронной плотности бислоя в виде суммы гауссовых функций (левая панель). Положительная электронная плотность с асимметричными пиками относится к полярным головным группам молекул липидов. Центральный отрицательный пик профиля плотности соответствует гидрофобным хвостам липидов. Внизу показано расстояние между ламеллярными повторяющимися мотивами (d).

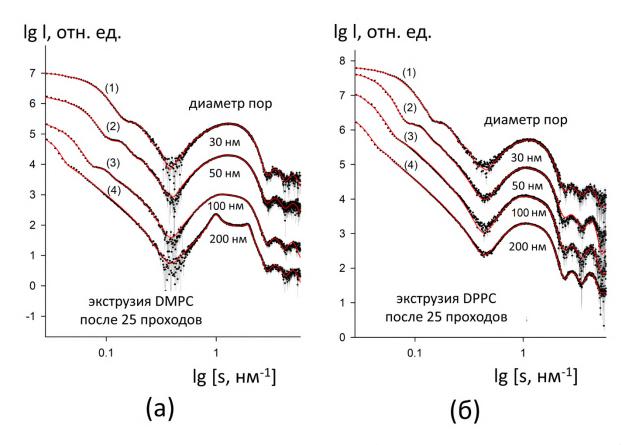


Рис. 4.20 Экспериментальные данные МУРР (кружки с интервалами ошибок) экструдированных везикул после 25 проходов через экструдер для DMPC (панель (а)) и DPPC (панель (б)) с использованием различных диаметров пор (кривая 1-30 нм, кривая 2 -50 нм, кривая 3 - 100 нм, кривая 4- 200 нм) и приближения, полученные с помощью программы LIPMIX (красные сплошные линии).

Восстановленные структурные параметры липосом приведены в Таблице 4.16, а на Рис. 4.21 представлены полученные профили электронной плотности липидных бислоев и распределения размеров везикул для DMPC и DPPC. Примечательно, что профили плотности и распределение размеров значительно отличаются для двух типов везикул, хотя толщина бислоя DMPC аналогична DPPC (~ 6 нм). Для DMPC профиль плотности имеет узкие пики для областей гидрофильной головной группы липидного бислоя. В центральной области бислоя, соответствующей гидрофобным углеводородным хвостам липидов, наблюдается широкий пик электронной плотности. Для DPPC широкие пики имеются как в областях головных групп липидного бислоя, так и на гидрофобном центральном участке бислоя. Профили плотности слабо чувствительны к изменению диаметра пор, используемых для

экструзии. Также стоит отметить, что восстановленные профили везикул DPPC на Рис. 4.21б (кривые 1-4) находятся в хорошем согласии с профилем, полученным стандартным прямым расчетом форм-факторов от дифракционных пиков кривой рассеяния с использованием преобразования Фурье (Рис. 4.216, кривая 5), описанного В работах [170, 174] ДЛЯ полидисперсного раствора неэкструдированных мультиламеллярных липосом от DPPC. Мультиламеллярная везикул после процесса экструзии соответствует организация ЛИПИДНЫХ однослойным везукулам.

Что касается распределения везикул по размерам, то здесь ситуация оказывается более сложной. Для частиц DMPC и DPPC наблюдается четкая корреляция между средними радиусами везикул и диаметрами пор мембранных фильтров, причем радиусы везикул пропорциональны диаметрам пор. Однако в среднем внешний диаметр везикул остается немного больше, чем диаметр пор, что отражает способность частиц сжиматься/деформироваться в определенной степени, оставаясь при этом неповрежденными при прохождении через поры мембраны. Такое поведение было отмечено ранее и зависит от давления, приложенного во время экструзии [175, 176].

Отдельно была проанализирована эволюция структурных параметров везикул в зависимости от количества проходов через мембранные фильтры экструдера. С этой целью были проведены эксперименты МУРР для растворов экструдированных частиц DMPC с различным количеством проходов мембраны (от 1 до 25) с использованием мембран с крупными (200 нм) и мелкими (30 нм) порами. Данные МУРР и расчеты, выполненные с помощью программы LIPMIX, представлены на Рис. 4.22а,б, рассчитанные структурные параметры приведены в Таблице 4.17.

Таблица 4.16 Структурные параметры экструдированных везикул DMPC/DPPC, описывающие структуру липидного бислоя, общие размеры и мультиламеллярную организацию везикул, после 25 проходов через экструдер, восстановленные программой LIPMIX по экспериментальным данным МУРР. Оценки ошибок для параметров модели были получены как стандартное отклонение оптимизированных параметров по набору успешных реконструкций LIPMIX.

Параметры	DMPC				DPPC				
липидного	Диаметр пор фильтра				Диаметр пор фильтра				
бислоя	30 нм	50 нм	100	200	30 нм	50 нм	100 нм	200 нм	
			НМ	НМ					
Z_{HI} , HM	1.86±	1.83±	1.82±	1.84±	2.14±	2.13±	2.12±	2.15±	
	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	
σ_{H1} , HM	0.27±	0.26±	0.25±	0.27±	0.32±	0.33±	0.34±	0.32±	
·	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	
A_2/A_1	1.78	1.78	1.76	1.74	1.78	1.74	1.75	1.76	
	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	
Z_{H2} , HM	1.71±	1.76±	1.80±	1.78±	2.16±	2.15±	2.14±	2.17±	
	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.03	
$\sigma_{\!H2}$, HM	0.13±	0.12±	0.12±	$0.14 \pm$	0.34±	0.33±	0.32±	0.35±	
	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	
ρ_r	1.53±	1.50±	1.54±	1.53±	1.74±	1.63±	1.65±	1.61±	
	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	
σ_C , HM	$0.42 \pm$	$0.40\pm$	$0.45 \pm$	$0.43 \pm$	$0.42 \pm$	$0.45 \pm$	$0.44 \pm$	0.43±	
	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	
Параметры распределения везикул по размерам									
R , $_{\rm HM}$	30 ± 1	41 ± 1	59 ± 1	100±2	30 ± 1	54 ± 1	65 ± 1	192 ± 3	
σ , HM	1.2±	1.3±	4.2±	21 ± 1	1.2±0.1	2.1±	9.4±0.5	25±1	
	0.1	0.1	0.2			0.2			
	My.	льтилам	иеллярн	ая оргаі	низация	везикул	I		
Параметр	-	-	-	0.10±	-	-	-	-	
Кайле η				0.01					
Максима-	SUV	SUV	SUV	SUV+	SUV	SUV	SUV	SUV	
льное				MLV					
число слоев				(до 8					
(SUV/MLV)				слоев)					

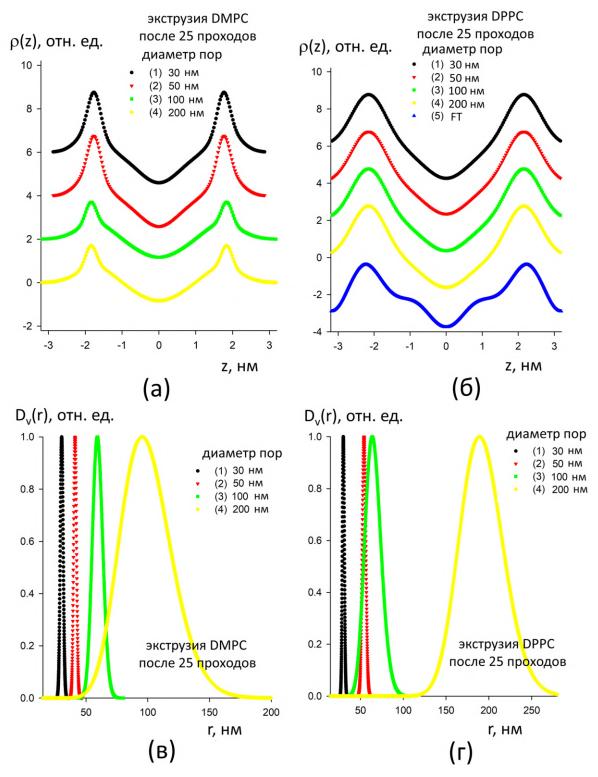


Рис. 4.21 Профили электронной плотности липидных бислоев (а,б - верхняя панель) и распределение размеров везикул (в, г - нижняя панель) для DMPC (левая часть) и DPPC (правая часть) после 25 проходов через экструдер с использованием различных диаметров пор. Обозначения кривых (1-4) те же, что и на Рис. 4.20. Электронная плотность DPPC, полученная с помощью прямого расчета по дифракционным пикам, обозначена как кривая 5 (панель (б)).

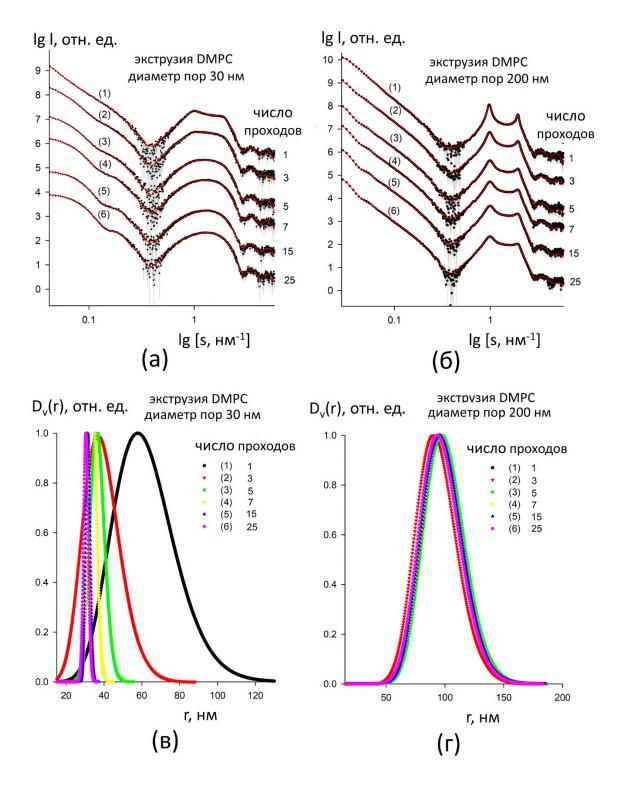


Рис. 4.22 Экспериментальные данные МУРР (A,B - верхняя панель) и восстановленные распределения размеров везикул (C, D - нижняя панель), полученные программой LIPMIX для везикул DMPC в процессе экструзии (от 1 до 25 проходов мембраны) с использованием фильтра с мембранными порами диаметром 30 нм (левая часть) и 200 нм (правая часть). Экспериментальные данные МУРР показаны в виде кружков с интервалами ошибок, а рассчитанные приближения - в виде красных сплошных линий.

Таблица 4.17 Эволюция везикул DMPC в процессе экструзии. Оптимизированные структурные параметры, описывающие структуру липидного бислоя, общие размеры и мультиламеллярную организацию везикул, полученные программой LIPMIX по данным МУРР от DMPC везикул в водном растворе для различного числа проходов через экструдер. Оценки ошибок для параметров модели были получены как стандартное отклонение оптимизированных параметров по набору успешных реконструкций LIPMIX.

Параметры	DMPC				DMPC			
липидного	Диаметр пор фильтра (30 нм)			Диаметр пор фильтра (200				
бислоя				нм)				
Число	1	5	9	25	1	5	9	25
проходов								
через								
экструдер								
Z_{HI} , HM	1.89±	1.92±	1.90±	1.86±	1.85±	1.82±	1.82±	1.84±
1117	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
$\sigma_{\!HI}$, HM	0.28±	0.27±	0.26±	0.27±	0.29±	0.28±	0.25±	0.27±
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
A_2/A_1	1.81±	1.82±	1.86±	1.78±	1.73±	1.71±	1.75±	1.74±
	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Z_{H2} , HM	1.65±	1.64±	1.65±	1.71±	1.79±	1.80±	1.75±	1.78±
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$\sigma_{\!H2}$, HM	0.17±	$0.14 \pm$	0.15±	0.13±	0.17±	0.16±	0.15±	$0.14 \pm$
,	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
$ ho_r$	1.58±	1.65±	1.64±	1.53±	1.56±	$1.56\pm$	1.55±	1.53±
,	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
σ_C , HM	$0.41 \pm$	$0.44 \pm$	$0.45 \pm$	$0.42 \pm$	$0.42 \pm$	$0.40\pm$	$0.41 \pm$	0.43±
	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
I	Тарамет	ры раст	гределе н	ния вези	кул по р	азмерам	Л	
R, HM	61± 1	37 ± 1	33 ± 1	30 ± 1	96 ± 2	100 ± 2	99 ± 2	100±2
σ , HM	14 ± 1	3.5±	1.2±	1.2±	14 ± 1	17 ± 1	17 ± 1	21 ± 1
,		0.2	0.1	0.1				
Мультиламеллярная организация везикул								
Параметр	-	-	-	-	0.11±	0.12±	0.09±	0.10±
Кайле η					0.01	0.01	0.01	0.01
Максимальное	SUV	SUV	SUV	SUV	SUV+	SUV+	SUV+	SUV+
число слоев					MLV	MLV	MLV	MLV
(SUV/MLV)					(до 8	(до 8	(до 8	(до 8
					слоев)	слоев)	слоев)	слоев)

Эволюция структурных параметров липосом при экструзии ясно показывает, что количество мультислойных везикул в растворе уменьшается с увеличением

количеством проходов через поликарбонатные мембраны. При многократной экструзии (20-25 проходов) с использованием мембраны с диаметром пор 30 нм в растворе остаются только однослойные везикулы (Таблица 4.17). Профили электронной плотности существенно не меняются в зависимости от числа проходов и аналогичны представленным на Рис. 4.21а, в то время как распределение везикул по размерам зависит как от диаметра пор мембраны, так и от числа проходов в процессе экструзии (Рис. 4.22в,г). Интересно, что для большого диаметра пор (200 нм) средний радиус везикул остается постоянным около 100 нм при относительно высокой степени полидисперсности (около 15-20%). Однако для малого диаметра пор (30 нм) средний радиус везикул значительно уменьшается с увеличением числа проходов (с 61 до 30 нм), а полидисперсность везикул падает с 25% до 5%, что свидетельствует о формировании однослойных везикул DMPC с малой степенью полидисперсности. Эти результаты хорошо согласуются с результатами предыдущих исследований [175].

В работе [A70] также была рассмотрена зависимость структурных параметров неэкструдированных мультиламеллярных везикул DPPC, а также экструдированных однослойных везикул DPPC (с диаметром пор 30 и 100 нм, соответственно) от температуры в пределах 10÷40 °C. Данные МУРР и расчеты, выполненные с помощью программы LIPMIX, представлены на Рис. 4.23.

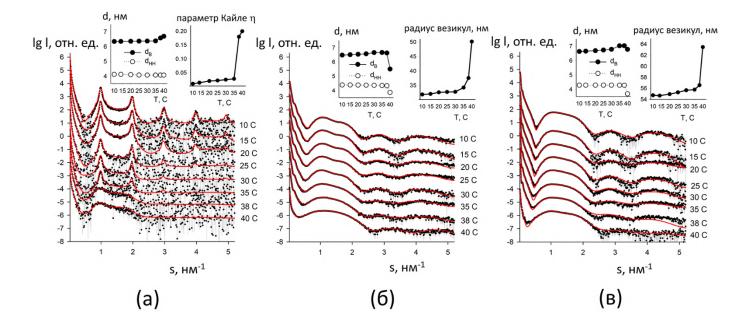


Рис. 4.23 Экспериментальные данные МУРР (кружки) неэкструдированных везикул DPPC (а) и экструдированных везикул DPPC, (б) - диаметр пор 30 нм, (в) - диаметр пор 100 нм, в диапазоне температур 10÷40 °C и приближения, полученные с помощью программы LIPMIX (красные сплошные линии). Вставки содержат температурную зависимость расстояния между повторами ламелл (d), расстояния между головками липидов (d_{HH}), и степень жесткости везикул (параметр Кайле) (для неэкструдированного DPPC) или среднего радиуса везикул (для экструдированного DPPC).

Для неэкструдированных образцов мультислойные везикулы присутствуют в растворе при всех исследованных температурах: они состоят из 10-11 упорядоченных бислоев при $10\div35$ °C, в то время при $38\div40$ °C остаются везикулы только с 2-3 упорядоченными бислоями (Рис. 4.23а). Расстояние между соседними бислоями в мультислойной везикуле (d) постепенно увеличивается с температурой от 6.3 нм до 6.7 нм, в то время как расстояние между пиками электронной плотности от фосфолипидных полярных головных групп бислоя (d_{HH}) остается почти постоянным около 4.3 нм (Рис. 4.23а, вставка). Средний радиус везикул составляет 110-120 нм, полидисперсность везикул по размеру достаточно высокая (около 20%). Степень изгиба поверхности бислоя (пропорциональная параметру Кайле η) также значительно увеличивается при температуре около 38 °C (Рис. 4.23а, вставка).

Для экструдированных DPPC везикул (диаметр пор 30 нм) везикулы с 2 упорядоченными слоями присутствуют при $10\div30$ °C, тогда как при $35\div40$ °C остаются только однослойные липосомы. Расстояние между ламеллярными слоями везикул (d) постепенно увеличивается с температурой от 6.4 нм до 6.6 нм (в пределах $10\div38$ °C) и снижается до 5.6 нм при 40 °C (Рис. 4.236, вставка) вблизи температуры фазового перехода DPPC (T_c =41.3 °C). В то же время, средний радиус везикул постепенно увеличивается с 32 нм (при 10 C) до 37.5 нм (при 38 °C), и затем демонстрирует скачок до 50 нм при 40 °C (Рис. 4.236, вставка). Полидисперсность везикул по размерам относительно низкая при $10\div35$ °C (около 5%) и увеличивается при $38\div40$ °C до 15%.

Экструдированные DPPC везикулы (диаметр пор 100 нм) образуют только однослойные везикулы во всем температурном диапазоне. Расстояние d монотонно увеличивается с 6.6 нм до 7.0 нм (в пределах 10÷38 °C) и немного уменьшается до 6.8 нм при 40 °C (Рис. 4.23в, вставка). В то же время, средний радиус везикул постепенно увеличивается с 54.7 нм (при 10 °C) до 56.5 нм (при 38 °C) и затем резко подскакивает до 63.5 нм при 40 °C (Рис. 4.23в, вставка). Полидисперсность везикул относительно низкая (около 3-4%) при 10÷38 °C и увеличивается до 25% при 40 °C. Полученные результаты согласуются с ожидаемым поведением систем DPPC [177, 178] и еще раз подтверждают, что подход, реализованный в программе LIPMIX, имеет высокую эффективность при анализе данных МУРР от полидисперсных растворов липидных везикул.

4.3 Применение разработанных методов для анализа взаимодействия липидных везикул с белками

В работе [A71] было исследовано влияние различных компонентов вирусной оболочки на структурную перестройку липидной мембраны в ответ на взаимодействие с матричным белком М1 вируса гриппа А. Анализ данных МУРР от нескольких типов однослойных липосом и протеолипосом в отсутствие белка

M1 и при его добавлении в раствор был проведен с помощью программы BILMIX. Были рассмотрены следующие липидные системы: 1) относительно простые двухкомпонентные отрицательно заряженные липосомы, состоящие (DOPC/DOPS); 2) синтетических липидов четырехкомпонентные (PS/POPC/SM/Chol) везикулы, содержащие компоненты, образующие липидные рафты, чтобы проверить влияние гидрофобных контактов и липидных рафтов; 3) так называемые "нативные" липосомы, состоящие из сложной смеси вирусных липидов, выделенных из вирусных частиц гриппа A/Puerto Rico/8/34 (H1N1) по методу Фольча [179]; 4) протеолипосомы, содержащие те же вирусные липиды, но вместе с якорными сегментами интегральных вирусных белков, в основном гемагглютинина (НА), как самого сильного гликопротеина вирусной оболочки. Эксперименты МУРР проводились на синхротронной станции P12 (EMBL c/o DESY, Гамбург, Германия). Растворы синтетических липосом, нативных липосом и протеолипосом, содержащих НА-пептид LI45, измеряли без белка M1 и с загруженным белком M1 при 283 K в диапазоне концентраций 1.0÷3.0 мг/мл в буфере 100 мМ NaCl, 50 мМ MES. Были использованы различные молярные соотношения липосом и белка М1: (67:1, 30:1, 20:1, 15:1, 10:5, 7,5:1, 5:1) для синтетических липосом; (80:1, 40:1, 20:1, 10:1, 4:1, 2,6:1) для нативных липосом; (5:1, 4:1, 2,6:1, 1:1) для протеолипосом, содержащих пептид LI45. Образцы измеряли сразу после смешивания и после 2 ч инкубации. Для синтетических липосом было исследовано пять различных составов: 10% DOPS + 90% DOPC; 30% DOPS + 70% DOPC; 20% bPS+13,3% POPC + 33,3% SM + 33,3% Chol; 30% bPS+10% POPC+40% SM+20% Chol, где PS - фосфатидилсерин, Chol - холестерин, SM - сфингомиелин.

Экспериментальные кривые МУРР, полученные от синтетических липосом без и с добавлением белка М1, представлены на Рис. 4.24 и 4.25. Данные, полученные от липосом без белка М1, содержали широкий максимум в средней части углового диапазона, что характерно для бислойных липидных везикул [176]. В то же время при малых углах наблюдался быстрый подъем интенсивности рассеяния, что

свидетельствует о наличии крупных частиц (диаметром до 1000÷1500 нм) и полидисперсности везикул по размерам. Такие характерные особенности кривых (подъем в области малых углов и широкий пик в средней части углового диапазона) присутствовали для всех образцов и могли быть учтены в рамках приближения форм-факторов, кривая разделенных где рассеяния представляется произведение форм-фактора от крупных сферических оболочек (дающего информацию о размере везикул) и форм-фактора от плоского липидного бислоя (описывающего профиль электронной плотности в бислое). Поэтому для анализа полученных данных МУРР хорошо подходит алгоритм, реализованный в программе BILMIX [A69], который также учитывает полидисперсность везикул по размерам и позволяет моделировать как симметричные, так и асимметричные профили электронной плотности.

С помощью программмы BILMIX удалось хорошо приблизить экспериментальные данные МУРР (Рис. 4.24-4.25а,в, сплошные кривые). Во всех случаях системы липосом оставались достаточно полидисперсными и состояли из нескольких фракций с характерными радиусами 40÷50 нм и 1000÷1500 нм, что хорошо светорассеяния согласовывалось данными динамического [A71]. Восстановленные профили электронной плотности липидного бислоя показаны на Рис. 4.24-4.256, г. Черные кривые на Рис. 4.24-4.256, г представляют профили электронной плотности для липосом без белка М1. Эти кривые симметричны с минимумом в центральной (межмонослойной) области липидного бислоя и двумя максимумами в области полярных головных групп липидов. Расстояние между этими максимумами составляет около 4 нм, что является типичной толщиной липидного бислоя [180].

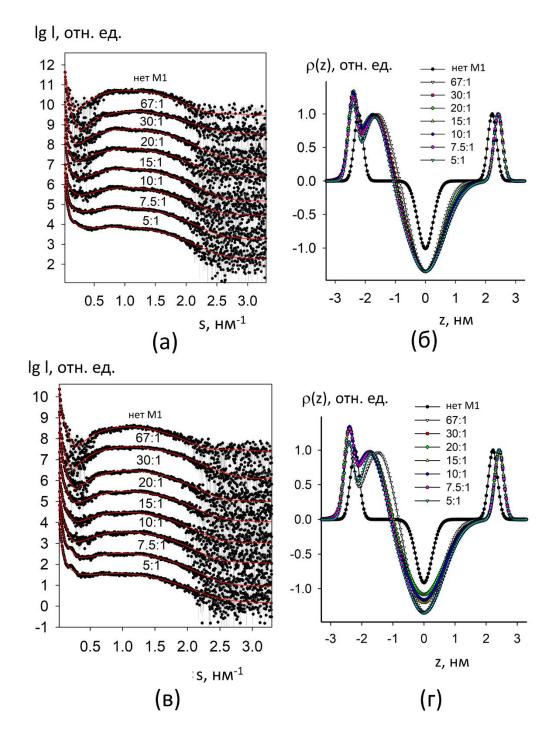


Рис. 4.24 Данные МУРР от смесей синтетических липосом, состоящих из 10 мол.% DOPS + 90% мол. DOPC (а,б) или 30 мол.% DOPS + 70 мол.% DOPC (в,г), в отсутствии белка М1 и с добавлением белка М1. Экспериментальные данные МУРР (кружки) и наилучшие приближения, полученные с помощью программы ВІLМІХ (сплошные красные линии). Кривые сдвинуты на один логарифмический порядок для большей наглядности; На панелях (б, г) показаны восстановленные профили электронной плотности липидного бислоя до загрузки М1 (черная кривая) и после загрузки М1 (цветные кривые). Молярные соотношения липидов и М1 следующие: без М1, 67:1, 30:1, 20:1, 15:1, 10:1, 7.5:1, 5:1.

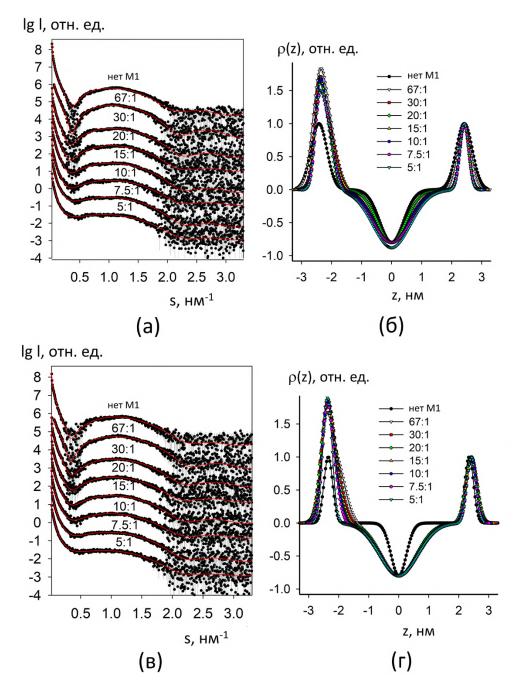


Рис. 4.25 Данные МУРР от смесей синтетических липосом, состоящих из 30 мол.% bPS + 10 мол.% POPC + 40 мол.% SM + 20 мол.% Chol (а,б) или 20 мол.% bPS + 13.3 мол.% POPC + 33.3 мол.% SM + 33.3 мол.% Chol (в,г), в отсутствии и с добавлением белка М1. Экспериментальные данные МУРР (кружки) и наилучшие приближения, полученные с помощью программы BILMIX (сплошные красные линии). Кривые сдвинуты на один логарифмический порядок для большей наглядности; На панелях (б, г) показаны восстановленные профили электронной плотности липидного бислоя до загрузки М1 (черная кривая) и после загрузки М1 (цветные кривые). Молярные соотношения липидов и М1 следующие: без М1, 67:1, 30:1, 20:1, 15:1, 10:1, 7.5:1, 5:1.

Для двухкомпонентных синтетических липосом добавление белка М1 уменьшило плотность электронов в межмонослойной области, одновременно расширив этот минимум. Правый максимум сместился дальше вправо, что означает увеличение эффективной толщины внутреннего монослоя, а левый максимум разделился на два подпика, один из которых сместился к центральной области мембраны, при этом расширившись, а другой сместился влево и увеличился по интенсивности. Такое поведение можно объяснить следующим образом: расщепление пика вызвано появлением белка М1, причем самый левый пик с повышенной интенсивностью соответствует адсорбированному на липосоме белку М1. Адсорбция белка приводит к конденсации липидов под ним, т.е. их площадь на одну головную группу липида уменьшается. Белок М1 конденсирует анионные липиды, т.е. DOPS, как было показано ранее [181]. Поэтому для состава 30 мол.% DOPS + 70 мол. % DOPC эффект сильнее, чем для 10 мол. % DOPS + 90 мол. % DOPC (Рис. 4.24 б,г). В результате такой конденсации диполи полярных головных групп липидов выпрямляются, и поэтому пик электронной плотности полярных головных групп для данного монослоя становится шире. В этом случае возникает дисбаланс в площади монослоев, что приводит к изгибу мембраны внутри липосомы в местах адсорбции белков. Можно предположить, что это приводит к образованию трубчатых образований (инвагинаций) внутри липидных везикул, аналогичных структурам, индуцированным белком М1 вируса гриппа С [182]. Такие локальные изгибы приводят к эффективному уменьшению толщины внешнего (контактирующего с белком М1) монослоя липида и увеличению толщины внутреннего монослоя.

Для четырехкомпонентных синтетических липосом (с более сложным составом, содержащим в том числе холестерин и сфингомиелин) (Рис. 4.25 б,г) центральный минимум электронной плотности бислоя был значительно ниже, что отражает повышенную упорядоченность липидных хвостов, поскольку выбранный состав липидов обеспечивает образование упорядоченных липидных доменов (рафтов)

[183]. В этой системе на восстановленном профиле электронной плотности липидного бислоя не наблюдалось разделенного белково-ассоциированного пика, однако интенсивность левого максимума увеличилась. Кроме того, правый пик не смещался со своего положения для липосом в отсутствии белка М1. Можно предположить, что присутствие холестерина и сфингомиелина стабилизирует липидный бислой М1-индуцированную И предотвращает конденсацию заряженных липидов. Любая деформация мембраны зависит от ее жесткости. С точки зрения теории упругости жидких кристаллов, энергия деформации липидного бислоя зависит от упругих модулей мембраны, который значительно больше для жидкоупорядоченной мембраны [184, 185]. В результате, инвагинации липидного бислоя внутри объема везикулы в рафтообразующих липидных смесях не образуются. Примечательно, что никакой разницы между 30 мол.% bPS + 10 мол.% POPC + 40 мол.% SM + 20 мол.% Chol и 20 мол.% bPS + 13,3 мол.% POPC + 33,3 мол.% SM + 33,3 мол.% не было обнаружено. Этот факт свидетельствует о том, что изменение соотношения SM/Chol не влияет на адсорбцию M1 и его способность образовывать мембранные выступы.

Экспериментальные кривые МУРР и наилучшие приближения ВІСМІХ для нативных липосом и протеолипосом без белка М1 и при его добавлении представлены на Рис. 4.26-4.27а. Восстановленные профили электронной плотности представлены на Рис. 4.26-4.27б. В случае нативных липосом профиль электронной плотности был ближе к двухкомпонентным синтетическим липосомам, чем к четырехкомпонентным рафтообразующим липосомам (Рис. 4.24б). Это интересный факт, поскольку принято считать, что липидная мембрана вируса гриппа А в основном находится в жидкоупорядоченном состоянии [186]. В то же время для липосом, содержащих пептид LI45, электронная плотность в межмонослойной (центральной) области бислоя была минимальной (Рис. 4.27б), как и для четырехкомпонентных синтетических липосом. Этот результат свидетельствует о том, что присутствие пептида LI45 с тремя пальмитатами делает

липидные хвосты более упорядоченными, придавая вирусной мембране рафтообразный вид.

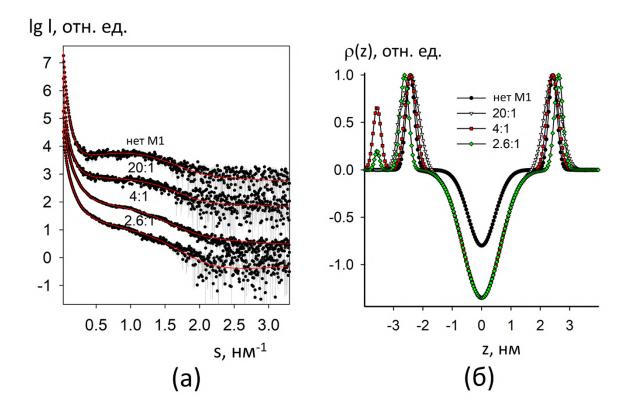


Рис. 4.26 Данные МУРР от нативных липосом, выделенных из оболочки вируса A/Puerto Rico/8/34 (H1N1), в отсутствии белка М1 и с добавлением белка М1. Экспериментальные данные МУРР (кружки) и наилучшие приближения, полученные с помощью программы BILMIX (сплошные красные линии). Кривые сдвинуты на один логарифмический порядок для большей наглядности; На панели (б) показаны восстановленные профили электронной плотности липидного бислоя до загрузки М1 (черная кривая) и после загрузки М1 (цветные кривые). Молярные соотношения липиды:М1 следующие: без М1, 20:1, 4:1, 2.6:1.

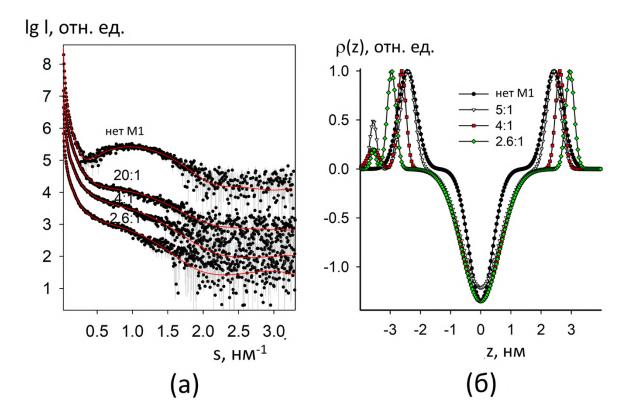


Рис. 4.27 Данные МУРР от смесей, состоящих из протеолипосом, полученных из оболочки вируса A/Puerto Rico/8/34 (H1N1) вместе с пептидами НА LI45, в отсутствии белка М1 и с добавлением белка М1. Экспериментальные данные МУРР (кружки) и наилучшие приближения, полученные с помощью программы ВІLМІХ (сплошные красные линии). Кривые сдвинуты на один логарифмический порядок для большей наглядности; На панели (б) показаны восстановленные профили электронной плотности липидного бислоя до загрузки М1 (черная кривая) и после загрузки М1 (цветные кривые). Молярные соотношения липиды:М1 следующие: без М1, 5:1, 4:1, 2.6:1

Как для нативных липосом, так и для протеолипосом, содержащих пептид LI45, на профиле электронной плотности липидного бислоя (Рис. 4.26-4.276) наблюдалось появление белково-ассоциированного пика, как и для двухкомпонентных отрицательно заряженных синтетических липосом с адсорбированным белком М1. Однако здесь он был отделен от второго (липид-ассоциированного) пика и имел меньшую интенсивность по сравнению с двухкомпонентными синтетическими липосомами. Это наблюдение позволяет предположить, что в случае нативных липосом с пептидом LI45 или без него М1 в меньшей степени внедряется в липидный бислой. Это согласуется с результатами криоэлектронной микроскопии

для неповрежденных (интактных) вирионов [187]. Для нативных липосом также имеет место падение электронной плотности на поверхности межмонослоя (Рис. 4.26б) при адсорбции белка М1, в то время как для липосом, содержащих пептид LI45, эта плотность была минимальной, как и для липосом, не содержащих M1 (Рис. 4.27б). Бислой нативных липосом снова продемонстрировал довольно схожее поведение двухкомпонентными синтетическими cлипосомами: пик взаимодействующего с М1 монослоя сместился влево, в то время как пик внутреннего монослоя сместился немного вправо. Однако смещения второго подпика ближе к центру бислоя (как это было в случае отрицательно заряженных двухкомпонентных липосом) в нативных липосомах при этом не наблюдалось. Таким образом, трубчатые инвагинации мембраны, вероятно, более специфичны для синтетических смесей, а не для сложных смесей нативных вирусных липидов, которые существуют в более упорядоченном состоянии.

Более того, можно видеть, что внутренний монослой липосомы утолщался при адсорбции М1, поскольку соответствующий пик смещался вправо после добавления белка М1. Примечательно, что этот эффект был более выражен в протеолипосомах по сравнению с нативными липосомами: ширина их липидного бислоя увеличилась на 2 нм (по 1 нм с каждой стороны). Это свидетельствует о дополнительном влиянии якорных сегментов НА на формирование плотных липидно-белковых нанодоменов. В связи с этим можно предположить, что пептиды LI45 стимулируют олигомеризацию белка М1 и образование комплексов LI45-М1, что в свою очередь усиливает самоассоциацию белка М1 и образование слоя молекул М1 на поверхности липидного бислоя, облегчая процесс сборки вирусной частицы.

Таким образом, анализ данных МУРР в работе [A71] впервые позволил оценить влияние вирусных якорных пептидов (в основном гемагглютинина НА) на формирование липопротеиновых нанодоменов в липидном бислое. И сделать предположения о физико-химические механизмах, лежащих в основе этого процесса, и его роли в процессах сборки и почкования вируса гриппа А.

Заключение к Главе 4

В Главе 4 был рассмотрен подход, позволяющий проводить количественный анализ состава липидных и липид-белковых смесей по данным МУРР. С помощью разработанного метода можно восстанавливать профили электронной плотности липидного бислоя, распределения везикул по размерам, оценивать мультислоевую организацию липидных везикул и исследовать их взаимодействие с белковыми C разработанных была макромолекулами. использованием программ проанализирована эволюция структурных параметров везикул DMPC/DPPC в зависимости от количества проходов через мембранные фильтры экструдера и температуры. Был изучен характер ассоциации матриксного белка М1 в липидном бислое для синтетических и нативных липосом, полученных из вирусных липидов, выделенных из вирусных частиц гриппа A/Puerto Rico/8/34 (H1N1). Разработанный подход может быть применен для анализа различных типов систем на основе липидов, в том числе для характеристики взаимодействий между молекулами лекарственных средств и системами их доставки.

Глава 5

Прямое определение формы промежуточного компонента в эволюционирующих многокомпонентных системах по данным МУРР

Одной из наиболее важных и сложных задач в современной структурной биологии является исследование сложных и эволюционирующих систем, в которых происходят динамические биологические процессы. Эти системы могут быть проанализированы in vitro, но методы высокого разрешения, такие как рентгеновская кристаллография, ядерный магнитный резонанс или электронная ограничения при исследовании имеют таких гетерогенных объектов, изменяющихся со временем. Метод МУРР позволяет получать структурную информацию о макромолекулярных системах в их естественных физиологических условиях с высоким временным разрешением. Как было показано в Главе 1, для монодисперсных растворов разработаны и широко используются методы интерпретации данных с использованием трехмерных моделей. Они включают в себя как *ab initio* методы восстановления формы частиц с низким разрешением, так и гибридные методы моделирования, использующие домены или субъединицы с известными или предсказанными структурами моделей [188]. высокого разрешения ДЛЯ построения составных Эволюционирующие системы, однако, обычно содержат нескольких видов или конформаций молекул, находящихся в динамическом равновесии, и измеренная интенсивность рассеяния содержит вклады от всех этих компонентов. Методы восстановления формы частиц оказываются неприменимыми, и типичная задача для таких многокомпонентных систем заключается уже не в реконструкции структуры, а в определении объемных долей компонентов. Как было показано в Главе 3, рассеяние от отдельных компонентов может быть известно *a priori*, либо может быть аппроксимировано компонентами в виде простых геометрических тел [А5]. В качестве примеров эволюционирующих систем, изучаемых с помощью метода МУРР, можно указать формирование амилоидных фибрилл [189, 190],

созревание вирусоподобных частиц [191, 192] или динамические олигомерные равновесные смеси в изменяющихся физико-химических условиях [А56, А57]. Для определения количества компонентов и их парциальных кривых рассеяния по данным МУРР от эволюционирующих систем существует несколько подходов, каждый из которых имеет свои преимущества и недостатки. Для анализа смесей типа мономер-димер был предложен метод хемометрического разложения с использованием альтернирующих наименьших квадратов MCR-ALS (Multivariate Curve Resolution-Alternating Least Squares) [193], тогда как для смесей мономермультимер форма мультимера может быть восстановлена в приближении псевдоаминокислотных остатков [А7]. Однако первый подход применим только к простым двухкомпонентным мономер-димерным смесям, тогда как второй подход (реализованный в программе GASBORMX) ограничен диссоциирующими белками и не может быть использован для нуклеиновых кислот или небиологических объектов (например, наночастиц). Недавно был предложен модифицированный метод хемометрического разложения, который был применен для анализа данных МУРР процессов формирования фибрилл белка амилоида [194] с использованием процедуры MCR-ALS на различных взвешенных наборах данных МУРР. Этот метод, реализованный в программе COSMICS, не требует знания кривых рассеяния от чистых компонентов, и при некоторых ограничениях, основанных на физической природе системы, с помощью данного метода могут быть определены кривые рассеяния от компонентов и их относительные объемные доли.

Также было предложено несколько подходов для выбора подмножеств из ансамблей возможных конформаций макромолекул. К ним относятся, например, метод оптимизации ансамбля (ЕОМ) [89, 90] для выбора лучшего подмножества моделей из большого ансамбля или подход ограниченного максимального правдоподобия для оценки относительных вкладов различных компонентов [195]. Однако эти методы требуют, чтобы (предварительные) модели были известны или предсказаны заранее.

Важно отметить, что очень часто при изучении эволюционирующих систем начальное и конечное состояние процесса достаточно стабильно, и картины рассеяния от этих состояний могут быть надежно измерены (или хорошо аппроксимированы теоретическими моделями). В ходе динамического процесса образуется основная промежуточная структура, которая исчезает в конце, так что ее рассеяние и, следовательно, форма не могут быть получены непосредственно из экспериментальных данных. Вместо этого регистрируют несколько кривых рассеяния от такой системы, представляющих собой смеси различных состояний. Наличие промежуточного состояния можно обнаружить с помощью независимых подходов, таких как сингулярное разложение (SVD: Singular Value Decomposition), которые показывают, что для описания собранных данных требуется более двух независимых компонентов (в случае четко определенного промежуточного состояния - три компонента) [120]. Однако метод SVD не позволяет восстановить кривую рассеяния от промежуточного состояния системы и, следовательно, сделать выводы о его структуре. Чтобы решить эту задачу, нами был разработан новый метод прямого восстановления формы неизвестного промежуточного эволюционирующей системе вместе с объемными состояния компонентов во всех измеренных временных точках. Метод основан на совместном анализе нескольких наборов данных МУРР и дает наилучшее общее соответствие всем имеющимся данным МУРР. В следующих разделах будут приведены детали данного подхода и примеры его практического применения.

5.1 Memod ab initio определения формы промежуточного компонента в эволюционирующей системе

Пусть имеется К кривых рассеяния, собранных от эволюционирующей системы (например, кинетические измерения с временным разрешением). В начале система обладает интенсивностью $I_m(s)$ (например, мономерное состояние); в конце есть конечное определенное состояние (например, крупный агрегат $I_a(s)$). Предполагается, что две интенсивности рассеяния $I_m(s)$ и $I_a(s)$ известны, но в ходе

реакции образуется промежуточный компонент, структура которого и кривая рассеяния $I_i(s)$ неизвестны. Интенсивность рассеяния в k-й точке реакции представляет собой линейную комбинацию

$$I_k(s) = v_{mk}I_m(s) + v_{ak}I_a(s) + v_{ik}I_i(s)$$
(5.1)

где v_{mk} , v_{ak} и v_{ik} - объемные доли компонентов, $v_{mk} + v_{ak} + v_{ik} = 1$.

Если картина рассеяния от промежуточного компонента $I_i(s)$ может быть получена из эксперимента, то форма частицы с низким разрешением может быть восстановлена ab initio, например, с помощью представления модели в виде плотноупакованных шариков [52]. В этом случае прямое пространство поиска представляется плотно упакованной сеткой маленьких сфер (фиктивных атомов) достаточно малого радиуса r << R, где R - характерный размер частицы. Каждая сфера может принадлежать либо частице (индекс=1), либо растворителю (индекс=0). Интенсивность рассеяния от конфигурации шариков $I_{DAM}(s)$ быстро вычисляется с использованием сферических гармоник [47]. Алгоритм имитации отжига (SA) [59] используется для поиска оптимальной формы частицы (т.е. вектора X, содержащего информацию о фазовой принадлежности шариков) путем случайного изменения фазовой принадлежности одного фиктивного атома за одну итерацию, чтобы в конечном итоге минимизировать целевую функцию:

$$F(X) = \chi^{2}(X) + P(X) \tag{5.2}$$

де $\chi^2(X)$ - невязка между экспериментальной и расчетной кривыми, а P(X) – штрафное слагаемое, гарантирующее, что форма частицы является физически обоснованной (компактной, взаимосвязанной, с центром масс, близким к началу координат; и если требуется, имеющей заданную анизометрию). Данный подход реализован в программах DAMMIN [52], использующей поиск формы в ограниченном пространстве с максимальным диаметром 2R, и DAMMIF [55], использующей неограниченное пространство поиска.

Для эволюционирующей системы интенсивность промежуточного компонента $I_i(s)$ не может быть измерена; вместо этого информация о ней зашифрована в измеренных данных с помощью уравнения (5.1). Поэтому нами был предложен

обобщенный метод использования этой информации и восстановления формы промежуточного компонента вместе с неизвестными объемными долями трех компонентов путем одновременного приближения всех экспериментальных картин рассеяния $I_k(s)$, k=1...K.

В этом случае, обобщенная функция, которую необходимо минимизировать F(X), рассчитывается как:

$$F(X) = \sum_{k=1}^{K} \chi_k^2(X) + \sum_j W_j * P_j(X)$$
 (5.3)

Здесь первое слагаемое - это общая невязка между экспериментальными и расчетными данными по К экспериментальным кривым. Второй член содержит физические ограничения, аналогичные ограничениям ДЛЯ программ DAMMIN/DAMMIF в уравнении (5.2), и включает штраф за рыхлость и компактность модели, а также штраф за геометрический центр и радиус инерции (R_g) , удерживающий модель вблизи начала координат [55]. Кроме того, этот член содержит штраф за минимальную долю, который гарантирует, что средняя объемная доля промежуточного (неизвестного) компонента не должна стать меньше 5%, чтобы избежать не имеющих физического смысла решений с пренебрежимо малым вкладом данного компонента в измеренные данные. Веса штрафов (W_i) подбирают так, чтобы обеспечить баланс между общей невязкой и физическими ограничениями моделей. По умолчанию веса для степени рыхлости модели и геометрического центра/ R_g такие же, как в DAMMIF [55], а значение для штрафа за объемную долю равно 0.5. Подобранные по умолчанию значения штрафных весов хорошо работают в большинстве случаев, но при необходимости они могут быть изменены пользователем.

Восстановление формы частицы осуществляется по алгоритму DAMMIF без ограничений на пространство поиска. Метод автоматически предлагает размер шарика, обеспечивающий адекватное представление относительно небольших промежуточных продуктов, таких как олигомеры исходного состояния. В качестве начального приближения для промежуточного состояния выбирается сфера с объемом $(V_m + V_a)/2$ $(V_m$ и V_a — объемы частиц начального и конечного состояния

системы, соответственно), и метод имитации отжига (SA) используется для поиска 3D формы промежуточного компонента, для того чтобы минимизировать F(X) в уравнении (5.3). Поскольку пространство поиска неограничено то, чтобы избежать получения слишком больших форм, "конкурирующих" с рассеянием от конечного состояния, объем промежуточной модели ограничивается (по умолчанию, он не должен превышать объем мономера более чем в 50 раз). На каждом шаге SA алгоритма форма частицы случайным образом изменяется, как и в программе DAMMIF, интенсивность $I_i(s)$ пересчитывается и определяются объемные доли компонентов, которые дают наилучшее общее соответствие измеренным данным МУРР. Для каждого набора данных это делается путем приближения экспериментальных данных с помощью уравнения (1) с ограничениями на неотрицательность объемных долей (аналогично программе OLIGOMER [A5]). Приведенный алгоритм реализован в виде компьютерной программы DAMMIX [A72], результаты его применения будут представлены в следующих разделах.

5.2 Проверка эффективности работы алгоритма на теоретических данных МУРР

Сначала метод был протестирован на теоретических данных МУРР, описывающих процессы ассоциации частиц. В примере, представленном на Рис. 5.1, для системы эллипсоидо-подобных частиц были рассчитаны десять теоретических кривых рассеяния, представляющих собой линейные комбинации от профилей рассеяния эллипсоидом с полуосями 20, 30 и 70 Å (соответствует начальному "мономерному" состоянию), G-подобной структурой, составленной из пяти эллипсоидов (соответствует промежуточному "олигомерному" состоянию) и агрегатом, составленным из пяти G-подобных структур (конечное "агрегатное" состояние системы). Используя такой теоретический набор данных МУРР с заданным изменением объемных долей компонентов и добавлением шума с относительным уровнем в 2 %, а также рассчитанные кривые рассеяния от мономеров и агрегатов, программа DAMMIX восстановила G-подобную форму промежуточного

состояния и получила объемные доли компонентов в пределах 3 % по сравнению с (заданными) (Рис. 5.1, вставка). значениями Изменчивость восстановленных моделей оценивалась по нормированному пространственному расхождению (NSD) с помощью программы SUPCOMB [57], аналогично процедуре для программы DAMAVER [58]. Среднее значение NSD было рассчитано по 15 отдельным запускам программы DAMMIX с использованием различных последовательностей генератора случайных чисел и составило <NSD>=(0.92±0.08). Поскольку значения NSD около единицы соответствует хорошей корреляции между моделями, можно сделать вывод о воспроизводимости полученных результатов. Процедура сравнения также автоматически выбирает наиболее типичную реконструкцию формы (т.е. модель, которая имеет наилучшее перекрытие, т.е. минимальный средний NSD относительно всех других моделей), они представлены на Рис. 5.1 и всех последующих Рисунках Главы 5. Вариации найденных объемных долей компонентов между разными запусками DAMMIX остаются в пределах 3%, число олигомеризации промежуточного компонента близко к ожидаемому пентамеру, а отклонения между восстановленными кривыми промежуточного компонента находятся в пределах ошибок (Таблица 5.1). Эти результаты подтверждают воспроизводимость и надежность реконструкции формы промежуточного компонента с помощью предложенного нами алгоритма.

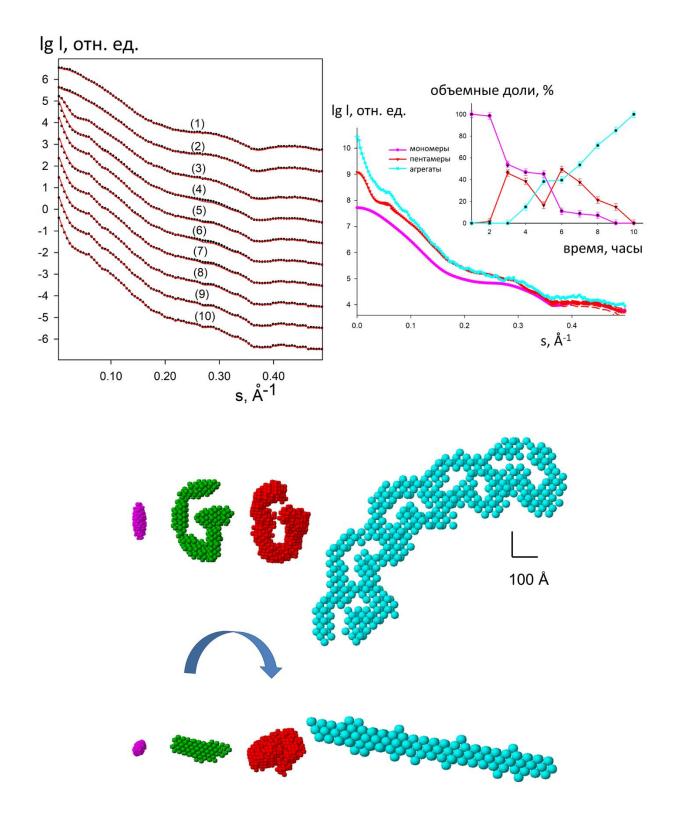


Рис. 5.1 Модельные кривые малоуглового рассеяния от эволюционирующей системы (начальное состояние - эллипсоид, промежуточное состояние - G-подобная структура из пяти эллипсоидов, конечное состояние - агрегат, состоящий из пяти G-подобных промежуточных структур). Кривые рассеяния для начального и конечного состояния системы были рассчитаны по шариковым моделям с помощью программы DAMMIN, их линейные комбинации с "проектными"

объемными долями были получены с помощью программы PRIMUS [A5]. К модельным данным был добавлен шум с относительным уровнем в 2%. На верхней левой панели модельные данные показаны точками, а результаты моделирования программой DAMMIX - красными сплошными линиями. Формы компонентов (эллипсоид, G-структура, агрегат) показаны на нижней панели малиновыми, зелеными и голубыми сферами соответственно. Типичная восстановленная форма промежуточного компонента с помощью программы DAMMIX показана на нижней панели красными сферами. Масштаб составляет 100 Å. Кривые рассеяния от компонентов показаны на верхней правой панели (две самые разные кривые рассеяния промежуточного компонента, полученные в результате нескольких запусков программы DAMMIX, показаны пунктирными красными линиями), их восстановленные объемные доли показаны на вставке (цвета те же, что и на нижней панели) вместе с фактическими объемными долями, использованными в моделировании (черные кружки); интервалы ошибок объемных долей показывают их средний разброс по нескольким запускам программы DAMMIX.

Нами было проведено еще несколько тестов на различных теоретических наборах данных МУРР, и для всех этих синтетических примеров программа DAMMIX восстанавливала промежуточные состояния системы с высокой степенью надежности. В качестве иллюстрации на Рис. 5.2 представлены результаты для системы сфера-цилиндр-призма, имитирующей процесс агрегации промежуточных компонентов (цилиндров), где программа DAMMIX успешно восстановила цилиндрическую форму промежуточного состояния и объемные доли всех компонентов.

Таблица 5.1 Результаты многократных запусков программы DAMMIX (для каждой системы было проведено 15 отдельных запусков) для модельных (теоретических) и экспериментальных наборов данных МУРР. Во втором столбце приведены средние значения пространственного несоответствия <NSD> между попарно наложенными моделями, полученными при разных запусках программы DAMMIX для промежуточных компонентов; $<\Delta v_{beg}>$, $<\Delta v_{int}>$ и $<\Delta v_{end}>$ - средние дисперсии объемных долей для начального, промежуточного и конечного состояний системы, соответственно; <N_{olig}> - средняя степень олигомеризации для промежуточных компонентов; $\langle \chi^2 \rangle$ - средние невязки для восстановленных кривых рассеяния от полученных промежуточных компонентов, программой DAMMIX. гле соответствующие ошибки были оценены согласно статистике Пуассона.

Набор данных	< <i>NSD</i> >	$<\Delta v_{beg}>$,	$<\Delta v_{int}>$,	$<\Delta v_{end}>$,	<nolig></nolig>	<\chi^2>
		%	%	%		
Эллипсоид	0.92 ± 0.08	1.9	2.1	1.6	5.1±0.3	1.13
(Рис. 5.1)						
Цилиндр	1.02±0.09	2.0	2.3	1.6	8.1±0.2	1.08
(Рис. 5.5)						
Инсулин	0.87 ± 0.06	1.8	2.2	1.5	4.6±0.4	1.05
(Рис. 5.2)						
Люмазин-	0.94 ± 0.05	1.79	2.4	1.8	N/A	1.03
синтаза						
(Рис. 5.3)						
hNGF	0.85±0.06	1.6	1.7	N/A	2.1±0.2	1.04
(Рис. 5.4)						
α-Synuclein	1.22±0.12	2.2	2.5	1.9	8.3±0.5	1.37
(Рис. 5.6)						

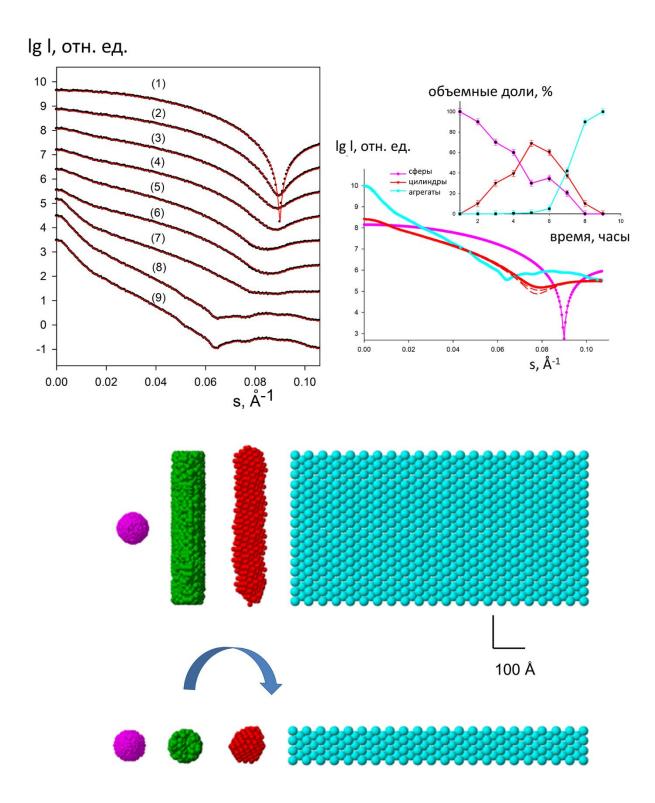


Рис. 5.2. Модельные кривые малоуглового рассеяния от эволюционирующей системы (начальное состояние - сфера диаметром 100 Å, промежуточное состояние - цилиндр диаметром 100 Å и высотой 500 Å, конечное состояние - прямоугольная призма с размерами сторон 100, 500 и 1000 Å). Кривые рассеяния для начального и конечного состояния системы были рассчитаны по шариковым моделям с помощью программы DAMMIN, их линейные комбинации с "проектными"

объемными долями генерировались с помощью программы PRIMUS [A5]. К модельным данным был добавлен шум с относительным уровнем в 2%. На верхней левой панели модельные данные показаны точками, а результаты моделирования программой DAMMIX - красными сплошными линиями. Идеальные формы компонентов (сфера, цилиндр, прямоугольная призма) показаны на нижней панели малиновыми, зелеными и голубыми шариками, соответственно. Типичная восстановленная форма промежуточного элемента с помощью программы DAMMIX показана на нижней панели красными шариками. Масштаб составляет 100 Å. Кривые рассеяния от компонентов показаны на верхней правой панели (две самые разные кривые рассеяния промежуточного компонента, полученные в результате нескольких запусков программы DAMMIX, показаны пунктирными красными линиями), их восстановленные объемные доли показаны на вставке (цвета те же, что и на нижней панели) вместе с фактическими объемными долями, использованными в моделировании (черные кружки); интервалы ошибок объемных долей показывают их средний разброс по нескольким запускам программы DAMMIX.

Таким образом, алгоритм, реализованный в программе DAMMIX показал свою высокую эффективность на ряде теоретических наборов данных (приближенных к реальным, благодаря добавлению шума), но конечно, особенно интересным представляется применение метода к реальным экспериментальным данным МУРР от эволюционирующих систем, которые будут рассмотрены в следующем разделе.

5.3 Применение метода для ряда трехкомпонентных эволюционирующих систем и двухкомпонентных олигомерных смесей

5.3.1. Процесс образования фибрилл белка инсулина

Процесс образования амилоидных фибрилл является нуклеационно-зависимым процессом, и характеристика ядер (зародышей) чрезвычайно важна для понимания его механизма. Временной набор данных МУРР, собранный во время процесса образования фибрилл белка инсулина [189], себя 15 включает экспериментальных кривых, измеренных во время четырехчасовой фазы роста, которая началась после пятичасового периода инкубации (Рис. 5.3). Сразу отметим, что данные МУРР не могут быть представлены линейными комбинациями только мономеров (начальное состояние) и зрелых фибрилл (конечное состояние). Сингулярное разложение набора данных МУРР однозначно указывает, что существуют минимум три компонента, вносящие значительный вклад в кривые рассеяния в процессе роста (удлинения) фибрилл.

Предложенный нами подход был применен к указанному набору данных МУРР и позволил хорошо приблизить кривые рассеяния во всех временных точках. Форма промежуточного компонента инсулина, объемные доли мономеров, промежуточных и зрелых фибрилл, полученные с помощью программы DAMMIX, и их соответствие экспериментальным данным МУРР представлены на Рис. 5.3. Интересно отметить, что восстановленная форма промежуточного компонента представляет собой вытянутую структуру, состоящую из пяти связанных доменов по размерам близких к мономерам инсулина. Эта форма прекрасно согласуется с гипотетической моделью олигомера инсулина, предложенной в работе [189], которая была построена из мономерных моделей белка инсулина с помощью процедуры итеративного моделирования. Полученная нами *ab initio* реконструкция промежуточного звена фибриллы очень хорошо соответствует более ранней модели и дает дополнительное подтверждение гипотезе о том, что олигомерное ядро является первичной удлиняющей единицей инсулиновых амилоидных фибрилл. Степень стабильности результатов восстановления при повторных DAMMIX представлена Таблице 5.1. Из нее запусках видно, что восстанавливаемые формы хорошо согласуются друг с другом.

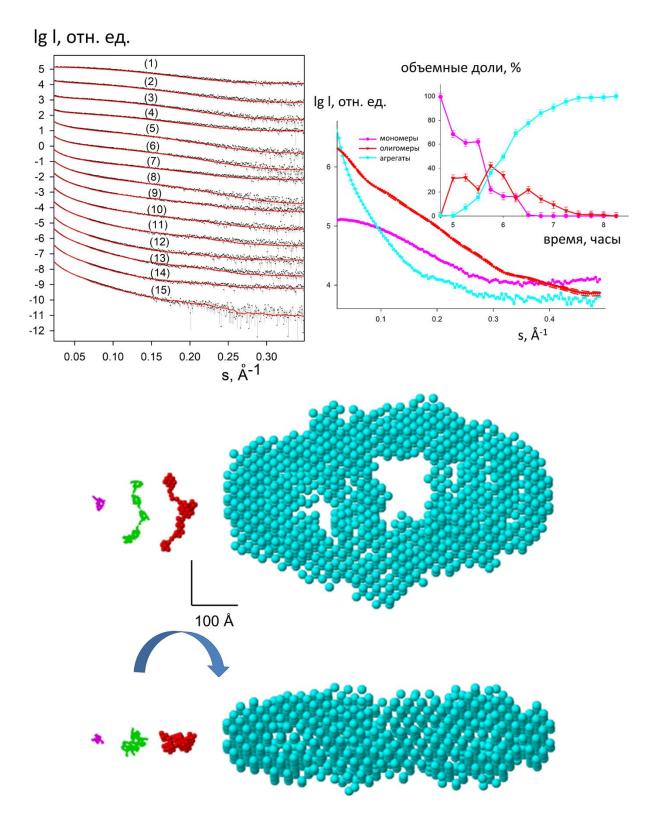


Рис. 5.3. Анализ промежуточного состояния во время процесса образования фибрилл белка инсулина. Начальное состояние системы соответствует мономеру инсулина (PDB ID: 1GUJ), конечное состояние аппроксимируется шариковой моделью зрелой фибриллы [189]. На верхней левой панели показаны экспериментальные данные (кружки с интервалами ошибок), а красными

сплошными линиями - результаты моделирования программой DAMMIX. Формы мономера и агрегатов показаны на нижней панели малиновыми и голубыми шариками, соответственно; ранее представленная промежуточная пентамерная модель показана зеленым С-образным следом [189]. Восстановленная форма неизвестного промежуточного компонента, полученная с помощью программы DAMMIX, показана на нижней панели (красные шарики). Масштаб составляет 100 Å. Кривые рассеяния от компонентов показаны на верхней правой панели (две самые разные кривые рассеяния промежуточного компонента, полученные в результате нескольких запусков программы DAMMIX, показаны пунктирными красными линиями), их восстановленные объемные доли показаны на вставке (цвета те же, что и на нижней панели); интервалы ошибок объемных долей показывают их средний разброс по нескольким запускам программы DAMMIX.

5.3.2. Множественные состояния само-сборки фермента люмазинсинтазы

Способность белков образовывать различные четвертичные структуры необходима для многих биологических процессов, таких как передача сигнала, регуляция клеточного цикла и катализ ферментов. Одним из таких представителей является белок-фермент люмазинсинтаза, катализирующий предпоследний этап биосинтеза рибофлавина (витамин В2). Совместное исследование МУРР и криоЭМ общей показало, что множественные состояния само-сборки являются особенностью этой образует [A13]. Люмазинсинтаза системы икосаэдрические капсулы с диаметром 160 Å (T1) и более крупные капсулы с диаметром около 300 Å (капсиды Т3). Относительное количество маленьких и больших капсул сильно зависит от буфера и рН раствора, а также от наличия мутаций в белке. В работе [А13] было показано, что помимо двух видов капсул в растворе могут присутствовать свободные грани диссоциированных или неполных капсул, объемные доли которых были оценены ранее с помощью программы MIXTURE (см. Главу 4). Таким образом, данная система может быть описана тремя независимыми компонентами и вполне подходит для применения разработанного подхода.

Набор экспериментальных данных МУРР, проанализированных с помощью программы DAMMIX, содержал в общей сложности двенадцать кривых рассеяния

люмазинсинтазы дикого типа и ее мутантных форм, полученных из Aquifex aeolicus в фосфатном буфере (при рН 6.0÷8.0) и Tris буфере (при рН 7.0÷9.0) и люмазинсинтазы дикого типа из Bacillus subtilis в боратном буфере (pH 7.0), содержащей преимущественно T1 капсулы и ее мутантной формы в Tris буфере (рН 7.0) с преобладанием ТЗ капсул (Рис. 5.4). Программа DAMMIX позволила приблизить весь набор данных МУРР и восстановить форму диссоциированных фрагментов капсул в растворе, а также объемные доли всех компонентов. Восстановленный компонент по форме и размеру очень похож на мономер люмазинсинтазы (правая панель Рис. 5.4), а объемные доли мономера хорошо согласуются с ранее полученными результатами (вставка Рис. 5.4). Этот пример демонстрирует высокий потенциал алгоритма при поиске форм неизвестных компонентов для систем, образующих в процессе само-сборки несколько состояний (например, вирусоподобные частицы ИЛИ наночастицы, стабилизированные полимерными цепями). Степень изменчивости модели и средняя дисперсия объемных долей по нескольким запускам программы DAMMIX показаны в Таблице 5.1.

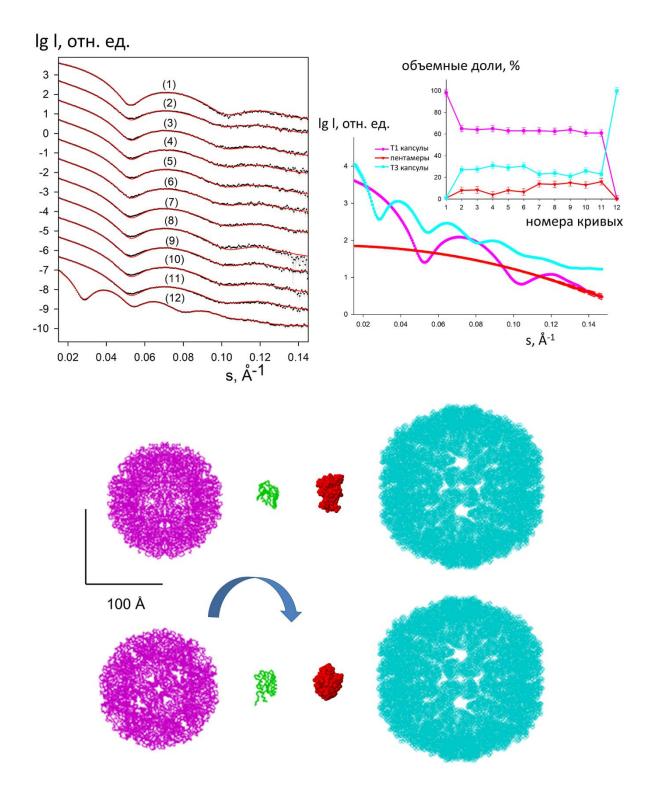


Рис. 5.4 Обнаружение промежуточного компонента в процессе формирования капсул фермента люмазинсинтазы. Левая верхняя панель показывает данные, измеренные при различных условиях (кривая 1, люмазинсинтаза дикого типа из Bacillus subtilis (LSBS) в боратном буфере (рН 7. 0); кривые 2-6, люмазинсинтаза дикого типа из Aquifex aeolicus (LSAQ) в фосфатном буфере (рН 6.0,6.5,7.0,7.5,8.0); кривые 7-11, LSAQ в трис-буфере (рН 7.0,7.5,8.0,8.5,9.0), кривая 12, мутант LSAQ-IDEA в трис-буфере (рН 8.0). На верхней левой панели показаны

экспериментальные данные (кружки с интервалами ошибок), а красными сплошными линиями - результаты моделирования программой DAMMIX. Известные "начальное" и "конечное" состояния, капсула Т1 диаметром 160 Å и капсула Т3 диаметром 300 Å, соответственно, показаны на нижней панели малиновыми и голубыми шариками. Экспериментальные данные МУРР от LSBS в боратном буфере и мутанта LSAQ-IDEA в Трис-буфере (кривые 1 и 12, соответственно), соответствующие этим моделям, были использованы в качестве входных данных в программе DAMMIX после их регуляризации с помощью программы GNOM [40]. Мономер люмазинсинтазы (PDB ID: 1RVV) показан зеленой С_α цепью, а типичная восстановленная форма неизвестного компонента (диссоциированные фрагменты капсул), полученная с помощью программы DAMMIX, показана красными шариками. Масштаб составляет 100 Å. Кривые рассеяния от компонентов показаны на верхней правой панели (две самые разные кривые рассеяния промежуточного компонента, полученные в результате нескольких запусков программы DAMMIX, показаны пунктирными красными линиями), их восстановленные объемные доли показаны на вставке (цвета те же, что и на нижней панели); интервалы ошибок объемных долей показывают их средний разброс по нескольким запускам программы DAMMIX.

5.3.3. Концентрационно-зависимая олигомеризация белка фактора роста нервов С помощью предложенного метода можно также изучать двухкомпонентные равновесные смесы (например, смесы типа мономер одигомер), в этом слугае

равновесные смеси (например, смеси типа мономер-олигомер), в этом случае конечное состояние системы в уравнении (5.1) отсутствует, а промежуточное состояние сосуществует только с начальным состоянием. Примером такой системы может служить фактор роста нервов (NGF) - белок, играющий ключевую роль в процессах выживания, дифференциации и поддержания популяций нейронов. NGF является функциональным гомодимером, состоящим из двух нековалентносвязанных цепей, и его олигомерный состав в растворе зависит от концентрации белка. Исследование данных МУРР [А24] показало, что белок NGF образует смесь димера и тетрамера (или димера димеров в конфигурации "голова к голове"). Система содержит всего два компонента, и программа DAMMIX может быть применена для нахождения неизвестной формы димера димеров NGF с учетом известной структуры димера NGF (начального состояния системы при низкой концентрации белка).

Экспериментальные кривые МУРР от растворов NGF, полученные в диапазоне концентраций белка от 0.43 до 5.5 мг/мл, представлены на Рис. 5.5. Рассеяние от начального состояния системы рассчитывалось по димерной структуре PDB (код PDB: 1BTG), а "конечное" состояние системы в рамках алгоритма DAMMIX устанавливалось отсутствующим (т.е. интенсивность и объемная доля третьего компонента системы были приравнены нулю). Восстановленный с помощью программы DAMMIX олигомер имел примерно вдвое больший размер по сравнению с димером "начального состояния" и форму, похожую на структуру димера димеров в конфигурации "голова к голове", о которой сообщалось в работе [A24]. Восстановленные объемные доли также хорошо согласуются с данными, полученными в более раннем исследовании (Рис. 5.5, вставка) [А22]. Степень воспроизводимости восстановленных структур димера димеров и средняя дисперсия объемных долей в олигомерной смеси, оцененные по нескольким независимым запускам программы DAMMIX показаны в Таблице 5.1. Эти результаты показывают, что и для двухкомпонентных систем программа DAMMIX получает разумные решения и может быть использована, например, для количественного анализа олигомерных равновесных смесей.

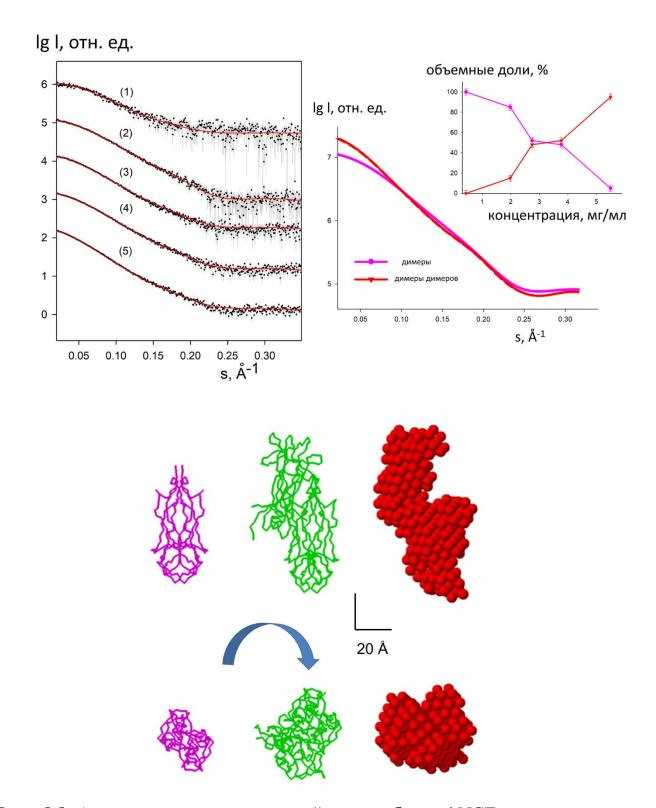


Рис. 5.5 Анализ состава олигомерной смеси белка hNGF в зависимости от концентрации. Начальное состояние: модель димера (PDB ID: 1BTG), промежуточное состояние - неизвестный олигомер, конечное состояние - установлено как 'none' для программы DAMMIX). На верхней левой панели показаны экспериментальные данные (кружки с интервалами ошибок), а красными сплошными линиями - результаты моделирования программой DAMMIX. Ранее

представленные модели (димер и димер димеров в конфигурации "голова к голове") показаны на нижней панели малиновыми и зелеными C_{α} цепями [A24]. Типичная восстановленная форма неизвестного компонента, полученная с помощью программы DAMMIX, показана на нижней панели красными шариками. Масштаб составляет 20 Å. Кривые рассеяния от компонентов показаны на верхней правой панели (две самые разные кривые рассеяния промежуточного компонента, полученные в результате нескольких запусков программы DAMMIX, показаны пунктирными красными линиями), их восстановленные объемные доли показаны на вставке (цвета те же, что и на нижней панели); интервалы ошибок объемных долей показывают их средний разброс по нескольким запускам программы DAMMIX.

5.3.4. Общие замечания по работе алгоритма

Интерпретация данных МУРР/МУРН с использованием трехмерных моделей неоднозначна для монодисперсных систем [60] и, тем более, для смесей. Поэтому неудивительно, что до сих пор не существовало методов, позволяющих в общем форму случае напрямую восстановить неизвестного эволюционирующей системе. В предложенном нами подходе мы рассмотрели случай трехкомпонентной системы с одним неизвестным промежуточным компонентом. Он представляет собой комбинацию ab initio определения формы с неотрицательной линейной минимизацией для разложения компонентов смеси и предназначен для восстановления формы промежуточного компонента и объемных долей компонентов. Метод, по сути, работает как обычное определение формы частицы только с двумя дополнительными параметрами (т.е. тремя объемными долями, связанными соотношением: $v_{mk}+v_{ak}+v_{ik}=1$ при одновременном приближении всего набора данных МУРР. Как было показано в предыдущих разделах на примерах теоретических и экспериментальных данных МУРР, метод действительно позволяет надежно восстанавливать форму для различных типов эволюционирующих систем путем приближения всего массива измеренных данных.

Очевидно, что перед использованием DAMMIX необходимо провести соответствующие тесты (например, использовать сингулярное разложение -

мSVD), чтобы убедиться, что система действительно может быть представлена тремя основными компонентами. Следующий пример наглядно демонстрирует, что произойдет, если применить DAMMIX к системе с бoльшим числом Ранее было эволюционирующих компонентов. показано, что процесс фибриллобразования α-синуклеина описывается четырьмя основными компонентами, и содержит, помимо мономеров и зрелых фибрилл, также димерные и олигомерные промежуточные продукты [190]. В этом случае, как показано на Рис. 5.6, DAMMIX восстанавливает форму, которая находится где-то посередине между формой димерных и олигомерных промежуточных продуктов, причем для части кривых рассеяния из всего набора данных МУРР наблюдается существенное расхождение между экспериментальными и рассчитанными кривыми, что указывает на неполную адекватность описания системы трехкомпонентной моделью. При анализе таких многокомпонентных эволюционирующих систем следует выбирать подмножества данных, в которых SVD-анализ указывает на наличие трех компонентов, и последовательный анализ соответствующих подмножеств данных может позволить восстановить уже несколько неизвестных промежуточных состояний системы. Хемометрический подход [194] также может оказать большую помощь в мониторинге более сложных путей динамических временных процессов и, в конечном итоге, найти подмножества, где применим алгоритм DAMMIX.

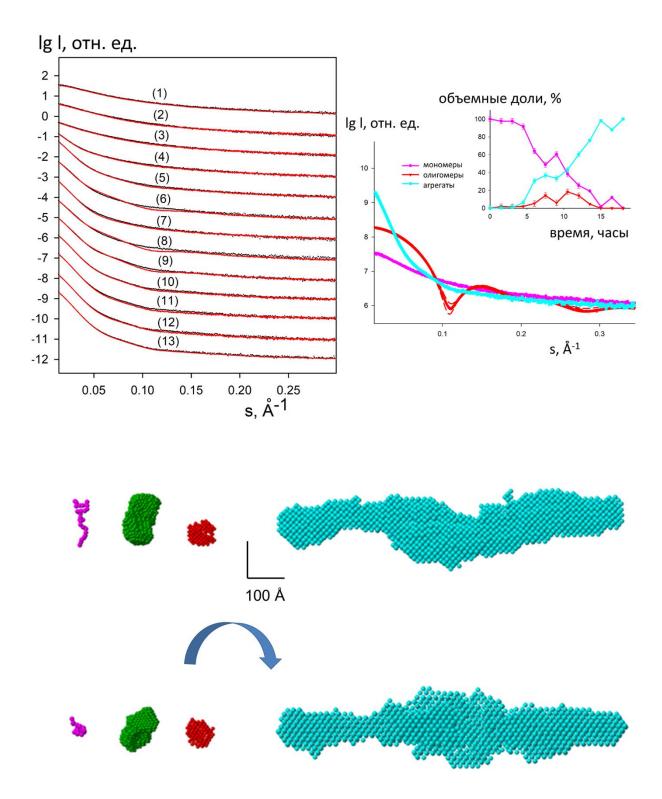


Рис. 5.6 Анализ промежуточных компонентов во время процесса образования фибрилл белка α-синуклеина. Начальное состояние соответствует мономеру α-синуклеина, а конечное состояние – шариковой модели для зрелой фибриллы [190]. На верхней левой панели показаны экспериментальные данные (кружки с интервалами ошибок), а красными сплошными линиями - результаты моделирования программой DAMMIX. Формы мономера и агрегатов показаны на

нижней панели малиновыми и голубыми шариками, соответственно; ранее упоминавшийся промежуточный олигомер (16-мер) показан зелеными шариками [190]. Присутствие четвертого компонента в системе подтверждается сингулярным разложением данных МУРР, и этот компонент, по-видимому, содержит димеры, что было подтверждено гель-фильтрацией [190]. Масштаб составляет 100 Å. Восстановленная форма промежуточного компонента, полученная с помощью программы DAMMIX, показанная на нижней панели (красные шарики), выглядит примерно в два раза меньше, чем 16-мер, являясь фактически средним значением двух видов промежуточных компонентов (димеров и 16-меров), о которых сообщалось в [190]. Кривые рассеяния от компонентов показаны на верхней правой панели (две самые разные кривые рассеяния промежуточного компонента, полученные в результате нескольких запусков программы DAMMIX, показаны пунктирными красными линиями), их восстановленные объемные доли показаны на вставке (цвета те же, что и на нижней панели); интервалы ошибок объемных долей показывают их средний разброс по нескольким запускам программы DAMMIX. Примечательно, что в приближениях программой DAMMIX имеются систематические отклонения от экспериментальных данных в области средних углов для нескольких временных точек во время процесса образования фибрилл, что указывает на недостаточность использования трех компонентов для адекватного описания системы.

Важно заметить, что анализ полидисперсных систем является весьма сложной задачей, и ни один из доступных методов, каждый из которых имеет свои ограничения, не дает уникального рецепта для решения неоднозначности, присущей данным МУРР. Для DAMMIX, который представляет собой автоматизированную процедуру определения формы, никогда не следует забывать, что энантиоморфная структура дает идентичное рассеяние, и энантиоморфы учитываются во всех процедурах усреднения и отбора. Также надо помнить, что даже небольшое 'загрязнение' кривой (кривых) начального и конечного состояний примесными видами может повлиять на качество восстановления формы промежуточного состояния. Для предотвращения таких эффектов потенциальные примеси должны быть тщательно проверены путем оценки молекулярного веса и исключенного объема образца, а также линейности области Гинье [5] и формы функции p(r) (в случае наличия 'загрязнения' p(r) обычно показывает дополнительный хвост на больших расстояниях) по соответствующим кривым рассеяния. Если в наличии имеются структуры высокого разрешения для

начальных/конечных состояний системы, то они могут быть использованы в качестве входных данных для программы DAMMIX, а теоретические кривые, рассчитанные программой CRYSOL [70], можно использовать в качестве основных компонентов исследуемой смеси.

Предложенный метод также применим для двухкомпонентных эволюционирующих систем, когда один из компонентов (например, "мономер") известен, что позволяет восстановить структуру другого компонента. Этот случай можно считать полезным дополнением к алгоритму GASBORMX [A7], который хоть и не требует знания структуры мономера, но ограничен симметричными белковыми олигомерами, тогда как предлагаемый подход свободен от ограничений симметрии и может быть использован также для небиологических объектов.

Заключение к Главе 5

Рассмотренный в Главе 5 новый подход по прямому анализу формы промежуточного состояния безусловно будет полезен для интерпретации широкого спектра кинетических экспериментов МУРР/МУРН с временным разрешением от эволюционирующих биологических систем. В целом, предложенный алгоритм дополняет хемометрические методы анализа, причем независимое применение различных методов особенно полезно для повышения надежности результатов.

Глава 6

Метод эволюционного факторного анализа для восстановления индивидуальных профилей рассеяния компонентов белковыми смесями по данным МУРР с использованием хроматографической колонки

Точное определение структурных параметров и анализ 3D формы биологических макромолекул с помощью малоуглового рентгеновского рассеяния (МУРР) требует наличия очищенных монодисперсных растворов. На практике ситуация осложняется тем, что биологические образцы часто присутствуют в виде смесей отдельных компонентов, что затрудняет анализ данных МУРР. Однако, если соединить этап гель-хроматографического разделения с измерениями МУРР (режим измерений SEC-SAXS, Size Exclusion Chromatography – Small-Angle X-ray Scattering), то становится возможным разделить вклады отдельных компонентов, присутствующих в системе. Хотя гель-хроматография чрезвычайно эффективна МУРР по-прежнему остается при работе со смесями, анализ данных нетривиальным, когда компоненты образца плохо разделяются, например, при наложении пиков выхода компонентов на хроматографическом профиле. Фактически данные МУРР представляют собой линейные комбинации вкладов эволюционирующей (выходящей рассеяния компонентов смеси ИЗ хроматографической колонки), взвешенных пропорционально их объемным долям в растворе, и, если компоненты перекрываются, то прямое разделение их сигналов во времени оказывается невозможным. Анализ таких данных МУРР требует процедуры декомпозиции для оценки количества компонентов смеси и дальнейшего восстановления их профилей рассеяния набору ПО всему экспериментальных данных МУРР.

Существует несколько хемометрических алгоритмов, которые могут справиться с такой задачей разделения вкладов компонентов, в частности, метод альтернирующих наименьших квадратов (ALS – Alternating Least Squares) [196, 197] и эволюционный факторный анализ (EFA – Evolutionary Factor Analysis) [198, 199]. Эти алгоритмы были успешно применены для анализа данных SEC-SAXS

[200-202] при исследовании динамики в амилоидных системах [194], слабосвязанных переходных комплексах белков [203], процессов сворачивания белков [204] и равновесных олигомерных смесей [193], а также данных ионнообменной хроматографии (IEC-SAXS, Ion-Exchange Chromatography-SAXS) и исследований МУРР с временным разрешением (TR-SAXS, Time Resolved-SAXS) [205]. В последнем случае метод ALS сочетался с регуляризацией по Тихонову [206].

Нами был разработан алгоритм, реализованный в виде программы EFAMIX [A73] для восстановления профилей рассеяния и концентрации отдельных компонентов по набору данных SEC-SAXS с использованием подхода EFA, сингулярного разложения [120] и метода вращения матрицы [199]. Подход на основе EFA с модифицированным гибридным методом вращения матрицы был ранее реализован в программе BioXTAS RAW [201]. Однако, реализованный в BioXTAS RAW итеративный алгоритм поиска имеет ограничения по сходимости и требует ручного выбора пользователем параметров для оценки количества компонентов. Тогда как программа EFAMIX требует минимального вмешательства пользователя при анализе данных SEC-SAXS и предоставляет возможность автоматической оценки компонентов. Разработанный параметров метод помогает разложить перекрывающиеся пики в профилях SEC-SAXS на индивидуальные составляющие компонентов, и как будет показано в следующих разделах, метод применим, в том числе, и к данным IEC-SAXS с умеренной степенью градиента солевого буфера. Эффективность работы алгоритма EFAMIX будет продемонстрирована на большом ряде теоретических и экспериментальных наборов данных МУРР.

6.1 Метод эволюционного факторного анализа (ЭФА) для анализа набора данных МУРР в сочетании с гель-хроматографией

6.1.1 Общая концепция метода эволюционного факторного анализа

Эволюционный факторный анализ (ЭФА) - это модельно-независимый подход для анализа матриц данных, составленных из одномерных кривых (например, спектров) от многокомпонентных систем, в которых наблюдается последовательное, но неполное разделение компонентов. Типичным примером таких данных являются измерения МУРР, записанные во время последовательного выхода (элюирования) раствора образца из хроматографической колонки, в которых происходит временное наложение компонентов.

Как уже было ранее сказано, в эксперименте МУРР одномерные кривые интенсивности рассеяния I(s) измеряются как функции момента передачи импульса (вектора рассеяния) $s=4\pi \sin\theta/\lambda$, где 2θ - угол рассеяния, а λ - длина волны рентгеновского излучения. Набор из нескольких профилей рассеяния данных МУРР описывается матрицей $\mathbf{A} = \{A_{ik}\} = \{I^{(k)}(s_i)\}, (i = 1,...,N, k = 1,...,K, где N$ число экспериментальных точек, K - число кривых МУРР, например, общее число временных кадров в наборе данных SEC-SAXS. С помощью сингулярного разложения (SVD) эта матрица может быть представлена как $A = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathsf{T}}$, где матрица S диагональная, а столбцы ортогональных матриц U и V - собственные векторы матриц AA^T и A^TA , соответственно. Матрица U дает набор левых сингулярных векторов, т.е. ортонормированных базисных кривых $U^{(k)}(s_i)$, который охватывает пространство столбцов матрицы \mathbf{A} , тогда как диагональ \mathbf{S} содержит их соответствующие сингулярные значения в порядке убывания (чем больше сингулярное значение, тем значительнее соответствующий U-вектор). Число значимых сингулярных векторов (неслучайно осциллирующих кривых со значимыми сингулярными значениями) дает число независимых кривых, необходимых ДЛЯ представления всего набора данных ИХ линейными комбинациями, т.е. число отдельных компонентов в смесях.

ДЛЯ

Алгоритм ЭФА использует SVD-разложение набора данных МУРР

нахождения начальных и конечных точек компонентов в процессе эволюции системы (так называемые "окна концентрации" компонентов). Каждый компонент не присутствует вне временного окна концентрации, поэтому его концентрация в этих временных точках равна нулю. Также предполагается, что компоненты выходят (элюируются) друг за другом, т.е. первый компонент, присутствующий в системе, исчезает первым, второй компонент исчезает следующим, и так далее. Основная идея ЭФА алгоритма заключается в том, чтобы проследить за рангом матрицы данных А в зависимости от количества учитываемых измерений. Оценка проводится путем выполнения SVD на матрице данных с последовательным увеличением ее размера. Для этого обычно проводят ЭФА в прямом и обратном направлениях, чтобы определить, когда компонент появляется/исчезает из системы (для наборов данных SEC-SAXS или IEC-SAXS это определяет точки, соответствующие появлению/исчезновению дополнительного компонента во временной шкале хроматографического профиля элюирования). ЭФА в прямом направлении состоит из разложений SVD, многократно проводимого на части матрицы A, а именно, на матрице A_m , где A_m определяется как первые m столбцов матрицы **A**. Для Θ **A** в обратном направлении используются последние *m* столбцов матрицы A (также с последовательно возрастающим числом m). По графикам зависимости значимых сингулярных значений полученных SVD разложений от номера кривой рассеяния в наборе данных МУРР (например, временном интервале данных SEC-SAXS) можно оценить "концентрационные окна" компонентов по временным моментам, когда соответствующие сингулярные значения начинают резко расти выше некоторой базовой линии или резко снижаться, приближаясь к базовой линии.

Следующим шагом алгоритма ЭФА является определение матрицы вращения, которая необходима для преобразования значимых сингулярных векторов в матрицу концентрации и профили рассеяния компонентов. Примем во внимание, что матрица данных имеет вид A=IC, где столбцы матрицы I представляют собой

профили рассеяния компонентов, а матрица концентрации С выражается как $C = (U^T I)^{-1} S V^T = R V^T$, где R - матрица вращения. Последняя матрица неизвестна, но может быть найдена с помощью информации о концентрационных окнах компонентов, полученной на предыдущем этапе из эволюционирующих графиков сингулярных значений. Взяв только ту часть сингулярных векторов \mathbf{V}^{T} и матрицы **C**, которые находятся вне диапазона концентрационного окна, онжом последовательно восстановить один столбец матрицы **R** за другим [199]. Найдя элементы матрицы \mathbf{R} , можно легко вычислить столбцы матрицы \mathbf{C} . Используя этот подход, матрица концентрации С оценивается неитеративным способом и без каких-либо допущений о ее виде. Последним шагом алгоритма ЭФА является расчет матрицы І, который может быть выполнен с помощью псевдоинверсии Мура-Пенроуза матрицы С, умноженной на матрицу данных А.

6.1.2~ Программная~ реализация~ алгоритма~ ЭФА~ и~ схема~ автоматизированного поиска

Алгоритм ЭФА был нами реализован в виде программы EFAMIX, которая входит в пакет ATSAS (http://www.embl-hamburg.de/biosaxs/software.html) [A73]. В качестве входных данных программа EFAMIX требует наборы данных МУРР (SEC-SAXS) в текстовом формате ASCII с перечислением файлов, соответствующих измеренным экспериментальным кадрам, в порядке возрастания, а также полный или относительный путь к каталогу файлов. Пользователю необходимо указать следующие параметры:

предполагаемое количество компонентов в системе (из практических соображений допускается использование от двух до четырех компонентов)

подмножество (выборка) экспериментальных кадров данных, которые будут обрабатываться по алгоритму ЭФА (т.е. номера начального и конечного кадров выборки, содержащих рассеяние от образца)

подмножество экспериментальных кадров сигнала от буфера для расчета среднего сигнала буфера (если не задано явным образом, то выбирается область набора данных до первого кадра выборки образца);

угловой диапазон по оси *s* (по умолчанию EFAMIX включает все экспериментальные точки данных, но на практике удобно отбрасывать нестабильное/паразитное рассеяние на очень малых углах вблизи ограничителя пучка, которое может быть искажено небольшим, но существенным дрейфом системы регистрации (что может дать относительно большую ошибку при вычитании буферного рассеяния), а также шумные точки данных на больших углах, где кривые выходят на фон атомного рассеяния (максимальное значение *s* полезного углового диапазона может быть оценено, например, с помощью программы SHANUM, см. Главу 2).

Количество компонентов для алгоритма ЭФА можно определить по форме профиля элюирования или независимо с помощью модуля SVDPLOT в программе PRIMUS [A5] и/или POLYSAS [A9], используя непараметрический статистический тест Вальда-Вольфовица [127] на случайное изменение значимых сингулярных векторов В качестве результата программа EFAMIX выдает восстановленные кривые рассеяния отдельных компонентов, профили их концентрации и приближения к экспериментальным данным для каждого временного интервала. EFAMIX написан на языке Фортран и использует подпрограммы из Фортранбиблиотеки LAPACK (открытого доступа) для проведения матричного SVD разложения (www.netlib.org/lapack/). В этой библиотеке особый интерес представляют два алгоритма. Первый, DGESVD, близок к классической процедуре Голуба и Рейнша [120] (входная матрица А преобразуется в верхнюю бидиагональную матрицу последовательностью преобразований Хаусхолдера с последующей диагонализацией по алгоритму QR [207]). Вторая процедура, DGESDD, использует алгоритм "разделяй и властвуй" (адаптивное блочное разбиение матрицы). Этот алгоритм в 2-4 раза быстрее первого при аналогичной

вычислительной точности и именно он был выбран нами в программе EFAMIX.

Для решения избыточной системы линейных уравнений (по которой восстанавливаются элементы матрицы вращения **R**) в программе EFAMIX используется линейный метод наименьших квадратов, работающий на основе разложения по сингулярным значениям [207]. Такой метод численно устойчив, поскольку решение получается исключительно путем ортогональных преобразований. Кроме того, сингулярное разложение позволяет стабилизировать решение путем неявного вычисления псевдоинверсной матрицы Мура-Пенроуза, сингулярных значений (диагональ ограничивая спектр матрицы отсортированных в порядке убывания, отношением $\frac{S_{min}}{S_{max}} \ge \sqrt{K \cdot N} \cdot EPS$, где EPS – константа машинной точности вычислений равная 1.12*10⁻¹⁶.

В соответствии с общей концепцией подхода ЭФА с использованием метода матрицы вращения, подробная схема процедуры поиска для разложения перекрывающихся пиков данных МУРР (SEC-SAXS), полученных в сочетании с гель-хроматографией, может быть записана следующим образом:

назначить количество компонентов в системе;

назначить исходную матрицу \mathbf{A} , представляющую набор данных SEC-SAXS, взвешенную на оценки ошибок интенсивности, усредненных по временным кадрам (это опция по умолчанию, но пользователь может отключить взвешивание ошибок матрицы \mathbf{A} , если не уверен в качестве оценок ошибок);

выполнить расчеты ЭФА в прямом и обратном направлениях и сохранить соответствующие собственные значения ЭФА (EV - EigenValues) в зависимости от временных интервалов для каждого компонента;

определить "окна концентрации" компонентов по зависимостям собственных значений ЭФА в интерактивном или автоматическом режиме. В автоматическом режиме точки "перегиба" собственных значений ЭФА определяются временными моментами, когда первые производные собственных значений ЭФА достигают 10% от своих максимальных значений (таким образом удается отследить значительное увеличение ранга матрицы **A**);

найти столбцы матрицы вращения ${\bf R}$ путем решения набора линейных уравнений, составленных из элементов сингулярных векторов ${\bf V}^{\rm T}$ исходной матрицы данных ${\bf A}$, где размерность матрицы M выбирается в соответствии с "окнами концентрации" компонентов;

вычислить матрицу концентрации ${\bf C}$ как произведение матрицы вращения ${\bf R}$ и матрицы собственных векторов ${\bf V}^T$: ${\bf C} = {\bf R} {\bf V}^T$;

получить псевдоинверсию Мура-Пенроуза матрицы \mathbf{C} , используя ее SVD-разложение: $\mathbf{C}^{\text{-1}} = \mathbf{V}_{\mathbf{C}} \mathbf{S}_{\mathbf{C}}^{\text{-1}} \mathbf{U}_{\mathbf{C}}^{\mathbf{T}}$;

восстановить матрицу \mathbf{I} с профилями рассеяния индивидуальных компонентов из произведения псевдоинверсии Мура-Пенроуза матрицы \mathbf{C} и матрицы данных \mathbf{A} : \mathbf{I} = $\mathbf{A}\mathbf{C}^{-1}$.

Необходимо также отметить следующие важные детали реализованного алгоритма:

- i) анализ ЭФА проводится только для наборов данных с вычитанием буфера, пользователю необходимо предоставить информацию об используемом подмножестве (выборке) кадров от буферного раствора;
- ii) в случае использования взвешивания исходной матрицы данных **A** на оценки ошибок интенсивностей, матрица **I** на последнем шаге схемы поиска также корректируется путем умножения на веса ошибок;
- ііі) описанный выше метод матрицы вращения не имеет ограничений на получаемые значения матрицы концентраций С и в некоторых случаях может отрицательным значениям концентраций. В таких случаях необходимо перепроверить выбор "окон концентрации" пользователю компонентов и уточнить их в интерактивном режиме. Однако альтернативные методы вращения (такие как итеративные или гибридные методы, реализованные программе BioXTAS RAW), которые накладывают ограничения положительность матрицы С в процессе поиска, не всегда сходятся к решению и имеют риск застрять в локальных минимумах, далеких от истинного решения;

iv) критерий, используемый для автоматизированной оценки "окон концентрации" компонентов, позволяет проводить анализ с минимальным вмешательством пользователя, однако не исключает получение неоптимальных результатов. Пользователи должны всегда проверять соответствующие графики ЭФА и уточнять выбор программы в автоматизированном режиме работы, если наблюдается несоответствие данным.

6.2 Возможности и ограничения метода ЭФА на примерах теоретических данных МУРР

Сначала метод был протестирован на теоретических наборах данных SEC-SAXS от белковых смесей. В примерах, представленных на Рис. 6.1-6.7, сгенерированы наборы данных SEC-SAXS от двухкомпонентных смесей мономердимер бычьего сывороточного альбумина (BSA: Bovine Serum Albumin) с различной степенью перекрытия хроматографических пиков (имитирующих различные типы хроматографических колонок (Рис. 6.2), различные соотношения концентраций компонентов в смеси (Рис. 6.4-6.7) и различные уровни сигнал/шум в данных (Рис. 6.1). Теоретические кривые рассеяния от мономерных и димерных моделей BSA (PDB ID: 4F5S) были рассчитаны с помощью программы CRYSOL [A10], а затем набор из 100 кривых рассеяния был рассчитан в виде линейных комбинаций теоретических кривых, взвешенных по объемным долям согласно профилям концентраций компонентов. Последние были представлены гауссовыми (симметричные пики) (Рис. 6.1-6.2 и Рис. 6.4-6.7) и экспоненциально-гауссовыми гибридными (EGH) функциями (асимметричные пики) (Рис. 6.3) [208], сдвинутыми друг относительно друга. Как и в реальном случае, при элюировании компонентов из хроматографической колонки, сначала на выходе появлялись димеры, а затем мономеры (Рис. 1-3).

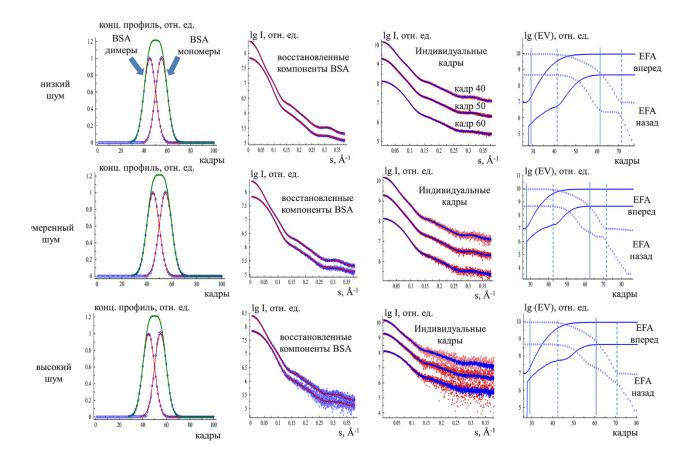


Рис. 6.1 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (оба компонента имеют равные доли). Концентрационные профили моделируются гауссовыми функциями. Слева направо, колонка 1 - концентрационные профили компонентов, восстановленные программой EFAMIX (синий), теоретические (идеальные) профили компонентов (красный), общий теоретический концентрационный профиль (зеленый); колонка 2 профили рассеяния компонентов, восстановленные программой EFAMIX (синий) и теоретические (идеальные) профили рассеяния, рассчитанные от известных кристаллических моделей BSA программой CRYSOL (красный); колонка 3 - отдельные временные кадры данных SEC-SAXS (кадры с номерами 40, 50 и 60) (красный) и приближения, полученные программой EFAMIX (синий); колонка 4 - графики прямого ЭФА (сплошные линии) и обратного ЭФА (кружки) для первых двух значимых сингулярных значений, появление и исчезновение первой и второй компонент показано сплошными и пунктирными вертикальными линиями, соответственно. Уровень шума пуассоновского типа, добавленный к данным, соответствует следующему количеству фотонов вблизи первичного пучка (с учетом радиального усреднения): сверху вниз, ряд 1 (низкий шум) - 10⁴ фотона, ряд 2 (умеренный шум) - 10^3 фотона, ряд 3 (высокий шум) - 10^2 фотона.

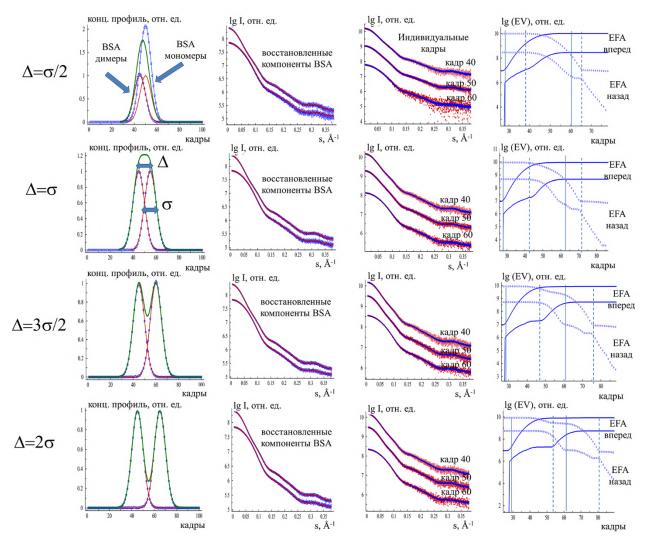


Рис. 6.2 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (оба компонента имеют равные доли) с различной степенью перекрытия пиков выхода (элюирования) компонентов. Концентрационные профили моделируются гауссовыми функциями. Обозначения и цветовые схемы - как на Рис. 6.1.

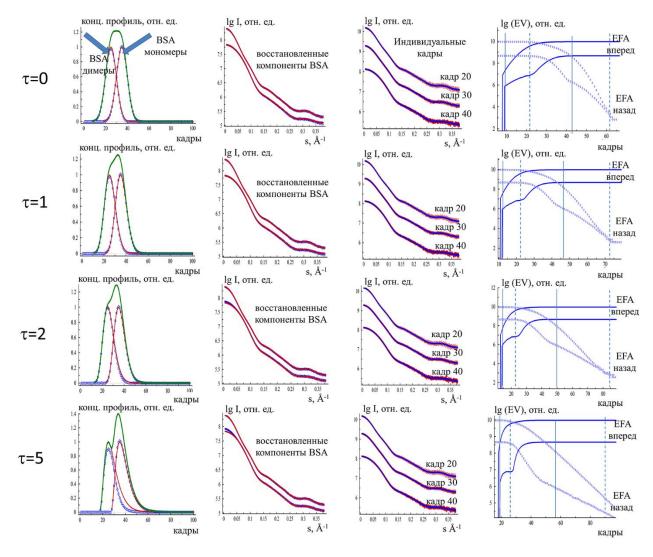


Рис. 6.3 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (оба компонента имеют равные доли) с различной асимметрией профиля концентрации компонентов. Концентрационные профили моделируются асимметричными функциями EGH [208]. Обозначения и цветовые схемы - как на Рис. 6.1.

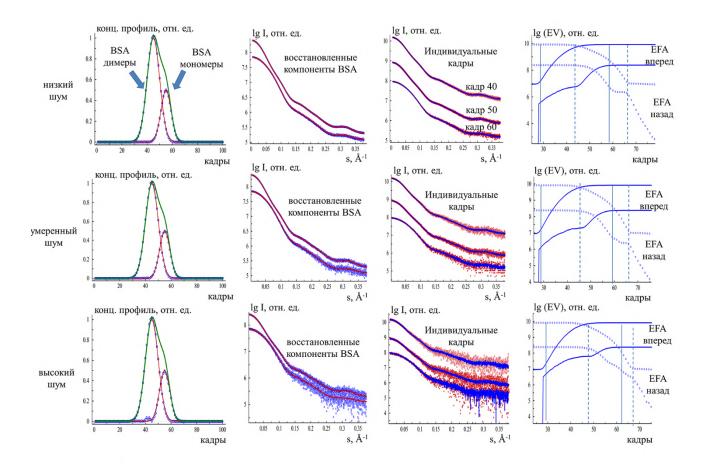


Рис. 6.4 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (основная фракция - димерная, соотношение концентраций 2:1). Обозначения и цветовые схемы - как на Рис. 6.1.

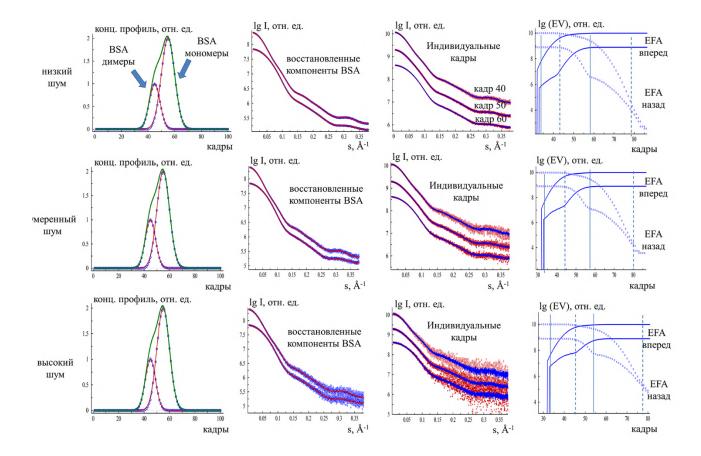


Рис. 6.5 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (основная фракция - димерная, соотношение концентраций 1:2). Обозначения и цветовые схемы - как на Рис. 6.1.

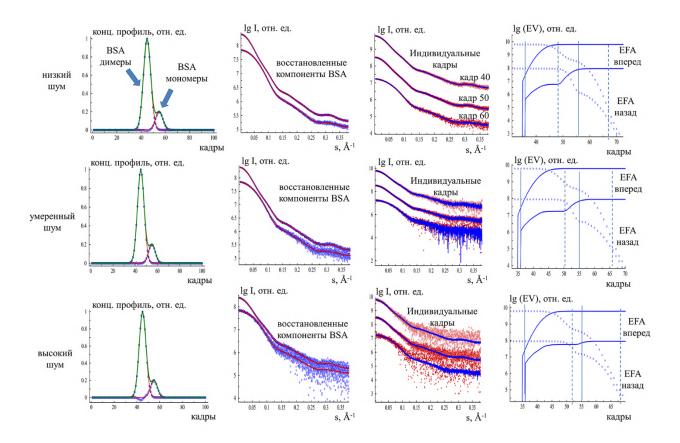


Рис. 6.6 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (основная фракция - димерная, соотношение концентраций 5:1). Обозначения и цветовые схемы - как на Рис. 6.1.

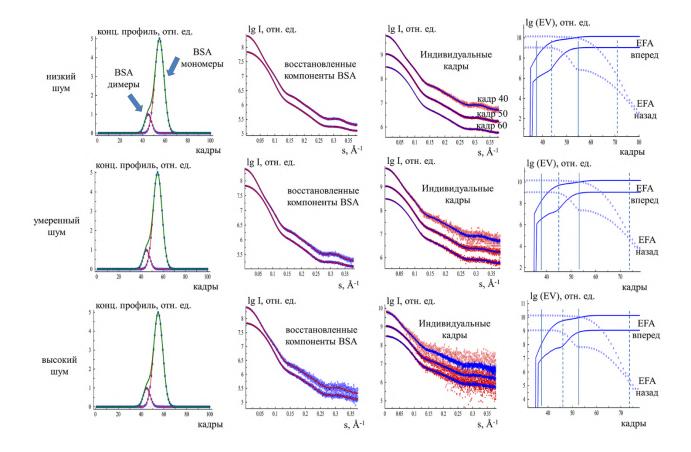


Рис. 6.7 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (основная фракция - димерная, соотношение концентраций 1:5). Обозначения и цветовые схемы - как на Рис. 6.1.

Поскольку любые экспериментальные данные содержат шум, то было необходимо соответствующим образом добавить его к теоретическим наборам данных МУРР. Современные 2D пиксельные детекторы (например, Pilatus, Eiger и др.) считаются счетчиками одиночных фотонов, и регистрируемые сигналы подчиняются статистике Пуассона. При этом во время процедуры моделирования шума проводить вычитание буфера не требуется, так как сумма двух случайных величин с пуассоновским распределением также будет иметь пуассоновское распределение. В то же время необходимо учесть распространение ошибок при азимутальном

усреднении двумерных данных. Нами был использован алгоритм генерации пуассоновского псевдослучайного шума, описанный в [209, 210]. Расчет распространения ошибок при азимутальном усреднении производился для двумерного детектора с числом пикселей 2500х2500. Для получения одномерных кривых рассеяния с различным отношением сигнал/шум мы масштабировали кривые (которые представляют собой линейные исходные комбинации теоретических кривых рассеяния от компонентов, взвешенных согласно их объемным долям) на коэффициент, соответствующий стандартным отклонениям пуассоновского шума в 1, 3 и 10 процентов в области максимальной интенсивности (начальная часть одномерной кривой). Рассматривая такие математические ожидания интенсивности, для каждого пикселя детектора были рассчитаны псевдослучайные реализации пуассоновского шума. Наконец, интенсивности были усреднены по плоскости детектора в азимутальном направлении. Полученные уровни шума можно классифицировать как низкий шум (соответствует 10^4 фотонам в области, близкой к области прямого пучка на детекторе); умеренный шум (10^3 фотона) и высокий шум (10^2 фотона).

Сгенерированные наборы данных SEC-SAXS были проанализированы с помощью программы EFAMIX. Сигналы рассеяния мономеров и димеров BSA были восстановлены вместе с их концентрационными профилями и сравнены с рассчитанными кривыми рассеяния исходных мономерных и димерных рентгеновских кристаллических структур BSA. Как видно из Рис. 6.1-6.2 и Рис. 6.4-6.7, в случае гауссовых (симметричных) профилей концентраций компонентов, программа EFAMIX успешно разложила перекрытые профили и восстановила информацию об отдельных компонентах с точностью, которая зависит как от степени перекрытия пиков, так и соотношения положений пиков и уровня шума в данных. При низком уровне шума (с числами фотонов 10³ и 10⁴) восстановленные профили концентрации и кривые рассеяния компонентов полностью совпали с теоретическими (идеальными) значениями. Однако, при более высоком уровне шума (количество фотонов 10²) восстановленные профили концентрации уже

содержали некоторые артефакты (несоответствия), а кривые рассеяния компонентов становились более шумными в области больших углов. Стоит отметить, что общее качество декомпозиции данных оставалось удовлетворительным, в том числе и при высоком уровне шума. Проведенное моделирование продемонстрировало высокую эффективность алгоритма ЭФА для анализа данных SEC-SAXS от двухкомпонентных систем с симметричными профилями концентрации даже при значительном перекрытии пиков и высоком уровне шума в данных МУРР.

На практике пики элюирования компонентов смеси могут быть несколько асимметричными с более резким подъемом на переднем крае и удлиненным хвостом на заднем крае. Такие профили можно смоделировать с помощью асимметричных функций EGH с дополнительным параметром "время релаксации" т [208]. Этот "хвост" может вызвать систематические отклонения при разложении пиков с помощью алгоритма ЭФА даже при низком уровне шума для двухкомпонентных систем, как видно из Рис. 6.3. Для сильно перекрывающихся пиков разложение все еще возможно при умеренной асимметрии профилей концентрации (когда параметр релаксации функции EGH т не превышает значения 2). При более высокой асимметрии профилей ПИКИ компонентов восстанавливаются со значительно перекрывающимися концентрационными окнами, что приводит к систематическим отклонениям от идеальных значений (например, при параметре релаксации т=5, Рис. 6.3).

Форма белков также может влиять на результаты процесса разложения, но в достаточно умеренной степени. Чтобы убедиться в этом, программа EFAMIX была применена к анализу теоретического набора данных SEC-SAXS от смеси сильно вытянутых димеров и тетрамеров белка фибриногена (Рис. 6.8) и смогла достаточно хорошо восстановить индивидуальные компоненты даже при высоком уровне шума (с количеством фотонов 10²).

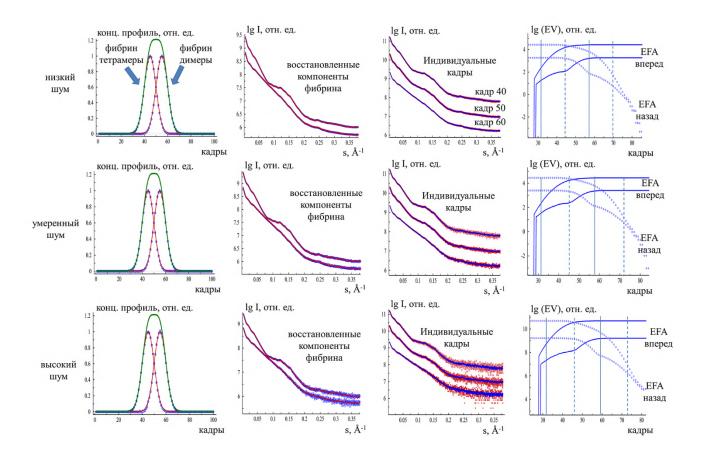


Рис. 6.8 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси димер-тетрамер белка фибриногена (PDB ID: 3GHG) (случай сильно вытянутых частиц). Обозначения и цветовые схемы - как на Рис. 6.1.

Можно также оценить, как будет зависеть вероятность успешного разложения по методу ЭФА в зависимости от соотношений концентраций компонентов для двухкомпонентных смесей с одним преобладающим и одним минорным компонентом. На Рис. 6.6-6.7 представлены результаты программы EFAMIX для мономер-димер BSA с соотношением концентраций 1:5 соответственно. Как видно из Рис. 6.6-6.7, восстановить димеры можно, даже если незначительную собой фракцию представляют смеси, проблематично восстановить мономерный компонент, когда рассеяние от димеров доминирует над сигналом от мономеров. Компоненты еще можно разрешить, если соотношение их концентраций составляет примерно 1:5/1:7, в то время как соотношение, превышающее 1:10, по-видимому, является пределом для успешного ЭФА. Расстояние между разложения ПО методу пиками также играет

определенную роль. Расстояние между перекрывающимися пиками концентрационных профилей компонентов во всех тестовых случаях было как минимум в два раза больше ширины индивидуального пика; при меньшем расстоянии разложение пиков становится неоднозначным, как видно из Рис. 6.2, где степень перекрытия пиков Δ варьировалась в широких диапазонах.

Затем нами были рассмотрены более сложные случаи - данные SEC-SAXS от трехи четырехкомпонентных смесей BSA (равновесных смесей мономер-димертетрамер и мономер-димер-тетрамер-октамер, соответственно) (Рис. 6.9-6.10). Пики элюирования появлялись и исчезали последовательно, начиная с более меньшими олигомерами высокомолекулярных видов И заканчивая ИЛИ мономерами. Программа EFAMIX смогла успешно разделить пики трехкомпонентной системы при низком, умеренном и высоком уровнях шума (с числами фотонов 10^4 , 10^3 и 10^2) (Рис. 6.9). Для четырехкомпонентной системы восстановление было возможно только при относительно низком уровне шума (с числом фотонов 10^4 и 10^3), тогда как при более высоком уровне шума (число фотонов 10^2) удалось получить правильную информацию только о самых крупных компонентах в смеси (тетрамерах и октамерах) (Рис. 6.10). Таким образом, метод ЭФА может быть полезен для анализа данных SEC-SAXS и в случае более сложных трех- и четырехкомпонентных систем, но порог шума, после которого реконструкция становится ненадежной, снижается с ростом числа компонентов в системе.

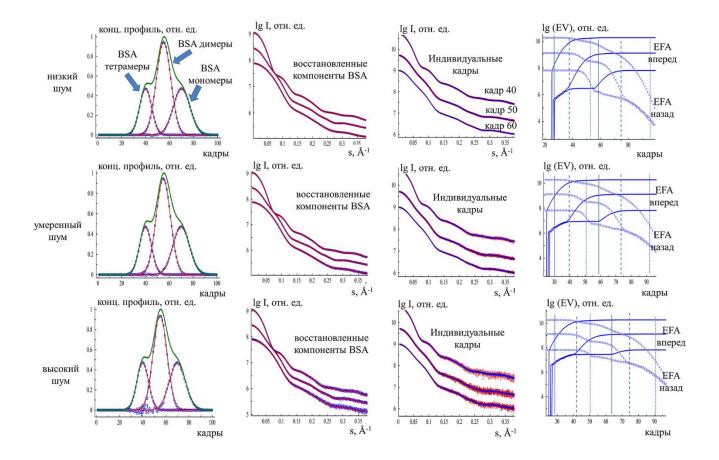


Рис. 6.9 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономер-димер-тетрамер белка BSA (трехкомпонентная система). Остальные обозначения и цветовые схемы - как на Рис. 6.1.

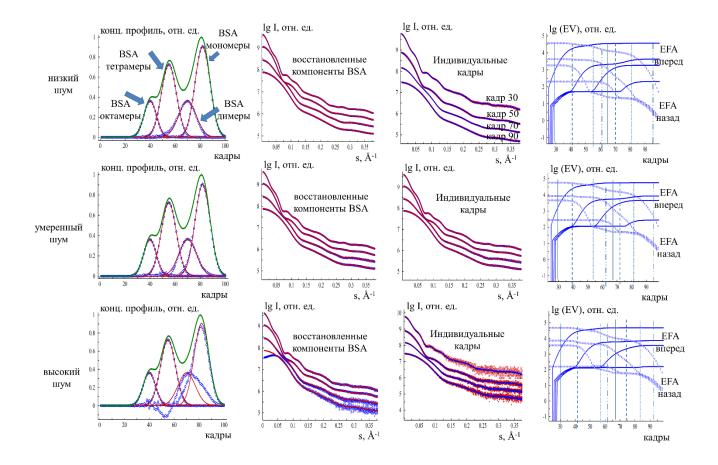


Рис. 6.10 Разделение компонентов программой EFAMIX на синтетических данных SEC-SAXS от смеси мономер-димер-тетрамер-октамер белка BSA (четырехкомпонентная система). Колонка 3- отдельные кадры данных SEC-SAXS (временные кадры № 30, 50, 70 и 90). Остальные обозначения и цветовые схемы - как на Рис. 6.1.

дополнительной независимой проверки эффективности работы алгоритма также был смоделирован набор данных SEC-SAXS с помощью программ IMSIM и IM2DAT [211] из пакета ATSAS, еще более приближенный к условиям реального эксперимента. В этом случае двумерные (2D) картины мономеров димеров **BSA** (соответствующие рассеяния смеси И хроматографическому профилю с перекрывающимися пиками, как на Рис. 6.1) были смоделированы программой IMSIM для экспериментальной установки с детектором Pilatus 6M, расположенном на расстоянии 3.0 м от образца, с потоком интенсивности падающего пучка 10¹² фотонов/сек, временем экспозиции 1 секунда и концентрацией белка 1 мг/мл (в максимуме пика элюирования). После этого

двумерные изображения преобразовывались в одномерные кривые рассеяния путем радиального усреднения с помощью программы IM2DAT, а буферный сигнал вычитался из каждого временного кадра. После этого, мы использовали программу EFAMIX для анализа сгенерированного набора данных SEC-SAXS и успешно восстановили кривые рассеяния индивидуальных компонентов и соответствующие профили концентрации (Рис. 6.11). Эти результаты еще раз доказывают высокую эффективность алгоритма ЭФА при анализе наборов данных SEC-SAXS, имитирующих реальные экспериментальные условия.

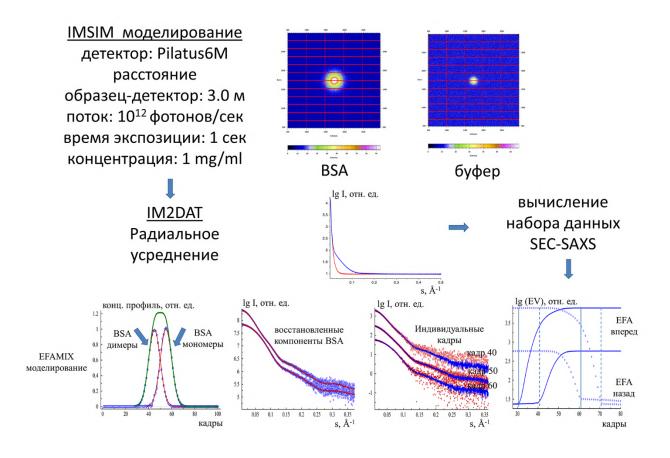


Рис. 6.11 Разделение компонентов программой EFAMIX на "квазиэкспериментальных" данных SEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (оба компонента имеют равные доли), полученных с помощью программ IMSIM и IM2DAT [211] из пакета ATSAS.

6.3 Применение метода ЭФА для экспериментальных наборов данных МУРР с использованием гель-фильтрации и ионно-обменной хроматографии

После проверки на теоретических наборах данных SEC-SAXS метод ЭФА был протестирован на ряде экспериментальных наборов данных SEC-SAXS, собранных от образцов, содержащих частицы разных размеров с различными концентрациями. Ниже будут представлены полученные результаты для данных МУРР, которые были записаны на синхротронной станции P12 (EMBL) [23, 212] накопительного кольца PETRA III (DESY, Гамбург) в сочетании со встроенной гель-хроматографической колонкой.

Первый набор данных SEC-SAXS, полученный от чистого мономерного образца ВSA, дал единственный элюирующий пик (Рис. 6.12а). Восстановленные программой EFAMIX профили рассеяния содержали только один значимый компонент, при этом кристаллическая структура BSA (PDB ID: 4F5S) четко соответствовала восстановленной кривой рассеяния от этого компонента. Второй набор данных SEC-SAXS, полученный от стандартного белка, часто используемого для калибровки молекулярной массы, глюкозо-изомеразы, также дает профиль элюирования с симметричным одиночным пиком (Рис. 6.12б). Разложение данных программой EFAMIX и в этом случае дало только один значимый компонент, а кривая, рассчитанная по кристаллографической модели глюкозо-изомеразы (РDB ID: 1OAD), соответствовала восстановленному сигналу рассеяния. Эти результаты показывают, что для систем с одним компонентом программа EFAMIX позволяет надежно отличить полезный сигнал рассеяния от шумового компонента.

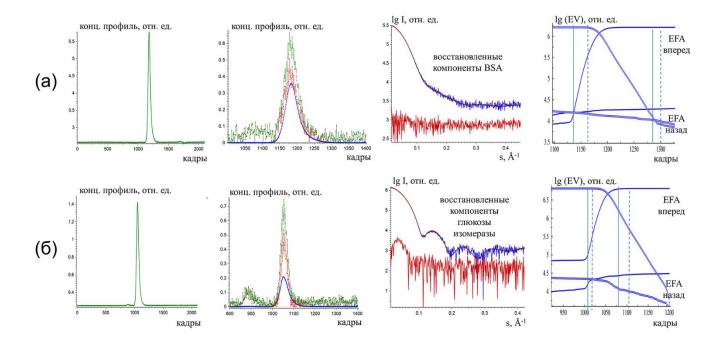


Рис. 6.12 Разделение компонентов программой EFAMIX на экспериментальных SEC-SAXS монодисперсных растворов белка **BSA** данных OT (a) глюкозоизомеразы (б); расчеты проводились в рамках двухкомпонентного приближения. Колонка 1 - профили элюирования данных SEC-SAXS, полученные с помощью программы CHROMIXS (зеленый). Колонка 2 - восстановленные программой EFAMIX профили концентрации компонентов, синяя и красная кривые - отдельные компоненты, зеленая кривая - общий профиль концентрации. восстановленные программой EFAMIX профили рассеяния компонентов (синяя и красная кривые, соответственно), кривые рассеяния, рассчитанные от кристаллографических моделей (коричневый) (BSA: 4F5S.pdb, Glucose Isomerase: 1OAD.pdb) совмещенные с восстановленным профилем рассеяния от наиболее значимого компонента (синий). Колонка 4 - графики прямого ЭФА (сплошные линии) и обратного ЭФА (кружки), для которых обозначения и цветовые схемы те же, что и на Рис. 6.1.

Третий набор данных SEC-SAXS, полученный от белка пируват альдолазы класса II [213], дал асимметричный хроматографический профиль выхода, указывающий на потенциальное присутствие двух значимых компонентов (Рис. 6.13а). Действительно, SVD-анализ указал на наличие двух значимых компонентов, а разложение EFAMIX восстановило два разных компонента, где кривая компонента меньшего размера хорошо воспроизводилась кристаллической моделью гексамера (РDB ID: 6R62). Более крупный компонент, вероятно, соответствует октамеру, поскольку отношение объемов частиц по инварианту Порода [33], рассчитанное

для двух восстановленных кривых рассеяния, составляет около 1.3. Кроме того, нами была дополнительно проверена стабильность работы алгоритма ЭФА путем учета только нечетных или четных экспериментальных кадров набора данных SEC-SAXS. Оказалось, что в обоих случаях восстановленные решения существенно не отличались от результатов, полученных с использованием полного набора данных SEC-SAXS.

Четвертый набор SEC-SAXS был данных получен otпредварительно подготовленной смеси двух белков, овальбумина и В-амилазы. Профиль элюирования этого раствора показал два частично перекрывающихся пика (Рис. 6.136). Хотя после первого пика наблюдалось небольшое плечо, SVD-анализ данных выявил только два значимых компонента в системе. Программа EFAMIX смогла разложить профиль и приблизить весь набор данных SEC-SAXS в виде линейной комбинации двух компонентов. Восстановленные кривые рассеяния хорошо согласуются с теоретическими кривыми, рассчитанными на основе кристаллических структур двух белков - мономерного овальбумина PDB ID: 10VA (MW=42 кДа) и тетрамерной β-амилазы PDB ID: 1FA2 (MW=223 кДа). Таким ЭФА стабильные образом, метод дает надежные И решения ДЛЯ **SEC-SAXS** наборов экспериментальных данных двумя различными компонентами даже в случае частично перекрывающихся профилей элюирования.

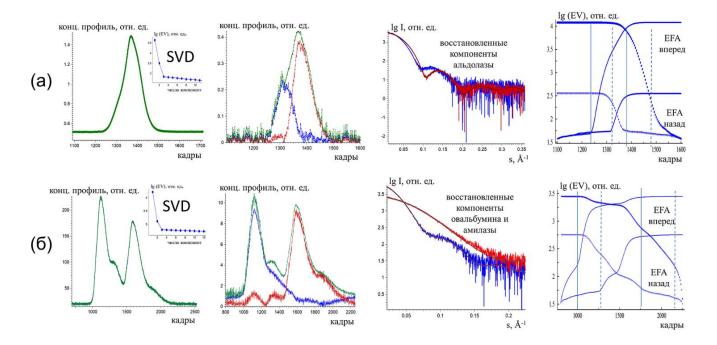


Рис. 6.13 Разделение компонентов программой EFAMIX на экспериментальных данных SEC-SAXS от белка альдолазы (а) и смеси белков овальбумина с бета-амилазой (б); расчеты проводились в рамках двухкомпонентного приближения. Колонка 1 - профили элюирования данных SEC-SAXS, полученных с помощью программы CHROMIXS (зеленый). Во вставках приведены сингулярные значения SVD-разложения данных SEC-SAXS (после вычитания буфера) в порядке убывания. Колонка 2 - восстановленные программой EFAMIX профили концентрации компонентов, синяя и красная кривые - отдельные компоненты, зеленая кривая - общий профиль концентрации. Колонка 3 - восстановленные программой EFAMIX профили рассеяния компонентов и кривые рассеяния, рассчитанные от кристаллографических моделей (коричневые кривые) (гексамер альдолазы: 6R62.pdb, мономер овальбумина: 1OVA.pdb; тетрамер -амилазы: 1FA2.pdb). Колонка 4 - графики прямого ЭФА (сплошные линии) и обратного ЭФА (кружки), для которых обозначения и цветовые схемы те же, что и на Рис. 6.1.

Рассмотрим теперь экспериментальные наборы данных МУРР в сочетании с ионно-обменной хроматографиией (IEC-SAXS). В процессе IEC образец из хроматографической колонки вымывается потоком буфера с возрастающей концентрацией соли. Основной проблемой при анализе данных IEC-SAXS является учет фона от буфера, меняющегося из-за градиента соли. Формально, меняющийся фон может нарушить неявные предположения метода ЭФА (наличие

непересекающихся областей в профилях концентраций компонентов), но на практике степень этого нарушения зависит от степени изменения содержания соли. Сначала нами были смоделированы данные IEC-SAXS от смеси мономер-димер BSA (такой же, как на Рис. 6.1), где мы учли градиент буфера в виде возрастающего постоянного слагаемого для каждого временного кадра. Мы выбрали случай с относительно высоким уровнем шума (с числом фотонов 10^2) и протестировали два буферных градиента с уровнями 12% и 25% (относительная разница в сигнале буфера до и после пика элюирования образца). На практике градиент буфера с 12% уровнем перепада соответствовал бы добавлению 1.2 M NaCl. Как видно из Рис. 6.14, программа EFAMIX позволяет успешно разложить данные IEC-SAXS при наличии 12% градиента буфера, но начинает испытывать трудности при 25% градиенте буфера. В последнем случае только димерный компонент может быть надежно восстановлен, тогда как кривая рассеяния от мономерной фракции имеет систематические отклонения от теоретической кривой в области высоких углов. Но в тоже время, полученные результаты показывают, что метод ЭФА хорошо применим для наборов данных IEC-SAXS с умеренной степенью градиента буфера.

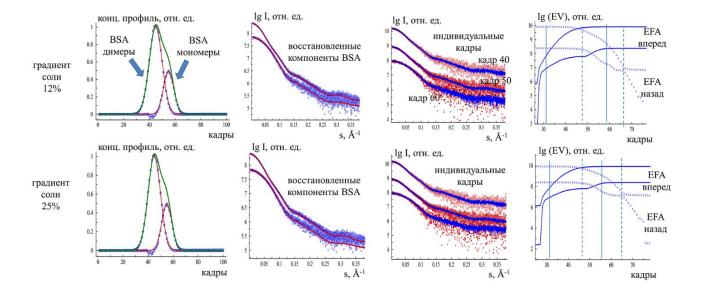


Рис. 6.14 Разделение компонентов программой EFAMIX на синтетических данных IEC-SAXS от смеси мономеров и димеров белка бычьего альбумина BSA (основная фракция - димерная, соотношение концентраций 2:1) с постоянным солевым

градиентом буфера в 12% (а) и 25% (б). Остальные обозначения и цветовые схемы - как на Рис. 6.1.

После этого мы попробовали применить метод ЭФА к экспериментальному набору данных IEC-SAXS, полученному для моноклонального антитела IgG1 после разбавления его папаином. В этом случае IgG1 разделяется на Fc-домен, а также два идентичных Fab-домена. Все эти домены имеют молекулярную массу около 50 кДа быть разделены помощью поэтому МОГУТ cстандартной хроматографической методики (SEC). Однако за счет наличия разного заряда на поверхности доменов ионно-обменная хроматография (IEC) позволяет провести их разделение, при котором Fc-домен элюируется раньше Fab-домена. Для этого катионообменная колонка, разработанная специально использовалась высокоразрешающего и высокоэффективного анализа моноклональных антител (Рис. 6.15). Относительное изменение буферного фона по всему набору IEC-SAXS (оцениваемое как отношение разницы между сигналом образца в максимуме элюирующего пика и фоном буфера после элюирующего пика к разнице между сигналом образца и фоном буфера перед элюирующим пиком) довольно высока (около 35%). Поэтому мы провели анализ методом ЭФА только для отдельно выделенных временных участков, соответствующих выходу Гс домена (первый пик элюции) и Fab домена (второй пик элюции), результаты представлены на Рис. 6.15. Видно, что в каждом случае программа EFAMIX выявила только один значимый компонент, присутствующий в системе (второй компонент имел незначительный сигнал интенсивности). Восстановленные профили рассеяния от этих компонентов можно сравнить с кривыми, полученными путем стандартного вычитания 'образец-буфер' в программе CHROMIXS [214] с использованием буферных сигналов до и после хроматографического пика выхода образца. В обоих случаях кривые хорошо согласуются друг с другом в области малых углов, но немного отличаются в области больших углов. Интересно, что кривые рассеяния, рассчитанные от кристаллографических структур Fc и Fab доменов IgG1, лучше соответствуют кривым, полученным программой EFAMIX, а не программой

СНКОМІХЅ. Таким образом, артефакты изменяющегося фона буфера влияют незначительно на эффективность разложения в программе EFAMIX, в то время как 'ручная' стандартная обработка данных программой CHROMIXЅ дает более смещенный сигнал. Этот результат показывает, что программу EFAMIX можно использовать и на данных IEC-SAXЅ с умеренно изменяющимся уровнем фона.

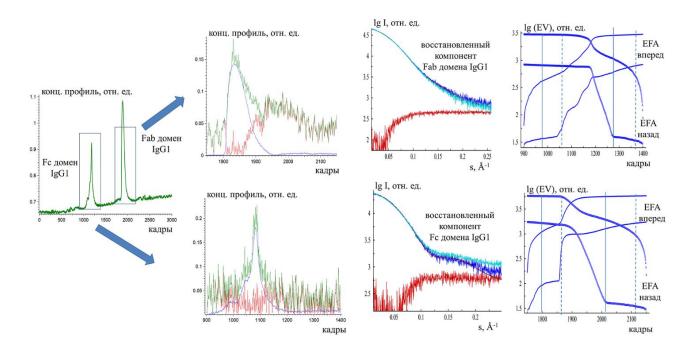


Рис. 6.15 Разделение компонентов программой EFAMIX на экспериментальных данных IEC-SAXS Fc и Fab доменов белка IgG1. Слева направо, колонка 1 профиль элюирования данных IEC-SAXS (зеленый), полученный CHROMIXS (первый пик элюирования соответствует Fc домену IgG1, второй пик элюирования принадлежит Fab домену IgG1). Колонка 2 - восстановленные программой EFAMIX профили концентрации компонентов, синяя и красная кривые - отдельные компоненты, зеленая кривая - общий профиль концентрации. Колонка 3 профили (синий восстановленные рассеяния компонентов красный, соответственно), кривые рассеяния, рассчитанные от кристаллографических моделей Fc и Fab доменов IgG1 (1hzh.pdb) (коричневые кривые) и кривая рассеяния, полученная программой CHROMIXS (голубая) с использованием сигналов буфера до и после пика элюции. Колонка 4 - графики прямого ЭФА (сплошные линии) и обратного ЭФА (кружки), для которых обозначения и цветовые схемы такие же, как на Рис. 6.1.

Эффективность работы алгоритма ЭФА при разделении вкладов компонентов в случае смесей, состоящих из частиц с одинаковой молекулярной массой, была

также подтверждена в работе [A74] на примере двух конформационных состояний (открытой и закрытой форм) фермента 3-изопропилмалатдегидрогеназы.

6.4 Автоматический и интерактивный режимы работы метода ЭФА

ЭФА - это общий хемометрический метод анализа множества наборов данных, описываемых систематическими эволюционирующими И отдельными компонентами. Метод не требует никаких априорных предположений о количестве компонентов, их форме или степени разделения между компонентами. Подход ЭФА был ранее успешно применен в аналитической химии и химии растворов, в частности, ДЛЯ высокоэффективной жидкостной хроматографии детектированием на фотодиодной матрице (HPLC-DAD) и ультрафиолетовой спектрометрии (UVS). Применение метода для анализа данных SEC-SAXS также показало его высокий потенциал и позволило не только разложить сигналы от олигомерных равновесных смесей белков [201, 215], но и охарактеризовать доменные движения фермента, участвующего в аллостерической активации [200]. Из литературы по хемометрическому разделению матриц смесей известно, что разделение является однозначным, если имеются непересекающиеся области в контурах спектров компонентов (что не всегда имеет место в эксперименте МУРР) или в контурах профилей концентрации (на этом основан принцип ЭФА). Хроматографическое разделение, как минимум, обеспечивает непересекающиеся начальные участки кривых концентрации, и этого должно быть достаточно для успешного разложения. То же самое справедливо и для хвостовых участков концентрационных кривых, но даже в случае их перекрытия все равно используются полнодиапазонные профили, поскольку они содержат полезную информацию, улучшающую статистику набора данных.

Следует отметить, что определение концентрационных "окон присутствия" компонентов в программе EFAMIX проводится в автоматическом режиме, однако поскольку эффективность работы метода ЭФА имеет определенные ограничения, зависящие от степени перекрывания концентрационных профилей компонентов, а

также степени их асимметрии, то в ряде случаев автоматический режим работы программы может приводить к смещенным оценкам, что уменьшает вероятность получения правильного решения. Поэтому важно иметь возможность использования интерактивного моделирования для быстрого сканирования различных вариантов "окон присутствия" компонентов и проверки результатов восстановления профилей рассеяния компонентов.

интерактивной обработки набора МУРР, Для данных полученных cиспользованием хроматографической колонки и содержащих только один компонент, ранее была разработана программа CHROMIXS [214], входящая в пакет ATSAS. Анализ данных в ней проводится путем усреднения временных кадров от образца и буфера и последующего вычитания буферного сигнала из образца. Удобный графический интерфейс, использующий данных OT кроссплатформенную графическую библиотеку Qt, позволяет быстро проводить манипуляции с данными.

Для проведения интерактивного моделирования данных МУРР от белковых смесей, содержащих от двух до четырех компонентов, программа CHROMIXS была нами модернизирована [A75]. В нее было добавлено специальное меню вызова программы EFAMIX, в котором можно задавать количество компонентов, "временные" окна присутствия компонентов и угловой диапазон используемых данных МУРР. После завершения работы программы EFAMIX восстановленные концентрационные профили и кривые рассеяния компонентов отображаются с помощью программы PrimusQt [A7]. Графический интерфейс модернизированной программы CHROMIXS представлен на Рис. 6.16. На ней показаны результаты разложения олигомерной смеси белка пируват альдолазы (гексамер-октамер) с помощью алгоритма ЭФА, которые хорошо согласуются с известными кристаллографическими структурами [213].

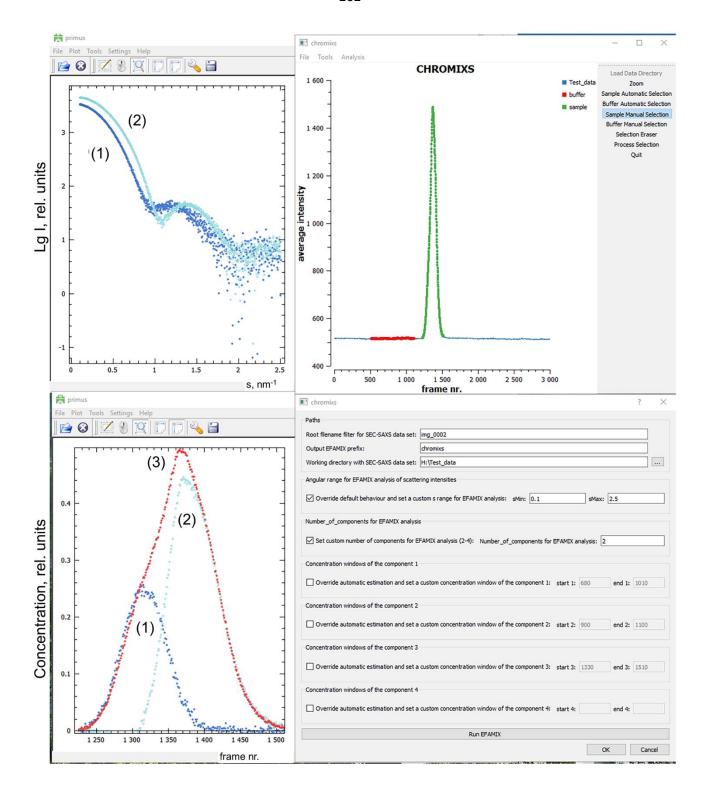


Рис. 6.16 Графический интерфейс вызова программы EFAMIX с помощью программы CHROMIXS, позволяющий проводить моделирование в интерактивном режиме. Данные МУРР от белка альдолаза, представляющего собой в растворе смесь октамеров (кривая 1) и гексамеров (кривая 2) [213]. Справа вверху представлен хроматографический ("временной") профиль выхода компонентов, область фонового ("буферного") сигнала соответствует временным

кадрам 500÷1000 (выделена красным цветом), область присутствия образца — временным кадрам 1200÷1500 (выделена зеленым цветом). Справа внизу находится меню вызова программы EFAMIX. Слева вверху показаны восстановленные программой EFAMIX профили рассеяния октамеров (синяя кривая 1) и гексамеров (голубая кривая 2) белка альдолазы. Слева внизу показаны найденные с помощью эволюционного факторного анализа концентрационные профили компонентов (синяя кривая 1 — октамеры, голубая кривая 2 - гексамеры) и их суммарный профиль (красная кривая 3).

Интерактивная обработка данных SEC-SAXS также может быть выполнена, например, с помощью таких программ, DATASW [216], DELA [217] и US-SOMO HPLC-SAXS [218]. Результаты могут зависеть от доступных угловых диапазонов в данных МУРР, от соотношения сигнал/шум и от степени перекрытия пиков, поэтому часто полезно проводить сравнение результатов, полученных с помощью различных подходов. Сравнение результатов, полученных при интерактивном и автоматическом разложении, также может обеспечить полезную перекрестную проверку, особенно в сложных случаях с сильно перекрывающимися пиками.

Заключение к Главе 6

В данной Главе нами был рассмотрен алгоритм эволюционного факторного анализа (ЭФА) для анализа данных МУРР, полученных в сочетании с гельхроматографической колонкой (SEC-SAXS). Алгоритм был реализован в виде программы EFAMIX. Проведено систематическое исследование чувствительности метода в зависимости от уровня шума в данных и количества компонентов в системах с перекрывающимися пиками выхода (элюирования) компонентов смеси. Показано, что ДЛЯ двухкомпонентных систем симметричными (гауссоподобными) профилями концентрации (например, равновесная смесь мономера и димера) метод ЭФА способен восстановить профили концентрации и кривые рассеяния отдельных компонентов даже при наличии значительного уровня шума в данных. При более высоких уровнях шума реконструкция ЭФА становится нестабильной, а для систем с большим числом компонентов этот шумовой порог неуклонно снижается. Интересно, что даже в случае высокого

уровня шума кривые рассеяния от более крупных компонентов все еще могут быть восстановлены, в то время как получаемые вклады от более мелких частиц уже имеют систематические отклонения от ожидаемых зависимостей. Метод ЭФА ограниченно применим к системам со значительно асимметричными профилями концентрации или в случае слишком сильного перекрывания хроматографических пиков выхода компонентов (когда расстояние между максимумами пиков меньше, чем удвоенная ширина отдельного пика). Такие случаи могут возникать, например, при неоптимальном давлении и скорости потока в хроматографической колонке (SEC) или из-за структурной неоднородности образца, когда специфические конформации исследуемого образца имеют тенденцию по-разному взаимодействовать с матрицей-фильтром в колонке SEC.

Метод ЭФА был протестирован на ряде экспериментальных наборах данных SEC-SAXS и IEC-SAXS и дал надежные восстановления кривых рассеяния отдельных компонентов смеси, совместимые с теоретическими кривыми, рассчитанными на основе известных кристаллографических структур. В частности, удалось описать необычную олигомерную смесь белка пируват альдолазы, состоящей из гексамеров и октамеров. Для данных IEC-SAXS показано, что несмотря на изменение фона из-за варьирования количества соли в буферном растворе, метод ЭФА эффективен в случае умеренных градиентов солевого буфера.

Глава 7

Разработка программного обеспечения для интерактивного моделирования малоугловых данных с помощью графических приложений

Обработка и последующее проведение моделирования данных МУРР от растворов макромолекул представляет собой многоэтапный процесс, включающий в себя применение большого набора математических алгоритмов и требующий строгого контроля на всех этапах анализа. В последние десятилетия в связи с бурным развитием инструментальной базы и вводом в строй большого количества синхротронных малоугловых станций с использованием роботизированных систем доставки образцов, появилась возможность проведения экспериментов в непрерывном поточном режиме и вопрос быстрой обработки большого объема экспериментальных данных МУРР в режиме реального времени стал чрезвычайно актуальным. На ряде станций были разработаны и успешно применены автоматические системы обработки данных МУРР (пайплайны) [А7], [23, 64, 65], позволившие значительной степени сократить время планирования экспериментов и улучшить эффективность работы пользовательских групп.

Однако, не все этапы анализа данных МУРР могут быть выполнены в автоматическом режиме во время проведения экспериментов на синхротронных станциях, и поэтому требуется наличие "портативного" программного обеспечения для анализа данных, которое пользовательские группы могут использовать у себя в лабораториях. Кроме того, автоматические оценки при выборе параметров модели не всегда могут оказаться оптимальными, и в ряде случаев требуется вмешательство пользователя и их ручное задание для получения приемлемого результата. В этой связи одним из важных требований при разработке программного обеспечения остается возможность проведения моделирования данных в интерактивном режиме. Такую задачу удобнее реализовать с помощью графических приложений, когда пользователь может видеть не только "статичные" цифры анализа, но и визуализировать полную картину данных на экране монитора.

Для проведения эффективного анализа данных МУРР нами был разработан ряд графических программ, позволяющих проводить как первичную обработку данных, так и последующее моделирование. Программы могут быть использованы как для анализа данных МУРР от растворов биологических многокомпонентных систем, так и для широкого ряда небиологических объектов и частично упорядоченных систем (наночастиц, полимеров, микроэмульсий и др.).

7.1 Первичная обработка экспериментальных данных в интерактивном режиме

Обработка данных МУРР в значительной степени остается нетривиальной многошаговой задачей. Новые расширенные возможности проведения экспериментов на современных синхротронных источниках требуют наличия гибких инструментов анализа данных МУРР.

С этой целью нами была разработана графическая программа PRIMUS [A5], которая представляет собой пакет обработки данных МУРР/МУРН, включающий основные этапы анализа данных для изотропных систем, начиная с первичной обработки исходных экспериментальных данных И заканчивая моделированием. Эта программа позволяет проводить манипуляции одномерными наборами данных МУРР/МУРН и вызывает другие программы через пользовательские интерфейсы. удобные PRIMUS содержит модули преобразования и обработки исходных ("сырых") данных МУРР с использованием форматов данных, которые ранее применялись на станции X33 синхротронного кольца DORIS-III (HASYLAB, DESY, Гамбург). В частности, PRIMUS позволяет проводить первичную обработку исходных данных, собранных линейным газовым детектором в формате ОТОКО [219], и двумерных данных, зарегистрированных детектором MAR345 image plate (http://www.marresearch.com/ip.htm). Кроме того, PRIMUS может работать с одномерными кривыми рассеяния от изотропных систем, полученными на любых других установках МУРР/МУРН, и заранее преобразованными в трех-колоночный текстовый (ASCII) формат данных. Такой ASCII файл должен содержать после своего заголовка угловую ось момента переноса импульса (вектора рассеяния) "s" в первой колонке, интенсивность "I(s)" во второй колонке и, при наличии, стандартную оценку ошибок интенсивности " $\sigma(s)$ " в третьей колонке. Напомним, вектор рассеяния определяется следующим образом: $s=4\pi\sin\theta/\lambda$, где 2θ - угол рассеяния, а λ - длина волны излучения.

Чистую интенсивность рассеяния от образца I(n) как функцию канала 1D детектора или бина n (при радиальном усреднения для 2D детектора) можно получить с помощью стандартного уравнения [220]

$$I(n) = \frac{1}{cDet(n)} \left[\frac{I_{s}(n)}{I_{0s}T_{s}} - \frac{I_{m}(n)}{I_{0m}T_{m}} - \frac{I_{e}(n)}{I_{0e}} \left(\frac{1}{T_{s}} - \frac{1}{T_{m}} \right) \right], \tag{7.1}$$

где подписи s, m и e обозначают рассеяние от образца, матрицы (буферного раствора) и пустой ячейки, соответственно, подпись θ означает интенсивность падающего пучка, T и c означают пропускание интенсивности образцом и концентрацию образца. В зависимости от типа системы, некоторые члены могут отсутствовать в уравнении (7.1), так, в случае $T_s = T_m$ третий член в уравнении (7.1) исчезает. Коррекция функции отклика детектора Det(n) может применяться либо к одномерным данным (для линейного газового детектора), либо к исходным двумерным данным МУРР перед их радиальным усреднением (в случае двумерного детектора).

Обработка файлов формата ОТОКО производится программой SAPOKO, вызываемой из меню PRIMUS. Данные по желанию маскируются, нормализуются по калибровочным каналам и корректируются с учетом отклика детектора. Для наборов данных, содержащих несколько кадров, они могут быть усреднены или сохранены по отдельности. Можно выбрать опорный кадр, и в усреднение будут включены только кадры, статистически совместимые с опорным (обычно для синхротронных исследований в качестве опорного выбирается первый кадр для контроля радиационного повреждения). Соответствующие ошибки вычисляются с помощью статистики Пуассона по значениям интенсивности (количеству отсчетов) в исходных данных [221]. Для получения угловой оси используются стандартные

образцы с известной периодичностью структуры (например, коллаген или Адбегенат). Используя информацию об угловой оси, обработанные данные МУРР преобразуются из бинарного ОТОКО формата в столбчатый текстовый формат ASCII.

Аналогичный пользовательский интерфейс доступен для обработки и радиального усреднения исходных данных, записанных детектором MAR Image Plate. В этом случае, программа FIT2D [222] используется для оценки положения центра пучка и создания файла маски, в котором отмечаются неактивные области детектора и область заслонки детектора (beamstop) вблизи центра падающего пучка. Бинарные данные MAR (TIFF формат) корректируются на отклик детектора, и затем пиксели с интенсивностями усредняются в радиальном направлении в соответствующие одномерные угловые бины. Показания калибровочных каналов извлекаются из заголовка файла MAR, а экспериментальные данные нормируются интенсивности прошедшего пучка и времени экспозиции. Выходной ASCII файл имеет формат, аналогичный формату обработанных файлов ОТОКО, но с дополнительной колонкой, содержащей стандартное отклонение, обусловленное радиальным усреднением.

Дальнейшие операции по манипулированию данными в программе PRIMUS, выполняемые с текстовыми ASCII файлами, уже не зависят от исходного формата данных. Все модули программы PRIMUS имеют считывающую процедуру, распознающую заголовочную запись и начало данных во входном ACSII файле. Пользовательский интерфейс программы PRIMUS, отображающий меню Tools для интерактивной работы с данными, представлен на Puc. 7.1. Угловые единицы во входных файлах могут быть преобразованы как между m^{-1} и m^{-1} и m

файлы в полулогарифмическом масштабе (Рис. 7.1). Более расширенная опция для графического показа использует программу SASPLOT, которая отображает файлы в отдельном окне с изменяемым размером, используя различные масштабы (например, фрактальный (логарифмический) график, график Порода и т.д.) и позволяет выбрать диапазон отрисовки и шкалирующий коэффициент с помощью манипулирования компьютерной мышью. Концентрация образца и шкалирующий множитель по умолчанию равны единице, но могут быть заданы пользователем (концентрация автоматически считывается из строки заголовка, если в ней содержится маркер "c=", за которым следует числовое значение).

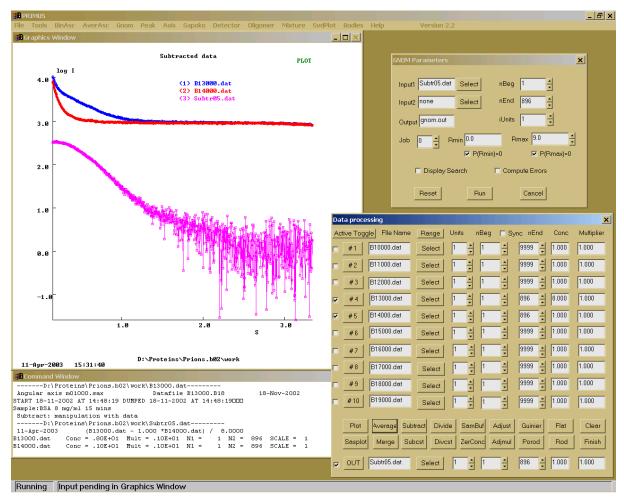


Рис. 7.1 Графический интерфейс программы PRIMUS, отображающий меню обработки данных МУРР/МУРН. В графическом окне изображены данные МУРР от белка бычьего альбумина BSA. Синяя кривая - профиль рассеяния от раствора BSA, красная - от буферного раствора, малиновая - кривая рассеяния BSA после вычитания фона. Также показано всплывающее диалоговое окно для вызова программы GNOM.

Меню 'Tools' для работы с данными включает в себя основные арифметические операции с двумя или несколькими наборами данных (вычитание, деление, усреднение и т.д.). Операции с двумя файлами также включают функцию корректировки, которая проверяет гипотезу о том, что два набора данных статистически различаются с точностью до масштабного коэффициента и (опционально) постоянного члена. Несколько файлов данных могут быть пришкалированы друг к другу и, при необходимости, объединены с помощью приближения методом наименьших квадратов в диапазоне их перекрытия. Для этого экспериментальные данные пересчитываются на новую угловую мастерсетку. Эта опция полезна, например, для построения составных кривых из данных МУРР/МУРН, записанных в разных угловых диапазонах. Для всех операций учет распространения ошибок выполняется с помощью стандартных уравнений [221]. Диалог манипулирования данными содержит двенадцать полей входных файлов и выходное поле, куда помещается результат каждой операции, который может быть использован в последующих манипуляциях с данными. Информация о выполненной операции (тип операции, имена исходных файлов, их веса, диапазоны используемых точек и т.д.) записывается в верхний и нижний колонтитулы выходного файла. Заголовки и колонтитулы файлов данных могут быть визуализированы в текстовом окне программы PRIMUS для отслеживания истории операций.

Для экстраполяции данных к нулевой концентрации образца в программе PRIMUS доступны две опции. Один вариант учитывает конечную объемную долю образца v_k путем решения системы линейных уравнений

$$I_{\text{exp}}^{(k)}(s) = v_k I_{\text{sample}}(s) + (1 - v_k) I_{\text{matrix}}(s), \tag{7.2}$$

где индекс k пробегает по имеющимся измерениям в серии концентраций для разделения двух функций, $I_{sample}(s)$ и $I_{matrix}(s)$ (это полезно для вычитания фона при измерениях белковых растворов в области больших углов для высоких концентраций растворителя). Другой вариант учитывает притягивающие или отталкивающие взаимодействия, которые в основном влияют на данные в области

очень малых углов. Для этого несколько моделей рассеяния при различных концентрациях соответствующим образом масштабируются, и начальная часть данных экстраполируется к нулевой концентрации в предположении линейной зависимости от концентрации.

Программа PRIMUS предоставляет удобный интерфейс для расчета общих параметров монодисперсных и полидисперсных систем по данным рассеяния. Радиус инерции R_g и рассеяние в нулевой угол I(0) вычисляются с использованием приближения Гинье, согласно уравнению (1.8), которое справедливо для (sR_g) < 1.3 [5]. Первая и последняя точки кривых рассеяния, используемые при расчете Гинье приближения, могут быть изменены интерактивно, а график Гинье при этом автоматически перерисовывается. Аналогичные опции доступны для вычисления радиусов инерции толщины И поперечного сечения ДЛЯ плоских И стержнеподобных частиц, соответственно. Исключенный объем V частицы [33] оцениваются с помощью инварианта Порода Q как

$$V = 2\pi^2 I(0)/Q = 2\pi^2 I(0) / \int_0^\infty s^2 (I(s) - K) ds.$$
 (7.3)

Константа K (необходимая для обеспечения спадания интенсивности по закону s^{-4}) автоматически вычисляется по асимптотике Порода ($I(s)s^4$ versus s^4) при больших углах.

Более сложные операции ПО вычислению структурных характеристик выполняются путем вызова внешних программных модулей из графического интерфейса программы PRIMUS. Во всех случаях, когда вызывается внешняя программа, файл из поля вывода (или, если оно пустое, первый активный файл из поля ввода) панели инструментов 'Tools' передается в эту программу для обработки. Для быстрой оценки общей формы частиц программа BODIES приближает начальную часть данных рассеяния рассеянию ПО OT трехпараметрических геометрических тел (трехосный эллипсоид, эллипсоид вращения, круглый, эллиптический или полый цилиндр, призма, полая сфера). Интенсивности рассеяния от тел I(s) оцениваются по стандартным формулам [3], и программа минимизирует невязку χ^2 согласно уравнению (1.20). В результате находится наилучшее приближение для каждого класса геометрических тел, и выдаются соответствующие кривые рассеяния вместе с найденными параметрами тел.

Для расчета характеристических функций от разбавленных монодисперсных или полидисперсных систем программа PRIMUS вызывает программу GNOM [40], которая использует метод непрямого Фурье преобразования для решения интегрального уравнения

$$I_{\exp}(s) = \int_{r_{\min}}^{r_{\max}} p(r) \cdot K(s, r) dr, \qquad (7.4)$$

где p(r) - функция распределения (например, распределение по расстояниям внутри частицы или распределение по размерам частиц), определенная в диапазоне [r_{min} , r_{max}], а K(s,r) - интегральное ядро соответствующего преобразования Фурье. Задаваемые пользователем параметры, включая тип системы (монодисперсная или полидисперсная), диапазоны в прямом и обратном пространстве и т.д., задаются в диалоговом окне. Критерии для выбора параметра регуляризации, встроенные в программу GNOM, обычно обеспечивают нахождение функции распределения по расстояниям p(r) без дополнительного вмешательства пользователя.

Структурные характеристики частично упорядоченных систем вычисляются по максимумам в профилях рассеяния с помощью программы РЕАК. Более подробно об этом будет рассказано в разделе 7.5. Для количественного анализа состава многокомпонентных полидисперсных систем и оценки числа компонентов в программе PRIMUS могут быть использованы модули 'OLIGOMER', 'MIXTURE' и 'SVDPLOT', алгоритмы и применение данных программ были подробно описаны в Главе 3.

Программа PRIMUS работает на операционной системе Windows и не требует дополнительных программных пакетов или языков, таких как Matlab, IgorPro или IDL. Имеется также расширенный кроссплатформенный вариант программы Primus (PrimusQt), который использует графическую библиотеку Qt [A8]. Оба

варианта входят в пакет программ ATSAS и широко используются научным сообществом.

Таким образом, программа PRIMUS позволяет проводить комплексный анализ и интерактивно обрабатывать одномерные рентгеновские и нейтронные малоугловые кривые рассеяния для различных типов изотропных и частично упорядоченных систем, таких как растворы биополимеров и коллоидных частиц, микроэмульсии, полимеры и др.

7.2 Интерактивное моделирование четвертичной структуры белковых комплексов

За последние десятилетия было получено огромное количество информации о структуре и функции отдельных белков. С наступлением эры "постгеномики" ожидается получение тысяч новых структур белков с помощью рентгеновской кристаллографии, криоэлектронной микроскопии и ядерного магнитного резонанса (ЯМР) [223, 224]. Однако наиболее важные клеточные функции выполняются не отдельными белками, а макромолекулярными комплексами. Такие комплексы обычно слишком велики для ЯМР-исследований, и они часто обладают присущей структурной гибкостью, ИМ ЧТО затрудняет ИХ кристаллизацию. Для анализа больших макромолекулярных комплексов необходимы альтернативные экспериментальные методы и подходы к анализу данных, чтобы определить, как свойства отдельных макромолекул связаны с общей структурой и функцией макромолекулярных машин, существующих в живых клетках.

Данные МУРР/МУРН чувствительны к изменениям четвертичной структуры макромолекул, поэтому эти методы особенно полезны для анализа макромолекулярных комплексов [4]. Если с помощью разработанных ранее *ab initio* методов [49, 52] удается надежно восстанавливать трехмерную форму частиц на низком разрешении, то гораздо более детальные модели могут быть построены

при наличии структур высокого разрешения отдельных доменов или субъединиц, составляющих комплекс.

Как правило третичная структура доменов в основном сохраняется при образовании комплекса, и для построения макромолекулярных комплексов можно использовать метод молекулярной тектоники (или rigid body modelling - моделирование жесткими телами) [225-227]. Например, для сборки из двух субъединиц структуру комплекса можно определить путем варьирования шести позиционных параметров, описывающих относительное положение и ориентацию второй субъединицы относительно первой.

Рассмотрим комплекс, состоящий из двух субъединиц (A и B). Амплитуды рассеяния от субъединиц с центром в начале координат в исходных ориентациях обозначаются как A(s) и B(s), соответственно. Произвольный комплекс может быть построен путем фиксации первой субъединицы и вращения и перемещения второй. Вращение описывается углами Эйлера [228], а сдвиг - вектором $\mathbf{u} = (\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z)$, так что вся операция описывается шестью параметрами.

Обозначим амплитуду рассеяния от смещенной второй субъединицы как C(s), тогда рассеяние от комплекса примет вид [229, 230]

$$I(s) = I_A(s) + I_B(s) + 2\langle A(s)C^*(s) \rangle_{\Omega}, \tag{7.5}$$

где $<>_{\Omega}$ обозначает сферическое усреднение в обратном пространстве.

Удобно представить амплитуды рассеяния с помощью сферических гармоник $Y_{lm}()$ в виде

$$A(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm}(s) Y_{lm}(\Omega)$$
 (7.6)

$$B(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B_{lm}(s) Y_{lm}(\Omega), \qquad (7.7)$$

$$C(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} C_{lm}(s) Y_{lm}(\Omega).$$
 (7.8)

Благодаря ортогональным свойствам сферических гармоник, сферическое среднее в уравнении (7.5) приводит к выражению

$$I(s) = 2\pi^{2} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(A_{lm}(s)^{2} + \left| B_{lm}(s)^{2} + 2 \operatorname{Re} \left[A_{lm}(s) C_{lm}^{*}(s) \right] \right)$$
 (7.9)

Когда структуры субъединиц известны, амплитуды рассеяния и соответствующие парциальные функции $A_{lm}(s)$ и $B_{lm}(s)$ могут быть вычислены из атомных моделей, как описано [70, 75]. Парциальные функции $C_{lm}(s)$ повернутой и перемещенной второй субъединицы выражаются аналитически как [229, 230]:

$$C_{lm}(s) = \sum_{k=-l}^{l} D_{mk}^{l}(\alpha, \beta, \gamma) C_{lk}^{0}(s, \mathbf{u}), \tag{7.10}$$

где $D_{mk}^{\ l}(\ ,\ ,\)$ обозначают элементы матрицы конечных вращений [228]. Если предположить без потери общности, что весь комплекс повернут так, что направление вектора **u** совпадает с осью Z, то функции $C_{lk}^{\ 0}(s,u)$ примут вид [48]

$$C_{lm}^{0}(s,\mathbf{u}) = (-1)^{m} \sum_{p=0}^{\infty} j_{p}(su) \sum_{k=|l-p|}^{l+p} d_{lm}(k,p) \sum_{j=-k}^{k} B_{kj}(s)$$
(7.11)

где $j_p(x)$ - сферические функции Бесселя, а коэффициенты $d_{lm}(k,p)$ представлены через 3j символы Вигнера [228]

$$d_{lm}(k,p) = i^{p}(2p+1)\sqrt{(2l+1)(2k+1)} \begin{pmatrix} l & p & k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & p & k \\ -m & 0 & m \end{pmatrix}.$$
 (7.12)

Таким образом, идея моделирования жесткими телами по данным малоуглового рассеяния от раствора макромолекулярного комплекса заключается в определении шести позиционных параметров второй субъединицы, минимизирующих невязку χ^2 согласно уравнению (1.20). Уравнение (7.9) может быть легко обобщено на систему из K жестких тел, которые в общем случае будут описываться 6(K-1) позиционными параметрами.

На основе использования рассмотренного выше алгоритма была разработана графическая программа MASSHA [A17], которая позволяет отображать 3D модели и проводить манипулирование атомными моделями высокого разрешения для уточнения четвертичной структуры макромолекулярных комплексов.

Программа MASSHA может одновременно представлять до 10 различных макромолекулярных структур (с общим числом атомов до 500000), причем для каждой из них предусмотрены свои особенности рисования, такие как цвет,

яркость и тип представления. Основные опции пользовательского интерфейса программы MASSHA включают в себя меню ADD / REMOVE BODY для загрузки и удаления графических объектов, меню ZOOM IN / OUT для масштабирования графического окна, меню COMMANDS для выполнения операций с объектами (например, преобразование координат, изменение атрибутов, сохранение текущей позиции), а также меню MODELLING, открывающее интерфейс для интерактивного и автоматического моделирования жестких тел. Также доступна интерактивная справка по основным командам и опциям.

Программа MASSHA представляет трехмерные объекты либо как твердые тела, либо как наборы точек. Первый вариант обычно используется для визуализации моделей низкого разрешения и отображает поверхность частиц с помощью функции угловой огибающей $r=F(\omega)$, где (r,ω) - сферические координаты в прямом пространстве. Информация хранится в ASCII-файле, содержащем мультипольные коэффициенты представления поверхности с помощью сферических гармоник [47]

$$F(\omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_{lm}(s) Y_{lm}(\omega)$$
 (7.13)

Файлы, содержащие коэффициенты flm (*.flm), генерируются, в частности, программой *ab initio* определения формы SASHA [48] и программой CRYSOL [70]. Для отображения структуры используется алгоритм триангуляции, где объекты твердого тела представляются либо в виде каркасов, либо в виде затененных поверхностей Гуро [79, 231].

Для второго типа объектов MASSHA считывает координаты в стандартном формате Protein Data Bank (PDB) [232]. Атомная модель представляется либо в виде C_{α} цепи (в случае белковых макромолекул), либо цепи из P атомов (для молекул ДНК/PHK). Формат PDB (*.pdb) также может использоваться для чтения и отображения шариковых моделей, полученных, например, путем *ab initio* определения формы [52]. Шарики могут быть представлены в виде открытых кругов или затененных сфер с переменным радиусом. На Рис. 7.2 показаны

различные режимы представления формы/атомной структуры с использованием моделей лизоцима высокого и низкого разрешения (PDB ID: 6LYZ).

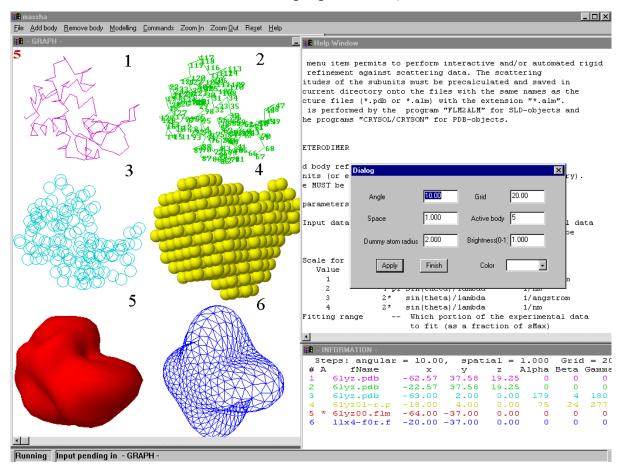


Рис. 7.2 Графический интерфейс программы MASSHA, иллюстрирующий различные режимы представления формы/атомной структуры лизоцима. (1) и (2) C_{α} -цепь без и с нумерацией аминокислотных остатков, соответственно; (3) C_{α} -цепь в виде набора сфер; (4) аb initio шариковая модель лизоцима, полученная программой DAMMIN; (5) заштрихованная оболочка Гуро *ab initio* формы полученной программой SASSHA; (6) проволочно-остовный вид *ab initio* формы.

Каждый графический объект (форма или структура) имеет свои собственные графические характеристики и текущую матрицу преобразования (последняя изменяется при поворотах и сдвигах). Информация о текущем местоположении объекта доступна через горячую клавишу. Относительные сдвиги и повороты могут выполняться как через диалоговое меню, так и с помощью горячих клавиш. Угловое приращение по умолчанию для преобразований с помощью горячих клавиш составляет $\Delta \phi = 5^{\circ}$, а пространственное приращение по умолчанию (X=Y=Z) около 1/20 окна экрана (оба приращения могут быть изменены

пользователем). Диалог меню трансформаций моделей позволяет позиционировать объект в абсолютной системе координат или сдвигать его в относительных единицах от текущего положения. Программа MASSHA позволяет сохранять PDB-объекты в их текущем положении, что удобно для хранения результатов интерактивного или автоматизированного моделирования. Опция "Align PDB" автоматически совмещает два произвольных объекта низкого или высокого разрешения (согласно алгоритму программы SUPCOMB [57].

Для моделирования жесткими телами гетеродимерной структуры в программе MASSHA мультипольные компоненты амплитуд рассеяния двух субъединиц должны быть предварительно вычислены и сохранены в виде отдельных файлов с расширением *.alm. Далее две субъединицы загружаются как отдельные тела, и моделирование начинается с выбора опции HETERODIMER из меню MODELLING. В диалоговом окне HETERODIMER должны быть заданы следующие параметры:

- -1 Входной файл, содержащий экспериментальные данные МУРР/МУРН. Имя файла можно выбрать с помощью меню выбора файлов (опция "Select").
- -2 единицы измерений для s-вектора (nm⁻¹ или A⁻¹)
- -3 Угловой диапазон (в долях от максимального значения *s*-вектора)
- -4 Имя выходного файла, который будет содержать наилучшее приближение к данным
- -5 Предварительно вычисленные *.alm файлы первой и второй субъединиц
- -6 Предварительные множители (например, контрасты) двух субъединиц.

Программа MASSHA предупреждает пользователя, если информация неверная или неполная. Возможны два режима работы при уточнении структуры: интерактивный (кнопка COMPUTE) и автоматический (кнопка REFINE). Кнопка COMPUTE проводит расчет интенсивности кривой рассеяния от текущей модели комплекса и затем отображает вычисленное приближение к экспериментальным данным в отдельном окне 'Curve Fit'. Пользователь может изменить положение двух субъединиц, манипулируя объектами в 3D окне, и пересчитать невязку χ^2 ,

нажав кнопку COMPUTE. Кнопка REFINE запускает программу автоматического уточнения, при этом вторая субъединица перемещается вокруг своего текущего положения на (ХҮZ) (т.е. зондируются три точки вдоль каждого направления X, Y и Z). Его ориентация изменяется вокруг текущей ориентации путем поворота вокруг X, Y и Z на $\pm 2*\Delta \varphi$, т.е. исследуется пять поворотов вокруг каждой оси. Пользователь может выбрать номер остатка для определения центра вращения (по умолчанию второе тело вращается вокруг своего геометрического центра). После уточнения второе тело перемещается положение, обеспечивающее наилучшее соответствие данным, а его исходное положение добавляется в качестве третьего тела для визуализации изменения модели. Пользователь может скорректировать решение вручную или продолжить автоматическую процедуру, начиная с нового положения второго тела.

Специальный режим предусмотрен в программе MASSHA для практически важного случая гомодимерной структуры, обладающей осью симметрии второго порядка. Если структура мономера известна, то модель гомодимера определяется только ориентацией мономера и расстоянием между центром последнего и осью симметрии, т.е. четырьмя параметрами, а не шестью, как для гетеродимера. Таким образом, симметрия частиц накладывает существенные ограничения на возможную модель, и интенсивность рассеяния от комплекса (уравнение 7.9) зависит в этом случае только от амплитуд первого тела (мономера). Если предположить, что ось симметрии совпадает с осью Y, а мономеры разделены вдоль Z на расстояние 2u_z, то интенсивность может быть выражена следующим образом [48]

$$I(u_{z},s) = 2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} |(-1)^{m} \sum_{p=0}^{\infty} \text{Re}[i^{p} j_{p}(su_{z})(2p+1)]$$

$$\sum_{k=|l-p|}^{l+p} \sqrt{(2l+1)(2k+1)} \binom{l}{0} \sum_{0}^{m} \binom{l}{0} \binom{p}{0} \binom{k}{-m} \binom{k}{0} \binom{k}{j}_{j=-k}^{k} A_{kj}(s)|^{2}$$

$$(7.14)$$

где $A_{kj}(s)$ - амплитуды повернутого мономера, вычисленные согласно уравнению (7.10).

Для использования опции HOMODIMER достаточно загрузки одного тела (мономера). Второй мономер генерируется автоматически, принимая ось Y за ось симметрии второго порядка. В этом случае необходимо указывать те же параметры, что и при моделировании в опции HETERODIMER, за исключением амплитуды и коэффициента второго тела. Отметим, что в режиме HOMODIMER движения и/или повороты первого тела приводят к соответствующим операциям симметрично связанного с ним второго тела.

Практический пример моделирования гомодимерной структуры тетрамерной пируватдекарбоксилазы дрожжей (PDC) представлен на Рис. 7.3. Белок PDC - это тиаминдифосфат-зависимый фермент, участвующий в некоторых спиртового брожения. PDC представляет собой тетрамер с молекулярной массой 236 кДа при низком рН и диссоциирует на димеры при высоком рН [233]. Теоретическая кривая рассеяния, рассчитанная по кристаллографической модели (PDB ID: 1PVD), значительно отличается от экспериментальных данных (χ^2 = 1.72). В работе [234] было проведено трудоемкое ручное уточнение этой структуры. Автоматический режим уточнения в программе MASSHA позволяет получить такую же модель, как в работе [234], всего за несколько минут (Рис. 7.3). Димеры в уточненной модели наклонены на 11° вокруг оси X, а расстояние между ними уменьшено на 0.4 нм. Модель хорошо согласуется с экспериментальными данными ($\chi^2 = 1.21$). Среднеквадратичное смещение атомов в уточненной модели составляет 0.57 нм по сравнению со структурой в кристалле. Такое большое конформационное изменение можно объяснить слабыми контактами между димерами в кристаллографической модели (Рис. 7.3).

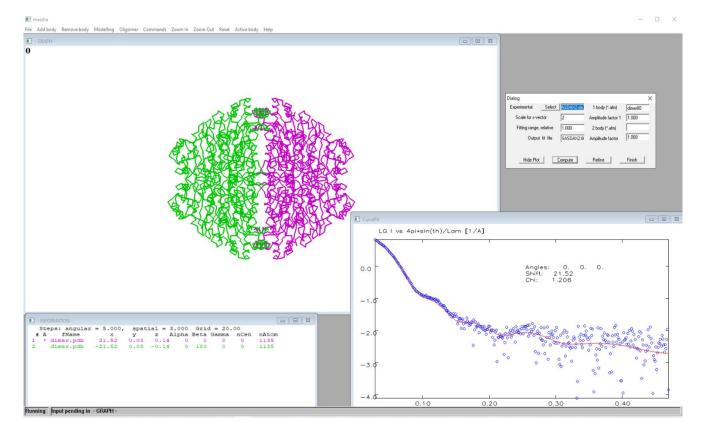


Рис. 7.3 Графический интерфейс программы MASSHA, меню 'Homodimer Modelling'. Моделирование по данным МУРР методом молекулярной тектоники тетрамерной структуры белка пируват декарбоксилазы. Две димерные молекулы изображены в виде C_{α} цепей зеленым и малиновым цветами. Экспериментальные данные МУРР показаны синими кружками, кривая рассеяния от текущей модели – сплошной красной линией. В информационном окне показаны координаты геометрического центра димеров, углы вращения Эйлера относительно исходной структуры и общее число аминокислотных остатков.

Кроме того, в программе MASSHA имеется возможность моделирования симметричных комплексов с точечной симметрией РЗ - Р6 и Р22 - Р62. Когда мономер перемещается/вращается пользователем, автоматически генерируются все его симметричные партнеры, рассчитывается рассеяние от всего комплекса и отображается его соответствие экспериментальным данным. Для несимметричных комплексов добавлена возможность мульти-субъединичного моделирования, позволяющая манипулировать семью отдельными структурами. Пользователь может изменить положение/ориентацию отдельной субъединицы или выбрать группу, состоящую из нескольких субъединиц, которая затем будет перемещаться как единое целое. При перемещении или повороте субъединицы/группы

интенсивность рассеяния от всего комплекса пересчитывается и отображается ее соответствие экспериментальным данным. Опции моделирования как симметричных, так и мульти-субъединичных тел также позволяют проводить автоматическое локальное уточнение путем поиска в окрестностях текущей конфигурации.

Отметим, что взаимодействия между субъединицами в макромолекулярных комплексах нековалентны по своей природе, и это часто позволяет перемещать и вращать субъединицы как жесткие тела. В частности, энергия таких взаимодействий сравнима с энергией взаимодействия между соседними молекулами в кристаллах. Таким образом, метод молекулярной тектоники оказывается очень полезным для анализа различий в четвертичной структуре мульти-субъединичных белков, вызванных силами кристаллической упаковки [234].

В заключение, программа MASSHA предоставляет расширенные возможности для сочетания интерактивного и автоматического режимов моделирования данных МУРР/МУРН методом молекулярной тектоники. Кроме того, важными дополнительными опциями программы MASSHA являются автоматическое выравнивание объектов и учет симметрии мульти-субъединичных комплексов.

7.3 Суперпозиция атомных моделей с использованием разложения на сферические гармоники

Методы анализа данных МУРР/МУРН позволяют восстанавливать трехмерные модели макромолекул различной природы и разрешения, и сравнение таких разнородных моделей часто требуется для перекрестной проверки (валидации) структурных результатов, полученных различными методами. Для таких сравнений обычно требуется автоматическая суперпозиция трехмерных структур с наилучшим соответствием. Наложение разнородных моделей, которые могут не только отличаться по разрешению и способам представления (атомные модели, шариковые модели, цепи псевдо-аминокислотных остатков, электронные карты

плотности, поверхности и т.д.), но и содержать части с существенно различной плотностью (например, нуклеопротеиновые комплексы), не является тривиальной задачей.

Ранее была разработана программа SUPCOMB [57] для сопоставления трехмерных структур высокого и низкого разрешения, которая использует нормализованное пространственное расхождение (Normalized Spatial Diescrepancy - NSD) в качестве меры близости между объектами. Для каждой точки (шарика или атома) в первой модели находится ближайшая соседняя точка во второй модели, и то же самое делается для всех точек во второй модели. Квадрат расстояния между ближайшими точками складывается и нормируется на квадрат среднего расстояния между соседними точками для двух моделей, таким образом можно рассчитать значение NSD для заданных относительных положений моделей. Вначале происходит выравнивание моделей вдоль их осей инерции, а затем программа SUPCOMB минимизирует величину NSD и, таким образом, находит наилучшее наложение структур. Однако процессорное время, используемое программой SUPCOMB, пропорционально произведению числа точек (шариков или атомов), представляющих два объекта, и для больших макромолекулярных комплексов время вычислений становится недопустимо большим.

Чтобы устранить этот недостаток, нами был разработан алгоритм для быстрого совмещения макромолекулярных моделей, основанный на представлении их амплитуд рассеяния в пространстве Фурье в виде сферических гармоник [A12]. Данный метод использует нормализованный интегральный перекрестный коэффициент произведения амплитуд рассеяния в качестве меры близости между двумя 3D-объектами низкого разрешения. Метод реализован в программе SUPALM, которая работает значительно быстрее, чем SUPCOMB, и вдобавок применима для сравнения форм низкого разрешения и гетерогенных моделей различной природы (атомы, шарики, ЭМ карты и т.д.).

Напомним, что для произвольного 3D объекта с плотностью рассеяния $\rho_A(r)$ в прямом пространстве, амплитуда рассеяния A(s) может быть представлена с помощью сферических гармоник как [47]

$$A(s) = \sum_{l=0}^{L} \sum_{m=-l}^{l} A_{lm}(s) Y_{lm}(\Omega)$$
 (7.15)

где вектор рассеяния $s=4\pi\sin\theta/\lambda$, 2θ - угол рассеяния, λ - длина волны, Ω - телесный угол в обратном пространстве, $s=(s,\Omega)$, а величина L определяет разрешение модели. Для объекта, представленного N точками (атомами, шариками или значениями электронной плотности), парциальные амплитуды $A_{lm}(s)$ вычисляются как

$$A_{lm}(s) = 4\pi i^{l} \sum_{j=1}^{N} f_{j}(s) j_{l}(sr_{j}) Y_{lm}^{*}(\omega_{j})$$
(7.16)

где сумма проходит по всем элементам с координатами $\mathbf{r}_j = (r_j, \omega_j) = (r_j, \theta_j, \varphi_j)$, а $f_j(s)$ - соответствующие атомные форм-факторы рассеяния.

Благодаря ортогональным свойствам сферических гармоник, сферически усредненная интенсивность рассеяния $I_A(s)$ (например, измеренная в малоугловом эксперименте от ансамбля случайно ориентированных частиц) выражается как сумма отдельных мультипольных вкладов

$$I_A(s) = 2\pi^2 \sum_{l=0}^{L} \sum_{m=-l}^{l} |A_{lm}(s)|^2$$
(7.17)

Рассмотрим теперь два трехмерных объекта с плотностями $\rho_A(r)$ и $\rho_B(r)$. Наилучшая суперпозиция этих объектов должна максимизировать интеграл корреляции, выраженный как $\int [\rho_A(r) + \rho_B(r)]^2 dr$. По теореме Парсеваля [34] этот интеграл равен полной интенсивности в обратном пространстве

$$T(A,B) = \int [A(s) + B(s)]^2 ds = \int [I_A(s) + I_B(s) + 2A(s)B^*(s)] ds$$
$$= 4\pi \int [I_A(s) + I_B(s) + 2I_{AB}(s)] s^2 ds \qquad (7.18)$$

где $I_A(s)$ и $I_B(s)$ - усредненные интенсивности рассеяния объектов A и B, соответственно, а $I_{AB}(s)$ обозначает перекрестное слагаемое. Заметим, что отдельные интегральные интенсивности не зависят от положения, и только интеграл от перекрестного слагаемого изменяется при перемещении/вращении

частиц. Используя разложение сферических гармоник (уравнения 7.15-7.17), три члена в уравнении (7.18) легко вычисляются через свои парциальные амплитуды рассеяния, и количественная мера согласия между двумя трехмерными объектами в прямом пространстве может быть выражена в обратном пространстве как нормализованный коэффициент корреляции (*NCC*):

$$NCC(A,B) = \frac{\int_0^{Sm} [A_{lm}(s)B_{lm}^*(s) + A_{lm}^*(s)B_{lm}(s)] s^2 ds}{2(\int_0^{Sm} I_A(s)s^2 ds)^{\frac{1}{2}} (\int_0^{Sm} I_B(s)s^2 ds)^{\frac{1}{2}}}$$
(7.19)

где $A_{lm}(s)$ и $B_{lm}(s)$ - амплитуды рассеяния объектов A и B, соответственно. Здесь все интегралы оцениваются В ограниченном **УГЛОВОМ** диапазоне определяющем разрешение объектов. Теоретически (при бесконечном верхнем пределе интегрирования), NCC изменяется между 0 и 1, последнее значение соответствует идеальному перекрытию двух идентичных структур. Таким образом, коэффициент корреляции NCC может служить удобной мерой близости, и его максимизация должна позволить получить наилучшее перекрытие двух объектов. Численная минимизация меры близости между трехмерными неоднородными объектами с учетом позиционных и вращательных параметров является нетривиальной задачей. Целевая функция может иметь множество локальных минимумов, в которых алгоритмы поиска могут застрять. Алгоритмы глобальной минимизации (например, метод имитации отжига) могут преодолеть локальные минимумы, но они зачастую неэффективны с вычислительной точки зрения. Практический подход к решению этой проблемы заключается в использовании локальной минимизации, начиная с предварительно выровненных позиций. Их можно получить путем процедуры выравнивания моделей вдоль осей инерции. Главные оси инерции находятся для обоих объектов как собственные векторы матрицы тензора инерции, представляющие собой линейные комбинации вторых центральных моментов распределения вокруг центров масс. Два объекта устанавливаются в такие канонические позиции, где они центрированы по началу координат и повернуты так, чтобы их главные оси инерции, взятые в порядке возрастания собственных значений, были выровнены по осям X, Y и Z,

соответственно. В зависимости от того, разрешены или нет энантиоморфные структуры, существует четыре или восемь комбинаций знаков для собственных значений.

Исходя из указанных соображений, был предложен следующий алгоритм для суперпозиции двух трехмерных объектов

(модель А и модель В):

- (i) Вычисляются тензоры инерции и их собственные векторы для модели A и модели B.
- (ii) Модель В поворачивается на $\mathbf{M}_{\mathbf{A}}\mathbf{M}_{\mathbf{B}}^{\mathbf{T}}$ и сдвигается на $\mathbf{T}_{\mathbf{B}}$ - $\mathbf{T}_{\mathbf{A}}$, чтобы совместить ее главные оси с осями модели А (здесь $\mathbf{M}_{\mathbf{A}}$ и $\mathbf{M}_{\mathbf{B}}$ матрицы поворота модели А и модели В соответственно, $\mathbf{T}_{\mathbf{A}}$ и $\mathbf{T}_{\mathbf{B}}$ соответствующие смещения центров масс от начала системы координат). Знаки столбцов матрицы вращения $\mathbf{M}_{\mathbf{B}}$ выбираются из четырех комбинаций знаков (или восьми, если разрешены энантиоморфы).
- (ііі) Оцениваются амплитуды рассеяния двух объектов $A_{lm}(s)$ и $B_{lm}(s)$. Для этого используется программа CRYSOL [70] для атомных моделей высокого разрешения или программа DAM2ALM (доступна в пакете ATSAS [A6-A10]) для *ab initio* форм низкого разрешения. Также предусмотрены опции для прямого чтения карт электронной плотности (файлы в формате EMD) или *ab initio* многофазных шариковых моделей (полученных программой MONSA [100]) и расчета амплитуд их рассеяния.
- (iv) Положение и ориентация модели В уточняются путем минимизации значения 1/NCC с помощью локальной минимизации методом нелинейных наименьших квадратов библиотеки NL2SOL [235]. Полученное наилучшее значение NCC выдается как оценка меры близости между объектами.
- Шаги (i) и (ii) аналогичны тем, что используются в программе SUPCOMB, однако шаги (iii) и (iv) отличаются и имеют ряд преимуществ. Во-первых, использование предварительно вычисленных амплитуд позволяет легко сопоставлять разнородные объекты и модели различной природы. Во-вторых, изменения $B_{lm}(s)$ при поворотах и смещениях объекта В быстро рассчитываются с помощью

матрицы элементов конечных вращений [48, 228]. Важно, что амплитуды рассеяния рассчитываются только один раз, и вычислительные затраты алгоритма не зависят от сложности накладываемых структур (в отличие от SUPCOMB, где время расчета пропорционально произведению числа элементов в двух объектах, N_A*N_B).

Предложенный метод был реализован в компьютерной программе SUPALM, и его эффективность была проверена на ряде тестовых случаев для различных моделей биологических макромолекул с высоким и низким разрешением и сравнена с работой программы SUPCOMB. Большинство моделей низкого разрешения, представленных Таблицах 7.1-7.3, являют собой экспериментально определенные ab initio формы белков, восстановленные по данным МУРР, полученным на синхротронных станциях X33 и P12 (EMBL c/o DESY) с помощью программ DAMMIN [52], DAMMIF [55] и GASBOR [53]. Все эти 3D формы накладывались на имеющиеся кристаллические структуры этих белков, взятые из базы данных Protein Data Bank (PDB). Чтобы проиллюстрировать работу SUPALM на картах ЭМ, шариковая модель, полученная на основе экспериментальной ЭМ карты электронной плотности, накладывалась на модель гамма-секретазы человека высокого разрешения, полученную с помощью криоЭМ [236]. В случае гетерогенного комплекса модель белок-РНК в 70S рибосоме *E.coli*, полученная на основе данных МУРР/МУРН [237], сравнивалась с недавно полученной крио-ЭМ структурой рибосомы [238]. Как видно из Таблиц 7.1-7.3, для всех этих различных объектов значения NSD, полученные программой SUPCOMB, хорошо согласуются с NSD, рассчитанными на основе моделей, найденных программой SUPALM. На Рис. 7.4-7.6 представлены опорные модели (высокого разрешения) и *ab initio* модели, суперпозированные программами SUPCOMB/SUPALM, для иллюстрации сходства результатов этих двух программ.

Таблица 7.1 Сравнение результатов суперпозиции моделей с использованием программ SUPALM и SUPCOMB. Шаблонные (опорные) объекты представляют собой структуры высокого разрешения (их амплитуды рассеяния были рассчитаны программой CRYSOL), а сопоставленные объекты - ab initio формы низкого разрешения (с амплитудами рассеяния, рассчитанными программой DAM2ALM), которые были получены с помощью программ DAMMIN [52], DAMMIF [55], GASBOR [55] и MONSA [100]. Значения NSD вычислялись между шаблоном и сопоставленными объектами, полученными с помощью программ SUPCOMB и SUPALM. Значения NCC оценивались с помощью программы SUPALM. Коэффициент выигрыша в скорости вычислений определяется как отношение расчета программой процессорного времени **SUPCOMB** времени, затрачиваемому программой SUPALM. Для всех тестов использовались параметры SUPALM по умолчанию (L_{max} =5 и N_{sh} =7).

№.	Шаблонный	Налагаемая	MW,	NSD,	NCC,	NSD,	NCC,	Фактор
	объект	ab initio	кДа	sup-	sup-	supalm	supalm	выигрыша в
		форма		comb	comb			скорости
								(Supalm <i>vs</i> .
								Supcomb)
1	2A38	z1z2	21	1.70	0.87	1.79	0.96	1.6
	(pdb)	(DAMMIF)						
2	1GOT	G protein	40	1.12	0.72	1.14	0.76	2.8
	(pdb)	(GASBOR)						
3	106V	Internalin	50	1.69	0.74	1.94	0.95	3.2
	(pdb)	(DAMMIN)						
4	4UIS	Emd-2974	140	3.65	0.86	3.75	0.93	70.0*
	(pdb)	(EMD map)*						
5	1HQK	Lumazine	960	1.22	0.83	1.24	0.87	10.1
	(pdb)	synthase						
		(DAMMIN)						
6	4V4W	70S	2150	1.10	0.81	1.27	0.94	25.2
	(pdb)	ribosome						
		(MONSA)						

^{*}Для электронной карты EMD (Emd-2974) сравнение с программой SUPCOMB проводилось следующим образом. Для программы SUPCOMB карта электронной плотности была сначала преобразована в шариковую модель (содержащую 29374 атома) с использованием порога электронной плотности 0.04 с помощью программы EM2DAM из пакета ATSAS, а затем наложена на pdb структуру (PDB ID: 4UIS). SUPALM выполнял прямое наложение файлов PDB и EMD. RMSD между депонированной моделью EMDB и моделью SUPALM составляет 5.5 Å.

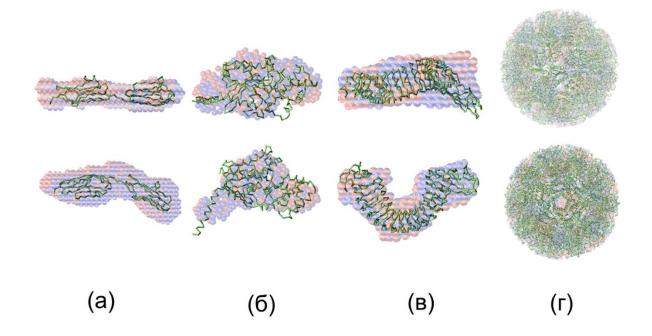


Рис. 7.4 Кристаллические структуры белков (зеленые C_{α} цепи) с суперпозированными *ab initio* формами, полученными с помощью программ SUPALM (красные сферы) и SUPCOMB (синие сферы). Панель (a) - белок Z1Z2 (PDB ID: 2A38), (б) - белок G (PDB ID: 1GOT), (в) – белок интерналин (PDB ID: 1O6V), (г) – фермент люмазинсинтаза (PDB ID: 1HQK). Нижний ряд моделей повернут против часовой стрелки на 90° вокруг горизонтальной оси.

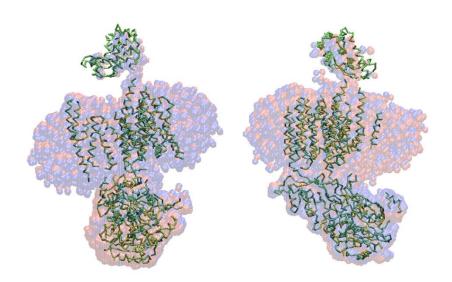


Рис. 7.5 Гибридная модель высокого разрешения гамма-секретазы человека, полученная из крио-ЭМ (PDB ID: 4UIS) (зеленые C_{α} цепи) и наложенные на нее аb initio формы, выделенные из карты электронной плотности (ID: EMD-2974.map). Результаты наложения программой SUPALM и SUPCOMB показаны красными и синими сферами, соответственно. Вид справа повернут против часовой стрелки на 90° вокруг вертикальной оси.

Таблица 7.2 Результаты тестирования работы программы SUPALM с использованием различных значений L_{max} . Тестовые объекты, наложенные программой SUPALM при различных максимальных числах гармоник L_{max} , использованных в расчетах амплитуд рассеяния. Объекты те же, что и в Таблице 7.1. Для тестов использовался угловой диапазон данных, соответствующий семи шенноновским каналам (N_{sh} =7).

N.	Шаблон	Налагаемая	MW	NSD/	NSD/	NSD/	NSD/	NSD/
	ный	ab initio	кДа	NCC	NCC	NCC	NCC	NCC
	объект	форма		(L _{max} =3)	(L _{max} =4)	$(L_{max}=5)$	(L _{max} =6)	(L _{max} =7)
1	2A38	z1z2	21	1.82/0.97	1.84/0.96	1.79/0.96	1.80/0.95	1.79/0.95
	(pdb)	(DAMMIF)						
2	1GOT	G protein	40	1.28/0.77	1.16/0.76	1.14/0.76	1.14/0.75	1.13/0.74
	(pdb)	(GASBOR)						
3	106V	Internalin	50	2.08/0.97	1.96/0.96	1.94/0.95	1.95/0.94	1.93/0.93
	(pdb)	(DAMMIN)						
4	4UIS	Emd-2974	140	4.10/0.95	3.89/0.94	3.75/0.93	3.75/0.92	3.74/0.91
	(pdb)	(EMD map)						
5	1HQK	Lumazine	960	1.27/0.88	1.26/0.87	1.24/0.87	1.25/0.87	1.24/0.86
	(pdb)	synthase						
		(DAMMIN)						
6	4V4W	70S	2150	1.30/0.96	1.28/0.95	1.27/0.94	1.28/0.94	1.26/0.93
	(pdb)	ribosome						
		(MONSA)						

Таблица 7.3 Результаты тестирования работы программы SUPALM с использованием различных угловых диапазонов (разным числом шенноновских каналов N_{sh}). Тестовые объекты, наложенные программой SUPALM с использованием различных угловых диапазонов данных МУРР для расчета амплитуд рассеяния. Объекты те же, что и в Таблице 7.1. Для тестов использовалось максимальное число сферических гармоник L_{max} =5.

N.	Шабло	Налагаемая	MW,	NSD/	NSD/	NSD/	NSD/	NSD/
	нный	ab initio	кДа	NCC,	NCC,	NCC,	NCC,	NCC,
	объект	форма		$(N_{sh}=5)$	(N _{sh} =6)	$(N_{sh}=7)$	$(N_{sh}=8)$	(N _{sh} =9)
1	2A38	z1z2	21	1.95/0.97	1.87/0.96	1.79/0.96	1.80/0.95	1.79/0.94
	(pdb)	(DAMMIF)						
2	1GOT	G protein	40	1.17/0.77	1.16/0.76	1.14/0.76	1.15/0.76	1.14/0.75
	(pdb)	(GASBOR)						
3	106V	Internalin	50	2.13/0.96	2.02/0.95	1.94/0.95	1.96/0.94	1.95/0.94
	(pdb)	(DAMMIN)						
4	4UIS	Emd-2974	140	5.33/0.95	3.91/0.94	3.75/0.93	3.74/0.93	3.73/0.92
	(pdb)	(EMD map)						
5	1HQK	Lumazine	960	1.28/0.88	1.25/0.87	1.24/0.87	1.24/0.87	1.23/0.86
	(pdb)	synthase						
		(DAMMIN)						
6	4V4W	70S	2150	1.30/0.97	1.29/0.9	1.27/0.94	1.28/0.94	1.27/0.93
	(pdb)	ribosome			5			
		(MONSA)						

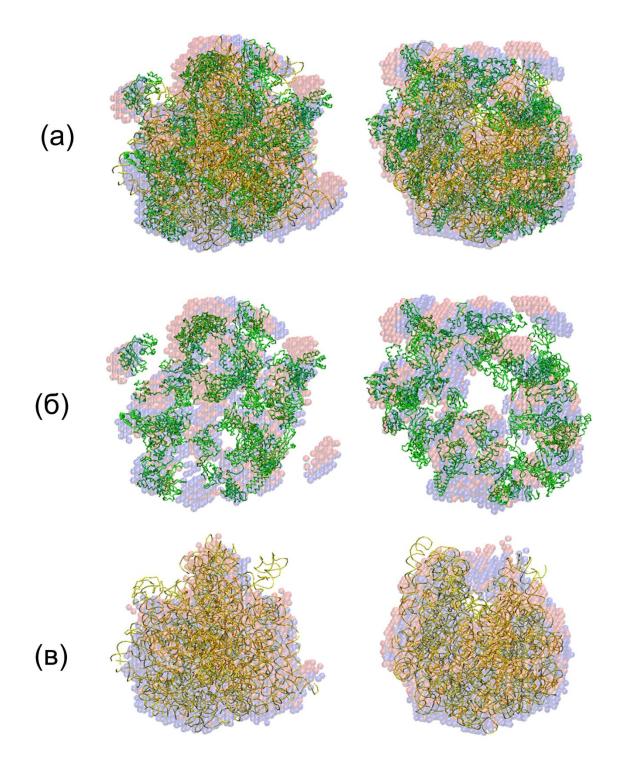


Рис. 7.6 Кристаллическая структура 70S рибосомы (PDB ID: 4V4W) (зеленые C_{α} цепи соответствуют белковым частям, части PHK показаны желтым цветом) с наложенными на нее двухфазными *ab initio* формами, наложение проводилось с помощью программ SUPALM (красные сферы) и SUPCOMB (синие сферы). Панель A содержит полные модели 70S рибосомы, панель B — только белковые части, панель C — только части PHK. Вид справа повернут против часовой стрелки на 90 вокруг вертикальной оси.

Результаты Таблицы 7.2 показывают, что во всех тестовых случаях максимальное число сферических гармоник L_{max} =5 было достаточным для качественного наложения двух объектов. При использовании меньшего числа гармоник качество совмещения становится значительно хуже (о чем можно судить по более высоким значениям NSD), в то время как более высокие значения L_{max} увеличивают время вычислений, но не улучшают качество совмещения.

Диапазон угловых данных, используемый для вычисления интегралов в уравнении (5), также может быть оптимально выбран. Как показано в Таблице 7.3, значения s_m , соответствующие семи каналам Шеннона N_{sh} (где $N_{sh}=D_{max}*s_m/\pi$ и D_{max} максимальный размер частицы), достаточны для надежной суперпозиции. Использование более широких угловых интервалов не приводит к значительному улучшению значений *NSD*, но, опять же, замедляет вычисления. На основании этих результатов значения по умолчанию для программы SUPALM выбраны $L_{max} = 5$ и $s_m = 7\pi/D_{max}$ (при необходимости ЭТИ значения МОГУТ быть изменены пользователем).

Программа SUPALM с этими параметрами, выбранными по умолчанию, работает примерно в 1.5÷2 раза быстрее, чем программа SUPCOMB для малых макромолекул с молекулярной массой менее 100 кДа (представленных примерно 10^3 атомами). Для больших макромолекулярных комплексов (например, 1 МДа, около 10^5 атомов) программа SUPALM работает более чем в десять раз быстрее по сравнению с программой SUPCOMB.

Следует отметить, что суперпозированные модели, полученные программами SUPCOMB и SUPALM, как и ожидалось, не идентичны, поскольку минимизируемые величины (*NSD* и 1/*NCC*, соответственно) различны. Тем не менее, совмещения моделей на Рис. 7.4-7.6 и сравнения значений *NSD* в Таблице 7.1 показывают, что, хотя решения, полученные программой SUPALM, имеют несколько более высокие значения *NSD*, их позиции в основном совпадают с теми, которые дает программа SUPCOMB. На Рис. 7.7 показан пример профилей контуров *NSD* и 1/*NCC* в окрестности решения программы SUPALM. На графиках

обеих функций хорошо видны минимумы, и обе ведут себя как аналитические функции, плавно приближающиеся к минимальным значениям. Таким образом, во всех рассмотренных тестовых примерах SUPALM дает практически те же результаты, что и SUPCOMB, но работает гораздо быстрее, особенно для больших макромолекулярных комплексов.

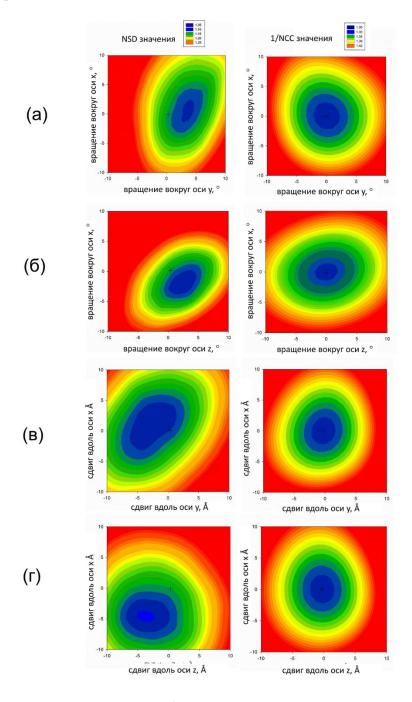


Рис. 7.7 Двумерные контурные графики значений NSD (левая колонка) и 1/NCC (правая колонка) в окрестности решения SUPALM для белка G (см. Рис. 7.4б). Решение программы SUPALM соответствует началу координат для всех графиков (отмечено крестом). На панелях (а) и (б) показаны контурные графики в

зависимости от вращения вокруг осей x/y и x/z, соответственно. На панелях (в) и (г) показаны контурные графики относительно сдвигов вдоль осей x/y и x/z, соответственно. "Истинное" решение программы SUPCOMB (минимум 2D контурного графика) расположено близко к решению программы SUPALM (RMSD = 3.3 Å).

Как правило, значения NSD около единицы указывают на хорошее сходство между двумя объектами, в то время как значительно большие значения NSD указывают на плохое сходство. Для NCC пороговые значения хорошего перекрытия зависят от размера молекулы, сложности формы и разрешения модели, используемой в расчетах. Из наших тестовых примеров следует, что нижний предел хорошего сходства моделей для NCC коэффициента составляет около $0.7 \div 0.8$ при используемом уровне разрешения (L_{max} =5 и 7-ми шенноновских каналов).

Программа Supalm входит в пакет программ ATSAS и успешно применяется в кроссплатформенном графическом приложении SASpy [239] (Рис. 7.8), позволяющем проводить моделирование с помощью гибридных моделей при анализе данных МУРР.

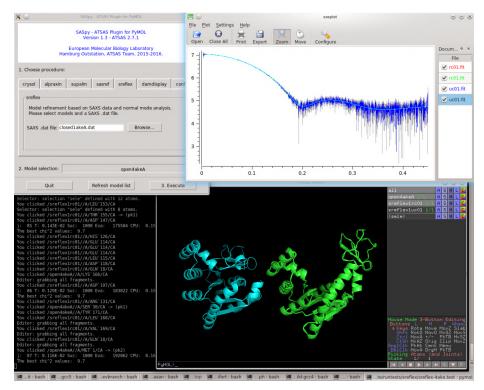


Рис. 7.8 Графический интерфейс программы SASPy для интерактивного локального уточнения четвертичной структуры комплексов макромолекул. Для

быстрого совмещения моделей в программе SASpy используется вызов модуля SUPALM.

Таким образом, нами был предложен и реализован в виде программы SUPALM быстрый алгоритм суперпозиции макромолекулярных моделей в обратном пространстве. Используя сферические гармоники для вычисления амплитуд, метод работает на порядок быстрее ранее реализованных алгоритмов в прямом пространстве. Важной особенностью программы SUPALM является возможность работы с картами электронной плотности (ЕМD файлы) и с многофазными *ab initio* моделями неоднородных частиц (например, комплексов белок-РНК и/или белок-ДНК).

7.4 Анализ структурного состава полидисперсных систем

Рассмотренные в разделах 7.1-7.3 программы интерактивного моделирования в основном работают с кривыми от одиночных экспериментальных данных МУРР, тогда как для больших наборов данных МУРР/МУРН (например, при сканировании по времени, концентрации или температуре) их обработка становится довольно утомительной задачей. Поэтому нами была разработана графическая платформа (реализованная в виде программы POLYSAS) [A76], которая позволяет в интерактивном режиме оперировать с большими наборами данных МУРР/МУРН и проводить комплексный анализ количественного состава многокомпонентных полидисперсных систем.

Программа POLYSAS имеет управляемый с помощью меню графический пользовательский интерфейс, вызывающий вычислительные модули из пакета ATSAS для выполнения обработки и анализа данных МУРР/МУРН. Графический интерфейс позволяет обрабатывать большие наборы данных (зависящие от времени, концентрации или температуры) и интерактивно изменять параметры моделирования данных с помощью графических слайдеров ("ползунков"). При этом для графического представления кривых рассеяния используется программа SASPLOT [A5].

Для последовательных наборов данных МУРР (состоящих из большого количества кривых рассеяния) программа POLYSAS позволяет в автоматическом режиме оценить общие структурные параметры, такие как радиус инерции (R_g) , максимальный размер частицы (D_{max}) , молекулярная масса (MM) и исключенный гидратированный объем частицы (V_{Porod}) . Для этих целей POLYSAS вызывает такие вычислительные модули, как DATRG, DATGNOM и DATPOROD [A7] из пакета ATSAS. Оцененные структурные параметры отображаются в виде покадровой зависимости с помощью программы SASPLOT в режиме реального времени и могут быть легко сохранены в формате ASCII или в виде картинки (скриншота). Список загруженных экспериментальных кривых МУРР/МУРН может быть оперативно изменен и переобработан с помощью кнопок графического меню.

В случае полидисперсных систем программа POLYSAS может предоставить распределение частиц по размерам. В зависимости от типа частиц оцениваются следующие распределения по размерам: 1) функции распределения по расстояниям внутри частицы p(r) для монодисперсных частиц произвольной формы; 2) функции объемного распределения по размерам $D_{V}(r)$ для сферических частиц; 3) функции распределения размера поперечного сечения для стержнеобразных частиц; 4) функции распределения размера длины для цилиндрических частиц. Это осуществляется путем вызова модуля GNOM из пакета ATSAS. Эффекты замазывания данных МУРР/МУРН для различных геометрических параметров установок могут быть учтены путем введения соответствующих коллимационных поправок. Графическое меню POLYSAS позволяет интерактивно изменять и/или структурные параметры частицы геометрические параметры экспериментальной установки и отслеживать, как это влияет на качество приближения (фита) и рассчитанной функции распределения по размерам.

В программе POLYSAS существует несколько возможностей для анализа смесей. В случае, когда доступны модели высокого разрешения компонентов смеси (например, в формате PDB), программа POLYSAS позволяет приближать наборы экспериментальных данных МУРР/МУРН с помощью линейных комбинаций

кривых рассеяния, рассчитанных по выбранным PDB моделям. Объемные доли компонентов рассчитываются с помощью модуля OLIGOMER из пакета ATSAS, а результаты отображаются с помощью программы SASPLOT.

POLYSAS Программа нелинейного имеет возможность анализа многокомпонентных систем с использованием простых геометрических тел трех типов (сферы, цилиндры и гантели) и заданием до пяти компонентов. При моделировании можно учитывать эффекты полидисперсности и межчастичные взаимодействия (последние могут быть применены для сферических частиц в рамках приближения потенциала притягивающихся жестких сфер). Уточненные параметры рассчитываются с помощью вызова модуля MIXTURE из пакета ATSAS. Параметры моделей можно интерактивно изменять с помощью слайдеров-"ползунков" графического меню, при этом рассчитанные приближения модели к экспериментальным данным будут обновляться в режиме реального времени. Верхние и нижние пределы (ограничения) параметров для каждого компонента смеси могут быть установлены перед началом процесса минимизации.

Помимо этого, в программе POLYSAS доступен модельно-независимый анализ числа компонентов смеси с использованием сингулярного разложения (SVD) набора данных МУРР/МУРН. Он может применяться для систем, измеренных при различных физико-химических условиях среды (рН, температура, состав буфера, концентрация образца, концентрация соли (осадителя) в растворе, добавление лиганда и т.д.). В этом случае вычисляются собственные значения и собственные векторы матриц сингулярного разложения и оценивается количество независимых компонентов в системе с помощью статистического теста на неслучайные осцилляции сингулярных векторов. Для этого вызывается модуль SVDPLOT из пакета ATSAS. Графическое меню программы POLYSAS предоставляет опции для интерактивного выбора полных наборов экспериментальных данных МУРР/МУРН или его частичных подмножеств и позволяет проверить стабильность оцененного числа компонентов в системе.

Для иллюстрации применения программы POLYSAS приведем несколько примеров экспериментальных наборов МУРР, полученных на синхротронной станции X33 [A4] накопительного кольца DORIS-III (DESY, Гамбург).

Первый пример касается изучения сложной равновесной олигомерной смеси АХН-домена белка атаксин-1 [A57]. Атаксин-1 - это человеческий белок, ответственный за атаксию типа 1, наследственное заболевание, связанное с агрегацией и неправильным сворачиванием белков. АХН-домен атаксина-1 содержит димеры в асимметричной (независимой) части кристаллографической ячейки. В растворе АХН-домен присутствует в виде сложной равновесной смеси мономерных, димерных и тетрамерных состояний. Такое поведение, наряду со способностью АХН-домена образовывать множественные димеризационные интерфейсы, позволяет признать АХН-домен необычным примером белка-хамелеона, свойства которого имеют потенциальное значение для агрегационных свойств атаксина-1 и, следовательно, для протекания заболевания.

На Рис. 7.9 представлен графический интерфейс пользователя программы POLYSAS для моделирования АХН-домена Атаксина-1 в виде олигомерной смеси. График с теоретическими кривыми рассеяния, рассчитанными по известным кристаллическим моделям (мономеры, димеры, тетрамеры и октамеры), показан в правом верхнем углу, график с приближениями к набору экспериментальных данных МУРР показан в правом нижнем углу, а график с оценкой объемных долей компонентов показан в левом нижнем углу. Использование программы POLYSAS позволяет количественно оценить олигомерный состав смеси и дает возможность интерактивного мониторинга различных гипотез в процессе моделирования данных.

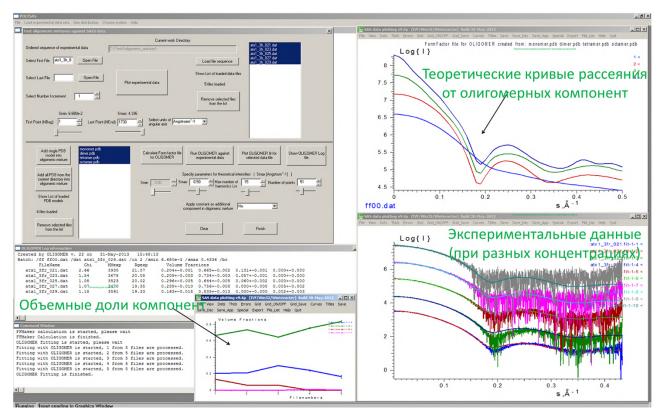


Рис. 7.9 Графический интерфейс программы POLYSAS при анализе количественного состава олигомерной смеси АХН домена белка Ataxin-1.

Второй пример относится к зависящим от времени структурам гидролитически деградирующих полимерных мицелл, предназначенных для контролируемой доставки лекарств [А77]. Высокогидрофобные холестериновые молекулы, а также противораковый препарат доксорубицин были присоединены к полимерной матрице с помощью чувствительной к изменениям рН гидразонной связи. Эксперименты МУРР в зависимости от времени проводились после внезапного изменения рН от типичного значения (рН=7.2) в крови до значения, характерного при наличии опухолевых клеток (рН=5.0). Целью исследования являлось изучение процесса высвобождения лекарства и изменения размера и формы частиц. Было обнаружено, наночастицы, состоящие что ИЗ конъюгатов, содержащих доксорубицин, в целом оказывались больше, чем наночастицы в отсутствие лекарства. Для большинства конъюгатов рост или распад наночастиц наблюдался в течение нескольких часов. Скорость роста/распада и стабильный размер наночастиц зависели от структуры гидразонной связи.

На Рис. 7.10 представлен графический интерфейс пользователя программы POLYSAS, которая применялась для моделирования зависимости от времени общих параметров (радиуса инерции R_g и максимального размера частиц D_{max}) полимерных мицелл с противораковым препаратом. График с набором экспериментальных данных, зависящих от времени, отображается в правом верхнем углу, зависимость R_g и D_{max} от времени - в правом нижнем углу. Эта зависимость была получена автоматически при нажатии на кнопку графического интерфейса программы POLYSAS.

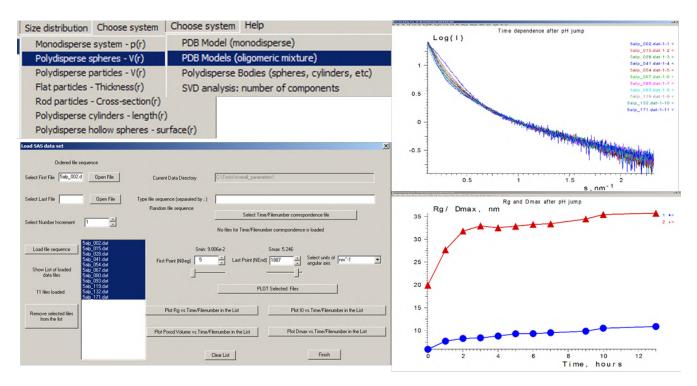


Рис. 7.10 Графический интерфейс программы POLYSAS при моделировании зависимости от времени структурных параметров полимерных мицелл с противораковым препаратом. Вверху справа показана временная зависимость профилей рассеяния, внизу справа — найденная временная зависимость радиуса инерции R_g и максимального размера частиц D_{max} . Меню программы POLYSAS позволяет работать с большими наборами данных и автоматически находить структурные параметры исследуемой системы.

Таким образом, программа POLYSAS предоставляет расширенные и удобные возможности для структурных исследований многокомпонентных растворов

макромолекул с использованием данных МУРР/МУРН. Как и рассмотренные выше программы, программа POLYSAS входит в пакет ATSAS.

7.5 Оценка структурных параметров частично упорядоченных систем

В случае частично упорядоченных систем кривые малоуглового рассеяния содержат уширенные дифракционные пики, по которым можно определить такие структурные параметры, как межплоскостное расстояние и средний размер кристаллитов.

Для анализа таких систем нами была разработана графическая программа Peak [A5]. На первом шаге проводится интерактивный выбор одного или несколько пиков на кривой рассеяния, затем выполняется приближение выбранной области пика с помощью гауссовой функции и двухпараметрического вычитания фона. К оцениваемым структурным параметрам относятся: межплоскостное расстояние, размер области кристалличности и степень разупорядочения, рассчитанные по положению пика и его ширине с использованием стандартных уравнений дифракционной теории [240].

Структурные параметры, которые можно оценить в программе Peak, определяются следующими выражениями:

- межплоскостное расстояние $d=2\pi/s_{\text{max}}$;
- размер дальнего порядка $L=\lambda/\beta_s$ cos $\theta_{\rm max}$
- радиус взаимодействия $R_m = (\pi/2.5)^2 \cdot (\lambda/\beta_s)$
- степень разупорядочения Δ/d = $(1/\pi)(\beta_s*d/\lambda)^{1/2}$

где λ - длина волны рентгеновского (нейтронного) излучения, β_s - полная ширина на половине максимума пика (в радианах), θ_{max} - угол рассеяния, соответствующий значению вектора рассеяния s_{max} , для которого наблюдается максимум пика.

Ниже приведен интерфейс программы PEAK при анализе данных МУРР от полимерной матрицы поли(октадецилсилоксана) с внедренными в нее металлическими наночастицами [A78] (Рис. 7.11).

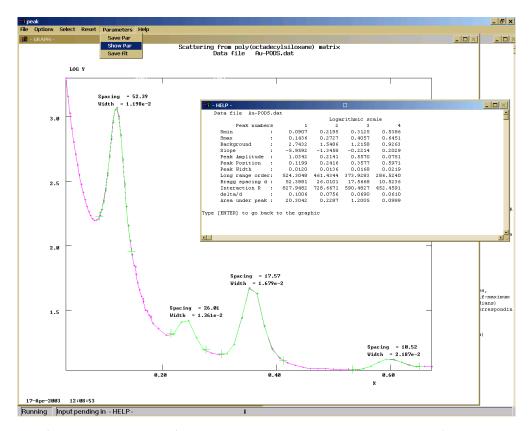


Рис. 7.11 Графический интерфейс программы PEAK [A5], отображающий кривую рассеяния МУРР от полимерной матрицы поли(октадецилсилоксана) с внедренными металлическими наночастицами [A78]. Экспериментальная кривая показана малиновым цветом, зеленые кривые - это приближенные гауссовскими функциями профили пиков. Найденные структурные параметры выведены во всплывающем окне.

Однако поскольку программа Peak (в первоначальной версии) использовала графическую библиотеку QuickWin, то она могла работать только на Windows-платформе, что в определенной мере ограничивало ее использование. Поэтому нами была также создана новая кроссплатформенная версия программы PeakQt [A75], которая использует графическую библиотеку Qt, а также улучшенную процедуру оптимизации структурных параметров, описывающих дифракционные пики.

На Рис. 7.12 показан графический интерфейс новой версии программы PeakQt с данными МУРР от композитного твердого электролита на основе блок сополимера Плюроник PL-L64 и SiO2/Al [A79].

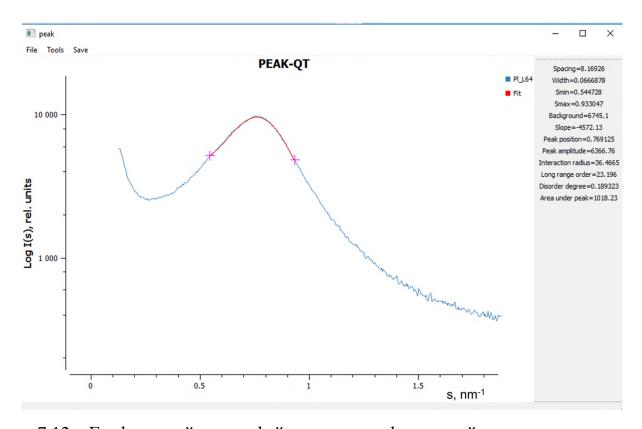


Рис. 7.12. Графический интерфейс кроссплатформенной версии программы PEAKQt [A75]. Синим цветом обозначены экспериментальные данные МУРР (композитный твердый электролит на основе блок сополимера Плюроник PL-L64 и SiO2/Al [A79]), красным цветом — приближение программы PEAKQt. Крестики показывают выбранную область данных для моделирования. Справа на панели указаны вычисленные структурные параметры (межлоскостное расстояние (spacing), размер кристаллита (long range order), степень разупорядочения структуры (disorder degree) и др.).

Помимо улучшенной графики и возможности масштабирования графика с помощью компьютерной мыши, программа получила более удобный пользовательский интерфейс, когда результаты моделирования мгновенно отображаются непосредственно на рабочей панели программы. Кроме того, в программе для оптимизации структурных параметров используется более современный **MINPACK** минимизационный пакет (https://www.math.utah.edu/software/minpack.html), что также повысило точность получаемых результатов.

Таким образом, программа Peak (и ее кроссплатформенная версия PeakQt) позволяют проводить интерактивное моделирование дифракционных пиков на экспериментальных кривых малоуглового рассеяния от частично упорядоченных систем и извлекать ценную структурную информацию о расстоянии между повторяющимися мотивами структуры и размерах области кристаллитов в образце.

Заключение к Главе 7

В Главе 7 было детально представлено разработанное автором программное обеспечение, позволяющее проводить интерактивную обработку и моделирование малоуглового рассеяния. Рассмотренные графические программы охватывают широкий спектр задач от первичной обработки данных МУРР/МУРН до моделирования трехмерных структур макромолекулярных комплексов и количественного анализа сложных многокомпонентных полидисперсных систем с наборов использованием больших экспериментальных данных. Данные программы, наряду с рассмотренными в предыдущих Главах, применялись для изучения по данным МУРР/МУРН структурных особенностей широкого круга биологических систем (растворов белков, нуклеиновых кислот, липидов) [А80-А105], а также анализа полимерных соединений [А106-А111], металлических наночастиц [А112-114], материалов, используемых в тканевой инженерии [А115-117] и др. Представленные интерактивные методы обработки и моделирования данных хорошо дополняют автоматизированные алгоритмы, рассмотренные ранее в Главах 2-6, и позволяют повысить надежность получаемых результатов.

Выводы

- 1. Разработаны подходы для количественного анализа данных малоуглового рассеяния многокомпонентными смесями белковых и липидных частиц с помощью параметрических методов нелинейной минимизации. Они включают в себя оценку объемных долей компонентов, поиск функции распределения частиц по размерам взаимодействия, параметров межчастичного восстановление электронной плотности липидного бислоя и оценку параметров мультислоевой организации липидных везикул, а также их взаимодействия с белковыми макромолекулами. Разработанные программы могут быть применены для широкого круга биологических, органических и композитных систем. Метод успешно апробирован на ряде белковых и липидных смесей, содержащих до 5 компонентов. Впервые удалось выявить случайно-столкновитель-ный механизм образования комплекса адренодоксина (Adx) с митохондриальным цитохромом с (Сс); показать способность фермента люмазинсинтазы самособираться в капсулы размером от 16 до 32 нм, причем их количественный состав зависит как от наличия мутаций белка. так И изменений физико-химических условий проанализировать эволюцию структурных параметров везикул DMPC/DPPC в процессе экструзии, позволяющую получать однослойные везикулы с радиусом (30±1) нм; для смесей, состоящих из вирусных липидов и матричного белка M1 вируса гриппа А, оценить влияние гемагглютинина на формирование липопротеиновых нанодоменов в липидном бислое.
- 2. Представлен и реализован в виде программы новый подход по прямому восстановлению трехмерной формы неизвестного промежуточного состояния в эволюционирующей трехкомпонентной системе в случае, когда начальное и конечное состояние системы известно или может быть аппроксимировано теоретическими кривыми рассеяния. Он представляет собой комбинацию *ab initio* алгоритма определения 3D формы частицы и решения системы линейных уравнений методом наименьших квадратов для нахождения объемных долей компонентов смеси (с точностью до 1-5%) в каждой временной точке

динамического процесса. Разработанный подход был успешно применен для изучения процесса образования фибрилл белка инсулина и позволил определить форму промежуточного компонента в процессе их роста, который представляет собой вытянутую структуру размером 20 нм, состоящую из 5 связанных доменов по размерам близких к мономерам инсулина, что подтверждает гипотезу об олигомерном ядре, являющемся первичной удлиняющей единицей инсулиновых амилоидных фибрилл. С помощью предложенного метода можно также изучать двухкомпонентные равновесные олигомерные смеси, когда известно только начальное состояние системы. Для белка фактора роста нервов была восстановлена трехмерная форма димера димеров по концентрационной серии данных малоуглового рассеяния.

3. Для белковых смесей с перекрывающимися пиками выхода (элюирования) компонентов из гель-хроматографической колонки при ее сочетании разработан малоугловыми измерениями автоматизированный метод восстановления индивидуальных профилей рассеяния компонентов на основе алгоритма эволюционного факторного анализа. Проведено систематическое исследование чувствительности метода в зависимости от количества компонентов (при наличии от 2 до 4 компонентов), уровня шума в данных, степени асимметрии и соотношения амплитуд пиков концентрационных профилей компонентов (в пределах от 1:5 до 5:1). Показано, что для двухкомпонентных систем с симметричными профилями концентрации метод эффективно работает даже при наличии значительного уровня шума в данных (при регистрации 10^2 фотонов вблизи заслонки первичного пучка). Подтверждено, что метод применим и для наборов малоугловых данных в сочетании с ионно-обменной хроматографией при умеренной степени (<12%) градиента солевого раствора. Метод успешно протестирован на ряде экспериментальных данных, в частности определено, что для смеси белка пируват-альдолазы ее олигомерными компонентами в растворе являются гексамеры И октамеры. Восстановленные профили рассеяния

компонентов хорошо соотносятся с известными кристаллографическими моделями белков.

- 4. Разработан новый подход для апостериорного определения полезного углового диапазона в экспериментальных данных малоуглового рассеяния от разбавленной монодисперсной системы. Предложенная объективная автоматизированная процедура основывается на использовании информационной теоремы Котельникова-Шеннона и учитывает, как отношение сигнал/шум, так и шаг угловой сетки малоугловых данных. Метод был успешно применен для данных МУРР/МУРН от ряда монодисперсных растворов макромолекул, а также от стабилизированных узкодисперсных наночастиц золота со средним размером 8 нм.
- 5. Предложен новый быстрый алгоритм суперпозиции макромолекулярных моделей с использованием разложения на сферические гармоники и критерием корреляции в обратном пространстве, позволивший более, чем в 10 раз ускорить расчеты по сравнению с имеющимися алгоритмами. Для надежной работы алгоритма требуется использовать не менее 5 сферических гармоник и угловой диапазон данных, включающий не менее 7 шенноновских каналов.
- 6. Разработан набор графических приложений, позволяющих проводить интерактивную обработку и моделирование малоугловых данных от изотропных систем. Они охватывают широкий спектр задач от первичной обработки данных МУРР/МУРН до моделирования трехмерных структур макромолекулярных комплексов и количественного анализа сложных многокомпонентных полидисперсных систем.

Приложение

Программное обеспечение для анализа данных малоуглового рассеяния П 1.1. Пакет для количественной оценки состава многокомпонентных полидисперсных систем по данным малоуглового рассеяния

1) Программа OLIGOMER поиска объемных долей компонентов смеси частиц при известных профилях рассеяния компонентов с помощью метода линейных наименьших квадратов с ограничением на неотрицательность решений.

<u>Вклад автора:</u> реализация автоматизированного командного режима работы программы для обработки больших наборов экспериментальных данных; автоматическая оценка всевозможных комбинаций формфакторов компонентов и последующая сортировка данных по критерию качества приближения; добавление ограничений на соотношения объемных долей компонентов в случае диссоциации белковых комплексов или олигомеров.

2) Программа MIXTURE поиска функций распределения частиц по размерам для полидисперсных многокомпонентных систем в случае аппроксимации формы компонентов простыми геометрическими телами с помощью нелинейных методов минимизации из пакета OPTIS.

<u>Вклад автора:</u> модернизация программы для учета полидисперсности в случае икосаэдрических полых капсул и эллипсоидов вращения; расчет структурного фактора в приближении Перкуса-Йовика для потенциала притягивающихся жестких сфер.

- 3) Программа LIPMIX для проведения количественной оценки состава и структуры растворов липидных смесей с помощью нелинейных методов минимизации из пакета OPTIS. LIPMIX позволяет восстановить функцию распределения сферических липидных везикул по размерам, электронную плотность липидного бислоя внутри везикулы и оценить мультислойную ламеллярную организацию везикул.
- 4) Программа BILMIX для анализа смесей однослойных везикул как сферической, так и эллипсоидной формы, с помощью нелинейных методов минимизации из

пакета OPTIS. BILMIX позволяет моделировать асимметричную электронную плотность липидного бислоя, которая может возникать при взаимодействия липосомы с белком или другим связывающим партнером.

П 1.2. Пакет для ab initio восстановления формы частицы неизвестного промежуточного состояния в эволюционирующей трехкомпонентной системе по данным малоуглового рассеяния

Программа DAMMIX предназначена для прямого восстановления формы частицы неизвестного промежуточного состояния в эволюционирующей трехкомпонентной системе наряду с объемными долями компонентов во всех состояниях (временных точках динамического процесса) системы. DAMMIX также применим для случая эволюционирующих двухкомпонентных систем олигомерных смесей с одним неизвестным компонентом.

П 1.3. Пакет для определения индивидуальных профилей рассеяния компонентов в белковых смесях по данным малоуглового рассеяния в сочетании с гель-хроматографией

Программа EFAMIX позволяет оценивать количество компонентов и восстанавливать индивидуальные профили рассеяния компонентов в белковых смесях с помощью хемометрического метода эволюционного факторного анализа (ЭФА) для набора данных малоуглового рентгеновского рассеяния, полученного с использованием гель-хроматографической колонки.

П 1.3. Пакет для интерактивного моделирования малоугловых данных с помощью графических приложений

- 1) Программа PRIMUS включает в себя основные этапы анализа малоугловых данных ДЛЯ изотропных систем, начиная c первичной обработки экспериментальных данных и заканчивая моделированием данных. PRIMUS имеет графические модули для оценки основных структурных параметров частиц и формы моделирования помощью аппроксимации простыми частиц геометрическими телами.
- 2) Программа MASSHA имеет интерактивный и автоматический режимы моделирования данных малоуглового рассеяния методом молекулярной тектоники, позволяет проводить автоматическое выравнивание структур, учитывать симметрию при моделировании мульти-субъединичных комплексов.
- 3) Программа POLYSAS позволяет в интерактивном режиме оперировать с большими наборами данных МУРР/МУРН и проводить комплексный анализ количественного состава многокомпонентных полидисперсных систем. POLYSAS содержит графические меню для нахождения функций распределения частиц по размерам и оценки числа компонентов в смесях.
- 4) Программа SUPALM для быстрой суперпозиции макромолекулярных моделей с использованием разложения на сферические гармоники и критерия корреляции в обратном пространстве. SUPALM используется для интерактивного манипулирования трехмерных структур с использованием данных малоуглового рассеяния в программе SASPY (устанавливается как плагин к графической программе PyMol).
- 5) Программа PEAK позволяет проводить интерактивное моделирование дифракционных пиков на экспериментальных кривых малоуглового рассеяния от частично упорядоченных систем и оценивать расстояние между повторяющимися мотивами структуры и размеры области кристаллитов.

Все программы, указанные в Приложении, входят в пакет ATSAS (находящийся в свободном доступе для академических пользователей на сайте: https://www.embl-hamburg.de/biosaxs/software.html) и были разработаны при участии Д.И. Свергуна, В.В. Волкова и М.В. Петухова в рамках многолетнего и плодотворного сотрудничества.

Публикации по теме диссертации

Статьи:

- A1. Roessle M. W., Klaering R., Ristau U., Robrahn B., Jahn D., Gehrmann T., Konarev P., Round A., Fiedler S., Hermes C., Svergun D. Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg // J. Appl. Cryst. 2007. V. 40. P. s190-s194.
- A2. Blanchet C. E., Zozulya A. V., Kikhney A. G., Franke D., Konarev P. V., Shang W., Klaering R., Robrahn B., Hermes C., Cipriani F., Svergun D. I., Roessle M. Instrumental setup for high-throughput small- and wide-angle solution scattering at the X33 beamline of EMBL Hamburg // J. Appl. Cryst. − 2012. − V. 45, № 3. − P. 489-495.
- A3. Peters G. S., Zakharchenko O. A., Konarev P. V., Karmazikov Y. V., Smirnov M. A., Zabelin A. V., Mukhamedzhanov E. H., Veligzhanin A. A., Blagov A. E., Kovalchuk M. V. The small-angle X-ray scattering beamline BioMUR at the Kurchatov synchrotron radiation source // Nuclear Inst. and Methods in Physics Research, A. 2019. V. 945. P. 162616.
- A4. Peters G. S., Gaponov Y. A., Konarev P. V., Marchenkova M. A., Ilina K. B., Volkov V. V., Pisarevsky Y. V., Kovalchuk M. V. Upgrade of the BioMUR beamline at the Kurchatov synchrotron radiation source for serial small-angle X-ray scattering experiments in solutions // Nuclear Inst. and Methods in Physics Research, A. 2022. V. 1025. P. 166170.
- A5. Konarev P. V., Volkov V. V., Sokolova A. V., Koch M. H. J., Svergun D. I. PRIMUS a Windows-PC based system for small-angle scattering data analysis // J. Appl. Cryst. 2003. V. 36. P. 1277-1282.
- A6. Konarev P. V., Petoukhov M. V., Volkov V. V., Svergun D. I. ATSAS 2.1, a program package for small-angle scattering data analysis // J. Appl. Cryst.— 2006. V. 39. P. 277-286.
- A7. Petoukhov M. V., Konarev P. V., Kikhney A. G., Svergun D. I. ATSAS 2.1 towards automated and web-supported small-angle scattering data analysis // J. Appl. Cryst. -2007. -V. 40, N_{2} s1. -P. s223-s228.

- A8. Petoukhov M. V., Franke D., Shkumatov A. V., Tria G., Kikhney A. G., Gajda M., Gorba C., Mertens H. D. T., Konarev P. V., Svergun D. I. New developments in the ATSAS program package for small-angle scattering data analysis // J. Appl. Cryst. 2012. V. 45, No. 2. P. 342-350.
- A9. Franke D., Petoukhov M. V., Konarev P. V., Panjkovich A., Tuukkanen A., Mertens H. D. T., Kikhney A. G., Hajizadeh N. R., Franklin J. M., Jeffries C. M., Svergun D. I. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions // J. Appl. Cryst. 2017. V. 50, № Pt 4. P. 1212-1225.
- A10. Manalastas-Cantos K., Konarev P. V., Hajizadeh N. R., Kikhney A. G., Petoukhov M. V., Molodenskiy D. S., Panjkovich A., Mertens H. D. T., Gruzinov A., Borges C., Jeffries C. M., Svergun D. I., Franke D. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis // J. Appl. Cryst. − 2021. − V. 54, № Pt 1. − P. 343-355.
- A11. Sjekloca L., Pudas R., Sjoblom B., Konarev P., Carugo O., Rybin V., Kiema T. R., Svergun D., Ylanne J., Djinovic Carugo K. Crystal structure of human filamin C domain 23 and small angle scattering model for filamin C 23-24 dimer // J Mol Biol. − 2007. − V. 368, № 4. − P. 1011-1023.
- A12. Konarev P. V., Petoukhov M. V., Svergun D. I. Rapid automated superposition of shapes and macromolecular models using spherical harmonics // J. Appl. Cryst.—2016. V. 49, № Pt 3. P. 953-960.
- A13. Zhang X., Konarev P. V., Petoukhov M. V., Svergun D. I., Xing L., Cheng R. H., Haase I., Fischer M., Bacher A., Ladenstein R., Meining W. Multiple assembly states of lumazine synthase: a model relating catalytic function and molecular assembly // J Mol Biol. -2006. -V. 362, N 4. -P. 753-770.
- A14. Heikkinen O. K., Ruskamo S., Konarev P. V., Svergun D. I., Iivanainen T., Heikkinen S. M., Permi P., Koskela H., Kilpelainen I., Ylanne J. Atomic structures of two novel immunoglobulin-like domain pairs in the actin cross-linking protein filamin // J Biol Chem. -2009. -V. 284, N 37. -P. 25450-25458.

- A15. Wong E., Vaaje-Kolstad G., Ghosh A., Hurtado-Guerrero R., Konarev P. V., Ibrahim A. F., Svergun D. I., Eijsink V. G., Chatterjee N. S., van Aalten D. M. The Vibrio cholerae colonization factor GbpA possesses a modular structure that governs binding to different host surfaces // PLoS Pathog. − 2012. − V. 8, № 1. − P. e1002373.
- A16. Zanchi D., Konarev P. V., Tribet C., Baron A., Svergun D. I., Guyot S. Rigidity, conformation, and solvation of native and oxidized tannin macromolecules in water-ethanol solution // J Chem Phys. − 2009. − V. 130, № 24. − P. 245103-245116.
- A17. Konarev P. V., Petoukhov M. V., Svergun D. I. MASSHA a graphic system for rigid body modelling of macromolecular complexes against solution scattering data // J. Appl. Cryst. 2001. V. 34. P. 527-532.
- A18. Prischi F., Konarev P. V., Iannuzzi C., Pastore C., Adinolfi S., Martin S. R., Svergun D. I., Pastore A. Structural bases for the interaction of frataxin with the central components of iron-sulphur cluster assembly // Nat Commun. 2010. V. 1 P. 95-105. A19. Sharapov M. G., Novoselov V. I., Samygina V. R., Konarev P. V., Molochkov A. V., Sekirin A. B., Balkanov A. S., Gudkov S. V. A chimeric recombinant protein with peroxidase and superoxide dismutase activities: Physico-chemical characterization and applicability to neutralize oxidative stress caused by ionizing radiation // Biochemical Engineering Journal –2020. V. 159. P. 107603.
- A20. Paoletti F., Konarev P. V., Covaceuszach S., Schwarz E., Cattaneo A., Lamba D., Svergun D. I. Structural and functional properties of mouse proNGF // Biochem Soc Trans. 2006. V. 34, № Pt 4. P. 605-606.
- A21. Paoletti F., Covaceuszach S., Konarev P. V., Gonfloni S., Malerba F., Schwarz E., Svergun D. I., Cattaneo A., Lamba D. Intrinsic structural disorder of mouse proNGF // Proteins. 2009. V. 75, № 4. P. 990-1009.
- A22. Covaceuszach S., Cassetta A., Konarev P. V., Gonfloni S., Rudolph R., Svergun D. I., Lamba D., Cattaneo A. Dissecting NGF interactions with TrkA and p75 receptors by structural and functional studies of an anti-NGF neutralizing antibody // J Mol Biol. 2008. V. 381, No. 4. P. 881-896.

- A23. Paoletti F., Malerba F., Konarev P. V., Visintin M., Scardigli R., Fasulo L., Lamba D., Svergun D. I., Cattaneo A. Direct intracellular selection and biochemical characterization of a recombinant anti-proNGF single chain antibody fragment // Arch Biochem Biophys. -2012.-V.522, No. 1.-P.26-36.
- A24. Covaceuszach S., Konarev P. V., Cassetta A., Paoletti F., Svergun D. I., Lamba D., Cattaneo A. The conundrum of the high-affinity NGF binding site formation unveiled? // Biophys J. $-2015.-V.\ 108,\ No.\ 3.-P.\ 687-697.$
- A25. Yan R., Yalinca H., Paoletti F., Gobbo F., Marchetti L., Kuzmanic A., Lamba D., Gervasio F. L., Konarev P. V., Cattaneo A., Pastore A. The Structure of the Pro-domain of Mouse proNGF in Contact with the NGF Domain // Structure. − 2019. − V. 27, № 1. − P. 78-89.
- A26. Covaceuszach S., Peche L. Y., Konarev P. V., Lamba D. A combined evolutionary and structural approach to disclose the primary structural determinants essential for proneurotrophins biological functions // Comput Struct Biotechnol J. 2021. V. 19. P. 2891-2904.
- A27. Covaceuszach S., Peche L. Y., Konarev P. V., Grdadolnik J., Cattaneo A., Lamba D. Untangling the Conformational Plasticity of V66M Human proBDNF Polymorphism as a Modifier of Psychiatric Disorder Susceptibility // Int J Mol Sci. − 2022. − V. 23, № 12. − P. 6596.
- A28. Paoletti F., Covaceuszach S., Cassetta A., Calabrese A. N., Novak U., Konarev P., Grdadolnik J., Lamba D., Golic Grdadolnik S. Distinct conformational changes occur within the intrinsically unstructured pro-domain of pro-Nerve Growth Factor in the presence of ATP and Mg(2) // Protein Sci. -2023. V. 32, No 2. P. e4563.
- A29. Sicorello A., Rozycki B., Konarev P. V., Svergun D. I., Pastore A. Capturing the Conformational Ensemble of the Mixed Folded Polyglutamine Protein Ataxin-3 // Structure. -2021.-V.29, No. 1.-P.70-81.
- A30. Falces J., Arregi I., Konarev P. V., Urbaneja M. A., Svergun D. I., Taneva S. G., Banuelos S. Recognition of nucleoplasmin by its nuclear transport receptor importin

- alpha/beta: insights into a complete import complex // Biochemistry. -2010. V.49, $N_{\odot}45. P.9756-9769$.
- A31. Shiozawa K., Konarev P. V., Neufeld C., Wilmanns M., Svergun D. I. Solution structure of human Pex5.Pex14.PTS1 protein complexes obtained by small angle X-ray scattering // J Biol Chem. − 2009. − V. 284, № 37. − P. 25334-25342.
- A32. Konarev P. V., Svergun D. I. Limitations of the iterative electron density reconstruction algorithm from solution scattering data // Nat Methods. -2021. V. 18, No. 3. P. 244-245.
- A33. Konarev P. V., Svergun D. I. A posteriori determination of the useful data range for small-angle scattering experiments on dilute monodisperse systems // IUCrJ. -2015. -V. 2, N Pt 3. -P. 352-360.
- A34. Taneva S. G., Banuelos S., Falces J., Arregi I., Muga A., Konarev P. V., Svergun D. I., Velazquez-Campoy A., Urbaneja M. A. A mechanism for histone chaperoning activity of nucleoplasmin: thermodynamic and structural models // J Mol Biol. − 2009. − V. 393, № 2. − P. 448-463.
- A35. Konarev P. V., Kachalova G. S., Ryazanova A. Y., Kubareva E. A., Karyagina A. S., Bartunik H. D., Svergun D. I. Flexibility of the linker between the domains of DNA methyltransferase SsoII revealed by small-angle X-ray scattering: implications for transcription regulation in SsoII restriction-modification system // PLoS One. 2014. V. 9, N 4. P. e93453.
- A36. Svergun D. I., Konarev P. V., Volkov V. V., Koch M. H. J., Sager W. F. C., Smeets J., Blokhuis E. M. A small angle X-ray scattering study of the droplet-cylinder transition in AOT microemulsions // J. Chem. Phys. 2000. V. 113. P. 1651-1665.
- A37. Kryukova A. E., Konarev, P. V., Volkov, V. V. Evaluation of solution stability for two-component polydisperse systems by small-angle scattering // J Phys Conf Ser. 2017. V. 941. P. 012069.
- АЗ8. Крюкова А. Е., Конарев, П. В., Волков, В. В. Исследование устойчивости решений при анализе полидисперсных систем методом малоуглового рассеяния // Кристаллография. 2018. Т. 63, № 1. С. 32-37.

- А39. Крюкова А. Е., Козлова, А. С., Конарев, П. В., Волков, В. В., Асадчиков, В. Е. Оценка устойчивости решений при восстановлении функции распределения частиц в растворе кремнезоля по размерам по данным малоуглового рассеяния // Кристаллография. − 2018. − Т. 63, № 4. − С. 524-529.
- A40. Kryukova A. E., Konarev, P. V., Volkov, V. V., Asadchikov, V. E. Restoring silicasol structural parameters using gradient and simulation annealing optimization schemes from small-angle X-ray scattering data // Journal of Molecular Liquids. 2019. V. 283. P. 221-224.
- А41. Крюкова А. Е., Конарев, П. В., Волков, В. В., Асадчиков, В. Е. Комбинирование методов минимизации в целях повышения эффективности восстановления распределений частиц по размерам для раствора кремнезоля по данным малоуглового рассеяния // Ядерная физика и инжиниринг. − 2019. − Т. 10, № 4. − С. 393-398.
- А42. Крюкова А. Е., Конарев, П.В., Волков, В.В. Поиск эффективного алгоритма для восстановления решений при анализе данных малоуглового рассеяния от раствора кремнезоля // Кристаллография. 2021. Т. 66, № 5. С. 730-737.
- А43. Волков В. В., Конарев, П. В., Крюкова, А. Е. Комбинированная схема восстановления функции распределения частиц по размерам с использованием данных малоуглового рассеяния // Письма в ЖЭТФ. 2020. Т. 112, № 9. С. 632-636.
- A44. Volkov V. V., Kryukova A.E., Konarev P. V. Approaches for improving the quality of particle size distribution reconstructions from small-angle scattering data // J Phys Conf Ser. 2020. V. 1686. P. 012059.
- A45. Varga A., Flachner B., Konarev P., Graczer E., Szabo J., Svergun D., Zavodszky P., Vas M. Substrate-induced double sided H-bond network as a means of domain closure in 3-phosphoglycerate kinase // FEBS Lett. − 2006. − V. 580, № 11. − P. 2698-2706.
- A46. Szabo J., Varga A., Flachner B., Konarev P. V., Svergun D. I., Zavodszky P., Vas M. Communication between the nucleotide site and the main molecular hinge of 3-phosphoglycerate kinase // Biochemistry. − 2008. − V. 47, № 26. − P. 6735-6744.

- A47. Szabo J., Varga A., Flachner B., Konarev P. V., Svergun D. I., Zavodszky P., Vas M. Role of side-chains in the operation of the main molecular hinge of 3-phosphoglycerate kinase // Febs Letters. − 2008. − V. 582, № 9. − P. 1335-1340.
- A48. Varga A., Szabo J., Flachner B., Roy B., Konarev P., Svergun D., Zavodszky P., Perigaud C., Barman T., Lionne C., Vas M. Interaction of human 3-phosphoglycerate kinase with L-ADP, the mirror image of D-ADP // Biochem Biophys Res Commun. 2008. V. 366, No. 4. P. 994-1000.
- A49. Graczer E., Konarev P. V., Szimler T., Bacso A., Bodonyi A., Svergun D. I., Zavodszky P., Vas M. Essential role of the metal-ion in the IPM-assisted domain closure of 3-isopropylmalate dehydrogenase // FEBS Lett. − 2011. − V. 585, № 20. − P. 3297-3302.
- A50. Graczer E., Pallo A., Olah J., Szimler T., Konarev P. V., Svergun D. I., Merli A., Zavodszky P., Weiss M. S., Vas M. Glutamate 270 plays an essential role in K(+)-activation and domain closure of Thermus thermophilus isopropylmalate dehydrogenase // FEBS Lett. − 2015. − V. 589, № 2. − P. 240-245.
- A51. Graczer E., Szimler T., Garamszegi A., Konarev P. V., Labas A., Olah J., Pallo A., Svergun D. I., Merli A., Zavodszky P., Weiss M. S., Vas M. Dual Role of the Active Site Residues of Thermus thermophilus 3-Isopropylmalate Dehydrogenase: Chemical Catalysis and Domain Closure // Biochemistry. − 2016. − V. 55, № 3. − P. 560-574.
- A52. Goettig P., Brandstetter H., Groll M., Gohring W., Konarev P. V., Svergun D. I., Huber R., Kim J. S. X-ray snapshots of peptide processing in mutants of tricorn-interacting factor F1 from Thermoplasma acidophilum // J Biol Chem. -2005. -V. 280, N_{\odot} 39. -P. 33387-33396.
- A53. Qazi O., Bolgiano B., Crane D., Svergun D. I., Konarev P. V., Yao Z. P., Robinson C. V., Brown K. A., Fairweather N. The HC fragment of tetanus toxin forms stable, concentration-dependent dimers via an intermolecular disulphide bond // J Mol Biol. -2007. V. 365, No. 1. P. 123-134.
- A54. Xu X., Reinle W., Hannemann F., Konarev P. V., Svergun D. I., Bernhardt R., Ubbink M. Dynamics in a pure encounter complex of two proteins studied by solution

- scattering and paramagnetic NMR spectroscopy // J Am Chem Soc. -2008. V. 130, No. 20. P. 6395-6403.
- A55. de Chiara C., Rees M., Menon R. P., Pauwels K., Lawrence C., Konarev P. V., Svergun D. I., Martin S. R., Chen Y. W., Pastore A. Self-assembly and conformational heterogeneity of the AXH domain of ataxin-1: an unusual example of a chameleon fold // Biophys J. -2013. V. 104, $N \ge 6. P. 1304-1313$.
- A56. Boikova A. S., Dyakova Y. A., Ilina K. B., Konarev P. V., Kryukova A. E., Kuklin A. I., Marchenkova M. A., Nabatov B. V., Blagov A. E., Pisarevsky Y. V., Kovalchuk M. V. Octamer formation in lysozyme solutions at the initial crystallization stage detected by small-angle neutron scattering // Acta Cryst. D Struct. Biol. − 2017. − V. 73, № Pt 7. − P. 591-599.
- А57. Дьякова Ю. А., Ильина К.Б., Конарев, П.В., Крюкова, А.Е., Марченкова, М.А., Благов, А.Е., Волков, В.В., Писаревский, Ю.В., Ковальчук, М.В. Исследование условий образования единиц роста белкового кристалла в растворах лизоцима методом малоуглового рассеяния рентгеновских лучей // Кристаллография. 2017. Т. 62, № 3. С. 364-369.
- А58. Бойкова А. С., Дьякова, Ю.А., Ильина, К.Б., Конарев, П.В., Крюкова, А.Е., Марченкова, М.А., Благов, А.Е., Писаревский, Ю.В., Ковальчук, М.В. Исследование влияния замены растворителя Н2О на D2О на начальную стадию кристаллизации лизоцима тетрагональной сингонии методом малоуглового рентгеновского рассеяния // Кристаллография. 2017. Т. 62, № 6. С. 876-881.
- А59. Попов А. М., Бойкова А. С., Волков В. В., Дьякова Ю. А., Ильина К. Б., Конарев П. В., Марченкова М. А., Петерс Г. С., Писаревский Ю. В., Ковальчук М. В. Микрофлюидная ячейка для изучения структуры предкристаллизационной стадии растворов белков методом малоуглового рассеяния рентгеновских лучей // Кристаллография. − 2018. − Т. 63, № 5. − С. 697-702.
- Або. Бойкова А. С., Дьякова, Ю.А., Ильина, К.Б., Конарев, П.В., Крюкова, А.Е., Марченкова, М.А., Писаревский, Ю.В., Ковальчук, М.В. Исследование предкристаллизационной стадии раствора (влияния температуры и типа осадителя)

- протеиназы К методом малоуглового рассеяния рентгеновского излучения // Кристаллография. -2018.-T.63, № 6.-C.857-862.
- Аб1. Дьякова Ю. А., Бойкова А. С., Ильина К. Б., Конарев П. В., Марченкова М. А., Писаревский Ю. В., Тимофеев В. И., Ковальчук М. В. Исследование влияния катиона осадителя на образование олигомеров в кристаллизационных растворах белка лизоцима // Кристаллография. 2019. Т. 64, № 1. С. 15-19.
- A62. Kovalchuk M. V., Boikova A. S., Dyakova Y. A., Ilina K. B., Konarev P. V., Marchenkova M. A., Pisarevskiy Y. V., Prosekov P. A., Rogachev A. V., Seregin A. Y. Structural Characteristics of Lysozyme Langmuir Layers Grown on a Liquid Surface from an Oligomeric Mixture Formed during the Early Stages of Lysozyme Crystallization // Thin Solid Films. 2019. V. 677. P. 13-21.
- A63. Kovalchuk M. V., Boikova A. S., Dyakova Y. A., Ilina K. B., Konarev P. V., Kryukova A. E., Marchenkova M. A., Pisarevsky Y. V., Timofeev V. I. Precrystallization phase formation of thermolysin hexamers in solution close to crystallization conditions // J Biomol Struct Dyn. − 2019. − V. 37, № 12. − P. 3058-3064. A64. Marchenkova M. A., Konarev P. V., Rakitina T. V., Timofeev V. I., Boikova A. S., Dyakova Y. A., Ilina K. B., Korzhenevskiy D. A., Yu Nikolaeva A., Pisarevsky Y. V., Kovalchuk M. V. Dodecamers derived from the crystal structure were found in the pre-crystallization solution of the transaminase from the thermophilic bacterium Thermobaculum terrenum by small-angle X-ray scattering // J Biomol Struct Dyn. − 2020. − V. 38, № 10. − P. 2939-2944.
- Аб5. Марченкова М. А., Конарев П. В., Бойкова А. С., Ильина К. Б., Писаревский Ю. В., Ковальчук М. В. Влияние хлоридов одно- и двухвалентных металлов на олигомерный состав кристаллизационных растворов лизоцима и дальнейший рост кристаллов // Кристаллография. 2021. T. 66, № 5. C. 723-729.
- A66. Marchenkova M. A., Konarev P. V., Kordoskaya Y. V., Ilina K. B., Pisarevsky Y. V., Soldatov A. V., Timofeev V. I., Kovalchuk M. V. The role of cations and anions in the formation of crystallization oligomers in protein solutions as revealed by combination

- of small-angle x-ray scattering and molecular dynamics // Crystals. 2022. V. 12. P. 751.
- A67. Kordonskaya Y. V., Timofeev V. I., Marchenkova M. A., Konarev P. V. Identification of the precursor cluster in the crystallization solution of proteinase k protein by molecular dynamics methods // Crystals. 2022. V. 12. P. 484.
- A68. Sukhanov A. E., Ilina K. B., Konarev P. V., Peters G. S., Pisarevsky Y. V., Smirnova E. S., Alekseeva O. A., Kovalchuk M. V. Temperature Behavior of Precursor Clusters at the Pre-Crystallization Phase of K(H2PO4) Studied by SAXS // Crystals. 2023. V. 13. P. 26.
- A69. Konarev P. V., Petoukhov M. V., Dadinova L. A., Fedorova N. V., Volynsky P. E., Svergun D. I., Batishchev O. V., Shtykova E. V. BILMIX: a new approach to restore the size polydispersity and electron density profiles of lipid bilayers from liposomes using small-angle X-ray scattering data // J. Appl. Cryst.— 2020. V. 53. P. 236-243.
- A70. Konarev P. V., Gruzinov A. Y., Mertens H. D. T., Svergun D. I. Restoring structural parameters of lipid mixtures from small-angle X-ray scattering data // J. Appl. Cryst. 2021. V. 54, № Pt 1. P. 169-179.
- A71. Kordyukova L. V., Konarev P. V., Fedorova N. V., Shtykova E. V., Ksenofontov A. L., Loshkarev N. A., Dadinova L. A., Timofeeva T. A., Abramchuk S. S., Moiseenko A. V., Baratova L. A., Svergun D. I., Batishchev O. V. The cytoplasmic tail of influenza A virus hemagglutinin and membrane lipid composition change the mode of M1 protein association with the lipid bilayer // Membranes (Basel). 2021. V. 11. P. 772.
- A72. Konarev P. V., Svergun D. I. Direct shape determination of intermediates in evolving macromolecular solutions from small-angle scattering data // IUCrJ. -2018. V. 5, No Pt 4. P. 402-409.
- A73. Konarev P. V., Graewert M. A., Jeffries C. M., Fukuda M., Cheremnykh T. A., Volkov V. V., Svergun D. I. EFAMIX, a tool to decompose inline chromatography SAXS data from partially overlapping components // Protein Sci. 2022. V. 31. P. 269-282. A74. Конарев П. В., Волков, В.В. Восстановление профилей интенсивности
- малоуглового рассеяния от двух конформационных состояний 3-

- изопропилмалатдегидрогеназы с помощью эволюционного факторного анализа // Кристаллография. -2023. T. 68, № 6. C. 948-952.
- A75. Konarev P. V., Volkov V. V. Development of algorithms for analysis of small-angle x-ray scattering data from polydisperse and partially ordered systems // Physics of Atomic Nuclei. 2022. V. 85. P. 104-109.
- A76. Konarev P. V., Volkov V. V., Svergun D. I. Interactive graphical system for small-angle scattering analysis of polydisperse systems // Journal of Physics: Conference Series. 2016. V. 747. P. 012036.
- A77. Filippov S. K., Franklin J. M., Konarev P. V., Chytil P., Etrych T., Bogomolova A., Dyakonova M., Papadakis C. M., Radulescu A., Ulbrich K., Stepanek P., Svergun D. I. Hydrolytically degradable polymer micelles for drug delivery: a SAXS/SANS kinetic study // Biomacromolecules. − 2013. − V. 14, № 11. − P. 4061-4070.
- A78. Svergun D. I., Shtykova E. V., Kozin M. B., Volkov V. V., Konarev P. V., Dembo A. T., Shtykova E. V. J., Bronshtein L. M., Chernyshov D. M., Platonova O. A., Yakunin A. N., Valetsky P. M., R. H. A. Small-angle X-ray scattering study of the structure of self-organized polymer matrices and formation of embedded metal nanoparticles // Cryst. Rep. 2001. V. 46. P. 586-595.
- A79. Bronstein L. M., Karlinsey R. L., Yi Z., Carini J., Werner-Zwanziger U., Konarev P. V., Svergun D. I., Sanchez A., Khan S. Composite solid polymer electrolytes based on pluronics: Does ordering matter? // Chemistry of Materials. − 2007. − V. 19, № 25. − P. 6258-6265.
- A80. Pinotsis N., Chatziefthimiou S. D., Berkemeier F., Beuron F., Mavridis I. M., Konarev P. V., Svergun D. I., Morris E., Rief M., Wilmanns M. Superhelical architecture of the myosin filament-linking protein myomesin with unusual elastic properties // PLoS Biol. -2012.-V.10, No 2.-P. e1001261.
- A81. King-Scott J., Konarev P. V., Panjikar S., Jordanova R., Svergun D. I., Tucker P. A. Structural characterization of the multidomain regulatory protein Rv1364c from Mycobacterium tuberculosis // Structure. − 2011. − V. 19, № 1. − P. 56-69.

- A82. Bonanomi M., Mazzucchelli S., D'Urzo A., Nardini M., Konarev P. V., Invernizzi G., Svergun D. I., Vanoni M., Regonesi M. E., Tortora P. Interactions of ataxin-3 with its molecular partners in the protein machinery that sorts protein aggregates to the aggresome // Int J Biochem Cell Biol. 2014. V. 51. P. 58-64.
- A83. Dadinova L. A., Shtykova E. V., Konarev P. V., Rodina E. V., Snalina N. E., Vorobyeva N. N., Kurilova S. A., Nazarova T. I., Jeffries C. M., Svergun D. I. X-Ray Solution Scattering Study of Four Escherichia coli Enzymes Involved in Stationary-Phase Metabolism // PLoS One. -2016. V. 11, No. 5. P. e0156105.
- A84. Cordeiro M., Otrelo-Cardoso A. R., Svergun D. I., Konarev P. V., Lima J. C., Santos-Silva T., Baptista P. V. Optical and Structural Characterization of a Chronic Myeloid Leukemia DNA Biosensor // ACS Chem Biol. − 2018. − V. 13, № 5. − P. 1235-1242.
- A85. Boyko K. M., Baymukhametov T. N., Chesnokov Y. M., Hons M., Lushchekina S. V., Konarev P. V., Lipkin A. V., Vasiliev A. L., Masson P., Popov V. O., Kovalchuk M. V. 3D structure of the natural tetrameric form of human butyrylcholinesterase as revealed by cryoEM, SAXS and MD // Biochimie. 2019. V. 156. P. 196-205.
- A86. Czech A., Konarev P. V., Goebel I., Svergun D. I., Wills P. R., Ignatova Z. Octarepeat domain of the mammalian prion protein mRNA forms stable A-helical hairpin structure rather than G-quadruplexes // Sci Rep. − 2019. − V. 9, № 1. − P. 2465-2475.
- A87. Bernardini A., Lorenzo M., Chaves-Sanjuan A., Swuec P., Pigni M., Saad D., Konarev P. V., Graewert M. A., Valentini E., Svergun D. I., Nardini M., Mantovani R., Gnesutta N. The USR domain of USF1 mediates NF-Y interactions and cooperative DNA binding // Int J Biol Macromol. − 2021. − V. 193, № Pt A. − P. 401-413.
- A88. Chukhontseva K. N., Berdyshev I. M., Safina D. R., Karaseva M. A., Bozin T. N., Salnikov V. V., Konarev P. V., Volkov V. V., Grishin A. V., Kozlovskiy V. I., Kostrov S. V., Demidyuk I. V. The protealysin operon encodes emfourin, a prototype of a novel family of protein metalloprotease inhibitors // Int J Biol Macromol. 2021. V. 169. P. 583-596.

- A89. Sokolova O. S., Pichkur E. B., Maslova E. S., Kurochkina L. P., Semenyuk P. I., Konarev P. V., Samygina V. R., Stanishneva-Konovalova T. B. Local Flexibility of a New Single-Ring Chaperonin Encoded by Bacteriophage AR9 Bacillus subtilis // Biomedicines. − 2022. − V. 10, № 10. − P. 2347.
- A90. Nowak E., Potrzebowski W., Konarev P. V., Rausch J. W., Bona M. K., Svergun D. I., Bujnicki J. M., Le Grice S. F., Nowotny M. Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid // Nucleic Acids Res. -2013. V. 41, No. 6. P. 3874-3887.
- A91. Mosbaek C. R., Konarev P. V., Svergun D. I., Rischel C., Vestergaard B. High concentration formulation studies of an IgG2 antibody using small angle X-ray scattering // Pharm Res. 2012. V. 29, № 8. P. 2225-2235.
- A92. Adrover M., Martorell G., Martin S. R., Urosev D., Konarev P. V., Svergun D. I., Daura X., Temussi P., Pastore A. The role of hydration in protein stability: comparison of the cold and heat unfolded states of Yfh1 // J Mol Biol. − 2012. − V. 417, № 5. − P. 413-424.
- A93. Manolaridis I., Mumtsidu E., Konarev P., Makhov A. M., Fullerton S. W., Sinz A., Kalkhof S., McGeehan J. E., Cary P. D., Griffith J. D., Svergun D., Kneale G. G., Tucker P. A. Structural and biophysical characterization of the proteins interacting with the herpes simplex virus 1 origin of replication // J Biol Chem. − 2009. − V. 284, № 24. − P. 16343-16353.
- A94. Werner A., Konarev P. V., Svergun D. I., Hahn U. Characterization of a fluorophore binding RNA aptamer by fluorescence correlation spectroscopy and small angle X-ray scattering // Anal Biochem. − 2009. − V. 389, № 1. − P. 52-62.
- A95. Weghofer M., Dall'Antonia Y., Grote M., Stocklinger A., Kneidinger M., Balic N., Krauth M. T., Fernandez-Caldas E., Thomas W. R., van Hage M., Vieths S., Spitzauer S., Horak F., Svergun D. I., Konarev P. V., Valent P., Thalhamer J., Keller W., Valenta R., Vrtala S. Characterization of Der p 21, a new important allergen derived from the gut of house dust mites $\frac{1}{2}$ Allergy. -2008. -V. 63, N 6. -P. 758-767.

- A96. Redecke L., von Bergen M., Clos J., Konarev P. V., Svergun D. I., Fittschen U. E., Broekaert J. A., Bruns O., Georgieva D., Mandelkow E., Genov N., Betzel C. Structural characterization of beta-sheeted oligomers formed on the pathway of oxidative prion protein aggregation in vitro // J Struct Biol. − 2007. − V. 157, № 2. − P. 308-320.
- A97. Raczynska J., Olchowy J., Konarev P. V., Svergun D. I., Milewski S., Rypniewski W. The crystal and solution studies of glucosamine-6-phosphate synthase from Candida albicans // J Mol Biol. − 2007. − V. 372, № 3. − P. 672-688.
- A98. Nardini M., Svergun D., Konarev P. V., Spano S., Fasano M., Bracco C., Pesce A., Donadini A., Cericola C., Secundo F., Luini A., Corda D., Bolognesi M. The C-terminal domain of the transcriptional corepressor CtBP is intrinsically unstructured // Protein Sci. -2006. V. 15, No 5. P. 1042-1050.
- A99. Riek U., Scholz R., Konarev P., Rufer A., Suter M., Nazabal A., Ringler P., Chami M., Muller S. A., Neumann D., Forstner M., Hennig M., Zenobi R., Engel A., Svergun D., Schlattner U., Wallimann T. Structural properties of AMP-activated protein kinase: dimerization, molecular shape, and changes upon ligand binding // J Biol Chem. − 2008. − V. 283, № 26. − P. 18331-18343.
- A100. Simpson P. J., Monie T. P., Szendroi A., Davydova N., Tyzack J. K., Conte M. R., Read C. M., Cary P. D., Svergun D. I., Konarev P. V., Curry S., Matthews S. Structure and RNA interactions of the N-terminal RRM domains of PTB // Structure. -2004. V. 12, N = 9. P. 1631-1643.
- A101. Beich-Frandsen M., Vecerek B., Konarev P. V., Sjoblom B., Kloiber K., Hammerle H., Rajkowitsch L., Miles A. J., Kontaxis G., Wallace B. A., Svergun D. I., Konrat R., Blasi U., Djinovic-Carugo K. Structural insights into the dynamics and function of the C-terminus of the E. coli RNA chaperone Hfq // Nucleic Acids Res. 2011. V. 39, $Noldsymbol{Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{Noldsymbol{P}}$ $Noldsymbol{P}$ $Noldsymbol{P}$ N
- A102. de Almeida Ribeiro E., Jr., Beich-Frandsen M., Konarev P. V., Shang W., Vecerek B., Kontaxis G., Hammerle H., Peterlik H., Svergun D. I., Blasi U., Djinovic-Carugo K. Structural flexibility of RNA as molecular basis for Hfq chaperone function // Nucleic Acids Res. − 2012. − V. 40, № 16. − P. 8072-8084.

A103. Ribeiro Ede A., Jr., Pinotsis N., Ghisleni A., Salmazo A., Konarev P. V., Kostan J., Sjoblom B., Schreiner C., Polyansky A. A., Gkougkoulia E. A., Holt M. R., Aachmann F. L., Zagrovic B., Bordignon E., Pirker K. F., Svergun D. I., Gautel M., Djinovic-Carugo K. The structure and regulation of human muscle alpha-actinin // Cell. − 2014. − V. 159, № 6. − P. 1447-60.

A104. Song J. G., Kostan J., Drepper F., Knapp B., de Almeida Ribeiro E., Jr., Konarev P. V., Grishkovskaya I., Wiche G., Gregor M., Svergun D. I., Warscheid B., Djinovic-Carugo K. Structural insights into Ca2+-calmodulin regulation of Plectin 1a-integrin beta4 interaction in hemidesmosomes // Structure. − 2015. − V. 23, № 3. − P. 558-570. A105. Adinolfi S., Puglisi R., Crack J. C., Iannuzzi C., Dal Piaz F., Konarev P. V., Svergun D. I. Martin S. Le Brun N. F. Pastore A. The Molecular Bases of the Dual

Svergun D. I., Martin S., Le Brun N. E., Pastore A. The Molecular Bases of the Dual Regulation of Bacterial Iron Sulfur Cluster Biogenesis by CyaY and IscX // Front Mol Biosci. – 2017. – V. 4. – P. 97-109.

A106. Filippov S. K., Chytil P., Konarev P. V., Dyakonova M., Papadakis C., Zhigunov A., Plestil J., Stepanek P., Etrych T., Ulbrich K., Svergun D. I. Macromolecular HPMA-based nanoparticles with cholesterol for solid-tumor targeting: detailed study of the inner structure of a highly efficient drug delivery system // Biomacromolecules. -2012.-V. 13, No. 8.-P. 2594-2604.

A107. Bogomolova A., Filippov S. K., Starovoytova L., Angelov B., Konarev P., Sedlacek O., Hruby M., Stepanek P. Study of complex thermosensitive amphiphilic polyoxazolines and their interaction with ionic surfactants. Are hydrophobic, thermosensitive, and hydrophilic moieties equally important? // J Phys Chem B. -2014. -V. 118, N 0 18. -P. 4940-4950.

A108. Sergeeva O., Vlasov P. S., Domnina N. S., Bogomolova A., Konarev P. V., Svergun D. I., Walterova Z., Horsky J., Stepanek P., Filippov S. K. Novel thermosensitive telechelic PEGs with antioxidant activity: synthesis, molecular properties and conformational behavior. // RSC Advances. – 2014. – V. 4. – P. 41763-41771.

A109. Filippov S. K., Verbraeken B., Konarev P. V., Svergun D. I., Angelov B., Vishnevetskaya N. S., Papadakis C. M., Rogers S., Radulescu A., Courtin T., Martins J.

- C., Starovoytova L., Hruby M., Stepanek P., Kravchenko V. S., Potemkin, II, Hoogenboom R. Block and Gradient Copoly(2-oxazoline) Micelles: Strikingly Different on the Inside // J Phys Chem Lett. − 2017. − V. 8, № 16. − P. 3800-3804.
- A110. Shtykova E. V., Shtykova E. V., Jr., Volkov V. V., Konarev P. V., Dembo A. T., Makhaeva E. E., Ronova I. A., Khokhlov A. R., Reynaers H., Svergun D. I. Small-angle X-ray scattering reveals hollow nanostructures in native kappa- and iota-carrageenan /surfactant complexes // J. Appl. Cryst.— 2003. V. 36. P. 669-673.
- A111. Bronstein L. M., Dixit S., Tomaszewski J., Stein B., Svergun D. I., Konarev P. V., Shtykova E., Werner-Zwanziger U., Dragnea B. Hybrid Polymer Particles with a Protective Shell: Synthesis, Structure, and Templating // Chem. Mater. 2006. V. 18. P. 2418-2430.
- A112. Malyutin A. G., Cheng H., Sanchez-Felix O. R., Carlson K., Stein B. D., Konarev P. V., Svergun D. I., Dragnea B., Bronstein L. M. Coat Protein-Dependent Behavior of Poly(ethylene glycol) Tails in Iron Oxide Core Virus-like Nanoparticles // ACS Appl Mater Interfaces. − 2015. − V. 7, № 22. − P. 12089-12098.
- А113. Мулаков С. П., Конарев П. В., Тимаева О. И., Кузьмичева Г. М. Применение рентгеновских методов для определения размеров наночастиц в системе наноразмерный анатаз—поли-n-винилкапролактам // Кристаллография. 2020. Т. 65, № 4. С. 644-653.
- A114. Timaeva O., Pashkin I., Mulakov S., Kuzmicheva G., Konarev P., Terekhova R., Sadovskaya N., Czakkel O., Prevost S. Synthesis and physico-chemical properties of poly(N-vinyl pyrrolidone)-based hydrogels with titania nanoparticles // J Mater Sci. -2020.-V.55, No 7.-P.3005-3021.
- А115. Крюкова А. Е., Шпичка, А.И., Конарев, П.В., Волков, В.В., Тимашев, П.С., Асадчиков, В.Е. Восстановление формы бычьего фибриногена в растворе по данным малоуглового рассеяния // Кристаллография. − 2018. − Т. 63, № 6. − С. 863-865.
- A116. Shpichka A. I., Konarev P. V., Efremov Y. M., Kryukova A. E., Aksenova N. A., Kotova S. L., Frolova A. A., Kosheleva N. V., Zhigalina O. M., Yusupov V. I.,

Khmelenin D. N., Koroleva A., Volkov V. V., Asadchikov V. E., Timashev P. S. Digging deeper: structural background of PEGylated fibrin gels in cell migration and lumenogenesis // RSC Adv. -2020.-V. 10, No 8. -P. 4190-4200.

A117. Konarev P. V., Grigorev, V. A., Bikmulina, P. Yu., Presnyakova, V. S., Kryukova, A. E., Volkov, V. V., Shpichka, A. I., Asadchikov, V. E., Timashev, P. S. The Structural Features of Native Fibrin and Its Conjugates with Polyethylene Glycol and Vascular Endothelial Growth Factor according to Small-Angle X-Ray Scattering // Review Journal of Chemistry. − 2020. − V. 10, № 3-4. − P. 158–163.

Патенты

- Попов А.М., Волков В.В., Дьякова Ю.А., Конарев П.В., Марченкова М.А., Писаревский Ю.В., Ковальчук М.В. – "Микрофлюидная ячейка для условий кристаллизации белков из характеристик малоуглового рассеяния рентгеновских лучей" – Патент на полезную модель № RU 182997 U1 от 22.06.2018.
- 2. Самыгина В.Р., Пичкур Е.Б., Конарев П.В., Ворович М.Ф., Ишмухаметов А.А., Егоров А.М. "Способ оценки качества образца флавивируса для получения трехмерной структуры с использованием лазера на свободных электронах" Патент на полезную модель № RU 2741124 C1 от 22.01.2021.

Тезисы докладов на конференциях

- Т1. Конарев, П.В. Исследование структурных изменений в трехкомпонентных микроэмульсиях типа вода—АОТ—изооктан методом малоуглового рассеяния / П.В. Конарев, Д.И. Свергун, В.В. Волков, М.Н.J. Косh, J. Smeets, W.F.C. Sager / II Национальная конференция по применению Рентгеновского, Синхротронного излучений, Нейтронов и Электронов для исследования материалов (РСНЭ—99), Москва, 1999, 23—27 мая, Материалы конференции. с. 280.
- T2. D.I.Svergun, P.V.Konarev, V.V.Volkov, M.H.J.Koch, J.Smeets, W.F.C.Sager // A small-angle scattering study of the droplet-cylinder transition in AOT microemulsions. Abstracts of the XI International Conference on Small-Angle Scattering New York, USA, 17-20 May, 1999, p. 147
- T3. P.V. Konarev, V.V. Volkov, M.V. Petoukhov, D.I. Svergun // Programs for non-linear modelling of small-angle scattering data. Abstracts of the 9th Annual Fibre Diffraction and Non-Crystalline Diffraction Workshop, Sheffield, England, 26-28 June, 2000, p. 16
- T4. Э.В.Штыкова, Д.И.Свергун, Е.В.Штыкова, В.В.Волков, П.В.Конарев, А.Т.Дембо, Е.Е.Махаева, А.Р.Хохлов, Г.А.Евмененко, Х.Райнерс. Сравнительное малоугловое рентгеновское исследование структуры комплексов - и -каррагинанов с ПАВ. // III Национальная конференция по применению Рентгеновского, Синхротронного излучений, Нейтронов и Электронов ДЛЯ исследования материалов (РСНЭ-2001), Москва, 2001, 21-25 мая, Материалы конференции, с. 180 Конарев, П.В. Massha – новый программный пакет для моделирования T5. трехмерных структур по данным малоуглового рассеяния. / П.В. Конарев, М.В. Петухов, В.В. Волков, Д.И. Свергун / РСНЭ–2001. – Москва, 2001. – с. 319
- T6. E.V.Shtykova, D.I.Svergun, E.V.Shtykova, Jr., V.V.Volkov, P.V.Konarev, A.T.Dembo, E.E.Makhaeva, A.R.Khokhlov. Comparative small-angle X-ray study and modelling of and -carrageenan/curfactant complexes. // The Third International Conference "Chemistry of Highly-Organized Substances and Scientific Principles of Nanotechnology", St-Peterburg, 26-28 June, 2001, p.452-453.

- T7. Konarev, P.V. WINTEL-based menu-driven program packages for small-angle scattering data analysis. / P.V. Konarev, M.V. Petoukhov, V.V. Volkov, D.I. Svergun / XIII International Conference On Small-Angle Scattering. Venice, Italy, August 25–29, 2002. p. 167.
- T8. E.V.Shtykova, E.V.Shtykova, Jr., V.V.Volkov, P.V.Konarev, A.T.Dembo, E.E.Makhaeva, A.R.Khokhlov, I.A.Ronova, H.Reynaers, D.I.Svergun. Small-angle X-ray scattering reveals superhelical nanostructures in and -carrageenan/surfactant complexes. // XII International Conference on Small-Angle Scattering, SAS-2002, August 25-29, 2002, Venice, Italy, Conference book, p.59
- T9. Э.В.Штыкова, Е.В.Штыкова, В.В.Волков, П.В.Конарев, А.Т.Дембо, А.Р.Хохлов, Д.И.Свергун. Е.Е.Махаева, И.А.Ронова, Полые спиральные наноструктуры в каррагинан/ПАВ комплексах, выявленные с помощью малоуглового рентгеновского рассеяния. // Международная конференция по Рентгеновского, Синхротронного излучений, Нейтронов применению Электронов для исследования материалов (РСНЭ-2003), Москва, 2003, 17-22 ноября, Материалы конференции, с. 308.
- Т10. П.В. Конарев, В.В. Волков, А.В. Соколова, М.Н.J. Косh, Д.И. Свергун. PRIMUS новый программный пакет анализа данных малоуглового рассеяния. // Международная конференция по применению Рентгеновского, Синхротронного излучений, Нейтронов и Электронов для исследования материалов (РСНЭ-2003), Москва, 2003, 17-22 ноября, Материалы конференции, с. 341.
- Т11. М.В. Петухов, А.В. Соколова, П.В. Конарев, М.Б. Козин, В.В. Волков, Д.И. Свергун. ATSAS 2.0 комплексная система обработки и интерпретации данных малоуглового рассеяния. // Международная конференция по применению Рентгеновского, Синхротронного излучений, Нейтронов и Электронов для исследования материалов (РСНЭ-2003), Москва, 2003, 17-22 ноября, Материалы конференции, с. 348.
- T12. P.V. Konarev, M.V. Petoukhov, V.V. Volkov, D.I. Svergun // ATSAS2.1 a program suite for small-angle scattering data analysis. // XX Congress of the International

- Union of Crystallography, (IUCr2005) 23-31 August 2005, Florence, Italy, Abstracts p. 411
- Т13. Д.И.Свергун, Э.В.Штыкова, П.В.Конарев, Л.М.Бронштейн. Гибридные полимерные наночастицы: исследование квазикристаллической структуры. // Международная конференция по применению Рентгеновского, Синхротронного излучений, Нейтронов и Электронов для исследования материалов (РСНЭ-2005), Москва, 2005, 14-18 ноября, Материалы конференции, с.131
- T14. Roessle, M.W., Klaering, R., Ristau, U., Robrahn, B., Jahn, D., Gehrmann, T., Konarev, P., Round, A., Fiedler, S., Hermes, C. & Svergun, D. Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg. // XIII International Conference on Small-angle Scattering (SAS2006). Japan, Kyoto, July 9 13, 2006, Abstracts, P.87.
- T15. Petoukhov, M.V., Konarev, P.V., Kikhney, A.G. & Svergun, D.I. ATSAS 2.1 towards automated and web-supported small-angle scattering data analysis. // XIII International Conference on Small-angle Scattering (SAS2006). Japan, Kyoto, July 9 13, 2006, Abstracts, P.126.
- T16. P.V.Konarev, B.Flachner, A.Varga, F.Paoletti, D.Lamba, M. Vas, D.I. Svergun // Investigation of quaternary structure of macromolecules using small-angle X-ray scattering data with program package ATSAS. // European meeting "Trends in Enzymology", 7-10 June 2006, Como, Italy Abstracts p. 39
- T17. B.Flachner, P.Konarev, A.Varga, J.Szabo, I.Hajdu, P.Zavodsky, D.Svergun, M.Vas // Role of the catalytic residues (R38 and K215) in the substrate-assisted domain closure of human 3-phosphoglycerate kinase // European meeting "Trends in Enzymology", 7-10 June 2006, Como, Italy, Abstracts p. 29
- T18. A. Kikhney, P. Konarev, D. Franke, D.I. Svergun //Software for Automated High-throughput Biological Small-Angle X-ray Scattering // 9th International Conference on Biology and Synchrotron Radiation 13-17 August 2007, Manchester, England, Abstracts p. 73

- T19. E.V.Shtykova, P.V.Konarev, L.M.Bronstein, D.I.Svergun. Conventional and anomalous SAXS studies of polymer matrices containing metal nanoparticles. // IUCr2008 satellite meeting on X-ray and neutron techniques for nano-structural research, Spring-8, Hyogo, Japan, 2008, August 21-23, p. 8.
- T20. E.V.Shtykova, P.V.Konarev, L.M.Bronstein, D.I.Svergun. Synchrotron SAXS reveals structural organization of iron oxide nanoparticles in aqueous solutions. // XXI Congress of the International Union of Crystallography (IUCr2008), Osaka, Japan, August 23-31, 2008, Book of Abstracts & Acta Crys., A64, C459-460
- T21. P.V. Konarev, M.V. Petoukhov, E.Mylonas, A.G. Kikhney, V.V. Volkov, D.I. Svergun. Analysis of solution small-angle scattering using the program package ATSAS // // XXI Congress of the International Union of Crystallography (IUCr2008), Osaka, Japan, August 23-31, 2008, Book of Abstracts & Acta Crys., A64, C554-554
- Т22. Э.В.Штыкова, П.В.Конарев, Л.М.Бронштейн, Д.И.Свергун. Исследование структуры метал-содержащих нанокомпозитов методом малоуглового рентгеновского рассеяния. // Конференция "Рентгеновское, Синхротронное излучения, Нейтроны и Электроны для исследования наносистем и материалов". Нано-Био-Инфо-Когнитивные технологии (РСНЭ-НБИК 2009), Москва, 16-21 ноября, 2009, с. 457.
- Т23. А.В. Зозуля, М. Roessle, D. Franke, А.Г. Кихней, П.В. Конарев, М.В. Петухов, W. Shang, C. Blanchet, R. Klaering, B. Robrahn, C. Hermes, Д.И. Свергун. Станция малоуглового рассеяния синхротронного излучения. Станция малоуглового рассеяния синхротронного излучения X33 для исследования биологических макромолекул. // Конференция "Рентгеновское, Синхротронное излучения, Нейтроны и Электроны для исследования наносистем и материалов". Нано-Био-Инфо-Когнитивные технологии (РСНЭ-НБИК 2009), Москва, 16-21 ноября, 2009, с. 581.
- T24. E.V.Shtykova, P.V.Konarev, L.M.Bronstein, D.I.Svergun. Structure of metal-containing nanocomposites revealed by SAXS. // XIV International Conference on

- Small-angle Scattering (SAS2009). UK, Oxford, September 13-18, 2009, Book of Abstract, p. 118.
- T25. P.V.Konarev, M.V.Petoukhov, A.G.Kikhney, M. Gajda, A.Shkumatov, D.Franke, D.I.Svergun. New developments in ATSAS // XIV International Conference on Small-angle Scattering (SAS2009). UK, Oxford, September 13-18, 2009, Book of Abstract, p. 125.
- T26. E.V.Shtykova, P.V.Konarev, L.M.Bronstein, D.I.Svergun. SAXS studies of iron oxide nanoparticles in solution. // 12th International Conference of Magnetic Fluids (ICMF12), Sendai, Japan, 2010, p 106
- T27. E.V.Shtykova, P.V.Konarev, L.M.Bronstein, D.I.Svergun. Monodisperse magnetic nanoparticles and their stabilization in solution by hydrophilic shells: SAXS study and modeling. // The International Chemical Congress of Pacific Basin Societies (Pacifichem 2010), Honolulu, Hawaii, USA, December 15-20, 2010, 09-MATL, p. 235. T28. A.G.Malyutin, G.J.Gumina, E.V.Shtykova, P.V.Konarev, D.I.Svergun, B.Dragnea, L.M.Bronstein. Magnetic core-shell virus-like nanoparticles as promising MRI contrast agents and agents for hyperthermia treatment. // 244th National Fall Meeting of the American-Chemical-Society (ACS). США, Филадельфия, 19-23 августа, 2012.
- T29. Konarev, P. POLYSAS: an interactive praphical system for small—angle scattering analysis of polydisperse systems and multiple data sets / P. Konarev, V. Volkov, D. Svergun / 11th International Conference Biology and Synchrotron Radiation 2013 (BSR–2013). Hamburg, 8 11th September 2013. Abstracts. p. 91.
- T30. Konarev P.V. Rapid automated superposition of high and low-resolution models using spherical harmonics / P.V. Konarev, M.V.Petoukhov, D.I. Svergun // 16th International Conference on Small-Angle Scattering, 13-18 September 2015, Berlin, Germany. c. 320 (341 c)
- T31. Kikhney A.G. ATSAS 2.7 automated processing and advanced interpretation of scattering from isotropic systems / A.G. Kikhney, D. Franke, P.V. Konarev, M.V. Petoukhov, A. Panjkovich, G. Tria, H.D. Mertens, M. Kachala and D.I. Svergun // 16th

- International Conference on Small-Angle Scattering, 13-18 September 2015, Berlin, Germany. c. 318 (341 c)
- T32. Konarev P.V. Quantitative assessment of the useful angular range for small-angle scattering data from solutions / P.V. Konarev, D.I. Svergun // 16th International Conference on Small-Angle Scattering, 13-18 September 2015, Berlin, Germany. c. 86 (123 c.)
- ТЗЗ. Конарев П.В Использование программного пакета atsas для исследования структуры биологических систем врастворах / П.В. Конарев, В.В. Волков, Д. Франке, Г. Триа., А.Г.Кихней, М.В. Петухов, Д.И. Свергун // VII Российский симпозиум «Белки и пептиды», 12-17 июля 2015, Новосибирск, Россия с. 146 (476 с.)
- T34. Konarev P.V. Interactive graphical system for small-angle scattering analysis of polydisperse systems. P.V. Konarev, V.V. Volkov and D.I. Svergun // II International Conference on Plasma & Laser Research and Technologies, 25-27 January 2016, Moscow, Russia. c. 119 (186 c.)
- Т35. Крюкова А.Е. Исследование усточивости решений при анализе полидисперных систем методом малоуглового рассеяния / А.Е. Крюкова, П.В.. Конарев, В.В. Волков // 1-й Российский Кристаллографический Конгресс, 21-26 ноября 2016, Москва, Россия с. 172 (452 с.)
- T36. Konarev P.V. Rapid automated superposition of high and low resolution models using spherical harmonics / P.V. Konarev, M.V. Petoukhov, D.I. Svergun // 8-й международный научный семинар и 6-я международная молодежная научная школа-семинар, 22 июня-2 июля 2016, Новгород, Россия с. 89 (256 с.)
- Т37. Конарев П.В. Развитие методов автоматической суперпозиции макромолекулярных моделей / П.В. Конарев, М.В.. Петухов, Д.И. Свергун // 1-й Российский Кристаллографический Конгресс, 21-26 ноября 2016, Москва, Россия с. 207 (452 с.)
- Т38. Волков В.В. Проблемы устойчивости и однозначности анализа данных малоуглового рассеяния от моно- и полидисперсных систем / В.В. Волков, П.В.

Конарев, А.Е. Крюкова // 1-й Российский Кристаллографический Конгресс, 21-26 ноября 2016, Москва, Россия с. 123 (452 с.)

ТЗ9. Дадинова Л.А. Малоугловое рентгеновское рассеяние в исследовании четырех ферментов, участвующих в метаболизме стационарной фазы роста клеток escherichia coli / Л.А. Дадинова, Э.В. Штыкова, П.В. Конарев, Е.В. Родина, Н.Н. Воробьева, С.А. Курилова, Т.И. Назарова, С.М. Jeffries, Д.И. Свергун // 1-й Российский Кристаллографический Конгресс, 21-26 ноября 2016, Москва, Россия с. 203 (452 с.)

Т40. Ильина К.Б. Исследование начальной стадии кристаллизации лизоцима тетрагональной сингонии методом малоуглового рассеяния нейтронов / К.Б. Ильина, А.Е. Благов, А.С. Бойкова, Ю.А. Дьякова, П.В. Конарев, А.Е. Крюкова, А.И. Куклин, М.А. Марченкова, Ю.В. Писаревский, М.В. Ковальчук // 1-й Российский Кристаллографический Конгресс, 21-26 ноября 2016, Москва, Россия с. 236 (452 с.)

Т41. Конарев, П.В. Анализ многокомпонентных гетерогенных белковых цистем по данным малоуглового рассеяния / П.В. Конарев, В.В. Волков, Д.И. Свергун / Научные труды объединенного научного форума: международной научной конференции по биоорганической химии «ХІІ чтения памяти академика Ю.А. Овчинникова» и VII российский симпозиум "Белки и пептиды" (18-22 сентября 2017) под ред. В.Т. Иванова и А.Г. Габибова, Аста Naturae — Спец. Выпуск — Москва, 2017, — с. 101 (192 с.)

Т42. Крюкова, А.Е. Устойчивость решений программы mixture при исследовании полидисперсных систем методом малоуглового рассеяния / А.Е. Крюкова, П.В. Конарев, В.В. Волков / Научные труды объединенного научного форума: международной научной конференции по биоорганической химии «ХІІ чтения памяти академика Ю.А. Овчинникова» и VII российский симпозиум "Белки и пептиды" (18-22 сентября 2017) под ред. В.Т. Иванова и А.Г. Габибова, Асtа Naturae — Спец. Выпуск — Москва, 2017, — с. 106 (192 с.)

- Т43. Петухов, М.В. Анализ структуры белков по данным малоуглового рассеяния с помощью программного пакета ATSAS / М.В. Петухов, D. Franke, П.В. Конарев, А. Panjkovich, A. Tuukkanen, H.D.T. Mertens, А.Г. Кихней, N.R. Hajizadeh, С.М. Jeffries, Д.И. Свергун / Научные труды объединенного научного форума: международной научной конференции по биоорганической химии «ХІІ чтения памяти академика Ю.А. Овчинникова» и VII российский симпозиум "Белки и пептиды" (18-22 сентября 2017) под ред. В.Т. Иванова и А.Г. Габибова, Асtа Naturae Спец. Выпуск Москва, 2017, с. 107-108 (192 с)
- T44. Petoukhov, M.V. Structural modeling of liposomes in a free state and with adsorbed Influenza virus M1 protein from synchrotron SAXS data / M.V. Petoukhov, P.V. Konarev, E.V. Shtykova, O.V. Batishchev / 42nd Congress of the Federation-of-European-Biochemical-Societies (FEBS) on From Molecules to Cells and Back Location, 10-14 September 2017, Jerusalem, Israel, FEBS Journal V. 284. Special Issue P. 311
- Т45. Конарев, П.В. Программный пакет ATSAS для анализа данных малоуглового рассеяния/ П. В. Конарев, В. В. Волков, М. В. Петухов, Д. Франке, А. Г. Кихней, А. Панджкович, Х. Мертенс, К. Керр, С. Джеффрис, Д. И. Свергун / Сборник тезисов LI Школы ПИЯФ по физике конденсированного состояния (ФКС-2017), 11–16 марта 2017 г., Санкт-Петербург с. 234 (264 с.)
- T46. Kryukova, A.E. Evaluation of solution stability for two-component polydisperse systems by small-angle scattering / A E Kryukova, P.V. Konarev, and V.V. Volkov / III International Conference on Plasma & Laser Research and Technologies, 24-27 January 2017, Moscow, Russia c. 116 (292 c.)
- T47. Pankovich, A. ATSAS 2.8 software for small-angle scattering from macromolecular solutions / A. Pankovich, D. Franke, M. V. Petoukhov, P. V. Konarev, A. Tuukkanen, H. D. T. Mertens, A. G. Kikhney, N. R. Hajizadeh, J. M. Franklin, C. M. Jeffries, D. I. Svergun / Abstracts from the 24th Congress of the International Union of Crystallography 2017 from 21 28 August 2017, Hyderabad, India, P. 339 (1350 p.)

- Т48. Крюкова, А.Е. Восстановление решений при анализе полидисперсной системы в зависимости от зашумленности данных малоуглового рассеяния / А. Е. Крюкова, П. В. Конарев, В. В. Волков / Сборник тезисов LI Школы ПИЯФ по физике конденсированного состояния (ФКС-2017), 11–16 марта 2017 г., Санкт-Петербург с. 235 (264 с.)
- Т49. Конарев, П.В. Моделирование многокомпонентных полидисперсных систем по данным малоуглового рассеяния / П.В. Конарев, А.Е. Крюкова, В.В. Волков, Д.И. Свергун / Сборник научных трудов IV международной конференции «Лазерные, плазменные исследования и технологии» (Лаплаз -2018), 30 января-1 февраля 2018 г. Москва, 2018, с. 101-102 (543 с.)
- Т50. Крюкова, А.Е. Восстановление функции распределения по размерам растворов кремнезолей по данным малоуглового рассеяния / А.Е. Крюкова, П. В. Конарев, В. В. Волков. В.Е. Асадчиков, А.С. Козлова / Сборник тезисов LII Школы ПИЯФ по физике конденсированного состояния (ФКС-2018), 12–17 марта 2018 г. Санкт-Петербург, 2018, с. 56 (260 с.)
- Т51. Волков, В.В. Стратегия моделирования формы частиц по данным малоуглового рассеяния / В.В. Волков, П.В.. Конарев, А.С. Козлова / Сборник тезисов LII Школы ПИЯФ по физике конденсированного состояния (ФКС-2018), 12–17 марта 2018 г. Санкт-Петербург, 2018, с. 224 (260 с.)
- Т52. Конарев, П.В. Современные возможности метода малоуглового рентгеновского рассеяния для структурного исследования биологических макромолекул в растворе / П.В. Конарев / Сборник тезисов XIII Международной научной конференции «Проблемы сольватации и комплексо-образования в растворах», 1-6 июля 2018 г. Суздаль, 2018, с. 60 (509 с.)
- T53. Franke, D. Recent progress in the validation and analysis of sas experiments from biological macromolecules in solution / D. Franke, M.V.Petoukhov, P. V. Konarev, A. Panjkovich, A. Tuukkanen, N. R. Hajizadeh, A. G. Kikhney, C. M. Jeffries, D. I. Svergun / Abstracts of XVII International Small-Angle Scattering Conference (SAS-2018), 7-12 October 2018. —Treverse City, Michigan, USA, 2018, —p. 93 (544 p.)

- T54. Hajizadeh, N.R. Atsas 3.0 software suite for small angle scattering from macromolecular solutions / N. R. Hajizadeh, M.V. Petoukhov, P. V Konarev, A. Tuukkanen, H. D.T Mertens, K. K.Manalastas, A. G. Kikhney, A. Panjkovich, C. M. Jeffries, D. Franke, D. I. Svergun / Abstracts of XVII International Small-Angle Scattering Conference (SAS-2018), 7-12 October 2018. Treverse City, Michigan, USA, 2018, p. 135 (544 p.)
- T55. Filippov, S.K. Block and gradient copoly(2-oxazoline) micelles: strikingly different on the inside / S. K. Filippov, B. Verbraeken, P. V. Konarev, D. I. Svergun, N. S. Vishnevetskaya, C. M. Papadakis, S. Rogers, A. Radulescu, T. Courtin, J. C.Martins, L. Starovoytova, M. Hruby, P. Stepanek, V. S. Kravchenko, I. I. Potemkin, R. Hoogenboom / Abstracts of XVII International Small-Angle Scattering Conference (SAS-2018), 7-12 October 2018. Treverse City, Michigan, USA, 2018, p. 490 (544 p.)
- Т56. Петухов, М.В. Трехмерное моделирование структуры макромолекул по данным рассеяния растворами / М.В. Петухов, D. Franke, П.В. Конарев, А. Panjkovich, A. Tuukkanen, H.D.T. Mertens, А.Г. Кихней, N.R. Hajizadeh, С.М. Jeffries, Д.И. Свергун / Конференция по использованию рассеяния нейтронов в конденсированных средах РНИКС-2018, 17-21 сентября 2018 г. Санкт-Петербург, 2018, с. 62 (260 с.)
- Т57. Конарев, П.В. Развитие методов моделирования многокомпонентных гетерогенных систем по данным малоуглового рассеяния / П.В. Конарев, А.Е. Крюкова, В.В. Волков / Сборник тезисов LII Школы ПИЯФ по физике конденсированного состояния (ФКС-2018), 12–17 марта 2018 г. Санкт-Петербург, 2018, с. 55 (260 с.)
- Т58. Петухов, М.В. Трехмерное моделирование структуры макромолекул по данным рассеяния растворами / М.В. Петухов, D. Franke, П.В. Конарев, А. Panjkovich, A. Tuukkanen, H.D.T. Mertens, А.Г. Кихней, N.R. Hajizadeh, С.М. Jeffries, Д.И. Свергун / Сборник тезисов LII Школы ПИЯФ по физике

- конденсированного состояния (ФКС-2018), 12–17 марта 2018 г. Санкт-Петербург, 2018, — с. 65 (260 с.)
- Т59. Конарев, П.В. Исследование формы и сольватации макромолекул таннина в водно-спиртовом растворе методом малоуглового рентгеновского рассеяния / П.В. Конарев / Сборник тезисов XIII Международной научной конференции «Проблемы сольватации и комплексо-образования в растворах», 1-6 июля 2018 г. Суздаль, 2018, с. 111 (509 с.)
- Т60. Крюкова, А.Е. Восстановление решений при анализе растворов кремнезолей методом малоуглового рассеяния / А.Е. Крюкова, П. В. Конарев, В. В. Волков. В.Е. Асадчиков, А.С. Козлова / Сборник тезисов XIII Международной научной конференции «Проблемы сольватации и комплексо-образования в растворах», 1-6 июля 2018 г. Суздаль, 2018, с. 136 (509 с.)
- Т61. Конарев, П.В. Развитие методов моделирования структуры многослойных липидных везикул по данным малоуглового рассеяния / П.В. Конарев, А.Ю. Грузинов, Д.И. Свергун / Конференция по использованию рассеяния нейтронов в конденсированных средах РНИКС-2018, 17-21 сентября 2018 г. Санкт-Петербург, 2018, с. 187 (260 с.)
- Т62. Конарев, П.В. К вопросу о восстановлении профилей рассеяния отдельных компонентов в белковых смесях по данным малоуглового рентгеновского рассеяния / П.В. Конарев, В.В. Волков / ІІ объединенный научный форум: vi съезд физиологов СНГ, VI съезд биохимиков России, IX российский симпозиум «белки и пептиды», Научные труды. Том 2. 1-6 октября 2019 г. Сочи, Дагомыс 2019, с. 59-60 (299 с.)
- Т63. А.В. Скобёлкина, А.В. Структурные особенности процессов самосборки фрагментов белка предшественника бета-амилоида в липидных мембранах / А.В. Скобёлкина, Е.В. Бочаров, О.В. Бочарова, М.В. Петухов, П.В. Конарев, О.В. Батищев, Э.В. Штыкова / II объединенный научный форум: vi съезд физиологов СНГ, VI съезд биохимиков России, IX российский симпозиум «белки и пептиды»,

Научные труды. Том 2. 1-6 октября 2019 г. — Сочи, Дагомыс 2019, — с. 56-57 (299 с.)

Т64. Демидюк, И.В. Новое семейство белковых ингибиторов протеаз / И.В. Демидюк, К.Н. Чухонцева, М.А. Карасева, Д.Р. Сафина, Т.Н. Бозин, Э.В. Бочаров, П.В. Конарев, В.В. Сальников, С.В. Костров / П объединенный научный форум: vi съезд физиологов СНГ, VI съезд биохимиков России, IX российский симпозиум «белки и пептиды», Научные труды. Том 2. 1-6 октября 2019 г. — Сочи, Дагомыс 2019, — с. 64 (299 с.)

Т65. Крюкова, А.Е. Поиск распределений частиц по размерам путем комбинирования методов минимизации на примере данных малоуглового рассеяния от раствора кремнезоля / А.Е. Крюкова, П.В. Конарев, В.В. Волков / Сборник научных трудов V международной конференции «Лазерные, плазменные исследования и технологии» (Лаплаз -2019), Часть 1, 12-15 февраля 2019 г. — Москва, 2019, — с. 111-112 (348 с.)

Т66. Козлова, А.С. Алгоритм поиска размерного распределения с оценкой стабильности решения по данным малоуглового рассеяния / А.С. Козлова, В. В. Волков, А.Е. Крюкова, П. В. Конарев / Сборник тезисов LIII Школы ПИЯФ по физике конденсированного состояния (ФКС-2019), 11–16 марта 2019 г. — Санкт-Петербург, 2019, — с. 221 (249 с.)

T67. Franke, D. Atsas 3.0 software suite for small angle scattering from macromolecular solutions / D. Franke, N.R. Hajizadeh, M.V. Petoukhov, P.V. Konarev, H. D.T Mertens, K. K. Manalastas, A.G. Kikhney, C. M. Jeffries, D. I. Svergun / Abstracts of 13th International Conference On Biology and Synchrotron Radiation (BSR-2019), 21-24 September 2019. —ShanghaiTech, China, 2019, — p. 39-40 (73 p.)

Т68. Крюкова, А.Е. Сравнение эффективности восстановления решении при анализе данных малоуглового рассеяния от раствора кремнезоля набором итеративных алгоритмов / А.Е. Крюкова, П. В. Конарев, В. В. Волков / Сборник тезисов LIII Школы ПИЯФ по физике конденсированного состояния (ФКС-2019), 11–16 марта 2019 г. — Санкт-Петербург, 2019, — с. 224 (249 с.)

- Т69. Скобёлкина, А.В. Изучение процессов самосборки фрагментов белка предшественника бета-амилоида в липидных мембранах / А.В. Скобёлкина, Е.В. Бочаров, О.В. Бочарова, М.В. Петухов, П.В. Конарев, О.В. Батищев, Э.В. Штыкова / Сборник тезисов LIII Школы ПИЯФ по физике конденсированного состояния (ФКС-2019), 11–16 марта 2019 г. Санкт-Петербург, 2019, с. 177 (249 с.)
- T70. Samygina, V.R. Quality control of tick-bone encephalitis virus samples using TEM and SAXS for XFEL studies / V. R. Samygina, E. B. Pichkur, M. F. Vorovich, P. V. Konarev, G. S. Peters, D. I. Osolodkin, A. A. Ishmukhametov, A. M. Egorov / Abstracts of Russian International Conference on Cryo-Electron Microscopy 2019, 2-5 June 2019.

 —Moscow, Russia, 2019, p. 97-98 (179 p.)
- Т71. Крюкова, А.Е. Сравнение эффективности решений, полученных рядом итеративных алгоритмов при анализе данных малоуглового рассеяния от раствора кремнезоля / А.Е. Крюкова, П.В. Конарев, В.В. Волков // Сборник научных трудов VI международной конференции «Лазерные, плазменные исследования и технологии» (Лаплаз -2020), Часть 1, 11 14 февраля 2020 г. Москва, 2020, с. 232-233 (463 с.)
- Т72. Волков, В.В. Поиск распределений частиц по размерам на примере модельных данных малоуглового рассеяния от полидисперсных систем сферических частиц / В.В. Волков, А.Е. Крюкова, П.В. Конарев // Сборник научных трудов VI международной конференции «Лазерные, плазменные исследования и технологии» (Лаплаз -2020), Часть 1, 11 14 февраля 2020 г. Москва, 2020, с. 180-181 (463 с.)
- Т73. Конарев, П.В. Методы восстановления отдельных компонент в белковых и липидных смесях по данным малоуглового рентгеновского рассеяния / П.В. Конарев // Сборник научных трудов VI международной конференции «Лазерные, плазменные исследования и технологии» (Лаплаз -2020), Часть 1, 11 14 февраля 2020 г. Москва, 2020, с. 230-231 (463 с.)
- Т74. Волков, В.В. Малоугловое рентгеновское и нейтронное рассеяние в исследовании структуры вещества: объекты, принципы, алгоритмы, программы /

- В. В. Волков, П. В. Конарев, М. В. Петухов // Сборник научных трудов 9-й международного научного семинара и 7-я международной молодежной научной школы-семинара "Современные методы анализа дифракционных данных и актуальные проблемы рентгеновской оптики", 1 11 июля 2020 г. Москва, 2020, с. 45 (164 с.)
- Т75. Григорьев, В.А. Исследование нативного и пегилированного фибрина методом малоуглового рентгеновского рассеяния / В.А. Григорьев, П.В. Конарев, В.В. Волков, В.С. Преснякова, П.Ю. Бикмулина, В.Е. Асадчиков, А.И. Шпичка, П.С. Тимашев // Сборник научных трудов 9-й международного научного семинара и 7-я международной молодежной научной школы-семинара "Современные методы анализа дифракционных данных и актуальные проблемы рентгеновской оптики", 1 11 июля 2020 г. Москва, 2020, с. 61-64 (164 с.)
- Т76. Конарев, П.В. Возможности метода малоуглового рентгеновского рассеяния для структурного анализа белковых и липидных смесей / П. В. Конарев // Сборник научных трудов 9-й международного научного семинара и 7-я международной молодежной научной школы-семинара "Современные методы анализа дифракционных данных и актуальные проблемы рентгеновской оптики", 1 11 июля 2020 г. Москва, 2020, с. 84-88 (164 с.)
- Т77. Крюкова, А.Е. Сопоставление эффективности модификаций программы mixture при анализе данных малоуглового рассеяния от раствора кремнезоля / А.Е. Крюкова, П.В. Конарев, В.В. Волков // Сборник научных трудов 9-й международного научного семинара и 7-я международной молодежной научной школы-семинара "Современные методы анализа дифракционных данных и актуальные проблемы рентгеновской оптики", 1 11 июля 2020 г. Москва, 2020, с. 95-96 (164 с.)
- Т78. Конарев, П.В. Устойчивость восстановления функции распределения частиц по размерам от системы полидисперсных сферических частиц по данным малоуглового рассеяния / П. В. Конарев, А.Е. Крюкова, В. В. Волков // Сборник

- тезисов LIV Школы ПИЯФ по физике конденсированного состояния (ФКС-2020), 16-21 марта 2020 г. Санкт-Петербург, 2020, с. 43 (211 с.)
- Т79. Волков, В.В. Методы моделирования экспериментальных данных малоуглового рассеяния от изотропных биологических систем / В. В. Волков, П. В. Конарев, М. В. Петухов // Сборник тезисов LIV Школы ПИЯФ по физике конденсированного состояния (ФКС-2020), 16 21 марта 2020 г. Санкт-Петербург, 2020, с. 179 (211 с.)
- Т80. Крюкова, А.Е. Устойчивость решений для полидисперсной системы вытянутых цилиндрических частиц по данным малоуглового рассеяния / А.Е. Крюкова, П.В. Конарев, В.В. Волков // Сборник научных трудов VII международной конференции «Лазерные, плазменные исследования и технологии» (Лаплаз -2021), Часть 1, 23 26 марта 2021 г. Москва, 2021, с. 247-248 (460 с.)
- Т81. Конарев, П.В. Восстановление профилей рассеяния отдельных компонент олигомерных белковых смесей по данным малоуглового рентгеновского рассеяния с онлайн хроматографической колонкой / П.В. Конарев, В.В. Волков, С.М. Jeffries, М. Graewert, Д.И. Свергун / Сборник тезисов XIV Международной научной конференции «Проблемы сольватации и комплексо-образования в растворах», 20-24 сентября 2021 г. Иваново, 2021, с. 20-21 (414 с.)
- T82. Kordyukova L. A new structural approach to study lipid-protein interactions within a viral envelope / L. Kordyukova, P. Konarev, N. Fedorova, E. Shtykova, A. Ksenofontov, O. Batishchev / Abstracts of the 8th ESWI Influenza Conference 2021 Virtual Edition, December 4-7, 2021 p. 207 (352 c.)
- Т83. Крюкова, А.Е. Сравнение эффективности восстановления решений, найденных рядом итеративных минимизационных методов, при анализе данных малоуглового рентгеновского рассеяния от раствора кремнезоля / А.Е. Крюкова, П.В. Конарев, В.В. Волков / Сборник тезисов XIV Международной научной конференции «Проблемы сольватации и комплексо-образования в растворах», 20-24 сентября 2021 г. Иваново, 2021, с. 102-103 (414 с.)

- T84. Samygina, V.R. SAXS-based control of microcrystallization for 4D enzymology experiments / V. Samygina, G. Peters, K. Dubova, P. Konarev / Abstracts of the 45th FEBS Congress, July 3–8, 2021 Ljubljana, Slovenia, 2021, Febs Open Bio V. 11 Issue S1 p. 235-236 (507 p.)
- T85. Lamba D. The Structure of the Pro-domain of Mouse proNGF in Contact with the NGF Domain / D. Lamba, P.V. Konarev, F.L. Gervasio, A. Pastore, A. Cattaneo, F. Paoletti / Abstracts of the 1st Conference on Crystallography, Structural Chemistry and Biosystems, October 4–6, 2021 Catania, Italy, 2021, Smart eLab V. 16 p. 64 (84 p.)
- T86. Covaceuszach S. Disclosing the Major Structural Determinants Essen tial for Proneurotrophins Biological Functions / S. Covaceuszach, L.Y. Peche, P.V. Konarev, D. Lamba / Abstracts of the 1st Conference on Crystallography, Structural Chemistry and Biosystems, October 4–6, 2021 Catania, Italy, 2021, Smart eLab V. 16 p. 18-19 (84 p.)
- Т87. К вопросу ab initio восстановления 3D формы и электронной плотности многокомпонентных систем по данным малоуглового рентгеновского рассеяния / П.В. Конарев, Д.И. Свергун / Сборник научных трудов III объединенного научного Форума физиологов, биохимиков и молекулярных биологов и X Российского симпозиума «Белки и Пептиды», Том 2, 3–8 октября 2021 г. Сочи, Дагомыс, 2021, с. 107 (313 с.)
- Т88. Конарев, П.В. Развитие кроссплатформенных программ с графическим интерфейсом для анализа данных малоуглового рентгеновского рассеяния / П.В. Конарев, В.В. Волков // Сборник научных трудов VIII международной конференции «Лазерные, плазменные исследования и технологии» (Лаплаз -2022), 22 25 марта 2022 г. Москва, 2022, с. 456 (463 с.)
- T89. Peters, G. The small-angle scattering beamline at the Kurchatov synchrotron radiation source current status and latest results / G. Peters, P. Konarev, Y. Gaponov, M. Marchenkova, K. Il'ina, V. Volkov, Y. Posarevskii, M. Kovalchuk / Abstracts of the International Conference "Synchrotron and Free Electron Laser Radiation: Generation

and Application" (SFR-2022), 27 - 30 июня 2022 г. — Новосибирск, 2022, — с. 29-30 (132 с.)

T90. Franke, D. ATSAS-3.1: An Overview of New Developments, / Franke, D., Markov, P., Borges, C., Mertens, H., Molodenskiy, D., Kikhney, A., Petoukhov, P., Konarev, P., Svergun, D. / Abstracts of XVIII International Small Angle Scattering Conference, 11-16 September 2022. — Campinas, Brasil, 2022, — P. 417 (448 P.)

Т91. Конарев, П.В. Разделение перекрывающихся вкладов в интенсивность малоуглового рассеяния от двух конформационных состояний фермента 3-изопропилмалат дегидрогеназы при использовании хроматографической колонки / П.В. Конарев, В.В. Волков / Сборник научных трудов ІХ международной конференции «Лазерные, плазменные исследования и технологии» (Лаплаз -2023), 28 - 31 марта 2023 г. — Москва, 2023, — с. 427 (440 с.)

Т92. Конарев, П.В. Оценка эффективности *ab initio* методов восстановления формы макромолекулярных комплексов с неоднородной электронной плотностью по данным малоуглового рассеяния / П.В. Конарев, В.В. Волков / Сборник тезисов VI конференции по малоугловому рассеянию и рефлектометрии (МУРомец-2023), 21 - 23 июня 2023 г. — Гатчина, 2023, — с. 22 (57 с.)

Цитируемая литература

- 1. Small Angle Scattering of X-Rays. / Guinier A., Fournet G. New York: Wiley, 1955.
- 2. Small Angle X-ray Scattering. / Glatter O., Kratky O. London: Academic Press, 1982. 515 P.
- 3. Рентгеновское и нейтронное малоугловое рассеяние. / Свергун Д. И., Фейгин, Л.А. Москва: "Наука" Главная редакция физико-математической литературы, 1986. 278 С.
- 4. Small angle x-ray and neutron scattering from solutions of biological macromolecules. International Union of Crystallography monographs on crystallography. / Svergun D. I., Koch M. H. J., Timmins P. A., May R. P.: Oxford University Press, 2013. T. 19: International Union of Crystallography monographs on crystallography. 358 P.
- 5. Guinier A. La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques // Ann. Phys. (Paris). 1939. V. 12. P. 161-237.
- 6. Stuhrmann H. B., Kirste R. G. Elimination der intrapartikulaeren Untergrundstreuung bei der Roentgenkleinwinkelstreuung am kompakten Teilchen (Proteinen) // Zeitschr. Physik. Chem. Neue Folge. 1965. V. 46. P. 247-250.
- 7. Ibel K., Stuhrmann H. B. Comparison of neutron and X-ray scattering of dilute myoglobin solutions // J Mol Biol. 1975. V. 93, № 2. P. 255-265.
- 8. Pardon J. F., Worcester D. L., Wooley J. C., Tatchell K., Van Holde K. E., Richards B. M. Low-angle neutron scattering from chromatin subunit particles // Nucleic Acids Res. -1975. V. 2, No 11. P. 2163-2176.
- 9. Luger K., Mader A. W., Richmond R. K., Sargent D. F., Richmond T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution // Nature. − 1997. − V. 389, № 6648. − P. 251-260.
- 10. Capel M. S., Engelman D. M., Freeborn B. R., Kjeldgaard M., Langer J. A., Ramakrishnan V., Schindler D. G., Schneider D. K., Schoenborn B. P., Sillers I. Y., et al. A complete mapping of the proteins in the small ribosomal subunit of Escherichia coli // Science. − 1987. − V. 238, № 4832. − P. 1403-1406.

- 11. May R. P., Nowotny V., Nowotny P., Voss H., Nierhaus K. H. Inter-protein distances within the large subunit from Escherichia coli ribosomes // Embo J. 1992. V. 11, N_{\odot} 1. P. 373-378.
- 12. Ban N., Nissen P., Hansen J., Moore P. B., Steitz T. A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution // Science. 2000. V. 289, № 5481. P. 905-920.
- 13. Jacques D. A., Trewhella J. Small-angle scattering for structural biology--expanding the frontier while avoiding the pitfalls // Protein Sci. -2010. -V. 19, N_2 4. -P. 642-657.
- 14. Neutrons in Soft Matter. / Imae T., Kanaya T., Furusaka M., Torikai N. New Jersy: John Wiley and Sons, 2011. 654 P.
- 15. Lake J. A. An iterative method of slit-correcting small angle X-ray data // Acta Cryst. 1967. V. 23. P. 191-194.
- 16. Vonk C. A procedure for desmearing X-ray small-angle scattering curves // J. Appl Cryst. 1971. V. 4. P. 340-342.
- 17. Glatter O. A new iterative method for collimation correction in small-angle scattering // J.Appl.Cryst. 1974. V. 7. P. 147-153.
- 18. Mogilevsky L. Y., Dembo A. T., Svergun D. I., Feigin L. A. Small-angle scattering diffractometer with a coordinate detector // Cryst. Rep. 1984. V. 29. P. 587-591.
- 19. Skarzynski T. Collecting data in the home laboratory: evolution of X-ray sources, detectors and working practices // Acta Crystallogr D Biol Crystallogr. 2013. V. 69, № Pt 7. P. 1283-1288.
- 20. Bruetzel L. K., Fischer S., Salditt A., Sedlak S. M., Nickel B., Lipfert J. A Mo-anode-based in-house source for small-angle X-ray scattering measurements of biological macromolecules // Rev Sci Instrum. − 2016. − V. 87, № 2. − P. 025103.
- 21. Hemberg O., Otendal M., Hertz H. M. Liquid-metal-jet anode electron-impact X-ray source // Appl. Phys. Lett. 2003. V. 83. P. 1483–1485
- 22. Larsson D. H., Takman P. A., Lundstrom U., Burvall A., Hertz H. M. A 24 keV liquid-metal-jet x-ray source for biomedical applications // Rev Sci Instrum. 2011. V. 82, № 12. P. 123701.

- 23. Blanchet C. E., Spilotros A., Schwemmer F., Graewert M. A., Kikhney A., Jeffries C. M., Franke D., Mark D., Zengerle R., Cipriani F., Fiedler S., Roessle M., Svergun D. I. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY) // J Appl Cryst. 2015. V. 48. P. 431-443.
- 24. Kirby N. M., Mudie S. T., Hawley A. M., Cookson D. J., Mertens H. D. T., Cowieson N., Samardzic-Boban V. A low-background-intensity focusing small-angle X-ray scattering undulator beamline // J.Appl.Cryst. 2013. V. 46. P. 1670-1680.
- 25. Hura G. L., Menon A. L., Hammel M., Rambo R. P., Poole F. L., 2nd, Tsutakawa S. E., Jenney F. E., Jr., Classen S., Frankel K. A., Hopkins R. C., Yang S. J., Scott J. W., Dillard B. D., Adams M. W., Tainer J. A. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS) // Nat Methods. − 2009. − V. 6, № 8. − P. 606-612.
- 26. Pernot P., Round A., Barrett R., De Maria Antolinos A., Gobbo A., Gordon E., Huet J., Kieffer J., Lentini M., Mattenet M., Morawe C., Mueller-Dieckmann C., Ohlsson S., Schmid W., Surr J., Theveneau P., Zerrad L., McSweeney S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution // J Synchrotron Radiat. − 2013. − V. 20, № Pt 4. − P. 660-664.
- 27. David G., Perez J. Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline // J. Appl. Cryst. -2009. V. 42, N_{\odot} 5. -P. 892-900.
- 28. Yoon J., Kim K.-W., Kim J., Heo K., Jin K. S., Jin S., Shin T. J., Lee B., Rho Y., Ahn B., Ree M. Small-angle x-ray scattering station 4C2 BL of pohang accelerator laboratory for advance in Korean polymer science // Macromolecular Research. 2008. V. 16. P. 575-585.
- 29. Fischetti R., Stepanov S., Rosenbaum G., Barrea R., Black E., Gore D., Heurich R., Kondrashkina E., Kropf A. J., Wang S., Zhang K., Irving T. C., Bunker G. B. The BioCAT undulator beamline 18ID: a facility for biological non-crystalline diffraction and

- X-ray absorption spectroscopy at the Advanced Photon Source // J Synchrotron Radiat. -2004. V. 11, No Pt 5. -P. 399-405.
- 30. Fujisawa T., Inoue K., Oka T., Iwamoto H., Uruga T., Kumasaka T., Inoko Y., Yagi N., Yamamoto M., Ueki T. Small-angle X-ray scattering station at the SPring-8 RIKEN beamline // J Appl Cryst. 2000. V. 33. P. 797-800.
- 31. Cowieson N. P., Edwards-Gayle C. J. C., Inoue K., Khunti N. S., Doutch J., Williams E., Daniels S., Preece G., Krumpa N. A., Sutter J. P., Tully M. D., Terrill N. J., Rambo R. P. Beamline B21: high-throughput small-angle X-ray scattering at Diamond Light Source // J Synchrotron Radiat. − 2020. − V. 27, № Pt 5. − P. 1438-1446.
- 32. Liu G., Li Y., Wu H., Wu X., Xu X., Wang W., Zhang R., N. L. Upgraded SSRF BL19U2 beamline for small-angle X-ray scattering of biological macromolecules in solution // J.Appl.Cryst. 2018. V. 51. P. 1633-1640.
- 33. Porod G. General theory // Small-angle X-ray scattering / Glatter O., Kratky O. London: Academic Press, 1982. P. 17-51.
- 34. Mémoire sur les séries et sur l'intégration complète d'une équation aux différences partielles linéaire du second ordre, à coefficients constants. / Parseval des Chênes M.-A. Paris, 1799.
- 35. Mittelbach P., Porod G. Die Röntgenkleinwinkelstreuung verdünnter kolloider Systeme. VI. Die Berechnung der Streukurven von elliptischen Zylindern und Hohizylindern. // Acta Phys. Austr. 1961. V. 14. P. 405-439.
- 36. Mittelbach P., Porod G. Zur Röntgenkleinwinkelstreuung verduennter kolloider System. VII. Die Berechnung der Streukurven von dreiachsigen Ellipsoiden. // Acta Phys. Austr. 1962. V. 15. P. 122-147.
- 37. Goodisman J. The correlation function, intersect distribution and scattering from a cube // J. Appl. Cryst. -1980. V. 13. P. 132-134.
- 38. Gille W. The small-angle scattering structure functions of the single tetrahedron. // J. Appl. Cryst. -2003.-V. 36. -P. 850-853.
- 39. Glatter O. A new method for the evaluation of small-angle scattering data // J. Appl. Cryst. -1977. V. 10. -P. 415-421.

- 40. Svergun D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria // J. Appl. Cryst. 1992. V. 25. P. 495-503.
- 41. Ilavsky J., Jemian P. J. Irena: tool suite for modeling and analysis of small-angle scattering // J.Appl.Cryst. 2009. V. 42. P. 347-353.
- 42. Kratky O., Porod G. Röntgenuntersuchung gelöster Fadenmoleküle // Recueil Des Travaux Chimiques Des Pays-Bas. 1949. V. 68. P. 1106–1122.
- 43. Doniach S. Changes in biomolecular conformation seen by small angle X-ray scattering // Chem Rev. -2001. V. 101, No 6. P. 1763-1778.
- 44. Perez J., Vachette P., Russo D., Desmadril M., Durand D. Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering // J Mol Biol. -2001.-V.~308, No.~4.-P.~721-743.
- 45. Durand D., Vives C., Cannella D., Perez J., Pebay-Peyroula E., Vachette P., Fieschi F. NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers // J Struct Biol. − 2010. − V. 169, № 1. − P. 45-53.
- 46. Stuhrmann H. B. Ein neues Verfahren zur Bestimmung der Oberflaechenform und der inneren Struktur von geloesten globularen Proteinen aus Roentgenkleinwinkelmessungen. // Zeitschr. Physik. Chem. Neue Folge. 1970. V. 72. P. 177-198.
- 47. Stuhrmann H. B. Interpretation of small-angle scattering of dilute solutions and gases. A representation of the structures related to a one-particle scattering functions // Acta Cryst. 1970. V. A26. P. 297-306.
- 48. Svergun D. I., Volkov V. V., Kozin M. B., Stuhrmann H. B., Barberato C., Koch M. H. J. Shape determination from solution scattering of biopolymers // J. Appl. Cryst. 1997. V. 30. P. 798-802.
- 49. Chacon P., Moran F., Diaz J. F., Pantos E., Andreu J. M. Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm // Biophys J. 1998. V. 74, No 6. P. 2760-2775.

- 50. Bada M., Walther D., Arcangioli B., Doniach S., Delarue M. Solution structural studies and low-resolution model of the Schizosaccharomyces pombe sap1 protein // J Mol Biol. 2000. V. 300, № 3. P. 563-574.
- 51. Vigil D., Gallagher S. C., Trewhella J., Garcia A. E. Functional dynamics of the hydrophobic cleft in the N-domain of calmodulin // Biophys J. − 2001. − V. 80, № 5. − P. 2082-2092.
- 52. Svergun D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing // Biophys J. − 1999. − V. 76, № 6. − P. 2879-2886.
- 53. Svergun D. I., Petoukhov M. V., Koch M. H. J. Determination of domain structure of proteins from X-ray solution scattering // Biophys J. − 2001. − V. 80, № 6. − P. 2946-2953.
- 54. Heller W. T., Abusamhadneh E., Finley N., Rosevear P. R., Trewhella J. The solution structure of a cardiac troponin C-troponin I-troponin T complex shows a somewhat compact troponin C interacting with an extended troponin I-troponin T component // Biochemistry. -2002. V. 41, No. 52. P. 15654-15663.
- 55. Franke D., Svergun D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering // J. Appl. Cryst. 2009. V. 42. P. 342-346.
- 56. Koutsioubas A., Pérez J. Incorporation of a hydration layer in the `dummy atom' ab initio structural modelling of biological macromolecules // J.Appl.Cryst. 2013. V. 46. P. 1884-1888.
- 57. Kozin M. B., Svergun D. I. Automated matching of high- and low-resolution structural models // J. Appl. Cryst. 2001. V. 34. P. 33-41.
- 58. Volkov V. V., Svergun D. I. Uniqueness of ab initio shape determination in small angle scattering // J. Appl. Cryst. 2003. V. 36. P. 860-864.
- 59. Kirkpatrick S., Gelatt C. D., Jr., Vecci M. P. Optimization by simulated annealing // Science. 1983. V. 220. P. 671-680.
- 60. Petoukhov M. V., Svergun D. I. Ambiguity assessment of small-angle scattering curves from monodisperse systems // Acta Cryst. D. 2015. V. 71. P. 1051-1058.

- 61. Volkov V. V. On the Tactics of Ab Initio Search for the Shape of Protein Particles from Small-Angle X-Ray Scattering Data // Cryst. Rep. 2021. V. 66. P. 819-827.
- 62. Saxton W. O., Baumeister W. The correlation averaging of a regularly arranged bacterial cell envelope protein // J Microsc. − 1982. − V. 127, № Pt 2. − P. 127-138.
- 63. Tuukkanen A. T., Kleywegt G. J., Svergun D. I. Resolution of ab initio shapes determined from small-angle scattering // IUCrJ. 2016. V. 3, № Pt 6. P. 440-447.
- 64. Franke D., Kikhney A. G., Svergun D. I. Automated acquisition and analysis of small angle X-ray scattering data // Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment. 2012. V. 689. P. 52-59.
- 65. Brennich M. E., Kieffer J., Bonamis G., De Maria Antolinos A., Hutin S., Pernot P., Round A. Online data analysis at the ESRF bioSAXS beamline, BM29 // J.Appl.Cryst. 2016. V. 49. P. 203-212.
- 66. Svergun D. I., Ekstrom F., Vandegriff K. D., Malavalli A., Baker D. A., Nilsson C., Winslow R. M. Solution structure of poly(ethylene) glycol-conjugated hemoglobin revealed by small-angle X-ray scattering: implications for a new oxygen therapeutic // Biophys J. -2008. -V. 94, Nolution 1. -P. 173-181.
- 67. Yadavalli S. S., Klipcan L., Zozulya A., Banerjee R., Svergun D., Safro M., Ibba M. Large-scale movement of functional domains facilitates aminoacylation by human mitochondrial phenylalanyl-tRNA synthetase // FEBS Lett. − 2009. − V. 583, № 19. − P. 3204-3208.
- 68. Cheng C. Y., Yang J., Taylor S. S., Blumenthal D. K. Sensing domain dynamics in protein kinase A-Ialpha complexes by solution X-ray scattering // J Biol Chem. 2009. V. 284, № 51. P. 35916-35925.
- 69. Schmidt C. Q., Herbert A. P., Mertens H. D., Guariento M., Soares D. C., Uhrin D., Rowe A. J., Svergun D. I., Barlow P. N. The central portion of factor H (modules 10-15) is compact and contains a structurally deviant CCP module // J Mol Biol. -2010. V. 395, Nole 1. P. 105-122.

- 70. Svergun D. I., Barberato C., Koch M. H. J. CRYSOL a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates // J. Appl. Cryst. 1995. V. 28. P. 768-773.
- 71. Schneidman-Duhovny D., Hammel M., Tainer J. A., Sali A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles // Nucleic Acids Res. − 2016. − V. 44, № W1. − P. W424-W429.
- 72. Grishaev A., Guo L. A., Irving T., Bax A. Improved Fitting of Solution X-ray Scattering Data to Macromolecular Structures and Structural Ensembles by Explicit Water Modeling // Journal of the American Chemical Society. − 2010. − V. 132, № 44. − P. 15484-15486.
- 73. Poitevin F., Orland H., Doniach S., Koehl P., Delarue M. AquaSAXS: a web server for computation and fitting of SAXS profiles with non-uniformally hydrated atomic models // Nucleic Acids Research. − 2011. − V. 39, № suppl 2. − P. W184-W189.
- 74. Virtanen J. J., Makowski L., Sosnick T. R., Freed K. F. Modeling the hydration layer around proteins: applications to small- and wide-angle x-ray scattering // Biophys J. -2011. V. 101, No. 8. P. 2061-2069.
- 75. Svergun D. I., Richard S., Koch M. H. J., Sayers Z., Kuprin S., Zaccai G. Protein hydration in solution: experimental observation by x-ray and neutron scattering // Proc Natl Acad Sci U S A. − 1998. − V. 95, № 5. − P. 2267-2272.
- 76. Yang S., Park S., Makowski L., Roux B. A rapid coarse residue-based computational method for x-ray solution scattering characterization of protein folds and multiple conformational states of large protein complexes // Biophys J. − 2009. − V. 96, № 11. − P. 4449-4463.
- 77. Chen P. C., Hub J. S. Validating solution ensembles from molecular dynamics simulation by wide-angle X-ray scattering data // Biophys J. -2014. -V. 107, No. 2. -P. 435-447.

- 78. Wriggers W., Chacón P. Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering // J.Appl.Cryst. 2001. V. 34. P. 773-776.
- 79. Kozin M. B., Volkov V. V., Svergun D. I. ASSA a program for three-dimensional rendering in solution scattering from biopolymers // J. Appl. Cryst. 1997. V. 30. P. 811-815.
- 80. Boehm M. K., Woof J. M., Kerr M. A., Perkins S. J. The Fab and Fc fragments of IgA1 exhibit a different arrangement from that in IgG: a study by X-ray and neutron solution scattering and homology modelling // J Mol Biol. − 1999. − V. 286, № 5. − P. 1421-1447.
- 81. Petoukhov M. V., Svergun D. I. Global rigid body modeling of macromolecular complexes against small-angle scattering data // Biophys J. -2005. V. 89, N 2. P. 1237-1250.
- 82. Sun Z., Reid K. B., Perkins S. J. The dimeric and trimeric solution structures of the multidomain complement protein properdin by X-ray scattering, analytical ultracentrifugation and constrained modelling // J Mol Biol. − 2004. − V. 343, № 5. − P. 1327-1343.
- 83. Webb B., Sali A. Protein structure modeling with MODELLER // Methods Mol Biol. 2014. V. 1137. P. 1-15.
- 84. Yoon T., Cowan J. A. Iron-sulfur cluster biosynthesis. Characterization of frataxin as an iron donor for assembly of [2Fe-2S] clusters in ISU-type proteins // J Am Chem Soc. 2003. V. 125, № 20. P. 6078-6084.
- 85. Cavadini P., O'Neill H. A., Benada O., Isaya G. Assembly and iron-binding properties of human frataxin, the protein deficient in Friedreich ataxia // Hum Mol Genet. − 2002. − V. 11, № 3. − P. 217-227.
- 86. Petoukhov M. V., Eady N. A., Brown K. A., Svergun D. I. Addition of missing loops and domains to protein models by x-ray solution scattering // Biophys J. -2002. V. 83, $N_{\odot} 6. P. 3113-3125$.

- 87. Zhou H. X. Polymer models of protein stability, folding, and interactions // Biochemistry. -2004. -V. 43, N_{\odot} 8. -P. 2141-2154.
- 88. Bernado P., Blackledge M. Structural biology: Proteins in dynamic equilibrium // Nature. 2010. V. 468, № 7327. P. 1046-1048.
- 89. Bernado P., Mylonas E., Petoukhov M. V., Blackledge M., Svergun D. I. Structural characterization of flexible proteins using small-angle X-ray scattering // J Am Chem Soc. 2007. V. 129, № 17. P. 5656-5664.
- 90. Tria G., Mertens H. D., Kachala M., Svergun D. I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering // IUCrJ. − 2015. − V. 2, № Pt 2. − P. 207-217.
- 91. Pelikan M., Hura G. L., Hammel M. Structure and flexibility within proteins as identified through small angle X-ray scattering // Gen Physiol Biophys. -2009. V. 28, No. 2. P. 174-189.
- 92. Yang S., Blachowicz L., Makowski L., Roux B. Multidomain assembled states of Hck tyrosine kinase in solution // Proc Natl Acad Sci U S A. − 2010. − V. 107, № 36. − P. 15757-15762.
- 93. Rozycki B., Kim Y. C., Hummer G. SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions // Structure. 2011. V. 19, № 1. P. 109-116.
- 94. Krzeminski M., Marsh J. A., Neale C., Choy W. Y., Forman-Kay J. D. Characterization of disordered proteins with ENSEMBLE // Bioinformatics. -2013.-V. 29, No. 3.-P. 398-399.
- 95. Nodet G., Salmon L., Ozenne V., Meier S., Jensen M. R., Blackledge M. Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings // J Am Chem Soc. -2009. V. 131, \mathbb{N}° 49. P. 17908-17918.
- 96. Moller M., Nielsen S. S., Ramachandran S., Li Y., Tria G., Streicher W., Petoukhov M. V., Cerione R. A., Gillilan R. E., Vestergaard B. Small angle X-ray scattering studies of mitochondrial glutaminase C reveal extended flexible regions, and link oligomeric state with enzyme activity // PLoS One. -2013. V. 8, No. 9. P. e74783.

- 97. Bernado P. Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering // Eur Biophys J. -2010. -V. 39, N = 5. -P. 769-780.
- 98. Svergun D. I., Volkov V. V., Kozin M. B., Stuhrmann H. B. New developments in direct shape determination from small-angle scattering 2. Uniqueness. // Acta Crystallogr. 1996. V. A52. P. 419-426.
- 99. Chacon P., Diaz J. F., Moran F., Andreu J. M. Reconstruction of protein form with X-ray solution scattering and a genetic algorithm // J Mol Biol. 2000. V. 299, № 5. P. 1289-1302.
- 100. Svergun D. I., Nierhaus K. H. A map of protein-rRNA distribution in the 70 S Escherichia coli ribosome // J Biol Chem. 2000. V. 275, № 19. P. 14432-14439.
- 101. Niemann H. H., Petoukhov M. V., Hartlein M., Moulin M., Gherardi E., Timmins P., Heinz D. W., Svergun D. I. X-ray and neutron small-angle scattering analysis of the complex formed by the Met receptor and the Listeria monocytogenes invasion protein InlB // J Mol Biol. -2008. V. 377, No. 2. P. 489-500.
- 102. Dingwall C., Dilworth S. M., Black S. J., Kearsey S. E., Cox L. S., Laskey R. A. Nucleoplasmin cDNA sequence reveals polyglutamic acid tracts and a cluster of sequences homologous to putative nuclear localization signals // EMBO J. -1987. V. 6, $N_{\odot} 1. P. 69-74$.
- 103. Grant T. D. Ab initio electron density determination directly from solution scattering data // Nat Methods. -2018. V. 15, No. 2. P. 191-193.
- 104. Fienup J. R. Reconstruction of an object from the modulus of its Fourier transform // Opt Lett. -1978. V. 3, № 1. P. 27-29.
- 105. Marchesini S., He H., Chapman H. N., Hau-Riege S. P., Noy A., Howells M. R., Weierstall U., Spence J. C. H. X-ray image reconstruction from a diffraction pattern alone // Phys Rev B. 2003. V. 68. P. 140101.
- 106. Jacques D. A., Guss J. M., Svergun D. I., Trewhella J. Publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution // Acta Crystallographica Section D. -2012.-V.68, No.60.-P.620-626.

- 107. Skou S., Gillilan R. E., Ando N. Synchrotron-based small-angle X-ray scattering of proteins in solution // Nat Protoc. 2014. V. 9, № 7. P. 1727-1739.
- 108. Wlodawer A., Minor W., Dauter Z., Jaskolski M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures // FEBS J. -2008. -V. 275, N0 1. -P. 1-21.
- 109. Valentini E., Kikhney A. G., Previtali G., Jeffries C. M., Svergun D. I. SASBDB, a repository for biological small-angle scattering data // Nucleic Acids Res. 2015. V. 43, № Database issue. P. D357-D363.
- 110. Kikhney A. G., Borges C. R., Molodenskiy D. S., Jeffries C. M., Svergun D. I. SASBDB: Towards an automatically curated and validated repository for biological scattering data // Protein Sci. -2020. V. 29, No. 1. P. 66-75.
- 111. The mathematical theory of communication. / Shannon C. E., Weaver W. Urbana: University of Illinois Press, 1949.
- 112. Damaschun G., Mueller J. J., Puerschel H. V. Ueber die Messstrategie bei der Untersuchung der Roentgen-Kleinwinkelstreuung von verduennten monodispersen Loesungen von Makromolekuelen // Monatshefte fuer Chemie. 1968. V. 99. P. 2343-2348.
- 113. Taupin D., Luzzati V. Informational content and retrieval in solution scattering studies I. Degrees of freedom and data reduction. // J. Appl. Cryst. 1982. V. 15. P. 289-300.
- 114. Moore P. B. Small-angle scattering: Information content and error analysis // J. Appl. Cryst. 1980. V. 13. P. 168-175.
- 115. Rambo R. P., Tainer J. A. Accurate assessment of mass, models and resolution by small-angle scattering // Nature. -2013. V.496, No 97446. P.477-481.
- 116. Grant T. D. Describing small-angle scattering profiles by a limited set of intensities // J Appl Cryst. − 2022. − V. 55, № Pt 5. − P. 1116-1124.
- 117. Solving Least Squares Problems. / Lawson C. L., Hanson R. J. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1974.

- 118. Schwartz L. Transformation de Laplace des distributions // Comm. Sém. Math. Univ. Lund. 1952. V. Tome suppl. P. 196-206.
- 119. Frieden B. R. Evaluation, design and extrapolation methods for optical signals, based on the use of the prolate functions // Progress in Optics / Wolf E. Amsterdam: North Holland, 1971. P. 312-407.
- 120. Golub G. H., Reinsh C. Singular Value Decomposition and Least Squares Solution // Numer. Math. 1970. V. 14. P. 403-420.
- 121. Moglianetti M., Ong Q. K., Reguera J., Harkness K. M., Mameli M., Radulescu A., Kohlbrecher J., Jud C., Svergun D. I., Stellacci F. Scanning tunneling microscopy and small angle neutron scattering study of mixed monolayer protected gold nanoparticles in organic solvents // Chemical Science. − 2014. − V. 5, № 3. − P. 1232-1240.
- 122. Proteins: structure and function. / Whitford D. New Jersy: John Wiley & Sons 2005. 534 p.
- 123. Physical biochemistry: principles and applications. / Sheehan D. New Jersy: John Wiley & Sons 2009. 433 p.
- 124. Putnam C. D., Hammel M., Hura G. L., Tainer J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution // Q Rev Biophys. -2007. V. 40, No. 3. P. 191-285.
- 125. Fischetti R. F., Rodi D. J., Gore D. B., Makowski L. Wide-angle X-ray solution scattering as a probe of ligand-induced conformational changes in proteins // Chem Biol. -2004. V. 11, No 10. P. 1431-1443.
- 126. Canady M. A., Tsuruta H., Johnson J. E. Analysis of rapid, large-scale protein quaternary structural changes: time-resolved X-ray solution scattering of Nudaurelia capensis omega virus (NomegaV) maturation // J Mol Biol. -2001. V.311, NomegaV = 4. P. 803-814.
- 127. Statistics: an Introduction. / Larson H. J. New York: John Wiley, 1975. 328-329 P.

- 128. Schulz G. V. Ueber die Beziehung zwischen Reaktiongeschwindigkeit und Zusammensetzung des Reaktionproduktes Macropolymerisationsvorgaemgen // Z. Phys. Chem. Abt. B. 1935. V. 30. P. 379-398.
- 129. Percus J. K., Yevick G. J. Analysis of Classical Statistical Mechanics by Means of Collective Coordinates // Phys. Rev. 1958. V. 110. P. 1-13.
- 130. Baxter R. J. Percus—Yevick Equation for Hard Spheres with Surface Adhesion // J. Chem. Phys. 1968. V. 49. P. 2770-2774.
- 131. Practical Optimisation. / Gill P. E., Murray W., Wright M. H.: London:Academic Press, 1981.
- 132. Bressler I., Kohlbrecher J., Thunemann A. F. SASfit: a tool for small-angle scattering data analysis using a library of analytical expressions // J. Appl. Cryst. 2015. V. 48. P. 1587-1598.
- 133. Forster S., Fischer S., Zielske K., Schellbach C., Sztucki M., Lindner P., Perlich J. Calculation of scattering-patterns of ordered nano- and mesoscale materials // Adv Colloid Interface Sci. − 2011. − V. 163, № 1. − P. 53-83.
- 134. Ben-Nun T., Ginsburg A., Szekely P., Raviv U. X+: a comprehensive computationally accelerated structure analysis tool for solution X-ray scattering from supramolecular self-assemblies // J. Appl. Cryst. -2010. V. 43, No. 6. P. 1522-1531.
- 135. Bressler I., Pauw B. R., Thunemann A. F. McSAS: software for the retrieval of model parameter distributions from scattering patterns // J Appl Cryst. -2015. -V. 48, N_{\odot} Pt 3. -P. 962-969.
- 136. Volkov V. V. VOLDIS: A Direct Search for the Size Distribution of Nanoparticles from Small-Angle Scattering Data and Stability of the Solution // Crystals. 2022. V. 12. P. 1659.
- 137. Matte A., Tari L. W., Delbaere L. T. How do kinases transfer phosphoryl groups? // Structure. -1998. V. 6, No. 4. P. 413-419.
- 138. Banks R. D., Blake C. C., Evans P. R., Haser R., Rice D. W., Hardy G. W., Merrett M., Phillips A. W. Sequence, structure and activity of phosphoglycerate kinase: a possible hinge-bending enzyme // Nature. − 1979. − V. 279, № 5716. − P. 773-777.

- 139. Harlos K., Vas M., Blake C. F. Crystal structure of the binary complex of pig muscle phosphoglycerate kinase and its substrate 3-phospho-D-glycerate // Proteins. 1992. V. 12, № 2. P. 133-144.
- 140. Davies G. J., Gamblin S. J., Littlechild J. A., Dauter Z., Wilson K. S., Watson H. C. Structure of the ADP complex of the 3-phosphoglycerate kinase from Bacillus stearothermophilus at 1.65 A // Acta Crystallogr D Biol Crystallogr. − 1994. − V. 50, № Pt 2. − P. 202-209.
- 141. Krishnan P., Fu Q., Lam W., Liou J. Y., Dutschman G., Cheng Y. C. Phosphorylation of pyrimidine deoxynucleoside analog diphosphates: selective phosphorylation of L-nucleoside analog diphosphates by 3-phosphoglycerate kinase // J Biol Chem. -2002. -V. 277, N0 7. -P. 5453-9.
- 142. Hayashi-Iwasaki Y., Oshima T. Purification and characterization of recombinant 3-isopropylmalate dehydrogenases from Thermus thermophilus and other microorganisms // Methods Enzymol. 2000. V. 324. P. 301-322.
- 143. Walz J., Tamura T., Tamura N., Grimm R., Baumeister W., Koster A. J. Tricorn protease exists as an icosahedral supermolecule in vivo // Mol Cell. 1997. V. 1, № 1. P. 59-65.
- 144. Ubbink M., Bendall D. S. Complex of plastocyanin and cytochrome c characterized by NMR chemical shift analysis // Biochemistry. 1997. V. 36, № 21. P. 6326-6335.
- 145. Worrall J. A. R., Reinle W., Bernhardt R., Ubbink M. Transient protein interactions studied by NMR spectroscopy: the case of cytochrome C and adrenodoxin. // Biochemistry. 2003. V. 42. P. 7068-7076.
- 146. Grinberg A. V., Hannemann F., Schiffler B., Muller J., Heinemann U., Bernhardt R. Adrenodoxin: structure, stability, and electron transfer properties // Proteins. 2000. V. 40, № 4. P. 590-612.
- 147. Müller A., Müller J. J., Muller Y. A., Uhlmann H., Bernhardt R., Heinemann U. New aspects of electron transfer revealed by the crystal structure of a truncated bovine adrenodoxin, Adx(4-108) // Structure. 1998. V. 6. P. 269-280.

- 148. Gebauer D., Kellermeier M., Gale J. D., Bergstrom L., Colfen H. Pre-nucleation clusters as solute precursors in crystallisation // Chem Soc Rev. 2014. V. 43, № 7. P. 2348-2371.
- 149. Askhabov A. M. Cluster (quatarone) self-organization of matter at the nanoscale level and the formation of crystalline and non-crystalline materials // Zap. Ross. Mineral. O-va –2004. V. 133. P. 108-123.
- 150. Huber R., Bennett W. S., Jr. Functional significance of flexibility in proteins // Biopolymers. 1983. V. 22, № 1. P. 261-279.
- 151. Caspar D. L., Klug A. Physical principles in the construction of regular viruses // Cold Spring Harb Symp Quant Biol. 1962. V. 27. P. 1-24.
- 152. Johnson J. E., Speir J. A. Quasi-equivalent viruses: a paradigm for protein assemblies // J Mol Biol. -1997. V. 269, No. 5. P. 665-675.
- 153. Bacher A., Mailander B. Biosynthesis of riboflavin in Bacillus subtilis: function and genetic control of the riboflavin synthase complex // J Bacteriol. -1978. V. 134, № 2. -P. 476-482.
- 154. Persson K., Schneider G., Jordan D. B., Viitanen P. V., Sandalova T. Crystal structure analysis of a pentameric fungal and an icosahedral plant lumazine synthase reveals the structural basis for differences in assembly // Protein Sci. 1999. V. 8, N0 11. P. 2355-2365.
- 155. Meining W., Mortl S., Fischer M., Cushman M., Bacher A., Ladenstein R. The atomic structure of pentameric lumazine synthase from Saccharomyces cerevisiae at 1.85 A resolution reveals the binding mode of a phosphonate intermediate analogue // J Mol Biol. -2000. V. 299, No 1. P. 181-197.
- 156. Sayle R. A., Milner-White E. J. RASMOL: biomolecular graphics for all. // Trends Biochem Sci. 1995. V. 20(9). P. 374-376.
- 157. Bourgaux C., Couvreur P. Interactions of anticancer drugs with biomembranes: what can we learn from model membranes? // J Control Release. 2014. V. 190. P. 127-138.

- 158. Malam Y., Loizidou M., Seifalian A. M. Liposomes and nanoparticles: nanosized vehicles for drug delivery in cancer // Trends Pharmacol Sci. 2009. V. 30, № 11. P. 592-599.
- 159. Schilt Y., Berman T., Wei X., Barenholz Y., Raviv U. Using solution X-ray scattering to determine the high-resolution structure and morphology of PEGylated liposomal doxorubicin nanodrugs // Biochim Biophys Acta. − 2016. V. 1860, № 1 Pt A. P. 108-119.
- 160. Josts I., Nitsche J., Maric S., Mertens H. D., Moulin M., Haertlein M., Prevost S., Svergun D. I., Busch S., Forsyth V. T., Tidow H. Conformational States of ABC Transporter MsbA in a Lipid Environment Investigated by Small-Angle Scattering Using Stealth Carrier Nanodiscs // Structure. − 2018. − V. 26, № 8. − P. 1072-1079.
- 161. Flayhan A., Mertens H. D. T., Ural-Blimke Y., Molledo M. M., Svergun D. I., Low C. Saposin Lipid Nanoparticles: A Highly Versatile and Modular Tool for Membrane Protein Research // Structure. − 2018. − V. 26, № 2. − P. 345-355.
- 162. Pabst G., Rappolt M., Amenitsch H., Laggner P. Structural information from multilamellar liposomes at full hydration: full q-range fitting with high quality x-ray data // Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics. -2000. V. 62, No 3 Pt B. -P. 4000-4009.
- 163. Pencer J., Hallett F. R. Small-angle neutron scattering from large unilamellar vesicles: an improved method for membrane thickness determination // Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics. -2000. V. 61, No. 3. P. 3003-3008.
- 164. Perez J., Koutsioubas A. Memprot: a program to model the detergent corona around a membrane protein based on SEC-SAXS data // Acta Crystallogr D Biol Crystallogr. 2015. V. 71, № Pt 1. P. 86-93.
- 165. Skar-Gislinge N., Kynde S. A., Denisov I. G., Ye X., Lenov I., Sligar S. G., Arleth L. Small-angle scattering determination of the shape and localization of human cytochrome P450 embedded in a phospholipid nanodisc environment // Acta Crystallogr D Biol Crystallogr. − 2015. − V. 71, № Pt 12. − P. 2412-2421.

- 166. Heftberger P., Kollmitzer B., Heberle F. A., Pan J. J., Rappolt M., Amenitsch H., Kucerka N., Katsaras J., Pabst G. Global small-angle X-ray scattering data analysis for multilamellar vesicles: the evolution of the scattering density profile model # J. Appl. Cryst. -2014.-V.47.-P.173-180.
- 167. Zemlyanaya E. V., Kiselev M. A., Zhabitskaya E. I., Aksenov V. L., Ipatova O. M., Ivankov O. I. The Small-Angle Neutron Scattering Data Analysis of the Phospholipid Transport Nanosystem Structure // Journal of Physics: Conference Series. 2018. V. 1023. P. 012017.
- 168. Pabst G., Kucerka N., Nieh M. P., Rheinstadter M. C., Katsaras J. Applications of neutron and X-ray scattering to the study of biologically relevant model membranes // Chem Phys Lipids. -2010. V. 163, No. 6. P. 460-479.
- 169. Caille A. Remarques sur la diffusion des rayons X dans les smectiques // C. R. Acad. Sci. Paris (Ser. B). 1972. V. 274. P. 891-893.
- 170. Zhang R. T., Suter R. M., Nagle J. F. Theory of the Structure Factor of Lipid Bilayers // Physical Review E. -1994. -V. 50, No. 6. -P. 5047-5060.
- 171. Kiselev M. A., Lesieur P., Kisselev A. M., Lombardo D., Aksenov V. L. Model of separated form factors for unilamellar vesicles // Applied Physics a-Materials Science & Processing. 2002. V. 74. P. S1654-S1656.
- 172. Pencer J., Krueger S., Adams C. P., Katsaras J. Method of separated form factors for polydisperse vesicles // J. Appl. Cryst. 2006. V. 39. P. 293-303.
- 173. Pabst G., Koschuch R., Pozo-Navas B., Rappolt M., Lohner K., Laggner P. Structural analysis of weakly ordered membrane stacks // J. Appl. Cryst. 2003. V. 36. P. 1378-1388.
- 174. Torbet J., Wilkins M. H. X-ray diffraction studies of lecithin bilayers // J Theor Biol. -1976. V. 62, No. 2. P. 447-458.
- 175. MacDonald R. C., MacDonald R. I., Menco B. P., Takeshita K., Subbarao N. K., Hu L. R. Small-volume extrusion apparatus for preparation of large, unilamellar vesicles // Biochim Biophys Acta. − 1991. − V. 1061, № 2. − P. 297-303.

- 176. Frisken B. J., Asman C., Patty P. J. Studies of vesicle extrusion // Langmuir. 2000. V. 16, № 3. P. 928-933.
- 177. Pabst G., Amenitsch H., Kharakoz D. P., Laggner P., Rappolt M. Structure and fluctuations of phosphatidylcholines in the vicinity of the main phase transition // Phys Rev E Stat Nonlin Soft Matter Phys. − 2004. − V. 70, № 2 Pt 1. − P. 021908.
- 178. Soloviov D. V., Gorshkova Y. E., Ivankov O. I., Zhigunov A. N., Bulavin L. A., Gordeliy V. I., Kuklin A. I. Ripple phase behavior in mixtures of DPPC/POPC lipids: SAXS and SANS studies // J. Phys. Conf. Ser. 2012. V. 351. P. 012010.
- 179. Folch J., Lees M., Sloane Stanley G. H. A simple method for the isolation and purification of total lipides from animal tissues // J Biol Chem. 1957. V. 226, № 1. P. 497-509.
- 180. Kucerka N., Nieh M.-P., Katsaras J. Small-Angle Scattering from Homogenous and Heterogeneous Lipid Bilayers // Advances in Planar Lipid Bilayers and Liposomes / Tien A. I. a. H. T. Cambridge: Academic Press, Elsevier, 2010. P. 201-235.
- 181. Ruigrok R. W., Barge A., Durrer P., Brunner J., Ma K., Whittaker G. R. Membrane interaction of influenza virus M1 protein // Virology. 2000. V. 267, № 2. P. 289-298.
- 182. Saletti D., Radzimanowski J., Effantin G., Midtvedt D., Mangenot S., Weissenhorn W., Bassereau P., Bally M. The Matrix protein M1 from influenza C virus induces tubular membrane invaginations in an in vitro cell membrane model // Sci Rep. 2017. V. 7. P. 40801.
- 183. Galimzyanov T. R., Lyushnyak A. S., Aleksandrova V. V., Shilova L. A., Mikhalyov, II, Molotkovskaya I. M., Akimov S. A., Batishchev O. V. Line Activity of Ganglioside GM1 Regulates the Raft Size Distribution in a Cholesterol-Dependent Manner // Langmuir. − 2017. − V. 33, № 14. − P. 3517-3524.
- 184. Hamm M., Kozlov M. M. Elastic Energy of Tilt and Bending of Fluid Membranes // Eur. Phys. J. E –2000. V. 3. P. 323–335.

- 185. Khelashvili G., Kollmitzer B., Heftberger P., Pabst G., Harries D. Calculating the Bending Modulus for Multicomponent Lipid Membranes in Different Thermodynamic Phases // J Chem Theory Comput. − 2013. − V. 9, № 9. − P. 3866-3871.
- 186. Veit M., Thaa B. Association of influenza virus proteins with membrane rafts // Adv Virol. 2011. V. 2011. P. 370606.
- 187. Fontana J., Steven A. C. At low pH, influenza virus matrix protein M1 undergoes a conformational change prior to dissociating from the membrane // J Virol. -2013. V. 87, $Noldsymbol{Noldsymbol{O}}$ 10. -P. 5621-5628.
- 188. Mertens H. D., Svergun D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering // J Struct Biol. 2010. V. 172, № 1. P. 128-141.
- 189. Vestergaard B., Groenning M., Roessle M., Kastrup J. S., van de Weert M., Flink J. M., Frokjaer S., Gajhede M., Svergun D. I. A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils // PLoS Biol. 2007. V. 5, № 5. P. e134.
- 190. Giehm L., Svergun D. I., Otzen D. E., Vestergaard B. Low-resolution structure of a vesicle disrupting & alpha;-synuclein oligomer that accumulates during fibrillation // Proc Natl Acad Sci U S A. -2011.-V.108, No. 8.-P. 3246-3251.
- 191. Matsui T., Tsuruta H., Johnson J. E. Balanced Electrostatic and Structural Forces Guide the Large Conformational Change Associated with Maturation of T=4 Virus // Biophysical Journal. − 2010. − V. 98, № 7. − P. 1337-1343.
- 192. Aramayo R., Merigoux C., Larquet E., Bron P., Perez J., Dumas C., Vachette P., Boisset N. Divalent ion-dependent swelling of tomato bushy stunt virus: A multi-approach study // Biochimica Et Biophysica Acta-General Subjects. -2005. -V. 1724, N_{\odot} 3. -P. 345-354.
- 193. Blobel J., Bernado P., Svergun D. I., Tauler R., Pons M. Low-resolution structures of transient protein-protein complexes using small-angle X-ray scattering // J Am Chem Soc. − 2009. − V. 131, № 12. − P. 4378-4386.

- 194. Herranz-Trillo F., Groenning M., van Maarschalkerweerd A., Tauler R., Vestergaard B., Bernado P. Structural Analysis of Multi-component Amyloid Systems by Chemometric SAXS Data Decomposition // Structure. − 2017. − V. 25, № 1. − P. 5-15.
- 195. Onuk A. E., Akcakaya M., Bardhan J. P., Erdogmus D., Brooks D. H., Makowski L. Constrained Maximum Likelihood Estimation of Relative Abundances of Protein Conformation in a Heterogeneous Mixture from Small Angle X-Ray Scattering Intensity Measurements // IEEE Trans Signal Process. 2015. V. 63, № 20. P. 5383-5394.
- 196. Tauler R. Multivariate curve resolution applied to second order data // Chemometrics and Intelligent Laboratory Systems. − 1995. − V. 30, № 1. − P. 133-146.
- 197. Jaumot J., Vives M., Gargallo R. Application of multivariate resolution methods to the study of biochemical and biophysical processes // Anal Biochem. -2004. V.327, No. 1. P. 1-13.
- 198. Maeder M. Evolving factor analysis for the resolution of overlapping chromatographic peaks // Anal. Chem. 1987. V. 59, № 3. P. 527-530.
- 199. Keller H. R., Massart D. L. Evolving factor analysis // Chemometrics and Intelligent Laboratory Systems. 1991. V. 12, № 3. P. 209-224.
- 200. Meisburger S. P., Taylor A. B., Khan C. A., Zhang S., Fitzpatrick P. F., Ando N. Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-ray Scattering // J Am Chem Soc. -2016.-V.138, No. 20.-P.6506-6516.
- 201. Hopkins J. B., Gillilan R. E., Skou S. BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis // J. Appl Cryst. 2017. V. 50. P. 1545-1553.
- 202. Tully M. D., Tarbouriech N., Rambo R. P., Hutin S. Analysis of SEC-SAXS data via EFA deconvolution and Scatter // J. Vis. Exp. 2021. V. 167. P. e61578.
- 203. Sagar A., Herranz-Trillo F., Langkilde A. E., Vestergaard B., Bernado P. Structure and thermodynamics of transient protein-protein complexes by chemometric decomposition of SAXS datasets // Structure. − 2021. − V. 29, № 9. − P. 1074-1090 e4.

- 204. Ayuso-Tejedor S., Garcia-Fandino R., Orozco M., Sancho J., Bernado P. Structural analysis of an equilibrium folding intermediate in the apoflavodoxin native ensemble by small-angle X-ray scattering // J Mol Biol. -2011.-V.406, No.40.-P.604-619.
- 205. Meisburger S. P., Xu D., Ando N. REGALS: a general method to deconvolve X-ray scattering data from evolving mixtures // IUCrJ. − 2021. − V. 8, № Pt 2. − P. 225-237.
- 206. Solutions of Ill-posed Problems. / Tikhonov A. N., Arsenin V. Y. New York: JOHN WILEY & SONS, 1977. 137 p.
- 207. Computer Methods for Mathematical Computations. Prentice-Hall Series in Automatic Computation. / Forsythe G. E., Malcolm M. A., Moler C. B. Englewood Cliffs, New Jersey Prentice-Hall, Inc, 1977. Prentice-Hall Series in Automatic Computation. 280 p.
- 208. Lan K., Jorgenson J. W. A hybrid of exponential and gaussian functions as a simple model of asymmetric chromatographic peaks // J Chromatogr A. -2001. V.915, No 1-2. -P.1-13.
- 209. Ahrens J. H., Dieter U. Computer generation of Poisson deviates from modified normal distributions // ACM Transactions on Mathematical Software. 1982. V. 8. P. 163-179.
- 210. Ahrens J. H., Kohrt K. D., Dieter U. Algorithm 599 Sampling from Gamma and Poisson distributions // ACM Transactions on Mathematical Software. 1983. V. 9. P. 255-257.
- 211. Franke D., Hajizadeh N. R., Svergun D. I. Simulation of small-angle X-ray scattering data of biological macromolecules in solution // J Appl Cryst. -2020. -V. 53, N_{\odot} Pt 2. -P. 536-539.
- 212. Graewert M. A., Franke D., Jeffries C. M., Blanchet C. E., Ruskule D., Kuhle K., Flieger A., Schafer B., Tartsch B., Meijers R., Svergun D. I. Automated pipeline for purification, biophysical and x-ray analysis of biomacromolecular solutions // Sci Rep. 2015. V. 5. P. 10734.

- 213. Mardsen S. R., Mestrom L., Bento I., Hagedoorn P.-L., McMillan D. G. G., Hanefeld U. CH-π Interactions Promote the Conversion of Hydroxypyruvate in a Class II Pyruvate Aldolase // Advanced Synthesis & Catalysis −2019. − V. 361, № 11. − P. 2649-2658.
- 214. Panjkovich A., Svergun D. I. CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data // Bioinformatics. − 2018. − V. 34, № 11. − P. 1944-1946.
- 215. Brookes E., Perez J., Cardinali B., Profumo A., Vachette P., Rocco M. Fibrinogen species as resolved by HPLC-SAXS data processing within the UltraScan Solution Modeler (US-SOMO) enhanced SAS module // J Appl Cryst. − 2013. − V. 46, № Pt 6. − P. 1823-1833.
- 216. Shkumatov A. V., Strelkov S. V. DATASW, a tool for HPLC-SAXS data analysis // Acta Crystallogr D Biol Crystallogr. 2015. V. 71, № Pt 6. P. 1347-1350.
- 217. Malaby A. W., Chakravarthy S., Irving T. C., Kathuria S. V., Bilsel O., Lambright D. G. Methods for analysis of size-exclusion chromatography-small-angle X-ray scattering and reconstruction of protein scattering // J Appl Cryst. -2015. -V. 48, \mathbb{N} Pt 4. -P. 1102-1113.
- 218. Brookes E., Vachette P., Rocco M., Perez J. US-SOMO HPLC-SAXS module: dealing with capillary fouling and extraction of pure component patterns from poorly resolved SEC-SAXS data // J Appl Cryst. − 2016. − V. 49, № Pt 5. − P. 1827-1841.
- 219. Boulin C., Kempf R., Koch M. H. J., McLaughlin S. M. Data appraisal, evaluation and display for synchrotron radiation experiments: hardware and software // Nucl. Instrum. Meth. A. -1986. V. 249. P. 399-407.
- 220. Strunz P., Saroun J., Keiderling U., Wiedenmann A., Przenioslo R. General formula for determination of cross-section from measured SANS intensities // J. Appl. Cryst. 2000. V. 33. P. 829-833.
- 221. Data Reduction and Error Analysis for the Physical Sciences. / Bevington P. B.: New York: McGraw-Hill, 1969. 336 p.
- 222. Hammersley A. P. FIT2D: a multi-purpose data reduction, analysis and visualization program // J.Appl.Cryst. 1995. V. 49. P. 646-652.

- 223. Burley S. K. An overview of structural genomics // Nat Struct Biol. 2000. V. 7 Suppl. P. 932-934.
- 224. Edwards A. M., Arrowsmith C. H., Christendat D., Dharamsi A., Friesen J. D., Greenblatt J. F., Vedadi M. Protein production: feeding the crystallographers and NMR spectroscopists // Nat Struct Biol. 2000. V. 7 Suppl. P. 970-972.
- 225. Svergun D. I., Aldag I., Sieck T., Altendorf K., Koch M. H. J., Kane D. J., Kozin M. B., Grueber G. A model of the quaternary structure of the Escherichia coli F1 ATPase from X-ray solution scattering and evidence for structural changes in the delta subunit during ATP hydrolysis // Biophys J. − 1998. − V. 75, № 5. − P. 2212-2219.
- 226. Ashton A. W., Boehm M. K., Gallimore J. R., Pepys M. B., Perkins S. J. Pentameric and decameric structures in solution of serum amyloid P component by X-ray and neutron scattering and molecular modelling analyses // J Mol Biol. -1997. V. 272, N = 3. P. 408-422.
- 227. Krueger J. K., Olah G. A., Rokop S. E., Zhi G., Stull J. T., Trewhella J. Structures of calmodulin and a functional myosin light chain kinase in the activated complex: a neutron scattering study // Biochemistry. − 1997. − V. 36, № 20. − P. 6017-6023.
- 228. Angular momentum in quantum mechanics. Investigations in physics; 4. / Edmonds A. R. Princeton, N.J.: Princeton University Press, 1957. Investigations in physics; 4. 146 p.
- 229. Svergun D. I. Mathematical methods in small-angle scattering data analysis // J. Appl. Cryst. 1991. V. 24. P. 485-492.
- 230. Svergun D. I. Solution scattering from biopolymers: advanced contrast variation data analysis // Acta Crystallogr. 1994. V. A50. P. 391-402.
- 231. Computer Graphics. Principles and practice. / Foley J. D., van Dam A., Feiner S. K., Hughes J. F. Second edition изд. New York: Addison-Wesley Publishing Company, 1990. 1175 p.
- 232. Bernstein F. C., Koetzle T. F., Williams G. J., Meyer E. E., Jr., Brice M. D., Rodgers J. R., Kennard O., Shimanouchi T., Tasumi M. The Protein Data Bank: a computer-based

- archival file for macromolecular structures // J Mol Biol. 1977. V. 112, № 3. P. 535-542.
- 233. Koenig S., Svergun D., Koch M. H. J., Hubner G., Schellenberger A. The influence of the effectors of yeast pyruvate decarboxylase (PDC) on the conformation of the dimers and tetramers and their pH-dependent equilibrium // Eur Biophys J. -1993. V. 22, N = 3. P. 185-194.
- 234. Svergun D. I., Petoukhov M. V., Koch M. H. J., Koenig S. Crystal versus solution structures of thiamine diphosphate-dependent enzymes // J. Biol. Chem. -2000. -V. 275, N0 1. -P. 297-302.
- 235. Dennis J., Gay D., Welsch R. An adaptive nonlinear least-squares algorithm // ACM. Trans. Math. Soft. 1981. V. 7. P. 348-383.
- 236. Sun L., Zhao L., Yang G., Yan C., Zhou R., Zhou X., Xie T., Zhao Y., Wu S., Li X., Shi Y. Structural basis of human gamma-secretase assembly // Proc Natl Acad Sci U S A. -2015. -V. 112, No 19. -P. 6003-6008.
- 237. Svergun D. I., Burkhardt N., Pedersen J. S., Koch M. H. J., Volkov V. V., Kozin M. B., Meerwink W., Stuhrmann H. B., Diedrich G., Nierhaus K. H. Solution scattering structural analysis of the 70 S Escherichia coli ribosome by contrast variation. II. A model of the ribosome and its RNA at 3.5 nm resolution // J Mol Biol. − 1997. − V. 271, № 4. − P. 602-618.
- 238. Mitra K., Schaffitzel C., Fabiola F., Chapman M. S., Ban N., Frank J. Elongation arrest by SecM via a cascade of ribosomal RNA rearrangements // Mol Cell. 2006. V. 22, № 4. P. 533-543.
- 239. Panjkovich A., Svergun D. I. SASpy: a PyMOL plugin for manipulation and refinement of hybrid models against small angle X-ray scattering data // Bioinformatics. -2016. V. 32, No 13. P. 2062-2064.
- 240. Diffraction of X-rays by Chain Molecules. / Vainshtein B. K.: Amsterdam:Elsevier, 1966. 440 P.