На правах рукописи

УДК 543.4; 539.2; 548.737

Волков Владимир Владимирович

"Спектроскопия и малоугловое рассеяние в решении обратных задач исследования многокомпонентных систем"

Специальность 01.04.18 – «кристаллография, физика кристаллов»

Автореферат диссертации на соискание ученой степени доктора химических наук

Москва 2013

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте кристаллографии имени А.В.Шубникова Российской академии наук (ИК РАН).

Официальные оппоненты:	Чвалун Сергей Николаевич , доктор химических наук, профессор. Отделение кристаллографии и материаловедения Курчатовского НБИКС-центра, НИЦ "Курчатовский институт".		
	Ткаль Валерий Алексеевич , доктор физико-математических наук, профессор, Новгородский филиал Санкт-Петербургского государственного университета сервиса и экономики НФСПбГУСЭ, кафедра технических и естественнонаучных дисциплин.		
	Бушуев Владимир Алексеевич , доктор физико-математических наук, профессор, МГУ имени М.В. Ломоносова, физический факультет, отделение физики твердого тела, кафедра физики твердого тела.		
Ведущая организация:	Федеральное государственное бюджетное учреждение науки Институт элементоорганических соединений им. А.Н.Несмеянова Российской академии наук ИНЭОС РАН.		

Защита состоится «____» ____ 2014 г. в _____ч. ____ мин. на заседании диссертационного Совета Д 002.114.01 при Федеральном государственном бюджетном учреждении науки Институте кристаллографии им. А. В. Шубникова Российской академии наук, по адресу 119333, г. Москва, Ленинский пр-т, д.59, конференц-зал.

С диссертацией можно ознакомиться в библиотеке ИК РАН, а также на сайте ИК РАН http://crys.ras.ru/index.php/dissertatsionnyj-sovet/zashchity-dissertatsij

Автореферат разослан «____»____ 2013 г.

Ученый секретарь диссертационного совета доктор физико-математических наук

1. Общая характеристика работы

Актуальность темы. Разработка эффективных методов определения химического состава и структуры вещества всегда была актуальной научной задачей, и успехи в этой области приобретают с развитием технологий все большее значение, так как изучение связей структура – свойство необходимо для решения всех практических задач по созданию новых материалов, мониторинга окружающей среды, синтеза веществ с заранее заданными функциями и т.д. Постоянно растущая сложность фундаментальных и практических задач делает развитие новых методов анализа неизменно важной частью современной химии. Хорошие перспективы в этой области имеют компьютерные методы, которые стали таким же инструментом исследования, как и химический эксперимент. Многие свойства материалов, в частности, включающих биомакромолекулы, определяются надатомной организацией. Поэтому одинаково важны как разработка эффективных спектроскопических методов химического анализа, так и определения строения исследуемых образцов при разрешении от единиц до десятков нанометров.

Современные материалы часто представляют собой системы, функции которых сохраняются в ограниченном диапазоне изменения внешних условий. Это требует применения неразрушающих физических методов исследований. Кроме того, многие материалы представляют собой не полностью упорядоченные системы, неприменимыми что делает такие высокоинформативные методы, как, например, рентгеноструктурный анализ. При этом экспериментально доступная прямая информация об интересующих исследователя параметрах объектов, которые, по существу, представляет собой физически или химически неразделяемые смеси, как правило, оказывается недостаточной. Единственный способ получения данных о чистых компонентах В этих случаях и/или построения физико-химических моделей объектов - компьютерная обработка результатов измерений С помощью алгоритмов анализа многомерных сигналов. Сложность возникающих при этом задач требует разработки численных методов, как общего назначения, так и учитывающих особенности конкретных случаев, что делает эти исследования Измеряемые сигналы (спектры неизменно актуальными. поглощения, магнитного и парамагнитного резонансов, испускания, флуоресценции, рассеяния электромагнитного излучения и элементарных частиц и т.д.) часто представляют собой аддитивные данные, то есть сумму характеристик отдельных компонентов. Химический анализ таких смесей, заключающийся в спектральных характеристик быть определении компонентов, может неоднозначен, что делает разработку математических методов разделения спектров с анализом устойчивости решения чрезвычайно актуальной для аналитиков, работающих в различных областях.

Данное направление получило развитие в рамках специализированной химической дисциплины, получившей название хемометрики (Chemometrics), или хемометрии (Шараф М.А., Иллмэн Д.Л., Ковальски Б.Р. Хемометрика. Л.: Химия, 1989. 272 с.). Многие разработанные в этой области методы не требуют априорных данных о составе объектов исследований. Такие методы называются безэталонными или "автомодельными" (self-modeling). В ряде случаев численный хемометрический анализ оказывается предпочтительнее хроматографических и масс-спектрометрических методов, которые хотя и обладают большими возможностями, не всегда пригодны, например, в случае дистанционного мониторинга или при изучении равновесных систем. К настоящему времени для разделения аддитивных сигналов разработано и активно используется несколько программ, во многих из которых реализованы основные принципы, аналогичные предложенным автором данной работы за несколько лет до широкого распространения этих процедур.

Определение размерных параметров формы И наночастиц, биомакромолекул, других наноразмерных структурных неоднородностей в неупорядоченных средах также является быстроразвивающейся и чрезвычайно востребованной областью исследований. В данной работе рассматриваются методологические и численные аспекты организации программ анализа данных рентгеновского И нейтронного рассеяния малоуглового для расчета распределений наночастиц по размерам, а также определения формы одно- и многокомпонентных макромолекул в растворах по данным рассеяния.

Особое внимание в работе уделено способам оценки надежности и устойчивости решений. Эта часть работы важна, поскольку рассмотренные выше задачи, как правило, плохо обусловлены. Приводимые в литературе оценки ошибок статистические решений оказываются, как правило, завышенными и, к тому же, не дают ответа на вопрос о надежности отдельных компонентов решения. Ситуация осложняется присутствием систематических ошибок в исходных данных, что создает дополнительные препятствия для физический решений, имеющих смысл. Поэтому получения оценка стабильности решений представляет собой актуальную проблему для обратных задач, в том числе и рассматриваемых в данной работе.

Цели и задачи работы. Целью настоящей работы была разработка методов химического анализа многокомпонентных систем по спектроскопическим определения данным И структурных параметров наночастиц в моно- и многокомпонентных по данным рентгеновского и нейтронного малоуглового рассеяния. Важной частью была разработка способов оценивания устойчивости решений и рассмотрение вопросов практической реализации предложенных алгоритмов.

Можно выделить две крупных группы задач наших исследований.

1. Разработка методов компьютерного анализа спектров (сигналов) неразделяемых смесей компонентов различной природы. При этом одной из целей была разработка методов, позволяющих легко учитывать дополнительную информацию, от простых требований неотрицательности решения – спектров компонентов и их относительных концентраций, ДО использования частично или полностью известных кандидатов на индивидуальные спектры, например, библиотечных спектров.

2. Разработка методов и программ определения формы и распределений по размерам наночастиц по данным малоуглового рентгеновского и нейтронного рассеяния в одно- и многокомпонентных системах.

Разрабатываемые методы решения этих задач, которые в ряде случаев не имеют единственного решения, должны были включать способы оценки устойчивости найденных спектроскопических и структурных моделей. Эти способы должны давать практически полезную информацию не только о разбросе, но и о надежности моделей в целом и их компонентов в отдельности.

Методы исследований. В работе применены два неразрушающих метода получения экспериментальных данных, позволяющих характеризовать вещество как со стороны состава и химических связей (спектроскопия), так и пространственного строения (малоугловое рентгеновское и нейтронное рассеяние). Конечные результаты получены с помощью новых компьютерных программ анализа данных, разработанных для решения соответствующих обратных задач.

Разработаны Научная новизна. методы спектроскопического компонентного анализа смесей химических соединений различной природы, позволяющие проводить исследование образцов с произвольным числом компонентов, которое ограничивается только уровнем шумов измерений. Новизна методов заключается также в разработанных приемах оценки стабильности и достоверности получаемых результатов. Развитые методы проводить компонентный анализ ПОЗВОЛЯЮТ достаточно просто как С привлечением дополнительной спектроскопической или другой информации, так и в условиях отсутствия априорных сведений об изучаемой системе. Данные методы применены для выделения индивидуальных спектров в смесях органических и неорганических соединений, а также впервые использованы для коррекции интенсивности в наборах данных измерений малоуглового рассеяния и определения числа рассеивающих компонентов.

Разработанные с непосредственным участием автора методы анализа данных малоуглового рассеяния развиты в части повышения стабильности результатов и эффективности поиска формы частиц в монодисперсных образцах и распределений по размерам в полидисперсных системах. Методы применены для изучения морфология макромолекул ряда белков и биологических комплексов, которые не поддаются исследованиям другими методами. Определены размерные параметры наночастиц и неоднородностей в системах различной природы. Новизна полученных результатов подтверждается высокими индексами цитирования работ с участием автора.

Научная и практическая значимость полученных результатов.

Ценность работы определяется актуальностью поставленных задач, универсальностью разработанных методов спектроскопического анализа смесей и структурного анализа данных рассеяния от систем наночастиц. Эффективность методов продемонстрирована на примере решения задач из разных областей химического и структурного анализа сложных объектов. Отметим основные применения разработанных методов.

1. Качественный и количественный спектральный анализ неразделяемых смесей любых сложных объектов по набору аддитивных неотрицательных данных измерений физических характеристик – спектров различной природы (поглощения, испускания, люминесценции, ядерного магнитного резонанса и т.д.) и интенсивности малоуглового рентгеновского и нейтронного рассеяния.

2. Анализ измерений смесей с целью определения числа компонентов и коррекции экспериментальных данных с целью приведения их в соответствие с физической моделью объекта и проверки адекватности самой модели.

3. Структурный анализ смесей наночастиц по данным малоуглового рентгеновского и нейтронного рассеяния.

4. Определение формы наночастиц в разбавленных системах, в частности, анализ строения белковых макромолекул, как в рамках однофазных, так и многофазных моделей.

Предложенные в работе методы анализа смесей и строения наночастиц реализованы в ряде компьютерных программ и программных комплексов, входящих в систему ATSAS [A39], которая использовалась на момент написания работы более чем в 2000 лабораторий и находящийся в свободном доступе (URL: http://www.embl-hamburg.de/research/unit/svergun/index.html).

разработок С учетом проведенных подготовлен И выпущен Российской (ГОСТ P 8.698-Национальный стандарт федерации 2010 Государственная система обеспечения единства измерений. Размерные параметры наночастиц и тонких пленок. Методика выполнения измерений с помощью малоуглового рентгеновского дифрактометра. –М. : Стандартинформ, 2010.-41 с.). Выпущено методическое пособие "Определение формы частиц по малоуглового рентгеновского И нейтронного рассеяния" данным (URL: http://nano.msu.ru/files/systems/ 4 2010/practical/41 full.pdf).

Степень достоверности полученных результатов. Достоверность представленных в диссертации результатов по разложению спектров подтверждается данными экспериментов по исследованию модельных смесей органических соединений. Результаты анализа данных малоуглового рассеяния от полидисперсных систем соответствуют данным, полученными, например,

методом электронной микроскопии. Результаты определения формы частиц биомолекул всегда сопоставлялись со структурами, найденными другими авторами методом белковой кристаллографии, если такая информация была доступна.

Апробация работы. Результаты работы докладывались на следующих конференциях: 4-й Всесоюзной конференции по аналитической химии органических соединений. (Москва, 1980); 5-й и 6-й Всесоюзных конференциях по использованию вычислительных машин в спектроскопии молекул и химических исследованиях. (Новосибирск, 1980, 1983); Всесоюзной конференции "Математические методы и ЭВМ в аналитической химии." (Москва, 1986); 4th International Conference on Laser Applications in Life Sciences(Jyvaskyla, Finland, 1992); X International Conference on Small-Angle Scattering (Campinas Sao Paulo Brazil, 1996); 1-й, 2-й, 3-й и 4-й Национальных конференциях по применению рентгеновского, синхротронного излучений, нейтронов и электронов для исследования материалов РСНЭ (Москва, 1997, 1999, 2001, 2003); 5-й 6-й, 7-й и 8-й Национальных конференциях по применению Рентгеновского, Синхротронного излучений, Нейтронов И Электронов для исследования наноматериалов и наносистем РСНЭ НАНО-НБИК (Москва, 2005, 2007, 2009, 2011); Second International Conference on Modern Trends in Computational Physics (Dubna, Russia, 2000); XX и XXII микроскопии Российских конференциях по электронной РКЭМ-2002 (Черноголовка 2002, 2008); XIII International Conference On Small-Angle Scattering (Venice, Italy, 2002); III Съезд биофизиков (Воронеж, 2004); I France-Russia Seminar "New Achievements in Materials Science" (Nancy, France, 2004); II France-Russia Seminar "New Achievements in Materials Science" (Moscow, Russia, 2005); European Polymer Congress EPF 2005 (M.V. Lomonosov Moscow State University (Moscow, Russia, 2005); XIII International Conference on Small-angle Scattering (SAS2006) (Japan, Kyoto, 2006); Electron Microscopy and Multiscale Modeling EMMM-2007 (Moscow, 2007); XII и XIII Национальных конференциях по росту кристаллов. НКРК-2006 (Москва, 2006, 2008); XI Симпозиуме "Современная химическая физика" (Туапсе, 2007); XXI Congress of the International Union of Crystallography (IUCr2008) (Osaka, Japan, 2008); 7th International Conference on X-Ray investigations of Polymer Structure (XIPS'2007) (Krakow, Poland, 2007); Международном форуме по нанотехнологиям (Москва, 2008); 6th International Symposium "Molecular Order and Mobility in Polymer Systems" (St.Peterburg, Russia, 2008); XIV International Conference on Small-angle Scattering (SAS2009) (UK, Oxford, 2009); International Conference on Supercritical Fluid "Supergreen 2009" (Tohoku University, Sendai, Japan, 2009); 15th Int. Small Angle Scattering Conference (Sydney, Australia, 2012); 11th International Conference Biology and Synchrotron Radiation BSR-2013 (Hamburg, 2013).

Разработанные подходы отражены в опубликованных лекциях и практикумах для студентов и аспирантов:

Workshop "Shape determination of biological macromolecules in solution and related tiopics" (Hambnurg, 1999); EMBO Practical Course on "Solution Scattering from Biological Macromolecules", European Molecular Bilogical Laboratory (EMBL) (Hamburg, 2000, 2003); Школе молодых ученых по физике наноструктурированных и кристаллических материалов. Нижегородский Государственный Университет им. Н.И.Лобачевского (Нижний Новгород, 2010); Всероссийской научной школе для молодежи «Образование в сфере нанотехнологий: современные подходы и перспективы. (Национальный исследовательский технологический университет «МИСиС», 2010).

Публикации. По теме диссертации опубликованы 63 статьи, список которых приведен в конце автореферата. Отметим, что данный список не включает публикации, прямо не связанные с темой диссертации.

Личный вклад автора. Вклад автора состоит в разработке всех алгоритмов и методик анализа экспериментальных спектроскопических данных, деталей алгоритмов и методик анализа формы наночастиц по данным устойчивости малоуглового рассеяния, алгоритмов оценки решений. используемых в данной работе. Под руководством автора и при его участии создан ряд программ, входящих в комплекс свободного программного обеспечения ATSAS [A39] по определению строения белковых макромолекул по данным малоуглового рассеяния и анализа смесей наночастиц. Автор принимал активное участие в проведении экспериментальных исследований, разработке методик измерений и интерпретации данных, в обработке результатов экспериментов.

Основные положения, выносимые на защиту.

1. Алгоритмы анализа наборов аддитивных неотрицательных спектров смесей, основанные на сингулярном разложении матрицы данных, линейных преобразованиях спектральных контуров и статистическом анализе результатов, в том числе:

– методы оценки числа компонентов, основанные на статистических критериях автокорреляции элементов сингулярных векторов матрицы спектроскопических данных;

 метод расчета спектров компонентов и их относительных концентраций в отсутствие дополнительной информации о задаче при условиях максимального различия и простоты контуров индивидуальных спектров.

2. Применение разработанных алгоритмов и методов анализа спектров для коррекции наборов данных малоуглового нейтронного рассеяния с вариацией контраста.

3. Выявление степени однозначности и устойчивости решения задач поиска формы частиц в монодисперсных системах по данным малоуглового

рентгеновского рассеяния с помощью модифицированных методов сферических гармоник и шариковых моделей.

4. Результаты исследования морфологии по данным малоуглового рассеяния от раствора макромолекул иммуноглобулина М и ревматоидного фактора человека с анализом стабильности решений.

5. Метод поиска размерных распределений частиц в полидисперсных смесях по данным малоуглового рассеяния при заданных формфакторах компонентов, отличающийся улучшенной стабильностью решений в случае широких распределений. Результаты анализа дисперсности ряда систем с наночастицами различной природы.

Структура диссертации. Диссертация состоит из введения, шести глав, выводов, списка цитируемой литературы и списка публикаций по материалам работы. Общий объем диссертации – 347 стр., включая 93 рисунка, 12 таблиц и библиографию из 207 цитируемых работ, 63 публикаций автора в реферируемых журналах и 80 тезисов выступлений на конференциях. Из-за разнообразия рассматриваемых вопросов литературный обзор проводится по мере необходимости в соответствующих разделах работы.

2. Основное содержание работы

Во Введении обосновывается актуальность темы работы, определяются ее задачи, предмет и объекты исследования; раскрывается научная новизна и практическая значимость; описываются применяемые методы, теоретикометодологическая основа и эмпирическая база исследования; выдвигаются основные положения, выносимые на защиту.

Диссертации состоит из двух частей. Первая часть посвящена методам решения обратных задач анализа спектроскопических данных и данных малоуглового рассеяния рентгеновских лучей И нейтронов ОТ многокомпонентных систем. Эта работы представляет собой часть продолжение и завершение исследований и разработок, описанных автором в кандидатской диссертации. Во второй части рассматриваются методы анализа малоуглового рассеяния от систем наночастиц. Разработанное данных программное обеспечение основано на единой библиотеке из более чем 500 процедур, написанных на языке Фортран.

Первая часть начинается с литературного обзора (**раздел 1.1**), из которого видно, что к моменту начала работы число публикаций, посвященных алгоритмам разложения спектров, было невелико, и основную их часть составляли методы факторного анализа. Можно отметить, что число компонентов во всех рассмотренных примерах было ограничено 2-3. Лишь в последнее десятилетие было разработано несколько универсальных методов, которые позволяют разделять наборы смешанных сигналов: Idependent

Component Analysis (ICA), Blind signal separation (BSS), методы автомодельного разделения кривых (self-modeling) - RADICAL, SIMPLISMA, MCR-ALS. Относительно недавно по Интернет стали доступны алгоритмы, реализующие методы независимых компонент - MILCA, (Mutual Information Least Dependent Component Analysis) и SNICA (Stochastic Non-Negative Independent Component Analysis). Многие базовые идеи этих подходов аналогичны предложенным в настоящей работе, но опубликованы после работ автора.

В <u>разделе 1.2</u> рассмотрены геометрическая интерпретация принципа методов разложения аддитивных спектров и условие единственности решения соответствующей обратной задачи. Аддитивность спектров означает, что форма индивидуальных спектров не зависит от присутствия других компонентов (компоненты не должны взаимодействовать, быть независимыми) и от условий съемки. Если это условие не соблюдается, число компонентов будет превышать число смесей при любом их количестве и задача будет недоопределенной. Но и в этом случае разделение при помощи разработанных процедур может давать полезную информацию о компонентном составе.

Если известна только матрица спектров смесей D, задача нахождения спектров компонентов X заключается в нахождении матричного разложения D = XC, (1)

относительно неизвестных спектров компонентов X и их концентраций C. Поскольку неизвестные входят в (1) в виде парных произведений, это соотношение представляет собой систему нелинейных уравнений, решение которой неединственно. Базисом пространства спектров смесей могут быть левые сингулярные векторы U матрицы данных из разложения

$$D = U \cdot \Lambda \cdot V^T \tag{2}$$

где U - это $N \ge M$ матрица собственных векторов матрицы $D \cdot D^T$ (в литературе по факторному анализу - ковариационной матрицы), *V* - матрица собственных $D^T \cdot D$, Λ – диагональная векторов матрица сингулярных чисел (положительных квадратных корней из ненулевых собственных значений матрицы $D \cdot D^T$ или $D^T \cdot D$), расположенных в невозрастающем порядке. В случае отсутствия нарушений аддитивности спектральных данных (шумов, систематических ошибок, зависимости контуров компонентов от состава смеси и др.) ранг матрицы D и, следовательно, число значимых (ненулевых) сингулярных чисел, равны числу компонентов. Для определенности здесь и далее будем считать, что число спектральных точек N, количество смесей M и компонентов **К** удовлетворяют неравенству N > M > K, хотя достаточными будут неравенства $N \ge K$, $M \ge K$. Далее обзор известных и рассмотрение предложенных в работе методов определения числа компонентов и поиска индивидуальных спектров проводится с использованием разложения (2).

Проблема неединственности в задаче анализа многокомпонентных

смесей по их спектрам решается путем привлечения той или иной дополнительной информации и ограничений на вид решения. Наиболее общими ограничениями являются требования **неотрицательности** спектров компонентов и их концентраций.



Рисунок 1 - Геометрическое представление задачи разложения 3-х точечных аддитивных спектров смесей для случая 4-х смесей из 2-х компонентов. Здесь $d_1 - d_4$ спектры смесей, x_1 и x_2 - возможные спектры компонентов, $x_1^{A\Phi}$ и $x_2^{A\Phi}$ - частично неперекрытые спектры компонентов (лежат в плоскостях I₁I₂ и I₂I₃, соответственно), t спектр, не являющийся компонентом смеси, p – его проекция на плоскость смесей. По осям I₁ – I₃ отложены интенсивности при трех значениях абсциссы (например, трех длинах волн). Все векторы нормированы для наглядности так, что их концы лежат на одной прямой AD.

Так как спектры смесей есть неотрицательные линейные комбинации спектров компонентов, области существования формально допустимых решений областями подпространства, определяемого ограничиваются векторами-(двумерной смесей плоскости AOD Рисунке спектрами на 1.1), расположенными в 1-м ортанте и ограниченными изнутри конусом, натянутым на набор векторов смесей (области AOd_1 и DOd_4 на Рисунке 1.1). Тогда условие единственности разложения (1) заключается в следующем: в каждой паре спектров компонентов должны существовать, по меньшей мере, две точки, в одной из которых ненулевую интенсивность имеет только первый компонент, в другой - только второй. Спектры компонентов при этом представляют собой совокупность векторов, соответствующих линиям пересечения гиперплоскости, образованной векторами смесей **(B** дальнейшем будем называть ee гиперплоскостью смесей) с координатными плоскостями (векторы OA и OD на Рисунок 1). Необходимость условия доказывается единственностью такого пересечения, достаточность – существованием процедуры поиска пересечения. В случае невыполнения таких условий неперекрытости для обеспечения единственности необходимо привлечение дополнительных ограничений.

Раздел 1.3, посвящен первому этапу анализа смешанных спектров определению числа компонентов по набору экспериментальных данных. Основной вывод из анализа литературных данных состоит в том, что на большинство методов оказывается малопригодным практике из-за невозможности учета систематических ошибок в исходных данных И отклонений распределения случайной составляющей ошибок от нормального закона. Практика автора показала, что закон распределения случайной составляющей в данных часто оказывается близок к лог-нормальному. Если систематическая составляющая ошибок относительно невелика, характер зависимости сингулярных чисел λ , ненулевые значения которых обусловлены наличием ошибок, оказывается приблизительно линейным в координатах $log(\lambda_k) - k$, где k – порядковый номер числа. За число компонентов принимают количество сингулярных чисел, расположенных выше прямой, проведенной через малые $log(\lambda_k)$. Рисунок 2 демонстрирует пример спектров смесей и сингулярных чисел. Качество соответствующий график линейной аппроксимации оценивают с помощью критерия R^2 , критическим значением которого при числе точек 3-6 и уровне значимости 0.05 можно принять 0.7.

Статистическую оценку значений сингулярных чисел при наличии даже незначительных систематических ошибок сделать проблематично. Поэтому нами предложен второй метод, хорошо зарекомендовавший себя на практике, а именно - определение числа компонентов по числу левых сингулярных U, представляют собой векторов элементы которых неслучайные Контуры сингулярных последовательности. векторов, соответствующие значимым сингулярным числам, представленные в виде "спектров", есть линейные комбинации спектров компонентов и при достаточно частом шаге оцифровки имеют неслучайный характер, выбирая корреляции в данных. Если систематические ошибки малы, элементы следующих далее по порядку сингулярных векторов обусловлены шумовыми случайными составляющими. Для оценки случайности элементов векторов предложено использовать непараметрические критерии серий Вальда-Вольфовица (WW)И модифицированный ранговый критерий Вилкоксона. Более эффективным оказался критерий автокорреляции Дарбина-Ватсона (DW):

$$DW = \frac{\sum_{i=2}^{M} (f_i - f_{i-1})^2}{\sum_{i=1}^{M} (f_i)^2},$$
(3)

где *f* представляет собой проверяемую последовательность. Величина критерия лежит в области {0 - 4}, принимая значение 2 в случае отсутствия корреляции. При уровне значимости гипотезы о наличии корреляции между последовательными элементами массива 0.05 ее отвергают, если величина (3)

лежит в диапазоне $\{1.7-2.2\}$. В случае слабой корреляции (из-за нарушения аддитивности спектров) нами предложено расширить диапазон до $\{1.5-2.5\}$, аналогично рекомендациям (Бородич, С. А. Вводный курс эконометрики: Учебное пособие / С. А. Бородич. – Мн.: БГУ, 2000. – 354 с.). Такое расширение, конечно, не является строгим, а относится к экспертным оценкам. На Рисунке 3 показаны левые сингулярные векторы матрицы *D*, в таблице 1 приведены значения критериев. При оценивании числа компонентов общим правилом может служить выбор минимальной из полученных оценок, так как оценки более 6 соответствуют слишком малым величинам сингулярных чисел и, соответственно, малым систематическим ошибкам в данных.

Предложенные критерии оценки числа компонентов применены на практике и для анализа изменений в образцах в процессе измерений малоуглового рентгеновского рассеяния. На Рисунке 4 приведен пример исходных данных измерений во времени белкового комплекса в растворе, которые показывают изменение состава во времени. Форма сингулярных векторов (Рисунок 5) свидетельствует о наличии двух основных компонентов, о чем говорят и приведенные в тексте диссертации величины критериев автокорреляции в последовательности элементов левых сингулярных векторов. В результате был сделать вывод, что число компонентов равно 2 и они, вероятно, соответствуют димерным и мономерным частицам в смеси.

Для неоднозначных случаев нами предложено дополнительно использовать анализ критерия Дарбина-Ватсона, который рассчитывают внутри окна, скользящего вдоль последовательности элементов. В общем случае можно использовать окно сканирования длиной 10% от всего интервала данных, но не менее 10 элементов. Как видно из Рисунка 6, получаемые более детально графики позволяют проследить наличие И характер систематических отклонений в данных измерений.

В конце раздела рассматриваются способы оценивания наличия компонентов, основанные на анализе разности между заранее известным индивидуальным спектром (например, взятого из базы данных) и его проекцией на пространство столбцов матрицы данных *D*. Анализ проводят с применением рассмотренных статистических критериев.

<u>Раздел 1.4</u> посвящен решению основной задачи анализа неразделяемых смесей - поиску контуров индивидуальных спектров по матрице исходных данных. В данной работе предложен сравнительно простой для использования метод разделения, использующий в качестве фильтров решений такие очевидные требования как неотрицательность компонентов и соответствующих им концентраций в совокупности с требованиями простоты формы и максимального различия индивидуальных контуров.



Рисунок 2 - Слева: ИК спектры родственных смесей 6-ти органических соединений (изооктана, толуола, бензола, диэтилового эфира, амилацетата и ацетона), полученные на выходе препаративного жидкостного хроматографа низкого разрешения конструкции автора. Спектры смещены по вертикали для лучшей визуализации. Справа: график логарифмов сингулярных чисел матрицы данных. R2 критерий для линейной аппроксимации правого участка графика равен 0.89 > 0.7. Область А соответствует значимым числам, В - шумовым.



Рисунок 3 - Левые сингулярные векторы *u*_j матрицы данных (Рисунок 2), представленные в виде спектральных контуров. Контуры смещены по вертикали, нулевая линия соответствует горизонтальной линии в серой закраске их площади. Таблица 1 - Значения критерия согласия Вилкоксона и Дарбина-Ватсона, рассчитанные для сингулярных векторов набора данных Рисунка 2.

Номер	Величина	Вывод о числе	Величина	Вывод о числе
сингуляр-	критерия	компонентов	критерия	компонентов
ного	Вилкоксона	(при уровне	Дарбина-	(при уровне
вектора		значимости)	Ватсона (3)	значимости 0.05)
1	1.0	-	0.15	-
2	2.0	-	0.32	-
3	8.0	-	0.30	-
4	14.0	-	0.68	-
5	11.0	-	0.53	-
6	17.0	-	1.00	-
7	25.0	6 (0.002)	1.52	6
8	28.0	7 (0.02)	1.56	7
9	28.0	8 (0.02)	1.72	8
10	30.0	9 (0.05)	1.59	9



Рисунок 4 - Данные 8 последовательных измерений интенсивности малоуглового рассеяния от образца раствора белкового комплекса в трис-буфере на синхротронной станции X33 (EMBL c/o DESY, Гамбург, Германия). Нумерация фреймов показана стрелкой. Ось абсцисс: логарифм модуля вектора рассеяния s = 4π sinθ/λ, λ – длина волны, 2θ - угол рассеяния.

Предложенный нами подход основан на построении последовательности пробных спектров компонентов x_i в виде линейной комбинации левых сингулярных векторов u (*K* - число компонентов, *M* - число смесей, *L* - число сингулярных векторов в сумме):

$$x_{ik} = \sum_{j=1}^{L} u_{ij} \cdot y_{jk} , \ K \le L \le M ,$$
 (4)

с коэффициентами y_{jk} , которые есть искомые параметры и аргументы минимизируемой целевой функции W(y):



Рисунок 5 - Слева - левые сингулярные векторы **u**_j матрицы данных (Рисунок 4), представленные в виде спектральных контуров. Контуры смещены по вертикали, нулевые линии показаны отдельно для каждого вектора. Единицы на осях аналогичны Рисунку 4. Справа: график логарифмов сингулярных чисел матрицы данных. R² критерий для линейной аппроксимации правого участка графика равен 0.69 ~ 0.7. Выше прямой линии располагается 2 сингулярных числа, которые соответствуют числу значимых компонентов.



Рисунок 6 - Графики оконной статистики Дарбина-Ватсона для элементов левых сингулярных векторов матрицы данных (Рисунок 4). Номера кривых 1 - 8 соответствуют номерам сингулярных векторов. Каждый график критерия показан со смещением по вертикали в рамке, соответствующей значениям 0 < DW < 4.0, а уровни критических значений 1.7, 2.0 и 2.2 показаны внутри

значений 1.7, 2.0 и 2.2 показаны внутри рамок тремя горизонтальными линиями. Значения критерия внутри диапазона 1.7 - 2.2 соответствуют отклонению гипотезы о наличии автокорреляций в окне сканирования. Длина окна сканирования 10% от всего углового диапазона.

$$W(\mathbf{y}) = \alpha \cdot \left(RES - RES_{\lim}\right) + \beta \cdot \left[cond(X) - 1\right]^2 + \gamma \cdot \frac{1}{N \cdot K} \cdot \sum_{\forall (x_{ik} < 0)}^{N,K} x_{ik}^2 + \left(\frac{1}{M \cdot K} \cdot \sum_{\forall (c_{k,m} < 0)}^{K,M} c_{ik}^2 + \varepsilon \cdot \frac{1}{N \cdot K} \cdot \sum_{k=1}^{K} len(\mathbf{x}_k) + \left(\frac{1}{N \cdot K} \cdot \sum_{k=1}^{K} cur(\mathbf{x}_k) + \xi \cdot \frac{2}{K \cdot (K-1)} \cdot \sum_{\forall (l \neq m)}^{K} \left[cos(\mathbf{x}_l, \mathbf{x}_m)\right] + \left(\frac{1}{N \cdot K} \cdot \frac{2}{K \cdot (K-1)} \cdot \sum_{\forall (l \neq m)}^{K} r(\mathbf{x}_l, \mathbf{x}_m) + \ldots\right)\right)$$

$$(5)$$

где X - матрица пробных спектров компонентов x_l, x_m - l-й и m-й векторспектры компонентов, составляющих матрицу X, i = 1,...,N – индекс точки в спектре, k = 1,...,K – индекс компонента, r – коэффициент корреляции между спектрами. RES означает сумму квадратов разностей двумя между нормированными экспериментальными и модельными спектрами смесей (невязка), *RES*_{lim} - минимальное ожидаемое значение невязки, рассчитываемое в виде суммы квадратов элементов матрицы остатков, рассчитанной по (2) с использованием сингулярных векторов с индексами K+1...M. cond(X) означает обусловленности матрицы пробных спектров компонентов число Х. вычисляемое как отношение максимального к минимальному сингулярным числам λ_x – диагональных элементов матрицы Λ_x из разложения

$$X = U_x \Lambda_x V_x^T, \quad cond(X) = \lambda_{\max} / \lambda_{\min}.$$
 (6)

Число обусловленности в данном случае отражает суммарную степень сходства вектор-спектров компонентов, достигая минимального значения 1.0 в случае их максимального формального различия. Третий и четвертый члены в (5) есть штрафы за отрицательность элементов в спектрах и соответствующих концентраций, определяемых по соотношению (1). $len(x_k)$ есть длина нормированного спектрального контура, аппроксимированного ломаной кривой (штраф за сложность контура).

Остальные слагаемые в (5) альтернативны соответствующим слагаемым, рассмотренным выше. Так, $cur(x_k)$ пропорционален кривизне спектрального контура, вычисляемой как сумма по вторым производным от контура (критерий сложности). 7-й и 8-й члены в (5) (суммы косинусов углов и коэффициентов корреляций между компонентами как критерии их парного сходства) соответствуют cond(X) в сумме (5) - критерию различия между спектрами. Введение альтернативных штрафных членов в целевую функцию полезно на практике, так как они обладают разной чувствительностью к деталям спектральных контуров. В разработанной автором программе REMIX, реализующей описанный возможность подбора подход, имеется

соответствующих весовых коэффициентов при решении конкретной задачи для достижения максимально эффективной работы алгоритма.

На Рисунке 7 показаны результаты поиска спектров компонентов по набору данных Рисунка 2 при разных числах искомых компонентов. Варьирование числа компонентов и сравнение контуров спектров-решений представляет собой одну из предлагаемых процедур оценивания однозначности и надежности решения. Данный эксперимент демонстрирует, что увеличение числа модельных компонентов не приводит к существенным искажениям спектральных контуров 1 - 6. Вместе c тем, основных введение дополнительных компонентов позволяет "очистить" некоторые из модельных спектров от полос, которые обусловлены нарушениями аддитивности из-за взаимодействия компонентов между собой. Распределение концентраций основных компонентов для решения K = 7 показано на Рисунке 8, на котором видно, что вклад 7-го компонента незначителен.



Рисунок 7 - Результаты разложения для спектров смесей (Рисунок 2). Жирными линиями показаны модельные спектры, линиями с точками – экспериментально полученные спектры чистых веществ: 1 – изооктана, 2 – толуола, 3 – бензола, 4 – диэтилового эфира, 5 - амилацетата, 6 – ацетона. *а) – в)* – решения для гипотез о 6, 7 и 8 компонентах, соответственно. Линии нулевой интенсивности показаны для модельных спектров, спектры чистых веществ смещены по вертикали для лучшей визуализации.

Эффективность предложенного подхода продемонстрирована в работе примером разделения 6-ти трехкомпонентных модельных спектров смесей при отсутствии точек неперекрытости в компонентах.

С помощью разработанной процедуры был изучен набор колебательных спектров воды при разных температурах. Этот пример иллюстрирует сложность



Рисунок 8 – Концентрационные вклады компонентов для решения б) (Рисунок 7), рассчитанные по уравнению (1). Номера компонентов соответствуют Рисунку 8.

интерпретации спектров смесей при большой ширине спектральных линий и недостаточном количестве исходных спектров.

Далее в работе рассматривается пример анализа ИК спектров смесей гексафторэтилена с ДМФА, демонстрирующий наиболее неблагоприятный случай сильных межмолекулярных взаимодействий, которые приводят к нарушениям аддитивности спектров. Но даже в этом случае варьирование числа компонентов при разложении позволяет выделить 3 основных спектра. К сожалению, в общем случае, методы анализа смесей взаимодействующих компонентов невозможно формализовать универсальным способом и необходимо разрабатывать индивидуальные алгоритмы.

<u>Раздел 1.5</u> посвящен рассмотрению способа учета априорной информации об относительных концентрациях компонентов при поиске разложения (1) с оценкой надежности решения по каждому из компонентов. Раздел содержит литературный обзор и пример анализа данных ИК-спектров.

В качестве критериев оценки качества (надежности, стабильности) решений задач, рассматриваемых в работе предложено использовать величины критериев, отражающих корреляции (или попарное сходство или различие) между решениями, а также статистические оценки автокорреляции в разностях между соответствующими компонентами в наборах решений, полученных путем варьирования комбинаций данных и параметров алгоритмов поиска.

Во второй части работы рассмотрено определение структурных параметров наноразмерных объектов по данным малоуглового рентгеновского и нейтронного рассеяния.

<u>Раздел 2.1</u> содержит краткий литературный обзор математического и физического аппарата методов интерпретации данных малоуглового рассеяния.

<u>Раздел 2.2</u> посвящен методам определения формы наночастиц и биомолекул в растворе по данным малоуглового рассеяния. В <u>разделе 2.2.1</u> проведено рассмотрение однозначности и неустойчивости решения обратной

19

задачи определения формы рассеивающих объектов по одномерным данным рассеяния. На практике находимые модели частиц могут отличаться друг от друга, если варьировать угловой диапазон данных рассеяния, производить поиск с разных стартовых приближений, изменять параметры самого метода минимизации. Однако разнообразие возможных форм частиц ограничено идентичностью их максимальных размеров, объемов и радиусов инерции, которые достаточно надежно определяются как инварианты кривых рассеяния.

<u>В разделе 2.2.2</u> рассмотрен разработанный с участием автора алгоритм определения формы однородных частиц по данным рассеяния с применением разложения в ряд по сферическим гармоникам. Согласно (Stuhrmann H. B. Interpretation of small-angle scattering functions of dilute solutions and gases // Acta Cryst. Sect. A. – 1970. – V. 26. – No. 3. – Р. 297–306.; Свергун Д(.И., Фейгин Л.А., Щедрин Б.М. Прямой метод интерпретации данных малоуглового рассеяния системами идентичных частиц. // ДАН СССР. –1981. –Т.261. –№4. – С.878–882) форму оболочки $F(\omega)$ представляют как расстояние от центра до границы частицы для каждого углового направления ω и параметризуют разложением в ряд по сферическим гармоникам:

$$F(\omega) = \sum_{l=0}^{L} \sum_{m=-l}^{l} f_{lm} Y_{lm}(\omega)$$
(7)

где f_{lm} - комплексные числа, а пространственное разрешение характеризуется параметром обрыва ряда *L*: $\delta r \approx \pi R_0 / (L+1)$, где R_0 - радиус эквивалентной сферы. Структура частицы представляется с помощью $(L+1)^2$ параметров, а расчет модельной интенсивности рассеяния проводят по формулам, предложенным в указанных работах:

$$I_{L}(s) = \sum_{l=0}^{L} \sum_{m=-l}^{l} |A_{lm}(s)|^{2}, \quad f_{lm}^{(q)} = \int_{\omega} [F(\omega)]^{q} \cdot Y_{lm}^{*}(\omega) \cdot d\omega, \quad A_{lm}(s) = (is)^{l} \cdot \binom{2}{\pi}^{l/2} \times \sum_{p=0}^{p_{max}} \left((-1)^{p} \cdot f_{lm}^{l+2p+3} \cdot \left\{ 2^{p} \cdot p! \cdot (l+2p+3) \cdot [2(l+p)+1]! \right\}^{-1} \cdot s^{2p} \right), \quad (8)$$

Разложение (7) наглядно продемонстрировано на Рисунке 9.



Рисунок 9 - Наглядное представление разложения функции формы частицы в ряд по некоторым сферическим гармоникам A_{lm} . f_{lm} – искомые коэффициенты разложения. Число независимых параметров модели = $(L+1)^2$ -6.

Для поиска параметров f_{lm} в (7) по данным рассеяния с участием автора была разработана программа SASHA [A9]. Целевая функция $\Phi(f_{lm})$ в программе состоит из слагаемого Δ , соответствующего задаче наименьших квадратов, и штрафных членов, накладывающих ограничения на решение:

$$\min_{f_{lm}} \left\{ \Phi(f_{lm}) = \Delta + \sum_{k=1}^{K} g_k^2 \cdot W_k^2 \right\}, \quad \Delta = \frac{\sum_{i=1}^{N} \left[\left(I_{\exp}(s_i) - \xi \cdot I(s_i) \right) \cdot W(s_i) \right]^2}{\sum_{i=1}^{N} \left[I_{\exp}^2(s_i) \cdot W^2(s_i) \right]}, \quad (9)$$

где W(s) - весовая функция, ослабляющая относительный вклад в суммарную невязку начального участка кривой рассеяния (которая может быть искажена рассеянием от агрегатов частиц):

ſ

$$W(s) = \begin{cases} s^{n} & ecnu \quad s > s \Big|_{\max} [I_{\exp}(s) \cdot s^{n}] \\ (n = 1, 2) & , \\ 0.5 \cdot \left\{ s^{n} + \frac{\max[I_{\exp}(s) \cdot s^{n}]}{I_{\exp}(s)} \right\} & ecnu \quad s \le s \Big|_{\max} [I_{\exp}(s) \cdot s^{n}] \end{cases}$$
(10)

i=1,..,*N*, *N* - число точек в экспериментальной кривой; ζ - МНК множитель, совмещающий кривые рассеяния перед вычислением невязки

$$\boldsymbol{\xi} = \left(\boldsymbol{I}_{\exp} \cdot \boldsymbol{I} \right) / \| \boldsymbol{I} \|^2 \,. \tag{11}$$

 $I_{\exp} \cdot I$ означает скалярное произведение векторов экспериментальной и теоретической интенсивностей рассеяния; g_k – набор штрафных членов за отрицательность функции формы, уход центра тяжести тела из начала координат и за негладкость формы (в этом случае K=3). Штрафные члены g_k умножены на соответствующие весовые коэффициенты W_k , величины которых были подобраны на основе решения различных тестовых задач. Важное свойство параметризации (7, 8) - возможность ограничения пространственного разрешения обрывом ряда (7) до некоторого значения l=L (высшие гармоники описывают вклад в рассеяние от внутренних неоднородностей частицы и искажают структурную модель в <u>однородном</u> приближении).

В качестве процедуры минимизации автором применена наиболее эффективная реализация процедуры Левенберга-Маркардта решения задач нелинейных наименьших квадратов, усиленная пересчетом матрицы Гессе целевой функции по схеме переменной метрики – алгоритм NL2SOL (Dennis J. E, Jr., Gay D. M., Welsh R. E. ALGORITHM 573 NL2SOL - An Adaptive Nonlinear Least-Squares Algorithm [E4]. // ACM Trans. on. Math. Soft. – 1981. – V.7. – No.3. – Р. 369–383). Этот алгоритм был модифицирован блоком адаптации

программы к ошибкам вычислений и автоматическим определением оптимальных значений параметров процедуры поиска.

Программа SASHA была опробована на большой серии модельных задач, что позволило выработать ряд практических рекомендаций для исследователей. В частности, было показано, что восстановление формы оболочки частицы может терять устойчивость при числе параметров модели, превышающем $1.5 \times N_{\rm S}$, где $N_{\rm S}$ - число шенноновских каналов

$$N_S = \frac{(s_{\max} - s_{\min}) \cdot D_{\max}}{\pi} + 1.$$
 (12)

С помощью разработанной программы определены формы молекул белков в растворе: гексокиназы, обратной транскриптазы, белкового комплекса SecA [A19] (Рисунок 10).



Рисунок 10 - Определение формы белкового комплекса SecA в растворе. А): экспериментальная и модельные кривые МУР. Пунктир: рассеяние от модели димера, представленного одной формой при *L* = 4; точки: рассеяние от модели димера, составленного из 2-х мономеров, форма которых одинакова и найдена при *L*=3 (тело 1 внизу). Составная модель точнее передает рассеяние в средних углах, увеличенный средний участок (помечен серой областью на рисунке A) показан на рисунке B). C): Полученная модель молекулы SecA из двух мономеров. Модель 2 - повернута относительно модели 1 на 90° вокруг горизонтальной оси, и 3 - вокруг вертикальной.

В конце раздела рассмотрено построение многокомпонентной оболочечной структуры рибосомы 70S *E.coli* в растворе по данным нейтронного малоуглового рассеяния с вариацией контраста. Данная работа выполнялась большим коллективом сотрудников из нескольких организаций. Математическое обеспечение для интерпретации данных рассеяния было создано в Европейской лаборатории молекулярной биологии под руководством Д.И. Свергуна. Вклад автора в данную работу состоял в применении техники сингулярного анализа аддитивных кривых, рассмотренной в разделе 1, для

фильтрации и коррекции экспериментального набора интенсивностей рассеяния, а также в проведении ряда расчетов.

Для вариации контраста в нейтронном рассеянии плотность отдельных компонент и растворителя можно менять путем замещения протонов в биополимерах на ядра дейтерия и меняя долю тяжелой воды в растворе. В данной работе модель рибосомы 70S представляли в виде 4-х компонентной частицы - 2 РНК-ядра и 2 белковые оболочки, для 50S и 30S субъединц, соответственно. Интенсивность рассеяния от такой 4-х компонентной частицы представляят собой сумму 10 слагаемых:

$$I(s) = \sum_{n=1}^{4} (\Delta \rho_n)^2 I_{nn}(s) + 2 \sum_{n>k} \Delta \rho_n \Delta \rho_k I_{nk}(s),$$

где n и k - индексы компонент. Тогда интенсивности рассеяния от образцов селективно дейтерированной рибосомы (различные комбинации дейтерированных и протонированных белковой и РНК-компонент с разными долями тяжелой воды в растворителе, всего 42 экспериментальные кривые) должны представлять собой 10-компонентные данные. Первые 15 сингулярных векторов матрицы данных рассеяния показан на Рисунке 11. Ошибки измерений были скорректированы с помощью восстановления матрицы данных по 10 значимым сингулярным векторам. Величина критерия Дарбина-Ватсона, рассчитанного для матрицы остатков, была равна 1.7, что свидетельствует о наличии малой автокорреляции в остатках с уровнем значимости не более 0.05.



Рисунок 11 - Первые 15 сингулярных векторов матрицы данных нейтронного рассеяния, представленные в виде контуров относительно вектора рассеяния. Кривые смещены по вертикали для большей наглядности. Первый сингулярный вектор не показан: он представляет собой усредненный контур кривых рассеяния. Из рисунка видно, что систематическое поведение контуров заметно до 10 вектора включительно.

Таким образом, стало возможным восстановление формы компонент рибосомы с точностью до шумов эксперимента. Это тем более важно, поскольку данные измерений представляли собой 42 кривые рассеяния, то есть степень переопределенности задачи была достаточно высокой.

Найденная по скорректированным данным 4-х компонентная оболочечная модель рибосомы показана на Рисунке 12.

Аналогичным образом были обработаны и скорректированы данные рассеяния от образцов рибосомы 30S *Thermus Thermophilus*, структура которой была найдена с помощью другого метода - шариковых моделей - адаптированного для 4-х компонентного случая (раздел 2.2.4).



Рисунок 12 - Найденная по данным нейтронного рассеяния оболочечная модель частицы рибосомы *E.coli* 70S. Полупрозрачные участки соответствуют областям, занятым белками, расположенными вокруг РНК-ядер. Вторая и третья модель: структуры повернуты на 22.5° и 50° вокруг вертикальной оси, соответственно.

Метод шариковых моделей для поиска формы частиц по данным малоуглового рассеяния рассмотрен в разделе 2.2.3. Основы подхода были опубликованы в (Svergun D. I. Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. // Biophys. J. -1999. - V. 76. - Р. 2879-2886). Автор диссертационной работы участвовал в разработке вариантов программы, названной DAMMIN, используемых в работе. Алгоритм основан на представлении структуры частиц с помощью набора малых сфер ("шариков") с заданной плотностью, и нахождении их взаимного расположения в пространстве методом случайного поиска. Набор координат центров сфер задается плотнейшей упаковкой внутри области поиска заданного определяют ПО данным рассеяния. размера, который Интенсивность малоуглового рассеяния от модели вычисляют по формуле

$$I(s) = 2\pi^2 \sum_{l=0}^{L_{\text{max}}} \sum_{m=-l}^{l} \left[A_{lm}(s) \right]^2$$
(13)

где

$$A_{lm}(s) = i^l \cdot \sqrt{2/\pi} \cdot f(s) \cdot \sum_{j=1}^{N_a} j_l(sr_j) \cdot Y_{lm}^*(\omega_j)$$
(14)

есть амплитуда рассеяния структурой, состоящей из N_a шариков единичной плотности, r_j, ω_j - их полярные координаты, f(s) - амплитуда рассеяния одним шариком, $j_l(x)$ - сферические функции Бесселя, $Y_{lm}(\omega)$ - сферические гармоники. Автором предложено задавать максимальный индекс гармоники L_{max} не менее $1.5 \times N_S$ (число шенноновских каналов (12)), чтобы ослабить систематические ошибки вычисления интенсивности из-за обрыва ряда (13).

Для поиска пространственного расположения объемных элементов в программе DAMMIN применен метод глобальной минимизации, известный под названием "моделирование отжига", или алгоритм Метрополиса, в котором структура модифицируется на каждом шаге путем вариации заселенности случайным образом выбранного пространственного узла шариком тела или растворителя. Виртуальная "температура" Τ. задаваемая как параметр алгоритма, выбирается автоматически из оценок дисперсии целевой функции при вариациях. Этот параметр, уменьшающийся по мере поиска, заставляет программу в качестве следующей модели принимать худшую с вероятностью, пропорциональную Т и, тем самым, выходить из локальных минимумов. Минимизируемая целевая функция задачи аналогична (9), но с другим набором штрафных членов:

$$\Phi(\mathbf{B}) = \Delta^2 + w_D \cdot P_D + w_L \cdot P_L + w_C \cdot P_C + w_G \cdot P_G, \qquad (15)$$

где **В** - вектор индексов фаз для каждого узла в области поиска, Δ^2 - невязка. Весовую функцию W(s) можно назначать в виде (10) при n = 0,1,2,3,4. На практике *n* рекомендуется выбирать таким образом, чтобы отношение максимальной интенсивности к минимальной было в интервале 10 - 50.

Для адекватного описания формы частицы и исключения влияния упаковки шариков, их диаметр должен быть достаточно мал, не более 1/20 от максимального размера частицы (области поиска). Это соответствует числу узлов в области поиска не менее 2 - 3 тысяч для однородных глобулярных частиц и до 20 - 30 тысяч для анизометричных тел сложной формы. Штрафные члены в (15), уменьшаются по мере поиска в той же степени, как виртуальная температура, и отражают требование неразрывности структуры w_DP_D , отсутствия отдельно расположенных шариков w_LP_L (штраф за "рыхлость"), а w_CP_C - уход центра связной части тела из начала координат. Штраф P_G за радиус инерции R_G представляет собой отношение разности радиуса инерции связного части текущей модели и радиуса инерции, предварительно найденного по данным рассеяния, к диаметру области поиска D_{max} .

Автором был введен также новый режим поиска [A63], который состоит в следующем. Как только число успешных испытаний становится при данной температуре меньше 0.3 от максимального числа проб, заданного в качестве критерия перехода к следующей, более низкой, температуре, программа начинает пропускать узлы растворителя, не контактирующие с шариками частицы. Аналогичный переход производится и в случае, когда относительное число изолированных друг от друга шариков структуры становится менее 3% от общего числа узлов. Такой подход позволил существенно повысить эффективность поиска (относительное число успешных проб выросло с 6 - 9% до 16 - 20% по отношению к общему числу испытаний).

В разделе 2.2.3.2 приведен анализ результатов численного исследования устойчивости определения строения частиц методом шариковых моделей. проведенного на большом наборе тел различной формы. Моделирование позволило сделать несколько выводов. В частности, плоские частицы, как правило, сложнее восстанавливать, чем удлиненные. Степень допустимой анизометрии для надежной реконструкции формы не должна превышать 10:1 для однородных вытянутых и 5:2 для плоских частиц. Концентрические пустоты в структуре восстанавливаются значительно лучше, чем ацентрические. Для численной оценки устойчивости и нахождения наиболее вероятной структуры с участием автора были разработаны программы, позволяющие проводить автоматическое совмещение и усреднение найденных структурных моделей с анализом разброса решений.

Со времени создания программ поиска формы частиц DAMMIN была применена различными исследователями во всем мире для определения морфологии многих десятков биомолекул. С помощью аналогичного алгоритма было проведено исследование структуры белковой фазы частицы 30S *Thermus Thermophilus* в растворе методом малоуглового рассеяния с вариацией контраста (раздел 2.2.4).

Раздел 2.2.5 посвящен исследованию морфологии макромолекул иммуноглобулинов и ревматоидного фактора человека методом малоуглового рентгеновского рассеяния. В структурном отношении как иммуноглобулин М (IgM), так и ревматоидный IgM-RF являются макромолекулами, состоящими из пяти IgG-подобных мономеров. На Рисунке 13 слева показана модель IgM, построенная Перкинсом и соавторами (Perkins S. J., Nealis A S., Sutton B. J., Feinstein A. Solution Structure of Human and Mouse Immunoglobulin M by Synchrotron X-ray Scattering and Molecular Graphics Modelling. // J. Mol. Biol. -1991. - V.221. - Р. 1345 - 1351) по данным малоуглового рассеяния. Как было показано, восстановление таких плоских структур с большим количеством деталей возможно, если на форму частиц накладывать определенные ограничения. В случае молекул IgM таким ограничением в наших работах было 5-го В наличие оси симметрии порядка. соответствии с нашей методологической рекомендацией, было проведено исследование устойчивости восстановления формы IgM по модельным данным рассеяния, рассчитанным от структуры 2RCJ по формулам (13, 14). Моделирование показало, что для данного типа плоских структур поиск достаточно вести более ограниченной (плоской) области поиска (Рисунок 13) и передать основные особенности формы молекулы достаточно устойчиво.

Работа по исследованию растворов молекул иммуноглобулинов G, M, RF и их фрагментов была выполнена в сотрудничестве с институтами органической химии РАН, молекулярной биологии РАН, биоорганической

химии РАН, и Гематологическим научным центром РАМН. Измерения малоуглового рентгеновского рассеяния и анализ данных проводились в Европейской лаборатории молекулярной биологии (EMBL c/o DESY, Гамбург) и в Институте кристаллографии РАН [A27,A31,A61,A63]. Автором были проведены измерения рассеяния, обработка и интерпретация основной части полученных данных.



Рисунок 13 - Пример восстановления формы 2RCJ (слева) по модельным кривым рассеяния при числе каналов $N_8 = 12$ (три структуры справа). Структура справа показана внутри области поиска. Модели показаны в двух ориентациях.

Из Рисунка 14 видно, что различие между кривыми малоуглового рассеяния от образцов IgM и IgM-RF незначительны, однако полученные модели отличаются: заселенность периферийных областей IgM-RF меньше, чем для IgM (Рисунок 15). Это подтверждается различиями радиальных распределений для структур (Рисунок 16). Был сделан вывод о большей гибкости Fab-фрагментов (Рисунок 15 слева) молекул IgM-RF, что приводит к их большей подвижности и, в результате, к меньшей заселенности этих участков в усредненных моделях, находимых по данным рассеяния.

Вывод о возможной большей гибкости Fab-областей в IgM-RF был подтвержден в работе результатами исследования выделенных Fab и Fab-RF методами ЭПР и масс-спектроскопии [A44]. Тем не менее, малоугловые исследования образцов показали, что средние формы отдельных Fab и Fab-RF фрагментов практически совпадают в растворе.

<u>Раздел 2.3</u> посвящен разработке и применению методов расчета распределений по размерам наночастиц в полидисперсных смесях различной природы. Расчет распределений проводится для заданных формфакторах частиц, так как одновременное определение распределений и формфакторов представляет собой математически недоопределенную задачу. В некоторых случаях (если возможно получение образцов с узким распределением, не превышающим по полуширине 5-10%) определение формы частиц может быть проведено с помощью методов, рассмотренных выше.

Автором был разработан вариант программы поиска параметров

распределений (POLYMIX), в которой применен метод минимизации NL2SOL (см. Раздел 2.2.2). Применение адаптивных алгоритмов численного интегрирования в расчетах модельной интенсивности рассеяния позволило в несколько раз повысить точность аппроксимации данных. Автором была проведена еще одна существенная модификация программы, состоящая в том, что расчет невязки В методе наименьших квадратов проводят на неэквидистантной угловой сетке.



Рисунок 14 - Сравнение экспериментальных данных рассеяния растворами IgM (1 - 3, черные точки) и IgM-RF (кружки). Для большей наглядности пары кривых, измеренных в одинаковых условиях, смещены по вертикали.

Рисунок 15 - Сравнение типичных моделей, найденных по данным рассеяния, с моделью 2RCJ (а). (b): два решения IgM (DAMMIN). (c, d): 4 решения для IgM-RF (DAMMIN). (е): модель IgM-RF из аминокислотных остатков (программа GASBOR). (f): моделирование IgM-RF доменами, из которых состоит структура 2RCJ.

Рисунок 16 - Усредненные функции радиального распределения (атомов в 2RCJ или шариков в найденных моделях), *r* - расстояние от центра тяжести частиц. (1): 2RCJ; (2): усредненные решения (см. Рисунок 15); (3): данные от растворов IgM1-IgM3; (4): образцы IgM-RF1 - IgM-RF3; (5): IgM-RFq. Разбросы рассчитаны по 10-30 решениям для каждого образца и отражают численную стабильность расчетов структурных моделей.

Эквидистантная угловая сетка может приводить к совпадению периода осцилляций модельных кривых рассеяния от частиц с периодом сетки, что резко уменьшает численное значения градиента целевой функции и приводит к необоснованно большой величине шага по параметрам модели. Это приводит к нарушению хода поиска и к большой зависимости решения от стартовых значений параметров модели. Практика показала, что хорошие результаты дает сетка с шагом, увеличивающимся по закону геометрической прогрессии. В программе POLYMIX реализовано определение начального и конечного шага угловой сетки по величинам шенноновских каналов, которые, в свою очередь, определяются по радиусам инерции (в предположении сферических частиц, которое, как показала практика, дает приемлемые результаты и в случае анизометричных тел). При этом радиусы инерции рассчитываются по начальному участку и в области максимальных углов. Расчет невязки по такой сетке обладает еще одним преимуществом, так как позволяет лучше описать начальный участок углов рассеяния, в котором располагается теперь значительное количество точек.

Малоугловое рассеяние от смеси, содержащей различные типы невзаимодействующих компонентов, представляет собой сумму парциальных интенсивностей рассеяния этими компонентами, взвешенных согласно их объемным долям и может быть представлено в виде:

$$I(s) = \sum_{k=1}^{K} V_k I_{k0}(s), \quad I_{k0}(s) = \int_{0}^{\infty} D_k(R) v_k(R) \left[\Delta \rho_k(R)\right]^2 i_{0k}(s,R) S_k(s) dR \quad (16)$$

где $I_{k0}(s)$ - интенсивность рассеяния *k*-ой компонентой. Функции объемного распределения $D_k(R)$ в данной работе представлены нормированным распределением Цимма-Шульца, описываемым двумя параметрами, средним значением R_{0k} и дисперсией ΔR_k :

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

где $z = (R_{0k} / \Delta R_k)^2 - 1$. В (16) межчастичную интерференцию учитывают путем расчета структурного фактора $S_k(s)$ в приближении Перкуса-Йовика.

Для определения объемных долей и размерных параметров частиц решают задачу подгонки данных эксперимента полидисперсной моделью с нелинейной минимизации функционала, помощью аналогичного (9). Использование масштабирующего множителя (11) обеспечивает сглаживание многих локальных минимумов целевой функции, делая решение задачи Для назначения оптимального практически возможным. динамического рассеяния W. диапазона применяют весовую функцию данных соответствующую степенному преобразованию интенсивности рассеяния, то есть невязку (9) рассчитывают в шкале I(s), $I^{1/2}(s)$, $I^{2/3}(s)$ или $I^{3/4}(s)$.

На Рисунке 17 показаны примеры анализа данных рассеяния от образца битума. Рассеяние в данной системе обусловлено в основном наличием кластеров непредельных соединений, которые расположены довольно близко друг от друга, о чем говорит широкий интерференционный пик в области 2.5 нм⁻¹. Поэтому при расчетах вводили структурные факторы для двух компонент (использовали 3 системы сферических частиц, 3-я система описывала фон рассеяния от малых образований с размерами менее 1 нм), параметры которых предметом поиска. Из также были рисунка видно преимущество предложенного подхода с расчетом невязки на неэквидистантной угловой сетке. χ^2 для эквидистантной сетки был 19.0, для неэквидистантной 3.0. от 1.0 связано с осциллирующим поведением функции Отличие его структурного фактора, которое свойственно аппроксимации Перкуса-Йовика. К сожалению, более общие выражения для функций межчастичной интерференции в до настоящего времени получить не удалось. С помощью разработанного программного обеспечения было проведено исследование нескольких десятков систем различной природы: микроэмульсий воды в органических растворителях, металлических наночастиц в полимерах И углеродных матрицах и т.д., что отражено в списке работ автора.



Рисунок 17 - Сравнение распределений по размерам кластеров непредельных органических соединений, рассчитанных по данным малоуглового рассеяния от образца битума. Слева показаны экспериментальные (точки) и модельные (сплошные линии) интенсивности рассеяния на угловых сетках, использованных при расчете невязки: а) - эквидистантная, в) - предложенная в работе неэквидистантная. Справа - показаны по два решения, полученные при разных стартовых аппроксимациях (наборы стартовых значений одинаковы для б) и *г*)). Модельные кривые рассеяния от двух показанных решений на графиках неразличимы, несмотря на различия распределений.

Основные результаты и выводы

1. Разработаны алгоритмы и программы решения обратных задач анализа наборов аддитивных неотрицательных спектров смесей, основанные на сингулярном разложении матрицы данных. Разработанные алгоритмы включают оценку числа независимых компонентов в смесях, определение контуров индивидуальных спектров и способы исследования устойчивости решений. Показана возможность анализа систем с числом априори неизвестных компонентов более 3-х, в зависимости от качества экспериментальных данных. Разработанные подходы реализованы в виде компьютерных программ и применены для анализа ИК спектров смесей органических соединений, колебательных спектров жидкой воды, анализа серий измерений интенсивности малоуглового рентгеновского рассеяния.

2. Разработанные способы разложения аддитивных спектров впервые применены, в частности, для анализа и коррекции данных малоуглового нейтронного рассеяния с вариацией контраста, по которым были определены структурные модели селективно дейтерированных частиц рибосомы 70S *E-Coli* и 30S *Thermus Thermophilus*.

3. С помощью модельных расчетов проведен анализ единственности и устойчивости решения задач поиска формы частиц в монодисперсных системах по данным малоуглового рентгеновского рассеяния. Определены параметры анизометрии простых геометрических форм, являющиеся пограничными для устойчивых решений.

Программа шариковых 5. поиска моделей частиц по данным рассеяния модифицирована способами малоуглового автоматического определения оптимальных значений параметров и режимов алгоритма поиска, что позволило уменьшить время расчетов в 3-5 раз и уменьшить разброс решений. С помощью разработанного программного обеспечения по данным малоуглового рассеяния раствора определены OT формы молекул иммуноглобулина М и ревматоидного фактора человека. Показано, ЧТО макромолекулами состоит В меньшей различие между заселенности периферийных участков молекул ревматоидных факторов. Выдвинута гипотеза о вероятной большей гибкости периферийных фрагментов ревматоидного фактора, что независимо подтверждено другими методами исследований.

Разработана модифицированная 6. схема поиска размерных распределений частиц по данным малоуглового рассеяния с повышенной устойчивостью решений. Повышение устойчивости достигнуто путем расчета невязки между экспериментальными И модельными интенсивностями рассеяния на неравномерной угловой сетке с определением шага сетки из оценок максимального и минимального радиусов инерции частиц в системе. Проведен анализ ряда систем с наночастицами – металлических катализаторов на органических И углеродных носителях, органико-неорганических композитов, полимерных систем.

Публикации по теме диссертации

- А1. Гречушников, Б.Н. Об однозначности решения задачи разделения аддитивных спектров смесей на спектры компонентов / Б. Н. Гречушников, И. Н. Калинкина, В. В. Волков, И. П. Петров // Журнал прикладной спектроскопии. 1984. Т. 40. Вып. 1. С. 89 92.
- А2. Волков, В. В. Применение преобразования проектирования в методе варьирования концентраций компонентов при анализе многокомпонентных смесей по спектрам поглощения / В. В. Волков, Б. Н. Гречушников // Журнал прикладной спектроскопии. – 1984. – Т. 40. – Вып. 2. – С. 264 – 267.
- А3. Волков, В. В. К вопросу оценивания спектров компонентов по спектрам родственных смесей / В. В. Волков // Журнал прикладной спектроскопии. 1988. Т. 48. Вып. 3. С. 436 442.
- А4. Волков, В. В. Коррекция решения задачи численного разложения спектров / В. В. Волков, Б. Н. Гречушников, И. Н. Калинкина // Журнал прикладной спектроскопии. 1988. Т. 49. Вып. 2. С. 320 323.
- А5. Волков, В. В. Оценивание спектров компонентов по набору аддитивных спектров смесей / В. В. Волков // Математические методы и ЭВМ в аналитической химии: сборник / Л. А. Грибов (ред.) – М. : Наука, 1989. – 302 с. – С. 42–50. – ISBN 5–02–001366–8.
- А6. Шевчук, Т. С. Изучение комплексообразования в системе NiCl₂– CO(NH₂)₂–H₂O методом электронной спектроскопии путем разложения спектральных контуров на аддитивные составляющие / Т. С. Шевчук, А. Ф. Борина, В. В. Волков // ЖНХ. – 1992. – Т. 37. – Вып. 6. – С. 1355–1361.
- A7. Schmidt, B. Small-angle X-ray solution scattering study on the dimerization of the FKBP25mem from *Legionella pneumophila* / B. Schmidt, S. Koenig, D. Svergun, V. Volkov, G. Fisher, M. H. J. Koch // FEBS Letters. 1995. V. 372. P. 169 172.
- A8. Volkov, V. V. Separation of Additive Mixture Spectra by a Self-Modeling Method / V. V. Volkov // Applied Spectroscopy. - 1996. - V. 50. - No. 3. -P. 320 - 326.
- A9. Svergun, D. I. New Developments in Direct Shape Determination from Small– Angle Scattering. 2. Uniqueness / D. I. Svergun, V. V. Volkov, M. B. Kozin, H. B. Stuhrmann // Acta Cryst. A. – 1996. – V. 52. – P. 419 – 426.
- А10. Юхневич, Г. В. Полоса валентных колебаний и структура жидкой воды / Г. В. Юхневич, В. В. Волков // Доклады АН. – 1997. – Т. 353. – № 4. – С. 465 – 468.
- A11. Svergun, D. I. Solution Scattering Structural Analysis of the 70S *Escherichia coli* Ribosome by Contrast Variation: I. Invariants and Validation of Electron Microscopy Models / D. I. Svergun, N. Burkhardt, J. S. Pedersen, M. H. J. Koch, V. V. Volkov, M. B. Kozin, W. Meerwink, H. B. Stuhrmann, G. Diedrich,

K. H. Nierhaus // J. Mol. Biol. – 1997. – V. 271. – P. 588 – 601.

- A12. Svergun, D. I. Solution Scattering Structural Analysis of the 70S *Escherichia coli* Ribosome by Contrast Variation: II. Structure modelling / D. I. Svergun, N. Burkhardt, J. S. Pedersen, M. H. J. Koch, V. V. Volkov, M. B. Kozin, W. Meerwink, H. B. Stuhrmann, G. Diedrich, K. H. Nierhaus // J. Mol. Biol. 1997. V. 271. P. 602 618.
- A13. Burkhardt, N. Architecture of the 70S *Escherichia coli* Ribosome / N. Burkhardt, G. Diedrich, K. H. Nierhaus, W. Meerwink, H. B. Stuhrmann, Skov J. Pedersen, M. H. J. Koch, V. V. Volkov, M. B. Kozin, D. I. Svergun // Physica B. – 1997. – V. 234–236. – P. 199 – 201.
- A14. Svergun, D. I. Shape Determination from Solution Scattering of Biopolymers / D. I.Svergun, V. V. Volkov, M. B. Kozin, H. B. Stuhrmann, C. Barberato, M. H. J. Koch // J. Appl. Cryst. 1997. V. 30. P. 798 802.
- A15. Kozin, M. B. ASSA a Program for Three–Dimensional Rendering in Solution Scattering from Biopolymers / M. B. Kozin, V. V. Volkov, D. I. Svergun // J. Appl. Cryst. – 1997. – V. 30. – P. 811 – 815.
- A16. Koenig, S. Small-Angle X-ray Solution-Scattering Studies on Ligand-Induced Subunit Interactions of the Thiamine Diphosphate Dependent Enzyme Pyruvate Decarboxylase from Different Organisms / S. Koenig, D. I. Svergun, V. V. Volkov, L. A. Feigin, M. H. J. Koch // Biochemistry. - 1998. - V. 37. -P. 5329 - 5334.
- A17. Svergun, D. I. Contrast Variation X–ray and Neutron Structural Studies of the 70S Ribosome from *E. coli* and *Thermus Thermophilus* / D. I. Svergun, V. V. Volkov, J. Skov Pedersen, M. H. J. Koch, M. B. Kozin, H. B. Stuhrmann, I. N. Serdyuk, L. Fan, V. L. Lamzin, K. H. Nierhaus // Structure Research of the Ribosome and its Functional Complexes / Ed. by R. Willumeit and S. Forthmann. GKSS: Geesthacht. 1998. P. 19 23.
- A18. Bilgin, N. Solution Structure of the Ternary Complex between AminoacyltRNA, Elongation Factor Tu, and Guanosine Triphosphate / N. Bilgin, M. Ehrenberg, C. Ebel, G. Zaccai, Z. Sayers, M. H. J. Koch, D. I. Svergun, C. Barberato, V. Volkov, P. Nissen, J. Nyborg // Biochemistry. – 1998. – V. 37. – No. 22. – P. 8163 – 8172.
- A19. Shilton, B. *Escherichia coli* SecA shape and dimensions / B. Shilton,
 D. I. Svergun, V. V. Volkov, M. H. J. Koch, S. Cusack, A. Economou // FEBS
 Lett. 1998. V. 436. P. 277 282.
- A20. Macheroux P. Evidence for a major structural change in *Escherichia coli* chorismate synthase induced by flavin and substrate binding / P. Macheroux, E. Schoenbrunn, D. I. Svergun, V. V. Volkov, M. H. J. Koch, S. Bornemann, R. N. F. Thorneley // Biochem. J. 1998. V.335. P. 319 327.
- A21. Svergun, D. I. Quaternary structure of V1 and F1 ATPase: significance of structural homologies and diversities / D. I. Svergun, S. Konrad, M. Huss, M. H. J.

Koch, H. Wieczorek, K.–H. Altendorf, V. V. Volkov, G. Grueber // Biochemistry. – 1998. – V. 37. – P. 17659 – 17663.

- А22. Свергун, Д. И. Автоматическое определение формы частицы в малоугловом рассеянии монодисперсными системами./ Д. И. Свергун, В. В. Волков, М. Б. Козин // Поверхность. Рентгеновские, синхротронные и нейтронные исследования. 1999. № 9. С. 3 7.
- A23. Svergun, D. I A small angle x-ray scattering study of the droplet-cylinder transition in oil-rich sodium bis(2-ethylhexyl) sulfosuccinate microemulsions / D. I. Svergun, P. V. Konarev, V. V. Volkov, M. H. J. Koch, W. F. C. Sager, J. Smeets, E. M. Blokhuis // J. Chem. Phys. 2000. V. 113. P. 1651 1665.
- A24. Svergun, D. I. Small–Angle X–Ray Scattering Study of Platinum–Containing Hydrogel/Surfactant Complexes / D. I. Svergun, E. V. Shtykova, M. B. Kozin, V. V. Volkov, A. T. Dembo, E. V. Jr. Shtykova, L. M. Bronstein, O. A. Platonova, A. N. Yakunin, P. M. Valetsky, A. R.Khokhlov // J. Phys. Chem. B. – 2000. – V. 104. – No. 22. – P. 5242 – 5250.
- A25. Svergun, D. I. Formation of metal Nanoparticles in Multilayered Poly(Octadecylsiloxane) as Revealed by Anomalous Small–Angle X–Ray Scattering / D. I. Svergun, M. B. Kozin, P. V. Konarev, E. V. Shtykova, V. V. Volkov, D. M. Chernyshov, P. M. Valetsky, L. M. Bronstein // Chemistry of Materials. 2000. V. 12. No. 12. P. 3552 3560.
- A26. Fan, L. Structural studies of the 30S subunit of ribosomes *Thermus thermophilus* by small–angle neutron and X–ray scattering / L. Fan, D. I. Svergun, V. V. Volkov, V. L. Aksenov, C. Ch. Agalarov, O. M. Selivanova, I. N. Shcherbakova, M. H. J. Koch, R. Gilles, A. Wiedenmann, R. May, I. N. Serdyuk // J. Appl. Crystallogr. 2000. V. 33. P. 515 518.
- A27. Volkov, V. V. Low-resolution structure of immunoglobulins IgG1, IgM and rheumatoid factor IgM-RF from solution X-ray scattering data / V. V. Volkov, V. A. Lapuk, R. L. Kayushina, E. V. Shtykova, E. Yu. Varlamova, M. Malfois, D. I. Svergun. // J.Appl.Cryst. 2003. V.- 36. P. 503 508.
- A28. Volkov, V. V. Uniqueness of *ab initio* shape determination in small–angle scattering / V. V. Volkov, D. I. Svergun // J. Appl. Cryst. 2003. V.36. P. 860 864.
- A29. Shtykova, E. V. Small-angle X-ray scattering reveals hollow nanostructures in iota- and kappa-carrageenan/surfactant complexes / E. V. Shtykova, E. V. Jr. Shtykova, V. V. Volkov, P. V. Konarev, A. T. Dembo, E. E. Makhaeva, I. A. Ronova, A. R. Khokhlov, H. Reynaers, D. I. Svergun // J. Appl. Cryst. 2003. V. 36. P. 669 673.
- A30. Sokolova, A. V. Prototype of a database for rapid protein classification based on solution scattering data / A. V. Sokolova, V. V. Volkov, D. I. Svergun // J. Appl. Cryst. – 2003. – V. 36. – P. 865 – 868.
- А31. Волков, В. В. Определение строения иммуноглобулинов человека IgG,

IgM и ревматоидного фактора IgM–RF в растворе / В. В. Волков, Р. Л. Каюшина, В. А. Лапук, Э. В. Штыкова, Е. Ю. Варламова, М. Мальфуа, Д. И. Свергун // Кристаллография. – 2003. – Т. 48. – № 1. – С. 103 – 110.

- A32. Starodoubtsev, S. G. Poly(acrylamide) gels with embedded magnetite nanoparticles / S. G. Starodoubtsev, E. V. Saenko, A. R. Khokhlov, V. V. Volkov, K. A. Dembo, V. V. Klechkovskaya, E. V. Shtykova, I. S. Zanaveskina // Microelectronic Engineering. 2003. V. 69. P. 324 329.
- A33. Dembo, A. T. Structure of Binary and Ternary Complexes Formed by Sodium Poly(2–acrylamide–2–methyl–1–1propanesulfonate) Gel in the Presence of Copper(II) Nitrate and Cetylpyridinium Chloride / A. T. Dembo, K. A. Dembo, V. V. Volkov, A. I. Kokorin, A. A. Lyubimov, E. V. Shtykova, S. G. Starodoubtsev, A. R. Khokhlov // Langmuir. 2003. V. 19. P. 7845 7851.
- A34. Konarev, P. V. PRIMUS: a Windows PC-based system for small-angle scattering data analysis / P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, D. I. Svergun. // J Appl. Crystallogr. – 2003. – V. 36: – P. 1277 – 1282.
- А35. Кононенко, А. В. Молекулярная морфология фактора терминации трансляции 1–го класса эукариот eRF1 в растворе / А. В. Кононенко, К. А. Дембо, Л. Л. Киселев, В. В. Волков // Молекуляр. биология. – 2004. – Т. 38, – С. 303 – 311.
- A36. Ozerin, A. N. The spatial structure of dendritic macromolecules / A. N. Ozerin, D. I. Svergun, V. V. Volkov, A. I. Kuklin, V. I. Gordeliy, A. Kh. Islamov, L. A. Ozerina, D. S. Zavoroynyuk // J. Appl. Crystallogr. 2005. V. 38, P. 996 1003.
- АЗ7. Юрков, Г. Ю. Железосодержащие наночастицы в матрице силоксановых каучуков / Г. Ю. Юрков, Д. А. Астафьев, Л. Н. Никитин, Ю. А. Кокшаров, Н. А. Катаева, Э. В. Штыкова, К. А. Дембо, В. В. Волков, А. Р. Хохлов, С. П. Губин // Неорганические материалы. 2006. Т. 42, № 5, С. 556 562.
- АЗ8. Кочервинский, В. В. Роль внутрицепных дипольных взаимодействий при формировании надмолекулярной структуры кристаллизующихся сегнетоэлектрических полимеров / В. В. Кочервинский, В. В. Волков, К. А. Дембо // Физика твердого тела. – 2006. – Т. 48. – Вып. 6. – С. 1019 – 1021.
- A39. Konarev, P. V. ATSAS 2.1, a program package for small–angle scattering data analysis / P. V. Konarev, M. V. Petoukhov, V. V. Volkov, D. I. Svergun // J. Appl. Crystallogr. – 2006. – V. 39. – P. 277 – 286.
- A40. Shtykova, E. V. Autoxidized Phospholipids in Hexane: New Types of Nanoparticles Studied by Synchrotron Small–Angle X–ray Scattering /

E. V. Shtykova, V. V. Volkov, H. Wang, T. Fujisawa, Jin–Ye Wang. // Langmuir. – 2006. – V. 22. – No. 19. – P. 7994 – 8000.

- А41. Никитин, Л. Н. Структура композитов, полученных формированием полипиррола в сверхкритическом CO₂ на микропористом полиэтилене / Л. Н. Никитин, М. О. Галлямов, А. Ю. Николаев, Э. Е. Саид–Галиев, А. Р. Хохлов, С. С. Букалов, Г. И. Магдануров, В. В.Волков, Э. В. Штыкова, К. А. Дембо, Г. К. Ельяшевич // Высокомолек. соед. А. 2006. Т. 48. № 8. С. 1431 1447.
- A42. Klechkovskaya, V. V. Network model of Acetobacter Xylinum Cellulose Intercalated by Drug Nanoparticles / V. V. Klechkovskaya, V. V. Volkov, E. V. Shtykova, N. A. Arkharova, Yu. G. Baklagina, A. K. Khripunov, R. Yu. Smyslov, L. N. Borovikova, A. A. Tkachenko // Nanomaterials for Application in Medicine and Biology / Eds. By M. Giersig and G. B. Khomutov. NATO Science for Peace and Security Series B: Physics and Biophysics. Springer Netherlands. – 2008. – P. 165 – 177.
- A43. Yevdokimov, Yu. M. Transition in DNA molecule's spatial ordering due to nano-scale structural changes / Yu. M. Yevdokimov, V. I. Salyanov, E. V. Shtykova, K. A. Dembo, V. V. Volkov, P. V. Spirin, A. S. Slusheva, V. S. Prassolov // The Open Nanoscience Journal. 2008. V. 2. P. 17 28.
- А44. Волков, В. В. Особенности структуры Fab-фрагментов ревматоидного фактора IgM-RF в растворе / В. В. Волков, В. А. Лапук, Э. В. Штыкова, Н. Д. Степина, К. А. Дембо, А. В. Соколова, С. В. Амарантов, В. П. Тимофеев, Р. Х. Зиганшин, Е. Ю. Варламова // Кристаллография. 2008. Т. 53. № 3. С. 476 484.
- A45. Aslanov, L. Mesoporous soot–supported palladium as a heterogeneous catalyst for the Heck reaction in ionic liquids / L. A. Aslanov, Y. A. Kabachii, S. Yu. Kochev, B. V. Romanovsky, P. M. Valetsky, V. V. Volkov, A. V. Yatsenko, V. N. Zakharov // Mendeleev Communications. 2008. V. 18. No. 6. P. 334 335.
- A46. Pastukhov, A. V. Magnetic Nanocomposites Based on Hypercrosslinked Polystyrenes / A. V. Pastukhov, V. A. Davankov, V. V. Volkov, K. A. Dembo, Y. V. Zubavichus, A. A. Korlyukov, A. G. Filatova. // Bulletin of the Russian Academy of Sciences: Physics. -2009. - V. 73. - No. 4. - P. 471 - 473.
- А47. Волков, В. В. Определение размера и фазового состава наночастиц серебра в гель–пленке бактериальной целлюлозы методами малоуглового рентгеновского рассеяния, электронной дифракции и электронной микроскопии / В. В. Волков, В. В. Клечковская, Э. В. Штыкова, К. А. Дембо, Н. А. Архарова, Г. И. Ивакин, Р. Ю. Смыслов // Кристаллография. – 2009. – Т. 54. – №2. – С. 197 – 201.
- A48. Starodubtsev, S. G. Montmorillonite-polycation multilayers incorporated in polyacrylamide / S. G. Starodubtsev, E. K. Lavrentyeva, E. V. Shtykova,

K. A. Dembo, V. V. Volkov. // Applied Clay Science. - 2009. - V. 46. - P. 88 - 94.

- А49. Штыкова, Э. В. Формирование наночастиц серебра в матрице ПВП в среде сверхкритического CO₂: малоугловое рентгеновское рассеяние и моделирование / Э. В. Штыкова, К. А. Дембо, В. В. Волков, Э. Е. Саид-Галиев, А. И. Стаханов, А. Р. Хохлов // Российские нанотехнологии. 2009. Т. 4. № 9–10. С. 64 71.
- A50. Starodubtsev, S. G. Comb–like poly(4–vilylpyridinium) salts with dodecylsulfate, sodium bis(2–ethylhexyl) sulfosuccinate and bromide counter ions. Small–angle X–ray scattering and dynamic light scattering study / S. G. Starodubtsev, T. V. Laptinskaya, A. S. Yesakova, A. R. Khokhlov, E. V. Shtykova, K. A. Dembo, V. V. Volkov // Polymer. 2010. V. 51. No. 1. P. 122 128.
- A51. Shtykova, E. V. SAXS-data-based structural modeling of DNA-gadolinium complexes fixed in particles of cholesteric liquid-crystalline dispersions / E. V. Shtykova, V. V. Volkov, V. I. Salyanov, Yu. M. Yevdokimov // Eur. Biophys. J. 2010. V. 39. No. 9. P. 1313 1322.
- A52. Priola, A. Effect of the swelling degree on the formation of magnetite nanoparticles in hydrogels / A. Priola, A. Di Gianni, R. Bongiovannia, S. G. Starodubtsev, S. S. Abramchuck, S. N. Polyakov, V. V. Volkov, E. V. Schtykova, K. A. Dembo // European Polymer Journal. 2010. V. 46. No. 11. P. 2105 2111.
- A53. Babushkina, T. A. Study of the gel films of Acetobacter Xylinum cellulose and its modified samples by 1 H NMR cryoporometry and small-angle X-ray scattering / T. A. Babushkina, T. P. Klimova, E. V. Shtykova, K. A. Dembo, V. V. Volkov, A. K. Khripunov, V. V. Klechkovskaya //Crystallography Reports. 2010. V. 55. No. 2. P. 312 317.
- А54. Логинова, Т. П. Синтез и исследование свойств наночастиц феррита кобальта в гибридных мицеллах полистирол –блок–полэтиленоксида и додецилсульфата натрия / Т. П. Логинова, О. В. Лыхина, Е. А. Юданова, И. А. Хотина, Г. И. Тимофеева, О. Л. Лепендина, В. В. Волков, К. А. Дембо, С. П. Солодовников // Высокомолек. соед. Серия А. – 2010. – Т. 52. – № 8. – С. 1458 – 1465.
- A55. Said–Galiev, E. E. Synthesis of Ag and Cu–Chitosan Metal–Polymer Nanocomposites in Supercritical Carbon Dioxide Medium and Study of Their Structure and Antimicrobial Activity / E. E. Said–Galiev, A. I. Gamzazade, T. E. Grigor'ev, A. R. Khokhlov, N. P. Bakuleva, I. G. Lyutova, E. V. Shtykova, K. A. Dembo, V. V. Volkov // Nanotechnologies in Russia. 2011. V. 6. No. 5–6. P. 341 352. ISSN 1995–0780.
- A56. Skuridin, S. G. Retention of gold nanoparticles in the structure of quasinematic layers formed by DNA molecules / S.G. Skuridin, V. A. Dubinskaya,

E. V. Shtykova, V. V. Volkov, V. M. Rudoy, O. V. Dement'eva, V. A. Kuzmin,
E. S. Lisitsyna, S. T. Zakhidov, I. A. Zelenina, Yu. M. Yevdokimov //
Biochemistry. Supplemental Series A. - 2011. - V. 5. - No. 2. - P. 191 - 197.

- А57. Божко, Н. Н. Исследование особенностей структурообразования в системах поливиниловый спирт – йодид калия – йод методом измерения малоуглового рентгеновского рассеяния / Н. Н. Божко, В. П. Столяров, Е. Б. Баблюк, В. Г. Назаров, В. В. Волков, С. В. Амарантов, К. А. Дембо // Высокомолекулярные соединения. Серия А. – 2011. – Т. 53. – № 9. – С. 1537 – 1541.
- А58. Свергун, Д. И. Рентгеновское малоугловое рассеяние, синхротронное излучение и структура био– и наносистем / Д. И. Свергун, Э. В. Штыкова, В. В. Волков, Л. А. Фейгин // Кристаллография. 2011. Т. 56. № 5. С. 777–804.
- А59. Владимиров, Ю. А. Белково–липидные наночастицы, от которых зависит "быть или не быть" живой клетке / Ю. А. Владимиров, Ю. Ц. Ноль, В. В. Волков // Кристаллография. 2011. Т. 56. № 4. С. 596 603.
- Механизм А60. Брик, В. В. образования, структура адсорбционные И характеристики микропористых нанокристаллических тонких пленок композитов (V,Ti)–N–He / В. В. Брик, Р. Л. Василенко, А. В. Гончаров, А. Г. Гуглия, В. Г. Колобродов, А. В. Грогорова, М. Л. Литвиненко, Е.С. Мельникова, И. Г. Марченко, I. V. Sassa, Ю. В. Никитенко, Ю. Н. Хайдуков, В. В. Волков, G. Chaboussant // Поверхность. - 2011. -No. 6. – C. 66 – 75.
- A61. Volkov, V. V. A comparative study of immunoglobulin IgM and rheumatoid factor IgMRF in solution by small-angle X-ray scattering / V. V. Volkov, V. A. Lapuk, A. I. Toropova, E. Yu. Varlamova, E. V. Shtykova, K. A. Dembo, V. P. Timofeev // Mendeleev Communications. 2012. V. 22. P. 159 161.
- А62. Коригодский, А. Р. Короткоблочные полимочевин-полидиметилоснове силоксановые блок-сополимеры на вторичных диаминов / С. А. Барулина, А. Р. Коригодский, В. В. Киреев, Н. П. Бессонова, В. В. Волков, К. А. Дембо, Э. В. Штыкова, М. А. Ванцян, Е. А. Симакина // Высокомолекулярные соединения. – 2013. – Т. 55. – № 4. – С. 488 – 495.
- A63. Volkov, V. V. Difference between conformations of immunoglobulin M and human rheumatoid factor based on small–angle X–ray solution scattering data / V. V. Volkov, V. A. Lapuk, D. I. Chekrygina, E. Yu. Varlamova, A. V. Chekushin // Mendeleev Comm. 2013. T. 23. P. 319 320.