

УДК: 577.3

## **Биологическая кристаллография без кристаллов**

**Лунин В.Ю.\* , Лунина Н.Л., Петрова Т.Е.**

*Институт математических проблем биологии РАН – филиал Института прикладной математики им. М.В. Келдыша РАН, Пущино, Московская область, Россия*

**Аннотация.** Основным препятствием на пути к определению атомной структуры биологических макромолекул методом рентгеновского структурного анализа является необходимость получения монокристаллов исследуемого объекта. Эта необходимость связана со сложностями экспериментальной регистрации рассеяния отдельной молекулой. Однако получить кристаллы удастся далеко не для всех исследуемых биологических объектов. Развитие техники эксперимента, в частности, ввод в строй рентгеновских лазеров на свободных электронах, позволяет подойти к практическому решению проблемы регистрации рассеяния изолированной частицей изучаемого объекта и, тем самым, к получению информации о трехмерной структуре некристаллических биологических объектов методами рентгеновской дифракции. Дискретизация экспериментальных данных рассеяния изолированной частицей делает задачу определения структуры объекта эквивалентной задаче биологической кристаллографии, что позволяет распространить методы биологической кристаллографии на исследование изолированных биологических частиц (отдельных клеток, органелл, молекулярных машин и, в перспективе, биологических макромолекул). В данной статье обсуждается положение дел в этой области, возникающие проблемы и пути их решения.

**Ключевые слова:** *рентгеновская кристаллография, фазовая проблема, XFEL, рассеяние изолированной частицей.*

### **ВВЕДЕНИЕ**

Рентгеноструктурный анализ монокристаллов является на сегодняшний день основным экспериментальным методом, позволяющим определять структуры биологических макромолекул и их комплексов с атомным разрешением. Этот метод широко применяется на практике и, в настоящее время, открытый для доступа банк белковых структур – Protein Data Bank – содержит информацию о более сотни тысяч макромолекулярных структур, определенных таким путем. Альтернативные подходы (электронная микроскопия, ядерный магнитный резонанс, дифракция нейтронов и электронов), хотя и достигли в последние годы колоссального прогресса, имеют ряд своих ограничений и сложностей и уступают еще по массовости применения рентгеноструктурному анализу. Основным препятствием на пути рентгеноструктурного исследования является необходимость приготовления образца исследуемого объекта в виде монокристалла, что является в ряде случаев весьма сложной задачей и служит основным фактором, лимитирующим применение метода. Развитие и автоматизация

---

\*lunin@impb.psn.ru

методов расшифровки структуры вывели этот метод на уровень, когда (с некоторой долей экстремизма) высказывается точка зрения, что получение кристалла – это единственная проблема при определении структуры. В последнее десятилетие наметились перспективы для решения этой проблемы. Появились результаты исследований, в которых объектом для рентгеноструктурного анализа выступали изолированные (некристаллизованные) копии исследуемого объекта. Хотя достигаемое в таких работах разрешение еще, как правило, далеко от того, что дают «классическая» биологическая кристаллография или электронная микроскопия, все же повышение мощности источников рентгеновского излучения и совершенствование детекторов позволяют рассчитывать на прогресс в этой области и приближение уровня получаемых результатов к тем, что достигаются в экспериментах с кристаллами. Данная статья посвящена обсуждению положения дел в этой области, возникающих проблем и путей их решения.

## РЕНТГЕНОВСКИЙ ДИФРАКЦИОННЫЙ ЭКСПЕРИМЕНТ

Проблема определения структуры сложных биологических объектов привлекает внимание исследователей, работающих в разных областях науки, и нередко аналогичные подходы разрабатываются независимо и выступают под разными названиями. Данная работа написана с позиций биологической кристаллографии и использует терминологию этого раздела науки со ссылками, по мере необходимости, на альтернативные названия. Предметом обсуждения является подход к определению структуры вещества, именуемый в различных источниках как рентгеноструктурный анализ, кристаллография белка, биологическая кристаллография, рентгеновская микроскопия, CDI (Coherent diffraction imaging) и т.д.

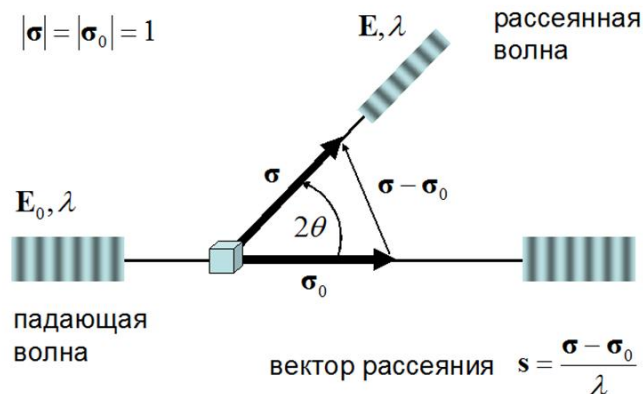


Рис. 1. Схема рентгеновского дифракционного эксперимента.

Во всех этих случаях основой является классический дифракционный эксперимент, схема которого приведена на рисунке 1. В этом эксперименте исследуемый объект помещается в пучок рентгеновских лучей и измеряется интенсивность волн, рассеянных во всевозможных направлениях. В процессе эксперимента варьируется направление  $\sigma_0$  первичного пучка по отношению к исследуемому объекту (например, путем вращения объекта в процессе эксперимента) и направление  $\sigma$  на детектор, регистрирующий интенсивность  $I(\sigma, \sigma_0)$  рассеянной (вторичной) волны. Результатом эксперимента является набор интенсивностей, измеренных для различных сочетаний направлений  $\sigma_0$  и  $\sigma$ . Мы будем говорить далее о картине рассеяния, подразумевая этот набор интенсивностей вторичных волн. Задачей, решаемой посредством этого эксперимента,

является восстановление распределения  $\rho(\mathbf{r})$  электронов в исследуемом объекте. Отметим сразу, что, говоря далее об исследовании изолированной частицы, мы подразумеваем, что в рентгеновском пучке находится только один экземпляр исследуемого объекта.

Связь распределения электронной плотности с результатами дифракционного эксперимента описывается (с приемлемой для практического определения структуры точностью) кинематической теорией рассеяния. В рамках этой теории первичный пучок считается плоской монохроматической волной (с длиной волны  $\lambda$  и амплитудой  $\mathbf{E}_0$ ). Под воздействием первичной волны электроны объекта начинают осциллировать и становятся источниками сферических волн. В точке детектора эти волны суммируются, интенсивность суммарной вторичной волны (часто называемой рефлексом), пропорциональная квадрату амплитуды суммарной вторичной волны, измеряется. Кинематическая теория рассеяния позволяет получить следующие выводы о картине рассеяния. Амплитуда суммарной волны зависит от комбинации направления первичного пучка и направления на детектор:

$$\mathbf{s} = \frac{\boldsymbol{\sigma} - \boldsymbol{\sigma}_0}{\lambda} \quad (1)$$

Эта комбинация играет важную роль в теории рассеяния и называется вектором рассеяния. Величина  $d = 1/|\mathbf{s}|$ , обратная длине вектора рассеяния, называется индивидуальным разрешением, соответствующим данному рефлексу. Для фиксированного вектора рассеяния амплитуда вторичной волны пропорциональна амплитуде первичной волны

$$\mathbf{E}(\mathbf{s}) = \varepsilon \mathbf{F}(\mathbf{s}) \mathbf{E}_0 \quad (2)$$

с двумя важными для дальнейшего изложения множителями. Комплексный множитель  $\mathbf{F}(\mathbf{s}) = F(\mathbf{s}) \exp[i\varphi(\mathbf{s})]$ , называемый структурным фактором, является не чем иным, как значением интегрального преобразования Фурье искомого распределения электронной плотности, вычисленного в точке  $\mathbf{s}$ :

$$\mathbf{F}(\mathbf{s}) = \int_{R^3} \rho(\mathbf{r}) \exp[2\pi i(\mathbf{s}, \mathbf{r})] dV_r \quad (3)$$

Распределение  $\rho(\mathbf{r})$  могло бы быть восстановлено посредством обратного преобразования Фурье

$$\rho(\mathbf{r}) = \int_{R^3} \mathbf{F}(\mathbf{s}) \exp[-2\pi i(\mathbf{s}, \mathbf{r})] dV_s \quad (4)$$

если бы комплексные значения структурных факторов были известны полностью. Однако, непосредственно, эксперимент позволяет измерить лишь интенсивности рефлексов  $I(\mathbf{s}) = F^2(\mathbf{s})$ , т. е. определить лишь значения модулей структурных факторов. Проблема восстановления необходимых для расчета значений фаз составляет центральную (фазовую) проблему рентгеноструктурного анализа. Еще одна проблема на пути восстановления распределения электронной плотности вызвана тем, что, формально, в расчет по формуле (4) должно быть включено бесконечное количество структурных факторов, в то время как набор экспериментально измеренных модулей всегда ограничен. Принято характеризовать объем экспериментально полученного набора значений модулей разрешением  $d_{\min}$ , подразумевая, что измерены интенсивности (почти) всех рефлексов с

индивидуальным разрешением  $d \geq d_{\min}$ . Из формулы (1) ясен геометрический смысл индивидуального разрешения рефлекса: это расстояние между ближайшими максимумами соответствующей ему в интеграле (4) гармоники Фурье  $\exp[-2\pi i(\mathbf{s}, \mathbf{r})]$ .

Множитель  $\varepsilon$  является комбинацией физических констант и для реальных условий эксперимента с биологическими макромолекулами может быть оценен величиной  $10^{-12}$ . Малость этого множителя создает основную сложность в практическом исследовании структур методом рентгеновской дифракции. Только ничтожная часть ( $10^{-24}$ ) энергии первичной волны переходит в энергию волны, рассеянной электроном. Это делает задачу практической регистрации рассеянных исследуемым объектом волн крайне сложной. Вся история развития метода рентгеноструктурного анализа (начиная с начала XX века) является историей борьбы за создание возможности экспериментальной регистрации рассеяния. Очевидные пути решения: повышение мощности источника излучения, повышение чувствительности детектора и увеличение времени экспозиции. Следует сразу заметить, что последний путь малоприменим в задачах биологической кристаллографии. Биологические объекты разрушаются под воздействием рентгеновского излучения, и предел экспозиции объекта сильно ограничен. В определенной мере это ограничение может быть ослаблено в рамках появившейся в последнее десятилетие техники эксперимента поточной кристаллографии (serial crystallography), при которой радиационная нагрузка на исследуемый объект распределяется среди большого количества идентичных копий объекта. При этом подходе картина рассеяния регистрируется отдельно для каждого из направлений  $\sigma_0$  первичного пучка относительно исследуемого объекта с использованием новой копии объекта для каждой ориентации объекта. Что же касается первых двух путей, то, несмотря на колоссальный прогресс в повышении мощности источников и чувствительности детекторов, регистрация рассеяния отдельной биологической частицей до последнего времени была не достижимой. Ключевым пунктом в практическом применении рентгеноструктурного анализа являлся дополнительный прием – многократное усиление интенсивности рассеяния в отдельных направлениях, достигаемое за счет изготовления и использования в эксперименте монокристалла исследуемого объекта.

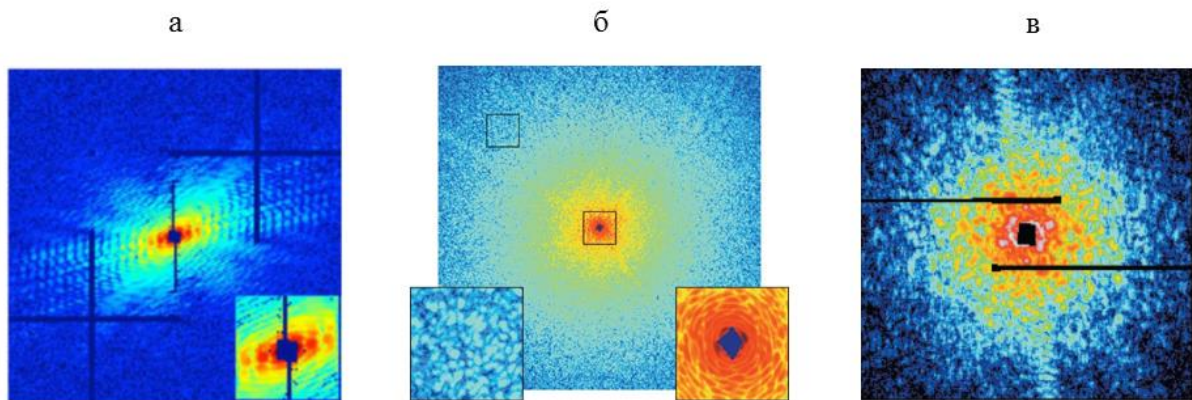
Остановимся еще на влиянии структурного фактора на интенсивность рефлекса. В предположениях о равномерном распределении позиций атомов в исследуемом объекте и эквивалентности рассеивающей способности атомов (что близко к реальности для биологических структур, где, в основном, присутствуют атомы С, N, O) среднюю величину модуля структурного фактора можно оценить величиной

$$\langle F(\mathbf{s}) \rangle = \sqrt{N} f(s) \quad . \quad (5)$$

Здесь  $N$  – число атомов в объекте, а  $f(s)$  – фактор атомного рассеяния. Факторы рассеяния для различных атомов табулированы и аппроксимируются, обычно, гауссовыми функциями. Быстрое убывание этих факторов при увеличении длины вектора рассеяния (или, что одно и то же, при увеличении угла рассеяния) создает дополнительные проблемы в регистрации интенсивностей рефлексов высоко разрешения и является фактором, в значительной мере определяющим разрешение собираемого набора экспериментальных данных. С другой стороны, формула (5) показывает, что с ростом размеров объекта задача регистрации рассеяния облегчается, поскольку интенсивность рефлексов возрастает, в среднем, пропорционально числу атомов в исследуемом объекте.

Важным этапом в развитии экспериментальной техники последнего десятилетия является достижение пределов, позволяющих в некоторых случаях зарегистрировать

рассеяние отдельными экземплярами макромолекулярных комплексов. На рисунке 2 приведены примеры экспериментально полученных картин рассеяния отдельными биологическими частицами [1–3]. Во всех проведенных экспериментах результаты были получены для крупных макромолекулярных комплексов, содержащих большое число атомов  $N$ , что, в соответствии с формулой (5), облегчает регистрацию рассеяния. Второй характерной чертой полученных результатов является чрезвычайно низкое, по меркам биологической кристаллографии, разрешение собранных наборов данных, обусловленное причинами, о которых шла речь выше. Среди полученных результатов выделяется недавно опубликованная работа [4], авторам которой удалось зарегистрировать рассеянные фотоны в зоне разрешения до  $5.9 \text{ \AA}$ , что кардинально меняет оценку возможностей при работе с изолированными частицами. Следует также отметить, что результаты были получены для существенно меньшего по размерам объекта (около  $700 \text{ \AA}$  в диаметре), нежели объекты, отраженные на рисунке 2. Эти результаты были достигнуты с использованием рентгеновских лазеров и экспериментальной схемы проточной кристаллографии, на возможностях которых мы остановимся ниже.



**Рис. 2.** Экспериментальная картина рассеяния для биологических изолированных частиц: а) Одноклеточный паразит *N. caninum* (из [2]). Диаметр клетки около  $5 \text{ мкм}$ . Съемка на синхротронном ускорителе ESRF в Гренобле (Франция). Зарегистрирована дифракция до разрешения  $600 \text{ \AA}$ . б) Дрожжевая клетка (из [1]). Диаметр клетки около  $2.5 \text{ мкм}$ . Съемка на синхротронном ускорителе ALS в Беркли (США). Зарегистрирована дифракция до разрешения  $207 \text{ \AA}$ . в) Хлоропласт (из [3]). Диаметр частицы около  $2 \text{ мкм}$ . Съемка на рентгеновском лазере SACLA (Япония). Зарегистрирована дифракция до разрешения  $175 \text{ \AA}$ .

Определенные надежды на продвижение в исследовании изолированных частиц связаны с вводом в строй новых мощных источников рентгеновского излучения – лазеров на свободных электронах (XFEL). Основные отличия нового источника излучения от предшествующих заключается в том, что XFEL-излучение распространяется в виде чрезвычайно мощных и чрезвычайно коротких ( $10\text{--}100$  фемтосекунд) импульсов. Импульсы характеризуются энергией фотонов  $E_{\text{photon}}$  и мгновенной плотностью потока фотонов  $j$  (фотонов/ $\text{\AA}^2 \text{ сек}$ ), связанных с длиной и амплитудой первичной электромагнитной волны соотношениями:

$$E_{\text{photon}} = \frac{hc}{\lambda}, \quad j = \frac{\lambda}{8\pi h} E_0^2. \quad (6)$$

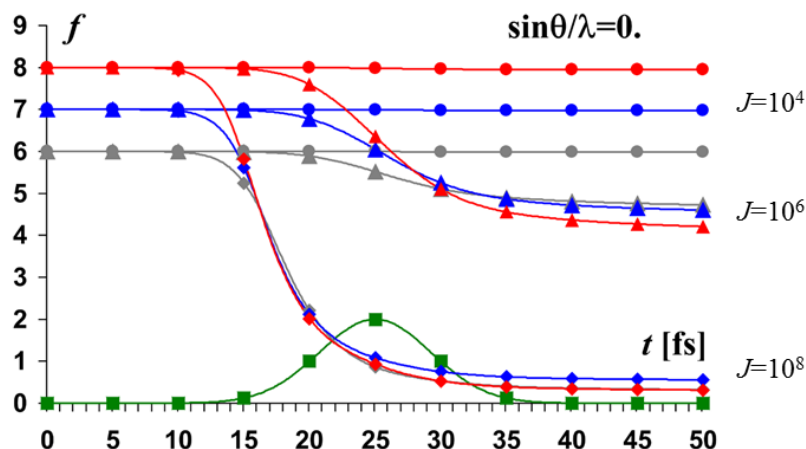
На практике, импульс часто характеризуется величиной плотности потока фотонов за импульс  $J$  (фотонов/ $\text{\AA}^2$ ), получающейся интегрированием мгновенной плотности

фотонов  $j$  по времени импульса. Предельную величину мощности имеющих в настоящее время рентгеновских лазеров можно охарактеризовать величиной  $J \approx 10^4$  фотонов/Å<sup>2</sup> за импульс. С вводом в строй новых XFEL планируется повышение этого предела на два, а в более отдаленной перспективе на четыре порядка.

Использование в эксперименте мощных рентгеновских импульсов XFEL делает более острой проблему радиационного повреждения объекта в процессе эксперимента. В основе процесса разрушения образца лежит явление фотоэффекта, при котором электроны внутренних оболочек могут спонтанно выбиваться налетающими фотонами первичного пучка. Процесс сопровождается переходами электронов с внешних оболочек на освободившиеся вакансии с испусканием дополнительного электрона (электрона Оже). Возникающая ионизация атомов приводит к частичному (или, в пределе, к полному) разрушению структуры. Такие процессы имеют место и при использовании других источников рентгеновского излучения, лимитируя допустимые дозы облучения объекта. Особенно важно учитывать эти процессы при повышении мощности источника. В настоящее время в биологической кристаллографии доминирующей является концепция “diffraction before destruction”. Согласно этой концепции исследуемый объект, даже когда разрушается в процессе эксперимента, до разрушения дает дифракционную картину, которая может быть зарегистрирована. В то же время можно отметить, что для используемых в настоящее время рентгеновских лазеров проблема разрушения биологического объекта не является острой. Модельные расчеты [5] показывают, что при плотности потока фотонов до  $10^4$  фотонов/Å<sup>2</sup> за импульс (предельная мощность существующих источников) у атомов С, N, O не происходит существенной потери электронов (рис. 3). Тем не менее, такие эффекты могут становиться существенными при возрастании плотности потока фотонов и требовать введения надлежащих поправок в расчетные схемы. Эта задача является актуальной, поскольку полученные оценки [6] показывают, что регистрация рассеяния для отдельных молекул белка требует существенного увеличения мощности существующих источников излучения.

Существенного снижения радиационной нагрузки на объекты и повышения точности регистрации данных позволяет достичь недавно предложенная схема поточной (serial) кристаллографии. При этой схеме эксперимента организуется поток экземпляров исследуемого объекта, например, микрокристаллов или отдельных частиц. Пролетая между источником излучения и детектором, объект может встретиться с рентгеновским импульсом и дать картину рассеяния (рентгенограмму), регистрируемую двумерным детектором. (На практике, частота таких встреч очень мала, что требует использования очень большого количества исследуемого вещества). Поскольку экземпляры объекта летят в произвольных ориентациях, набор полученных рентгенограмм позволяет получить трехмерную картину рассеяния. В этом подходе управляемое вращение кристалла в гониометре заменяется большим количеством случайно ориентированных кристаллов (или изолированных частиц). Преимуществами такого подхода является то, что с каждого экземпляра снимается только одна рентгенограмма, и весь радиационный ресурс объекта может быть использован для этой одной рентгенограммы, а не делиться на набор рентгенограмм. Далее, рентгенограммы, отвечающие идентичным ориентациям объектов, могут усредняться, повышая точность экспериментальных измерений. Недостатком подхода (помимо большого расхода экспериментального материала) является то, что ориентации пролетающих экземпляров заранее не известны, что создает проблему при сведении набора двумерных рентгенограмм в трехмерный набор. Дополнительной проблемой является проблема отсутствия идентичности объектов в потоке. Эта проблема решается (в определенных пределах) при исследовании в эксперименте микрокристаллов,

выращенных в одинаковых условиях. В этом случае различия сводятся к размерам кристаллов, что может быть компенсировано надлежащим шкалированием данных. Проблема идентичности объектов становится более острой при исследовании отдельных частиц, например, бактериальных клеток.



**Рис. 3.** Изменение во времени среднего числа рассеивающих электронов в процессе воздействия рентгеновского импульса при разной плотности потока фотонов  $J$  (фотонов на  $\text{\AA}^2$  за импульс) (из [5]). Профиль изменения локальной плотности потока фотонов в течение импульса показан зеленым цветом в относительных единицах. Красный, синий и серый цвета соответствуют атомам кислорода, азота и углерода соответственно.

### ФАЗОВАЯ ПРОБЛЕМА ПРИ РАССЕЙАНИИ ИЗОЛИРОВАННОЙ ЧАСТИЦЕЙ

Следующей проблемой, после регистрации картины рассеяния и сведения данных в трехмерный набор модулей структурных факторов, является фазовая проблема – определение значений фаз структурных факторов. При исследовании изолированной частицы фазовая проблема эквивалентна фазовой проблеме для кристаллической структуры, и весь арсенал методов биологической кристаллографии может быть применен для ее решения. Однако существует и отличие, делающее фазовую проблему для изолированной частицы более легкой.

При использовании в рентгеновском эксперименте монокристалла возникают выделенные направления рассеяния лучей, для которых волны, рассеянные разными копиями молекулы в кристалле, имеют одинаковую фазу, что приводит к многократному усилению суммарной волны и делает возможным экспериментальную регистрацию ее интенсивности. В прочих направлениях волны от разных копий молекулы ослабляют друг друга, что приводит к невозможности практической регистрации интенсивности. Условия, при которых происходит такое усиление (условия Лауэ – Брэгга – Вульфа), имеют вид:

$$(\mathbf{s}, \mathbf{a}) = h, \quad (\mathbf{s}, \mathbf{b}) = k, \quad (\mathbf{s}, \mathbf{c}) = l, \quad h, k, l - \text{целые числа}, \quad (7)$$

где  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  – базис элементарной ячейки кристалла. Таким образом, при использовании кристалла происходит принудительная дискретизация результатов эксперимента в пространстве векторов рассеяния (обратном кристаллографическом пространстве). Появляется сетка (решетка обратного пространства), определяемая параметрами кристаллической ячейки, в узлах которой амплитуда рассеянной волны есть не что иное, как  $M$ -кратная амплитуда волны, рассеянной отдельной копией молекулы, и только значения для этих узлов интенсивности рефлексов можно измерить в эксперименте. В

случае ортогональной ячейки кристалла эта сетка тоже ортогональная, с длинами сторон, обратными длинам сторон кристаллической ячейки. В общем случае, решетка построена на базисе  $\{\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*\}$ , являющемся сопряженным к базису  $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$  элементарной ячейки кристалла. Распределение электронной плотности в этом случае является периодической функцией и может быть представлено в виде ряда Фурье:

$$\rho(\mathbf{r}) = \frac{1}{|V|} \sum_{\mathbf{s} \in \mathcal{R}} F(\mathbf{s}) \exp[i\varphi(\mathbf{s})] \exp[-2\pi i(\mathbf{s}, \mathbf{r})] \quad (8)$$

Здесь суммирование идет по векторам решетки обратного пространства,  $|V|$  – объем элементарной ячейки кристалла. Фазовая проблема сводится к определению фаз  $\{\varphi(\mathbf{s})\}$  структурных факторов, участвующих в расчете ряда (8), в то время как модули  $\{F(\mathbf{s})\}$  этих структурных факторов определяются из эксперимента.

При изучении рассеяния изолированной частицей, формально, можно изучать рассеяние для произвольных векторов рассеяния. Однако на практике для работы отбираются значения интенсивностей, измеренные для узлов некоторой равномерной сетки  $\mathcal{R}'_{sp}$  в пространстве векторов рассеяния. Такая сетка определяется некоторым базисом дискретизации  $\{\mathbf{a}^*_{sp}, \mathbf{b}^*_{sp}, \mathbf{c}^*_{sp}\}$ , при этом величина шага дискретизации, хотя и может варьироваться, не может быть выбрана произвольно малой. Она ограничена снизу, например, физическими размерами пиксела детектора. Рассмотрим воображаемый кристалл с параметрами элементарной ячейки  $\{\mathbf{a}_{sp}, \mathbf{b}_{sp}, \mathbf{c}_{sp}\}$ , сопряженными базису дискретизации (обратными шагу дискретизации в случае ортогональной сетки дискретизации), в ячейке которого находится одна исследуемая частица. Если ячейка воображаемого кристалла достаточно велика, чтобы вместить в себя частицу целиком, то структурные факторы  $\mathbf{F}_{sp}(\mathbf{s})$ , соответствующие такому воображаемому кристаллу, совпадают со значениями интегрального преобразования Фурье (3), определяющего рассеяние изолированной частицей:

$$\mathbf{F}_{sp}(\mathbf{s}) = \int_{V_{sp}} \rho(\mathbf{r}) \exp[2\pi i(\mathbf{s}, \mathbf{r})] dV_r = \int_{R^3} \rho(\mathbf{r}) \exp[2\pi i(\mathbf{s}, \mathbf{r})] dV_r = \mathbf{F}(\mathbf{s}), \quad \mathbf{s} \in \mathcal{R}'_{sp} \quad (9)$$

Здесь  $V_{sp}$  – ячейка, построенная на векторах  $\{\mathbf{a}_{sp}, \mathbf{b}_{sp}, \mathbf{c}_{sp}\}$ ,  $\rho(\mathbf{r})$  – распределение электронной плотности в изолированной частице.

Таким образом, дискретный набор интенсивностей, измеренных в эксперименте с изолированной частицей, может трактоваться как набор значений квадратов модулей структурных факторов для воображаемого кристалла, содержащего один экземпляр частицы в каждой ячейке. Задача восстановления распределения электронной плотности в частице может быть сформулирована как задача восстановления электронной плотности в ячейке воображаемого кристалла в виде ряда Фурье с известными модулями структурных факторов

$$\rho(\mathbf{r}) = \frac{1}{|V_{sp}|} \sum_{\mathbf{s} \in \mathcal{R}'_{sp}} \mathbf{F}(\mathbf{s}) \exp[i\varphi(\mathbf{s})] \exp[-2\pi i(\mathbf{s}, \mathbf{r})]. \quad (10)$$

Эта задача совпадает со стандартной задачей биологической кристаллографии, и весь арсенал методов биологической кристаллографии может быть привлечен к ее решению.



Следует сразу отметить, что хотя формально задачи определения кристаллической структуры и структуры отдельной частицы эквивалентны (определить функцию по значениям модулей ее коэффициентов Фурье), количество привлекаемой для их решения экспериментальной информации может существенно различаться. При работе с кристаллическим образцом шаг дискретизации данных рассеяния жестко определен параметрами кристалла, и тем самым для некоторой выбранной зоны разрешения жестко определено количество экспериментальных измерений, отвечающих этому разрешению. При работе с изолированной частицей шаг дискретизации определяется исследователем, и при достаточно малом шаге количество экспериментальных измерений для того же объекта и в той же зоне разрешения может становиться существенно большим. На эту избыточность можно взглянуть и с другой стороны. При работе с воображаемым кристаллом с большим размером ячейки значения электронной плотности заранее известны для большинства точек ячейки – они равны нулю для всех точек вне области, занятой частицей. Т.е. формальная задача определения значений плотности во всей ячейке сводится к определению этих значений лишь в небольшой ее части. Аналогичная избыточность данных может возникать и используется при работе с кристаллами, когда значительная часть кристаллической ячейки занята растворителем [7, 8]. Однако свобода выбора шага дискретизации данных при работе с изолированной частицей позволяет сводить задачу к исследованию воображаемых кристаллов с аномально большим объемом «растворителя».

В биологической кристаллографии проблема недостаточности экспериментальных данных для однозначного восстановления распределения электронной плотности преодолевается, в основном, либо путем получения дополнительных данных в экспериментах с модифицированными объектами (изоморфное замещение), либо с модифицированными условиями эксперимента (аномальное рассеяние), либо из известных гомологичных структур (молекулярное замещение). Существует также ряд прямых или *ab initio* подходов, основанных на использовании дополнительной информации общего вида: атомность, гистограммы электронной плотности, связность и т.п. [9, 10]. Однако при исследовании макромолекулярных структур роль этих методов сильно ограничена, и они используются, в основном, на стадиях уточнения уже каким-либо образом полученных значений фаз. Существенная избыточность экспериментальных данных, достигаемая при исследовании изолированных частиц, кардинально меняет ситуацию и делает прямые методы основным подходом в решении фазовой проблемы.

## МЕТОД МОДИФИКАЦИИ ЭЛЕКТРОННОЙ ПЛОТНОСТИ

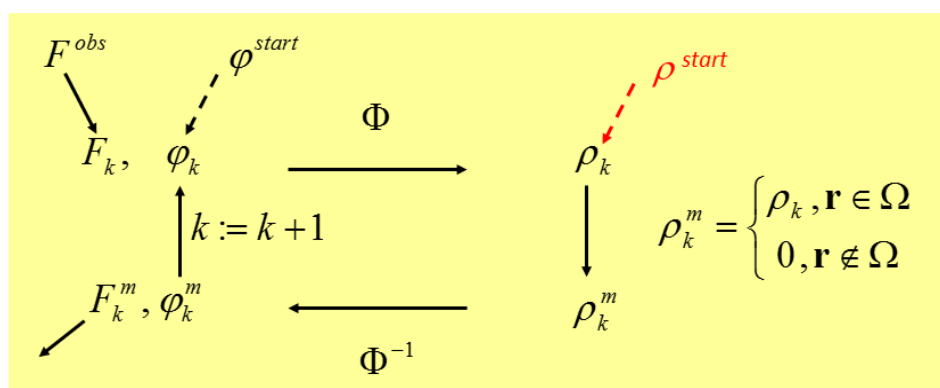
Наличие в ячейке кристалла значительного объема, занятого растворителем (имеющим более низкую, нежели в молекуле белка, и почти постоянную плотность), вносит в процесс расшифровки структуры дополнительную информацию, помимо экспериментально определяемых модулей структурных факторов. Эта информация может быть использована в процессе расшифровки структуры в рамках предложенного в 70-е годы итерационного метода модификации электронной плотности [7, 8]. Схема метода приведена на рисунке 4.

Процедура предполагает, что нам известна область  $\Omega$ , занятая объектом (область молекулы), или, что эквивалентно, область, занятая растворителем. В кристаллографии процедура стартует с некоторого набора приближенных значений фаз структурных факторов, которые в совокупности с экспериментальными значениями модулей позволяют рассчитать стартовой синтез Фурье электронной плотности. Полученная функция  $\rho_k$  не

отвечает требованию равенства нулю в области растворителя, поэтому к ней применяется операция  $P_s$  принудительного обнуления значений в области растворителя

$$\rho_k^m = P_s \rho_k = \begin{cases} \rho_k, & \mathbf{r} \in \Omega \\ 0, & \mathbf{r} \notin \Omega \end{cases} . \quad (11)$$

Полученная функция  $\rho_k^m$  равна нулю в области растворителя, но модули ее структурных факторов не совпадают с экспериментальными значениями. Поэтому к ней применяется операция  $P_m$  подмены модулей. По функции  $\rho_k^m$  рассчитываются структурные факторы, величины их модулей заменяются экспериментально определенными значениями, и рассчитывается новый синтез Фурье  $\rho_{k+1}$  с экспериментальными значениями модулей и модифицированными значениями фаз. Этот шаг – последовательное применение операторов  $P_s$  и  $P_m$  – повторяется итерационно до достижения сходимости. Такой подход широко применяется на практике в биологической кристаллографии для уточнения значений фаз структурных факторов [11].



**Рис. 4.** Схема метода модификации электронной плотности.  $\Phi$  – операция суммирования ряда Фурье.  $\Phi^{-1}$  – расчет структурных факторов (коэффициентов Фурье) для заданного распределения плотности.

Аналогичный подход, называемый Error Reduction Algorithm, был независимо предложен примерно в те же годы [12] для некоторых разделов оптики, где возникает та же задача – восстановление функции по значениям модулей ее коэффициентов Фурье. Единственно отличие заключается в том, что стартовой точкой является непосредственно некоторое распределение  $\rho^{\text{start}}$ , а не набор значений фаз. Далее на протяжении нескольких десятилетий эти два подхода развивались практически независимо друг от друга.

Предложенная итерационная процедура подкупает своей простотой, но обладает всеми недостатками, потенциально присущими итерационным процедурам: она может не сходиться или сходиться чрезвычайно медленно, она может сходиться к постороннему решению и т.п. Кроме того, обнаружился ряд дополнительных сложностей, связанных с тем, что задача восстановления функции по модулю преобразования Фурье априори имеет неоднозначность, связанную с выбором энантиомера (функции  $\rho(\mathbf{r})$  и  $\rho(-\mathbf{r})$  имеют одинаковый модуль преобразования Фурье). Возможно также наличие более сложных неоднозначностей, а именно гометрических решений. В связи с этим были предложены многочисленные модификации первоначальной процедуры. Два существенных элемента таких модификаций, используемых в кристаллографических приложениях –

комбинированные синтезы Фурье и «опрокидывание плотности» (density flip). Первый подход связан с тем, что при наличии значений модулей и фаз структурных факторов  $\{F^c(\mathbf{s}), \varphi^c(\mathbf{s})\}$ , рассчитанных по некоторой модели, и экспериментально определенных значений  $\{F^o(\mathbf{s})\}$ , синтез Фурье, рассчитанный с коэффициентами  $\{2F^o(\mathbf{s}) - F^c(\mathbf{s}), \varphi^c(\mathbf{s})\}$ , является лучшим приближением к искомому распределению электронной плотности, нежели простой синтез, рассчитанный с коэффициентами  $\{F^o(\mathbf{s}), \varphi^c(\mathbf{s})\}$ . Мы обозначим  $R_m$  операцию построения такого синтеза. Второй подход [13, 14] связан с тем, что при желании получить в рамках обсуждаемой итерационной процедуры функцию, равную нулю в некоторой области, ускорение сходимости может достигаться при замене оператора обнуления плотности  $P_s$  оператором  $R_s$ :

$$\rho_k^m = R_s \rho_k = \begin{cases} \rho_k, & \mathbf{r} \in \Omega \\ -\rho_k, & \mathbf{r} \notin \Omega \end{cases} \quad (12)$$

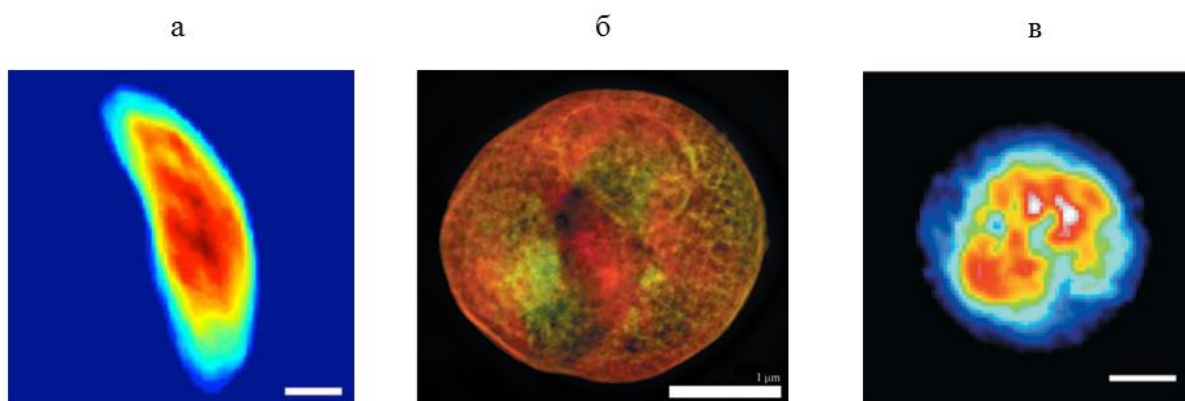
Операторы  $P_s$  и  $P_m$  могут, математически, рассматриваться как проекторы на многообразие функций, отвечающих поставленным условиям: равенства нулю вне области  $\Omega$  и наличия предписанных значений модулей структурных факторов  $\{F^o(\mathbf{s})\}$  соответственно. Операторы  $R_s$  и  $R_m$  являются связанными с ними рефлекторами, т.е. операторами, определяемыми для проектора  $P$  равенством  $R = 2P - I$  ( $I$  – тождественное преобразование). Действие рефлектора можно представлять себе как "отражение" точки в соответствующем многообразии. Было предложено множество разнообразных итерационных схем на базе этих проекторов и рефлекторов, некоторые из которых содержат дополнительные числовые параметры, определяющие комбинацию этих операторов [15, 16]. На практике выбор подходящей процедуры осуществляется эмпирически. Пробуются разные схемы и останавливаются на той, которая дает в результате распределение электронной плотности, визуальное обладающее ожидаемыми свойствами.

Существенного повышения эффективности итерационной процедуры удалось достичь с введением в нее процедуры автоматического выделения области молекулы. Итерационная процедура модификации электронной плотности предполагает, что нам известна область, в которой располагается исследуемая частица. Точность задания этой области может существенно влиять на ход итерационного процесса. Ключевой для развития подхода явилась дополнительная процедура автоматического выделения области молекулы [17–20]. Такая процедура состоит из двух шагов [19]. На первом шаге, исходя из текущей итерации распределения электронной плотности, выделяются опорные точки, с наибольшей вероятностью принадлежащие области молекулы. Это могут быть точки с максимальными значениями плотности, точки с минимальными значениями (которые также характерны для области молекулы [19]), точки, имеющие максимальную вариацию плотности в их окрестности, и т.п. Практически это реализуется назначением каждой точке веса  $p(\mathbf{r})$  (в простейшем случае 0 или 1), отражающего вероятность принадлежности точки области молекулы. На втором шаге система опорных точек «обволакивается» гладкой оболочкой. Это осуществляется путем свертки функции  $p(\mathbf{r})$  со сглаживающим ядром (например, гауссовым) и выбором одной из поверхностей уровня результирующей функции в соответствии с предполагаемым объемом области молекулы.

## ЭКСПЕРИМЕНТАЛЬНОЕ ОПРЕДЕЛЕНИЕ СТРУКТУРЫ ИЗОЛИРОВАННЫХ ЧАСТИЦ

### 1. Двумерные проекции

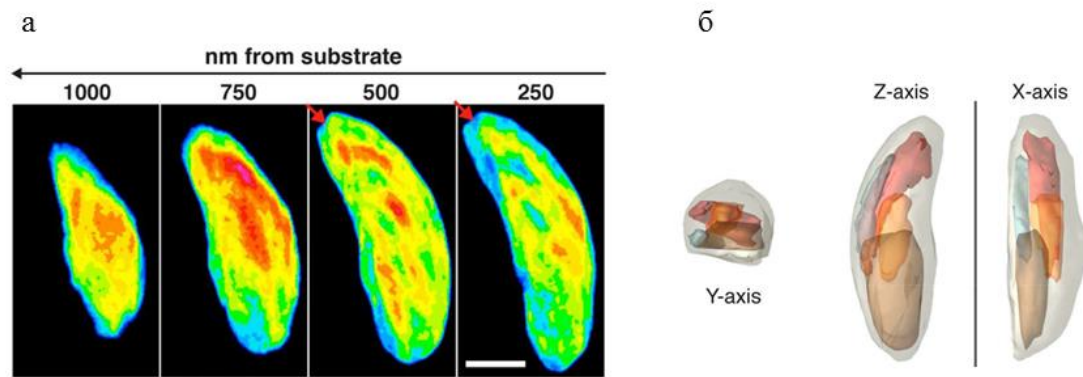
На рисунке 5 показаны результаты применения итерационных алгоритмов восстановления электронной плотности к данным, показанным на рисунке 2. На вход процедуры подавались двумерные рентгенограммы, результатом восстановления являются двумерные проекции трехмерного распределения плотности в направлении, перпендикулярном плоскости рентгенограммы. Следует отметить, что во всех случаях разрешение достигнутого результата уступает потенциально достижимому разрешению. Это связано в значительной мере с несовершенством процедуры восстановления значений фаз, и возможные пути преодоления этого разрыва – совершенствование методов решения фазовой проблемы.



**Рис. 5.** Реконструкция проекций электронной плотности по дифракционным картинкам, показанным на рисунке 2: а) Одноклеточный паразит *N. caninum* (из [2]). Оценка достигнутого разрешения 740–990 Å, в зависимости от направления. б) Дрожжевая клетка (из [1]). Оценка достигнутого разрешения 300 Å; в) Хлоропласт (из [3]). Оценка достигнутого разрешения 192 Å.

### 2. Трехмерная реконструкция

Наиболее интересной перспективой для задач структурной биологии является возможность восстановления трехмерного распределения электронной плотности. Мы остановимся здесь на двух примерах трехмерной реконструкции. В первом случае [2] исследуемым объектом являлся одноклеточный паразит *N. caninum*. Его клетки имеют продолговатую форму с размерами от 1–3 до 4–6 мкм в разных направлениях. Техника эксперимента была близка к стандартной кристаллографической практике. Набор дифракционных данных собирался на синхротронном ускорителе ESRF в Гренобле. Клетка в капле жидкости высаживалась на фрагмент синтетической мембраны с линейными размерами порядка 5 мкм. Образец замораживался жидким азотом и помещался в гониометрическую головку. В процессе съемки осуществлялось вращение образца в углах от –60 до +60 градусов. Это позволило снять набор из 72 рентгенограмм с экспозицией 100 секунд для каждой из них. Максимальное разрешение собранных данных было оценено величиной 600 Å. Восстановление электронной плотности по этим данным осуществлялось при помощи модифицированного итерационного алгоритма [21]. Оценка разрешения восстановленного изображения составила от 740 до 990 Å, в зависимости от направления. Восстановленное распределение плотности показано на рисунке 6.



**Рис. 6.** Трехмерная реконструкция распределения электронной плотности в одноклеточном паразите *N. caninum* (из [2]). Оценка достигнутого разрешения 740–990 Å, в зависимости от направления. а) Двумерные сечения. б) Трехмерная модель.

Альтернативная техника была использована в работе по трехмерной реконструкции гигантского мимивируса [22]. Диаметр капсиды этого вируса составляет около 0.45 мкм. Экспериментальные данные были собраны на рентгеновском лазере Linac Coherent Light Source. Эксперимент был организован по схеме поточной кристаллографии. Поток частиц подавался в пучок в виде гелиевого аэрозоля. Продолжительность импульса составляла 70 фс. Из 2000 рентгенограмм, содержащих сигнал, для работы было отобрано 198. Предельное разрешение отобранных данных составляло около 830 Å. Для сведения в трехмерный набор двумерных рентгенограмм, отвечающих неизвестным ориентациям частицы, был применен специально разработанный алгоритм [23]. Для решения фазовой проблемы использовались итерационные алгоритмы, реализованные в комплексе программ Hawk [24]. Разрешение полученного изображения было оценено как 1250 Å.

Следует отметить, что, как и при работе с двумерными проекциями, разрешение, достигнутое для восстановленного распределения электронной плотности, существенно уступает потенциально доступному разрешению, определяемому разрешением набора экспериментальных данных.

## СКАНИРОВАНИЕ КОНФИГУРАЦИОННОГО ПРОСТРАНСТВА РЕШЕНИЙ

Итерационные процедуры типа метода модификации электронной плотности, при их внешней простоте, имеют тот недостаток, что фактически они работают в области некоторого локального минимума, определяемого стартовой точкой. Этот недостаток может быть преодолен в рамках Монте-Карловской процедуры сканирования конфигурационного пространства распределений электронной плотности. Помимо этого, упомянутые ранее итерационные методы используют только один тип дополнительной информации об объекте – наличие области с постоянным значением плотности (области растворителя). Использование дополнительных типов ограничений может повысить стабильность процедуры и расширить область ее применения. В качестве такого типа дополнительной информации нами были использованы ограничения, связанные с топологическими характеристиками области высоких значений электронной плотности, например, с линейной связностью таких областей.

В целом, предложенная процедура решения фазовой проблемы имеет структуру, показанную на рисунке 7 [25–27]. Процедура может повторяться итерационно, с расширяющимся от шага к шагу набором экспериментальных данных, при этом каждый шаг состоит из двух этапов. На первом этапе генерируется значительное количество

случайных связанных масок (наборов узлов сетки, на которой рассчитывается электронная плотность) области частицы. Узлы сетки выбираются с равной вероятностью на первом шаге процедуры и с предписанной вероятностью – на последующих шагах. Количество узлов в маске определяется оценкой размеров объекта и шагом сетки. Для каждой сгенерированной маски вычисляются соответствующие структурные факторы, модули которых сравниваются с экспериментальными значениями. Если соответствие достаточно хорошее, маска считается «допустимой», и набор соответствующих ей фаз отбирается для последующего анализа. Процедура генерации повторяется до достижения заданного количества отобранных допустимых наборов фаз.

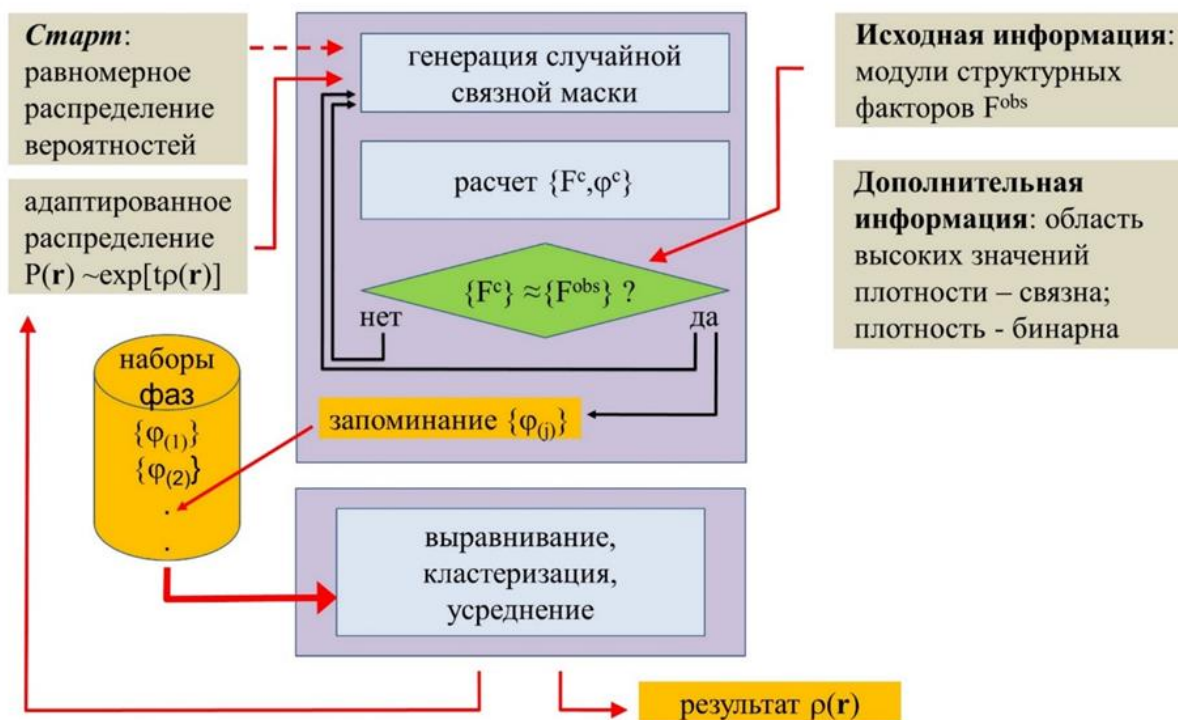
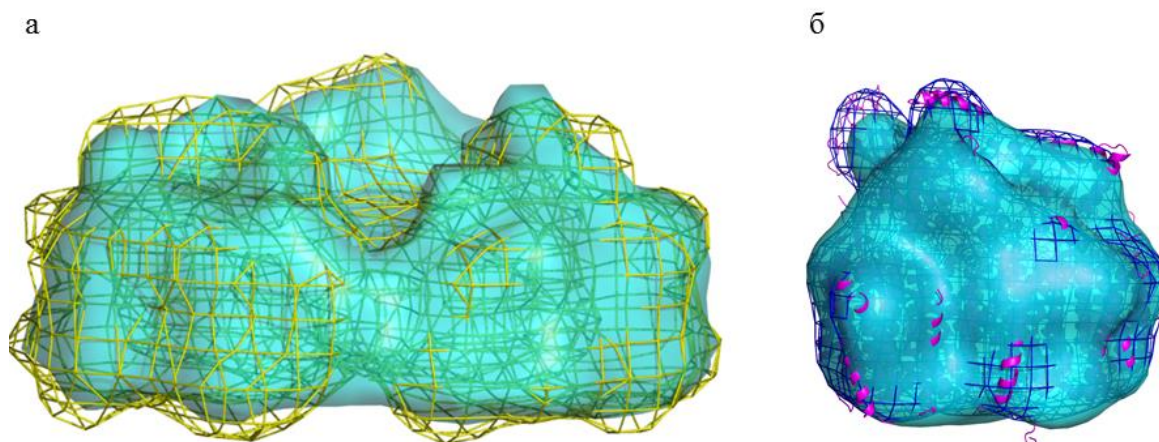


Рис. 7. Схема сканирования конфигурационного пространства решений.

На втором этапе строится кластерное дерево, выбирается компактный кластер вариантов и проводится усреднение вариантов внутри кластера. Полученный усреднённый вариант может рассматриваться либо как окончательное решение, либо использоваться для построения априорного распределения вероятностей, используемого при генерации масок на следующем шаге процедуры.

Следует отметить, что как при вычислении расстояний между фазовыми наборами на этапе кластеризации, так и перед усреднением проводится «выравнивание» фазовых наборов. Необходимость такого выравнивания связана с принципиальной неоднозначностью решения фазовой проблемы биологической кристаллографии. Среди возможных неоднозначностей можно выделить два типа «тривиальных» неоднозначностей, наличествующих априори. Первый тип тривиальной неоднозначности связан с тем, что сдвиг начала координат на произвольный вектор меняет значения фаз, но не меняет значений модулей структурных факторов. Поэтому два формально разных набора фаз могут давать (в совокупности с экспериментальными значениями модулей) одинаковые изображения объекта, различающиеся лишь сдвигом в пространстве. Поэтому, прежде чем сравнивать формально два изображения, эти изображения должны

быть оптимальным образом совмещены посредством произвольных сдвигов в пространстве.



**Рис. 8.** Тестовая реконструкция трехмерных распределений электронной плотности при разрешении 25 Å (из [25–27]). а) Тример фотосистемы I [28]. б) Мономер фотосистемы II [29]. «Сеткой» показаны поверхности, построенные на основе точного синтеза Фурье.

Второй тип тривиальной неоднозначности связан с тем, что два энантиомера  $\rho(\mathbf{r})$  и  $\rho(-\mathbf{r})$  имеют разные наборы фаз, но одинаковые наборы модулей структурных факторов. В отличие от случая со сдвигом изображения, эти функции не являются эквивалентными, поскольку они дают разную хиральность элементов структуры (аминокислотных остатков, спиралей и т. п.). Однако эта неоднозначность не может быть снята при работе на низком разрешении. Для ее снятия необходимы либо изображения достаточно высокого разрешения, когда, например, мы можем определить визуально хиральность аминокислот и выбрать в качестве правильного решения то, которое приводит к *L*-аминокислотам, либо использование дополнительных экспериментальных методов (например, связанных с аномальным рассеянием) для снятия неоднозначности. Поэтому при работе на низком разрешении принудительно фиксируется некоторый выбор энантиомера за счет того, что при выравнивании смена энантиомера рассматривается как допустимая операция. Однако вопрос о том, получено ли нами правильное решение или его энантиомер, остается открытым до более поздних этапов исследования.

Нетривиальная неоднозначность связана с тем, что две разных функции (в том числе, являющиеся суммой конечного числа атомов) могут не сводиться друг к другу сдвигом начала координат и сменой энантиомера. Однако при этом такие функции могут иметь одинаковые наборы модулей структурных факторов, обращаться в нуль вне заданной области пространства и иметь надлежащие характеристики связности. Такие решения называются в кристаллографии гомометрическими. В отличие от тривиальной неоднозначности, наличие гомометрических решений не является обязательным, и на данный момент не существует методов, позволяющих предсказать их наличие заранее. Возможным признаком присутствия гомометрических решений может являться появление нескольких разделенных компактных кластеров на кластерном дереве. В таком случае, работа должна вестись параллельно для нескольких потенциальных решений, пока достигнутое разрешение не позволит сделать выбор на основании дополнительных стереохимических соображений.

На рисунке 8 приведены примеры восстановленных изображений изолированных частиц, полученные в тестах со смоделированными данными рассеяния [25–27].

## ЗАКЛЮЧЕНИЕ

В заключение следует сделать несколько замечаний. Достигнутый прогресс в развитии экспериментальной техники делает возможной экспериментальную регистрацию рассеяния отдельными крупными биологическими частицами (на сегодняшний день – отдельными клетками и крупными вирусами). Развитие рентгеновских лазеров и совершенствование техники эксперимента позволяет рассчитывать на снижение ограничений как на размер частиц, так и на достигаемое разрешение.

С математической точки зрения, фазовая проблема при работе с изолированными частицами эквивалентна фазовой проблеме для кристаллических образцов. Поэтому для восстановления изображения изолированной частицы может применяться весь набор методов, разработанных в биологической кристаллографии, не сводящихся к процедуре сглаживания плотности в области растворителя. В отличие от работы с кристаллическими образцами, при работе с отдельной частицей возможно достижение существенной избыточности данных, что существенно повышает возможности прямых (или *ab initio*) методов решения фазовой проблемы.

Полученные на сегодняшний день результаты отличаются чрезвычайно низким (по меркам биологической кристаллографии или электронной крио-микроскопии) разрешением. Полученные изображения не приносят пока новой существенной структурной информации. Однако темпы совершенствования технических возможностей позволяют рассчитывать на существенное расширение возможностей метода в не слишком далеком будущем.

Работа поддержана грантом РФФИ 16-04-01037.

## СПИСОК ЛИТЕРАТУРЫ

1. Thibault P., Elser V., Jacobsen C., Shapiro D., Sayre D. Reconstruction of a yeast cell from X-ray diffraction data. *Acta Crystallographica Section A: Foundations of Crystallography*. 2006. V. 62. P. 248–261.
2. Rodriguez J.A., Xu R., Chen C.C., Huang Z., Jiang H., Chen A.L., Raines K.S., Pryor A. Jr, Nam D., Wiegart L., Song C., Madsen A., Chushkin Y., Zontone F., Bradley P.J., Miao J. Three-dimensional coherent X-ray diffractive imaging of whole frozen-hydrated cells. *IUCrJ*. 2015. V. 2. P. 575–583.
3. Takayama Y., Yonekura K. Cryogenic coherent X-ray diffraction imaging of biological samples at SACLA: a correlative approach with cryo-electron and light microscopy. *Acta Crystallographica Section A: Foundations and Advances*. 2016. V. 72. P. 179–189.
4. Munke A., Andreasson J, Aquila A., Awel S., Ayyer K., Barty A., Bean R.J., Berntsen P., Bielecki J., Boutet S. *et al.* Coherent diffraction of single Rice Dwarf virus particles using hard X-rays at the Linac Coherent Light Source. *Sci. Data*. 2016. V. 3. Article No. 160064. doi: [10.1038/sdata.2016.64](https://doi.org/10.1038/sdata.2016.64)
5. Lunin V.Y., Grum-Grzhimailo A.N., Gryzlova E.V., Sinitsyn D.O., Petrova T.E., Lunina N.L., Balabaev N.K., Tereshkina K.B., Stepanov A.S., Krupyanskii Y.F. Efficient calculation of diffracted intensities in the case of non-stationary scattering by biological macromolecules under XFEL pulse. *Acta Crystallographica, Section D: Biological Crystallography*. 2015. V. 71. P. 293–303.



6. Синицын Д.О., Лунин В.Ю., Грум-Гржимайло А.Н., Грызлова Е.В., Балабаев Н.К., Лунина Н.Л., Петрова Т.Е., Терешкина К.Б., Абдулнасыров Э.Г., Степанов А.С., Крупянский Ю.Ф. Новые возможности рентгеновской нанокристаллографии биологических макромолекул с использованием рентгеновских лазеров на свободных электронах. *Химическая физика*. 2014. Т. 33. № 7. С. 21–28.
7. Bricogne G. Geometric sources of redundancy in intensity data and their use for phase determination. *Acta Crystallographica Section A: Foundations of Crystallography*. 1974. V. 30. P. 349–405.
8. Bricogne G. Methods and programs for direct-space exploitation of geometric redundancies. *Acta Crystallographica Section A: Foundations of Crystallography*. 1976. V. 32. P. 832–847.
9. Lunin V.Y., Urzhumtsev A.G., Podjarny A.. Ab initio phasing of low-resolution Fourier syntheses. In: *International Tables for Crystallography vol. F, Second Edition*. Ed. Arnold E., Himmel D.M., Rossmann M.G. Chichester: John Wiley and Sons, 2012. P. 437–442.
10. Lunin V.Y., Lunina N.L., Urzhumtsev A.G. Ab initio low resolution phasing. In: *Advancing methods for biomolecular crystallography*. Ed. Read R., Urzhumtsev A.G., Lunin V.Y. Springer, 2013. P. 181–192.
11. Zhang K.Y.J., Cowtan K.D., Main P. Phase improvement by iterative density modification. In: *International Tables for Crystallography vol. F, Second Edition*. Ed. Arnold E., Himmel D.M., Rossmann M.G. Chichester: John Wiley and Sons, 2012. P. 385–400.
12. Fienup J.R. Reconstruction of an object from the modulus of its Fourier transform. *Optics Letters*. 1978. V. 3. N. 1. P. 27–29.
13. Abrahams J.P. Bias reduction in phase refinement by modified interference functions: introducing the  $\gamma$  correction. *Acta Crystallographica Section D: Biological Crystallography*. 1997. V. 53. P. 371–376.
14. Oslányi G., Sütő A. Ab initio structure solution by charge flipping. *Acta Crystallographica Section A: Foundations of Crystallography*. 2004. V. 60. P. 134–141.
15. Marchesini S. A unified evaluation of iterative projection algorithms for phase retrieval. *Rev. Sci. Instrum.* 2007. V. 78. Article No. 011301. doi: [10.1063/1.2403783](https://doi.org/10.1063/1.2403783).
16. Millane R., Lo V.L. Iterative projection algorithms in protein crystallography. I. Theory. *Acta Crystallographica Section A: Foundations of Crystallography*. 2013. V. 69. P. 517–527.
17. Уржумцев А.Г. Использование локального усреднения при анализе изображений макромолекул на картах распределения электронной плотности: препринт НЦБИ Пуццино. 1985.
18. Wang B.C. Resolution of phase ambiguity in macromolecular crystallography. *Methods in Enzymology*. 1985. V. 115. P. 90–112.
19. Urzhumtsev A.G., Lunin V.Y., Luzyanina T.B. Bounding a Molecule in a Noisy Synthesis. *Acta Crystallographica Section A: Foundations of Crystallography*. 1989. V. 45. P. 34–39.
20. Marchesini S., He H., Chapman H.N., Hau-Riege S.P., Noy A., Howells M.R., Weierstall U., Spence J.H.C. X-ray image reconstruction from a diffraction pattern alone. *Phys. Rev. B*. 2003. V. 68. Article No. 140101(R). doi: [10.1103/PhysRevB.68.140101](https://doi.org/10.1103/PhysRevB.68.140101).
21. Rodriguez J.A., Xu R., Chen C.-C., Zou Y., Miao J. Oversampling smoothness: an effective algorithm for phase retrieval of noisy diffraction intensities. *J. Applied Crystallography*. 2013. V. 46. P. 312–318.
22. Ekeberg T., Svenda M., Abergel C., Maia F.R.N.C., Seltzer V., Claverie J.M., Hantke M., Jönsson O., Nettelblad C., van der Schot G. *et al.* Three-Dimensional Reconstruction of the

- Giant Mimivirus Particle with an X-Ray Free-Electron Laser. *Phys. Rev. Lett.* 2015. V. 114. Article No. 098102. doi: [10.1103/PhysRevLett.114.098102](https://doi.org/10.1103/PhysRevLett.114.098102).
23. Loh N.-T.D., Elser V. Reconstruction algorithm for single-particle diffraction imaging experiments. *Phys. Rev. E.* 2009. V. 80. Article No. 026705. doi: [10.1103/PhysRevE.80.026705](https://doi.org/10.1103/PhysRevE.80.026705).
  24. Maia F.R.N.C., Ekeberg T., Spoel D., Hajdu J. *Hawk*: the image reconstruction package for coherent X-ray diffractive imaging. *J. Applied Crystallography*. 2010. V. 43. P. 1535–1539.
  25. Лунин В.Ю., Лунина Н.Л., Петрова Т.Е. Использование связанных масок в задаче восстановления изображения изолированной частицы по данным рентгеновского рассеяния. *Математическая биология и биоинформатика*. 2014. Т. 9. № 2. С. 543–562. doi: [10.17537/2014.9.543](https://doi.org/10.17537/2014.9.543).
  26. Лунина Н.Л., Петрова Т.Е., Уржумцев А.Г., Лунин В.Ю. Использование связанных масок в задаче восстановления изображения изолированной частицы по данным рентгеновского рассеяния. II. Зависимость точности решения от шага дискретизации экспериментальных данных. *Математическая биология и биоинформатика*. 2015. Т. 10. № 2. С. 508–525. doi: [10.17537/2015.10.508](https://doi.org/10.17537/2015.10.508).
  27. Lunin V.Y., Lunina N.L., Petrova T.E., Baumstark M.W., Urzhumtsev A.G. Mask-based approach to phasing of single-particle diffraction data. *Acta Crystallographica Section D: Structural Biology*. 2016. V. 72. P. 147–157.
  28. Jordan P., Fromme P., Witt H.T., Klukas O., Saenger W., Krauß N. Three-dimensional structure of cyanobacterial photosystem I at 2.5Å resolution. *Nature*. 2001. V. 411. P. 909–917.
  29. Broser M., Gabdulkhakov A., Kern J., Guskov A., Müh F., Saenger W., Zouni A. Crystal structure of monomeric Photosystem II from *Thermosynechococcus elongatus* at 3.6 Å resolution. *J. Biol. Chem.* 2010. V. 285. P. 26255–26262.

Рукопись поступила в редакцию 19.12.2016.  
Дата опубликования 30.01.2017.