

# ВОССТАНОВЛЕНИЕ ФОРМЫ НАНОЧАСТИЦЫ ПО РЕШЕНИЮ ПРЯМОЙ И ОБРАТНОЙ ЗАДАЧ МАЛОУГЛОВОГО РАССЕЯНИЯ ДЛЯ ЕДИНИЧНОГО ПОТЕНЦИАЛА ОГРАНИЧЕННОГО В ОБЪЕМЕ ТОРА

C. B. Амарантов\*

Институт кристаллографии имени А. В. Шубникова Российской академии наук  
119333, Москва, Россия

Поступила в редакцию 1 декабря 2008 г.

На примере простейшей двусвязной поверхности вращения — тора — продемонстрирована возможность устойчивого восстановления трехмерного однородного тела, заданного внутри своего объема единичным потенциалом  $U(\mathbf{r})$  по пространственно-усредненной кривой малоуглового рассеяния (МУР)  $I(s)$ . Моделирование проводилось методом Монте-Карло с «температурным отжигом» без априорной информации о форме и размере наночастицы. Получено точное и приближенное выражение для формфактора тора. Графически показано совпадение точной и приближенной кривых МУР формфактора тора для экспериментально достижимой области рассеяния. Приведены примеры.

PACS: 02.70.-c, 61.05.cf

## 1. ВВЕДЕНИЕ

Как известно, если задана форма потенциала  $U(\mathbf{r})$ , то пространственно-усредненная интенсивность рассеяния  $I(s)$ , где  $s$  — модуль вектора рассеяния,  $s = 4\pi \sin\theta/\lambda$ , определена однозначно. Здесь  $\theta$  — половина угла рассеяния (см. рис. 1),  $\lambda$  — длина волн излучения. Более трудную задачу представляет обратная задача — восстановление формы тела, определяемого областью  $U(\mathbf{r}) \neq 0$  в прямом пространстве, по известной функции  $I(s)$ . Для практики идентификации наночастиц в разбавленных растворах по данным рассеяния важно, что вид потенциала  $U(\mathbf{r})$  можно однозначно связать с формой исследуемых наночастиц через так называемый формфактор. Под формфактором мы будем понимать пространственно-усредненную трехмерную фурье-трансформанту от  $U(\mathbf{r})$ . В аналитическом (не интегральном) виде явно выражены лишь предельные формфакторы: шара, бесконечно тонких диска и стержня [1]. В ряде задач, таких как исследование структуры белков в растворе, возникает необходимость использования формфактора тора. В дан-

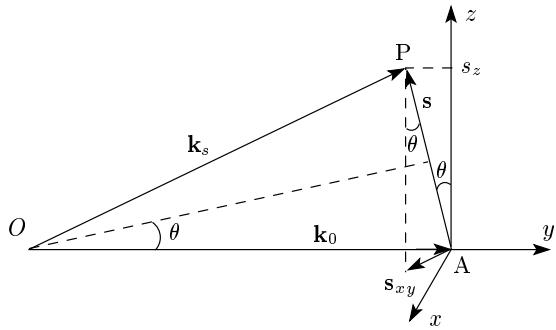
ной работе представлена прямая и обратная задачи восстановления трехмерной формы тора (как тела вращения) по кривой интенсивности малоуглового рассеяния в первом борновском приближении. Для полученного точного трехмерного пространственно-усредненного преобразования Фурье по объему, ограниченному поверхностью тора, в области малых углов ( $0 < \theta < 5^\circ$ ) методом Монте-Карло с «отжигом» было получено устойчивое восстановление формы тора по кривой малоуглового рассеяния (МУР)  $I(s)$ .

## 2. ИСПОЛЬЗУЕМЫЕ МЕТОДЫ

### 2.1. Однократное рассеяние на частице с однородной плотностью, дифракция Фраунгофера и первое борновское приближение

Рассмотрим изотропное рассеяние в точке О (см. рис. 1) падающего потока  $j_{inc}$  квазимонохроматических ( $\Delta E/E \approx 10^{-3}-10^{-4}$ ) рентгеновских квантов с энергией  $E$  на концентрированную систему типа «разбавленного раствора» невзаимодействующих  $N$  идентичных частиц. Рассеянный регистрируемый

\*E-mail: amarantov\_s@mail.ru, amarantov@ns.crys.ras.ru



**Рис.1.** Треугольник рассеяния, рассеяние упругое,  $|k_0| = |k_s|$ , где  $k_0$  — волновой вектор до рассеяния,  $k_s$  — волновой вектор после рассеяния,  $s$  — вектор рассеяния,  $\theta$  — половина угла рассеяния

поток в точке наблюдения  $P$ , расположенной на расстоянии  $L = |OP|$  от центра рассеяния, можно представить в виде

$$j_{scat} = \frac{k}{k_0} \frac{j_{inc}}{L^2} N A(s, \Omega) A^*(s, \Omega), \quad (2.1)$$

где  $k_0$ ,  $k_s$  — волновые векторы до и после рассеяния,  $A(s, \theta, \varphi)$  — амплитуда волны, рассеянной одной частицей, символ «\*» означает комплексное сопряжение,  $\Omega$  — телесный угол с образующей  $|OP|$  от оси падающего луча. Для дважды дифференциального (по углу  $\Omega$  и по энергии  $E$ ) когерентного сечения рассеяния имеем

$$\begin{aligned} \frac{d^2\sigma_{coh}}{j_{scat} dE} &= \frac{j_{scat} L^2}{j_{inc}} = \\ &= I(s) = \frac{k_s}{k_0} N A(s, \Omega) A^*(s, \Omega), \end{aligned} \quad (2.2)$$

где  $s$  — модуль вектора рассеяния, определяемый как модуль вектора разности падающего и рассеянного волновых векторов:

$$s \equiv k_0 - k_s, \quad |s| = 2k_0 \sin \theta = \frac{4\pi}{\lambda} \sin \theta. \quad (2.3)$$

В случае, если частицы рассеивают независимо друг от друга и  $|k_0| = |\mathbf{k}|$  (случай упругого рассеяния), для интенсивности упругого рассеяния  $I(s)$  можно записать

$$I(s) = I_e N A(s, \Omega) A^*(s, \Omega). \quad (2.4)$$

Здесь  $I_e$  представляет собой интенсивность «томсоновского» рассеяния или рассеяния на «свободном электроне»:

$$I_e(\theta) = r_e^2 \frac{I_0}{L^2} \mu(\theta), \quad (2.5)$$

$I_0$  — интенсивность падающего на образец излучения,  $r_e = e^2/m_e c^2 = 2.818 \cdot 10^{-13}$  см — классический радиус электрона. Для случая рассеяния рентгеновских лучей единица измерения интенсивности (материальный фактор) задается множителем  $r_e^2 = 7.94 \cdot 10^{-2}$  барн. Возможную некогерентную составляющую, комптоновский фон, мы не будем учитывать, поскольку доля энергии, сообщаемая рассеянному фотону в области энергий  $5 \text{ кэВ} < E < 50 \text{ кэВ}$ , для интересующего нас интервала углов рассеяния  $0 \leq 2\theta \leq 20^\circ$  является постоянной величиной. Здесь

$$\mu(\theta) = \frac{1 + \cos 2\theta}{2} \quad (2.6)$$

— поляризационный множитель для неполяризованного излучения.

В классической схеме малоуглового эксперимента характерное расстояние  $L$  много больше длины волны  $\lambda$  и размера рассеивающих объектов  $D_{max}$ . Полагая  $D_{max} \approx 100 \text{ нм}$ ,  $L \approx 1 \text{ м}$ , получим для числа зон Френеля в области регистрации рассеянного излучения:  $D_{max}^2/L\lambda \leq 10^{-4}$ . Отсюда следует, что рассеивающая неоднородность закрывает малую долю центральной зоны Френеля, поэтому мы имеем случай дифракции в параллельных лучах (Фраунгофера).

Следует заметить, что в данной работе не ставится задача единственности восстановления амплитуды рассеяния по сечению рассеяния, достаточно показать, что для системы типа «разбавленного раствора» правомерно использовать первое борновское приближение. Рассмотрим задачу рассеяния как задачу возмущения фронта плоской падающей волны, описываемой волновой функцией  $\Psi_0 = \exp(i\mathbf{k}_0 \cdot \mathbf{r})$  для стационарного уравнения Шредингера:

$$\left( -\frac{\hbar^2}{2\mu} \Delta + U(\mathbf{r}) \right) \Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \quad (2.7)$$

с ограниченной областью действия возмущающего потенциала  $U(\mathbf{r}) \in R^3$ . Здесь  $\mu \sim E/c^2$ , где  $c$  — скорость света. Если  $U(\mathbf{r}) \ll E \approx 50 \text{ кэВ}$  (энергия взаимодействия атомов частицы с падающим излучением много меньше энергии квантов, что типично для рентгеновских экспериментов), то, подставляя в уравнение (2.7) волновую функцию

$$\Psi(r) = \Psi_0(r) + \Psi_{scat}(r), \quad |\Psi_{scat}| \ll |\Psi_0| = 1,$$

получим для поправки  $\Psi_{scat}(r)$  поля рассеянной волны уравнение Гельмгольца:

$$\Delta\Psi_{scat}(\mathbf{r}) + k^2\Psi_{scat}(\mathbf{r}) = C_0 U(\mathbf{r}) \Psi_0(\mathbf{r}). \quad (2.8)$$

Здесь  $C_0 = E/(c\hbar)^2$ ,  $C_0 U_0/k^2 \ll 1$ ,  $C_0 U_0/k^2 \approx \approx 10^{-4} U_0 [\text{эВ}] \ll 1$  для длины волны  $\lambda(\text{СиK}_\alpha) = 1.54 \text{ \AA}$ . В этом случае общее решение для поправки  $\Psi_{scat}(\mathbf{r})$  можно записать в виде

$$\Psi_{scat}(\mathbf{r}) = -\frac{C_0}{4\pi} \frac{e^{ikL}}{L} \frac{n_e}{V'} \int_{V'} U(\mathbf{r}) e^{i\mathbf{s}\cdot\mathbf{r}} dV', \quad (2.9)$$

где  $n_e$  — эффективное число электронов, участвующих в упругом рассеянии рентгеновских квантов на одной частице. Решение (2.9) представляет собой «первое борновское приближение».

В дальнейшем нас будет интересовать не полный поток или  $\Psi(\mathbf{r})$ , а малая поправка  $\Psi_{scat}(\mathbf{r})$  или амплитуда рассеяния  $A(s, \Omega)$ :

$$\Psi_{scat} = -\frac{C_0}{4\pi} \frac{e^{ikL}}{L} A(s, \Omega). \quad (2.10)$$

Амплитуда рассеянной сферической волны связана с рассеивающим потенциалом преобразованием Фурье следующего вида:

$$\begin{aligned} A(s, \theta, \varphi) &= \frac{n_e}{V'} \int_{V'} U(\mathbf{r}) e^{i\mathbf{s}\cdot\mathbf{r}} dV' = \\ &= \int_{V'} \rho_e(\mathbf{r}) e^{i\mathbf{s}\cdot\mathbf{r}} dV', \end{aligned} \quad (2.11)$$

где  $\rho_e(\mathbf{r}) = (n_e/V')U(\mathbf{r})$  — электронная плотность рассеивающей частицы,  $V'$  — ее объем. Обозначим фурье-амплитуду от одной частицы ( $N = 1$ ) как

$$\Phi(s, \Omega) = \int_{V'} \rho_e(\mathbf{r}) e^{i\mathbf{s}\cdot\mathbf{r}} dV'. \quad (2.12)$$

Тогда для интенсивности малоуглового рассеяния в относительных единицах (материальный фактор не учитываем) получим так называемый формфактор частицы:

$$I(s) = \langle \Phi^2(s) \rangle = \int_{\Omega} \Phi^2(s, \Omega) d\Omega. \quad (2.13)$$

Область рассматриваемых углов рассеяния в малоугловой схеме ограничивается диапазоном  $0 \leq \theta \leq 5^\circ$ , где  $\theta$  — половина угла рассеяния. В этом случае поляризационный множитель  $\mu(\theta)$  будем считать равным единице, поскольку для данного углового диапазона он не приводит к сколько-нибудь заметному искажению кривой малоуглового рассеяния.

## 2.2. Расчет функции распределения по расстояниям, максимальному размеру частицы и ее радиуса инерции

Введем безразмерную электронную плотность  $\eta(\mathbf{r})$ :

$$\eta(\mathbf{r}) = \frac{V'}{n_e} \rho(\mathbf{r}). \quad (2.14)$$

В этом случае интенсивность рассеяния можно записать в безразмерном виде [1]:

$$\begin{aligned} I(s) &= \frac{n_e^2}{(V')^2} \times \\ &\times \int_{V'} \int_{V'} \langle \eta(\mathbf{r}_1) \eta(\mathbf{r}_2) \rangle e^{i\mathbf{s}\cdot\mathbf{r}} d\mathbf{r}_1 d\mathbf{r}_2. \end{aligned} \quad (2.15)$$

Угловые скобки означают усреднение по телесному углу  $\Omega$ . Преобразуем выражение для формфактора через автокорреляционную (характеристическую) функцию  $\langle \gamma_0(r) \rangle$  как самосвертку электронной плотности  $\eta_e(\mathbf{r})$ , полагая  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ ,  $d\mathbf{r}_1 = d\mathbf{r}_2 = dV'$ :

$$\begin{aligned} \langle \gamma_0(\mathbf{r}) \rangle &= \frac{1}{V'} \int_{V'} \langle \eta_e(\mathbf{r}_1) \eta_e(\mathbf{r}_1 - \mathbf{r}) \rangle d\mathbf{r}_1 \equiv \\ &\equiv \frac{1}{V'} \langle \eta_e(\mathbf{r}) \otimes \eta_e(\mathbf{r}) \rangle, \\ \langle \gamma_0(0) \rangle &= 1, \quad \langle \gamma_0(D_{max}) \rangle = 0. \end{aligned} \quad (2.16)$$

Геометрически  $\langle \gamma_0(r) \rangle$  трактуется как вероятность найти отрезок длины  $r$ , концы которого принадлежат объему частицы  $V'$  с однородной рассеивающей электронной плотностью  $\rho_e(\mathbf{r}) = \bar{\rho}_e = \text{const}$ , который также будет располагаться внутри частицы, т. е. не пересечет ее поверхности. Таким образом,  $\langle \gamma_0(r) \rangle$  является корреляционной функцией распределения по расстояниям внутри частицы. Для нее выполняется условие нормировки:

$$4\pi \int_0^{D_{max}} r^2 \gamma_0(r) dr = 1.$$

Пространственно-усредненная функция распределения по расстояниям  $p(r)$  вводится через автокорреляционную функцию как

$$p(r) \equiv \langle \gamma_0(\mathbf{r}) \rangle r^2. \quad (2.17)$$

Функция  $p(r)$  допускает следующую наглядную интерпретацию. Разобъем объем частицы  $V'$  на элементы объема  $v_i$ , тогда количество элементов  $v_i$  пропорционально  $V'$ . Выделим внутри частицы сферу

радиуса  $r$ , на поверхности которой расположим элементы  $v_i$ , тогда их количество будет пропорционально площади поверхности этой сферы  $4\pi r^2$ . Таким образом, произведение

$$4\pi r^2 V' \gamma_0(r) = 4\pi p(r)/\rho$$

дает количество отрезков длины  $r$ , соединяющих два произвольных элемента объема  $v_i$  и  $v_j$ .

Подставляя (2.16) в (2.15) и используя (2.17), выразим интенсивность рассеяния через  $p(r)$ :

$$\begin{aligned} I(s) &= \langle \Phi^2(s) \rangle = \frac{1}{V'} \int_{V'} \langle \gamma_0(\mathbf{r}) \rangle e^{i\mathbf{s} \cdot \mathbf{r}} d\mathbf{r} = \\ &= \frac{1}{V'} \int_{V'} \langle \gamma_0(\mathbf{r}) \rangle e^{i\mathbf{s} \cdot \mathbf{r}} r^2 d\mathbf{r} d\Omega = \\ &= \frac{1}{V'} \int_{V'} \langle p(\mathbf{r}) \rangle e^{i\mathbf{s} \cdot \mathbf{r}} d\mathbf{r} d\Omega. \end{aligned} \quad (2.18)$$

В частности, для разбавленной системы идентичных невзаимодействующих частиц, например, раствора белковых молекул, выражение для интенсивности рассеяния, приходящейся на одну молекулу, можно записать в виде [2]:

$$\begin{aligned} I(s) &= \langle \Phi(sD_{max}) \rangle = \frac{1}{V'} \int_0^{2\pi} \int_0^{D_{max}} \int_0^\pi p(r) e^{isr \cos \theta} \times \\ &\times \sin \theta d\theta dr d\varphi = \frac{4\pi}{V'} \int_0^{D_{max}} p(r) \frac{\sin(sr)}{sr} dr. \end{aligned} \quad (2.19)$$

Соотношение между интенсивностью рассеяния  $I(s)$  и функцией распределения по расстояниям  $p(r)$  определяет связь между обратным и прямым пространствами, выражаемую через преобразование Фурье для следующих трансформант:

$$sI(s) \xrightarrow{F} p(r), \quad p(r) \xrightarrow{F} \frac{2}{\pi} sI(s), \quad (2.20)$$

$$\begin{aligned} \frac{V'}{4\pi} sI(s) &= \int_0^\infty P(r) \frac{\sin(sr)}{r} dr \xleftarrow{F} \\ &\xrightarrow{F} P(r) = \frac{V'}{2\pi^2} \int_0^\infty I(s) \frac{\sin(sr)}{sr} s^2 ds. \end{aligned} \quad (2.21)$$

Таким образом, интенсивность рассеяния  $I(s)$  и функция  $p(r)$  для молекулы, имеющей максимальный размер  $D_{max}$ , связаны через интегральное уравнение Фредгольма первого рода:

$$\begin{aligned} I(s) &= \frac{4\pi}{V'} \int_0^{D_{max}} p(r) \frac{\sin(sr)}{sr} dr \equiv \\ &\equiv C \int_0^{D_{max}} p(r) K(s, r) dr, \end{aligned} \quad (2.22)$$

где  $C$  — постоянная,  $K(s, r)$  — ядро интегрального преобразования (2.22). Переход из обратного пространства в прямое и вычисление функции распределения по расстояниям внутри частицы,  $p(r)$ , в данной работе осуществлялись с помощью программы косвенного фурье-преобразования, в которой использован метод  $\alpha$ -регуляризации Тихонова [3]. Предполагалось, что искомое решение  $p(r)$  при  $r \in [0, D_{max}]$  должно иметь внутри интервала производную  $p'(r)$ , интегрируемую с квадратом, и  $p(0) = 0$ ,  $p'(0) = 0$ ,  $p(D_{max}) = 0$ ,  $p'(D_{max}) = 0$ . Эти равенства определяют для нашей задачи множество корректности, существенно сужающее класс возможных решений. В данной программной реализации [4] функционал Тихонова имел следующий вид:

$$\begin{aligned} \Phi_\alpha [p(r), I(s)] &= \frac{1}{N-1} \times \\ &\times \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[ I(s_i) - \int_0^{D_{max}} p(r) K(s_i, r) dr \right]^2 + \\ &+ \alpha \int_0^{D_{max}} \left[ \frac{dp(r)}{dr} \right]^2 dr, \end{aligned} \quad (2.23)$$

где  $\sigma_i$  — стандартное отклонение интенсивности рассеяния в данной экспериментальной точке  $s_i$ .

Преобразование Фурье осуществлялось на ограниченном интервале как обратного  $s \in [0, s_{max}]$ , так и прямого  $r \in [0, D_{max}]$  пространств, что допускается для функций с финитным спектром по теореме Котельникова–Шеннона [5].

В рамках однородного приближения для электронной плотности по кривой МУР можно найти радиус инерции частицы как геометрический инвариант аналогично механическому радиусу инерции [1, 2, 4]:

$$R_g^2 = \frac{\int_V \bar{\rho}_e r^2 dV}{\int_V \bar{\rho}_e dV} = \frac{\int_0^{D_{max}} p(r) r^2 dV}{2 \int_0^{D_{max}} p(r) dr}. \quad (2.24)$$

В отличие от традиционного подхода, когда делается оценка по тангенсу угла наклона начального участка кривой рассеяния в координатах Гинье:  $\ln I \propto s^2$ , вычисление по интегральному соотношению (2.13) устраниет произвол при выборе начального участка кривой рассеяния. Это даст более точную оценку для радиуса инерции, поскольку при вычислении  $\gamma(r)$  через функцию  $p(r)$  используется вся кривая рассеяния.

### 2.3. Решение обратной задачи рассеяния с помощью моделирования молекулярной структуры низкого разрешения по данным малоуглового рассеяния

Интенсивность рассеяния системой монодисперсных частиц в разбавленных растворах пропорциональна пространственно усредненному рассеянию от одной частицы [1, 2]:

$$\begin{aligned} I(s) &= \langle A^2(\mathbf{s}) \rangle_\Omega = \\ &= 4\pi \int_{r=0}^{\infty} r^2 \langle \eta(\mathbf{r}) \otimes \eta(\mathbf{r}) \rangle_\omega \frac{\sin(sr)}{sr} dr, \end{aligned} \quad (2.25)$$

где  $A(\mathbf{s})$  — амплитуда рассеяния,  $\langle \dots \rangle_\Omega$  и  $\langle \dots \rangle_\omega$  означает усреднение по углу соответственно в обратном и прямом пространствах, обусловленное хаотической ориентацией частиц в растворе,  $\otimes$  означает операцию свертки. Дискретным аналогом выражения (2.12) является формула Дебая (см. разд. 5).

Пространственная модель структуры частицы описывается функцией электронной плотности  $\eta(r)$ , которая в случае рентгеновского излучения пропорциональна разности электронной плотности молекулы и растворителя (контрасту). Из соотношения (2.22) видно, что непосредственное восстановление функции структуры  $\eta(r)$  по экспериментальным данным с помощью обратного фурье-преобразования невозможно. Однако можно рассчитать интенсивность рассеяния  $I(s)$  от любой трехмерной модельной структуры, затем вычислить разность между интенсивностью рассеяния на модели и экспериментальными данными и, минимизируя эту разность путем варьирования параметров модели методом нелинейной оптимизации, восстановить пространственное распределение рассеивающей плотности. В данной работе было использовано моделирование структуры частиц с помощью сферических объемных элементов («метод шаров»), которое позволяет восстановить структуру частиц в растворе с низким разрешением [6]. Метод

основан на представлении структуры с помощью набора небольших шариков, определенным образом расположенных в пространстве. Основным преимуществом данного подхода является отсутствие ограничений на сложность строения частицы. В случае однородных частиц интенсивность малоуглового рассеяния может быть вычислена по формуле [7]

$$I(s) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{lm}(s) A_{lm}^*, \quad (2.26)$$

где

$$A_{lm}(s) = i^l \sqrt{\frac{2}{\pi}} f(s) \sum_{j=1}^N j_l(sr_j) Y_{lm}^*(\Omega_j) \quad (2.27)$$

— амплитуда рассеяния структурой, состоящей из  $N$  сфер единичной плотности,  $r_j$ ,  $\omega_j$  — полярные координаты,  $f_k(s)$  — амплитуда рассеяния одной сферой,  $j_l(x)$  — сферические функции Бесселя,  $Y_{lm}(\Omega)$  — сферические гармоники. Для поиска пространственного расположения объемных элементов используется вариант программы решения нелинейной задачи наименьших квадратов с помощью глобальной минимизации, основанной на методе Монте-Карло с «имитацией отжига» в виде алгоритма Метрополиса [8]. Метод Метрополиса представляет собой один из вариантов марковского порогового алгоритма глобального случайного поиска [7, 9]. В процессе поиска программа минимизации инвертирует плотность случайным образом выбранного шарика, т. е. изменяет ее с плотности частицы (1) на плотность растворителя (0) или наоборот. После каждой модификации программа рассчитывает от текущей структуры кривую интенсивности малоуглового рассеяния по формуле Дебая (см. п. 5.1) и вычисляет взвешенное суммарное квадратичное отклонение от экспериментальной кривой рассеяния по  $M$  точкам:

$$\chi^2 = \sum_{j=1}^M \left[ \frac{I_{exp}(s_j) - I(s_j)}{\sigma(s_j)} \right]^2, \quad (2.28)$$

где  $\sigma(s_j)$  — оценка экспериментальной ошибки в  $j$ -й точке. Число  $M$  узлов на кривой рассеяния определяется на основании уже упоминавшейся теоремы Котельникова–Шеннона о передаче сигнала без потерь [5]. Согласно ей количество узлов

$$M = 2 \left[ \frac{(s_{max} - s_{min})D_{max}}{2\pi} \right] + 1,$$

где  $[ \dots ]$  — целая часть числа, оказывается достаточным для однозначного восстановления финитной функции по ее фурье-спектру.

По алгоритму, изложенному в работе [7], к сумме (2.27) добавляют штрафные члены, которые отражают требования неразрывности структуры ( $w_D P_D$ ) и отсутствия отдельно расположенных шаров ( $w_L P_L$ ),  $w_D$  и  $w_L$  — веса при штрафах  $P_D$  и  $P_L$  для функционала  $F = \chi^2 + w_D P_D + w_L P_L$ . Штраф  $P_D$  вычисляют как отношение общего числа шаров структуры к числу шаров, составляющих наибольший домен, понимаемый как полученный в процессе моделирования цельный «кусок» моделируемого тела, т. е. образованный непосредственно контактирующими атомами. Если значение целевой функции  $F$  меньше предыдущего значения ( $\Delta = F_k - F_{k-1} < 0$ ,  $k$  — номер шага), то структура запоминается. Процедура моделирования отжига состоит в том, что на каждом шаге в качестве «наилучшей» структуры с некоторой вероятностью программа может принимать ту, которая имеет худшее значение  $F$ . Вероятность задается параметром «температуры»  $T$ : если  $\Delta > 0$ , структура принимается с вероятностью  $\exp(-\Delta/T)$ . Со временем после каждого  $100K$  шагов (или  $10K$  шагов с  $\Delta < 0$ ) параметр  $T$  уменьшается ( $T_{k+1} = 0.9T_k$ ), что позволяет программе все реже запоминать неудачные структуры [7, 9, 10].

Необходимость процедуры отжига связана с тем, что целевая функция  $F$  может иметь несколько локальных минимумов. Возможность принятия худшего шага в качестве решения помогает программе не останавливаться на промежуточных решениях. Финальную пространственную конфигурацию шаров рассматривают как найденную модель структуры молекулы в растворе. В силу нелинейного характера зависимости параметров структурной модели (координат центров шариков) от величины интенсивности (2.26) в общем случае невозможно аналитически оценить степень близости решения к глобальному минимуму. Поэтому обычно проводится несколько независимых поисков решения для каждой кривой рассеяния с последующим их анализом.

### 3. РАСЧЕТ ИНТЕНСИВНОСТИ МАЛОУГОЛОВОГО РАССЕЯНИЯ ДЛЯ ЭЛЛИПТИЧЕСКОГО ТОРА

#### 3.1. Точное выражение фурье-трансформанты

Пусть задан потенциал, такой что  $U(r) = 1$  при  $\mathbf{r} \in V'$  и  $U(r) = 0$  при  $\mathbf{r} \notin V'$ , где  $V'$  — объем эллиптического тора, рис. 2, определяемый в цилиндрической системе координат выражением

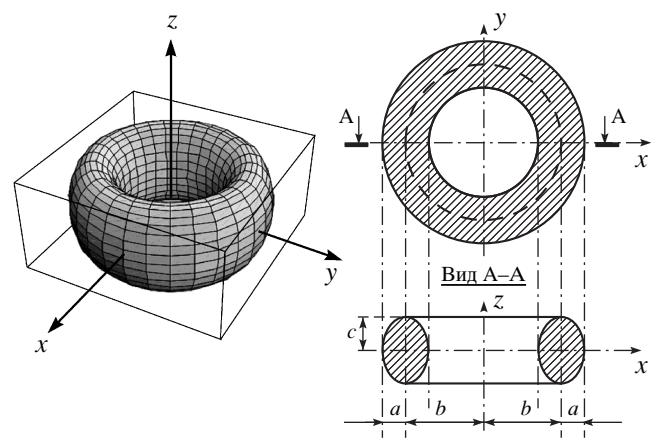


Рис. 2. Эллиптический тор и его главная плоскость сечения

$$\begin{aligned} V' \in \frac{(r-b)^2}{a^2} + \frac{z^2}{c^2} \leq 1, \quad b-a \leq r \leq b+a, \\ z_1 = -z_0, \quad z_2 = z_0, \quad z_0 = c\sqrt{1 - \frac{(r-b)^2}{a^2}}. \end{aligned} \quad (3.1)$$

Вычисление объема  $V'$  и радиуса инерции молекулы  $R_g$  дают следующие результаты:

$$\begin{aligned} V' &= \int_0^{2\pi} d\varphi \int_{b-a}^{b+a} r dr \int_{-z_0}^{z_0} dz = 2\pi^2 abc, \\ R_g^2 &= \frac{1}{2\pi^2 abc} \int_0^{2\pi} d\varphi \int_{b-a}^{b+a} r dr \int_{-z_0}^{z_0} (r^2 + z^2) dz = \\ &= \frac{3a^2 + c^2}{4} + b^2. \end{aligned} \quad (3.2)$$

Запишем радиус-вектор и вектор рассеяния в цилиндрической системе координат:  $\mathbf{r} = \{r \cos \psi, r \sin \psi, z\}$ ,  $\mathbf{s} = \{s_{xy} \cos \Psi, s_{xy} \sin \Psi, s_z\}$ ,  $s_{xy} = s \sin \theta$ ,  $s_z = s \cos \theta$ . В этой системе координат скалярное произведение примет вид  $\mathbf{r} \cdot \mathbf{s} = rs_{xy} \cos \alpha + zs_z$ , где  $\alpha \equiv |\psi - \Psi|$ . Подставляя это равенство в выражение (2.12) для амплитуды рассеяния, получим [11]

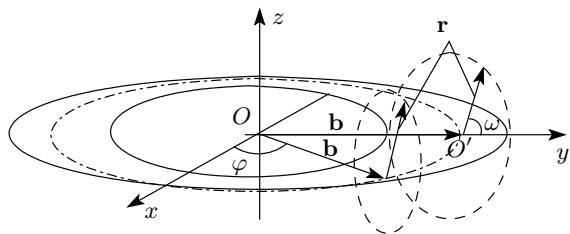


Рис. 3. Система координат для приближенного вычисления формфактора тора

$$\begin{aligned}
 \Phi(s, \theta) &= \frac{1}{V'} \times \\
 &\times \iiint_{V'} \exp(i(rs_{xy} \cos \alpha + zs_z)) r dr d\alpha dz = \\
 &= \frac{2\pi}{2\pi^2 abc} \int_{b-a}^{b+a} \left\{ \int_{-z_0}^{z_0} e^{izs_z} dz \right\} J_0(rs_{xy}) r dr = \\
 &= \frac{1}{\pi abc} \int_{b-a}^{b+a} \left\{ \frac{2}{s_z} \sin(s_z z_0) \right\} J_0(rs_{xy}) r dr = \\
 &= \frac{1}{\pi abc} \frac{2}{s \cos \theta} \int_{b-a}^{b+a} \sin \left( sc \sqrt{1 - \frac{(r-b)^2}{a^2}} \cos \theta \right) \times \\
 &\quad \times J_0(rs \sin \theta) r dr. \quad (3.3)
 \end{aligned}$$

Отсюда точное выражение для интенсивности рассеяния

$$I(s) = \int_0^{\pi/2} |\Phi(s, \theta)|^2 \sin \theta d\theta$$

примет вид

$$\begin{aligned}
 I(s) &= \left( \frac{2}{\pi abcs} \right)^2 \int_0^{\pi/2} \frac{1}{\cos^2 \theta} \times \\
 &\times \left\{ \int_{b-a}^{b+a} \sin \left( sc \sqrt{1 - \frac{(r-b)^2}{a^2}} \cos \theta \right) \times \right. \\
 &\quad \left. \times J_0(sr \sin \theta) r dr \right\}^2 \sin \theta d\theta. \quad (3.4)
 \end{aligned}$$

Интенсивность рассеяния от кругового тора получается как частный случай при  $c = a$ . Результаты моделирования для конкретных значений параметров тора по формуле (3.4) приведены на рис. 4 (кривая 1), см. далее.

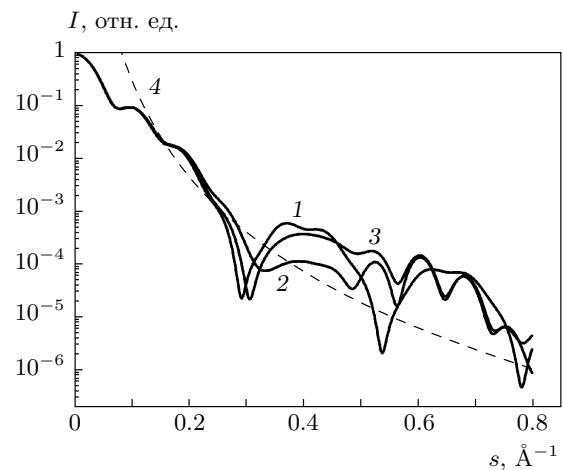


Рис. 4. Кривые МУР от кругового тора с параметрами  $a = c = 13 \text{ \AA}$ ,  $b = 38 \text{ \AA}$ , полученные через фурье-трансформанту: 1 — по точному выражению (3.4), 2 — по приближенному выражению (3.10), 3 — по приближенному выражению с поправкой  $\Delta$  (3.11), 4 — породовская асимптотика по (3.15)

### 3.2. Приближенное выражение фурье-трансформанты тора

Запишем параметрическое уравнение тора в координатах, показанных на рис. 3,

$$\mathbf{r} : \{(b + ar \cos \omega) \cos \varphi \mathbf{i}, (b + ar \cos \omega) \sin \varphi \mathbf{j}, cr \sin \omega \mathbf{k}\},$$

$$0 \leq r \leq 1, \quad 0 \leq \varphi \leq 2\pi, \quad 0 \leq \omega \leq 2\pi$$

и вектор рассеяния в уже используемой цилиндрической системе координат. Амплитуда рассеяния формфактора согласно выражению (2.12) при  $\rho_e \equiv 1$  примет вид

$$\begin{aligned}
 \Phi(s, r, \Omega) &= \frac{1}{V'} \int_{V'} e^{is \cdot \mathbf{r}} dV' = \frac{1}{2\pi^2 abc} \times \\
 &\times \int_0^{2\pi} \int_0^{2\pi} \int_0^{r_0} e^{ig(r, \omega)} acr(b + ar \cos \omega) dr d\alpha d\omega, \quad (3.5)
 \end{aligned}$$

$$g(r, \omega, \alpha) = s_{xy}(b + ar \cos \omega) \cos \alpha + rcs_z \sin \omega.$$

Далее мы полагаем, что  $\Phi(s, r, \Omega) = \Phi_1 + \Phi_2$ , где

$$\begin{aligned}
 \Phi_1 &= \frac{1}{2\pi^2} \int_0^{2\pi} \int_0^{2\pi} \int_0^{r_0} e^{ig(r, \omega, \alpha)} r dr d\alpha d\omega, \\
 \Phi_2 &= \frac{1}{2\pi^2} \frac{a}{b} \int_0^{2\pi} \int_0^{2\pi} \int_0^{r_0} e^{ig(r, \omega, \alpha)} r^2 \cos \omega dr d\alpha d\omega,
 \end{aligned} \quad (3.6)$$

причем  $\Phi_1 \gg \Phi_2$ , поскольку  $|e^{ig_{max}}| \leq 1$  и

$$\Phi_2 \approx |e^{ig_{max}}| \int_0^{r_0} \int_0^{2\pi} \cos \omega d\omega dr = 0.$$

Используя формулу сложения для функций Бесселя с нулевым по индексу функции Бесселя приближением:

$$\begin{aligned} J_0(z_1 + z_2) &= J_0(z_1)J_0(z_2) + \\ &+ 2 \sum_{n=1}^{\infty} (-1)^n J_n(z_1)J_n(z_2) \approx J_0(z_1)J_0(z_2), \end{aligned}$$

для интеграла  $\Phi_1$  получим приближение  $\Phi_{1,0}$ , которое после интегрирования по переменной  $\alpha$  примет вид

$$\begin{aligned} \Phi_{1,0} &= \frac{1}{2\pi^2} J_0(bs_{xy}) \times \\ &\times \int_0^1 r \int_0^{2\pi} 2\pi J_0(ars_{xy} \cos \omega) e^{irs_z \sin \omega} d\omega dr. \quad (3.7) \end{aligned}$$

Взять этот интеграл точно представляется затруднительным, поэтому используем следующее известное разложение для экспоненты:

$$\begin{aligned} e^{irs_z \sin \omega} &= J_0(r cs_z) + 2 \sum_{m=1}^{\infty} J_{2m}(r cs_z) \cos(2m\omega) + \\ &+ 2i \sum_{m=1}^{\infty} J_{2m-1}(r cs_z) \sin((2m-1)\omega). \quad (3.8) \end{aligned}$$

Ограничивааясь опять «нулевым по индексу приближением», получим

$$\begin{aligned} \Phi_{1,0,0} &= \frac{1}{\pi} J_0(bs_{xy}) \times \\ &\times \int_0^{2\pi} \int_0^1 J_0(ars_{xy} \cos \omega) J_0(r cs_z) r dr d\omega. \quad (3.9a) \end{aligned}$$

Поскольку  $s_z \gg s_{xy}$  и  $J_0(rs_z) \gg J_0(s_{xy})$ , основной вклад в интеграл по  $r$  дает  $J_0(rs_z)$ , поэтому, пользуясь теоремой о среднем, получим

$$\begin{aligned} \Phi_{1,0,0} &\approx \frac{1}{\pi} J_0(bs_{xy}) \cdot 2 \int_0^{\pi} J_0(as_{xy} \cos \omega) d\omega \times \\ &\times \int_0^1 r J_0(r cs_z) dr = \\ &= \frac{1}{\pi} J_0(bs_{xy}) \cdot 2\pi J_0^2\left(\frac{as_{xy}}{2}\right) \frac{J_1(cs_z)}{cs_z}. \quad (3.9b) \end{aligned}$$

При интегрировании по углу  $\omega$  мы воспользовались свойством интеграла от периодической функции, заменив промежуток интегрирования  $(0, 2\pi)$  промежутком  $(-\pi, \pi)$ , а поскольку подынтегральная функция четная, интеграл равен удвоенному значению на половине интервала. Погрешность вычисления при сведении кратного интеграла (3.9a) к повторному (3.9b) есть величина порядка  $\tilde{o}(\theta^2)$ , ее можно оценить, воспользовавшись табличным интегралом  $\int x J_v(ax) J_v(bx) dx$  [12] при  $v = 0$  и учитывая, что при малых углах  $s_{xy}/s_z = \operatorname{tg} \theta \approx \theta$ . Таким образом, в области малых углов при ограничениях на параметры тора  $a, b, c$  (в единицах длины) и модуля вектора рассеяния  $s$  (в единицах обратной длины):  $0 \leq as \leq 1$ ,  $0 \leq bs \leq 1$ ,  $0 \leq cs \leq 1$ , приближенное выражение для формфактора тора имеет вид

$$\begin{aligned} I(s)_{app} &= 4 \int_0^{\pi/2} J_0^2(bs \sin \theta) \times \\ &\times \left[ J_0^2\left(\frac{as}{2} \sin \theta\right) \frac{J_1(cs \cos \theta)}{cs \cos \theta} \right]^2 \sin \theta d\theta. \quad (3.10) \end{aligned}$$

Это «нулевое приближение» можно немного уточнить, используя первый член в сумме для действительной части в разложении (3.8). Обозначим поправку через  $\Delta$ , тогда

$$\begin{aligned} I(s)_{app} &= 4 \int_0^{\pi/2} J_0^2(bs \sin \theta) \times \\ &\times \left[ J_0^2\left(\frac{as}{2} \sin \theta\right) \frac{J_1(cs \cos \theta)}{cs \cos \theta} - \Delta \right]^2 \sin \theta d\theta, \quad (3.11) \end{aligned}$$

где

$$\begin{aligned} \Delta &= J_1^2\left(\frac{sa}{2} \sin \theta\right) \times \\ &\times \left( \frac{2(1 - J_0(sc \cos \theta))}{(sc \cos \theta)^2} - \frac{J_1(sc \cos \theta)}{sc \cos \theta} \right). \quad (3.12) \end{aligned}$$

Для наглядного сравнения точного выражения кривой интенсивности малоуглового рассеяния, рассчитанной по формуле (3.4), с приближенными выражениями (3.10), (3.11) на рис. 4 они графически представлены для тора с параметрами  $a = c = 13 \text{ \AA}$ ,  $b = 38 \text{ \AA}$ . Из него видно, что на участке, соответствующем первым 3.5 порядкам уменьшения интенсивности от «рассеяния в нулевой угол», все три кривые совпадают. В эксперименте по рассеянию рентгеновского излучения в разбавленном растворе, например, растворе белка, после вычитания рассеяния от раствора максимальное отношение сигнал/шум на кривой МУР обычно не превышает трех порядков

ков, поэтому в практических целях достаточно использовать соотношение (3.10). Кроме того, в отличие от точного выражения (3.4) выражение (3.11) позволяет вычислить асимптотику Порода при относительно больших углах рассеяния  $\theta \approx 10^\circ\text{--}15^\circ$ .

### 3.3. Асимптотики: область Гинье и область Порода

Для области Гинье разложим приближенное выражение (3.11) в степенной ряд

$$I(s)_{app} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{k+l+m} [(2l)!]^2 (2(m+1))!}{(k!)^4 (l!)^6 (m!)^3 ((m+1)!)^2} \times \\ \times \frac{(2m)!(k+l)!(k+l+m+1)!}{(m+1)(m+2)(2(k+l+m+1))! 2^{2(l+m)-1}} \times \\ \times (as)^{2l} (bs)^{2k} (cs)^{2m}. \quad (3.13)$$

Взяв сумму до квадратичных членов ряда (3.13), получим приближение Гинье (Guinier approximation) [1] ( $(as)(bs)(cs) \ll 1$ ). Выписывая слагаемые до квадратичных, получим электронный радиус инерции частицы:

$$I(s)_G \approx 1 - \frac{s^2}{3} \left( \frac{a^2 + b^2}{2} + \frac{c^2}{4} \right) \approx \\ \approx \exp \left( -\frac{s^2 \tilde{R}_g^2}{3} \right), \quad (3.14)$$

$$\tilde{R}_{g app}^2 = \frac{a^2 + b^2}{2} + \frac{c^2}{4}.$$

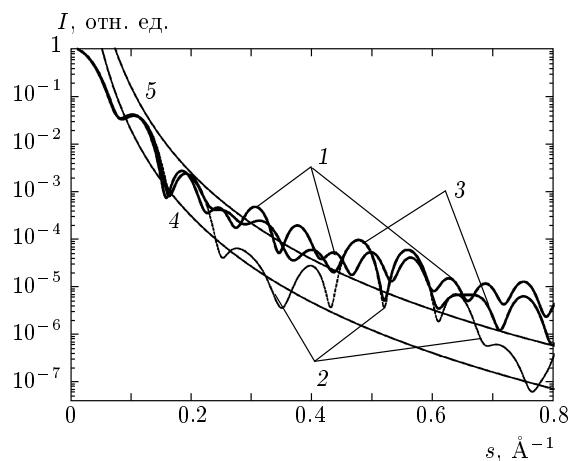
Воспользуемся полученным нами приближенным выражением (3.10) и известными асимптотиками для функций Бесселя при больших значениях аргумента [13]

$$J_v(z)|_{z \rightarrow \infty} \sim \sqrt{\frac{2}{\pi z}}$$

для приближения Порода (Porod) при  $sD_{max} \gg 1$ :

$$I(s)_P \approx \frac{4}{(bs)(as)^2(cs)^3}. \quad (3.15)$$

Это важный для экспериментальной практики результат, поскольку часто, не вдаваясь в тонкие особенности кривой рассеяния, по характеру спада можно сразу сказать, к какой форме ближе рассеивающий объект: глобуллярной частице (шар, эллипсоид), сильно вытянутой (стержень) или сплюснутой (диск). К тому же, если вспомнить породовские асимптотики от диска  $I(s)_{app} \sim s^{-2}$  и цилиндра  $I(s)_{app} \sim s^{-3}$ , то полученная асимптотика для тора не лишена некоторой физической интерпретации. В самом деле, проекция образующей тора на плоскость  $xy$  характеризуется в прямом пространстве одной из главных осей эллипса  $a$ , а в обратном пространстве — проекцией модуля вектора рассеяния



**Рис. 5.** Кривые МУР от эллиптического тора с параметрами  $a = 15 \text{ \AA}$ ,  $b = 35 \text{ \AA}$ ,  $c = 30 \text{ \AA}$ , рассчитанные: 1 — по точному выражению (3.4), 2 — по приближенной формуле (3.10), 3 — по приближенной формуле с поправкой  $\Delta$  (3.11), 4 — породовская асимптотика по (3.15), 5 — обобщенная породовская асимптотика (3.16)

$s_{xy} = s \sin \theta$  (см. рис. 1) и лежит в одной плоскости с волновым вектором  $\mathbf{k}_0$  падающего на образец излучения, величины этих проекций и определяют «степень сплюснутости» формы тора. В знаменателе выражения (3.15) стоит произведение  $as$  в квадрате, это и соответствует асимптотике сплюснутой частицы.

Аналогично проекция образующей тора на вертикальную плоскость  $xz$  характеризуется второй главной осью эллипса  $c$ , а в обратном пространстве —  $s_z = s \cos \theta$ , и соответственно произведение  $cs$  входит в (3.15) с асимптотикой цилиндра.

Следует заметить, что для кругового тора с параметрами  $b = 38 \text{ \AA}$ ,  $a = c = 13 \text{ \AA}$  ход асимптотики хорошо соответствует точной кривой рассеяния, рассчитанной по формуле (3.4) (см. рис. 5). Но поскольку при выводе «породовской асимптотики» (3.15) мы использовали приближенное выражение (3.10) для формфактора тора, величины  $a$ ,  $b$ ,  $c$ , входящие в (3.15), строго говоря, несколько отличны от параметров тора. Например, для эллиптического тора, вытянутого вдоль оси  $z$  ( $b = 35 \text{ \AA}$ ,  $a = 15 \text{ \AA}$ ,  $c = 30 \text{ \AA}$ ), асимптотика, вычисленная по (3.15), расположена несколько ниже кривой рассеяния (рис. 5, кривая 4). Поэтому представляется целесообразным использовать следующее выражение в «области Порода»  $sD_{max} \gg 1$ :

$$I(s)_P \sim \frac{4}{bs(ps)^5}, \quad (3.16)$$

где  $p = \min(a, c)$ . Это замечание не является принципиальным, поскольку, во-первых, нам важна «степень уменьшения» в зависимости от модуля вектора рассеяния, а она определена в (3.15) как  $I(s)_P \sim s^{-6}$  и, во-вторых, в эксперименте после решения задачи «редукции к идеальному прибору», а затем вычисления функции  $p(r)$  мы знаем только максимальный размер частицы.

#### 4. ИСПОЛЬЗОВАНИЕ ФОРМФАКТОРА ТОРА ДЛЯ МОДЕЛИРОВАНИЯ МОЛЕКУЛЫ БЕЛКА ШАПЕРОНИНА GroEL

Молекула шаперонина GroEL [14, 15] представлена на рис. 10 (см. далее). Для использования формфактора тора для моделирования молекулы белка шаперонина GroEL обобщим выражения (3.3), (3.4) для случая двойного тора:

$$\Phi(s, \theta) = \frac{1}{\pi abc} \int_{b-a}^{b+a} \left\{ \int_{-z_1}^{z_1} e^{izs_z} dz - \int_{-z_2}^{z_2} e^{izs_z} dz \right\} \times J_0(rs_{xy}) r dr, \quad (4.1)$$

где

$$\begin{Bmatrix} z_1 \\ z_2 \end{Bmatrix} = c(1 \pm \delta), \quad \delta = \sqrt{1 - \frac{(r-b)^2}{a^2}}, \quad (4.2)$$

$$I(s) = \left( \frac{2}{\pi abcs} \right)^2 \int_0^{\pi/2} \frac{1}{\cos^2 \theta} \times \times \left\{ \int_{b-a}^{b+a} \{ \sin(sc(1+\delta) \cos \theta) - \sin(sc(1-\delta) \cos \theta) \} \times \times J_0(sr \sin \theta) r dr \}^2 \right\} \sin \theta d\theta. \quad (4.3)$$

Общий вид двойного тора представлен на рис. 11а (см. далее). Так же хорошей моделью этого белка может служить полый цилиндр с внешним радиусом  $r_{max} = b + a$ , внутренним радиусом  $r_{min} = b - a$  и высотой  $H = 2L = 4c$ , его объем

$$V' = \pi (r_{max}^2 - r_{min}^2) H.$$

Используя первое полученное в (3.3) выражение для формфактора в цилиндрической системе координат, запишем формфактор полого цилиндра в виде

$$\Phi(s, \theta) = \frac{1}{\pi (r_{max}^2 - r_{min}^2) H} \int_{-L}^L \exp(izs_z) dz \int_{r_{min}}^{r_{max}} r dr \int_0^{2\pi} \exp(i rs_{xy} \cos \alpha) dz = \frac{1}{\pi (r_{max}^2 - r_{min}^2) 2L} \frac{2 \sin(s_z L)}{s_z} \times \times \int_{r_{min}}^{r_{max}} 2\pi J_0(s_{xy} r) r dr = \frac{2}{(r_{max}^2 - r_{min}^2) L} \frac{\sin(sL \cos \theta)}{s \cos \theta} \frac{r_{max} J_1(sr_{max} \sin \theta) - r_{min} J_1(sr_{min} \sin \theta)}{s \sin \theta}. \quad (4.4)$$

Соответственно интенсивность рассеяния от цилиндра имеет вид

$$I(s) = \frac{4}{(r_{max}^2 - r_{min}^2)^2} \times \times \int_0^{\pi/2} \left( \frac{r_{max} J_1(sr_{max} \sin \theta) - r_{min} J_1(sr_{min} \sin \theta)}{s \sin \theta} \right) \times \times \left( \frac{\sin(Ls \cos \theta)}{Ls \cos \theta} \right)^2 \sin \theta d\theta. \quad (4.5)$$

Таким образом, мы получили формфакторы шаперонина GroEL в двух аппроксимациях двойным тором и полым цилиндром.

#### 5. ПРЯМАЯ И ОБРАТНАЯ ЗАДАЧИ РАССЕЯНИЯ

##### 5.1. Формула Дебая и восстановление тора по неполным данным

Под прямой задачей понимается расчет по формуле Дебая интенсивности рассеяния от атомной структуры молекулы или шариковой модели молекулы. Для этого нужно знать из данных рентгеноструктурного анализа для данной молекулы координаты пространственного расположения каждого ее атома и его заряд  $Z$ . В этом случае статическое или мгновенное значение результирующей амплитуды  $A(s)$  рассеяния для некоторого произвольного, дискретного набора рассеивателей с амплитудами  $f_j^0(s)$  (так называемые атомные формфакторы) имеет вид

$$A(s) = \sum_{j=1}^n f_j^0 e^{is \cdot \mathbf{r}_j},$$

где  $\mathbf{s} = \mathbf{k} - \mathbf{k}_0$ ,  $|s| = 4\pi \sin \theta / \lambda$  — вектор рассеяния и его модуль,  $\mathbf{k}$  и  $\mathbf{k}_0$  — волновые векторы рассеянной и падающих волн,  $n$  — число атомов в молекуле или частице. Интенсивность рассеяния получим умножением на комплексно сопряженную величину. Поскольку по отношению к падающему лучу молекулы в растворе не ориентированы, полученную интенсивность следует усреднить по телесному углу:

$$\begin{aligned} \langle I(s) \rangle_\Omega &= I_e \sum_{i=1}^n \sum_{j=1}^n f_i^0(s) f_j^0(s) \langle e^{i\mathbf{s} \cdot \mathbf{r}_{ij}} \rangle_\Omega = \\ &= I_e \sum_{i=1}^n \sum_{j=1}^n f_i^0(s) f_j^0(s) \frac{\sin(sr_{ij})}{sr_{ij}}, \end{aligned} \quad (5.1)$$

где вектор  $\mathbf{r}_{ij}$  — разность между радиус-векторами  $\mathbf{r}_i$ ,  $\mathbf{r}_j$   $i$ -го и  $j$ -го атомов, а суммирование ведется по всем атомам, составляющим частицу. Следует заметить, что формула Дебая в виде (5.1) не учитывает тепловые колебания атомов; атомный фактор (случай аномального рассеяния мы не рассматриваем) для рентгеновского излучения имеет вид

$$\begin{aligned} f_i^0(s) &= 4\pi \int_0^\infty \rho_{ai}(r) \frac{\sin(sr)}{sr} r^2 dr, \\ \rho_i(r) &= 2 \sum_{a=1}^{n_a/2} \psi_{ai}^*(r) \psi_{ai}(r), \end{aligned} \quad (5.2)$$

где  $\rho_i(r)$  — атомная электронная плотность, выраженная через полную волновую функцию  $i$ -го атома  $\psi_{ai}(r)$ , вычисленную по методу самосогласованного поля Хартри–Фока (Hartree–Fock);  $n_a$  — полное число электронов в атоме. В данном случае в приближении заполненных оболочек учитывается только радиальная часть волновой функции. Вычисленные по формуле (5.2) значения атомных факторов как функций модуля вектора рассеяния были взяты из международных таблиц для рентгеновской кристаллографии [16]. Двойную сумму (5.1) можно представить в виде двух частей: первой — простой суммы, для которой  $i = j$ , т. е. суммы квадратов атомных амплитуд (в ней содержится  $n$  слагаемых), и второй — остаточной суммы с  $i \neq j$  ( $n(n-1)$  слагаемых):

$$\begin{aligned} I(s) &= \sum_{i=1}^n f_i^2(s) + \\ &+ 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n f_i(s) f_j(s) \frac{\sin(sr_{ij})}{sr_{ij}}. \end{aligned} \quad (5.3)$$

Если атомную систему заменить системой шаров радиуса  $a$  с одинаковой рассеивающей плотностью, то интенсивность рассеяния (4.7) от такой модельной системы примет более простой вид [5]:

$$I_M(s) = n\Phi^2(sa) \left[ 1 + \frac{2}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\sin(sr_{ij})}{sr_{ij}} \right], \quad (5.4)$$

где  $\Phi^2(sa)$  — формфактор шара,

$$\Phi^2(sa) = \left( 3 \frac{\sin(sa) - sa \cos(sa)}{(sa)^3} \right)^2. \quad (5.5)$$

Используемая программная реализация (5.4) представлена в работе [17].

На рис. 6а представлена стартовая модель кругового тора с параметрами  $b = 38 \text{ \AA}$ ,  $a = c = 13 \text{ \AA}$ . Внутренний объем тора набивается плотноупакованными «шариками-атомами» (см. рис. 6б). Далее для этой модели численно вычислялись кривые МУР по формуле Дебая (5.1) и по точному выражению (3.4) (рис. 7а). На рис. 7а видно хорошее совпадение кривых рассеяния, вычисленных по формуле (3.4) и по формуле Дебая для шариковой модели (5.4).

Компьютерное моделирование восстановления формы тора по кривой рассеяния проводилось по двум участкам полученной ранее кривой рассеяния. Был выбран круговой тор с параметрами  $b = 38 \text{ \AA}$ ,  $a = c = 13 \text{ \AA}$  и два различных по ширине интервала для модуля вектора рассеяния для одной и той же кривой рассеяния. Первый — это область  $0 \leq s \leq 0.5 \text{ \AA}^{-1}$  (см. рис. 7а), второй —  $0 \leq s \leq 0.27 \text{ \AA}^{-1}$  (см. рис. 7б), что соответствует пяти и трем порядкам уменьшения интенсивности рассеяния по оси ординат, и числу шенноновских каналов по оси модуля вектора рассеяния соответственно 15 и 9.

Как известно, для тел вращения, к которым относится и тор, прямая задача имеет единственное решение. При этом наличие единственного трехмерного фурье-образа, к тому же усредненного по полярному углу, что соответствует классической съемке с однокоординатным детектором, не гарантирует единственности восстановления формы тела в трехмерном пространстве по скалярной функции  $I(s)$ .

Вычисление функции распределения по расстояниям (функция парных расстояний)  $p(r)$  для кругового тора с параметрами  $b = 38 \text{ \AA}$ ,  $a = c = 13 \text{ \AA}$  проводилось по методике, описанной в п. 2.2. Ее график представлен на рис. 7в, из которого видно, что функция парных расстояний для тора имеет два сильно выраженных максимума. По кривой рассеяния,

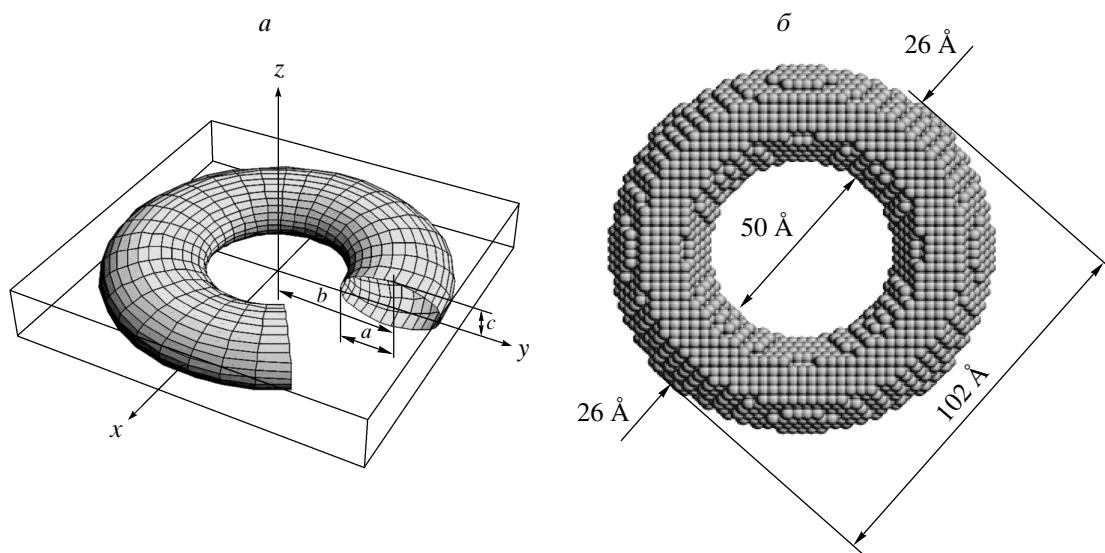


Рис. 6. а) Поверхность эллиптического тора с параметрами  $a$ ,  $b$ ,  $c$ . б) Шариковая модель тора при  $a = c$ , составленного из плотноупакованных «шариков-атомов»

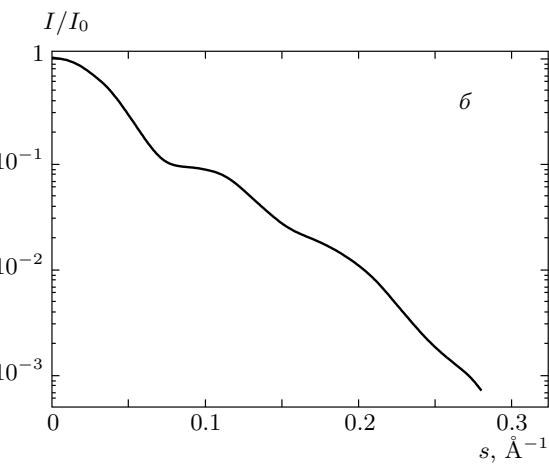
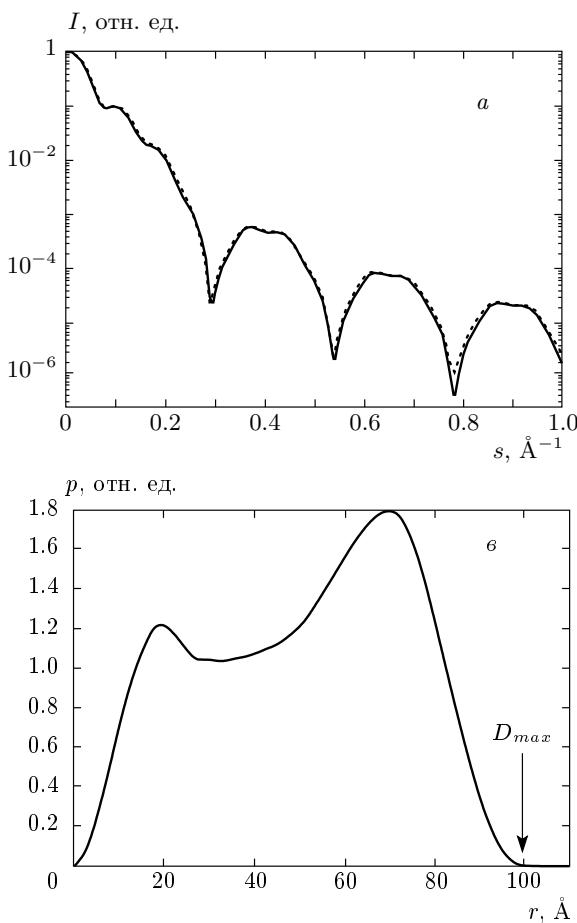
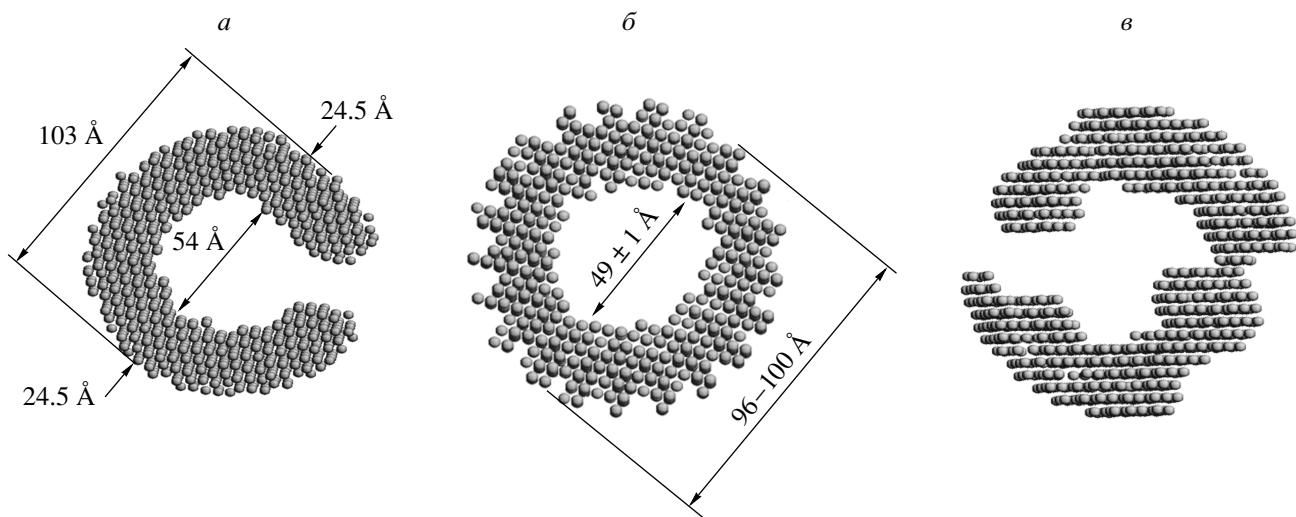


Рис. 7. а) Кривые рассеяния, вычисленные для кругового тора с параметрами  $b = 38 \text{ \AA}$ ,  $a = c = 13 \text{ \AA}$  по формуле Дебая (5.1) — точки, по точному выражению (3.4) — сплошная кривая. б) Кривая рассеяния (сплошная линия на рис. 7а) на интервале  $0 \leq s \leq 0.28 \text{ \AA}^{-1}$ . Число шенноновских каналов 9. в) Функция парных расстояний, вычисленная от кривой МУР рис. 7а



**Рис. 8.** Решение обратной задачи. Типичное (а) и наилучшее (б) восстановление формы тора по кривой рис. 7а на интервале  $0 \leq s \leq 0.5 \text{ \AA}^{-1}$  с числом шенноновских каналов 15. Характерное восстановление (в) формы тора по кривой рис. 7б

взятой на указанных выше интервалах, были получены одни и те же величины для радиуса инерции  $R_g = (40 \pm 1.5) \text{ \AA}$ , см. (2.24), и максимального размера тора  $D_{max} = (102 \pm 3) \text{ \AA}$ . Радиус инерции, вычисленный по приближенной формуле (3.15) с параметрами  $b = 38 \text{ \AA}$ ,  $a = c = 13 \text{ \AA}$ , равен  $R_{gapp} = 38.02 \text{ \AA}$ , а по точному выражению (3.2) —  $R_g = 40.23 \text{ \AA}$ , что также согласуется со значением радиуса инерции, полученным через функцию распределения по расстояниям.

На следующем этапе после определения максимального размера через функцию распределения по расстояниям осуществлялось моделирование формы тора согласно методу, изложенному в п. 2.3. Стартовая область задавалась в виде сферы диаметром  $D_{max}$ , ее внутренний объем набивался мелкими шариками — моделями атомов. В данном случае число моделей атомов в стартовом шаре было равно 7000, с одинаковым диаметром  $2.5 \text{ \AA}$  каждый.

Типичная и наилучшая восстановленные формы для кругового тора представлены на рис. 8а, б, а характерное восстановление формы тора по ограниченному участку  $0 \leq s \leq 0.3 \text{ \AA}^{-1}$  по первым трем порядкам уменьшения интенсивности МУР — на рис. 8в.

Аналогичная процедура была проделана для восстановления формы эллиптического тора с параметрами  $b = 35 \text{ \AA}$ ,  $a = 15 \text{ \AA}$ ,  $c = 30 \text{ \AA}$  (рис. 9а). Кривая МУР вычислялась аналитически по выражению (3.3) на интервале  $0 \leq s \leq 0.5 \text{ \AA}^{-1}$  (см.

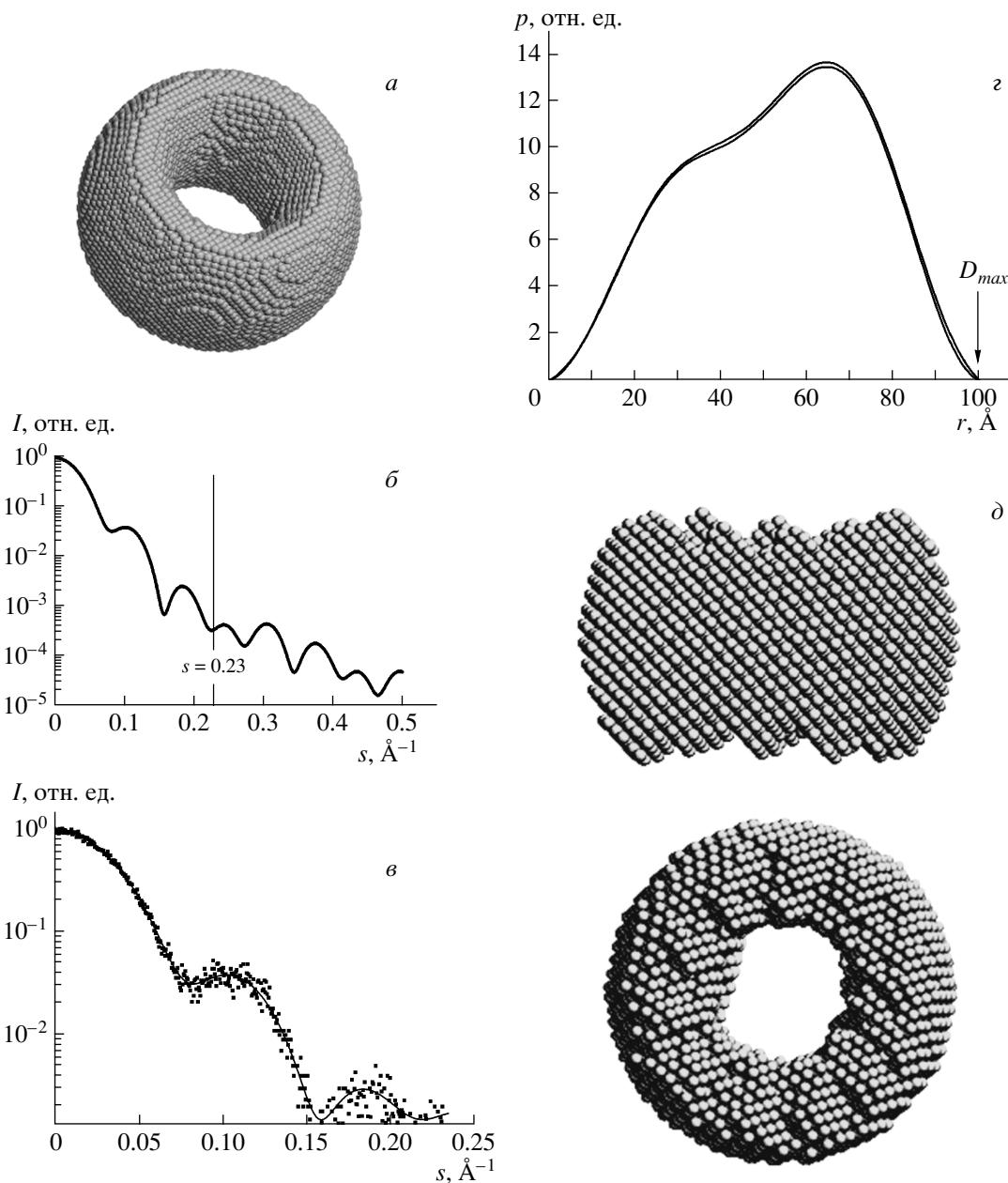
рис. 9б). Восстановление тора проводилось по ограниченной в интервале  $0 \leq s \leq 0.23 \text{ \AA}^{-1}$  и искусственно зашумленной по Пуассону кривой рассеяния (см. рис. 9б). Из функции парных расстояний, представленной на рис. 9г, определен максимальный размер  $D_{max} = (100 \pm 3) \text{ \AA}$ .

Результаты моделирования формы эллиптического тора представлены на рис. 9д, исходя из которого можно сделать вывод о достаточно высокой точности ( $< 10 \text{ \AA}$ ) передачи размеров эллиптического тора, об отсутствии деформаций в главных сечениях восстановленного тора и о наличии неровностей с торцевых сторон.

## 5.2. Восстановление формы молекулы белка шаперонина GroEL

Олигомерный белок теплового шока из клеток *Escherichia coli* GroEL является одним из представителей семейства шаперонов,участвующим в сворачивании других белков. Из данных по электронной микроскопии и рентгеноструктурному анализу известно, что функционирование GroEL как молекулярного шаперона связана с его конформационной изменчивостью. Закристаллизованная структура GroEL представляет собой цилиндр высотой  $145 \text{ \AA}$  и диаметром  $135 \text{ \AA}$  с центральным каналом, диаметр которого составляет  $45 \text{ \AA}$  [12, 13].

На рис. 10а, б представлена кристаллическая структура шаперонина GroEL (Protein Data Bank



**Рис. 9.** а) Стартовая модель эллиптического тора с параметрами  $a = 15 \text{ \AA}$ ,  $b = 35 \text{ \AA}$ ,  $c = 30 \text{ \AA}$ . б) Кривая МУР, вычисленная по (3.3). в) Начальный участок  $0 \leq s \leq 0.23 \text{ \AA}^{-1}$  с пуассоновским шумом кривой МУР, по которому далее будет проходить восстановление формы тора. г) Функция распределения по расстояниям, вычисленная на основе кривой МУР рис. 9в. д) Решение обратной задачи рассеяния: восстановление формы тора по кривой рис. 9в и функции распределения по расстояниям рис. 9г

(PDB) с Digital Object Identifier (DOI) 1kpo), который является олигомерным белком, состоящим из 14 идентичных субъединиц (57 кДа каждая), образующая два тороидальных кольца (по 7 субъединиц в каждом), лежащих друг на друге. Исследование структуры этого белка в растворе, в частности, методом малоуглового рассеяния,

представляет интерес для выяснения механизма шаперонзависимого сворачивания белков. Известно, что пространственная структура белка в растворе в его функционально активном состоянии может несколько отличаться от закристаллизованной фазы, в данном случае возможны пространственные подвижки субъединиц шаперонина.

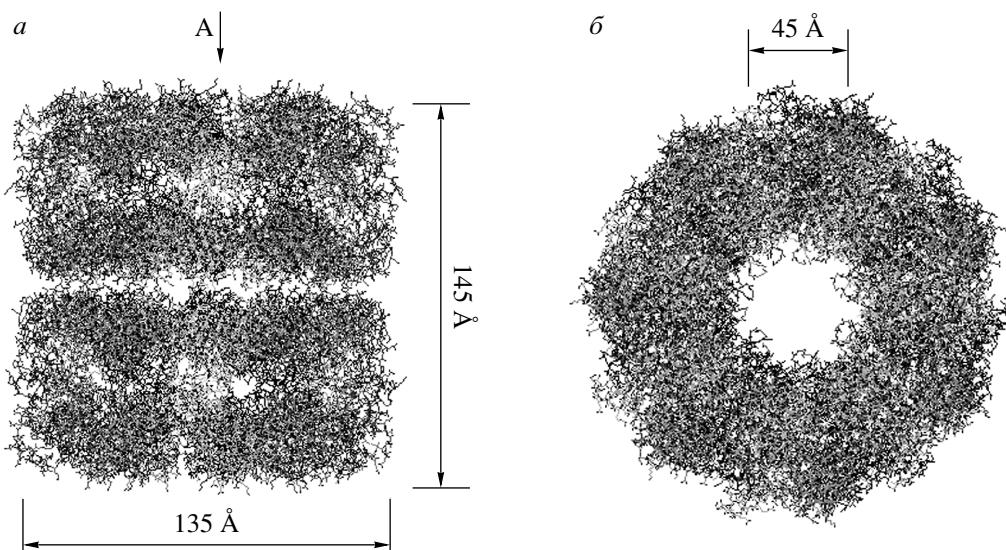


Рис. 10. Пространственная структура молекулы белка шаперона GroEL. Вид сбоку (a), вид А сверху (b)

Расчет кривых МУР для этого олигомерного белка проводился по полученным в разд. 4 формфакторам двойного эллиптического тора (рис. 11 $a$ ) с параметрами  $a = 22.5 \text{ \AA}$ ,  $b = 45 \text{ \AA}$ ,  $c = 36.5 \text{ \AA}$  и полого цилиндра с внешним радиусом  $r_{max} = 67.5 \text{ \AA}$ , внутренним радиусом  $r_{min} = 22.5 \text{ \AA}$  и высотой  $L = 146 \text{ \AA}$ . Результаты этих вычислений представлены на рис. 11 $b$ , кривая 3. Кривая 1 рассчитана непосредственно от атомной структуры (см. рис. 10) по формуле Дебая, в которой учитывается различный вклад в суммарную амплитуду для различных атомов (атомный формфактор). Кривые рассеяния 2 и 3 на рис. 11 $b$  вычислены по полученным ранее аналитическим выражениям для формфакторов двойного эллиптического тора (4.3) и полого цилиндра (4.5). В дальнейшем вычисление функции распределения по расстояниям и восстановление формы молекулы GroEL проводилось по кривой 2 на рис. 11 $b$  без ввода какой-либо априорной информации о форме молекулы (вытянутая или сплюснутая форма) или наличии оси или плоскости симметрии у предполагаемой молекулы. По второму нулю функции парных расстояний был определен максимальный размер молекулы  $D_{max}$ , он оказался равным  $180 \pm 3 \text{ \AA}$  (рис. 11 $b$ ). В данном случае максимальным размером или отрезком максимальной длины  $D_{max}$  является диагональ в любом из главных сечений молекулы, содержащая ось  $z$ .

Результаты восстановления формы молекулы шаперона GroEL, полученные по описанной выше методике, представлены на рис. 11 $c$ .

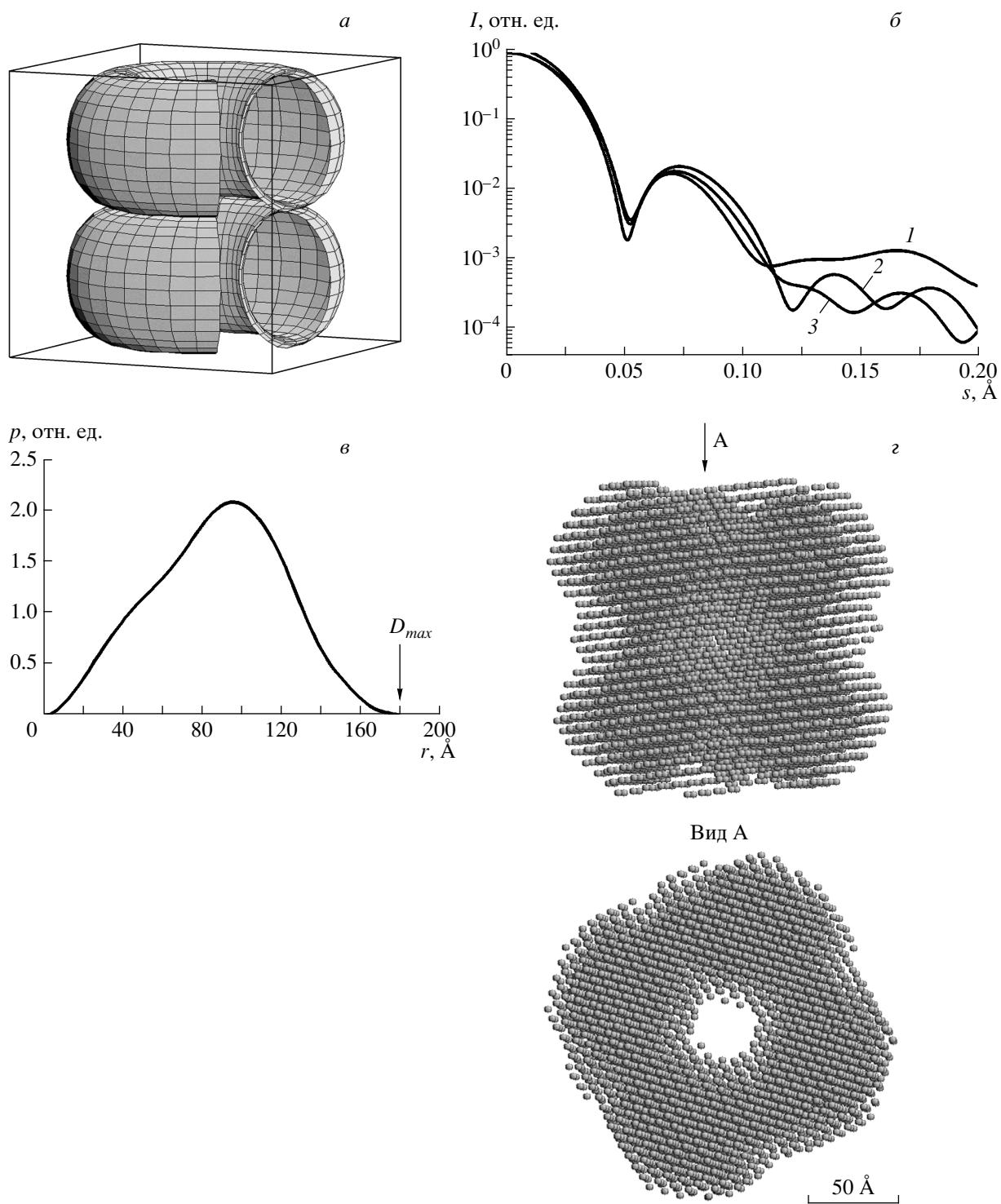
## 6. ВЫВОДЫ

1. По полученной в работе [11] точной фурье-трансформантой тора (см. (3.3)) была написана компьютерная программа расчета кривой интенсивности малоуглового рассеяния  $I(s)$  при заданных параметрах тора  $a$ ,  $b$ ,  $c$ , длине волны падающего излучения  $\lambda$  и интервале изменения половины угла рассеяния  $\theta \in [0, \theta_{max}]$ .

2. Получено приближенное выражение для формфактора тора, содержащее лишь интеграл по углу  $\theta$  (см. (3.10)) и вычислена уточняющая к нему поправка  $\Delta$  (см. (3.11) и (3.12)). Выведенное выражение (3.11) может быть реализовано в эксперименте по малоугловому рассеянию «от разбавленных растворов» для идентификации наночастиц, имеющих топологически двусвязную поверхность.

3. В разд. 2, 3 показано, что кривые рассеяния, вычисленные по приближенному выражению (3.11) и точному (3.3), совпадают в области первых четырех порядков уменьшения интенсивности от оси падающего луча, что вполне достаточно для практического использования в обработке экспериментальных кривых.

4. По выражению (3.10) удалось получить приближение Гинье  $(as)(bs)(cs) \ll 1$  (3.14) и асимптотику Порода (3.15), которые соответственно позволяют проверить закон Гинье на начальном участке кривой рассеяния вблизи оси падающего на частицу луча,  $sD_{max} \leq 1$ ,  $s \in [s_{min}, s_{max}]$ , где  $s_{min} \leq 1/R_g$ , и закон уменьшения интенсивности в области Порода  $sD_{max} \gg 1$ .



**Рис. 11.** *a)* Поверхность двойного эллиптического тора как модель молекулы GroEL. *б)* Кривые рассеяния: 1 — от атомной структуры молекулы белка GroEL, рассчитанные по формуле Дебая (5.1); 2 — от двойного тора рис. 11*a* по формуле (4.3) при  $a = 22.5 \text{ \AA}$ ,  $b = 45 \text{ \AA}$ ,  $c = 36.5 \text{ \AA}$ ; 3 — от полого цилиндра по формуле (4.5) при  $r_{max} = 67.5 \text{ \AA}$ ,  $r_{min} = 22.5 \text{ \AA}$ ,  $L = 146 \text{ \AA}$ . *в)* Функция распределения по расстояниям, вычисленная по кривой 2, рис. 11*б*. *г)* Восстановленная форма молекулы шаперона GroEL. Получена по кривой 2 рис. 11*б* и вычисленной из нее функции распределения по расстояниям (рис. 11*в*)

5. Продемонстрировано решение обратной задачи рассеяния для кругового и эллиптического торов по различным начальным участкам кривой рассеяния  $I(s)$ , что наравне с результатами разд. 4 важно для прикладной реализации восстановления формы частицы. По результатам прямого моделирования определена угловая область и число шенноновских каналов на кривой интенсивности рассеяния, по которой происходит восстановление формы тора его шариковой моделью без существенных искажений.

6. Получено устойчивое восстановление формы эллиптического тора по ограниченной и зашумленной кривой рассеяния.

7. По формуле Дебая была рассчитана кривая интенсивности МУР от имеющейся в Protein Data Bank атомной структуры белка GroEL и проведено моделирование этой молекулы двойным эллиптическим тором и полым цилиндром. Затем по этим моделям при конкретных значениях параметров были приближенно найдены в аналитическом виде форм-факторы белка GroEL и проведено восстановление пространственной формы этого белка по кривой МУР, полученной из модели двойного эллиптического тора. При компьютерном моделировании никакая другая информация (наличие осей или плоскостей, симметрии, вытянутость или сплюснутость молекулы) не вводилась.

## ЛИТЕРАТУРА

1. O. Glatter and O. Kratky, *Small Angle X-ray Scattering*, Acad. Press, London (1982).
2. Д. И. Свергун, Л. А. Фейгин, *Рентгеновское и нейтронное малоугловое рассеяние*, Наука, Москва (1987).
3. А. Н. Тихонов, В. Я. Арсенин, *Методы решения некорректных задач*, Наука, Москва (1986).
4. D. I. Svergun, J. Appl. Cryst. **25**, 495 (1992).
5. К. Э. Шенон, *Работы по теории информации и кибернетике*, Изд-во иностр. лит., Москва (1963).
6. D. I. Svergun, Biophys. J. **76**, 2879 (1999).
7. H. B. Stuhrmann, Acta Cryst. A **26**, 297 (1970).
8. N. Metropolis et al., J. Chem. Phys. **21**, 1087 (1953).
9. А. А. Жиглявский, А. Г. Жлинская, *Методы поиска глобального экстремума*, Наука, Москва (1991).
10. Г. Винклер, *Анализ изображений, случайные поля и динамические методы Монте-Карло. Математические основы*, Изд-во СО РАН, филиал «Гео», Новосибирск (2002).
11. T. Kawaguchi, J. Appl. Cryst. **34**, 580 (2001).
12. И. С. Градштейн, И. М. Рыжик, *Таблицы интегралов, сумм, рядов и произведений*, Физматлит, Москва (1962).
13. Г. Н. Ватсон, *Теория бесселевых функций*, Изд-во иностр. лит., Москва (1949).
14. K. Braig, Z. Otwinowski, R. Hegde, D. C. Boisvert, A. Joachimiak, A. L. Horwich, and P. B. Sigler, Nature **371**, 578 (1994).
15. K. Braig, P. D. Adams, and A. T. Brunger, Nat. Struct. Biol. **2**, 1083 (1995).
16. International Tables for X-ray Crystallography, Vol. IV, Kynoch Press, Birmingham (1973).
17. D. Svergun, C. Barberato, and H. J. Koch, J. Appl. Cryst. **28**, 768 (1995).