

Исследование структуры однослойных везикул DMPC с использованием параллельного алгоритма асинхронной дифференциальной эволюции

Е. И. Жабицкая[†], Е. В. Земляная[†], М. А. Киселёв^{*}

^{*} Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980

[†] Международный университет «Дубна»,
Дубна, Московская область, Россия, 141980

Модель разделённых форм-факторов (Separated Form Factors, SFF), разработанная ранее для обработки спектров малоуглового рассеяния нейтронов, в настоящей работе адаптирована для численного анализа структуры полидисперсной популяции однослойных везикул по данным малоуглового синхротронного рассеяния (Small Angle X-ray Scattering, SAXS). Параметры SFF-модели, определяющие структуру везикулярной системы (средний радиус везикул, полидисперсность, толщина бислоя и др.), вычисляются путём фитирования к экспериментальным данным SAXS. Процедура фитирования основана на использовании алгоритма Асинхронной Дифференциальной Эволюции (АДЭ) – эффективного метода поиска глобального минимума. Разработана параллельная реализация предложенного подхода на базе технологии MPI (Message Passing Interface). Проведено численное исследование структуры полидисперсной везикулярной системы димиристоилфосфатидилхолина (DMPC) в 40% водном растворе сахарозы. Показано, что для согласия правой части спектра рассчитанных SAXS-кривых с имеющимися экспериментальными данными необходим учёт флуктуации толщины бислоя. На основе расчётов с разными моделями внутренней структуры бислоя сделаны заключения о наиболее адекватной форме плотности длины синхротронного рассеяния поперёк мембраны. Представлены результаты расчётов на многопроцессорном кластере ЛИТ ОИЯИ (Дубна), подтверждающие эффективность параллельной реализации метода АДЭ.

Ключевые слова: везикулярные системы, малоугловое синхротронное рассеяние, Асинхронная Дифференциальная Эволюция, глобальная минимизация, параллельные вычисления.

1. Введение и постановка задачи в рамках SFF

Информация о структуре и свойствах однослойных везикул играет важную роль в структурной биологии, биохимии и фармакологии. В работах [1, 2] развит и успешно использован метод разделённых форм-факторов (SFF) для анализа структуры полидисперсных везикулярных систем по данным малоуглового рассеяния нейтронов (SANS). В настоящей работе метод SFF адаптирован для моделирования структуры полидисперсной популяции однослойных везикул димиристоилфосфатидилхолина (DMPC) в 40% водном растворе сахарозы на основе данных малоуглового синхротронного рассеяния (SAXS), измеренных на установке A2 (DESY, Германия).

В рамках SFF макроскопическое когерентное сечение монодисперсной популяции везикул определяется следующей формулой

$$\frac{d\Sigma}{d\Omega_{\text{mon}}}(q) = n_x F_s(q, R) F_b(q, \Theta_b), \quad (1)$$

где n_x — число везикул в единице объёма, q — вектор рассеяния, $F_s(q, R)$ и $F_b(q, \Theta_b)$ — соответственно формфактор сферы радиуса R и формфактор симметричного липидного бислоя:

Статья поступила в редакцию 30 декабря 2013 г.

Работа выполнена при частичной финансовой поддержке РФФИ (гранты 13-01-00060а, 12-01-00396а).

$$F_s(q, R) = \left(4\pi R^2 \frac{\sin(qR)}{qR} \right)^2, \quad F_b(q, \Theta_b) = \left(\int_{-\frac{d}{2}}^{\frac{d}{2}} \rho(x, \Theta_b) \cos(qx) dx \right)^2. \quad (2)$$

Здесь $\rho(x, \Theta_b) = \rho_C(x, \Theta_b) - \rho_{\text{suc}}(x)$ — разность плотности длины рассеяния липидного бислоя везикулы $\rho_C(x, \Theta_b)$ и плотности длины рассеяния раствора сахарозы $\rho_{\text{suc}}(x)$, называемая контрастом; Θ_b — набор параметров, характеризующих распределение плотности длины рассеяния внутри липидного бислоя.

В силу физических особенностей SAXS, модель плотности, предложенная в [1] и подтвердившая в [1, 2] свою эффективность для SANS-анализа структуры везикулярных систем в избытке жидкости, не применима для анализа SAXS-спектров. Поэтому, с целью выбора оптимальной формы функции плотности, мы провели исследования с использованием целого ряда моделей $\rho(x, \Theta_b)$, которые представлены на рис. 1 и упорядочены в табл. 1 по возрастанию числа варьируемых параметров.

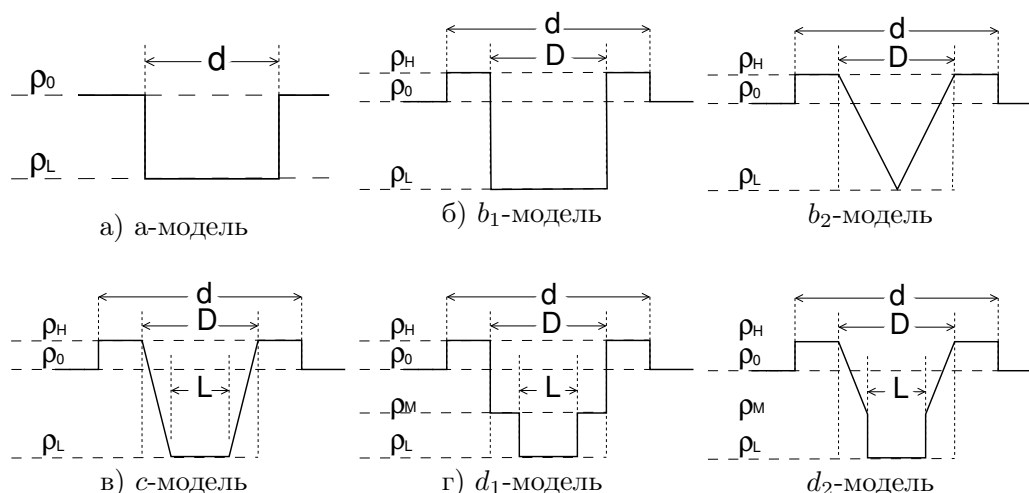


Рис. 1. Исследуемые модели функций плотности длины синхротронного рассеяния на липидном бислое $\rho(x, \Theta_b)$.

Таблица 1
Зафиксированные и варьируемые параметры моделей. Θ_b — параметры, относящиеся к внутренней структуре бислоя. Θ_s — параметры, описывающие размер, концентрацию везикул и некогерентный фон.

Модель	Зафиксированные параметры, $\times 10^{12} \text{ см}^{-1}$	Варьируемые параметры Θ	
		Θ_b	Θ_s
<i>a</i>	$\rho_0 = 10,94; \rho_{L_a} = 7,95$	d, k_{pd}	R, m, n_x, I_B
b_1, b_2	$\rho_0 = 10,94; \rho_L = 4,71;$	d, D, k_{pd}, ρ_H	R, m, n_x, I_B
<i>c</i>	$\rho_0 = 10,94; \rho_L = 4,71$	d, D, L, k_{pd}, ρ_H	R, m, n_x, I_B
d_1, d_2	$\rho_0; \rho_L; \rho_M = 7,95$	$d, D, L, k_{pd}, \rho_H,$	R, m, n_x, I_B

Полидисперсность радиуса везикул описывается распределением Шульца

$$G(\tilde{R}, R) = \frac{\tilde{R}^m}{m!} \left(\frac{m+1}{R} \right)^{m+1} \exp \left[-\frac{(m+1)\tilde{R}}{R} \right], \quad (3)$$

где R — средний радиус везикулы, m — коэффициент полидисперсности, относительное среднеквадратичное отклонение $\sigma = \sqrt{1/m + 1}$.

Тогда

$$F_s^{\text{poly}}(q, R, m) = \int_{R_{\min}}^{R_{\max}} F_s(q, \tilde{R}) G(\tilde{R}, R) d\tilde{R} / \int_{R_{\min}}^{R_{\max}} G(\tilde{R}, R) d\tilde{R}. \quad (4)$$

Полидисперсность параметров внутренней структуры бислоя учтём, используя усечённое нормальное распределение $f(x) = N(0, 1)$:

$$F_b^{\text{poly}}(q, \Theta_b) = \int_{-K}^K F_b(q, \tilde{\Theta}_b(\Theta_b, k_{pd}, x)) f(x) dx / \int_{-K}^K f(x) dx, \quad (5)$$

где компоненты $\tilde{\Theta}_b$ равны соответственно: $\tilde{d} = d + k_{pd}dx$, $\tilde{D} = D + k_{pd}dx$, $\tilde{L} = L - k_{pd}dx$; d , D и L — средние значения параметров внутренней структуры бислоя везикулы; k_{pd} — коэффициент полидисперсности параметров внутренней структуры бислоя везикулы. В расчётах $R_{\min} = 10$ нм, $R_{\max} = 100$ нм, $K = 3$.

Макроскопическое сечение полидисперсной везикулярной системы имеет вид:

$$\frac{d\Sigma}{d\Omega}(q) = I_m(q) = n_x F_s^{\text{poly}}(q, R; m) F_b^{\text{poly}}(q, \Theta_b). \quad (6)$$

С учётом разрешения спектрометра измеряемое в эксперименте значение макроскопического сечения $I(q)$ может быть записано как

$$I(q) = I_m(q) + \frac{1}{2} \Delta^2 \frac{d^2 I_m(q)}{dq^2} + I_B, \quad (7)$$

где $\Delta^2 = 3,6 \cdot 10^{-7} \text{ см}^2$ — второй момент функции разрешения спектрометра, I_B — величина некогерентного фона.

2. Численные результаты и выводы

Подгонка параметров моделей бислоя осуществлялась минимизацией среднеквадратичного относительного отклонения теоретических сечений $\hat{I}_i(q_i, \Theta)$ от соответствующих экспериментальных значений $I_i(q_i)$:

$$\chi^2 = \sum_{i=1}^n \left(\frac{I_i(q_i) - \hat{I}_i(q_i, \Theta)}{I_i(q_i)} \right)^2, \quad (8)$$

где Θ — вектор варьируемых параметров (см. столбцы 3 и 4 табл. 1).

Число варьируемых параметров в SAXS-расчётах существенно возрастает по сравнению с SANS, что приводит к снижению устойчивости итерационного процесса на основе обобщённого метода наименьших квадратов, успешно применявшегося для обработки SANS-спектров в [1, 2].

В нашей работе значения параметров Θ , обеспечивающие минимум χ^2 -функционала, определяются с помощью параллельного алгоритма Асинхронной Дифференциальной Эволюции (АДЭ) [3–5].

Результаты фитирования для разных моделей плотности длины рассеяния представлены в табл. 2 и на рис. 2. Все модели дают близкие оценки для среднего радиуса везикул R и его относительного отклонения σ . Соответственно, отличия

Таблица 2

Результаты оптимизации с целевой функцией (8) (округлены до последней статистически значащей цифры). Поскольку спектр не нормирован, параметр n_x включает в себя нормировочный фактор. Здесь $k = n - m$, число экспериментальных точек $n = 894$, число варьируемых параметров $m = 7$ для модели "a", $m = 8$ для "b" и $m = 9$ для "c" и "d". (*) — результат фитирования без учёта полидисперсности параметров внутренней структуры бислоя

Мо- дель	R, нм	σ	d, нм	D, нм	L, нм	k_{pd}	$\rho_H,$ $\times 10^{10} \text{ см}^{-1}$	$n_x,$ $\times 10^{18}$ см^{-3}	$I_B,$ a.e	χ^2 $\frac{\chi^2}{k}$
a	20,1	0,230	1,31	—	—	0,0	$\rho_L=7,95$	8,7	7,7	0,078
b ₁	19,0	0,267	3,53	2,38	—	0,18	17,0	1,6	0,0	0,00715
b ₂	18,8	0,272	3,89	3,89	—	0,18	14,4	3,6	9,6	0,0184
c	18,5	0,253	4,59	2,41	1,27	0,0	13,2	2,0	9,5	0,0180
d ₁	19,1	0,265	3,48	2,72	0,39	0,10	16,3	3,8	6,0	0,00688
d ₁ (*)	18,9	0,269	4,12	3,0	0,52	0,0(*)	12,7	3,6	10,4	0,0180
d ₂	18,9	0,268	3,97	3,97	1,41	0,0	15,1	2,1	10,5	0,0187

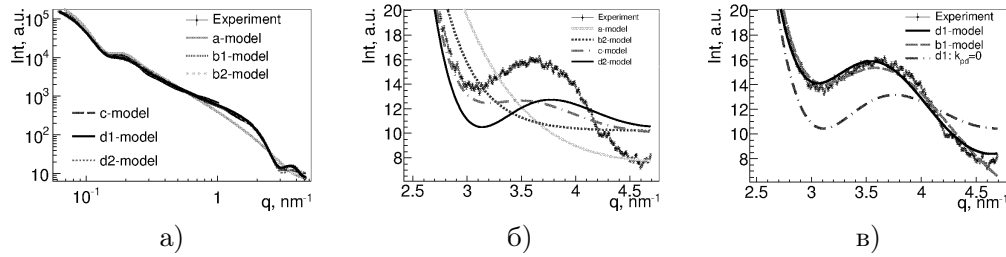


Рис. 2. а) Экспериментальный и теоретические спектры, соответствующие оптимальным параметрам модели SFM с использованием различных функций плотности $\rho_C(x, \Theta_b)$. б) $q > 2,5 \text{ нм}^{-1}$, модели a, b₂, c и d₂; в) $q > 2,5 \text{ нм}^{-1}$, модели b₁, d₁ и d₁ с $k_{pd} = 0$.

в описании спектра рассеяния при малых q несущественны (рис. 2 а)). Различия в моделях, связанные с описанием внутренней структуры бислоя оболочки везикул, сказываются в области больших значений q .

Варианты b_2 , c и d_2 (рис. 2 б)) обеспечивают качественное воспроизведение экспериментальной кривой, однако не могут описать всю амплитуду колебаний интенсивности в области больших значений q . Модели b_1 и d_1 (рис. 2 в)) имеют наилучшее согласие с экспериментальными данными.

Оптимизация модели d_1 с $k_{pd} = 0$ (штрихпунктирная кривая на рис. 2 в)) показывает, что именно введение полидисперсности внутренних параметров бислоя (5) позволило описать различия в глубине минимумов на экспериментальной кривой в области больших значений q .

Толщина бислоя и средний радиус везикул в наших расчётах получились ниже по сравнению с оценками, сделанными в [1] на базе SANS для везикул DMPC в D₂O. С другой стороны, в работе [6] показано, что значения этих параметров должны уменьшаться с ростом концентрации сахарозы в растворе. Поэтому в целом можно считать, что полученные нами параметры везикулярной системы DMPC в 40% растворе сахарозы не противоречат результатам предыдущих работ. Для уточнения оценки параметров, отвечающих за структуру везикул, необходимы систематические расчёты по SAXS- и SANS-спектрам, измеренным с разной концентрацией сахарозы.

Эффективность параллельной MPI-реализации алгоритма АДЭ тестировалась на LINUX-кластере ЛИТ ОИЯИ. Расчёт для d_1 -модели, имеющей максимальное число варьируемых параметров $m = 9$, запускался многократно с одними

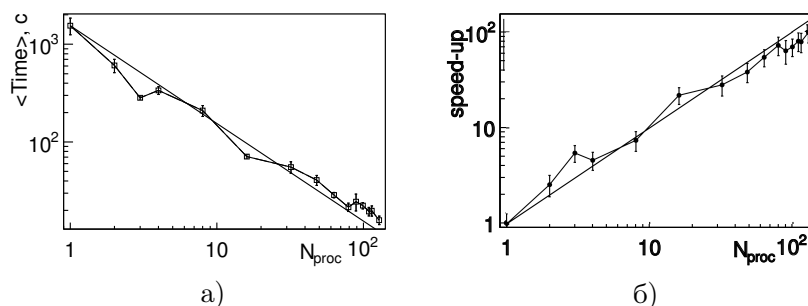


Рис. 3. Среднее время $\langle T \rangle$, затраченное на вычисления (а) и ускорение ($\text{speed-up} = \frac{T^{N_{\text{proc}}=2}}{T^{N_{\text{proc}}}}$) (б) в зависимости от числа задействованных процессоров N_{proc} .

и теми же границами начальных и допустимых значений варьируемых параметров на разном количестве вычислительных узлов: $N_{\text{proc}} \in [2; 128]$. В представленной статистике каждая комбинация запускалась $N = 10$ или 20 раз.

На рис. 3 а) видно, что среднее время, затраченное на вычисления, падает практически линейно при росте числа задействованных вычислительных узлов. Ускорение вычислений (speed-up) с учётом возрастания необходимого количества вычислений функции при $N_{\text{proc}} = 128$ составило 98 ± 21 раз.

Рис. 3 б) демонстрирует близкое к линейному ускорение вплоть до $n_{\text{proc}} = 128$ параллельно задействованных вычислительных узлов.

Таким образом, численное исследование показало, что модель SFF, модифицированная для учёта полидисперсности параметров внутренней структуры бислоя оболочки везикул, позволила воспроизвести экспериментальные данные в области больших значений q для случая полидисперсной везикулярной системы DMPC в 40% растворе сахарозы. Сравнительный анализ моделей плотности длины рассеяния выявил преимущество модели d_1 . Проведённое тестирование подтвердило эффективность параллельной компьютерной реализации подхода на базе SFF-АДЭ.

Литература

1. What Can We Learn About the Lipid Vesicle Structure from the Small-Angle Neutron Scattering Experiment? / M. A. Kiselev, E. V. Zemlyanaya, V. K. Aswal, R. H. N. Neubert // *European Biophysics Journal*. — 2006. — Vol. 35, No 6. — Pp. 477–493.
2. Structure of Unilamellar Vesicles: Numerical Analysis Based on Small-Angle Neutron Scattering Data / E. V. Zemlyanaya, M. A. Kiselev, J. Zbytovska et al. // *Crystallography Reports*. — 2006. — Vol. 51, No 1 Supplement. — Pp. S22–S26.
3. *Zhabitskaya E. I., Zhabitsky M. V.* Asynchronous Differential Evolution // *Lecture Notes in Computer Science*. — 2012. — Vol. 7125. — Pp. 328–333.
4. *Жабицкая Е. И., Жабицкий М. В.* Решение оптимизационных задач на распределенных вычислительных системах с помощью алгоритма асинхронной дифференциальной эволюции // *Математическое моделирование*. — 2012. — Т. 24, № 12. — С. 33–37. [*Zhabitskaya E. I., Zhabitsky M. V.* Parallel Solving of Optimization Problems on Distributed Systems by Asynchronous Differential Evolution // *Mat. Model.* — 2012. — Vol. 24, No 12. — Pp. 33–37. — (in russian).]
5. *Zhabitskaya E. I., Zhabitsky M. V.* Asynchronous Differential Evolution with Adaptive Correlation Matrix // *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation*. — 2013. — Pp. 455–462.

6. Sucrose Solutions as Prospective Medium to Study the Vesicle Structure: SAXS and SANS Study / M. A. Kiselev, P. Lesieur, A. M. Kisselev et al. // Journal of Alloys and Compounds. — 2001. — Vol. 328. — Pp. 71–76.

UDC 519.6 517.3

Unilamellar DMPC Vesicles Structure Analysis using Parallel Asynchronous Differential Evolution

E. I. Zhabitskaya[†], E. V. Zemlyanaya[†], M. A. Kiselev^{*}

** Joint Institute for Nuclear Research*

6, Joliot-Curie str., Dubna, Moscow region, Russia, 141980

† International university “Dubna”

Dubna, Moscow region, Russia, 141980

The Separated Form Factors model (SFF) developed previously for analysis of the small angle neutron scattering data, has been extended for numerical study of structure of polydispersed population of unilamellar DMPC vesicles using the small angle synchrotron scattering spectra (SAXS). Parameters of vesicle structure (average radius, polydispersity, bilayer thickness etc.) are fitted by means of the parallel Asynchronous Differential Evolution (ADE) method – the effective global minimization algorithm. Parallel computer implementation of our approach has been done using the MPI technique (Message Passing Interface). The numerical investigation of structure of polydispersed population of unilamellar vesicles of DMPC in the 40% water solution of sucrose has been performed. We show that accounting for the fluctuation of the bilayer thickness provides an agreement of our calculations with experimental data in the right part of SAXS spectra. On the basis of calculations with different models of internal bilayer structure, an appropriate form of the X-ray scattering length density across bilayer has been chosen. We present results of methodical calculations on the multiprocessor cluster (LIT, JINR, Dubna) demonstrating efficiency of the MPI-based parallel computer code of ADE.

Key words and phrases: vesicles, small angle X-ray scattering, Asynchronous Differential Evolution, global minimization, parallel computing.