

Научная статья

УДК 53.093, 53.096, 57.031, 57.033, 57.038

DOI: <https://doi.org/10.18721/JPM.15404>

ФУРЬЕ-СПЕКТР ИНТЕГРАЛЬНОГО ДИПОЛЬНОГО МОМЕНТА СУПЕРКОМПЬЮТЕРНЫХ МОДЕЛЕЙ РЯДА АМИНОКИСЛОТ

М. А. Баранов ✉, Э. К. Непомнящая, О. Ю. Цыбин

Санкт-Петербургский политехнический университет Петра Великого,

Санкт-Петербург, Россия

✉ baranovma1993@gmail.com

Аннотация. Реализован метод анализа молекулярных колебаний, основанный на вычислении Фурье-спектра частот суперкомпьютерных амплитудно-временных реализаций интегрального дипольного момента при суперкомпьютерном моделировании аминокислот глицина, дифенил-L-аланина и триптофана. Спектры верифицированы путем сопоставления с известными литературными данными квантовохимических вычислений, компьютерного моделирования и экспериментальной спектроскопии. Показано, что предложенные комплекс и методика позволяют эффективно вычислять достоверные спектры локальных колебаний атомной подсистемы молекулы. Полученные результаты полезны при разработке прототипов гибридных полупроводниковых микроэлектронных приборов с встроенными биомолекулярными компонентами.

Ключевые слова: аминокислота, гибридная биомолекулярная электроника, компьютерное моделирование, молекулярные колебания

Финансирование: Работа выполнена при финансовой поддержке Российского научного фонда в рамках программы «Суперкомпьютерное моделирование и технология биомолекулярных пленочных структур» (грант № 21-72-20029).

Для цитирования: Баранов М. А., Непомнящая Э. К., Цыбин О. Ю. Фурье-спектр интегрального дипольного момента суперкомпьютерных моделей ряда аминокислот // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2022. Т. 15. № 4. С. 55–68. DOI: <https://doi.org/10.18721/JPM.15404>

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

Original article

DOI: <https://doi.org/10.18721/JPM.15404>

FOURIER SPECTRUM OF THE INTEGRAL DIPOLE MOMENT OF A NUMBER OF AMINO ACIDS SUPERCOMPUTER MODELS

М. А. Baranov ✉, E. K. Nepomnyashchaya, O. Yu Tsybin

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ baranovma1993@gmail.com

Abstract. In order to analyze molecular vibrations, a method based on the calculation of the Fourier frequency spectrum of supercomputer amplitude-time realizations of the integral dipole moment in supercomputer simulation of glycine, Diphenyl-L-alanine, and tryptophan has been implemented. Under conditions of a zero external electromagnetic field, the frequency spectra of natural local vibrations of the atomic subsystem of the molecule were established to be the results of spectral analysis of these realizations. The spectra were verified by comparison with the known literature data on quantum chemical computing, computer simulation and experimental spectroscopy. It was shown that the proposed complex and technique made it possible to efficiently calculate reliable spectra of local amino acid vibrations. The results obtained may be useful for the development of prototypes of hybrid semiconductor microelectronic devices with built-in biomolecular components.

Keywords: amino acid, hybrid biomolecular electronics, computer modeling, molecular oscillation

Funding: The reported study was funded by Russian Science Foundation (grant No. 21-72-20029 “Supercomputer Simulation & Technology of Biomolecular Film Structures”)

For citation: Baranov M. A., Nepomnyashchaya E. K., Tsybin O. Yu., Fourier spectrum of the integral dipole moment of a number of amino acids supercomputer models, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 15 (4) (2022) 55–68. DOI: <https://doi.org/10.18721/JPM.15404>

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Гибридные полупроводниковые микроэлектронные устройства (ГЭУ) с встроенными молекулярными компонентами, созданные в последние годы, представляют собой новые транзисторы, фото- и светодиоды, биосенсоры, кремнийорганические гибриды и другие приборы [1 – 6]. Известны многочисленные исследования, в которых встроенные биомолекулярные компоненты (БМК) представлены ДНК, белками, пептидами и другими молекулами [5 – 8]. Установлено, что такие БМК дают уникальные функциональные возможности для микроэлектроники (в том числе в таких отраслях, как энергетика, электротехника, связь), которые не обеспечиваются какими-либо другими способами. Первостепенное значение для использования БМК в качестве электронных компонентов имеет сохранение биологической функции белков и характеристик переноса носителей заряда, получение БМК с достаточно большой площадью на поверхности интерфейса, иммобилизация и структурные свойства как одиночных молекул, так и их кластеров, кристаллов, пленочных и объемных образований (см., например, обзор [9]). Однако, несмотря на значительные усилия исследователей, в этой, несомненно перспективной, области большая часть актуальных задач далека не только от решения, но даже от их обоснованной постановки. Осталось много принципиальных вопросов, не решенных в гибридной электронике и тем самым сдерживающих ее развитие. С позиций электродинамики и электроники биомолекулы предстают в основном не как носители и исполнители биологических функций, а как физические осцилляторы с особыми свойствами локализации динамических состояний в пространственной структуре и в амплитудно-частотном континууме.

Имеются два конструктивных подхода к реализации таких осцилляторов: один из них развит на основе белков, другой предполагает использование аминокислот (АК); при этом для обоих характерны проблемы и возможности, существенно разные для ГЭУ. Если принимать во внимание большее количество атомов, повышение разнообразия и значения вторичной и более высоких структур, а также возможное наличие иных (кроме аминокислотных остатков) включений, по сравнению с простыми физическими осцилляторами, то белковые БМК оказываются гораздо сложнее. Они имеют множество конформационных и структурных различий, а воспроизводятся с большой точностью только статистически. Аминокислоты биомолекулярных компонентов выглядят в целом не только более простыми, но и более универсальными, причем характеризуются высокой повторяемостью структур.

Двадцать две природные аминокислоты, из которых строятся все белки в живых клетках, охарактеризованы детальным физическим описанием, экологически безопасны, обеспечивают возможность их получения в чистом виде, в условиях относительно простого производства и с невысокой стоимостью. В первичных цепях структуры биомолекул, пептидные группы в виде аминокислотных остатков служат основными «строительными блоками», однако собственно этим блокам в исследованиях ГЭУ уделено пока значительно меньшее внимание, чем построенным из них белковым молекулам. В перспективных разработках элементной базы электроники интегральный электронный дипольный момент (ЭДМ) АК может служить параметром, информативным и удобным для выявления конформационной динамики в адсорбированных состояниях и в растворах, для

электрических потенциалов и температурных зависимостей, энергии взаимодействия с поверхностью твердого тела и т. п. Например, с помощью компьютерного моделирования олигопептидов аланина (Ala 2-Ala 20), различных по размеру, в водной среде и вакууме, было получено большое количество данных о наиболее вероятных конформационных состояниях молекулы и значениях ЭДМ при определенных (варьируемых) внешних условиях [10]. Закономерности изменения ЭДМ, выявленные в работе [10] одним из авторов данной статьи, образовали физически обоснованный, внутренне непротиворечивый комплекс, сопоставимый с литературными данными.

Наиболее точными методами получения значений частот колебаний изолированных свободных белковых молекул считаются квантовомеханические, например метод теории функционала плотности (ТФП), реализованный в современных комплексах программ GAUSSIAN, имеющие, однако, высокий уровень сложности. Кроме того, при таком моделировании колебательных спектров обычно требуется использовать эмпирические поправочные числовые факторы для корректировки вычисленных частот нормальных колебаний и интенсивностей колебательных переходов.

Для разработок же гибридных полупроводниковых микроэлектронных устройств необходимы более эффективные методы моделирования неравновесной динамики связанных молекул, находящихся в адсорбированном состоянии или в растворах в неравновесных состояниях. Получение амплитудно-временных реализаций ЭДМ в работе [10] с помощью суперкомпьютера имело большую временную продолжительность, которую, однако, можно считать допустимой для регистрации достаточно точного частотного спектра с использованием Фурье-анализа. Действительно, высокое разрешение линий в Фурье-спектре достигается в том случае, когда на протяжении записи амплитудно-временной реализации укладывается достаточно большое (порядка $10^5 - 10^6$ и выше) количество периодов колебаний.

С учетом представленной информации, в данной работе проведена апробация предложенного метода получения достаточно длительных амплитудно-временных реализаций интегрального ЭДМ и последующего их Фурье-анализа.

Метод моделирования молекулярной динамики

В данном исследовании моделирование молекулярной динамики было направлено на вычисление амплитудно-временных реализаций интегрального ЭДМ и исходило из решения уравнений Ньютона для каждого атома молекулярной системы:

$$m_a \ddot{\mathbf{r}}_a = - \frac{\partial}{\partial \mathbf{r}_a} U_{total}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N), \quad a = 1, 2, \dots, N, \quad (1)$$

где m_a – масса атома, $\ddot{\mathbf{r}}_a$ – ускорение, U_{total} – интеграл по всем компонентам потенциальной энергии системы, зависящий от взаимного расположения всех атомов и вычисленный в точке нахождения данного атома.

Получаемые траектории движения атомов в силовом поле эмпирического атом-атомного потенциала создавали детальную микроскопическую картину внутренней тепловой подвижности макромолекулы. Уравнения движения (1) в программе вычислений по временным шагам представляли в конечно-разностном виде. Задавая координаты и скорости всех частиц в начальный момент времени, вычисляли на каждом временном шаге новые значения действующих сил, координат и скоростей частиц. Температуру определяли как среднюю кинетическую энергию, приходящуюся на одну степень свободы системы:

$$T(t) = \frac{1}{3Nk_B} \sum_{i=1}^n m_i \mathbf{v}_i^2, \quad \mathbf{v}_i = \frac{d\mathbf{r}_i}{dt}, \quad (2)$$

где N – полное число степеней свободы молекулы, $N = 3n - 6$ (n – полное число атомов); k_B – постоянная Больцмана.

В изолированной системе сохраняются полный импульс и момент импульса; сохраняется и полная энергия адиабатически изолированной системы, а температуру получают усреднением ее мгновенных значений $T(t)$ (см. уравнение (2)) по заданному интервалу времени.

Потенциальную энергию молекулы задавали в следующем виде:

$$U_{total} = \sum_{bonds} k_b (b - b_0)^2 + \sum_{angle} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi [1 + \cos(n\phi - \delta)] + \sum_{impropers} k_\omega (\omega - \omega_0)^2 + \sum_{LJ} (-E_{min}) \cdot \left[\left(\frac{R_{min_{ij}}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min_{ij}}}{r_{ij}} \right)^6 \right] + \epsilon_{14} \frac{q_i q_j}{\epsilon_0 r_{ij}}, \quad (3)$$

где b , b_0 – коэффициенты растяжения молекулярных связей; θ , θ_0 – углы относительно плоскости; ϕ , δ – углы скручивания; ω , ω_0 – неправильные двугранные углы; R_{min} – расстояние между i -ым и j -ым атомами, на котором энергия взаимодействия становится равной нулю; параметр E_{min} – минимум потенциальной энергии взаимодействия Леннарда – Джонса, $E_{min} = U_{LJ}(R_{min})$ (U_{LJ} – потенциал Леннарда – Джонса); ϵ_{14} – параметр, который представляет собой безразмерный масштабный коэффициент, $\epsilon_{14} = 1$, за исключением модифицированного взаимодействия 1 – 4, когда пара атомов разделена последовательностью из трех ковалентных связей (так что атомы также могут участвовать во взаимодействии под углом кручения); r_{ij} – расстояния между i -ым и j -ым атомами; q – заряд атома; ϵ_0 – диэлектрическая постоянная; k_b , k_θ , k_ϕ , k_ω – коэффициенты взаимодействия.

Представленные в формуле (3) атомные взаимодействия включают растяжение связей и их изгибание под двугранным углом (proper and improper dihedrals).

Вероятность реализации ω_τ данного микроскопического состояния с энергией E_τ зависела только от значения энергии, и задавалась распределением Гиббса:

$$\omega_\tau = \frac{1}{Z} e^{-E_\tau/k_B T}, \quad (4)$$

где τ – порядковый номер реализации.

Постоянную нормировки Z выбирали при условии, в котором сумма вероятностей равна 1; она приобретала вид

$$Z = \sum_{\tau} e^{-E_\tau/k_B T}.$$

Мгновенные значения ЭДМ на каждом временном шаге рассчитывали путем суммирования парциальных моментов:

$$\boldsymbol{\mu} = \sum_{n=1}^N q_n \mathbf{r}_n. \quad (5)$$

Отметим, что в уравнении (1) производится суммирование по всем зарядам в молекуле, с учетом их индивидуальных радиусов-векторов. Последовательность значений $\boldsymbol{\mu}(t)$ в выражении (5) составляла амплитудно-временную реализацию мгновенных значений интегрального ЭДМ молекулы.

Выполнение программы вычислений для систем с большим числом атомов оказалось сложной и ресурсоемкой задачей. Поэтому для ее решения был разработан специальный программно-вычислительный комплекс (ПВК) для суперкомпьютерных расчетов, в состав которого вошли пакеты прикладных программ Avogadro, Visual Molecular Dynamics и NAMD, а также дополнительные оригинальные программы, написанные на языке Python.

Суперкомпьютерный программно-вычислительный комплекс

Для выполнения ПВК был использован кластер «РСК-Торнадо» суперкомпьютерного центра «Политехнический» Санкт-Петербургского Политехнического университета Петра Великого. Этот кластер с пиковой производительностью 1015 терафлопс содержит 668 двухпроцессорных узлов (Intel Xeon E5 2697 v3), из них 56 узлов имеют два ускорителя вычислений NVIDIA K40 [11].

На рис. 1 представлена блок-схема разработанного и адаптированного к суперкомпьютерному кластеру «РСК-Торнадо» алгоритма ПВК в плане подготовки и выполнения моделирования молекулярной динамики, а также спектральной обработки данных и представления результатов.

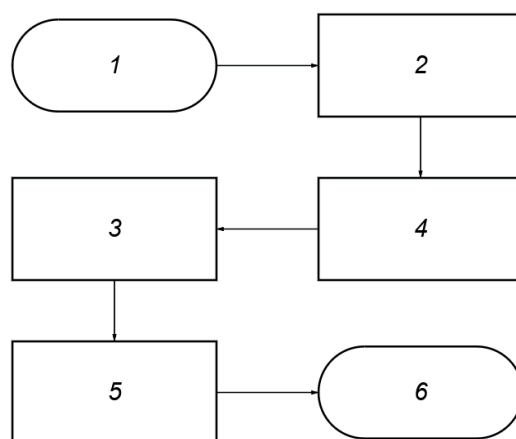


Рис. 1. Блок-схема алгоритма подготовки и выполнения моделирования молекулярной динамики, а также анализа полученных расчетных данных:

- 1 – определение параметров эксперимента, 2 – задание функций параметров силового поля,
 3 – определение координат атомов молекул аминокислот, 4 – учет параметров окружения молекул аминокислот, 5 – проведение частотного анализа дипольных моментов молекул аминокислот,
 6 – представление результатов

В рамках разработанного алгоритма для запуска этапа вычислений требовалось задать исходные параметры силового поля, определить координаты атомов в конкретной молекулярной модели, задать уравнения для учета параметров окружения молекул и представления параметров физико-математической модели колебательной системы.

Для задания моделей молекулярных систем были использованы лицензированные программы Avogadro и Visual Molecular Dynamics (VMD). Программа Avogadro позволяет создавать файлы, содержащие координаты атомов различных биологических молекул, в том числе аминокислот, а также выполнять предварительную оптимизацию геометрии молекулярной структуры. С помощью Avogadro были созданы первичные файлы с координатами атомов молекул глицина (GLY), дифенил-L-аланина (FF) и триптофана (TRP). Программа VMD позволяла создавать структурные файлы на основе топологии молекулярных связей, добавлять молекулы воды и ионы солей, визуализировать молекулярные системы, выполнять вычисления энергии и дипольных моментов конформационно-подвижных молекул.

На рис. 2 схематично показаны использованные в расчетах пространственно-структурные формулы молекул глицина, дифенил-L-аланина и триптофана.

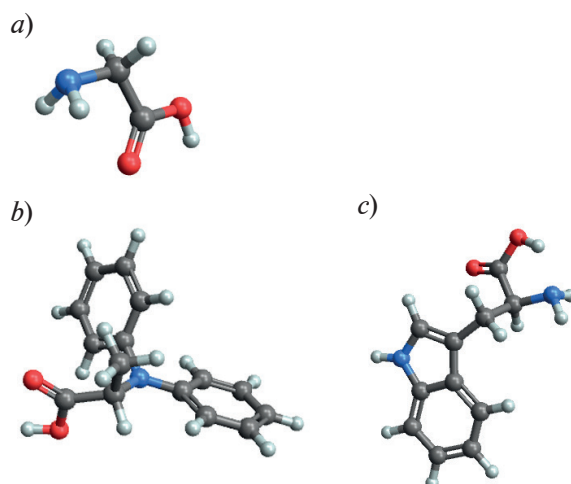


Рис. 2. Структурные формулы молекул глицина (a), дифенил-L-аланина (b) и триптофана (c). Атомы водорода, углерода, кислорода и азота обозначены серым, черным, красным и голубым цветами, соответственно

Выбор глицина для исследования обусловлен тем, что он является простейшей органической алифатической аминокислотой; это позволяет его использование в качестве модельного объекта для изучения и идентификации самых распространенных типов колебаний в АК. Дифенил-L-аланин и триптофан относятся к классу ароматических альфа-аминокислот. Дифенил-L-аланин по химическому строению можно представить как аминокислоту аланин, в которой два атома водорода замещены фенильной группой. Триптофан содержит ароматическое ядро индола, за счет чего является гидрофобным. Структуру триптофана можно представить в виде двух радикальных остатков: индола (ароматическое гетероциклическое соединение, образованное двумя сочлененными кольцами: бензольным и пиррольным) и аланина, соединенных через углеродную связь C–C. В основе обеих молекул находится аланин, что позволяет сравнивать частоты колебаний и при этом относить различия в колебательных спектрах к ароматической группе.

Модель молекулы каждой АК из экспериментального набора помещали в виртуальную ячейку в виде водного куба. Размер такого куба варьировался от 20 до 30 Å, в зависимости от размеров и положения биомолекул, а ячейки составляли периодическую последовательность. Для моделирования растворения молекул была использована модель воды TIP3W, соответствующая трем атомам молекулы воды в трех точках взаимодействия с молекулой. Подобные модели обеспечивают высокую вычислительную эффективность и применяются во многих приложениях молекулярной динамики. Такая модель воды была реализована в силовом поле программы CHARMM, в которой параметры Леннарда – Джонса присваиваются атомам водорода в дополнение к параметрам кислорода в молекуле воды.

Для создания моделей, близких к реальным системам, в периодические водные ячейки, содержащие молекулы АК, были добавлены ионы соли NaCl с концентрацией 0,15 моль/л. При этом максимальное количество ионов на одну молекулу в водном кубе составляло около 30. Подобные методы обоснованы в экспериментах, описанных в работах [12 – 15].

Молекулярное моделирование проводили при условии наложения силового поля CHARMM27 с использованием пакета NAMD v.2.14 (Nanoscale Molecular Dynamics) в последовательности четырех этапов:

- минимизация энергии,
- нагрев системы,
- установление равновесных величин кинетической и интегральной потенциальной энергии,
- моделирование молекулярной динамики.

В пакете NAMD использовали модель параллельного программирования Charm++, обладающую высокой эффективностью распараллеливания для использования суперкомпьютерных вычислений. Ставили периодические граничные условия, задавали значения радиуса отсечки (для несвязанных взаимодействий с функцией переключения), начиная с 9 Å и до 12 Å (подробные параметры моделирования представлены в табл. 1). Для моделирования момента установления равновесия в системе ставили условия минимизации энергии, медленного нагрева системы (0,01 К за 1 фс) и стабилизации конформационного состояния молекул. В результате минимизации энергии получали стабильные состояния молекул, необходимые для дальнейшего вычисления интегрального ЭДМ молекул аминокислот, находящихся в вакууме или в растворе.

В рамках данной работы вычисляли ЭДМ одиночной молекулы аминокислоты в водном растворе, стабилизированном ионами натрия.

Для запуска этапа расчетов использовали специальный скрипт, содержащий указания для начала вычислений. Такой скрипт (вместе с необходимыми конфигурационными файлами для моделирования молекулярной динамики) отправляли в локальную директорию кластера. Отправка всех необходимых файлов происходила с использованием пакета SSHFS (Secure SHell File System) для систем LINUX. Набор необходимых файлов включал скрипт для запуска вычислений, файл .pdb, содержащий координаты всех атомов молекулярной системы, файл .psf, содержащий описание молекулярных связей, а также четыре файла namd конфигурации моделирования для минимизации и стабилизации энергии.

Таблица 1

Параметры, использованные при моделировании

Параметр	Значение
Интервал записи фреймов, фс	1–20
Время реализации, пс	100–2000
Температура, К	300
Давление, бар	1,01325
Невалентные взаимодействия, Å	18
Список Верде, Å	20
Частота обновления списка, ПГц	1
Количество фреймов, тыс.	100
Вес dcd-файла координат, Гб	0,015–20,0

Примечание. Использовали ансамбль частиц NPT, тип силового поля CHARMM и явный водный тип растворителя.

Машинное время вычислений составляло от единиц до сотен часов, в зависимости от количества атомов и уровня выделяемой мощности. В качестве выходных данных анализировали координаты атомов молекулярной системы в каждый момент времени реализации. Размер файла координат для каждого записанного фрейма варьировался от 14 Мб для вакуумных систем до 20 Гб для больших систем с растворителем. Анализ частот колебаний выполняли при помощи пакета VMD IR Spectral Density Calculator. Этот пакет позволял анализировать координаты атомов из файла скоростей .dcd и строить спектральную диаграмму изменения дипольного момента выбранных атомов. Для вычисления распределения плотности энергии колебаний по спектру частот использовали амплитудно-временные реализации мгновенных значений интегрального ЭДМ каждой молекулы. Для этого в программу загружали структурный файл и файл с наборами координат атомов молекулярной системы. Задавали выборку по атомам биомолекул, исключая ионы солей и молекулы воды, температуру, при которой выполнялось моделирование, шаг записи координат, соответствующий шагу в конфигурационном файле моделирования, и верхнюю граничную частоту вычислений в интервале 6000 см^{-1} . В результате расчета получали файл со значениями частот и спектральной плотности колебаний. Колебания молекул аминокислот визуально непрерывно контролировали в программе VMD.

Для дальнейшего анализа полученных расчетных данных использовали оригинальный разработанный алгоритм на языке Python, совмещенный с программным пакетом Big Data – Pandas. Этот алгоритм позволил обрабатывать файлы .csv и формировать набор данных, удобный для проведения различных операций над строками и столбцами таблицы.

Частотный спектр амплитудно-временных характеристик интегрального дипольного момента исследованных аминокислот

Типичные зависимости мгновенных значений полного интегрального ЭДМ и его проекций на оси координат от времени, показывающие пространственно-угловую архитектуру молекулы в ее динамике, представлены на рис. 3.

Проекции ЭДМ на оси координат dip_x , dip_y , dip_z флуктуируют по амплитуде относительно медленно, и они скоррелированы с течением времени; модуль же интегрального ЭДМ остается примерно постоянным в области относительно низких частот. Скорее всего, это указывает на наличие относительно медленных нелокальных (в том числе вращательных) движений молекулы приблизительно как целого, при длительном сохранении конформационной структуры. Амплитудно-временные реализации полного интегрального ЭДМ (см. рис. 3) содержат колебания на существенно более высоких частотах. Соот-

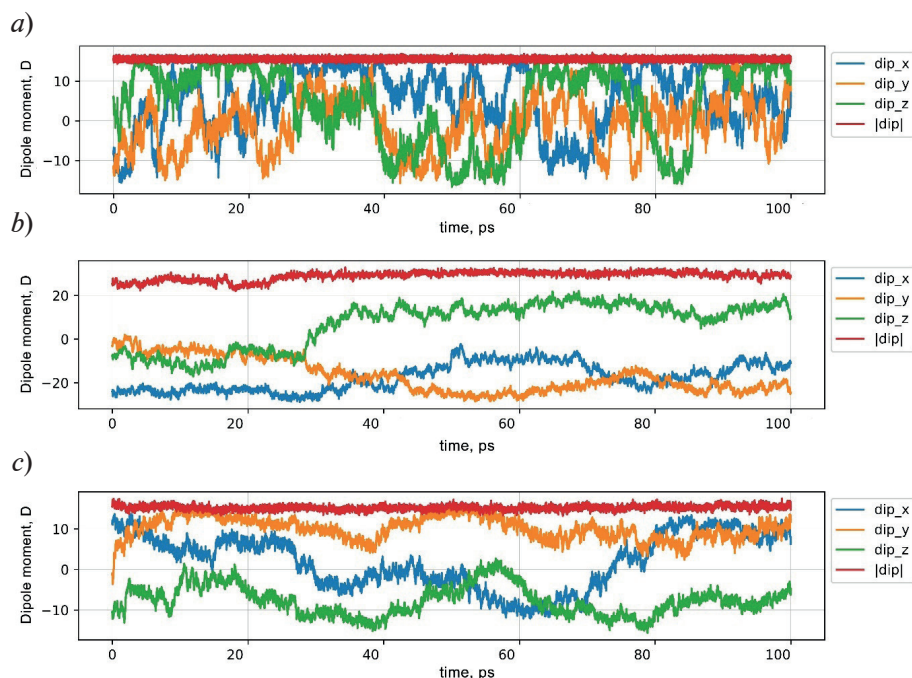


Рис. 3. Интегральные ЭДМ молекул АК глицина (a), дифенил-L-аланина (b) и триптофана (c), а также соответствующие проекции вектора ЭДМ на оси координат как функции времени (температура 300 К)

ветствующие частотные спектры представлены на рис. 4. Они имели квазидискретный характер и стабильную структуру пиков, что позволило выявить преимущественные частоты и дать соответствующую интерпретацию их связи с локальными внутримолекулярными колебаниями в атомной подсистеме.

Полученные результаты моделирования представлены в табл. 2 в соответствии с нормальными модами колебаний и данными, известными из публикаций в научно-технической литературе. Типичная пространственная локализация некоторых колебаний, соответствующая разным частотам в спектре, показана для молекулы триптофана на рис. 5.

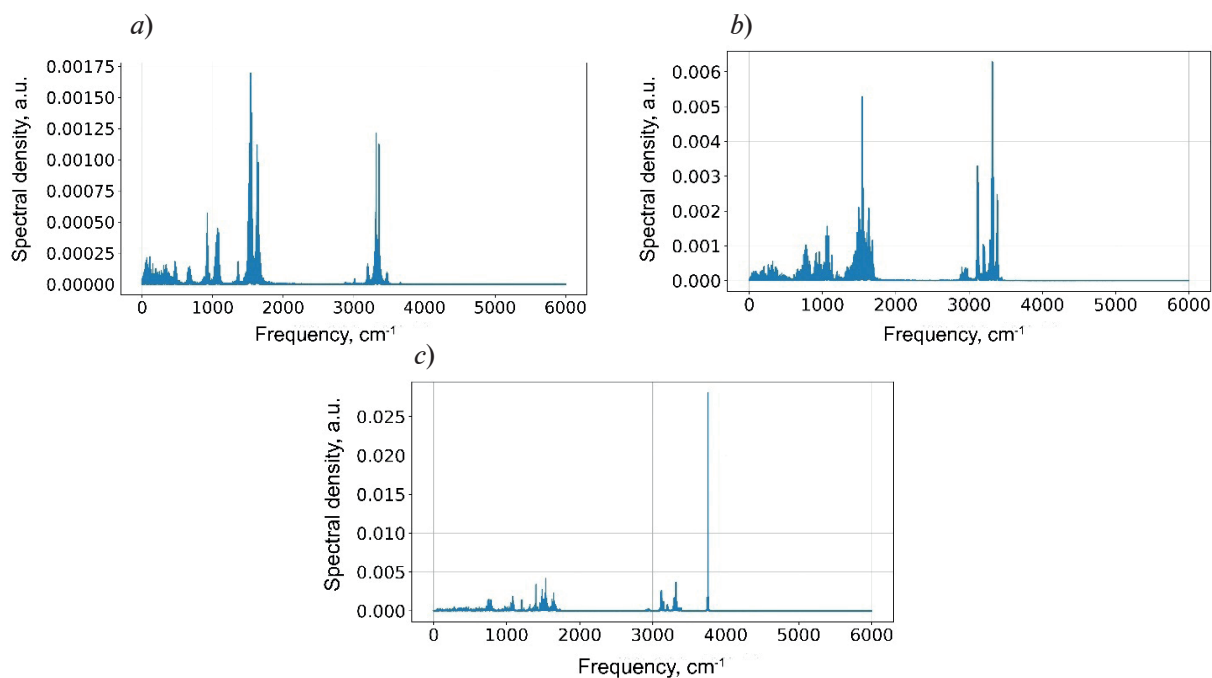


Рис. 4. Фурье-спектры амплитудно-временных реализаций полного интегрального ЭДМ АК глицина (a), дифенил-L-аланина (b) и триптофана (c)

Таблица 2

**Сопоставление вычисленных спектральных компонент
трех аминокислот с литературными данными**

Связь и тип колебания	Частота колебания связи, см ⁻¹					
	Результат моделирования			Литературные данные и источник		
	TRP	FF	GLY	TRP	FF	GLY
Водородная	38–200	57–201	43–210	40–215 [16]	59–216 [16]	22–233 [16]
СС ^α N, деформационное	240–242	278	–	238–269 [17]	258,278 [17]	–
СО ₂ , крутильное	–	260	–	–	258,261 [18]	–
NH ₃ ⁺	337–341	301–334	322–357	325–347 [17]	293,324 [18]	356 [18]
COO ⁻ , изгибное	582–597	479, 542–564 640–659	459, 543–563	499–655 [16]	535–647 [16,19]	504–607 [16]
TRP δ _{о.о.р.} , кольцо {ν ₄ }	683–687	–	–	660–692 [20]	–	–
νCC, Изменение кольца {ν ₁ }	745–766	–	–	716–795	–	–
δCH ₂ + t пиррол	844–874	–	–	790–828 [20]	–	–
δCOO ⁻ , ωCOO νCC	–	–	677	–	–	679 [20]
OCO ⁻	–	748–770	–	–	778 [20]	–
δCOO ⁻ , νCC, νCN	–	846	–	–	837–859 [19]	–
νCC, δCOO ⁻	–	–	855	–	–	882–916 [20]
νCC, νCN, r CH ₃	–	892, 949–955	–	–	913–931 [19, 20]	–
C–C	–	–	922–928	–	–	930, 988 [18, 20]
C–N	1036–1053 1060–1098	1039–1062 1134	1009–1061 1158	1018–1137 [18]		
C–O/OH	1202–1208	1200–1225	1229	1200–1300 [17]		
C–N (амидные III)	1321	1338	–	1200–1340 [17, 21]		
CH ₂ /COO ⁻	1401, 1486	1494	1362	1343/1330 [17]		
N–H (амидные II)	1536	1541–1548	1526–1552	1530–1580 [17, 21]		
Ножичные колебания NH ₂ /NH ₃	1620–1647	1606–1642	1629–1646	1604–1660 [17]		
Валентные C–H	2948	2887–2979	3013	2810–2900 [23]		
Асимметричные NH (амидные A)	3117–3122, 3205–3227, 3295–3332	3116, 3192–3206, 3318, 3384	3199, 3314, 3356, 3472	3300 [23]		
Растяжение O–H	3747–3765	–	3670	3700 [23]		

Обозначения: TRP, FF, GLY – молекулы триптофана, дифенил-L-аланина и глицина.

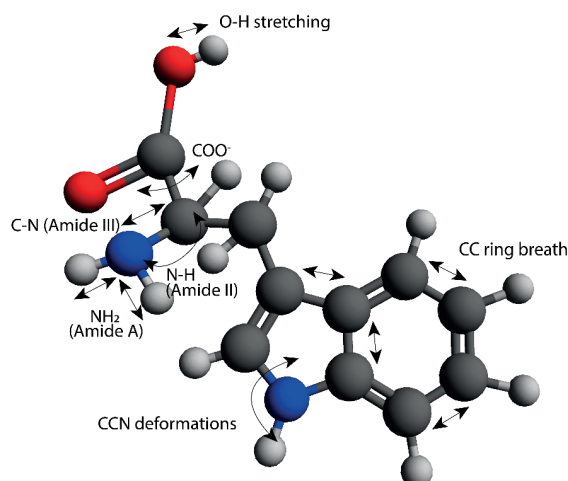


Рис. 5. Схематические изображения некоторых локальных колебательных движений в молекуле триптофана

Обсуждение

Значения частот собственных локальных колебаний, вычисленные с использованием амплитудно-временных реализаций мгновенного полного интегрального ЭДМ, лежат в области дальнего инфракрасного диапазона электромагнитных волн. Весь набор почти полностью совпадает с известными литературными результатами (как статистически, так и в деталях), которые были получены методами экспериментальной спектроскопии и вычислительными способами в широком диапазоне частот.

В табл. 2 можно выделить детальные соответствия полученных нами расчетных значений частот с экспериментальными. Например, для АК глицина наибольшее количество совпадений приходится на высокочастотную область колебаний (от 988 см^{-1}), где преобладают колебания аниона глицина, а также валентные колебания углеродной связи C–C скелета молекулы. В диапазоне $1083\text{--}1154\text{ см}^{-1}$ основной вклад вносит колебание связи C–N. Совпадения данных наблюдаются по всем трем изученным АК. Практически полное соответствие для трех АК наблюдается также в области $1200\text{--}1300\text{ см}^{-1}$, где присутствуют две спектральные полосы колебаний. Это валентные колебания C–O и (менее интенсивная полоса) деформационные колебания группы OH. Кроме того, данная частотная область может включать в себя «веерные» колебания группы CH_2^- . Согласно литературным данным, на границе этой частотной области ($\sim 1337\text{ см}^{-1}$) могут присутствовать колебания связи $\text{C}_\alpha\text{--H}$, однако нашими расчетами не выявлено какого-либо пика спектральной плотности на указанной частоте.

Наибольшие совпадения по частотам пришлись на такие связи как N–H (амидные II), ножничные колебания связей NH_2/NH_3 , ассиметричный NH (амидные A). Каких-либо «избыточных» колебаний, не находящих соответствия среди известных мод, в вычисленных частотных спектрах не было обнаружено. Например, согласно золотому правилу « $(3n - 6)$ », молекула глицина должна иметь набор из 24 собственных частот. В табл. 2 можно видеть полученные нами значения, близко отвечающие этому требованию. Такое же соответствие получено для аминокислот FF и TRP.

Если оценивать в целом полученные результаты, то достигнутый в нашем исследовании большой и достаточно полный комплекс совпадений и соответствий можно считать надежным аргументом эффективности разработанной методики моделирования и вычислений.

Отметим, что имеются также некоторые расхождения вычисленных значений с известными частотами локальных мод, что может быть связано с такими факторами, как исходное состояние молекул и их окружение, в том числе наличие молекул растворителя или ионов солей. Присутствие молекул воды с растворенными в ней ионами солей может сдвигать значение частот за счет спонтанного образования атомных комплексов и электрического поля ионной среды.

Заключение

В настоящем исследовании разработан суперкомпьютерный программно-вычислительный комплекс, в состав которого вошли пакеты прикладных программ Avogadro, Visual Molecular Dynamics и NAMD, а также дополнительные оригинальные программы, написанные на языке Python. Комплекс реализован на суперкомпьютерном кластере «РСК-Торнадо» в СПбПУ Петра Великого и использован для получения длительных амплитудно-временных реализаций интегрального ЭДМ продолжительностью до 100 пс. В результате получены частотные спектры АК глицина, дифенил-L-аланина и триптофана в вакууме и в водно-солевом окружении в условиях нулевого внешнего электромагнитного поля.

Аминокислоты глицина, дифенил-L-аланина и триптофана в вакууме и в водно-солевом окружении, наряду с полиаланинами, могут представлять первостепенный интерес для исследований, благодаря их особым структурным и функциональным свойствам, а также возможности сопоставления получаемых результатов с известными результатами исследований.

Проведена верификация методики и полученных результатов путем сравнительного анализа и сопоставления с известными характеристиками, представленными в научнотехнической литературе, в том числе с квантовохимическими расчетами методом теории функционала плотности. Методом Фурье-анализа этих реализаций получены частотные спектры собственных локальных колебаний трех выбранных молекул АК: глицина, дифенил-L-аланина и триптофана. Проведена верификация метода, программного комплекса и полученных результатов путем сравнительного сопоставления частотных спектров с известными литературными данными. На основе сопоставительного анализа сделан вывод, что предложенный и апробированный метод позволяет эффективно вычислять достоверные спектры локальных колебаний АК. По сравнению с квантовомеханическими подходами, данный метод отличается более простыми вычислениями динамики связанных молекул в неравновесных состояниях в адсорбированном состоянии и в растворе. Полученные данные представляются полезными для разработки прототипов гибридных полупроводниковых микроэлектронных приборов с встроенными биомолекулярными компонентами. Разработанный программный комплекс можно использовать для суперкомпьютерного вычисления более сложных и представительных молекулярных сценариев. В частности, этот подход предполагается применить для моделирования динамики молекул, находящихся во внешнем высокочастотном электромагнитном поле.

Результаты работы были получены с использованием вычислительных ресурсов суперкомпьютерного центра Санкт-Петербургского политехнического университета Петра Великого (www.scc.spbstu.ru).

СПИСОК ЛИТЕРАТУРЫ

1. **Dimitrakopoulos C. D., Malenfant P. R. L.** Organic thin film transistors for large area electronics // *Advanced Materials*. 2002. Vol. 14. No. 2. Pp. 99–117.
2. **Heni W., Kutuvantavida Y., Haffner C., et al.** Silicon-organic and plasmonic-organic hybrid photonics // *ACS Photonics*. 2017. Vol. 4. No. 7. Pp. 1576–1590.
3. **Frommer W. B., Davidson M. W., Campbell R. E.** Genetically encoded biosensors based on engineered fluorescent proteins // *Chemical Society Reviews*. 2009. Vol. 38. No. 10. Pp. 2833–2841.
4. **Rodriguez-Mozaz S., De Alda L. M. J., Barcelo D.** Biosensors as useful tools for environmental analysis and monitoring // *Analytical and Bioanalytical Chemistry*. 2006. Vol. 386. No. 4. Pp. 1025–1041.
5. **Lee M. Y., Lee H. R., Park C. H., Han S. G., Oh J. H.** Organic transistor-based chemical sensors for wearable bioelectronics // *Accounts of Chemical Research*. 2018. Vol. 51. No. 11. Pp. 2829–2838.
6. **Ruiz M. P., Aragonès A. C., Camarero N., Vilhena J. G., Ortega M., Zotti L. A., Pérez R., Cuevas J. C., Gorostiza P., Díez-Pérez I.** Bioengineering a single-protein junction // *Journal of the American Chemical Society*. 2017. Vol. 139. No. 43. Pp. 15337–15346.
7. **Amdursky N., Głowacki E.D., Meredith P.** Macroscale biomolecular electronics and ionics // *Advanced Materials*. 2019. Vol. 31. No. 3. P. 1802221.

8. **Torculas M., Medina J., Xue W., Hu X.** Protein-based bioelectronics // *ACS Biomaterials Science & Engineering*. 2016. Vol. 2. No. 8. Pp. 1211–1223.
9. **Bostick C. D., Mukhopadhyay S., Pecht I., Sheves M., Cahen D., Lederman D.** Protein bioelectronics: A review of what we do and do not know // *Reports on Progress in Physics*. 2018. Vol. 81. No. 2. P. 026601.
10. **Зезина Т. И., Цыбин О. Ю.** Субпикосекундная динамика дипольного момента молекулярных полиаланинов // *Научно-технические ведомости СПбГПУ. Физико-математические науки*. 2017. Т. 10. № 4. С. 100–110.
11. Политехник – РСК Торнадо. Технические характеристики. Режим доступа: <https://rscgroup.ru/project/spbstu-politechnic/> (Дата обращения: 15. 07. 2022).
12. **Баранов М. А., Цыбин О. Ю., Величко Е. Н.** Компьютерный анализ релаксации свободной энергии молекул глицина, триптофана и альбумина в ионизованном водном растворе // *Научно-технические ведомости СПбГПУ. Физико-математические науки*. 2021. Т. 14. № 4. С. 135–146.
13. **Баранов М. А., Цыбин О. Ю., Величко Е. Н.** Структурированные биомолекулярные пленки для микроэлектроники // *Научно-технические ведомости СПбГПУ. Физико-математические науки*. 2021. Т. 14. № 1. С. 85–99.
14. **Baranov M., Nepomyashchaya E., Velichko E.** Computer simulation of biomolecules around metallic nanoparticle for biomolecular electronics // *Proceedings of the 2021 International Conference on Electrical Engineering and Photonics (EExPolytech 2021)*. October 14–15, 2021. St. Petersburg: SPbPU. Pp. 171–174.
15. **Velichko E., Zezina T., Baranov M., Nepomyashchaya E., Tsybin O.** Dynamics of polypeptide cluster dipole moment for nano communication applications // *Lecture Notes in Computer Science (including book subseries “Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics”)*. Vol. 11118. Berlin: Springer Science + Business Media, 2018. Pp. 675–682.
16. **Matei A., Drichko N., Compf B., Dressel M.** Far-infrared spectra of amino acids // *Chemical Physics*. 2005. Vol. 316. No. 1–3. Pp. 61–71.
17. **Barth A., Zscherp C.** What vibrations tell us about proteins // *Quarterly Reviews of Biophysics*. 2002. Vol. 35. No. 4. Pp. 369–430.
18. **Бутырская Е. В., Нечаева Л. С., Шапошник В. А., Дроздова Е. И.** Отнесение полос в ИК спектрах водных растворов глицина на основе квантово-химического расчета // *Сорбционные и хроматографические процессы*. 2012. Т. 12. № 4. С. 501–512.
19. **Tortonda F. R., Pascual-Ahuir J.-L., Silla E., Tucyn I., Ramírez F. J.** Aminoacid zwitterions in solution: Geometric, energetic, and vibrational analysis using density functional theory-continuum model calculations // *The Journal of Chemical Physics*. 1998. Vol. 109. No. 2. Pp. 592–602.
20. **Wolpert M., Hellwig P.** Infrared spectra and molar absorption coefficients of the 20 alpha amino acids in aqueous solutions in the spectral range from 1800 to 500 cm^{-1} // *Spectrochimica Acta. Part A: Molecular and Biomolecular Spectroscopy*. 2006. Vol. 64. No. 4. Pp. 987–1001.
21. **Itoh K., Ikeda A., Iwamoto T., Nishizawa S.** DFT calculation analysis of terahertz time-domain spectra of polyalanines // *Journal of Molecular Structure*. 2011. Vol. 1006. No. 1–3. Pp. 52–58.
22. **Barth A.** The infrared absorption of amino acid side chains // *Progress in Biophysics and Molecular Biology*. 2000. Vol. 74. No. 3–5. Pp. 141–173.
23. **Mohamed M. E., Mohammed A. M. A.** Experimental and computational vibration study of amino acids // *International Letters of Chemistry, Physics and Astronomy*. 2013. Vol. 15. Pp. 1–17.

REFERENCES

1. **Dimitrakopoulos C. D., Malenfant P. R. L.,** Organic thin film transistors for large area electronics, *Adv. Mater.* 14 (2) (2002) 99–117.
2. **Heni W., Kutuvantavida Y., Haffner C., et al.,** Silicon-organic and plasmonic-organic hybrid photonics, *ACS Photonics*. 4 (7) (2017) 1576–1590.
3. **Frommer W. B., Davidson M. W., Campbell R. E.,** Genetically encoded biosensors based on engineered fluorescent proteins, *Chem. Soc. Rev.* 38 (10) (2009) 2833–2841.
4. **Rodriguez-Mozaz S., De Alda L. M. J., Barcey D.,** Biosensors as useful tools for environmental analysis and monitoring, *Anal. Bioanal. Chem.* 386 (4) (2006) 1025–1041.

5. Lee M. Y., Lee H. R., Park C. H., et al., Organic transistor-based chemical sensors for wearable bioelectronics, *Acc. Chem. Res.* 51 (11) (2018) 2829–2838.
6. Ruiz M. P., Aragonès A. C., Camarero N., et al., Bioengineering a single-protein junction, *J. Am. Chem. Soc.* 139 (43) (2017) 15337–15346.
7. Amdursky N., Głowacki E.D., Meredith P., Macroscale biomolecular electronics and ionics, *Adv. Mater.* 31 (3) (2019) 1802221.
8. Torculas M., Medina J., Xue W., Hu X., Protein-based bioelectronics, *ACS Biomater. Sci. Eng.* 2 (8) (2016) 1211–1223.
9. Bostick C. D., Mukhopadhyay S., Pecht I., et al., Protein bioelectronics: A review of what we do and do not know, *Rep. Progr. Phys.* 81 (2) (2018) 026601.
10. Zezina T. I., Tsybin O. Yu., Subpicosecond dynamics of the molecular polyalanine dipole moment, *St. Petersburg State Polytechnical University Journal. Physics and Mathematics.* 10 (4) (2017) 100–110.
11. Polytechnic – RSK Tornado, Technical Characteristics, URL: <https://rscgroup.ru/project/spbstu-politechnic/> (The date of access is 15.07.2022).
12. Baranov M. A., Tsybin O. Yu., Velichko E. N., Free energy relaxation of glycine, tryptophan and albumin molecules in the ionized aqueous solution: Computer analysis, *St. Petersburg State Polytechnical University Journal. Physics and Mathematics.* 14 (4) (2021) 135–146.
13. Baranov M. A., Tsybin O. Yu., Velichko E. N., Structured biomolecular films for microelectronics, *St. Petersburg State Polytechnical University Journal. Physics and Mathematics.* 14 (1) (2021) 85–99.
14. Baranov M., Nepomyashchaya E., Velichko E., Computer simulation of biomolecules around metallic nanoparticle for biomolecular electronics, *Proc. 2021 Int. Conf. on Electr. Eng. & Photonics (EExPolytech 2021).* Oct. 14–15, 2021, SPbPU, St. Petersburg (2021) 171–174.
15. Velichko E., Zezina T., Baranov M., et al., Dynamics of polypeptide cluster dipole moment for nano communication applications, In book: *Lecture Notes in Computer Science (including book subseries “Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics”)*. Vol. 11118, Springer Science + Business Media, Berlin (2018) 675–682.
16. Matei A., Drichko N., Compf B., Dressel M., Far-infrared spectra of amino acids, *Chem. Phys.* 316 (1–3) (2005) 61–71.
17. Barth A., Zscherp C., What vibrations tell us about proteins, *Q. Rev. Biophys.* 35 (4) (2003) 369–430.
18. Butyrskaya E. V., Nechaeva L. S., Shaposhnik V. A., Drozdova E. I., Solutions of glycine on the basis of quantum-chemical calculations, *Sorption and Chromatography Processes.* 12 (4) (2012) 501–512 (in Russian).
19. Tortonda F. R., Pascual-Ahuir J.-L., Silla E., et al., Aminoacid zwitterions in solution: Geometric, energetic, and vibrational analysis using density functional theory-continuum model calculations, *J. Chem. Phys.* 109 (2) (1998) 592–602.
20. Wolpert M., Hellwig P., Infrared spectra and molar absorption coefficients of the 20 alpha amino acids in aqueous solutions in the spectral range from 1800 to 500 cm⁻¹, *Spectrochim. Acta, Part A. Mol. Biomol. Spectrosc.* 64 (4) (2006) 987–1001.
21. Itoh K., Ikeda A., Iwamoto T., Nishizawa S., DFT calculation analysis of terahertz time-domain spectra of polyalanines, *J. Mol. Struct.* 1006 (1–3) (2011) 52–58.
22. Barth A., The infrared absorption of amino acid side chains, *Prog. Biophys. Mol. Biol.* 74 (3–5) (2000) 141–173.
23. Mohamed M. E., Mohammed A. M. A., Experimental and computational vibration study of amino acids, *Int. Lett. Chem. Phys. Astron.* 15 (2013) 1–17.

СВЕДЕНИЯ ОБ АВТОРАХ

БАРАНОВ Максим Александрович – ассистент Высшей школы прикладной физики и космических технологий, младший научный сотрудник научной лаборатории лазерной фотометрии и спектроскопии Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.

195251, Россия, г. Санкт-Петербург, Политехническая ул., 29
 baranovma1993@gmail.com
 ORCID: 0000-0003-4555-0009

НЕПОМНЯЩАЯ Элина Константиновна – ассистент Высшей школы прикладной физики и космических технологий, младший научный сотрудник научной лаборатории лазерной фотометрии и спектроскопии Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.

195251, Россия, г. Санкт-Петербург, Политехническая ул., 29

elina.nep@gmail.com

ORCID: 0000-0002-4416-9380

ЦЫБИН Олег Юрьевич – доктор физико-математических наук, профессор Высшей инженерно-физической школы Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия.

195251, Россия, г. Санкт-Петербург, Политехническая ул., 29

oleg.tsybin@gmail.com

ORCID: 0000-0002-8308-059X

THE AUTHORS

BARANOV Maksim A.

Peter the Great St. Petersburg Polytechnic University

29 Politechnicheskaya St., St. Petersburg, 195251, Russia

baranovma1993@gmail.com

ORCID: 0000-0003-4555-0009

НЕПОМНЯЩАЯ Элина К.

Peter the Great St. Petersburg Polytechnic University

29 Politechnicheskaya St., St. Petersburg, 195251, Russia

elina.nep@gmail.com

ORCID: 0000-0002-4416-9380

ТСЫБИН Олег Ю.

Peter the Great St. Petersburg Polytechnic University

29 Politechnicheskaya St., St. Petersburg, 195251, Russia

oleg.tsybin@gmail.com

ORCID: 0000-0002-8308-059X

Статья поступила в редакцию 22.07.2022. Одобрена после рецензирования 23.08.2022. Принята 23.08.2022.

Received 22.07.2022. Approved after reviewing 23.08.2022. Accepted 23.08.2022.