Theoretical modeling of two-dimensional molecular structure by laser diffraction of electrons from polyatomic molecules

A. Kharlamova[a, 1] , *D. Makarov*[a]

[a] Higher School of Natural Sciences and Technologies Northern (Arctic) Federal University named after M.V. Lomonosov, Arkhangelsk, Russia

Представлены методы теоретического моделирования спектра взаимодействия ультракороткого лазерного импульса с молекулой ClCF3. Методы позволяют численно решить задачу теоретического моделирования спектра взаимодействия ультракороткого импульса с веществом. Результаты расчетов сравниваются с результатами экспериментальной работы другой группы исследователей. Цель данной работы в сравнении и проверке предложенных методов, нахождении в методе ошибок и недостатков.

Methods of theoretical modeling of the spectrum of interaction of an ultrashort laser pulse with a ClCF3 molecule are presented. The methods allow us to numerically solve the problem of theoretical modeling of the spectrum of interaction of an ultrashort pulse with matter. The results of the calculations are compared with the results of the experimental work of another group of researchers. The purpose of this work is to compare and verify the proposed methods, to find errors and shortcomings in the method.

PACS: 44.25.+f; 44.90.+c

Introduction

In the study of molecular compounds, the most promising tool for analyzing atomic structures is a free electron laser [1]. This is due to the difficulties of analyzing molecules using other methods. There are several ways to study the structure of biomolecules, for example, the Sanger method. Such sequencing methods do not provide accurate results at the atomic level of research, while they are lengthy and time-consuming [2]. NMR spectroscopy also has a number of problems in the study of protein compounds, for example, the overlap of lines in the one-dimensional NMR spectrum, which occurs due to the larger number of atoms compared to a simple organic compound. X-ray radiation also does not allow for ultra-high resolution analysis [3]. The disadvantages of existing methods have prompted scientists to search for new ways to study complex molecular structures, namely the use of ultrashort laser pulses in the analysis [4, 5, 6]. In practice, visualizations of various molecules are already being obtained, mostly simple ones such as N2 and O2. The distances between atoms are determined by fitting experimental diffraction images over spectra using the model of independent atoms [7]. However, as the authors of the article [8] note, in order to apply the method

[1] E-mail: kharlamova.anastasya2015@yandex.ru

to more complex polyatomic molecules and dynamical systems, it is necessary to complicate the procedure for restoring the molecular structure using fitting. This article continues the development of the method presented earlier [9] for calculating the model of the scattering spectrum of an ultrashort laser pulse with polyatomic structures. The method is based on the search for symmetries in the structure of the sample under study, i.e. repeating atoms. An example would be the repeating nitrogenous bases of DNA. We will simulate the pulse interaction based on a perfectly one-dimensionally aligned ensemble of ClCF3 molecules. The calculation will be performed in two ways. The first method is as if there were no symmetries and repetitions in the ensemble. In this case, the laser pulse is scattered on each atom. The second method is a symmetry calculation, where scattering modeling is performed only on a repeating part of the overall molecular structure, and then summation occurs. The results obtained will be compared with the experiments of [10], where diffraction patterns were obtained by other methods.

Theory

Two methods of numerical simulation of interaction spectra are used for the calculation. Both are based on the Dirac model.Here, the molecule is a multi-atomic structure and this is taken into account in the calculation. An ultrashort pulse falls on this molecule in the direction of $\mathbf{n_0}$. It is assumed that the duration of such a pulse is many times shorter than the characteristic atomic time $\tau_a \sim 1$ so $\tau \ll \tau_a$. It is well known that this condition applies in the approximation of a sudden disturbance. In the approximation of a sudden perturbation, the system's own Hamiltonian can be neglected, since the electron in the atom does not have time to evolve under the action of the momentum field. The condition $\tau \ll \tau_a$ can be extended for X-ray ultrasound and we take into account that the approximation of a sudden disturbance is applicable for $\omega_0 \tau_a \gg 1$, where ω_0 carrier frequency of the incident pulse [5].

The strength of the electromagnetic field is used for the calculation, which, to a first approximation, is considered inhomogeneous in space

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E_0} h\left(t - \frac{\mathbf{n_0 r}}{c}\right), \tag{1}$$

where $\mathbf{E_0}$ - field amplitude, $h\left(t - \frac{\mathbf{n_0 r}}{c}\right)$- some arbitrary function defining the shape of an ultrashort pulse, c - the speed of light.

Next, the electron wave function found for this case is used. In [11], the solution of the Dirac equation is considered in more detail. In this work, the magnetic component is not considered, i.e. it is assumed that the fields are weak $\frac{\mathbf{E_0}}{c^2} \ll 1$. Thus, the wave function used is

$$\Psi(t) = \varphi_0\left(\{\mathbf{r_a}\}\right) e^{-\sum_a i \int_{-\infty}^{t} \mathbf{E}(\mathbf{r_a}, t')\mathbf{r_a} dt'}, \tag{2}$$

where $\sum_a$ - the sum of all the electrons in the system under consideration, $\varphi_0\left(\{\mathbf{r_a}\}\right)$ - the initial wave function.

The calculations use the quantum theory of scattering of an ultrashort laser pulse, without restrictions on the number of atoms in the object under consideration [12]. In [12], the basic expressions for calculating the required scattering characteristics have already been obtained. Using them, we find an expression for calculating the scattering energy ε per unit solid angle $\Omega_{\mathbf{k}}$ ($\mathbf{k} = \frac{\omega}{c}\mathbf{n}$, where $\mathbf{n}$ – the direction of the scattered pulse) in a single frequency range ω (hereinafter referred to as the spectrum)

$$\frac{d^2\varepsilon}{d\Omega_{\mathbf{k}} d\omega} = \frac{[\mathbf{E_0 n}]^2 \left|\hbar(\omega)\right|^2}{(2\pi)^2 c^3} \left\langle \varphi_0 \left| \sum_{a,a'} e^{-i\mathbf{p}(\mathbf{r_a} - \mathbf{r_{a'}})} \right| \varphi_0 \right\rangle, \tag{3}$$

where $\tilde{h}(\omega) = \int_{-\infty}^{+\infty} h(\eta) e^{i\omega\eta} d\eta$ и $\mathbf{p} = \frac{\omega}{c}(\mathbf{n} - \mathbf{n_0})$ - A re-emission or recoil pulse is when an ultrashort laser pulse is scattered by a bound electron. The problem of finding the electron density is solved using the model of independent atoms [5]. We consider atoms separately from each other, i.e. in isolation. We divide the equation by two. In this case, the first term in parentheses corresponds to the summation for $a = a'$, and the second one $a \neq a'$, вas a result, it turns out

$$\frac{d^2\varepsilon}{d\Omega_{\mathbf{k}} d\omega} = \frac{[\mathbf{E_0 n}]^2 \left|\hbar(\omega)\right|^2}{(2\pi)^2 c^3} \left[\sum_{\mathrm{i}=1}^{A} N_{e,i} N_{A,i} \left(1 - |F_i|^2\right) + \sum_{i,j=1}^{A} \delta_{i,j} N_{e,i} N_{e,j} F_i F_j \right] \tag{4}$$

where A - the number of varieties of atom, $N_{e,i}$ - the number of electrons in a grade I atom, $N_{A,i}$ - number of atoms of the I-th grade, $F_i = \frac{1}{N_{e,i}} \int \rho_{e,i}(\mathbf{r}) e^{-i\mathbf{pr}} d^3\mathbf{r} = 1$ - it is the form factor of a grade I atom with an electron density $\rho_{e,i}(\mathbf{r})$, ω - pulse frequency, $\delta_{i,j} = \sum_{A_i A_j} e^{-i\mathbf{p}(\mathbf{R}_{A_i} - \mathbf{R}_{A_j})}$ - hereafter, we will call the interference factor, because it depends only on the spatial position of the atoms in the system. In the expression, the interference factor is written for atoms of the i-th grade. The position of the atoms is determined by the radius vector. In this expression, the interference factor is the main component, since it shows the spatial change in the position of atoms in the system. This formula underlies both methods of numerical search of the scattering spectrum. The difference between the methods is only in the method of calculating the interference factor. In the first method, the interference factor is calculated as shown above $\delta_{i,j} = \sum_{A_i A_j} e^{-i\mathbf{p}(\mathbf{R}_{A_i} - \mathbf{R}_{A_j})}$. All the atoms included in the system are taken into account here and the calculation is carried out for a long time. The second method reduces the calculation time of the interference factor. If the molecule in question has a certain frequency of repetition of atoms in the structer, then you can use the following formula to calculate

$$\delta_{i,j} = \sum_{\alpha=1}^{S} \sum_{n_\alpha=0}^{N_\alpha} e^{i\mathbf{p}\mathbf{R}_{n_\alpha}} \sum_{A_i \in R_{\alpha,1}} e^{i\mathbf{p}\mathbf{R}_{A_i}} \sum_{\beta=1}^{S} \sum_{n_\beta=0}^{N_\beta} e^{i\mathbf{p}\mathbf{R}_{n_\beta}} \sum_{A_{j \in R_{\beta,1}}} e^{-i\mathbf{p}\mathbf{R}_{A_j}} \tag{5}$$

where α or β - some symmetry in the system, s - the number of symmetries in the system, $\mathbf{R}_{n_\alpha}$ и $\mathbf{R}_{n_\beta}$ - radius is a vector specifying the position of the symmetry α or β accordingly, N_α - the number of movements with a given symmetry,$\mathbf{R}_{A_i}$ - radius is a vector that defines the position of atoms of type i within a region. $R_{\alpha,1}$, (similarly $\mathbf{R}_{A_j}$).

This method has already been tested in the calculation of various structures, mainly DNA and RNA, as it was originally developed for fast calculations on these molecules.

In previous works, this calculation method was called the symmetry method, and it performed well in calculations based on DNA, RNA, and simple carbon chain structures. The results of calculations using both methods were the same. In this paper, we will check the calculations on the ClCF3 molecule and compare the obtained diffraction patterns with data from another group of scientists. This will help to identify errors and shortcomings in the developed methods.

Modeling

An ensemble of ClCF3 molecules was selected as the molecule for irradiation. When choosing an object for theoretical modeling, the article [10] aroused interest. The authors also modeled diffraction patterns, but the model included experimental data on the number of errors and errors that are present in the real experiment.

The molecules are elongated along the y-axis, while some deviation error of the molecules from the axis is set, the molecules can precess. This was introduced by the authors of the article [10] based on experiments that showed that molecules cannot be perfectly aligned. In the simulation proposed here, in order to account for precession, the ensemble of molecules was arranged as follows Fig.1.

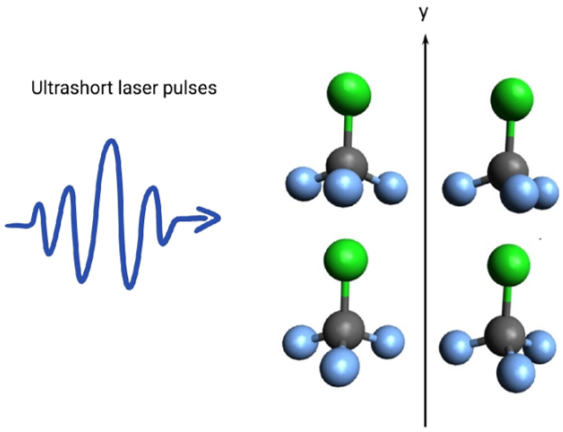


Fig. 1. The object of theoretical modeling of the interaction of ultrashort pulses with matter: the ensemble of ClCF3 molecules

The simulation was performed in two ways: the numerical method and the numerical method based on the search for symmetries. The result of the simulation is a spectrum constructed in a spherical coordinate system. The main axis of DNA runs along the molecule, orthogonally to its bases. The pulse falls on the system at an angle of 90 degrees to the main axis system parameter θ – is the angle between the helix axis and the scattering direction ϕ – is the angle between the helix axis and the projection on the plane perpendicular to the helix axis, the calculations were performed in spherical coordinates. The results in all figures are normalized to the maximum value of the scattering spectrum (see Fig. 2).

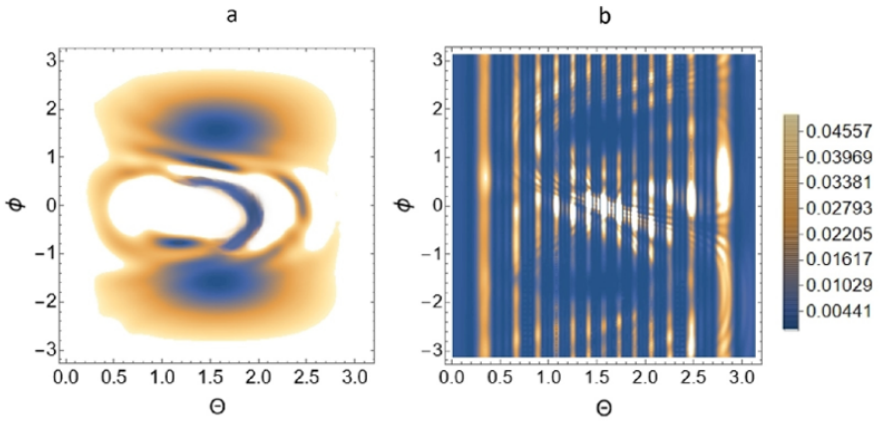


Fig. 2. Diffraction patterns of an ensemble of molecules: a) Numerical method b) Symmetry search method

As a result of the calculation, the authors of the article [10] obtain a 3D model of the scattering spectrum and show it as an expanded diffraction pattern. Our proposed method does not allow us to represent the spectrum in total over all scattering angles, and this is a flaw that needs to be worked out in the future. In our proposed method, the result of the calculation is a 2D model of the diffraction pattern. Therefore, only a fragment of the diffraction pattern from the article [10] is used for comparison.

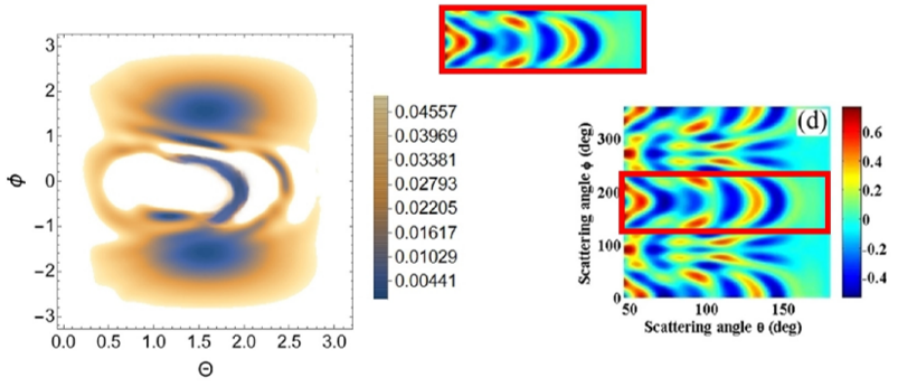


Fig. 3. Comparison of the obtained results with the research data

Conclusion

As a result of the simulation, diffraction patterns of the interaction spectra of an ultra-short laser pulse with an ensemble of ClCF3 molecules were obtained. A comparison of the obtained models showed that the first method of numerical modeling correlates with the data obtained by other research groups. Even with insufficient data on the reference model, it was possible to obtain a spectrum similar to the results of the work of another group. At the same time, it became clear how it was necessary to refine the method for greater accuracy and ease of operation in real experimental conditions. It is necessary to summarize the scattering angles and obtain a 3D diffraction pattern.

The second numerical calculation method, using the symmetry search method, has shown its inconsistency for this molecule. This is an important result, since no such errors were observed when modeling on DNA and RNA molecules. But modeling on an ensemble of molecules has shown that the calculation method requires serious refinement. This is due to the fact that the period of repetition of the symmetry region in the ensemble, i.e. of one ClCF3 molecule, is associated not only with the y axis, but also with the x axis, in this case. This problem was helped to identify such an interesting structure of molecules as an ensemble that precedes along an axis. In the future, it is necessary to refine the symmetry method by solving the problem of the arrangement of repeating sections of the molecule in the direction of different axes.

Acknowledgments

The study was supported by a grant from the Russian Federation, No. FSRU-2024-0005.

REFERENCES

1. *Tet I. et al.* Compact X-ray free-electron laser emitting in the sub-angstrom region // Nat. Photon — V. 6. — P. 540–544. — 2012.

2. *Chait E., Page G., Hunkapiller M.* Battle of the DNA sequencers // Nature — V. 333. — P. 477—478 . — 1988.

3. *Franklin R., Gosling R.* Evidence for 2-Chain Helix in Crystalline Structure of Sodium Deoxyribonucleate // Nature — V. 172.— P. 156—157 . — 1953.

4. *Deisenhofer J., Epp O., Miki K., Huber R.* X-ray structure analysis of a membrane protein complex. Electron density map at 3 A resolution and a model of the chromophores of the photosynthetic reaction center from Rhodopseudomonas viridis // J Mol Biol. — V. 180. — P. 385 . — 1984.

5. *Chapman H., Fromme P., Barty A. et al.* Femtosecond X-ray protein nanocrystallography. // Nature — V. 470. — P. 73—77 . — 2011.

6. *Jones N.* Crystallography: Atomic secrets. // Nature — V. 505. — P. 602—603 . — 2014.

7. *Hargittai I., Hargittai M.* Stereochemical applications of gas-phase electron diffraction. // Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials — V. 45. — P. 519—520 . — 1989.

8. *Sarah M., Wann D., Rankin D.* Structure by theory and experiment: One nationality, two languages // International Journal of Quantum Chemistry — V. 101. — P. 878–884 . — 2005.

9. *Makarov D.N., Kharlamova A.A.* Calculation of the Spectrum of Interaction Model with Macromolecules by the Method of Finding Symmetries // Phys. Part. Nucl. Lett. — V. 21. — P. 815–818. — 2024.

10. *Yu C., Wei H., Wang X. et al.* Reconstruction of two-dimensional molecular structure with laser-induced electron diffraction from laser-aligned polyatomic molecules // Sci Rep — V. 5. — P. 15753 . — 2015.

11. *Harker D., Kasper I.* Direct determination of the phases of Fourier coefficients from X-ray data //Acta Crystallographica — V. 35. — P. 271–275 . — 1948.

12. *Zachariasen W.H.* A new Analytical Metod for Solving Complex Crystal Structures //Acta Cryst. — V. 5. — P. 68–73 . — 1952.