Ultrafast isolated molecule imaging without crystallization

Zhuoran Ma\textsuperscript{a,b,1}, Xiao Zou\textsuperscript{a,b,1}, Lingrong Zhao\textsuperscript{a,b,1}, Pengfeng Qi\textsuperscript{a,b}, Tao Jiang\textsuperscript{a,b}, Pengfei Zhu\textsuperscript{a,b}, Dao Xiang\textsuperscript{a,b,c,d,2} and Jie Zhang\textsuperscript{a,b,c,2}

Contributed by Jie Zhang; received December 18, 2021; accepted March 10, 2022; reviewed by Martin Centurion and Jerome Hastings

Crystallography is the standard for determining the atomic structure of molecules. Unfortunately, many interesting molecules, including an extensive array of biological macromolecules, do not form crystals. While ultrafast and intense X-ray pulses from free-electron lasers are promising for imaging single isolated molecules with the so-called “diffraction before destruction” technique, nanocrystals are still needed for producing sufficient scattering signal for structure retrieval as implemented in serial femtosecond crystallography. Here, we show that a femtosecond laser pulse train may be used to align an ensemble of isolated molecules to a high level transiently, such that the diffraction pattern from the highly aligned molecules resembles that of a single molecule, allowing one to retrieve its atomic structure with a coherent diffraction imaging technique. In our experiment with CO$_2$ molecules, a high degree of alignment is maintained for about 100 fs, and a precisely timed ultrafast relativistic electron beam from a table-top instrument is used to record the diffraction pattern within that duration. The diffraction pattern is further used to reconstruct the distribution of CO$_2$ molecules with atomic resolution. Our results mark a significant step toward imaging noncrystallized molecules with atomic resolution and open opportunities in the study and control of dynamics in the molecular frame that provide information inaccessible with randomly oriented molecules.

Structure determination of molecules at the atomic level is essential for understanding molecules’ functions. Since the structural elucidation of myoglobin (1) and hemoglobin (2) in the 1950s, X-ray crystallography played a central role in identifying the structure of macromolecules. The periodically oriented identical molecules in large, high-quality crystals provide a coherent amplification of the diffraction signals forming discrete Bragg peaks for structure retrieval. However, it is challenging or impossible to produce large, defect-free crystals for many interesting molecules. To circumvent the need for crystallization, ultrashort and intense X-rays from free-electron lasers (FELs) have been proposed to image individual biomolecules through the “diffraction before destruction” method (3). The physics behind this technique is that if the X-ray pulse is short enough, the destruction occurs after the pulse has traversed the sample; and if the X-ray pulse is intense enough, useful, single-shot diffraction data could be obtained from scattering by a single isolated molecule.

In the past two decades, there have been worldwide efforts in extending the methodology of crystallography to allow imaging of the structure of single isolated molecules through the development of hard X-ray FELs (4–8), sample handling techniques (9, 10), and structure reconstruction algorithms (11, 12). However, the X-ray flux available today still requires nano-to-microsized crystals to produce useable scattering signals, as implemented in serial femtosecond (fs) crystallography (13–18). Because the intensity at the Bragg peaks is proportional to the square of the number of molecules in the crystal, even a nanocrystal would provide a several-orders-of-magnitude-higher scattering signal than a single molecule. With an upgrade to FELs in peak power and repetition rate, extending this methodology to crystals in the tens-of-nanometers size seems feasible, but there is still a long way to go to achieve atomic imaging of isolated noncrystallized macromolecules.

Alternatively, the structure of a molecule may be obtained with gas electron diffraction (19). Recently, it was proposed (20, 21) that a single macromolecule in doped liquid helium droplets may be aligned by a continuous laser, which enables the addition of a large number of diffraction patterns from successive molecules, eliminating the need for obtaining diffraction patterns in a single shot. The combination of laser alignment and gas electron diffraction also makes it possible to realize single molecule imaging on a table-top transmission electron microscope, taking advantage of the orders-of-magnitude-higher scattering cross-section of electrons. Instead of using a continuous laser to align a single molecule, here we used a train of ultrafast laser pulses to align an ensemble.

**Significance**

Excitation of molecules by an ultrashort laser pulse creates rotational wave packets that lead to transient alignment of the molecules along the laser polarization direction. Here, we show that a train of ultrashort laser pulses can be used to enhance the degree of alignment to a high level such that the diffraction from precisely timed ultrashort electron beams may be used to reconstruct the structure of the isolated molecules with atomic resolution through a coherent diffraction imaging technique. Our results mark a great step toward imaging noncrystallized molecules with atomic resolution and pave the way for creation of three-dimensional “molecular movies” at the femtosecond time scale and atomic spatial scale.

Author affiliations: *Key Laboratory for Laser Plasmas (Ministry of Education), School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China; **Collaborative Innovation Center of IFSA, Shanghai Jiao Tong University, Shanghai 200240, China; *Tsung-Dao Lee Institute, Shanghai Jiao Tong University, Shanghai 200240, China; and *Zhejiang Institute for Advanced Study, Shanghai Jiao Tong University, Shanghai 200240, China

Author contributions: D.X. and J.Z. designed research; Z.M., X.Z., L.Z., F.Q., and T.J. performed research; P.Z. contributed new reagents/analytic tools; Z.M., X.Z., and L.Z. analyzed data; and Z.M., D.X., and J.Z. wrote the paper.

Reviewers: M.C., University of Nebraska-Lincoln; and J.H., SLAC National Accelerator Laboratory. The authors declare no competing interest.

Copyright © 2022 The Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2122793119/-/DCSupplemental.

Published April 6, 2022.
of a large number of isolated molecules transiently and obtained a high-quality diffraction pattern with precisely timed ultrashort relativistic electron beams (Methods). The angular spread of the molecules with imperfect alignment is taken into account in the structural retrieval algorithm, and the atomic distribution of the molecule is reconstructed with subÅngström resolution using coherent diffraction imaging (CDI) methodology (22–25).

**Results**

The experiment was performed on a table-top ultrafast electron diffraction (UED) facility where an ultrashort electron beam with mega-electron-volt energy is used to produce the diffraction pattern (26). The spatial coherence length of the electron beam is longer than the size of the molecule such that the condition for CDI is satisfied, but it is significantly shorter than the average distance between molecules, which ensures that no interference between the electrons scattered from different molecules will occur. As a result, the diffraction pattern is an incoherent sum of the diffraction intensities over many molecules. With randomly oriented molecules, the fine features of the diffraction pattern from individual molecules are smeared out, and only 1D information (i.e., the interatomic distance) can be obtained from the isotropic diffraction distribution. In contrast, by aligning the molecules to the same direction, the diffraction distribution is the same for each individual molecule, and the sum over all the molecules simply enhances the signal by the number of molecules with the fine features maintained. Furthermore, the continuous supply of gas sample enables the scattering signal to be integrated over many pulses until a diffraction pattern with a sufficient signal-to-noise ratio is obtained.

Molecules may be aligned with a laser in two different regimes (27). When the laser pulse duration is much longer than the rotational period of the molecule, the alignment is considered adiabatic, and the degree of alignment follows the intensity envelope of the laser pulse. In contrast, if the laser pulse duration is much shorter than the molecular rotational period, the alignment is considered impulsive, where excitation of molecules by an ultrashort laser creates rotational wave packets that lead to transient alignment of the molecules along the laser polarization direction. Intuitively, the alignment results from the torque generated by the laser field and the induced dipole moment of the molecule. Impulsive alignment has the advantage of achieving the highest alignment state after the laser pulse turnoff, allowing the alignment to be achieved in a field-free condition and avoiding any distortions to the molecular structure.

In this experiment, CO$_2$ was chosen as the model system for illustrating the potential of CDI for molecules transiently fixed in space. CO$_2$ is a linear molecule with a relatively large polarizability anisotropy, and rotation around the molecular axis does not affect the diffraction distribution. The experiment is schematically shown in Fig. 1. The electron beam with a kinetic energy of about 3 MeV (Methods) is diffracted from the CO$_2$ molecules delivered to the interaction point by a pulsed nozzle with a backing pressure of about 1.3 bar. The 800-nm alignment laser pulse is divided into four pulses with equal energy and different time delays (Methods and SI Appendix, Fig. S1) before it is directed to the molecules. The measured diffraction pattern is used to reconstruct the atomic distribution of the molecule using the CDI methodology.

The diffraction pattern before laser excitation expressed as a function of the momentum transfer $s = (4\pi/\lambda)\sin\phi/2$ is shown in Fig. 2A, where $\lambda$ is the de Broglie wavelength of the electron beam, and $\phi$ is the angle between the scattered and incident electrons. The diffraction intensity consists of two parts (28): the atomic scattering intensity $I_a$ from each atom in the molecule and the molecular scattering intensity $I_m$ from interference for atom pairs. Because the diffraction intensity quickly decays with momentum transfer, throughout this paper, the diffraction pattern is divided by $I_a$ to highlight the features at high momentum transfer. The molecular scattering intensity is converted into a probability distribution in real space by inverse Fourier transform followed by Abel inversion, as shown in Fig. 2B. This procedure returns a pair distribution function (PDF) containing the information on both the interatomic distance of all of the atom pairs in the molecules and the angular distribution of the molecules (29). The isotropic inner and outer rings in Fig. 2B represent the C–O and O–O atom pairs for randomly oriented molecules.

A vertically polarized ultrafast laser will impulsively excite a rotational wave packet; the molecules will be transiently aligned, followed by rotational revivals (29–31). The evolution of the rotational wave packet can be traced by the anisotropy $(S_h - S_v)/(S_h + S_v)$ in the diffraction intensity as shown in Fig. 2C, where $S_h$ and $S_v$ are the integrated intensity in the horizontal and vertical cones, respectively, as indicated by the dashed line in Fig. 2A. The measurement was done with a time step of 100 fs, and each data point was recorded with an

![Fig. 1. CDI of molecules transiently fixed in space. The molecules, supersonically expanded into a vacuum by a pulsed nozzle, were aligned by a train of laser pulses and scattered by MeV electron beams to yield diffraction intensity that allows reconstruction of the atomic structure through the CDI technique.](https://doi.org/10.1073/pnas.2122793119)
acquisition time of 90 s. The simulated alignment factor $<\cos^2\theta>$ (i.e., the ensemble-averaged degree of alignment) is in excellent agreement with the experimental results (Methods), where $\theta$ is the polar Euler angle between the laser polarization and molecular axis. The diffraction pattern and the corresponding PDF distribution at the full revival (at a delay time of about 42.7 ps) with the highest alignment are shown in Fig. 2D and E, respectively. Fig. 2D indicates a laser alignment anisotropy, and Fig. 2E indicates that the molecules are preferably confined in the vertical cone along the laser polarization direction. For a symmetric top molecule with a heavy atom, the diffraction under such partial alignment may be used to determine the bond lengths and angles of the molecule (32); however, generally, the degree of alignment from a single ultrashort laser pulse is not sufficiently high to allow robust retrieval of the structure of molecules.

The degree of alignment can be improved by increasing the laser intensity and reducing the molecules’ temperature. While the temperature of the molecules may be reduced to a few Kelvin by increasing the distance from the nozzle to the interaction point, it nonetheless reduces the density of the molecules significantly, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34).

The degree of alignment can be improved by increasing the laser intensity and reducing the molecules’ temperature. While the temperature of the molecules may be reduced to a few Kelvin by increasing the distance from the nozzle to the interaction point, it nonetheless reduces the density of the molecules significantly, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34). The laser intensity can be increased only to a certain limit, making it difficult to obtain a high-quality diffraction pattern (33, 34).

We begin by measuring the temporal evolution of the rotational wave packet from 0 to 700 ps, limited by the range of our delay stage. The evolution of the anisotropy from the measured electron diffraction pattern following excitation by a 3-mJ laser pulse is shown in Fig. 3A, where revival at the 16th period with similar amplitude is observed, implying that the decoherence effect of rotational states is negligible. We then tested the feasibility of coherent superposition with a two-pulse alignment experiment. Enhancement and suppression of anisotropy were observed, depending on the time delay of the two pulses, as shown in Fig. 3B. The time delay for maximal enhancement of anisotropy is determined to be 42.7 ps, consistent with the rotational period. Finally, a train of four pulses (3 mJ each) was used to repetitively drive the molecules to a high degree of alignment, as shown in Fig. 3C.

A high degree of alignment is maintained for about 100 fs (from $t = 128.1$ ps to $t = 128.2$ ps in Fig. 3C), and the diffraction pattern obtained within such a short time window by ultrashort electron pulse is shown in Fig. 4A, with the corresponding PDF shown in Fig. 4B. Large anisotropy and features that are otherwise smeared out at a low degree of alignment are clearly seen. Such a high degree of alignment allows us to retrieve the diffraction pattern for well-aligned molecules. For an ensemble of molecules with a specific angular distribution, the diffraction intensity is a linear superposition of all the molecules with different orientations. Mathematically, the measured diffraction distribution $I_D$ is a convolution of that for a single molecule $I_s$ within the angular distribution. We employ an iterative algorithm (Methods) to retrieve the diffraction distribution of a single molecule from the measured diffraction pattern and angular distribution (38). The retrieved diffraction pattern for a single molecule $I_s$ is shown in Fig. 4C where details washed out by the angular spread in partially aligned molecules are retrieved.

Taking the distribution in Fig. 4C as the Fourier transform of the molecule, the molecule’s structure may be reconstructed with the CDI method (22–25), where the missing phase

---

**Fig. 2.** Electron diffraction of laser-aligned molecules. (A and B) Diffraction pattern and PDF distribution for randomly oriented molecules. (C) Evolution of a rotational wave packet following impulsive excitation. (D and E) Diffraction pattern and PDF distribution for laser-aligned molecules.
information is recovered from the intensity via oversampling (39) (Methods). Such a lensless method has been applied to solid samples (22, 23) and large viruses (40, 41), where details on the order of a few tens of nanometers can be reconstructed. Here, we apply the CDI method to the small gas molecule CO₂. The real space potential map of the molecule reconstructed with an oversampling smoothness algorithm (42) by averaging over 200 independent runs is shown in Fig. 4D. It is seen that CO₂ is a linear molecule with C–O and O–O distances of 1.13 Å and 2.26 Å, respectively, in good agreement with the literature. The resolution is estimated to be about 0.7 Å from the maximal momentum transfer in this experiment.

**Discussion**

It should be noted that previous studies on aligned molecules with UED technique used only one ultrashort laser pulse (29, 31, 43). As a result, the degree of alignment is relatively low, such that the anisotropy is clearly seen only with the diffraction-difference method, and only PDFs are obtained. The degree of alignment may be significantly increased with a nanosecond laser pulse in the adiabatic alignment regime together with quantum state selection (33, 34). However, the quantum state selection unavoidably increases the distance from the nozzle to the interaction point, and therefore the density of the molecules is significantly reduced. This makes it challenging to obtain a high-quality diffraction pattern even with a state-of-the-art X-ray FEL (33, 34). Furthermore, the momentum transfer (about 4 Å⁻¹) in typical FEL measurements (33, 34, 44) needs to be significantly increased in order to achieve reliable reconstruction of molecular structure with atomic resolution.

Based on the fact that the rotational wave packet created by ultrashort laser pulse leads to periodic revival of the alignment, we used a train of laser pulses to significantly enhance the degree of alignment, such that the fine details related to the molecular structure are encoded and preserved in the measured diffraction pattern. The large scattering cross-section of the electrons and the relatively high molecule density, together with the improved temporal resolution (26), enabled us to obtain a high-quality diffraction pattern for transiently aligned isolated molecules in a 100-fs time window with a table-top

---

**Fig. 3.** Enhancement of degree of alignment with a laser pulse train. (A) Periodic revival of a rotational wave packet following impulsive excitation. (B) Enhancement (pulse delay at 42.7 ps) and suppression (pulse delay at 41.9 ps) of anisotropy by two laser pulses with various pulse delays. (C) Repetitive excitation for increasing the degree of alignment with a train of four pulses.

---

| A | B | C |
|---|---|---|
| ![Graph A](https://example.com/graphA.png) | ![Graph B](https://example.com/graphB.png) | ![Graph C](https://example.com/graphC.png) |

*Time delay (ps) vs. \( \cos^2(\theta) \) for different pulse delays.*
MeV UED instrument. The short de Broglie wavelength of electrons also provides a large momentum transfer (about $8\text{Å}/C_0$), crucial for reliable reconstruction of the atomic distribution in real space. The diffraction pattern for a single isolated molecule is retrieved with an iterative algorithm and further used to reconstruct the atomic distribution of the molecule with the CDI method.

The ultrafast lensless imaging method demonstrated here may be extended for three-dimensional (3D) imaging of more complex molecules, which requires 3D alignment with an elliptically polarized laser (45). A higher degree of alignment is likely required, which may be achieved by increasing the number of laser pulses and doping the molecules in helium droplet to reduce the temperature to sub-Kelvin (46). Quantum state selection may also be employed to further increase the degree of alignment for large molecules measured by UED and FELs with MHz repetition rates, which help to compensate the low scattering signal from the reduced density of molecules. Furthermore, a high degree of alignment demonstrated in a field-free condition, and the high temporal and spatial resolution of UED may enable studies of ultrafast dynamics through the pump-probe technique in the molecular frame (47, 48), marking a significant step toward the creation of a 3D “molecular movie” at the femtosecond time scale and atomic spatial scale.

Materials and Methods

Laser Pulse Train. The 800-nm laser with a pulse width of 70 fs (full width at half maximum [FWHM]) has an energy of 20 mJ, operating at 50 Hz. The four laser pulses were generated by two nested interferometers, as shown in SI Appendix, Fig. S1. Two translation stages adjust the interpulse delay. Because two of the pulses have slight offsets at the exit of the interferometer, a specially designed aspheric lens was employed to focus the four pulses to the gas jet with a similar size (about 0.4 mm FWHM) and position. The laser pulse train is directed to the gas jet at an angle of $4^\circ$ with respect to the electron beam.

MeV UED. The electron beam with a kinetic energy of about 3 MeV is produced in a photocathode radiofrequency gun. The electron beam pulse width is compressed in a double-bend achromat consisting of two dipole magnets and three quadrupole magnets (26). The electron beam is scattered by the molecules delivered to the interaction point by an Even-Lavie nozzle. The interaction region is $0.6\text{mm}$ away from the nozzle exit, and the electron beam size is about $0.3\text{ mm FWHM}$ at the interaction point. A phosphor screen captures the diffraction pattern imaged onto an electron-multiplying charge-coupled device camera. The electron pulse width is about 40 fs (FWHM) measured with a terahertz dector (49), and the temporal resolution in the measurement is estimated to be about 80 fs (FWHM).

Simulation of Rotational Wave Packet Evolution. The temporal evolution of the angular distribution of the molecules was simulated by solving the time-dependent Schrödinger equation. A theoretical diffraction pattern was calculated with this angular distribution, and its anisotropy was extracted in the same manner as the experiment. The evolution of the anisotropy curve—in particular, the amplitude and position of the peaks and valleys—is mainly determined by four parameters: the laser fluence, the initial rotational temperature of the molecules, the time resolution of the measurement, and a rescaling factor that accounts for the spatial overlap of the beams and the variation of the laser intensity across the sample (31). We performed a four-parameter fitting between the theoretical curve and experimental results, which returned a laser fluence of $1.43\text{ J/cm}^2$, rotational temperature of 53 K, rescaling factor of 0.36, and time resolution of about 100 fs FWHM.

Fig. 4. Reconstruction of molecular structure with lensless imaging. (A and B) Diffraction pattern and PDF distribution for highly aligned molecules with a train of four laser pulses. (C) Retrieved diffraction distribution for a single molecule. (D) Reconstructed molecular structure with the CDI method.
Iterative Algorithm to Retrieve Diffraction Pattern for a Well-Aligned Molecule. The iterative algorithm starts with an uniform distribution as an initial guess of $f_i$. A random small change is made for the guessed retrieved diffraction pattern in each iteration, which is then convoluted with the angular distribution to generate a diffraction pattern $I_g$. Only changes that reduce the difference between $I_g$ and $I_q$ are kept. The result converges after about 10^4 iterations, and the result shown in Fig. 4C is averaged over 10 independent runs. It is worth pointing out that the feasibility of the iterative algorithm has been tested with simulated diffraction pattern and angular distribution of the molecules in Yang et al. (38) and with measured diffraction pattern and simulated angular distribution in Hensley et al. (32). In this work, the algorithm is implemented with all the parameters measured in the experiment.

Reconstruction with CDI. The real space potential map of CO$_2$ is reconstructed from the retrieved single molecule diffraction pattern in Fig. 4C by solving the phase problem with an oversampling smoothness algorithm (42). In Fourier space, a constraint is applied to match the amplitude in the retrieved diffraction distribution for a well-aligned molecule. No explicit constraints are used for the regions of missing data at the center of the diffraction. In real space, a positivity constraint is applied. In total, 200 independent runs with random starting distributions are performed, and the result shown in Fig. 4D is the average over the 200 runs. Each run iterates between real and reciprocal space for a total of 2000 times. All 200 reconstructions had Fourier errors below 0.28.

Data Availability. All study data included in the article and SI Appendix. The data are available on the Science Data Bank at https://www.scidb.cn/s/3Mqqm.

Acknowledgments. We thank Dan Li, Huaidong Jiang, and Jie Yang for useful discussions. This work was supported by the National Key R&D Program of China (No. 2021YFA1400202), the National Natural Science Foundation of China (Grants Nos. 11925505, 12005132, 11504232, and 11721091), the Office of Science and Technology, Shanghai Municipal Government (No. 16D222260200), and the China National Postdoctoral Program for Innovative Talents (No. BX20200220).