Doctoral Dissertation (Censored)

博士論文（要約）

Stochastic Dynamic Behaviors of Single Molecules Revealed by Sub-Millisecond Electron Microscopic Imaging

（サブミリ秒電子顕微鏡動画撮影による確率論的な単分子動的挙動の解析）

A Dissertation Submitted for the Degree of Doctor of Philosophy

December 2020

令和２年１２月 博士（理学）申請

Department of Chemistry, Graduate School of Science,
The University of Tokyo

東京大学大学院理学系研究科化学専攻

Toshiki Shimizu

清水　俊樹

Abstract

This thesis describes a new methodology to observe and analyze stochastic dynamic behaviors of single molecules in situ at maximum sub-millisecond and sub-angstrom resolution. By combining an atomic-resolution electron microscope, a fast shutter speed camera, and the optimized denoising algorithm for electron microscopy imaging, we can unveil the single molecular dynamics, opening up a new field of nanoscale science.

Chapter 1 describes the general introduction of the molecular imaging method. Various methods for imaging molecules from the point of view of spatial and temporal resolution are discussed. Especially, single-molecule atomic-resolution real-time electron microscopy (SMART-EM) imaging is a key to observe stochastic molecular events in situ. With the development of cameras in the field of transmission electron microscopy (TEM), we now can capture electron microscopic images at 1600 frame/s. However, these images are full of noises, which prevents us from reaching the molecular world.

Chapter 2 describes the investigation of various noise reduction methods for noisy TEM videos taken by a fast shutter speed camera. Chambolle total variation denoising algorism is revealed to be an optimized method for current setups, considering signal to noise ratio and molecular edge preservation of the denoised images. With the denoising method in hand, I can capture various molecular events at a sub-millisecond level.

第３章については, 5 年以内に雑誌等で刊行予定のため，非公開.

Chapter 4 describes the success of recording a video of a single molecular shuttle encapsulated in a carbon nanotube at the time resolution of a sub-millisecond. The sub-millisecond SMART-EM imaging method revealed the rich molecular dynamics of a single fullerene molecule shuttling, rotating, and interacting with a vibrating carbon nanotube.

Finally, Chapter 5 summarizes present studies and future perspectives.

Acknowledgments

I am very grateful to my supervisor, Prof. Eiichi Nakamura, who guided me through the process of writing this thesis. His enthusiasm, motivation, immense knowledge, and some ambitious orders helped me in not only research but also in the decision for my futures.

This thesis would not be possible without the support of Dr. Koji Harano. I am very grateful for all the discussions, which were essential for understanding the experimental results, and for the guidance in the whole research process.

I also wish to express my best acknowledgment to Dr. Takayuki Nakamuro for checking my thesis thoroughly, giving me accurate advice, discussing both research and life.

I also thank Dr. Rui Shang for his helpful advice, suggestions, and encouragements.

I would also like to thank Prof. Kaoru Yamanouchi and Prof. Mitsuhiro Murayama for fruitful discussion related with my research.

I appreciate the support of all members from the subgroup studying electron microscopy; Mr. Hiroki Hanayama, Mr. Ryosuke Sekine, Mr. Ko Kamei, Mr. Hikaru Uchida, Mr. Dongxin Liu, Mr. Issei Tomotsuka, Mr. Keishi Takeuchi, Mr. Masaya Sakakibara, and Takato Ogata. I also thank the support from former members; Dr. Stuckner Joshua, Dr. Dominik Lungerich, Dr. Junfei Xing, Dr. Yuki Itabashi, Dr. Luca Schweighauser, Ms. Satori Kowashi, and Mr. Takuya Tsubota. I would like to thank Mr. Takumi Sakamaki and Mr. Takahiro Doba for eating lunch with me.

I am indebted to all of my laboratory colleagues who showed immense support and patience to me. I also thank Dr. Takao Kaneko, Ms. Akemi Maruyama and Mr. Koshi Chiba, who have supported the research in this laboratory.

I would like to thank ALPS program (MEXT) that gave me financial support and providing an opportunity to research in a global view through my Ph.D. study.

Finally, I would like to express sincere gratitude to Mr. Yasuyuki Shimizu and Ms. Miho Shimizu for their daily help, advice, and encouragement and for their continuous love.

Table of Contents

Abbreviations

AFM	atomic force microscopy
CCD	charge-coupled device
CMOS	complementary metal oxide semiconductor
CNT	carbon nanotube
DFT	density functional theory
DS	downsampling
e^-	electron
EDR	electron dose rate
EM	electron microscopy
fps	frame per second
MD	molecular dynamics
ms	millisecond
NMR	nuclear magnetic resonance
OT	Osawa and Tománek
PSF	point spread function
ROI	region of interest
SNR	signal to noise ratio
SMART-EM	single-molecule atomic-resolution real-time electron microscopy
TEM	transmission electron microscopy
vdW	van der Waals
XCF	cross-correlation factor

－Chapter 1－

General Introduction

1.1 Science innovated by imaging

"A picture is worth a thousand words." As the saying goes, people can obtain a lot of information by seeing images. From a historical view, the invention of the microscope and telescope sparked people's interest in invisible objects in the 17th century.[1] Following these innovations, the invention of motion pictures opened a new door in 1878 when a galloping horse was photographed in a series of frames with a shutter speed of 2 ms/frame as shown in Figure 1.1.[2,3] This is the first example of video, and by capturing the decisive moment of horse gaits, people solved the mystery of the leg patterns whether one or more hooves leave the ground at a time. The same can be said for scientific study, especially against the nanoscale world. If we can observe and record the decisive moment of scientifically interesting phenomena as sequential images or videos, it will even have the power of changing the history of science.

With the tremendous efforts of scientists over the centuries, we are now able to spectroscopically image molecules. The static molecular pictures have been revealed by crystallographic techniques, and dynamic behaviors can be partially studied by NMR or other spectroscopic methods. However, in situ observation of molecular events such as chemical reactions and molecular dynamics have remained to be a big challenge for decades. To understand these events deeply, both spatial and temporal resolutions are essential. Especially, time-resolved imaging of molecular events is of great importance to unveil the behavior of an individual molecule in the nanoscale system and thus to break into an unexplored field of science. Here in this thesis, I show a single molecular imaging methodology at a level of sub-millisecond and sub-angstrom, which no one has ever seen before. I believe this imaging method will make an innovation to nanoscale science.

Figure 1.1. The horse in motion taken by a shutter speed of 2 ms/frame. The picture is obtained from ref. 2. Copyright, 1878 by Muibridge.

1.2 Computational studies in molecular science

Scientists have been investigating various ways to understand molecular events such as chemical reactions. One such approach is computational chemistry such as density functional theory (DFT) for calculating the electron density of molecules and molecular dynamics (MD) for simulating the molecular behaviors. Chemical reactions are extremely fast, and electrons involving chemical reactions are transferred from an atom to an atom in a very short time. Since the electron transfer process cannot be observed with an electron microscope, it has been thought that it is impossible to follow the details of chemical reactions without computational supports. Therefore, many researchers have participated in this computational field, and with the development of computer technology, the study of chemical reactions using computer science has made significant progress.[4] The verification of the differences between the ideal chemical reaction calculated by computational chemistry and the real one caused by the complex combination of various external stimuli is expected to provide important insights for future chemical reaction researches. Plus, it is now possible to observe structural changes associated with atomic transfer by microscopy, which is described in detail below.

1.3 Single molecular imaging with a high spatial resolution

We cannot observe molecules freely moving around in space. This is because molecular motion far exceeds the spatial and temporal resolution of human eyes. However, with the remarkable development of microscope technology in recent years, it has become possible to identify each molecule at the atomic level. One of the methods that has a high spatial resolution up to sub-angstrom level is atomic force microscopy (AFM). As shown in Figure 1.2, hexagonal shapes that came from benzene rings of a pentacene molecule fixed on the substrate can be clearly observed and analyzed.[5,6]

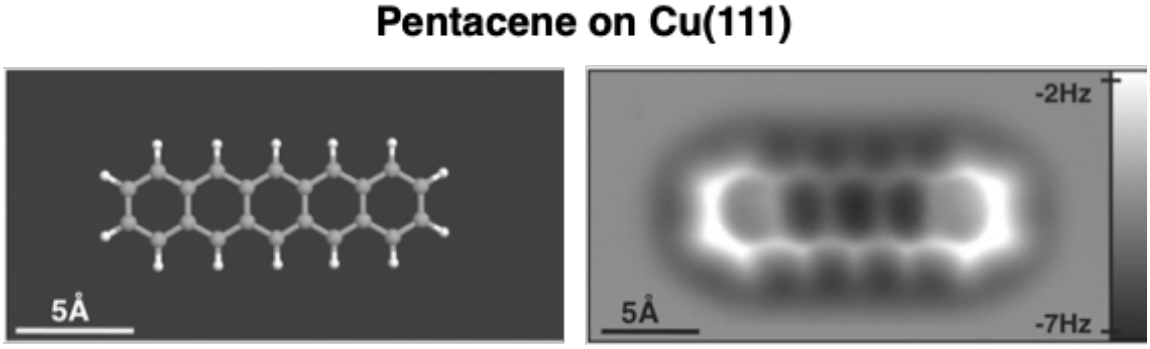

Figure 1.2. AFM imaging of pentacene on Cu(111). Ball-and-stick model of the pentacene molecule (left). Constant-height AFM images of pentacene acquired with a carbon monoxide-modified tip (right). Adapted with permission from ref. 5. Copyright 2009 The American Association for the Advancement of Science.

However, this method requires immobilization of molecules, usually limited to planer molecules, to be observed on the surface of a substrate like a metal. Furthermore, due to the limit of screening speed of a tip for imaging, the “before” and “after” stage of the dynamic phenomena can only be imaged. In other words, it is difficult to observe the dynamic behavior of molecules such as structural changes and reactions in real-time by using AFM techniques. What we want to observe is the “during” stage of molecular behaviors and structural changes, leading to the importance of real-time imaging of molecular dynamics. Note that since these methods only focus on the static processes of molecules, blurring of images is not a problem.

1.4 Single molecular imaging with a high temporal resolution

For single molecular imaging, not only a high spatial resolution but also a high temporal resolution is indispensable to understand molecular phenomena deeply. There is pump-probe spectroscopy as a method to maximize the time resolution up to a level of femtosecond.[7] In this method, a laser pulse (typically femtosecond or picosecond laser) pumps the population of a molecule into an excited state or causes its structural change, while another laser pulse with a delay probes the spectrum of the molecules to track its change as shown in Figure 1.3. In addition, a microscope applying this principle has also been developed.[8] Because the molecular phenomena are tracked by the change of spectrum, direct structural information of the molecule is not provided. This method requires the probe pulse to be applied to the sample at the moment when something happens. It means that molecular events should be easily reproducible and repeatable for imaging. However, any dynamic events seen at an atomistic or molecular level are stochastic, which makes us continuously capture dynamic phenomena until the events occur. The biggest problem in the visualization of molecular phenomena is that it is completely unpredictable, and this is the point that needs to be verified experimentally rather than computationally. However, it is important to mention that pump-probe spectroscopy and other single molecular methods with a relatively low temporal resolution give complementary information to unveil the behavior of an individual molecule in the nanoscale system.

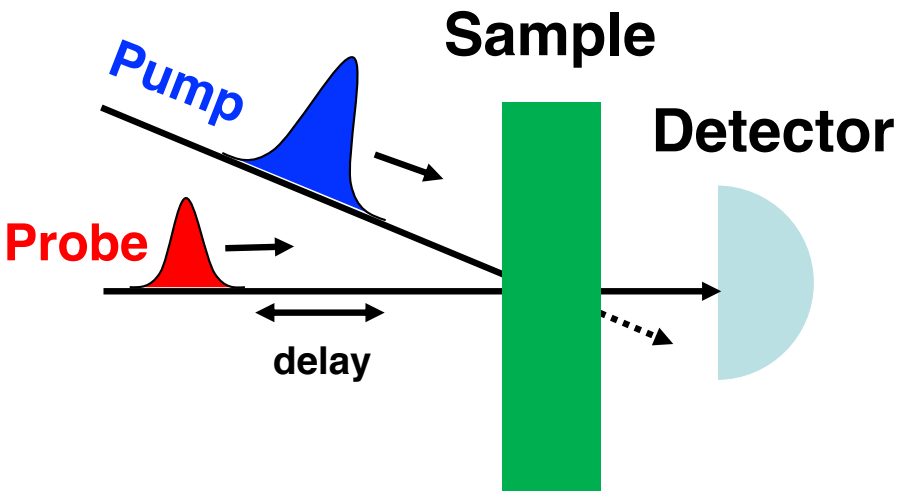

Figure 1.3. Schematic image of pump-probe spectroscopy.

1.5 Single-molecule atomic-resolution real-time electron microscopy (SMART-EM)

Under these circumstances, our group has developed single-molecule atomic-resolution real-time electron microscopy (SMART-EM) to observe and analyze dynamic behaviors of a single molecule in situ.[9] By encapsulating the target molecules in a single-walled carbon nanotube (CNT) as a substrate, we can observe various molecules with a transmission electron microscope (TEM) at an angstrom resolution. This confined nano-space limits molecular behaviors to two dimensions and facilitate the study of real-time molecular dynamics. For example, our group has achieved a kinetic study of [60]fullerene (C_{60}) dimerization by analyzing the reaction events in situ one by one.[10] Therefore, the SMART-EM imaging method is an appropriate method to conduct a real-time observation of stochastic dynamic behaviors of single molecules. By recording TEM videos of molecular events such as chemical reactions and mechanical motions continuously, it is possible to capture the decisive moment of scientifically meaningful phenomena. However, one big problem has remained since the advent of SMART-EM technology reported in 2007.[11] We could not obtain the high-quality molecular TEM images that we envisioned. For example, as shown in Figure 1.4, C_{60} molecules were often imaged like rugby balls, which should be imaged as hollow circles in principle. This can be due to the motions of the molecules in the horizontal-axis direction, and it means that we can only obtain less dynamical information from the blurred images. The struggle in our laboratory continued for 10 years, whether the resolution of the electron microscope itself was low, out of focus during TEM imaging, blurring due to rapid molecular motion, or whether the molecule being viewed had decomposed.

Here, in this thesis, I have identified the possible problems one by one and proved that one important cause is due to the slow shutter speed of TEM camera or low temporal resolution of TEM images for molecular studies. Generally, TEM images are taken by the TEM camera which has a shutter speed of 1–2 frame/s (fps). On the other hand, high-speed video recording at 1600 fps discussed in this thesis made it possible to image such a molecule clear enough to analyze in detail in the SMART-EM study, leading to the conclusion that with conventional slow TEM cameras, it was possible to obtain only blurred TEM images as a whole.

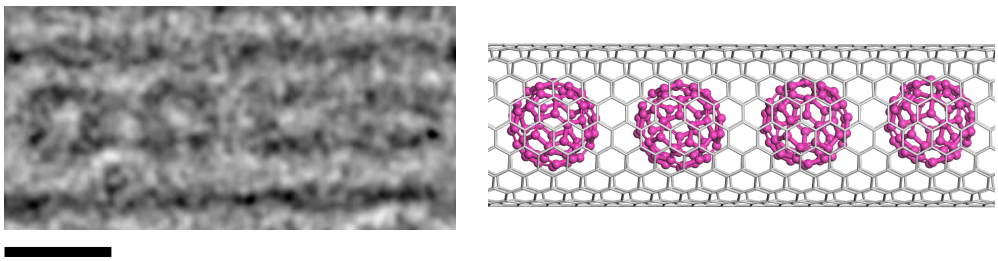

Figure 1.4. Blurred TEM image (2 frame/s) of van der Waals (vdW) dimers of C_{60} in a CNT (left) and its model (right). The purple structures refer to C_{60}. Scale bar: 1 nm.

1.6 Real-time observation of molecular events with a fast imaging rate

In the macroscopic world that we can directly recognize with our eyes, we can find scientifically interesting phenomena by imaging them with a fast imaging rate. For instance, a milk crown is a hydrodynamical phenomenon in which a beautiful crown-like shape is formed when a drop of milk is dropped on the liquid surface of milk. This phenomenon, which is so fast that it cannot be seen with our eyes, can be seen in the crown-like droplets as shown in Figure 1.5 when this phenomenon is taken by a high-speed camera. High-speed imaging clearly shows the process of milk crown formation, leading to further understanding of the mechanism.

On the other hand, in the molecular world, it is difficult to achieve both spatial resolution for viewing atoms and molecules and temporal resolution for tracking frequent high-speed motions of individual targets. For example, it was realized in the 1970s that molecules were confined in crystals and observed precisely with TEM over time.[12,13] However, when the target molecule moves around, it is necessary to use a high-speed camera to prevent blurring, but the electron dose irradiated to molecules per frame is insufficient, and a signal of TEM image will be buried by noise, and hence we cannot see the molecules. Recently, the progress of high-speed AFM for biomolecules that move relatively slowly and small organic molecules has been remarkable, and it has become possible to continuously capture their motion by the imaging rate of up to about 20 fps.[14] However, the spatial resolution of high-speed AFM imaging at this speed is limited, and the structure determination at the atomic level has not been achieved yet.

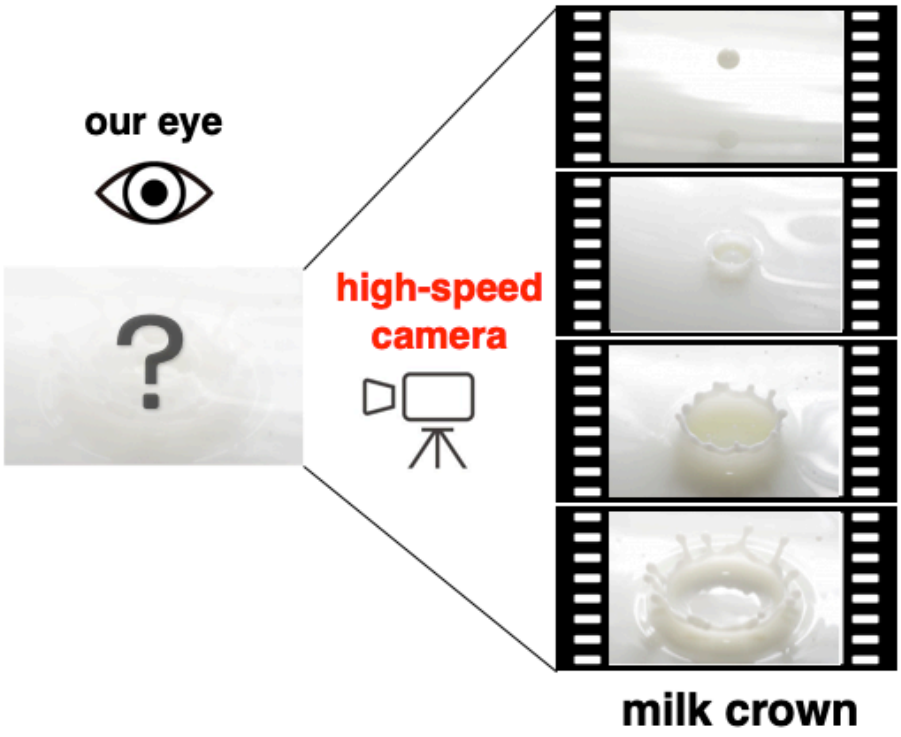

Figure 1.5. Capturing the process of a milk crown formation by a high-speed camera.

1.7 Development of a camera determining the performance of electron microscopy

With the development of image sensors from photosensitive films to charge-coupled device (CCD) and complementary metal oxide semiconductor (CMOS), high-sensitivity, high-speed imaging cameras have been developed one after another in recent years. The clear image without blurring obtained by these cameras proved to be effective in cryo-electron microscopy, which led to the explosive expansion of its application to the structural analysis of biomolecules from the 2010s to the present, as exemplified by the Nobel Prize in Chemistry 2017.[15] Along with this application, interest in the high-speed video recording of molecules increased, and the British group achieved TEM video recording of the motion of molecules in a CNT at an imaging rate of 12 fps in 2008, which was limited by the speed of the camera.[16] In response to the request of researchers who want to elucidate high-speed phenomena using a faster camera, the high-speed camera called K2-IS was launched in 2014. The camera can capture images at a maximum imaging rate of 1600 fps. By using this fast camera, there is a report that, it is possible to record the growth of the atomic layer of heavy atoms such as platinum on the crystal surface at an imaging rate of 400 fps to give the pictures of atomic-level growth mechanism.[17]

1.8 Image processing in molecular science

Images form important information and data in a variety of fields of science. Until recently, photographs or images were simply to give information about the appearances, and it is difficult to treat or quantify them mathematically. However, due to the advancement of computers and microelectronics, digital image processing and image analysis technology solved these problems related to the photography methods. Now, by using a K2-IS camera, we can obtain roughly about 10 GB of data within just a second. With large-scale data acquisition, the challenge of information extraction and image interpretation has also become more and more difficult. It is important to analyze them by not only visual inspection but also automated methods.[18]

In terms of molecular science, how clearly the target molecule can be seen is a key for further understanding of molecular behaviors. Generally speaking, against noisy images of molecules or particles, image processing such as filtering and denoising should be conducted to enhance the signal.[19] But, for fast imaging of molecular events, where molecules are mobile, it is difficult to apply the widely used enhancing methods for static objects.[20] To solve this problem, proper computational image processing of raw TEM images should be investigated.[21] However, there is no systematic study so far on image processing which is appropriate for noisy TEM images taken at a fast imaging rate or with low electron dose. Therefore, to innovate molecular science further, combining technology of information processing, microscopy, and scientific knowledge is crucial.

1.9 This work

With the development of high shutter-speed cameras for electron microscopy, we are now able to take videos at 1600 fps. However, their latent potential has not been fully realized because a high imaging rate makes the raw image noisy due to less electron dose. Especially, against molecules composed only from light elements such as carbon, it is difficult to obtain high contrast TEM images.

In Chapter 2, I developed the image processing for molecular science study with a collaborator, Dr. Joshua Stuckner, by introducing Chambolle total variation denoising algorism to successfully obtain sub-millisecond sub-angstrom videos of molecules. As target molecules, C_{60} molecules encapsulated in a CNT were used. This sub-millisecond SMART-EM imaging method makes it possible to visualize and analyze stochastic molecular events with a high spatial and temporal resolution, which cannot be conducted by other methods.

In Chapter 3, I applied this imaging technique to observe and analyze the C_{60} dimerization reaction in situ. I found the fusion reaction of C_{60} dimers takes place with the time resolution of a millisecond to give the short-lived intermediates which have never been captured experimentally. A TEM simulation analysis and automated cross-correlation image matching analysis, which was a theme of my master course studies, enabled me to characterize the intermediates with some accuracies to get insights into the chemical reactions.

In Chapter 4, I succeeded in recording of the world's fastest video of a single molecular shuttle encapsulated in a CNT at the time resolution of sub-millisecond. I analyzed these videos deeply to reveal the rich molecular dynamics of a single C_{60} molecule shuttling, rotating, and interacting with a vibrating CNT. I have quantitatively evaluated that the behaviors of the C_{60} molecule are driven by the supporting CNTs rather than its thermal motion.

Finally, Chapter 5 summarizes the present studies and does perspectives to the future application of sub-millisecond SMART-EM imaging.

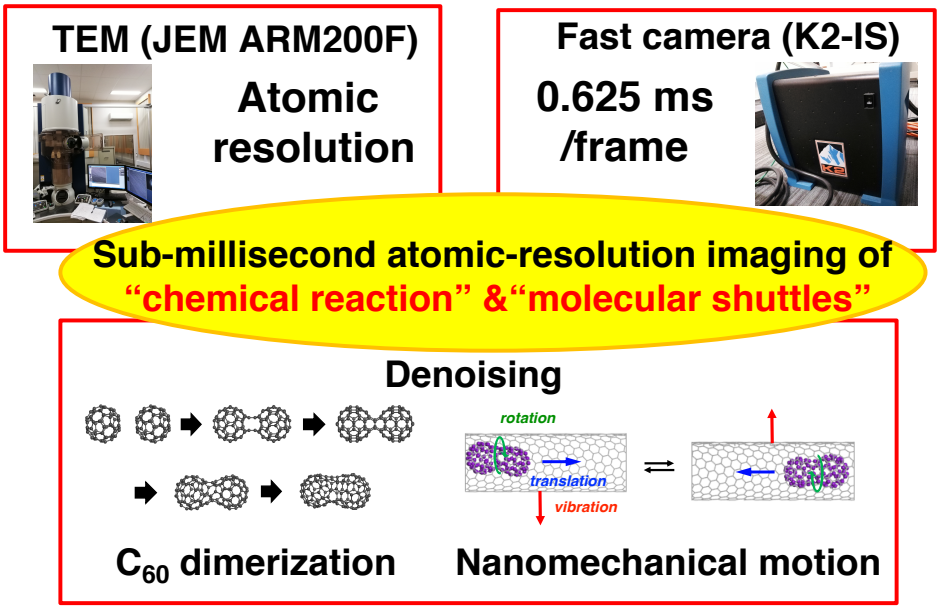

Figure 1.6. Schematic image of this work.

1.10 References

[1] P. Fara, *Nature* **2009**, *459*, 642–644.
[2] http://hdl.loc.gov/loc.pnp/cph.3a45870.
[3] A.H. Zewail, *Ultrafast Light and Electrons: Imaging the Invisible.* In: M. Al-Amri, M. El-Gomati, M. Zubairy (eds) Optics in Our Time. Springer, Cham, **2016**, 43–68.
[4] N. Rubin *et al.*, *Science* **2020**, *369*,1084–1089.
[5] L. Gross, F. Mohn, N. Moli, P. Liljeroth, G. Meyer, *Science* **2009**, *325*, 1110–1114.
[6] K. Iwata, S. Yamazaki, P. Mutombo, P. Hapala, M. Ondracek, P. Jelinek, Y. Sugimoto, *Nat. Commun.* **2015**, *6*, 7766.
[7] K. Nakagawa, A. Iwasaki, Y. Oishi, R. Horisaki, A. Tsukamoto, A. Nakamura, K. Hirosawa, H. Liao, T. Ushida, K. Goda, F. Kannari, I. Sakuma, *Nat. Photon*. **2014**, *8*, 695–700.
[8] A. H. Zewail, J. M. Thomas, *4D Electron Microscopy: Imaging in Space and Time*, Imperial College Press, **2010**.
[9] E. Nakamura, K. Harano, *Proc. Jpn. Acad., Ser. B* **2018**, *94*, 428–440.
[10] S. Okada, S. Kowashi, L. Schweighauser, K. Yamanouchi, K. Harano, E. Nakamura, *J. Am. Chem. Soc.* **2017**, *139*, 18281–18287.
[11] M. Koshino, T. Tanaka, N. Solin, K. Suenaga, H. Isobe, and E. Nakamura, *Science* **2007**, *316*, 853.
[12] N. Uyeda, T. Kobayashi, K. Ishizuka, Y. Fujiyoshi, *Chemica Scripta.***1978/79**, *14*, 47–61.
[13] N. Uyeda, T. Kobayashi, K. Ishizuka, Y. Fujiyoshi, *Nature* **1980**, *285*, 95–97.
[14] T. Ando, *Biophys Rev.* 2018, **10**, 285–292.
[15] J. Dubochet, M. Adrian, J. J. Chang, J. C. Homo, J. Lepault, A. W. McDowall, P. Schultz, *Q. Rev. Biophys*. **1988**, *21*, 129–228.
[16] J. H. Warner, Y. Ito, M. H. Ruemmeli, B. Buechner, H. Shinohara, G. A. D. Briggs, *ACS Nano* **2009**, *3*, 3037–3044.
[17] H.-G. Liao, D. Zherebetskyy, H. Xin, C. Czarnik, P. Ercius, H. Elmlund, M. Pan, L.-W. Wang, H. Zheng, *Science* **2014**, *345*, 916–919.
[18] A. E. Carpenter, *Nature Methods* **2007**, *4*, 120–121.
[19] J. B. Heymann, D. M. Belmap, Journal of Structural Biology **2007**, *157*, 3–18.
[20] Y. Zhu, J. Ciston, B. Zheng, X. Miao, C. Czarnik, Y. Pan, R. Sougrat, Z. Lai, C.-E. Hsiung, K. Yao, I. Pinnau, M. Pan, Y. Han, *Nat. Mater.* **2017**, 16, 532–536.
[21] J. M. Ede, R. Beanland, *Ultramicroscopy* **2019**, *202*, 18–25.

－Chapter 2－

Sub-Millisecond Molecular Imaging Enabled by Denoising

2.1 Introduction

Video recording of molecular events such as chemical reactions and dynamical motions with a single-molecule atomic-resolution real-time electron microscopy (SMART-EM) imaging method has emerged as a new technology for studying a variety of single molecules. By encapsulating molecules inside a single-walled carbon nanotube (CNT) or chemically attaching them on an outer surface of a CNT,[1] we are able to record the time evolution of individual molecular events with transmission electron microscopy (TEM). We have targeted so far chemical reactions,[2] molecular crystal formation,[3] and minute reaction intermediates.[4] These studies guarantee the potential of the SMART-EM imaging method in nanoscience.

As discussed in chapter 1, TEM observation is conducted at the frame rate of 1 or 2 frame per second (fps), causing blurring of TEM images because the molecules are always in motion during observation. Ideally, one can quantitatively study the dynamics of the molecular events by acquiring video images of reacting or moving molecules by using a fast shutter speed camera, which has a high temporal resolution. With the recent development of complementary metal oxide semiconductor (CMOS) cameras, it is possible to take TEM images at the frame rate of as many as 1600 fps.[5] In other words, each frame has information of 0.625 ms. However, their latent potential has not been fully realized because a high frame rate imaging reduces the electron dose per frame, causing a very low signal to noise ratio (SNR) of TEM images. For example, I showed a raw TEM image of [60]fullerene (C_{60}) molecules in a CNT taken at 0.625 ms/frame in Figure 2.1a. Although maximum electron dose rate, threshold the limit value of the detector, was used for imaging, it is impossible to recognize the molecules from this single noisy image. Therefore, no one has ever seen molecular behaviors directly with TEM in the time scale of sub-millisecond.

I hypothesized that the key to break through this wall of sub-millisecond imaging of TEM is denoising. In terms of fast molecular imaging, I cannot apply the widely used contrast enhancement methods like symmetry-imposing and lattice-averaging protocols for static molecular assemblies such as metal-organic frameworks to mobile molecules.[6] Although computational image processing of raw TEM images has been studied,[7] there has been no systematic investigation on denoising algorithms suitable for noisy TEM videos. In chapter 2, I investigated various noise reduction methods with a collaborator, Dr. Joshua Stuckner, and we

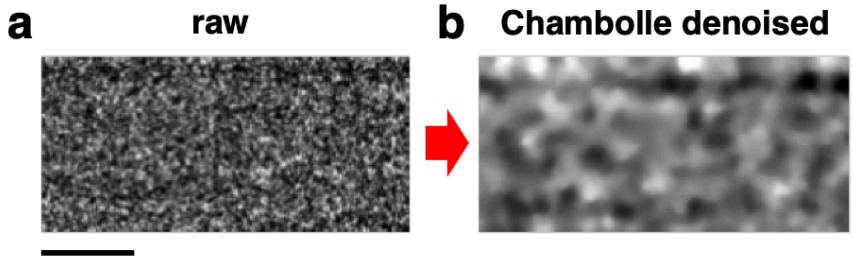

Figure 2.1. TEM images of C_{60} molecules in a CNT (0.625 ms/ frame, 80 kV, electron dose rate (EDR) = 210×10^5 e^- nm^{-2} s^{-1}). (a) A raw TEM image. (b) An identical TEM image with Chambolle denoising. Scale bar: 1 nm. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

succeeded in TEM molecular imaging at the level of sub-millisecond[8] as shown in Figure 2.1b. We concluded that the most appropriate denoising method for TEM molecular imaging is Chambolle total variation denoising algorism.[9] This sub-millisecond SMART-EM imaging techniques will reveal detailed dynamics of single molecules and open up a new field of nanoscience.

2.2 Ideal target molecules for denoising analysis

In this work, C_{60} molecules in a CNT [2] were used as the model of denoising analyses, and the detailed sample preparation methods are summarized in the experimental section. C_{60} molecules encapsulated in a CNT (C_{60}@CNT) were chosen as the target specimen because the hollow and spherical morphology of C_{60} is appropriate for the quantitative evaluation of TEM images by SNR and edge analyses. In addition, it is worth noting that because of its circular shape, I can determine its center position from noisy TEM images, giving its location information. In light of these facts, using C_{60}@CNT for TEM image analysis is the best way to evaluate the effect of denoising for quantitative analyses in SMART-EM.

In general, a single-walled CNT is imaged as two parallel lines and a C_{60} molecule as a hollow circle in TEM observation as shown in Figure 2.2. The simplest way to explain the reason why these molecules look these ways can be drawn as a cartoon in Figure 2.3. The electrons that have passed through the target specimen are greatly affected by its electrostatic potential of atoms.[10] For a spherical C_{60} molecule and a cylindrical CNT, their periphery is darkly emphasized because the TEM image gives a two-dimensional projection of the three-dimensional objects. Therefore, in under-focus condition, a C_{60} molecule can be seen as a black hollow circle and a CNT as two black lines.

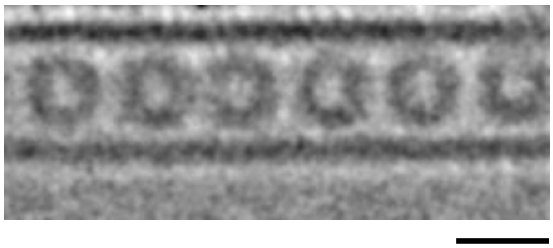

Figure 2.2. A TEM image of C_{60}@CNT (62.5 ms/ frame, 80 kV, EDR = 22×10^5 e^- nm^{-2} s^{-1}). Scale bar: 1 nm.

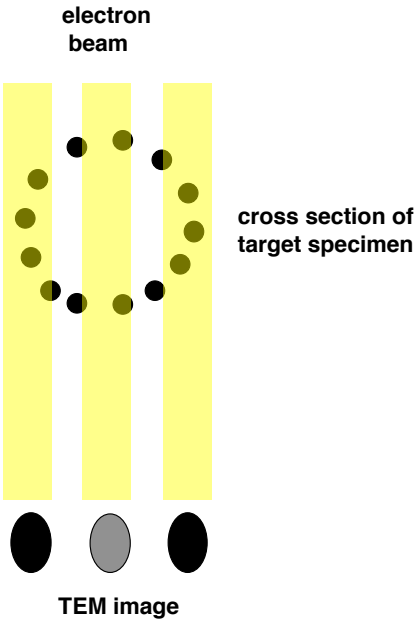

Figure 2.3. A cartoon showing how molecular species are imaged.

2.3 Methods for the evaluation of the effect of denoising in TEM images

2.3.1 SNR calculation

SNR is one of the metrics for image quality. There are many definitions for calculating the SNR depending on spectroscopic methods, although no agreed-upon standard exists. In the medical imaging field,[11] the vacuum area of images is taken as the noise while a region of interest (ROI) is taken as the signal. I incorporated the idea in SMART-EM analysis, in detail, I assume that the region of TEM containing the C_{60} molecules represents the signal as shown in a red square in Figure 2.4. And, the vacuum area in this TEM is defined as the noise shown in a blue square. I herein defined SNR as in the equation 1, and the values can be calculated from intensity values measured from each area. S_{mean} and N_{mean} are the mean pixel intensity in the ROI and vacuum, respectively, and N_{RMS} is the root mean square of pixel intensities in the vacuum. By analyzing these areas and calculating the SNR of TEM images before and after denoising procedures, it is able to qualitatively evaluate the quality of experimental images.

$$\mathrm{SNR} = \frac{|S_{\mathrm{mean}} - N_{\mathrm{mean}}|}{N_{\mathrm{RMS}}} \qquad (1)$$

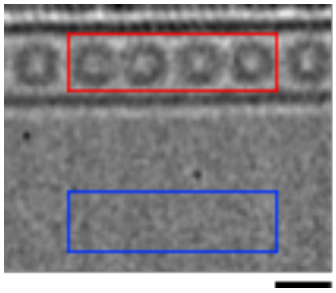

Figure 2.4. A TEM image of C_{60} molecules in CNT (62.5 ms/ frame, 80 kV, EDR = 22 $\times 10^5$ e^- nm^{-2} s^{-1}). Red square: ROI, Blue square: the area of vacuum. Scale bar: 1 nm.

2.3.2 Evaluation of signals preservation

Although SNR can be a metric for judging the qualities of an image, the image with a high SNR does not necessarily show that the signal has been preserved. Applying too much denoising can corrupt the signal in images by severely-reducing edge contrast or changing the apparent shape and size of features while the SNR is increased. Therefore, it is necessary to quantify the signal preservation after applying denoising to TEM images. Ideally, a denoising protocol will remove noise while it fully preserves molecules' features. In this work, the signal preservation was quantified by analyzing feature edges of C_{60} molecules, as shown in Figure 2.5.

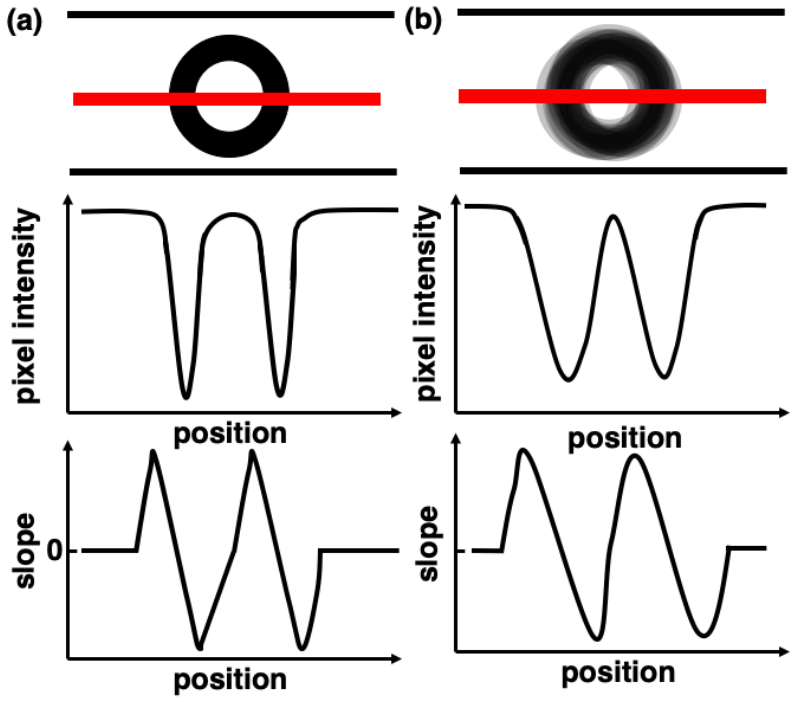

Figure 2.5. Schematic images of analyzing edge preservation of a C_{60} molecule in a CNT. (a) Ideal sharp image of a C_{60}. (b) Denoised and blurred image of a C_{60} (top) with pixel intensity profiles (middle) and slope of intensity or intensity gradient profiles (bottom). Red lines indicate the location of the line intensity profiles.

Edge preservation is judged from the derivatives of the pixel intensity profile. A red line is placed parallel to the CNT crossing along the center of the C_{60} molecules as shown in Figure 2.5. The pixel intensity profile can extract the pixel value along the line. From the point of view of the pixel value in an image, an edge is the location of a sudden change in contrast. Specifically, the edges of the molecules or the CNT can be evaluated by using this method. Denoising can make these edges of a molecule blurred, which means a sharp change of pixel value will become a gradual contrast change as shown in Figure 2.5 (b). The gradual contrast change at an edge makes it difficult to precisely locate the edge of a molecule, greatly reducing the accuracy of size or distance measurement. Therefore, edge sharpness should be also investigated by the slope or gradient of the intensity profile across the molecular edge as listed in Figure 2.5 (bottom). In other words, when the slope of intensity profile is sharp, the molecular edge is preserved well, and hence I can conclude that the denoising method is suitable for the current purpose.

2.4 Denoising algorisms for noisy TEM images

Nine denoising algorithms in total were applied and evaluated on each image in each dataset. The tested denoising algorithms were the mean, median, Gaussian, and bilateral[12] filters; nonlocal means denoising[13]; Wiener–Hunt deconvolution[14]; low rank approximation with singular value decomposition,[15,16] and the Chambolle[9] and Bregmann[17] total variation denoising algorithms. The image quality depends on how much noise and blur there is in the TEM images. While noise is the quasi-random change of individual pixel values, blur is the spreading of a point source of brightness written by the point spread function (PSF). These factors limit the inherent resolution of the microscope. Denoising algorithm basically assumes a certain model as a mathematical equation, which defines the relationship between the obtained image and the ideal image. The ideal image here assumes that it would have been taken with a perfect electron microscope system where aberration effects are removed. By making full use of computers, the denoising algorithms try to find the solution to the assumed ideal image.

The mean filter reduces noise by replacing each pixel value in an image with the average value of its specified neighbors, including itself. The median filter replaces each pixel with the median value of pixels within a specified neighborhood, a technique that is particularly effective against salt-and-pepper noise and speckle noise, which have extreme pixel value pixels. The Gaussian filter outputs a weighted average of each pixel's neighborhood, with the average weighted more towards the value of the central pixels. The weight is determined by a specified sigma of radial distance from the central pixels, considering a normalized Gaussian distribution. Therefore, the Gaussian filter provides gentler smoothing and preserves edges better than mean filter, which considers uniformly weighted average. The bilateral filter uses a weighted averaging method that weights nearby pixels based on not only their distances but also their intensity differences. Unlike other filters which indiscriminately blur edges, this method preserves image feature edges like molecular edges by considering intensity information. The non-local means algorithm takes a mean of all pixels in the image, weighted

by how similar these pixels are to the target pixel, instead of averaging together neighboring pixels. This method works well for denoising images which have specific repeated textures. Deconvolution algorithms such as Wiener–Hunt deconvolution seek to reduce blur. In general, the goal of deconvolution is to find the solution or ideal image f in the following equation (2).

$$(f * g) + \varepsilon = h \tag{2}$$

In the equation, g is blur, ε is noise, and h is an obtained noisy image. However, these methods require prior knowledge of the PSF, the function of g, and perform poorly on noisy images. Some works have been done to estimate the function g from microscopy images, but the work is ongoing.[18,19,20] For example, a numerical Bayesian method, which combines probability models and inference algorisms, is applied to repeatedly estimate the PSF for the Wiener–Hunt filter.[21] If there are no advances in estimating the PSF and in performing deconvolution on noisy images, this method is not expected to perform well. However, it is included in this experiment for comparison with past work on denoising filter against EM images of nanomaterials.[7] In addition, low rank approximation with singular value decomposition has been successful in denoising. The singular value decomposition takes an m x n image matrix **I** (n and m are the number of pixel rows and columns in the image) and decomposes it into $\mathbf{I} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathbf{T}}$. **U** and **V** are orthogonal matrices, where **U** is a m x m matrix and **V** is a n x n matrix. $\mathbf{\Sigma}$ is a diagonal matrix whose entries are called singular values of **I**. The number of non-zero singular values is equal to the rank of **I**. A denoised image can be reconstructed from a low rank matrix which is obtained by reducing the rank of $\mathbf{\Sigma}$. Therefore, solving a minimization problem in mathematics, to find a new **I**, leads to denoising of EM images, which can be thought of as matrices of pixel values.

Total variation denoising methods assume that image noise has sharp intensity difference or variation between neighboring pixels in the image. To obtain the desired output image, minimizing the variation is key. On the other hand, it is also important to minimize the difference between the noisy input image and the output image because the contrast due to the signal, such as molecular edges, can be preserved. Therefore, the total variation solution is formulated as a co-minimization problem as shown in equation (3):

$$u = \min[D(u, g) + \lambda V(g)] \tag{3}$$

where u is the denoised image, g is the noisy image, λ is a weight parameter, D is a difference function, and V is a total variation function. This is a balancing act because if we make no changes then D is 0, but V is very large. If we set all the pixels in the output image to be the same (one shade of gray), then V is 0, but there is a large difference between the output and input image. The weight parameter λ determines the preference for more strongly reducing pixel variation or preserving the original image. To summarize, equation (3) wants to minimize variation in g, while keeping the difference between u and g as small as possible. A common way to determine D is a simple sum of squares as described in equation (4).

$$D(u,g) = \sum_{i=0}^{n}\sum_{j=0}^{m}\left(u_{i,j} - g_{i,j}\right)^2 \tag{4}$$

n and *m* are the number of pixel rows and columns, respectively, and *i* and *j* are the pixel indexes. In terms of variation function, it can be written equation (5). The variation is the sum of the differences between each pixel in the proposed output image and the pixels next to it.

$$V(u) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\sqrt{\left|u_{i+1,j} - u_{i,j}\right|^2 + \left|u_{i,j+1} - u_{i,j}\right|^2} \tag{5}$$

This algorism has been in development since it was first proposed in 1992.[22] However, equation (3) is difficult to solve because it is non-differentiable and an infinite-dimensional minimization problem.[23] Many improvements have also been made to the equations for both D and Vor the minimization method. The Chambolle method uses a projection algorithm based on a dual formulation. On the other hand, the Bregman method uses an operator splitting method. Both methods tackle this problem from different mathematical approaches, giving a solution to obtain denoised images. More details can be found in their respective papers.[9,17]

Prior to applying these denoising methods to noisy TEM images, Gaussian downsampling can be used to resample the images to a proper sampling frequency. According to the Nyquist–Shannon Sampling Theorem,[24] the pixel edge length should be 2.3–3 times smaller than the point-to-point resolution of the electron microscope. Before resampling the images to a larger pixel size, a Gaussian filter is applied with an appropriate kernel size to eliminate image frequencies higher than the Nyquist cutoff frequency of the resampled image. For example, in Figure 2.6, a 6 × 6 image, which is a set of pixel values, is downsampled by a kernel matrix which has a kernel size of 3 to obtain an output 4 × 4 image. The light blue area of the 6 × 6 image is downsampled to one pixel colored in light blue in the 4 × 4 image by applying the kernel matrix. By screening 3 × 3 area as shown in dotted arrows in Figure 2.6 and applying the kernel matrix to each area like a red or green area, the whole images can be downsampled. All denoising algorithms are applied to both the original and downsampled datasets for comparison in this work.

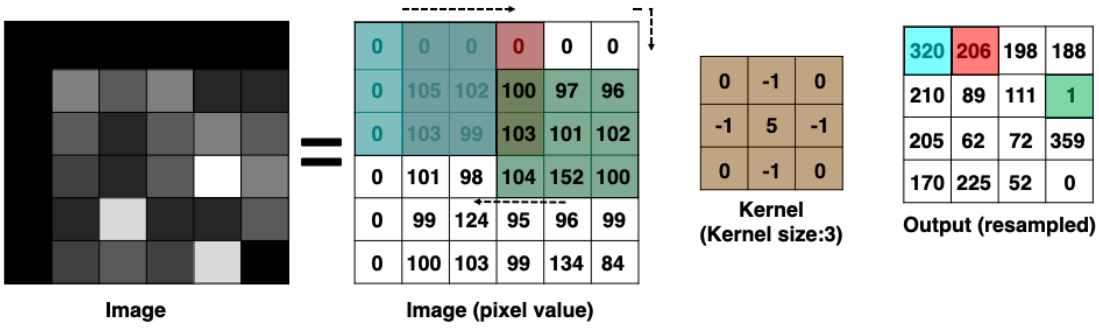

Figure 2.6. Schematic image of downsampling.

2.5 Qualitative analysis of TEM images obtained by various denoising methods

A final goal of this chapter is to find the best denoising algorism for noisy TEM images for sub-millisecond molecular imaging of C_{60}@CNT. To evaluate the performance of the denoising algorithms, the videos of C_{60}@CNT were recorded at an electron dose per image ranging from 0.5×10^5 to 65×10^5 e^- nm^{-2} by Gatan OneView camera. The electron dose per image was changed because it corresponds to exposure time and the quality of TEM images greatly affected by the electron beam irradiation. Qualitative results of applying select denoising algorithms are summarized in Figure 2.7. The Chambolle denoising algorithm applied after downsampling increased the visibility of the C_{60} molecules in CNT even at an extremely low electron dose irradiated condition. A Chambolle weight, λ in the equation (3), of 0.2 seemed best at high electron doses (first image column) and 0.5 seemed best at low electron doses (second image column) by visual inspections. However, these values are dependent on each dataset of TEM, so through this work, optimal weight values are used respectively.

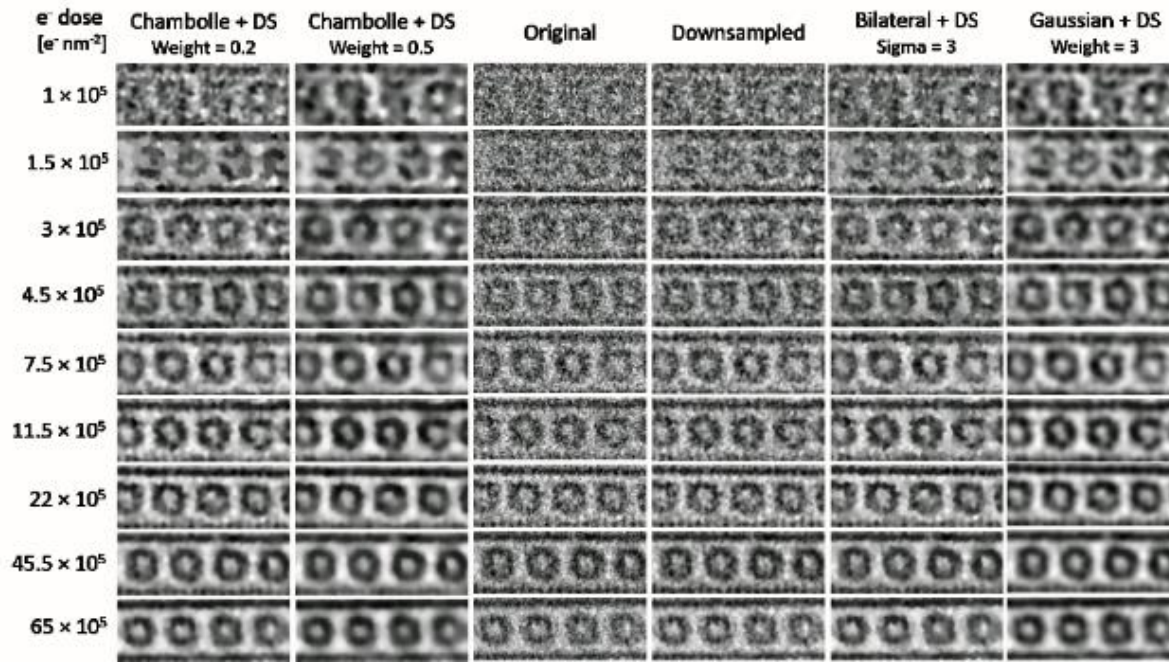

Figure 2.7. TEM images of C_{60}@CNT at various electron dose per frame before denoising (original) and after denoising with downsampling (DS) and the specified algorithm. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

From visual aspect, the effects of downsampling can be evaluated by comparing the third and fourth image columns. The results seem to be insignificant, but the downsampling of TEM images is often an important first step for denoising process. The bilateral filter also increased clarity compared with the original images, but not as well as the Chambolle algorithm. In case of the Gaussian filter, it appeared to remove the most of noise, however, it also blurred the edges of the molecules.

2.6 Quantitative analysis of TEM images obtained by various denoising methods

To compare a performance of each denoising algorithm, the SNR of TEM images denoised by each method was analyzed quantitatively. In each denoising method, the optimum parameter such as weight parameter was applied while considering signal preservation. Figure 2.8 shows the SNR of each select methods when applied with the optimum parameter. Under these conditions, the TEM image denoised by the Chambolle algorithm had the highest SNR, which was nearly six times better than the original image. The Chambolle denoising can make the same SNR of TEM images taken at an electron dose per frame of 1.5×10^5 e^- nm^{-2} as the dataset without any denoising at an electron dose per frame of 65×10^5 e^- nm^{-2}. It means that the Chambolle denoising can increase the SNR of unprocessed TEM images by nearly 1.5 orders of magnitude. The Bregman total variation denoising and Gaussian filter performed nearly as well as the Chambolle by the SNR. In addition, although the bilateral filter only improved the SNR about half as much as Chambolle denoising, the bilateral filter with downsampling still performed similarly at an electron dose per frame of 4.5×10^5 e^- nm^{-2} as

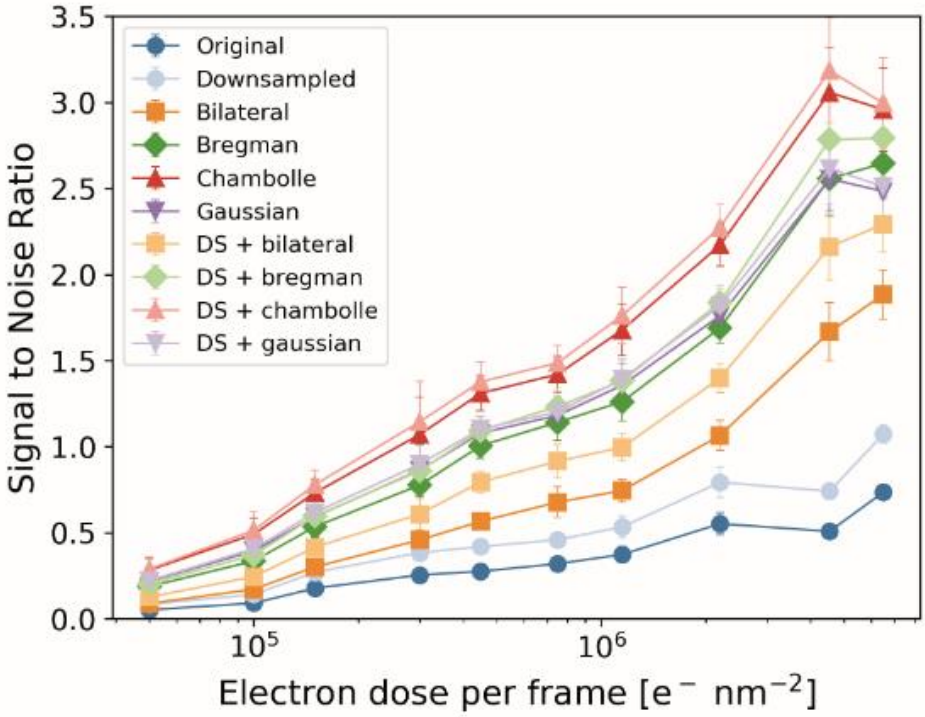

Figure 2.8. SNR of original images and with several applied denoising algorithms at varying electron dose per frame. Denoising was applied with parameters set to remove the most noise while optimally preserving the signal. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

the original images at an electron dose per frame of 65×10^5 e^- nm^{-2}, which was over an order of magnitude reduced electron dose. Further details of SNR analysis against each denoising methods are summarized in Figure 2.9 and 2.10. In these figures, TEM images processed by the strongest parameter setting (upper inset) and the optimum parameter setting are shown. It is quite important to note that the TEM image denoised by too much strong parameters, as exemplified by the Gaussian filter upper inset image gives no structural information of molecules.

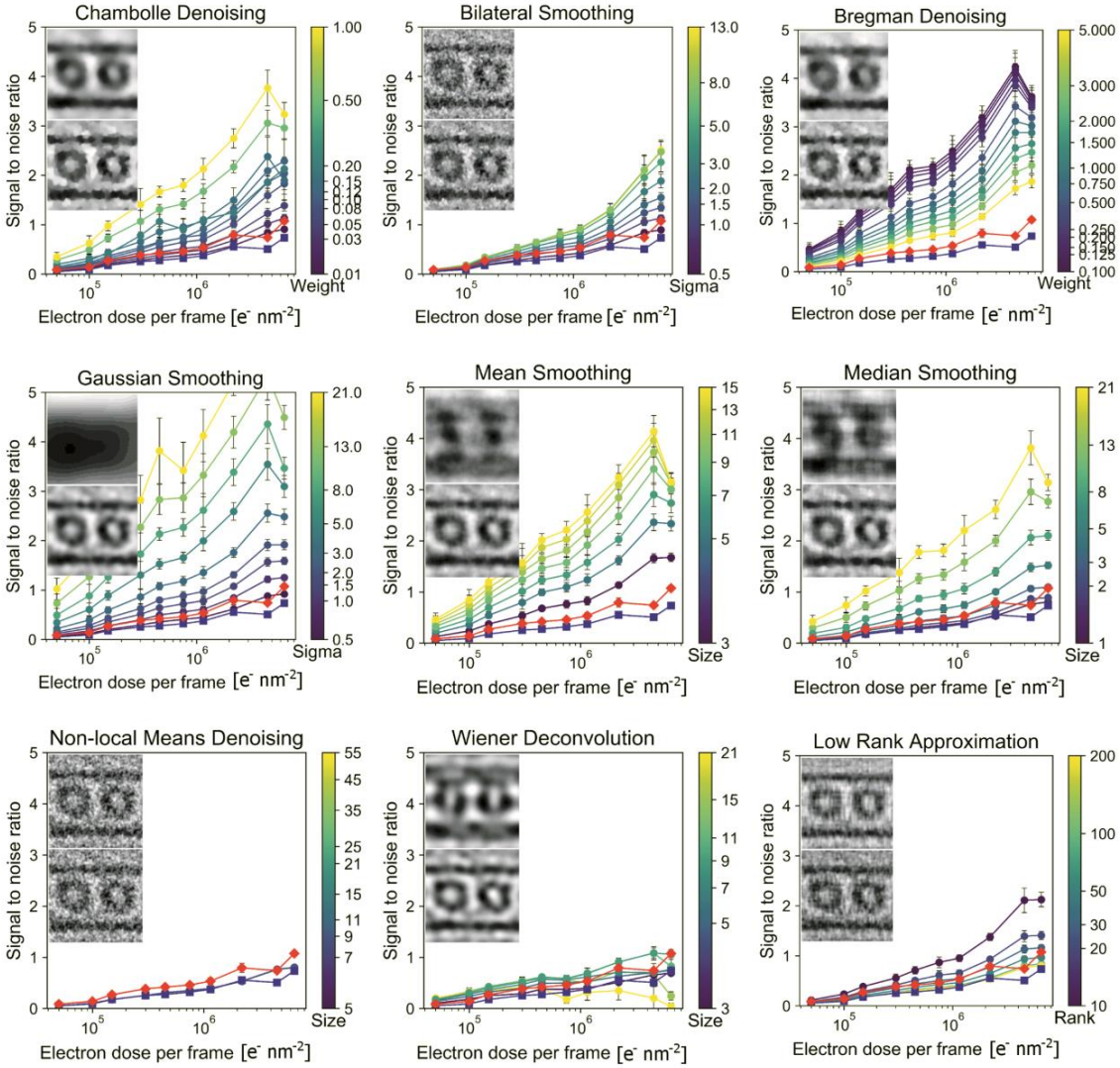

Figure 2.9. SNR plots for TEM images of C_{60}@CNT without downsampling. Unprocessed images shown in blue line with square markers. Downsampling only shown in red line with diamond markers. Blue and red lines are the same in each plot. SNR after denoising is shown by the lines with circle markers. The line color or shade represents the set denoising parameter indicated by the color bar. Upper inset image shows image processed by the strongest parameter setting. Lower inset shows image processed by optimum parameter. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

Although too much denoised images show high SNR, these kinds of images are meaningless in terms of molecular studies, so they were excluded from a point of view of signal preservation. From the same reason, the mean and median filter are also removed from the screening. Detailed signal preservation analysis is discussed in 2.7. By comparing Figures 2.9 and 2.10, it can be concluded that downsampling increased the SNR in all case. In conclusion, Chambolle algorithm is the optimized method for denoising current TEM images among the nine denoising algorithms judging from the quantitative SNR analysis.

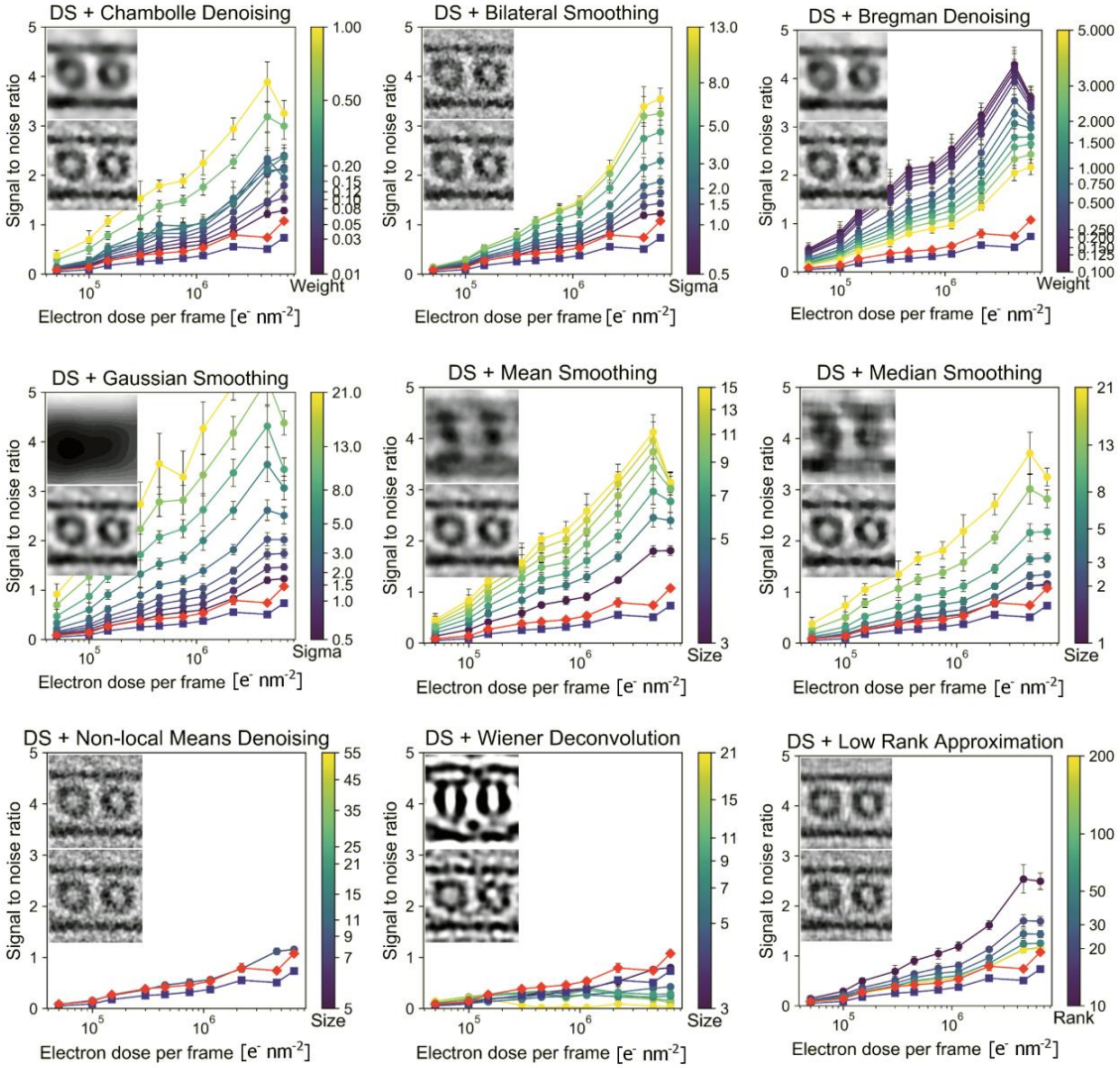

Figure 2.10. SNR plots for TEM images of C_{60}@CNT with downsampling. Unprocessed images shown in blue line with square markers. Downsampling only shown in red line with diamond markers. Blue and red lines are the same in each plot. SNR after denoising is shown by the lines with circle markers. The line color or shade represents the set denoising parameter indicated by the color bar. Upper inset image shows image processed by the strongest parameter setting. Lower inset shows image processed by optimum parameter. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

2.7 Signal preservation analysis for molecular studies

To study molecular events by the SMART-EM imaging methodology, locating the edges precisely is necessary for size or distance measurements. In this work, various denoising algorisms were investigated for noisy TEM images, and I found that too much strong parameter setting in the denoising algorithms corrupted clarity of images, followed by structural information loss. Therefore, it is indispensable to judge how well the denoising methods preserve edges. An edge can be characterized by a pixel intensity gradient shown in Figure 2.5, which I calculated by obtaining the derivative of a pixel intensity profile across an edge as explained in 2.3. When it has a large pixel intensity gradient, the images of molecules are preserved well even after denoising. The results of the pixel intensity profiles for the nine denoising algorithms with and without downsampling are summarized in Figures 2.11 and 2.12. The line intensity profiles were taken from TEM images of C_{60}@CNT (an electron dose per frame of 7.5×10^5 e^- nm^{-2}) denoised under the optimum parameters. The intensity profiles were plotted in blue lines, and their intensity gradients, obtained from the derivative of a pixel intensity profile, were plotted in dashed red lines. While the smooth blue line indicates less noise, a sharp peak of red dotted line represents a good molecular edge. From a comparison between Figure 2.11 and 2.12, it can be concluded that downsampling reduced edge sharpness. This is because downsampling re-bins the signal into larger pixels, causing the loss of precise

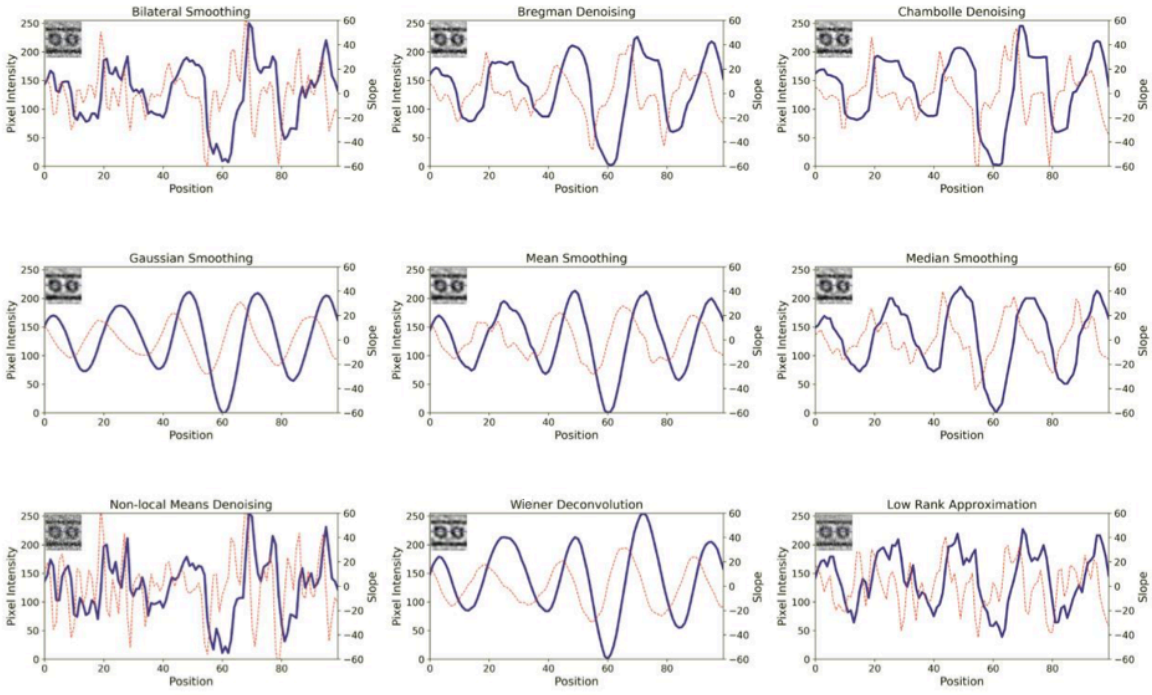

Figure 2.11. Line intensity profiles of TEM images of C_{60}@CNT (an electron dose per frame of 15×10^5 e^- nm^{-2}) after denoising with optimum parameters set. Solid blue lines indicate pixel intensity profile. Dotted red lines indicate intensity gradient. Blue line on inset image indicates location of the line intensity profile. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

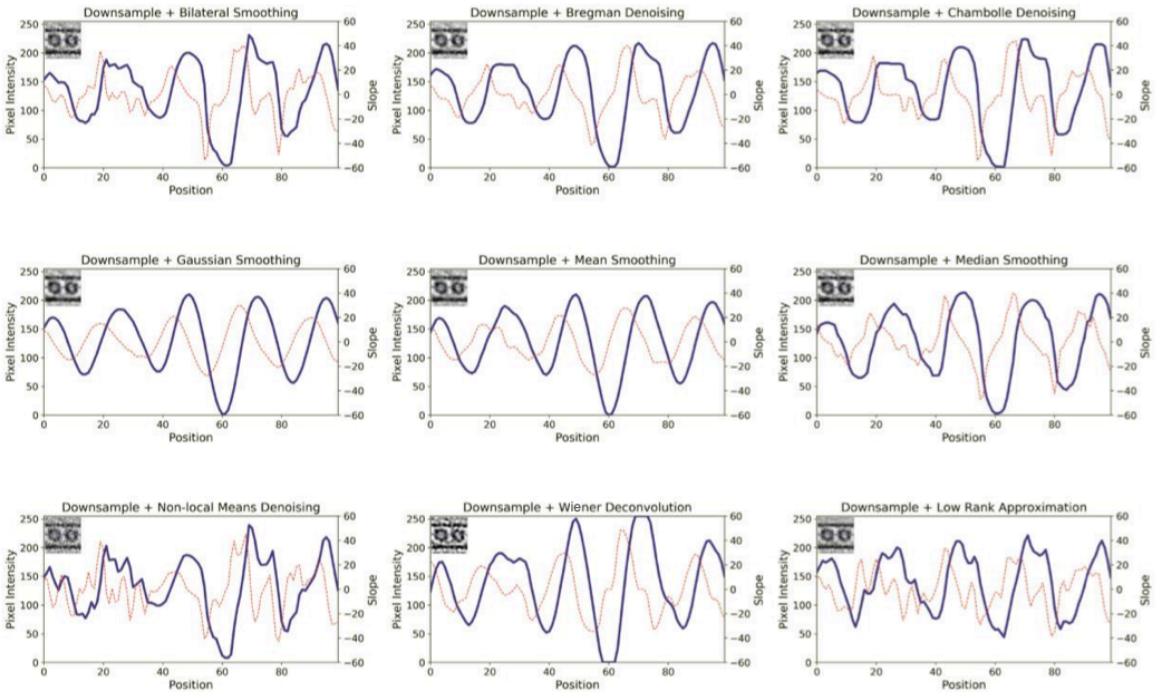

Figure 2.12. Line intensity profiles of TEM images of C_{60}@CNT (an electron dose per frame of 15×10^5 e^- nm^{-2}) after downsampling and denoising with optimum parameters set. Solid blue liens indicate pixel intensity profile. Dotted red lines indicate intensity gradient. Blue line on inset image indicates location of the line intensity profile. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

spatial information. Therefore, from a molecular analysis point of view, downsampling is not appropriate although it makes the SNR of TEM images higher.

The mean, median, and Gaussian filters performed very well against TEM images when measured by their SNR and signal strength. Namely, these denoising algorithms can remove a large amount of noise, resulting in the smooth blue lines. However, these algorithms make blurring points and edges, leading to broad peaks with a very small pixel intensity gradient across the edge in the red dashed lines. Instead of having a sharp edge, the molecular shape gradually fades into the background, making it difficult to perform high accuracy measurements of the molecule's size or position. By these denoising algorithms, important information of the signal such as the size and shape of the molecule was corrupted. Therefore, mean, median, and Gaussian filters are not appropriate denoising methods for molecular images. The profile of Wiener–Hunt deconvolution also shows the similar result with these filters. The rough intensity profiles of the nonlocal means denoising and low rank approximation with singular value decomposition along with small derivative magnitudes show that these methods are not effective in noise removal or edge preservation. On the other hand, the bilateral filter, having been designed specifically for edge preservation, performed the best in this regard. The derivative magnitudes of bilateral filtering are 2–3 times larger when comparing with the result of Gaussian filter. In the end, the Chambolle and Bregmann total variation denoising algorithms performed almost as well as the bilateral filter in terms of edge preservation.

2.8 Chambolle denoising for high frame rate TEM images

From the qualitative and quantitative analysis of TEM images denoised by various denoising algorithms, I found that the Chambolle denoising produced the optimal balance of noise removal and signal preservation against the noisy TEM images. Gaussian downsampling is an important first step if the data are sampled with a higher frequency than the ideal Nyquist sampling rate. Although it was shown to increase the SNR when applied after denoising, it also decreases edge sharpness. Therefore, only applying Chambolle denoising against the raw TEM images seems to be the best for molecular imaging.

Taking all of these things into consideration, I challenged to apply this method to high frame rate TEM images for sub-millisecond molecular imaging. By using Gatan K2-IS direct electron detection camera, I succeeded in recording TEM videos at 1600 fps or 0.625 ms/frame at the detector-tolerable limit of EDR of 213×10^5 e^- nm^{-2}. C_{60}@CNT dataset was used to judge the performance of the denoising algorithms on high frame rate, low-dose TEM images. The TEM images or videos captured at such a high frame rate reduce an electron dose per frame to only 0.133×10^5 e^- nm^{-2}, resulting in a very low SNR due to the short exposure time for each frame. The conventional method of improving the SNR is to superimpose consecutive frames by averaging each pixel value between frames as explained in Figure 2.2. However, such superimposition reduces the time resolution of the dataset by a factor equal to the number of frames stacked. In addition, simple superimposition causes motion blur if a target molecule has moved within the superimposed frames. It has been proven that Chambolle denoising can improve the SNR of each frame, but the algorithm can also show its power in terms of superimposition. When Chambolle superimposition is applied to TEM images, only pixels that are sufficiently similar between frames are combined, minimizing motion-based blur of molecules. Figure 2.13 compares the SNR after applying superimposition and Chambolle denoising to 1600 fps TEM images of C_{60}@CNT. The red curve represents the SNR increase by stacking, obtained by the typical averaging method to the dataset without any denoising. The SNR of 100 frames stacked image increased from 0.1 to 0.6 with a 100-fold decrease in time resolution. For comparison, the SNR of TEM images taken by the OneView camera at varying electron doses shown in Figure 2.8 is plotted in blue with square markers. The SNR difference between images taken by the OneView camera and K2-IS may be derived from the performance of camera detectors. It is important to note that the higher the electron dose per frame gives the higher SNR. This result means that increasing the electron dose and superimposing the images are almost the same effects to the value of SNR. On the other hand, applying Chambolle denoising obviously improves the SNR of TEM images. For example, Chambolle denoising at a weight of 0.5 without any stacking produced an SNR of 0.49 with no loss of time resolution. Using the superimposition method, 68 frames stacking of 0.625 ms/frame images were needed to generate an SNR of 0.5 showing a nearly two orders of magnitude improvement in time resolution at the same SNR by using the Chambolle algorithm. Without stacking, Chambolle denoising using a weight parameter of 0.5 produces an SNR equivalent to the OneView images obtained at a nearly equivalent EDR of 44×10^5 e^- nm^{-2}. Furthermore, while the solid curves with circle markers show the effects of simultaneous Chambolle denoising and Chambolle

superimposition, the dotted curves represent the results of superimposition by the averaging method followed by Chambolle denoising of the superimposed frames. Although the SNR of TEM images obtained by the average superimposition followed by Chambolle denoising is the same or higher value, Chambolle superimposition is recommended when it is possible because it generally preserves temporal changes of signals that would be lost by average superimposition.

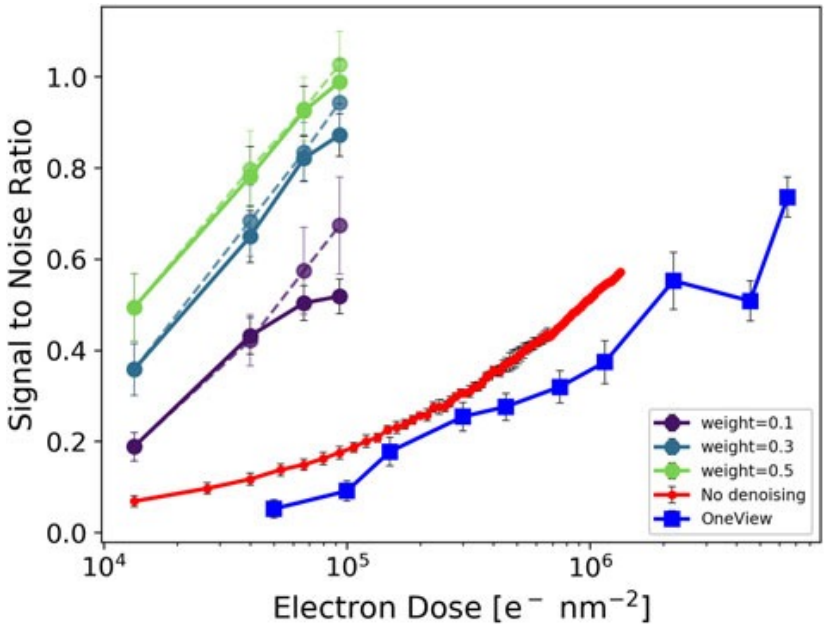

Figure 2.13. Plot showing the improvement of SNR of C_{60}@CNT TEM image by frame stacking. Solid lines with circle markers were simultaneously stacked and denoised by the Chambolle algorithm. Dashed lines with circle markers were stacked by the average stacking method, and finally denoised by the Chambolle algorithm. The red line with small circle markers shows the improvement of SNR by using average stacking at the expense of time resolution. For comparison, the SNR of OneView images at varying electron dose are plotted in blue with square markers. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

2.9 Sub-millisecond molecular imaging enabled by Chambolle denosing

The efficacy of the Chambolle denoising and Chambolle superimposition method for improving the image quality of high frame rate TEM videos is obvious from a visual aspect as summarized in Figure 2.14. Without any image process (Figure 2.14a), a single-frame original TEM image of C_{60}@CNT taken at 0.625 ms/frame with a K2-IS camera appear to be full of noise, giving a low SNR of 0.05. It is difficult to recognize molecules from this raw image. It is only after superimposing 50 frames of consecutive 0.625 ms/frame images by a pixel averaging method that the molecules become clear enough to detect (Figure 2.14b, SNR = 0.20). However, information of faster dynamic phenomena of molecules is lost because 50-frame

stacking is equivalent to an image taken with an exposure time of 31.25 ms. In other words, because the SNR value is proportional to the square root of the exposure time, the time resolution is inversely proportional to the exposure time. To obtain maximum information from the videos, Chambolle denoising was applied to the image to obtain the denoised image shown in Figure 2.14c. It has a SNR of 0.15, which is comparable to that of the 50-frame superimposition without denoising. Therefore, I can insist that we have succeeded in molecular imaging at a sub-millisecond level. In addition, by using simultaneous 3-frame Chambolle superposition and denoising, I could obtain the further improved image (Figure 2.14d, SNR = 0.30), with an exposure time of 1.875 ms. I successfully improved the SNR value from 0.20 in Figure 2.14b to 0.30 in Figure 2.14d with 1/17 of the exposure time. Although it is possible to gain SNR by more superimposition or longer exposure time, considering a final goal to study fast molecular events, temporal resolution should be prioritized, and the image quality obtained by this denoising method is clear enough to analyze molecular positions or their distances.

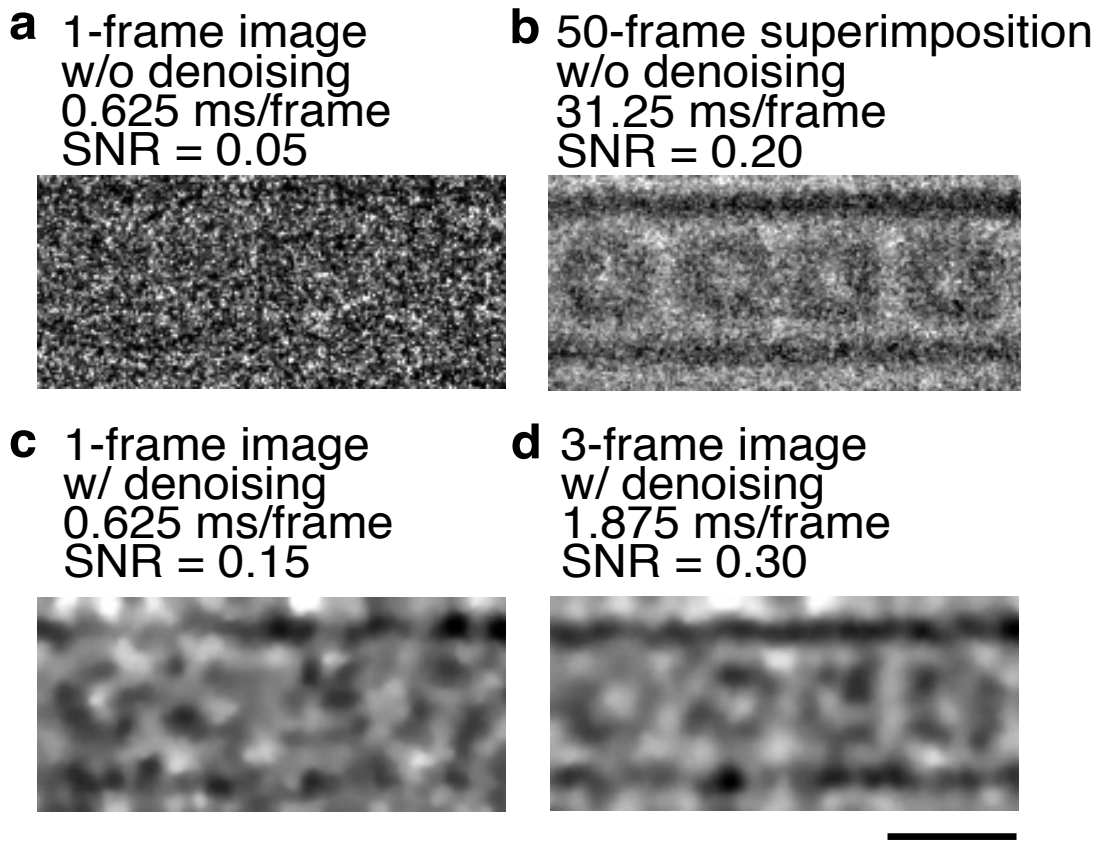

Figure 2.14. Fast video imaging of C_{60}@CNT with Chambolle denoising and superimposition. (a) A single-frame original image of C_{60}@CNT taken at 0.625 ms/frame with K2-IS camera at an EDR of 210×10^5 e^- nm^{-2} s^{-1} at 423 K. (b) Fifty-frame superimposition without denoising. (c) Chambolle denoised single-frame image of the molecule (a weight parameter of 0.5) shown in (a). (d) 3-frame Chambolle superimposition and denoised image (1.875 ms exposure). Scale bar: 1 nm. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

2.10 Conclusion

Atomic resolution electron microscopy makes us possible to conduct SMART-EM imaging with a high spatial resolution, while high shutter speed cameras have improved the time resolution to the sub-millisecond range. However, due to the high frame rate, the obtained images are noisy and have low SNR, making us difficult to study molecular events. Therefore, in this work, we examined various methods to denoise the TEM images and show that the Chambolle total variation denoising algorithm can demonstrate the optimal balance of noise removal and signal preservation. Considering optimal weight parameter settings, it produced denoised images with the highest SNR and was the next best, comparable to the bilateral filter, in edge preservation. More importantly, the Chambolle denoising revealed to be effective in denoising TEM images taken by a high shutter speed camera. I succeeded in obtaining molecular images at the level of sub-millisecond, which nobody has ever seen before in TEM. In addition, the Chambolle superimposition allowed images to be stacked with less motion blur than a simple averaging method. As a result, the Chambolle total variation denoising algorithm can produce TEM images with equal or better SNR while preserving morphology features of molecules, even when an electron dose rate is reduced by more than an order of magnitude compared with the images without any process. This methodology leads us to further studies of molecular events at a millisecond level, which are described in Chapters 3 and 4.

2.11 Experimental section

2.11.1 General

Unless otherwise noted, single-walled CNTs (Meijo Arc SO, produced by arc-discharge using Ni and Y catalysts, >90% purity, average diameter 1.4 nm, Lot # 6601316) were purchased from Meijo Nano Carbon Co. Ltd. C_{60} powder (nanom purple SUH, >99.9% purity, sublimed) was purchased from Frontier Carbon Corporation. TEM grids precoated with a lacy carbon (NS-C15 for experiments at 298 K, pore size 1.5–8 μm and carbon thickness 15 nm: RO-C15 for experiments above 298 K; pore size 3–8 μm and carbon thickness 70 nm) were purchased from Okenshoji Co., Ltd. Toluene (special grade) was purchased from Wako Pure Chemical Industries. Bath sonication was carried out on a Honda Electronics WT-200-M instrument.

2.11.2 Preparation of C_{60}@CNT samples

C_{60}@CNT samples were prepared as follows: CNT powder was heated in air in an oven gradually from 296 to 793 K for 12 min, kept at 793 K for 1 min, heated from 793 to 823 K for 20 min, and kept at 823 K for 20 min to remove the terminal caps of CNTs oxidatively. For encapsulation of C_{60} molecules, the opened CNTs (0.2 mg) and C_{60} powder (0.2 mg) were sealed in a glass tube (Pyrex φ6 mm) under a pressure of 2×10^{-4} Pa and gradually heated from 296 to 573 K over 1 h, then to 673 K over 1 h, and kept at 673 K for 72 h. The resulting C_{60}-containing CNTs were separated mechanically from remaining C_{60} powder, washed with toluene to remove C_{60} from the surface, and dried in vacuum. C_{60}@CNT thus obtained was a

black solid (0.3 mg). I dispersed the C_{60}@CNT in toluene (0.05 mg/mL) in a vial in a bath sonicator for 1 h to soften the sample, so that I could secure good contact between the CNTs and the carbon surface of the grid (essential for temperature control).[2] A 10-μL solution of the dispersion was deposited on a copper grid mesh with a lacy carbon (NS-C15, Okenshoji Co., Ltd.) placed on a paper that absorbs excess toluene. The TEM grid was dried in vacuum (60 Pa) to remove solvent for 2 h. To ensure reproducibility, we used the same sample grid of C_{60}@CNTs in a series of experiments.

2.11.3 SMART-EM imaging

Atomic-resolution TEM observation was carried out on a JEOL JEM-ARM200F TEM equipped with a spherical aberration corrector for imaging, Gatan OneView and K2-IS cameras, and at an acceleration voltage of 80 kV, under 1×10^{-5} Pa in the specimen chamber. Experiments were carried out on a double-tilt holder (JEOL EM-01030RSTH). To remove volatile impurities from the specimen, the holder was heated at 573 K for 30–60 min without electron irradiation before setting a desired temperature. After the stage temperature settled to the target value, we waited for an additional 30 min to minimize thermal drift.

The C_{60}@CNT datasets were imaged using the OneView camera under the varying EDR. The raw images were 2,048 × 2,048 pixels (binning 2 mode) with a 32-bit depth and a pixel edge length of 0.0213 nm at 1,000,000× magnification. Ten images were taken at ten different EDRs ranging from 1×10^5 to 130×10^5 electrons nm^{-2} s^{-1} for a total of 100 images. For another dataset, shown in 2.11.4, ten images were taken at 18 different EDRs ranging from 1×10^5 to 123×10^5 e^- nm^{-2} s^{-1} for a total of 180 images. Each image was taken with an exposure time of 0.5 s (2 fps). The C_{60}@CNT dataset with a high frame rate was recorded by the Gatan K2-IS direct electron detection camera. The raw images were 414 × 1,920 pixels with a pixel edge length of 0.021 nm at 400,000× magnification. Each image was taken with an exposure time of 0.000625 s (1,600 fps). The OneView camera uses a scintillator to convert electrons to photons during the acquisition process, which introduces an intrinsic convolution and blurs the resulting image. The K2-IS uses direct electron detection to produce clearer images.

2.11.4 Quantitative analysis of another dataset of C_{60}@CNT

For supporting data, we conducted quantitative analysis against another dataset of C_{60}@CNT. SNR plots (Figure 2.15 and 2.16) and line intensities profile (Figure 2.17 and 2.18) are obtained from the analysis. These all data show similar results with Figure 2.11, 2.12, 2.13, 2.14, respectively, indicating that our experiments are reproducible.

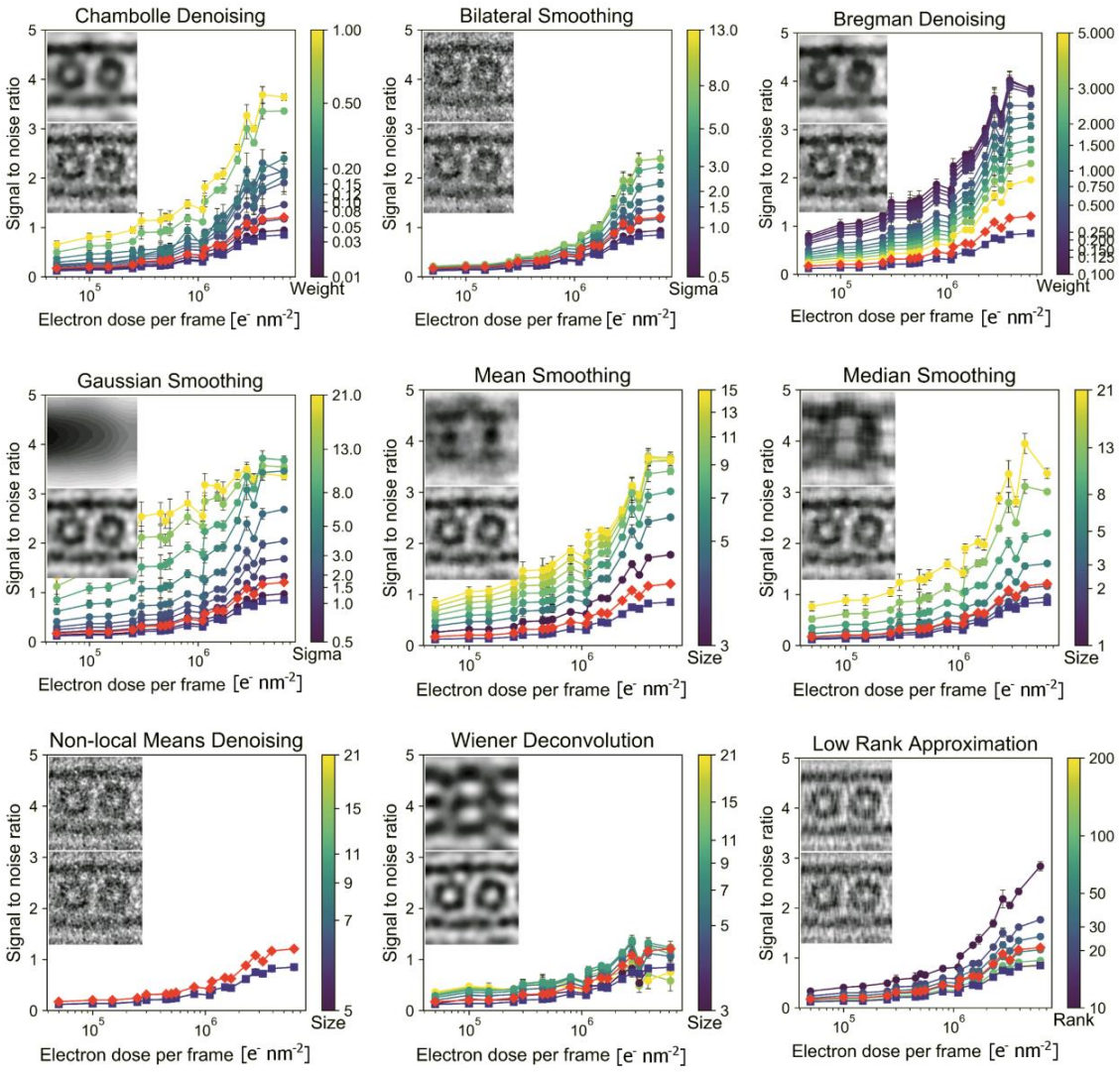

Figure 2.15. SNR plots for another dataset of C_{60}@CNT without downsampling. Unprocessed images shown in blue line with square markers. Downsampling only shown in red line with diamond markers. Blue and red lines are the same in each plot. SNR after denoising is shown by the lines with circle markers. The line color or shade represents the set denoising parameter indicated by the color bar. Upper inset image shows image processed by the strongest parameter setting. Lower inset shows image processed by optimum parameter. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

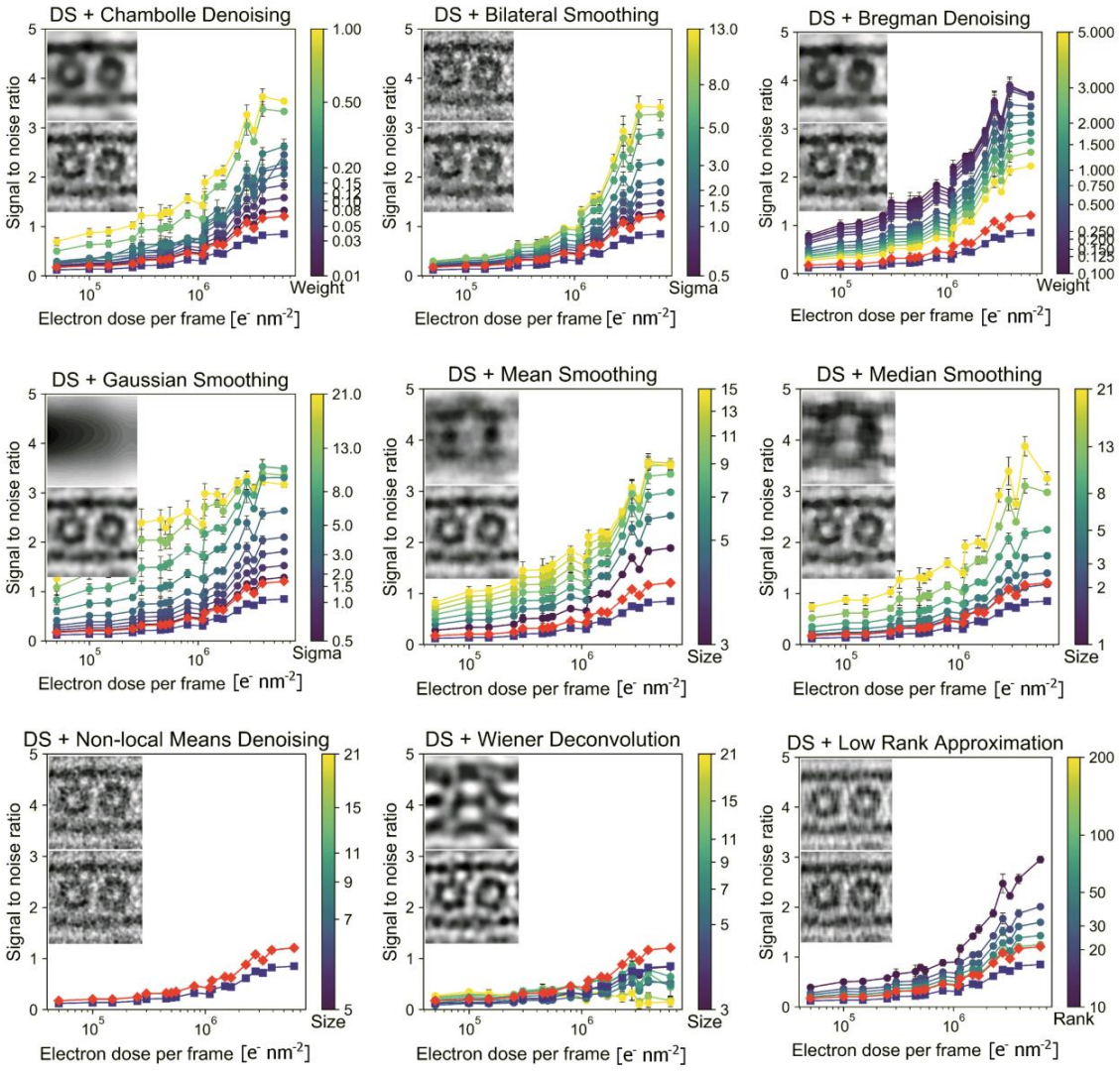

Figure 2.16. SNR plots for another dataset of C_{60}@CNT with downsampling. Unprocessed images shown in blue line with square markers. Downsampling only shown in red line with diamond markers. Blue and red lines are the same in each plot. SNR after denoising is shown by the lines with circle markers. The line color or shade represents the set denoising parameter indicated by the color bar. Upper inset image shows image processed by the strongest parameter setting. Lower inset shows image processed by optimum parameter. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

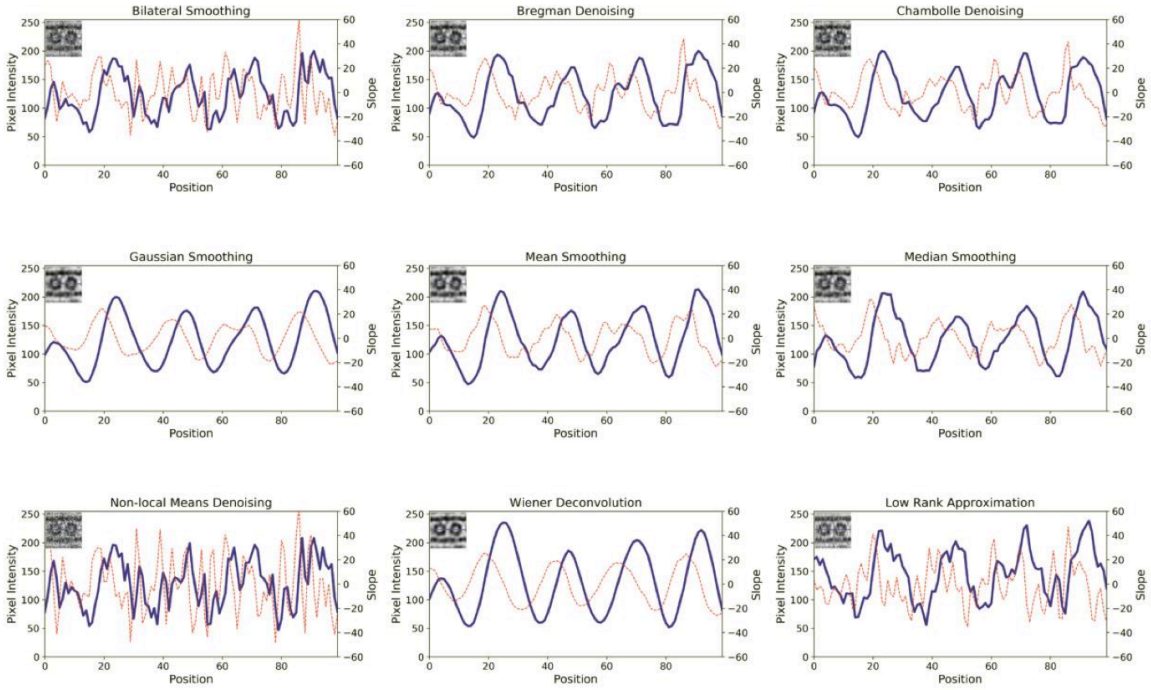

Figure 2.17. Line intensity profiles of another dataset of C_{60}@CNT (an electron dose per frame of 10×10^5 e^- nm^{-2}) after denoising with optimum parameters set. Solid blue lines indicate pixel intensity profile. Dotted red lines indicate intensity gradient. Blue line on inset image indicates location of the line intensity profile. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

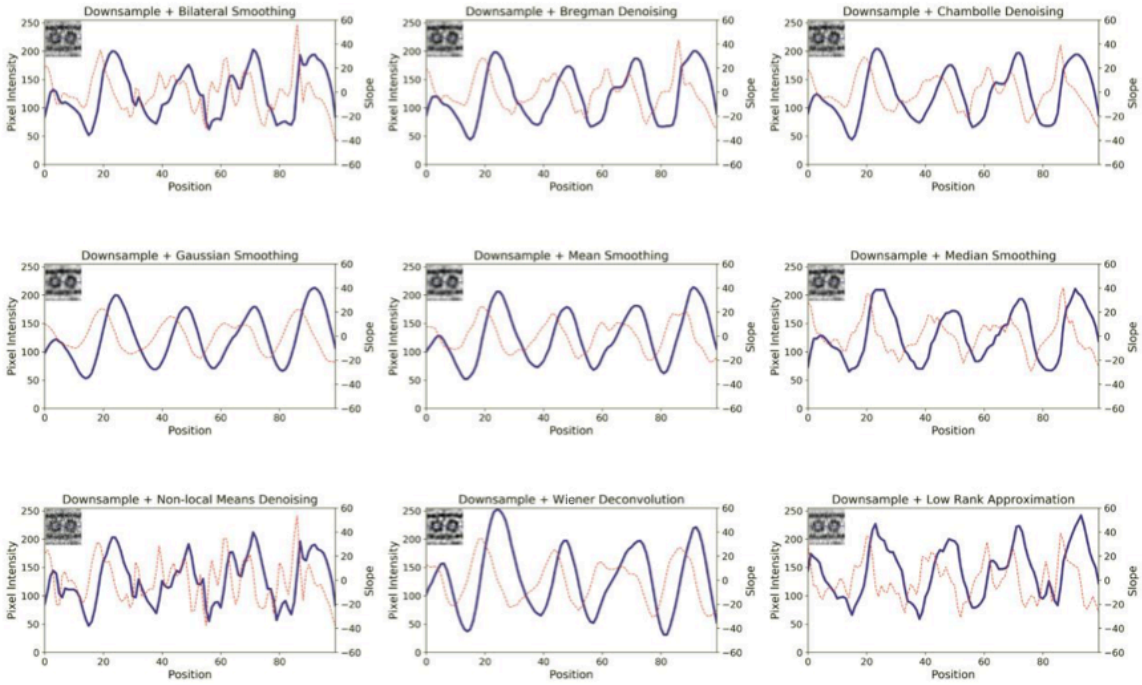

Figure 2.18. Line intensity profiles of another dataset of C_{60}@CNT (an electron dose per frame of 10×10^5 e^- nm^{-2}) after downsampling and denoising with optimum parameters set. Solid blue liens indicate pixel intensity profile. Dotted red lines indicate intensity gradient. Blue line on inset image indicates location of the line intensity profile. Adapted with permission from ref. 8. Copyright 2020 Cambridge University Press.

2.12 References

[1] E. Nakamura, *Acc. Chem. Res.* **2017**, *50*, 1281–1292.
[2] S. Okada, S. Kowashi, L. Schweighauser, K. Yamanouchi, K. Harano, E. Nakamura, *J. Am. Chem. Soc.* **2017**, *139*, 18281–18287.
[3] K. Harano, T. Homma, Y. Niimi, M. Koshino, K. Suenaga, L. Leibler, E. Nakamura, *Nat. Mater.* **2012**, *11*, 877–881.
[4] J. Xing, L. Schweighauser, S. Okada, K. Harano, E. Nakamura, *Nat. Commun.* **2019**, *10*, 3608.
[5] H.-G. Liao, D. Zherebetskyy, H. Xin, C. Czarnik, P. Ercius, H. Elmlund, M. Pan, L.-W. Wang, H. Zheng, *Science* **2014**, *345*, 916–919.
[6] Y. Zhu, J. Ciston, B. Zheng, X. Miao, C. Czarnik, Y. Pan, R. Sougrat, Z. Lai, C.-E. Hsiung, K. Yao, I. Pinnau, M. Pan, Y. Han, *Nat. Mater.* **2017**, 16, 532–536.
[7] H. S. Kushwaha, S. Tanwar, K. S. Rathore, S. Srivastava, *In 2012 Second International Conference on Advanced Computing & Communication Technologies*, 276–281. Los Angeles, California, USA: IEEE.
[8] J. Stuckner, T. Shimizu, K. Harano, E. Nakamura, M. Murayama, *Microsc. Microanal.* **2020**, *26*, 667–675.
[9] A. Chambolle, *J. Math. Imaging Vis*. **2004**, *20*, 89–97.
[10] E. Nakamura, M. Koshino, T. Tanaka, Y. Niimi, K. Harano, Y. Nakamura, H. Isobe, *J. Am. Chem. Soc.* **2008**, *130*, 7808–7809.
[11] O. Dietrich, J. G. Raya, S. B. Reeder, M. F. Reiser, S. O. Schoenberg, *J. Magn. Reson. Imaging* **2007**, *26*, 375–385.
[12] C. Tomasi, R. Manduchi, *IEEE Cat.*, **1998**, *98CH36271*, Bombay, India.
[13] A. Buades, B. Coll, J.-M. Morel, *In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, **2005**, *2*, 60–65, San Diego, California, USA: IEEE.
[14] B. Hunt, *IEEE Trans. Audio Electroacoust.*, **1971**, *19*, 285–288.
[15] S. Wold, K. Esbensen, P. Geladi, *Chemon. Intell. Lab. Syst.* **1987**, *2*, 37–52.
[16] Z. Lin, M. Chen, Y. Ma, *arXiv* **2010**, *preprint arXiv:1009.5055*.
[17] S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Yin, *Multiscale Model. Simul.* **2005**, *4*, 460–489
[18] G. Liu, S. Yousefi, Z. Zhi, R. K. Wang, *Opt. Express.* **2011**, *19*, 18135–18148.
[19] C. Dalitz, R. Pohle-Frohlich, T. Michalk, *IEEE Trans. Ultrason. Ferroelctr. Freq. Control.* **2007**, *62*, 531–544.
[20] J. Roels, J. Aelterman, J. De Vylder, H. Luong, Y. Saeys, W. Philips, *In 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, **2016**, 443–447, Orlando, Florida, USA: IEEE.
[21] F. Orieux, J.-F. Giovannelli, T. Rodet, *J.O.S.A A* **2010**, *27*, 1593–1607.
[22] L. I. Rudin, S. Osher, E. Fatemi, *Physica D* **1992**, *60*, 259–268.
[23] J. Duran, B. Coll, C. Sbert, *Image process On Line* **2013**, *3*, 311–331.
[24] A. J. Jerri, *Proc. IEEE* **1977**, *65*, 1565–1596.

－Chapter 3－

第３章については, 5 年以内に雑誌等で刊行予定のため，非公開.

本章については, 5 年以内に雑誌等で刊行予定のため，非公開.

－Chapter 4－

Real-Time Video Imaging of Mechanical Motions of a Single Molecular Shuttle

4.1 Introduction

Molecular machines are a very hot topic, as is the theme of the 2016 Nobel Prize in Chemistry.[1] A molecular shuttle is a typical molecular machine studied for a long time, which was proposed by Stoddart and his co-workers.[2] Like the model of a molecular shuttle shown in Figure 4.1, the macrocycle (purple ring) moves between two stations (e.g. [2] rotaxane). In the case of [2] rotaxane, it is believed that the activation energy of a single process of shuttling back and forth is 13 kcal/mol from ^{1}H NMR analysis as an average over numerous molecules. However, no one has ever directly observed the individual migration processes of the molecular shuttle in situ. Thus, single-molecule mechanics of molecular machines[3] is still an unexplored field of chemistry.[4] In other words, a system, where mechanical motions at a single molecule level are strongly coupled with fluctuations of its environment, is rarely studied quantitatively because a suitable way to observe the nanomechanical motions in real time is lacking. More importantly, the mechanism of the molecular shuttling in nanoscale from a perspective of molecular dynamics has been unknown because this kind of mechanics is stochastic and even non-repeatable.

Here I elucidated the mechanism of a nanomechanical behavior of a molecular shuttle by a recording of the 1600 fps video of a single molecular shuttle, enabled by the sub-millisecond SMART-EM imaging using Chambolle total variation denoising discussed in Chapter 2. C_{60} encapsulated in CNT reacted to each other under an electron beam irradiation to form a C_{60} dimer or oligomer, as I mentioned in Chapter 3. Due to the reaction, the distance between each C_{60} molecule got shorter and some void spaces were made in CNT, which would be suitable for observing dynamic behaviors. Then, I found a C_{60} dimer or oligomer moving back and forth in CNT during TEM observation. It means that we can intentionally make a molecular shuttle system to capture a decisive moment of the molecular motion at the nano-confined space. In general, molecular shuttling is initiated by external stimuli such as pH,[5] light,[6] and ions.[7] However, it was not clear in spite of its long history of research that why C_{60} molecules in CNT translate or rotate in CNT. According to few reports,[8,9,10] it was hypothesized that the motions of molecules or atoms in CNT were driven by electron beam

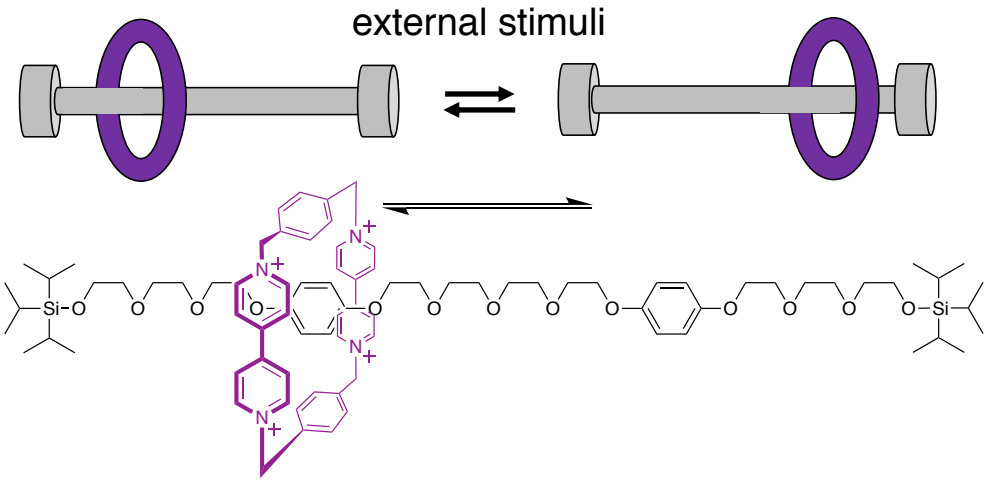

Figure 4.1. Model of molecular shuttle (top) and chemical structure of [2] rotaxane with a station compound (bottom).

irradiation or thermal energy, but there was no direct evidence. It is important to note that these studied were conducted by TEM with time resolution in seconds. Therefore, I herein hypothesized that time-resolved information of molecular dynamics is a key to analyze the detailed mechanism behind the molecular motion and discover in nanoscience.

In this work, I have revealed molecular dynamics at the single molecule level where the molecular motion is coupled with the mechanical motion of the CNT container as illustrated for a shuttling C_{60} dimer in Figure 4.2,[11] previously undetected by time-averaged measurements or microscopy. The sub-millisecond SMART-EM imaging technique opens a new dimension in the study of stochastic mechanics of a single molecule.[12]

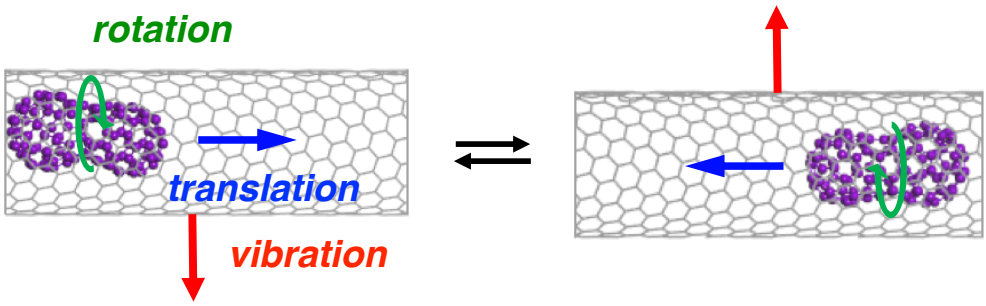

Figure 4.2. A rigid body model showing motions of a C_{60} dimer shown in purple coupled with a vibrating CNT. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

4.2 SMART-EM imaging of a single molecular shuttle

To capture the decisive moment of molecular motion by a camera with sub-milli seconds time resolution, the molecular shuttle system should be constructed intentionally. By using the knowledge gained from the reaction analysis of C_{60} dimer discussed in Chapter 3, I hypothesized that it would be possible to obtain molecular shuttling videos. In particular, as shown in Figure 4.3, during TEM observation of C_{60} molecules encapsulated in CNT, they start to react each other, making the distance between each C_{60} molecule shorter. After several seconds of electron beam irradiation, some void spaces were made in a CNT (red dashed square in Figure 4.3), allowing for stochastic molecular shuttling. Through TEM observation of roughly 400 C_{60} molecules in CNTs for a total observation time of 10 minutes at 0.625 ms/frame, equivalent to 380 million images of C_{60} molecules in total, I found a few tens translation events occurring, out of which four events captured on two molecules (a dimer C_{60} case analyzed in section 4.4 and an C_{60} oligomer case analyzed in section 4.5) were suitable for precise analysis. In other cases, overlapping with neighboring CNTs or severe CNT vibration made it difficult to conduct quantitative analyses at a molecular level. Figure 4.4 shows a molecular motion of a C_{60} dimer at a frame rate of 78.125 ms/frame or 125 stacks of 0.625 ms/frame images, revealing that the dimer (**6**–**7**, colored in light blue) translated in frame ii to the right side in the image. The frames are referenced to an arbitrary time zero, which is 6 s after the start of the TEM observation. Although these images have a low temporal resolution, lattice patterns of

CNT and clear edges of C_{60} molecules are obtained. Therefore, I conducted simulation analysis described in section 4.3

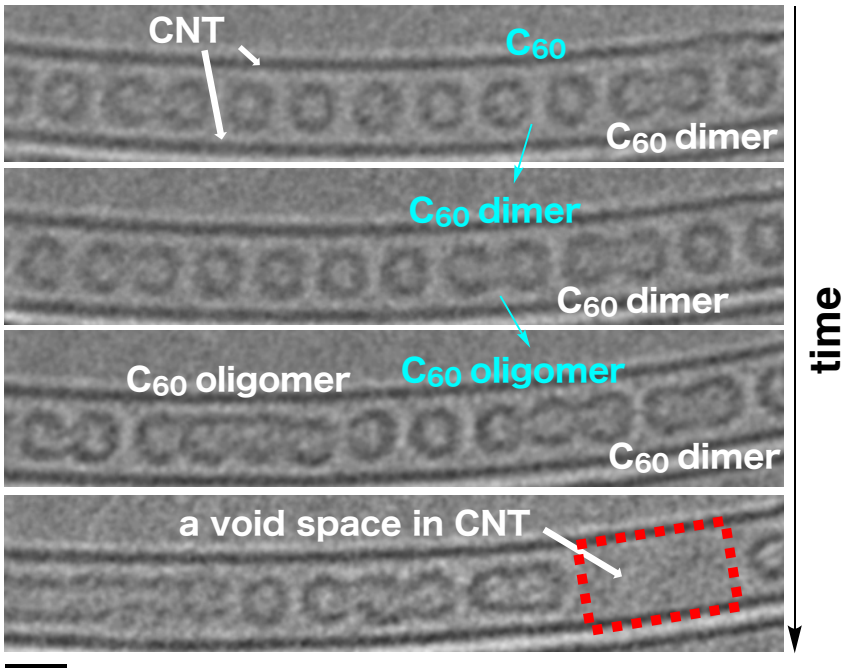

Figure 4.3. Representative TEM images of C_{60} dimerization reaction over time (62.5 ms/frame, 80 kV, EDR = 1.0×10^7 e^- nm^{-2} s^{-1}, 298 K). Red dashed square indicates a void space in a CNT. Scale bar: 1 nm.

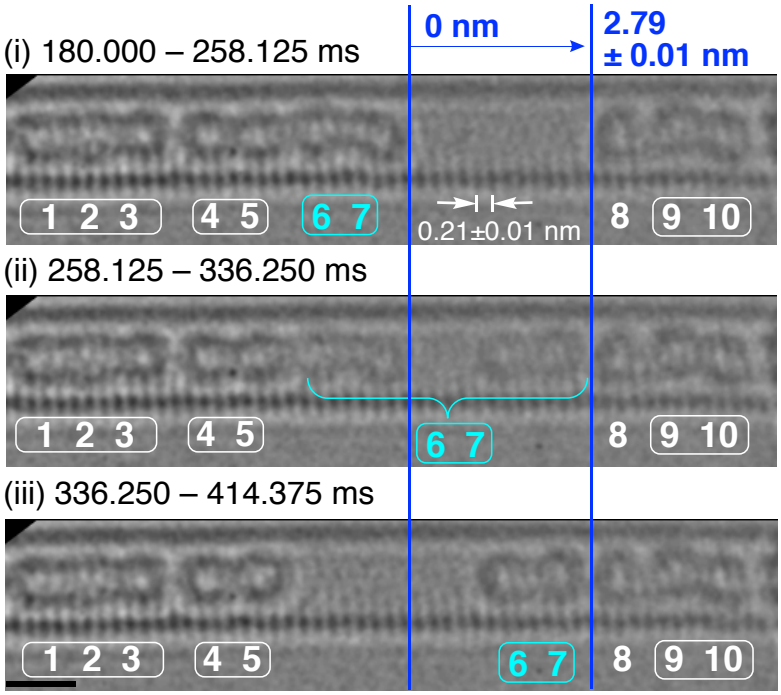

Figure 4.4. SMART-EM video imaging of translation of a C_{60} dimer in CNT (78.125 ms/frame, 80 kV, EDR = 1.4×10^7 e^- nm^{-2} s^{-1}, 298 K). Scale bar: 1 nm. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

4.3 Simulation analysis of CNT and the C_{60} dimer

4.3.1 TEM image simulation analysis of a chiral CNT

TEM simulation images were generated by using a multi-slice procedure implemented in a Bionet elbis software.[13] Parameters for TEM simulation were set to be the same as TEM observation conditions. As a model of a chiral CNT, (15,5) CNT structure with *P* helicity was used considering its dimeter and lattice pattern as shown in Figure 4.5. In TEM images shown in Figure 4.4 or Figure 4.5 (left), the upper sidewall of CNT appears as a straight line and the lower sidewall as vertical pitches with a periodicity of 0.21 ± 0.01 nm (average in 35 images) with a confidence level of 99%. As explained in Chapter 3, from my master course studies,[14] automated cross-correlation image matching software called XMAS was also applied to this work, revealing that CNT was inclined by 15° against the x-y plane (Figure 4.6). By introducing the inclination, vertical fringes with 0.21 nm periodicity appear in the simulation image, derived from the graphitic lattice due to the zigzag chains. The fringe contrast is captured only at the lower sidewall, indicating the absolute stereochemistry of this CNT being *P* helical.

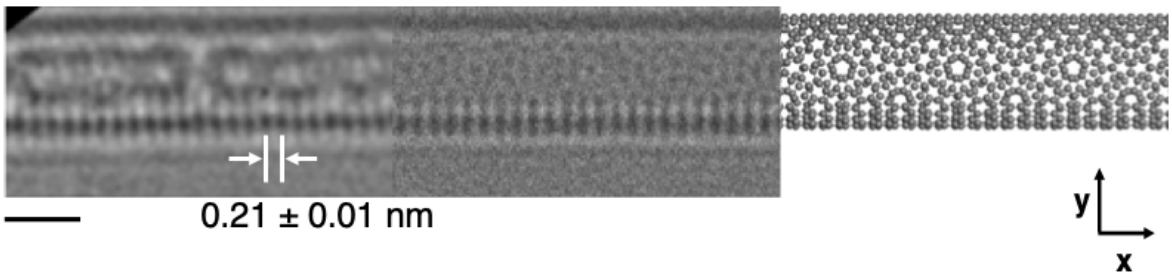

Figure 4.5. Comparison of a TEM image in this Figure 4.4. TEM image (left), a simulation image (middle), and a model corresponding to the simulation image (right) of CNT. Scale bar: 1 nm. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

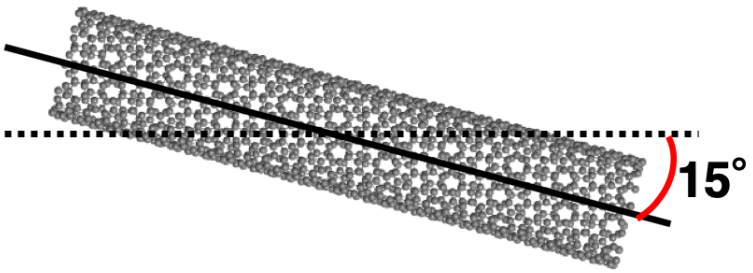

Figure 4.6. A side view of (15,5) CNT possessing P helicity with 15° inclination used for the model and simulation in Figure 4.5. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

4.3.2 TEM image simulation analysis of a C_{60} dimer

The structure of the kidney shaped C_{60} dimer in Figure 4.4 was adapted from previous publications[15,16] followed by geometry optimization on the semi-empirical AM1 level as discussed in Chapter 3 (Figure 4.7). Simulation images from the dimer corresponding to compound #14 in ref 15 or OT-14 in Chapter 3 where a series of intermediates involving the [2 + 2] cycloadduct and Stone-Wales rearrangement product are in agreement with the experimental TEM images. It is important to note that from the simulation analysis, I found that the dimer not only translated but also rotated. In addition, by using XMAS, it is possible to determine the precise position of the dimer without arbitrariness. The position of the dimer (i.e., x in Figure 4.9b) was identified by cross-correlation matching in each frame by using the simulated image as a template image. I defined the standard position (x = 0) as the position of the C_{60} dimer when it was located in the leftmost position in the video. As a result, the dimer translated by 2.79 ± 0.01 nm in the CNT as shown in Figure 4.4.

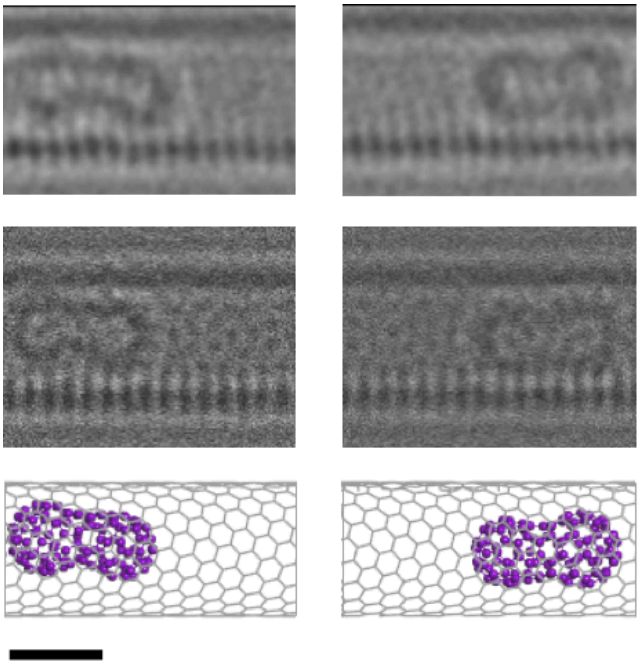

Figure 4.7. Simulation analyses of a C_{60} dimer at CNT. TEM images (top), simulation images (middle), and models (bottom) of the dimer @ (15,5) CNT before (left) and after (right) the translation and rotation shown in Figure 4.4. Scale bar: 1 nm. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

4.4 Millisecond-level analysis of C_{60} dimer motion in a CNT

To obtain high temporal resolution videos, the Chambolle total variation denoising method explained in Chapter 2 was investigated. I obtained a molecular motion of a C_{60} dimer at a frame rate of 4.375 ms/frame, revealing that the dimer (colored in light blue) labeled as **6**–**7** translated in the 297.8125–302.1875 ms frame where the faintly visible dimer can be recognized as illustrated in Figure 4.8. Furthermore, the dimer translation can be recognized even at a fast frame rate of 1.875 ms/frame in Figure 4.9a.

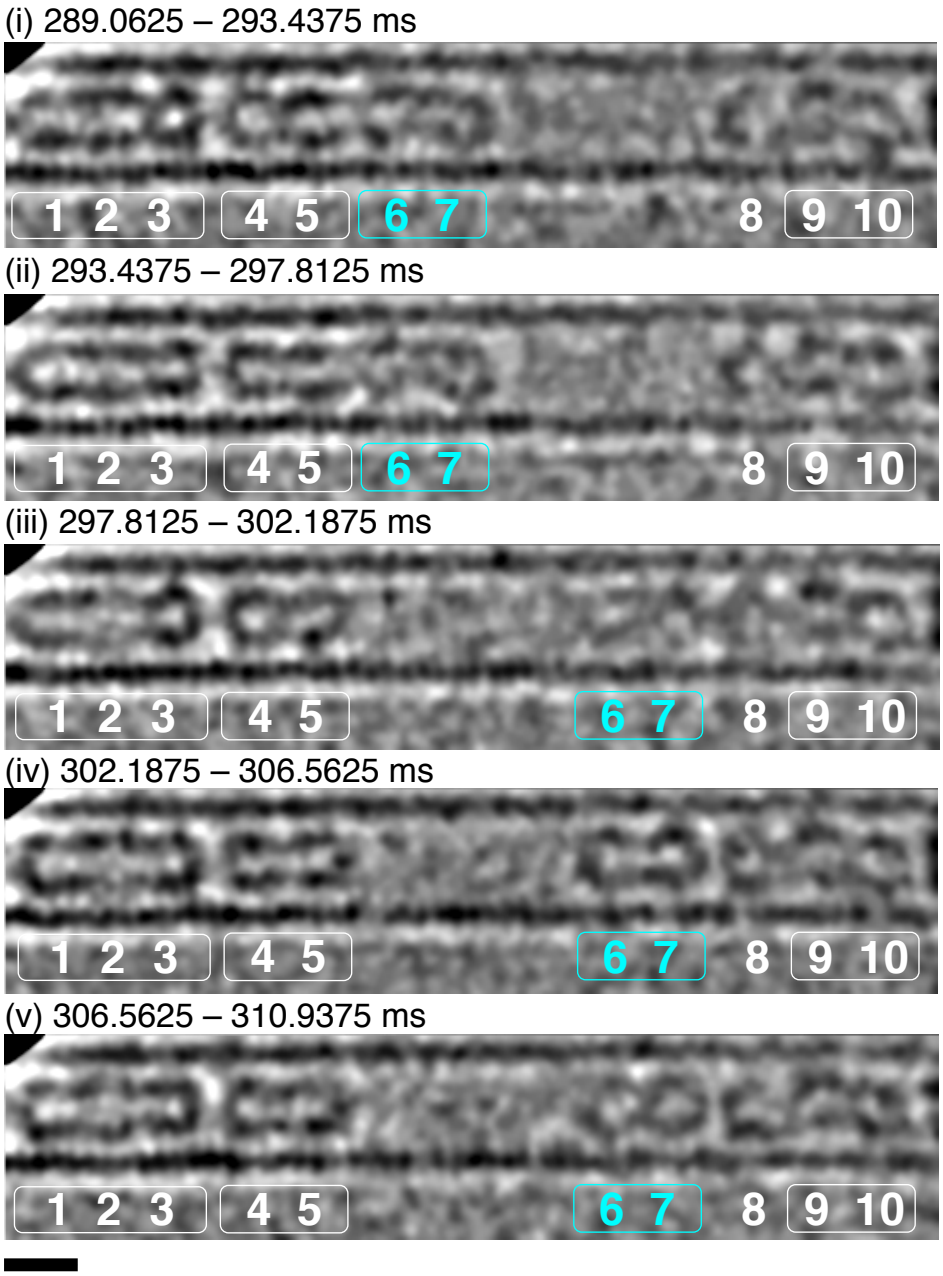

Figure 4.8. SMART-EM video imaging of translation of a C_{60} dimer in CNT (4.375 ms/frame with CTV denoising, 80 kV, EDR = 1.4×10^7 e^- nm^{-2} s^{-1}, 423 K). Scale bar: 1 nm. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

From these high-speed video images of the molecular shuttle of a C_{60} dimer, one question came out: why did the molecules in a CNT translate or rotate? I hypothesized that CNT vibration induces the molecular motion of a C_{60} dimer inside. To prove my hypothesis, the C_{60} dimer motions are quantified with parameters x (translation of the molecule from original position), y (lateral displacement of the center of CNT), and z (CNT diameter) in nm. As mentioned in 4.3.2, x is obtained by the automated cross-correlation matching. On the other hand, the plot profile method (same as Figure 2.11) was used to analyze the vibration and the diameter of CNTs, and details can be found in experimental section of this chapter. In a sequence of frames ii to iv in Figure 4.9a, the dimer **6**–**7** translated between 299.0625 and 300.9375 ms (300 ms ± 0.9375 ms, colored in light blue) until it hit the other C_{60} molecule **8**. Figure 4.9b quantitatively summarizes the dimer motion (x) in blue, the CNT vibration (y) in red, and the diameter of CNT (z) in green by analyzing 6.25 ms/frame images. At first, the CNT vibration event observed in real time at a millisecond and sub-Å level took place stochastically, roughly 10 times during one second with an amplitude of 0.1 nm or less. Importantly, the frequency of the CNT vibration is too small, several Hz, in comparison with the frequency of thermal motions, which has generally an order of 10^6 Hz.[17] It means that the CNT vibration observed with a millisecond order is a mechanical vibration. Here, one of the CNT vibration events caused the translation event of the dimer at 300 ms. As magnified in Figure 4.9c, the translation of the dimer (blue dashed line) took place 10 ms after the CNT vibration at its maximum amplitude at 290 ms (pink solid line). In other words, when the direction of CNT vibration changes from a negative to a positive direction (Figure 4.9b, blue and red, representative TEM images shown in Figure 4.9d), the dimer receives the energy from CNT, initiating the rightward translation. At this point, longitudinal rotation of the dimer **6**–**7** also occurred (cf. Figure 4.7). A schematic potential energy of a C_{60} molecule in CNT can be drawn like Figure 4.10 considering the C_{60} molecule as a particle in a box where the particle has a finite mass.[18] Due to vdW interactions between the dimer **6**–**7** and other C_{60} molecules such as a dimer **4**–**5** and a C_{60} molecule **8**, two stable positions exist in the system, which is a typical potential for causing molecular shuttling motion (cf. Figure 4.1). The dimer should have enough energy to break the vdW contact with its neighbor **4**–**5**, estimated to be ca. 0.3 eV.[19] At the points of (A), (B) and (D) in Figure 4.9b, where the direction of the CNT changed, the dimer did not move. It means that the energy is stochastically given to the dimer. Plus, the diameter of CNT changed little (< ±0.05 nm) during translational motion (green, Figure 4.9b).

In the structural analysis, high spatial precision of molecular distances can be determined by combination of fast SMART-EM imaging and automated cross-correlation image matching. From a 100-distance dataset on x in Figure 4.9b (after time 300 ms), I obtained x = 2.79 nm with a standard error of 0.003 nm with a confidence level of 68%, and x = 2.79 ± 0.01 nm with a confidence level of 99%. From these results, I conclude that in situ observation of the mechanical motions of a molecule coupled with mechanical vibration of a CNT with standard error as small as 0.9 ms in time and 0.01 nm in space has been achieved. Plus, the molecular video recording at a 0.625 ms/frame exceeds by 100 times the previous records of continuous recording of molecular motions.[20]

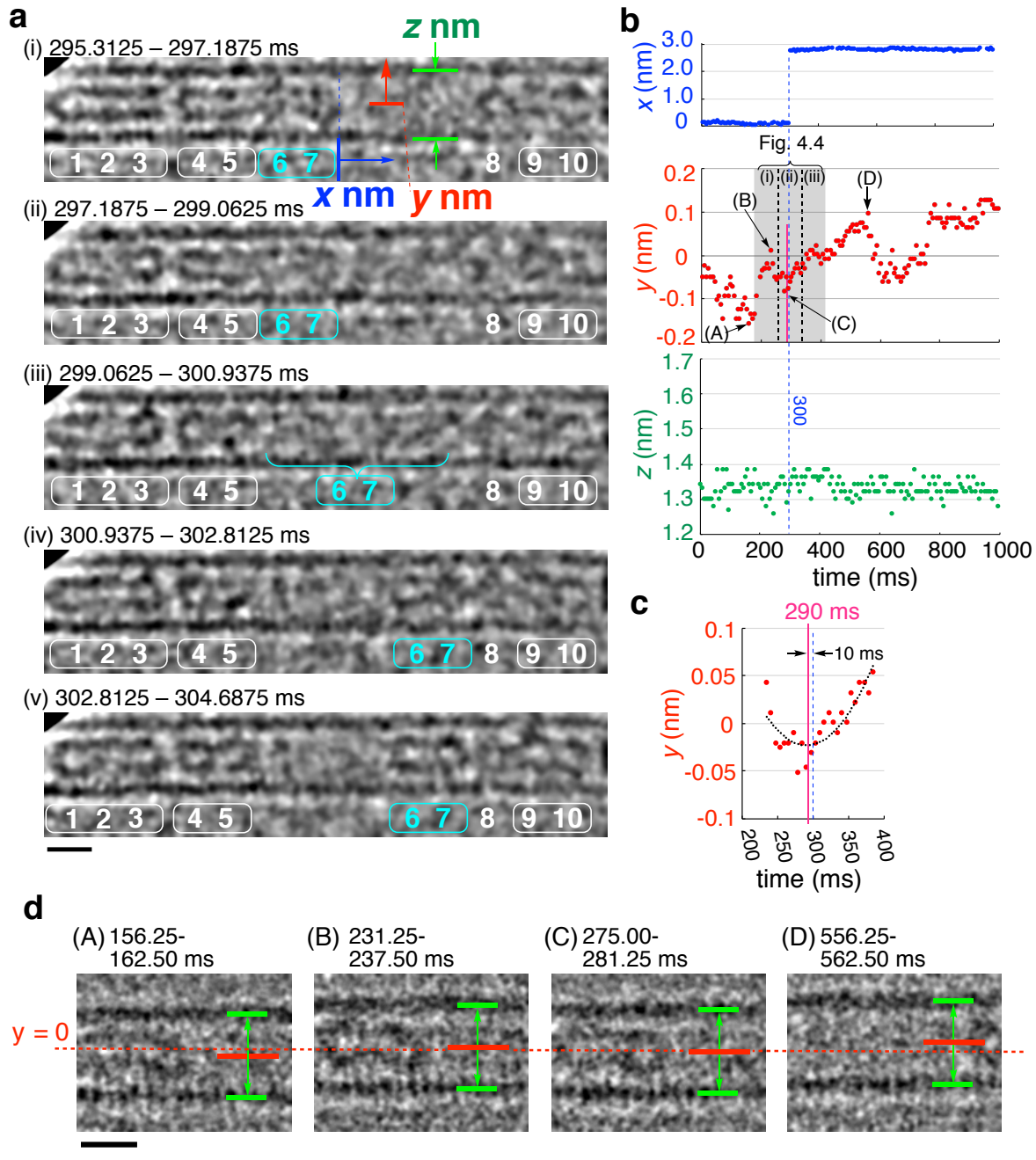

Figure 4.9. SMART-EM video frames showing the motions of a C_{60} dimer in a vibrating CNT. (a) The same dimer translation as Figure 4.8, reanalyzed at 1.875 ms/frame. (80 kV, EDR = 1.4×10^7 e^- nm^{-2} s^{-1}, 298 K). (b) Dimer translation (blue), CNT vibration (red), and deformation (green) analyzed every 6.25 ms. The time frames highlighted in gray corresponds to time frames of Figure 4.4 (i)-(iii). (c) Expansion of 290 ms area of CNT vibration with quadratic function curve fitting, dimer translation (blue dashed), and extremum point of vibration (pink). (d) Four representative frames (A)–(D) in Figure 4.9b (6.25 ms/frame) illustrating the vibration of the CNT. Red bar indicates the center of CNT (y) calculated from the two edges of CNT (light green). Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

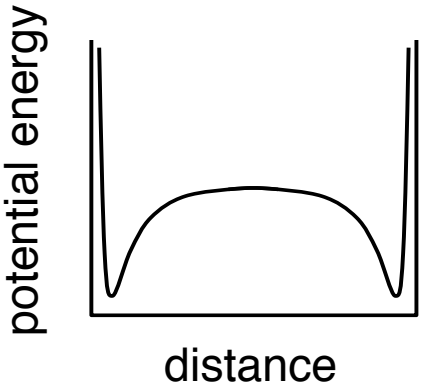

Figure 4.10. A schematic potential energy of a fullerene molecule in CNT. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

4.5 Millisecond-level analysis of C_{60} oligomer motion in a CNT

Figure 4.11 provides further evidence on the coupling of molecular translation with a mechanical vibration of a CNT. As shown in Figure 4.11a, I succeeded in recording a TEM video of C_{60} oligomer, made of at least seven C_{60} molecules (cf. Figure 4.11 iv), shuttled back and forth for three times in a CNT. In each time, molecular rotation was accompanied unidirectionally by the translational motions. Most importantly, it can be concluded that each of the three translation events was triggered by a single CNT vibration by comparing Figure 4.11b in blue and red. From expanded graphs in Figure 4.11c, d, and e, all translation events of C_{60} oligomer (blue dashed lines) took place within 11–40 ms of the time when the absolute displacement of CNT vibration is maximum (pink solid line). In detail, the two rightward translation events occurred when the CNT was changing its direction of vibration along an upwardly convex orbit (Figure 4.11c and e), and one leftward translation event (Figure 4.11d) occurred when the CNT was showing movement of a downwardly convex orbit. It is worth to mention that in case of C_{60} oligomer, the translational motions caused CNT deformation by roughly 20% during the 1.8-s observation time as shown in Figure 4.11b (green). It indicated that the energy which the oligomer possessed was large enough to deform CNT. Namely, a work and energy relationship at a molecular level was revealed by the sub-millisecond sub-Å precision in situ video imaging, which was previously undetected by other time-averaged measurements and microscopy. Furthermore, from the fact that the translation events of the dimer (at 300 ms, Figure 4.9b, green) and the oligomer (at 650 ms, Figure 4.11b, green) took place without change of the CNT diameter, the CNT deformation appeared at 1025 ms and 1519 ms in Figure 4.11b is the result of the translation and not vice versa.

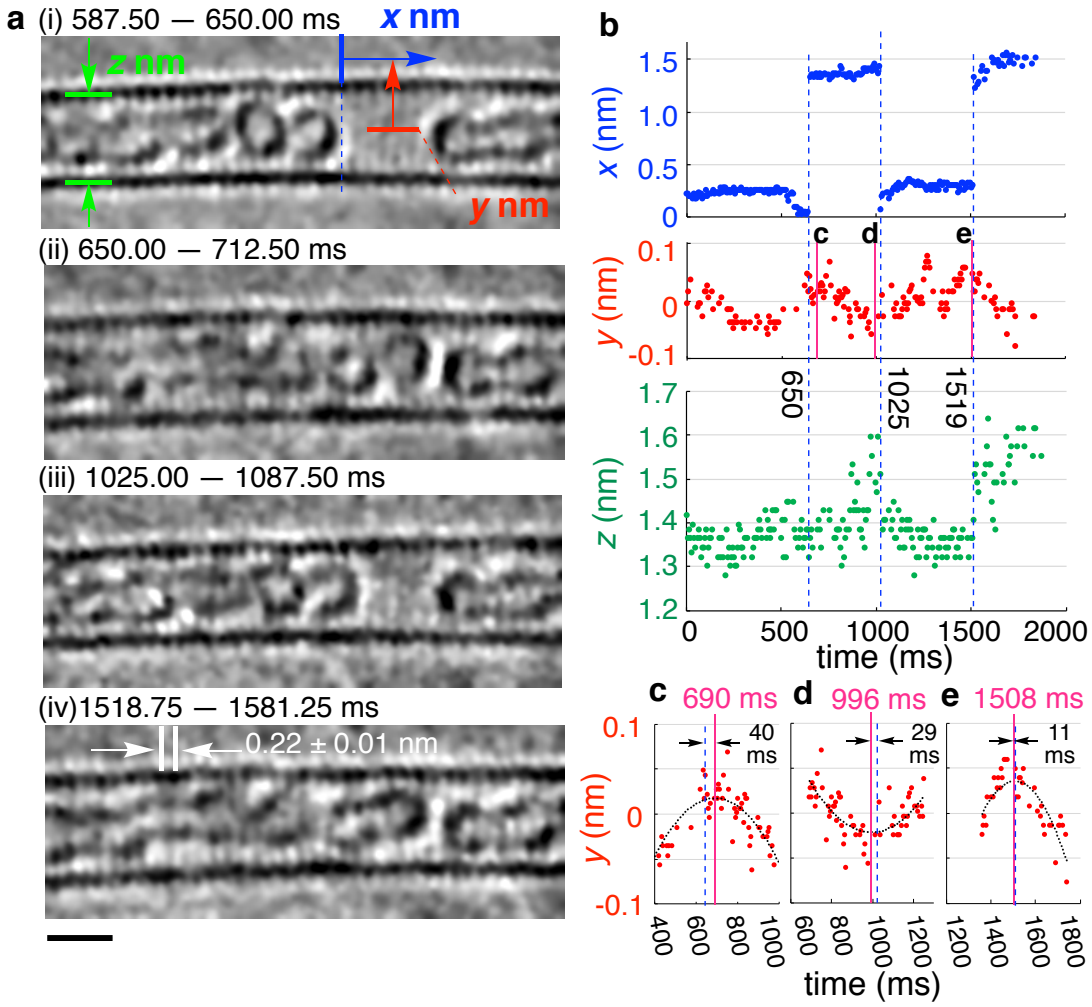

Figure 4.11. SMART-EM video imaging at 16 fps showing the motions of a C_{60} oligomer in a vibrating CNT (80 kV, EDR = 2.3×10^7 e^- nm^{-2} s^{-1}). (a) Oligomer translation at 298 K. See also the experimental section for larger images. The time zero is set to be 9 s after the start of the observation. (b) The *x*, *y*, and *z* values are plotted against time. (c–e) Expansion of 690, 996, and 1508 ms areas of CNT vibration with quadratic function curve fitting, oligomer translation (blue dotted), and extremum point of CNT vibration (pink). Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

From experimental data obtained by real-time video imaging of a single molecular shuttle, several interferences in terms of the mechanics of the molecular shuttles can be proposed. First, the translation events are coupled with the CNT vibration, occurring stochastically a few to several times per minute not only at 298 K but also at 423 K. Through TEM observation, it turns out that the frequency of the molecular shuttling events is insensitive to temperature. It can be interpreted that the key to initiate molecular motions is a large mechanical motion of CNT, which has a direction in the force and not thermal vibration of CNT.

Thermal energy of C_{60} molecules itself might be large enough to make molecular motions but without any directional driving force, there will be no anisotropic movement. Second, the strong correlation between the translation of molecules and the CNT vibration excludes a possibility that the translation was caused by direct electron bombardment to the C_{60} molecule. When I applied a high EDR to the system, molecular shuttling was not initiated but the C_{60} dimerization progressed to fix in one position in a CNT. Third, when C_{60} molecules translate in a CNT, its rotation also occurs. According to theoretical and computational investigations of the motion of a C_{60} molecule over a graphene sheet[21], it is diffusive in the lateral directions. In a CNT case, which is rolled up material of a graphene sheet, a C_{60} molecule might move along the array of its helical structure, causing rotational movement. Forth, there was a correlation between the direction of molecular translation and that of CNT vibration. It also suggests the helical potential of CNT may play a role for the molecular shuttling motion.

4.6 Conclusion

In summary, by combining an atomic-resolution electron microscope, a fast shutter speed camera, and a suitable denoising algorithm for SMART-EM imaging, I succeeded in sub-millisecond sub-Å precision in situ video imaging of a single fullerene molecule shuttling, rotating, and interacting with a vibrating CNT at 0.625 ms/frame. I discovered that the molecule and the CNT container as a whole system behave as a mechanically coupled oscillator. In other words, I elucidated the mechanism of molecular shuttling motion, where the molecular motion is coupled with the mechanical vibration of the CNT container with standard error as small as 0.9 ms in time and 0.01 nm in space. I have revealed rich molecular dynamics, where motions are stochastic and often non-repeatable, and a work and energy relationship at a molecular level previously being difficult to analyze by time-averaged measuring methods. In addition, the molecular video recording at a sub-millisecond level exceeds by 100 times the previous records of continuous recording of molecular motions, while achieving sub-angstrom localization precision. These results also provide an answer to the long-standing question of why the molecules attached to a CNT move so infrequently and stochastically. The sub-millisecond SMART-EM imaging will pave the way to an ultimate goal of the imaging of atoms and molecules undergoing stochastic motions. It will provide us further rich nanoscale dynamics information.

4.7 Experimental section

4.7.1 General

Unless otherwise noted, single-walled CNTs (Meijo Arc SO, produced by arc-discharge using Ni and Y catalysts, >90% purity, average diameter 1.4 nm, Lot # 6601316) were purchased from Meijo Nano Carbon Co. Ltd. C_{60} powder (nanom purple SUH, >99.9% purity, sublimed) was purchased from Frontier Carbon Corporation. EM grids precoated with a lacy carbon (NS-C15 for experiments at 298 K, pore size 1.5–8 μm and carbon thickness 15 nm: RO-C15 for experiments above 298 K; pore size 3–8 μm and carbon thickness 70 nm) were

purchased from Okenshoji Co., Ltd. Toluene (special grade) was purchased from Wako Pure Chemical Industries. Bath sonication was carried out on a Honda Electronics WT-200-M instrument.

4.7.2 Preparation of C_{60}@CNT samples

C_{60}@CNT samples were prepared as follows: CNT powder was heated in air in an oven gradually from 296 to 793 K for 12 min, kept at 793 K for 1 min, heated from 793 to 823 K for 20 min, and kept at 823 K for 20 min to remove the terminal caps of CNTs oxidatively. For encapsulation of C_{60} molecules, the opened CNTs (0.2 mg) and C_{60} powder (0.2 mg) were sealed in a glass tube (Pyrex φ6 mm) under a pressure of 2×10^{-4} Pa and gradually heated from 296 to 573 K over 1 h, then to 673 K over 1 h, and kept at 673 K for 72 h. The resulting C_{60}-containing CNTs were separated mechanically from remaining C_{60} powder, washed with toluene to remove C_{60} from the surface, and dried in vacuum. C_{60}@CNT thus obtained was a black solid (0.3 mg). We dispersed the C_{60}@CNT in toluene (0.05 mg/mL) in a vial in a bath sonicator for 1 h to soften it, so that we could secure good contact between the CNTs and the carbon surface of the grid (essential for temperature control).[2] A 10-μL solution of the dispersion was deposited on a copper grid mesh with a lacy carbon (NS-C15, Okenshoji Co., Ltd.) placed on a paper that absorbs excess toluene. The TEM grid was dried in vacuum (60 Pa) to remove solvent for 2 h. To ensure reproducibility, we used the same sample grid of C_{60}@CNTs in a series of experiments.

4.7.3 SMART-EM imaging

Unless otherwise noted, atomic-resolution TEM observation was carried out on a JEOL JEM-ARM200F instrument equipped with an aberration corrector and at an acceleration voltage of 80 kV, under 1×10^{-5} Pa in the specimen chamber. The microscope is placed on an anti-vibration base plate which in turn is placed on a five-meter-thick concrete bedding to minimize mechanical vibrations. Experiments at 298 K were carried out on a double-tilt holder (JEOL EM-01030RSTH), and those above 298 K were carried out on a heating holder (JEOL EM-21130). The accuracy of the grid temperature is ± a few degrees according to the instrumental specification. To remove volatile impurities from the specimen, the holder was heated at 573 K for 30–60 min without electron irradiation before setting a desired temperature. After the stage temperature settled to the target value, we waited for an additional 30 min to minimize thermal drift.

A series of TEM images were recorded at a rate of 0.625 ms/frame on a direct electron detection camera (Gatan K2-IS, output image size: 414 × 1,920 pixels) with pixel resolution 0.021 nm at × 400,000 magnification. We employed a spherical aberration (Cs) value of 1–3 μm and an electron dose rate of $1.4–2.3 \times 10^7$ e^- nm^{-2} s^{-1} at x 400,000 magnification of the K2-IS camera. All the images were automatically processed on Gatan DigitalMicrograph software.

I first surveyed numerous C_{60} molecules encapsulated in CNTs as described below at × 400,000 magnification on a complementary metal oxide semiconductor camera (Gatan OneView, 4,096 × 4,096 pixels, operated at binning 2 mode) at low EDR ($<10^4$ e^- nm^{-2} s^{-1}) to

find specimen molecules in CNT bundles suitable for video imaging of motions. Upon finding suitable specimens, I stopped beam irradiation by using the beam blanking function and waited for 1 min until thermal drift of the specimen stopped. Then, I set the EDR to the target value, changed the camera to K2-IS, and started the video recording. I kept focus manually on the specimen during the video imaging at under-focus conditions (defocus value: 12 ± 2 nm).

The images were collected in the .dm4 format on Gatan DigitalMicrograph software and processed using ImageJ 1.47t software. To remove the unevenness of the electron irradiation, all images were processed by a bandpass filter (filtering structures smaller than 3 pixels and larger than 40 pixels, tolerance of direction: 5%) and adjusted with brightness and contrast. The orientation of CNTs in all TEM images, simulations and models were rotated so that CNTs were shown to be horizontal in each frame. When applying Chambolle total variation denoising, a weight parameter was determined within the range of 0.2–0.5 from visual aspects.

4.7.4 Vibration and diameter analysis of CNT (measurement of *y* and *z* in Figures 4.9 and 4.11)

I used the plot profile method to analyze the vibration, the diameter, and the periodicity pitch of CNTs, which are seen as two dense lines (A and B in Figure 4.12). For an area (3 nm × 1 nm) shown in a red box (Figure 4.12a), I obtained an intensity profile plot shown in Figure 4.12b. The two local minima in the plot (green lines) are due to the CNT walls and the center of CNT (*y* in Figure 4.9b and 4.11b, red) is defined as the midpoint of the two green lines (red line), and CNT diameter (*z* in Figures 4.9b and 4.11b, green) calculated as the distance between the two green lines. I defined $y = 0$ as the averaged lateral position of the center of CNT over the whole video.

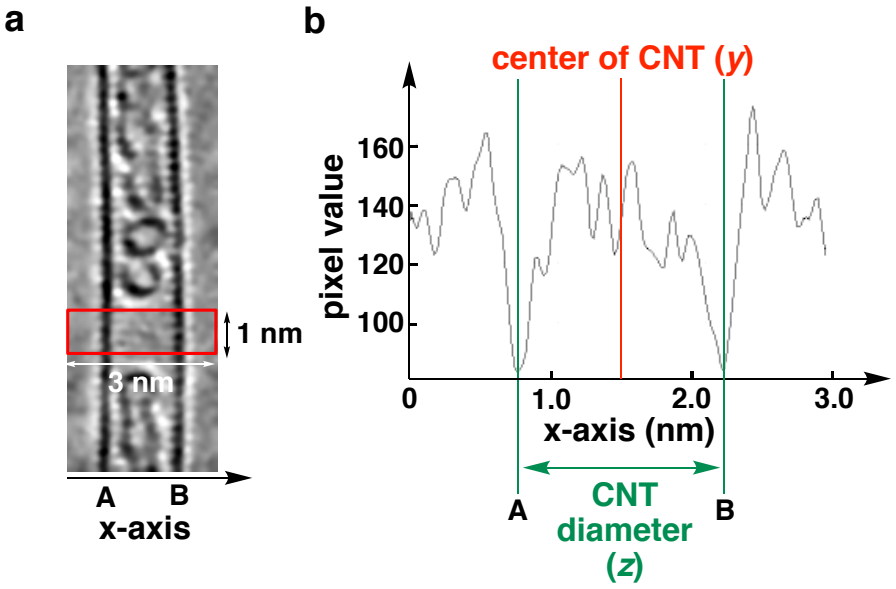

Figure 4.12. Determination of the center and the diameter of CNT. (a) TEM image of C_{60} oligomers in a CNT, where the selected area for vibration and diameter analysis is shown in the red box. (b) The intensity profile along the x-axis. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

Similarly, the periodic contrast of CNT wall originating in the overlap of carbon atoms was obtained to be 0.21 ± 0.01 nm (average in 35 images) with a confidence level of 99% (Figure 4.13), and the pitch corresponds to the graphitic structure in the CNT. Note that appearance of such periodic images is coincidental depending on the chirality and the direction of the CNT relative to the electron beam.

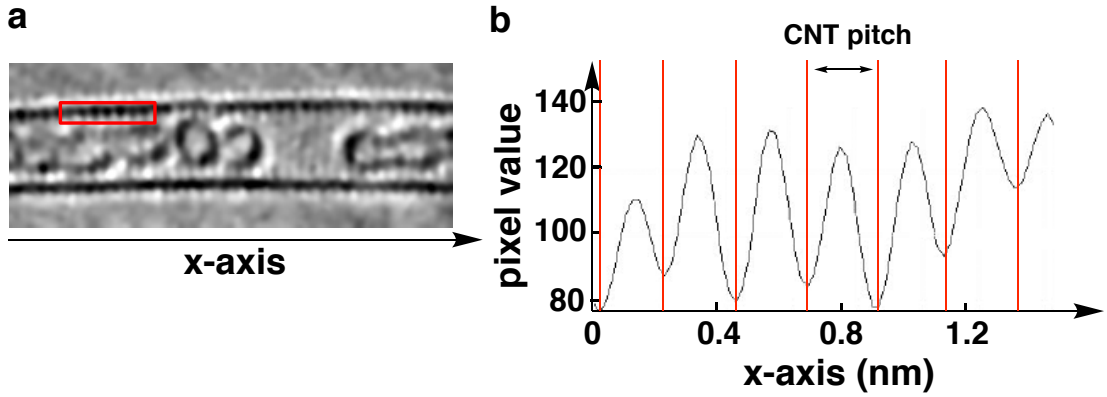

Figure 4.13. Determination of the periodicity observed on the wall of CNT. (a) TEM image of C_{60} oligomers in a CNT, where the selected area for periodic contrast analysis is shown in the red box. (b) The intensity profile along the x-axis. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

4.7.5 Detailed SMART-EM video frames of molecular shuttles

In Figures 4.14–4.17, SMART-EM video frames of a C_{60} molecules (a dimer: Figure 4.14, an oligomer: Figures 4.15–4.17) showing translational motion in CNT at 298 K are summarized. In most of the images, the denoising method discussed in Chapter 2 was applied.

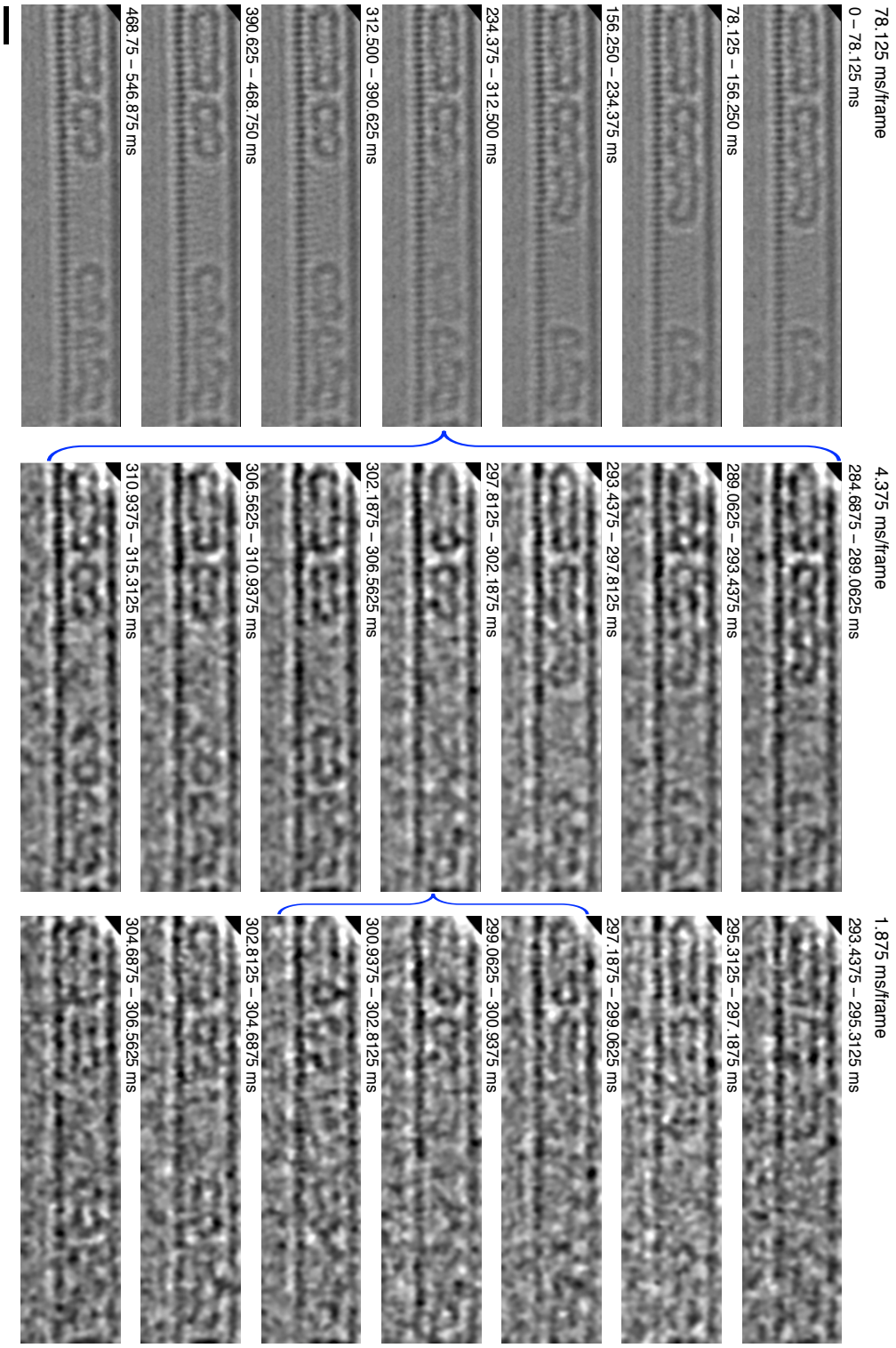

Figure 4.14. SMART-EM video frames (78.125 ms/frame without denoising, 4.375 and 1.875 ms/frame with denoising) of a C_{60} dimer showing translational motion in CNT at 298 K. Scale bar, 1 nm. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

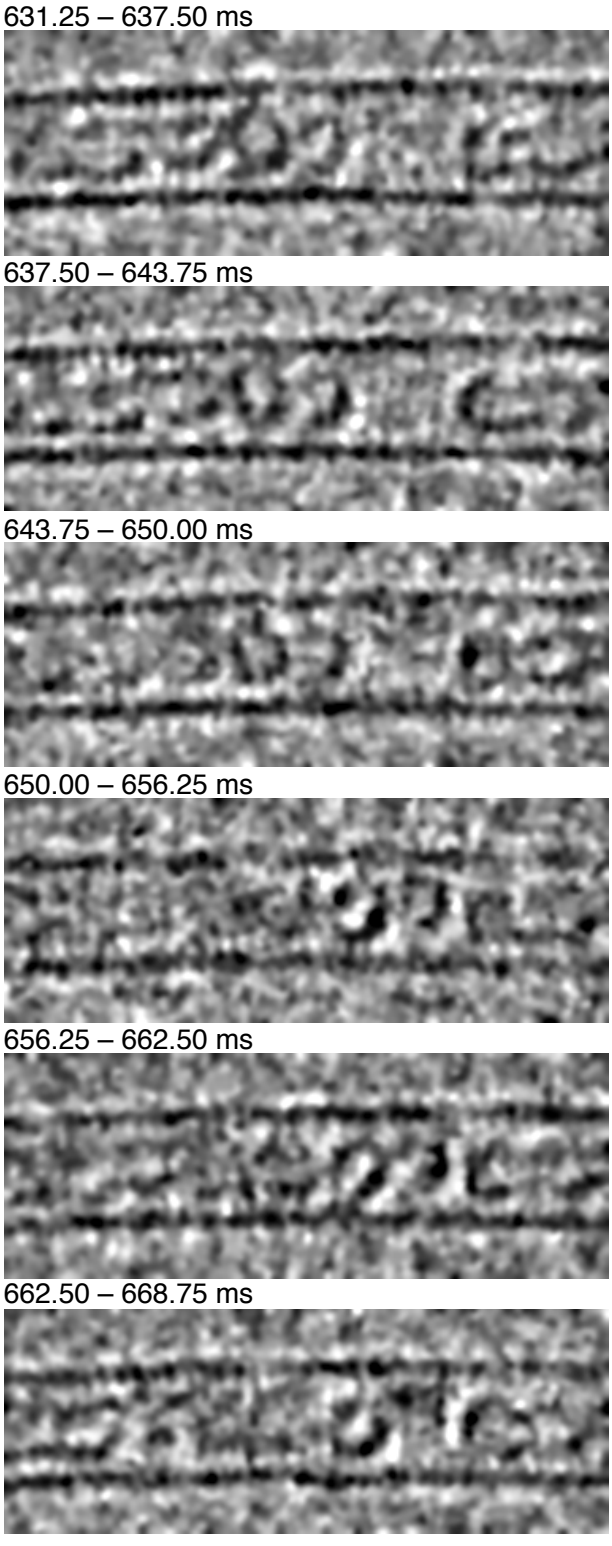

Figure 4.15. SMART-EM video frames of a C_{60} oligomer (6.25 ms/frame with denoising) showing translational motion corresponding to the analysis shown in Figure 4.11c. Numbers denote the time in milliseconds after starting video recording. Scale bar, 1 nm. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

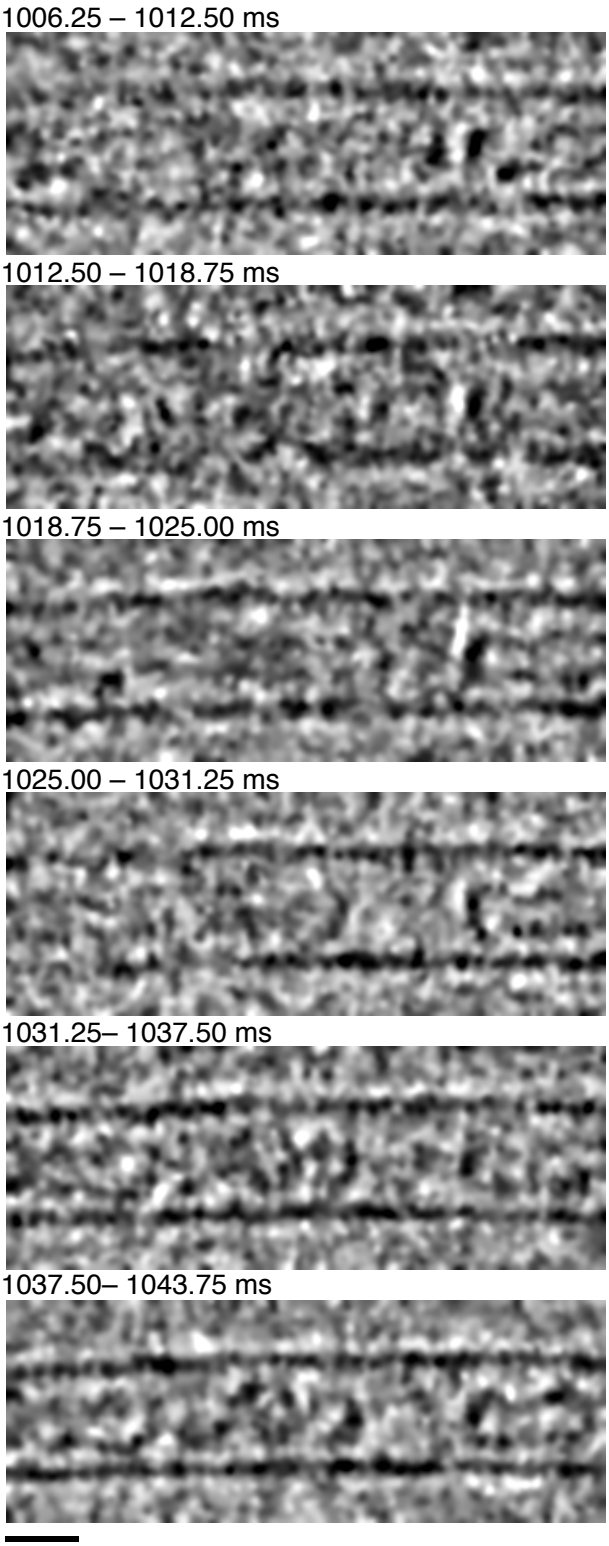

Figure 4.16. SMART-EM video frames of a C_{60} oligomer (6.25 ms/frame with denoising) showing translational motion corresponding to the analysis shown in Figure 4.11d. Numbers denote the time in milliseconds after starting video recording. Scale bar, 1 nm. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

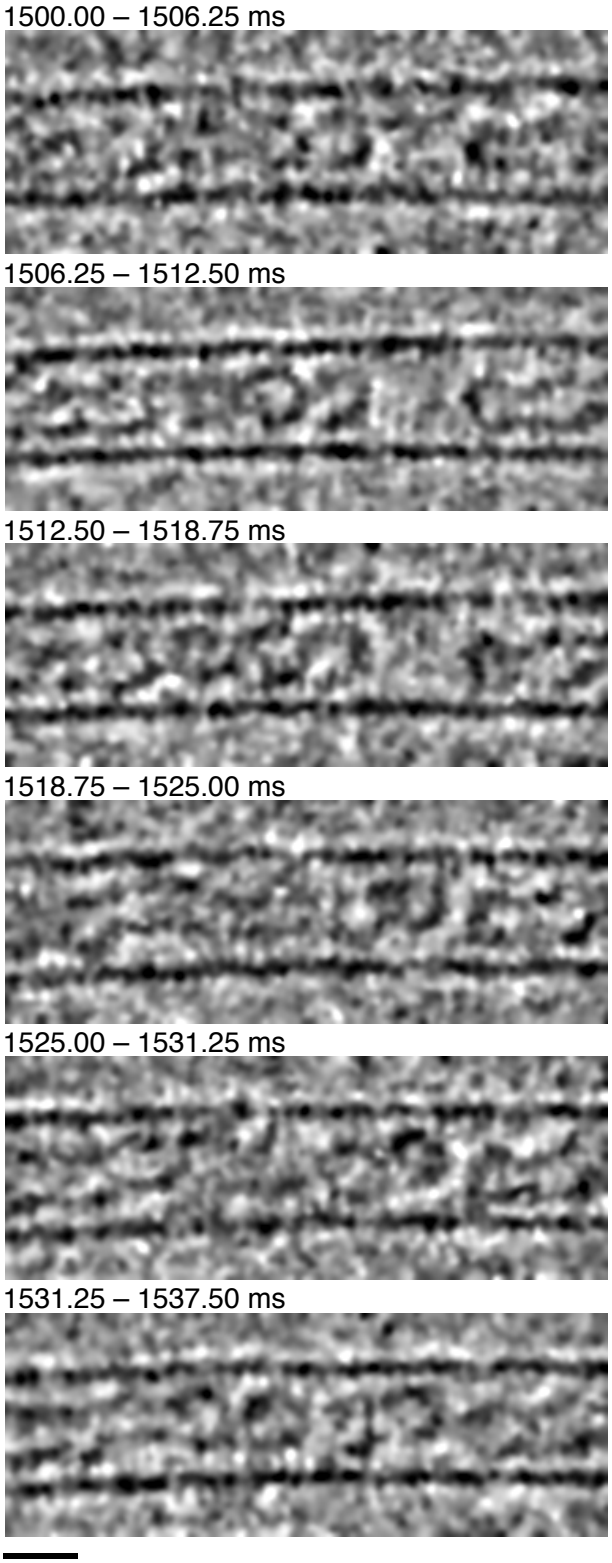

Figure 4.17. SMART-EM video frames of a C_{60} oligomer (6.25 ms/frame with denoising) showing translational motion corresponding to the analysis shown in Figure 4.11e. Numbers denote the time in milliseconds after starting video recording. Scale bar, 1 nm. Adapted with permission from ref.11. Copyright 2020 The Chemical Society of Japan.

4.8. References

[1] V. Richards, *Nat. Chem*. **2016**, *8*, 1090.
[2] P.L Anelli, N. Spencer, J. F. Stoddart, *J. Am. Chem. Soc.* **1991,** *113*, 5131–5133.
[3] S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan, A. L. Nussbaumer, *Chem. Rev*. **2015**, *115*, 10081–10206.
[4] T. Kudernac, N. Ruangsupapichat, M. Parschau, B. Maciá, N. Katsonis, R. S. Harutyunyan, K. Ernst, B. L. Feringa, *Nature* **2011**, *479*, 208–211.
[5] R. A. Bissell, E. Cordova, A. E. Kaifer, J. F. Stoddart, *Nature* **1994**, *369*, 133−137.
[6] V. Balzani, M. Clemente-Leon, A. Credi, B. Ferrer, M. Venturi, A. H. Flood, J. F. Stoddart, *Proc. Natl. Acad. Sci. U. S. A*. **2006**, *103*, 1178–1183.
[7] S. A. Vignon, T. Jarrosson, T. Iijima, H. R. Tseng, J. K. M. Sanders, J. F. Stoddart, *J. Am. Chem. Soc*. **2004**, *126*, 9884−9885.
[8] B. W. Smith, M. Monthioux, D. E. Luzzi, *Chem. Phys. Lett.* **1999**, *315*, 31–36.
[9] H. Somoda, K. Hirahara, S. Akita, Y. Nakayama, *Nano. Lett.* **2009**, *9*, 62–65.
[10] J. H. Warner, Y. Ito, M. H. Rümmeli, T. Gemming, B. Büchner, H. Shinohara, G. A. D. *Briggs, Phys. Rev. Lett.* **2009**, *102*, 1995504.
[11] T. Shimizu, D. Lungerich, J. Stuckner, M. Murayama, K. Harano, E. Nakamura, *Bull. Chem. Soc. Jpn.* **2020**, *93*, 1079–1085.
[12] M. Sotomayor, K. Schulten, *Science* **2007**, *316*, 1144–1148.
[13] F. Hosokawa, T. Shinkawa, Y. Arai, T. Sannomiya, *Ultramicroscopy*, **2015**, *158*, 56-64.
[14] T. Shimizu, *Master Thesis*, **2018**.
[15] S. Han, M. Yoon, S. Berber, N. Park, E. Osawa, J. Ihm, D. Tománek, *Phys. Rev. B* **2004**, *70*, 113402.
[16] M. Koshino, Y. Niimi, E. Nakamura, H. Kataura, T. Okazaki, K. Suenaga, S. Iijima, *Nat. Chem.* **2010**, *2*, 117–124.
[17] A. W. Barnard, M. Zhang, G. S. Wiederhecker, M. Lipson, P. L. McEuen, *Nature* **2019**, *566*, 89–93.
[18] J. W. Kang, H. J. Hwang, *J. Phys. Soc. Jpn.* **2004**, *73*, 1077–1081.
[19] C. Ginard, P. Lambin, A. Dereux, A. A. Lucas, *Phys. Rev. B*. **1994**, *49*, 11425–11432.
[20] J. H. Warner, Y. Ito, M. H. Ruemmeli, B. Buechner, H. Shinohara, G. A. D. Briggs, *ACS Nano* **2009**, *3*, 3037.
[21] M. Neek-Amal, N. Abedpour, S. N. Rasuli, A. Naji, M. R. Ejtehadi, *Phys. Rev. E* **2010**, *82*, 051605.

－Chapter 5－

Summary and Perspectives

In the present thesis, by combining an atomic-resolution electron microscope, a fast shutter speed camera, and the optimized denoising algorithm for electron microscopy imaging, I have succeeded in observing and analyzing in situ individual nanomechanical events such as a single molecular shuttle coupled with mechanical vibration of the CNT container, as well as chemical reactions of individual molecules at maximum sub-millisecond and sub-angstrom resolution.

In Chapter 2, I investigated various noise reduction methods for TEM images with low signal to noise ratio (SNR) values and concluded that the most appropriate denoising method for electron microscopy molecular imaging is Chambolle total variation denoising algorism with a collaborator, Dr. Joshua Stuckner. While the Chambolle denoising preserves molecular edges of C_{60} molecules, which is indispensable for precise size and distance measurements, noises are reduced to give a good SNR, improving the image contrast high. A fast shutter speed camera gives noisy image (Figure 5.1a) due to less electron dose irradiated to the specimen, but by applying the Chambolle denoising, I obtained clear molecular images suitable for structural analyses (0.625 ms/frame image in Figure 5.1b and 1.875 ms/frame image in Figure 5.1c) with keeping a time resolution of a millisecond. With this denoising method, I have successfully achieved a temporal precision of 0.9 ms and a localization precision of 0.01 nm for imaging of non-periodic molecular structures such as C_{60} molecules in a CNT. (Figure 5.1d). Sub-millisecond-level time-resolution electron microscopy images described here are the highest time resolution with electron microscopy so far for imaging of dynamic molecular events.

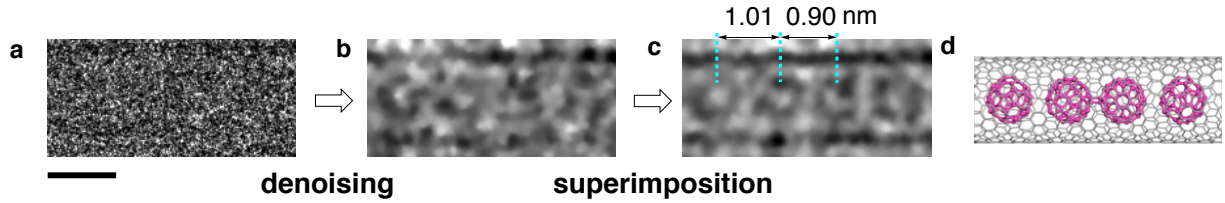

Figure 5.1. Fast video imaging of C_{60} molecules with Chambolle total variation denoising and superimposition. (a) A single-frame image of C_{60} molecules at 0.625 ms/frame (1600 fps) without any image processing and (b) with Chambolle denoising. (c) A three-frame superimposed image (1.875 ms/ frame) with Chambolle denoising. (d) Corresponding molecular model of C_{60} molecules in a CNT. Scale bar: 1 nm.

Chapter 3

本章については, 5年以内に雑誌等で刊行予定のため，非公開.

第３章については, 5 年以内に雑誌等で刊行予定のため，非公開.

In Chapter 4, I elucidated the mechanism of a nanomechanical behavior of a molecular shuttle by a recording of the 1600 fps video of a single molecular shuttle. From a series of TEM images of an oligomer moving back and forth for three times (Figure 5.3a), I found that the translation of the molecule occurred when the direction of the CNT vibration changes (Figure 5.3b and c), suggesting that the motion of molecules inside was induced by receiving an energy from the vibrating CNT. In other words, the molecule and the CNT container as a whole behave as a mechanically coupled oscillator, where the molecular motion is coupled with the mechanical motion of the CNT.

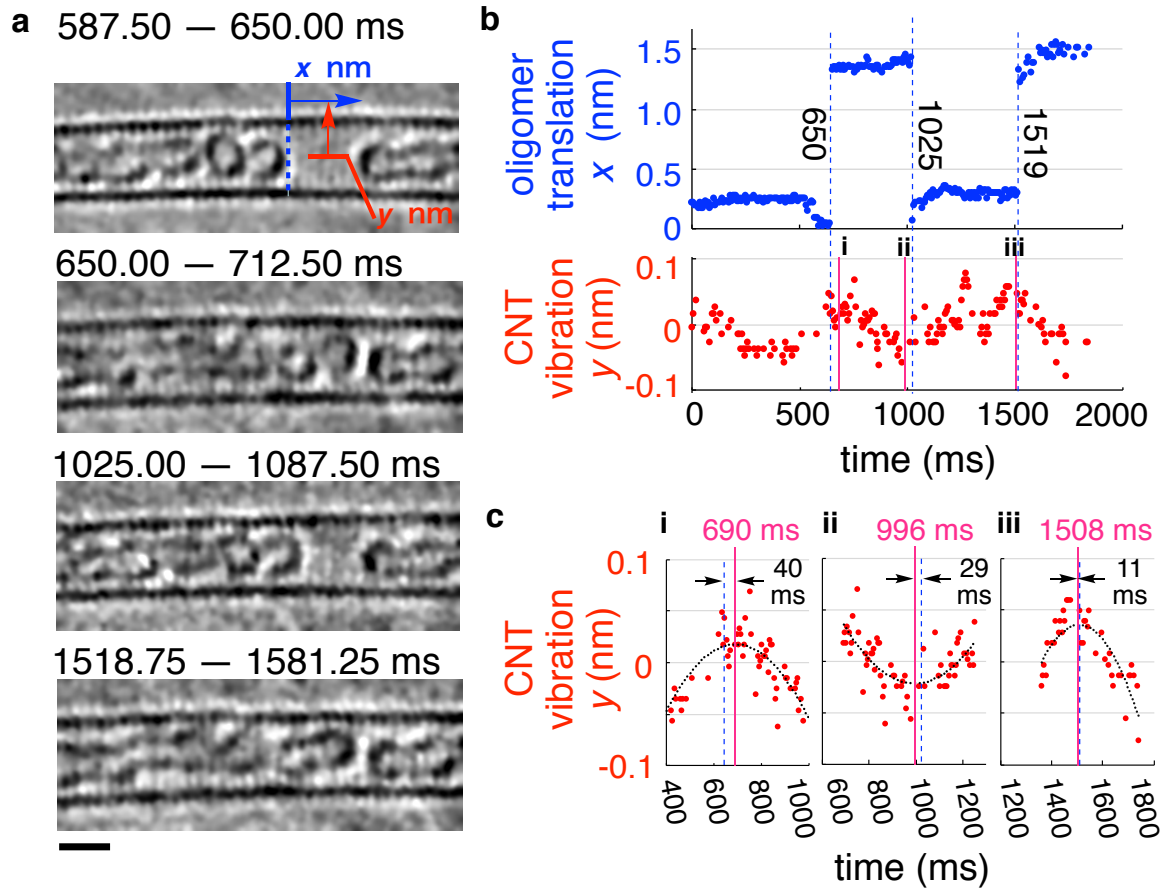

Figure 5.3. SMART-EM video frames showing the motions of a C_{60} oligomer in a vibrating CNT. (a) TEM imaging and (b) distance the analysis of shuttling C_{60} oligomer in a CNT. (c) Expansion of 690, 996, and 1508 ms areas of CNT vibration. Scale bar: 1 nm.

In this thesis, by developing the sub-millisecond electron microscopic imaging method, I explored the molecular world of sub-millisecond, which has never been observed before by other conventional methodologies. The sub-millisecond SMART-EM technique will open up a new field of stochastic dynamics of single molecules, elucidating the mechanism of molecular conformational changes and the equilibrium state of chemical reactions, which are more fundamental molecular behaviors. It is expected to be applied for the study of various scientific phenomena from materials science to life science, which until now could only be conducted in theoretical calculations. It will also give a great impact on the education of students because unlike boring texts or simple figures, molecular videos are thought to be more motivating students. There is a saying that “seeing is believing”. I strongly believe what I observed with sub-millisecond SMART-EM imaging will innovate the world and open up the science of the future.