

1 Introduction on Single-Molecule Science

Krishnarao Appasani and Raghu K. Appasani

Classical chemistry and biochemistry experiments in solution measure the properties of many molecules and/or interrogate them simultaneously – these are called *ensemble measurements* and tend to mask the underlying molecular dynamics. Studies at *single-molecule level* provide random, stochastic dynamics, and allow access to an incredible wealth of molecular information. Most importantly, previously “unanswerable” questions in the physical, chemical, and biological sciences can now be answered. The field of single-molecule science (SMS) can be roughly divided into two general areas: (1) improvements in single-molecule methodologies (technology development); and (2) use of these methodologies to address important scientific questions in fundamental biological research (applied research). Over the past decades, single-molecule research has fostered excellent collaboration and interdisciplinary research with input from biology, chemistry, and physics.

Scientists are progressively aiming to discover and describe the secrets in the index of the book of life. We all know that life is a dynamic process unfolding in three dimensions, and there still remains the daunting task of fully describing the spatiotemporal location and conformation of all these indexed components, as well as their complex interactions. In order to understand such complex molecular interactions, a novel technological platform is needed. Techniques such as gene or protein chips, phage display, X-ray diffraction, nuclear magnetic resonance (NMR), and several others allow researchers to uncover part of this information; however, they lack needed sensitivity to unravel the details of individual molecular events.

1.1 The Birth of Single-Molecule Science

The initial attempt to probe individual fluorescent molecules is probably due to Jean Perrin almost a century ago (Perrin, 1918). However, E. A. Syngé proposed an experimental scheme allowing in principle the performance of nanometer resolution microscopy in 1928, thus predating near-field optical microscopy by over 50 years. In 1948, Theodor Förster published a theoretical paper quantitatively

describing the probability that light absorbed by one molecule could be transferred to another molecule and be subsequently emitted (Förster, 1948). Possibly the first measurement of enzymatic activity at the single-molecule level was the observation by Rotman in the 1960s of fluorescent reaction products generated by a single β -galactosidase enzyme acting on a substrate analogue (Rotman, 1961). In the 1970s, Hirschfeld detected a single antibody, albeit labeled with ~ 80 fluorophores (Hirschfeld, 1976). At IBM/Bell Labs, Arthur Ashkin had discovered optical tweezers (Ashkin, 1970), a development that earned him the Nobel Prize for Physics in 2018; he found that a tightly focused laser beam could be used to trap and move micron-sized particles, but their application to biological problems took a few years. The first single-molecule measurements were first successfully performed with nonoptical methods such as the patch clamp technique, which was developed by Neher and Sakmann. This technique allowed recording of ion translocation through a single ionic channel embedded in a cell membrane (Neher and Sakmann, 1976). Later on in the early 1980s, a scanning tunneling microscope was developed by Gerd Binnig and Heinrich Rohrer at IBM Research, Zurich (Binnig and Rohrer, 1982), a development that also earned them the Nobel Prize for Physics in 1986. Binnig et al. (1986) described the first atomic force microscopy (AFM) in 1986. Since then, the applications of these single-molecule methods in solving biochemical and molecular biological problems have been growing rapidly.

Subsequently, few optical methods reached the sensitivity required to detect, image, manipulate, and follow the spectroscopic evolution of single fluorophores on surfaces, in solids at low temperatures (Moerner and Kador, 1989; Orrit and Bernard, 1990). These methods also helped to detect signals directly inside live cells and to image individual DNA molecules in water (Lindsay et al., 1988, 1989; Engel, 1991) and gliding filament and optical tweezers measurements of individual kinesin movement along microtubules (Howard et al., 1989; Block et al., 1990). The early years of the field resulted in several existing developments and improvements in a variety of fluorescence and manipulation methods, and these methods were generally applied to probe simpler chemical and biological systems. In this context, near-field scanning optical microscopy was developed, and its implementation with a simplified confocal optical microscope opened the field and made single-molecule fluorescence measurements more widely accessible and imaging was translated from low to room temperature (Betzig and Chichester, 1993).

Steven Chu, who later used optical tweezers to manipulate DNA, shared the Nobel Prize in 1997 for trapping individual gas-phase molecules with the tweezers (Perkins et al., 1994). Then, in the 1990s, methods were also developed to observe individual molecules in solution. First, fluorescence from a single dissolved molecule (Funatsu et al., 1997) was detected; this was quickly followed by measurement of fluorescence resonance energy transfer (FRET) between two molecules (Ha et al., 1996); the observation of single-molecule events inside living cells (Mashanov et al., 2003); and recently, super-resolution *in vivo* imaging beyond the fundamental limit imposed by the diffraction of light waves as demonstrated by Stefan Hell (Hell, 2007; Huang

et al., 2010). In a nutshell, for this groundbreaking discovery Moerner, Betzig, and Hell were awarded the Nobel Prize in Chemistry in the year 2014, “for their work on developing super-resolution fluorescence microscopy.”

1.2 Applications of Single-Molecule Techniques in Biology

Fluorescence microscopy coupled with green fluorescent proteins brought a *green revolution* culminating in single fluorescent protein molecule detection. In addition, *super-resolution fluorescence microscopy* technique in live cells became a routine and cutting-edge tool to reveal new insights into long-standing puzzles in biology (Shashkova and Leake, 2017). Single-molecule techniques are not able to provide atomic resolution pictures of large molecules such as proteins; their strengths lie in probing detailed structural distributions, as well as real-time and stochastic dynamics. Scope and applications of single-molecule science in various fields have been summarized in Figure 1.1. Few general areas

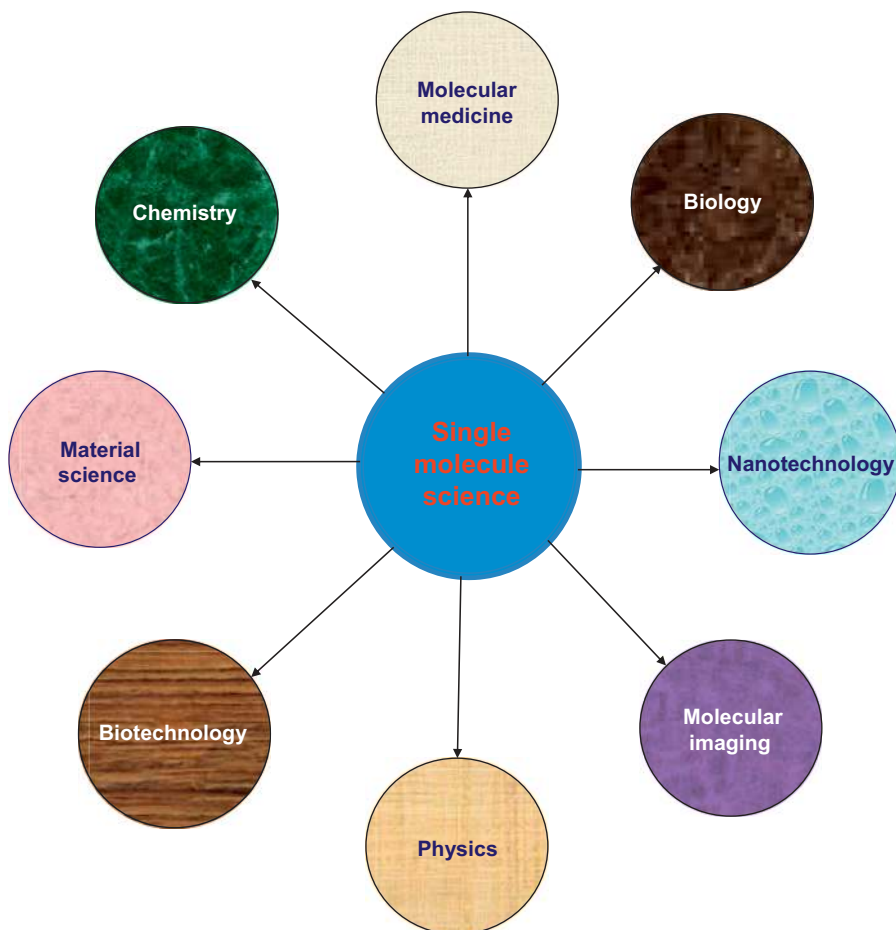


Figure 1.1 Applications of single-molecule science in various disciplines.

will be discussed in four different parts in this volume, therefore a brief introduction is provided in this section.

Protein folding: The question of how an unfolded polypeptide chain is converted into a natively folded and functional protein continues to fascinate biologists, chemists, and physicists. Conceptual breakthroughs and technological developments have flourished in this area during the past four decades. Early views of protein folding relied on theoretical descriptions developed for small-molecule reactions, involving one or a small number of specific folding pathways and intermediates (Kim and Baldwin, 1982).

Biological assemblies and functions: In addition to being high-sensitivity probes of structural properties of individual biological molecules, single-molecule methods are also well suited for studying the assembly of molecular complexes and their resulting function. One important area uniquely amenable to single-molecule investigations is how conformational changes are coupled with binding. An early example came from the study of a minimal RNA junction that binds a protein S15 from the 30S small ribosomal subunit (Ha et al., 1999). Ha et al. used smFRET to deduce that protein binding to the junction was accompanied by a large conformational change in the RNA junction.

Single-molecule enzymology: Soon after, Xie and coworkers reported a pioneering study of the activity of the single-molecule cholesterol oxidase enzymes that catalyze the oxidation of cholesterol by oxygen (Lu et al., 1998), taking advantage of the intrinsically fluorescent cofactor, flavin adenine dinucleotide (FAD).

Study of motor proteins: Systems studied include the rotary F_1 ATPase and synthase; cytoskeletal proteins such as kinesins, dyneins, and myosins; viral DNA-packaging motors; RNA and DNA polymerases; helicases; and chromatin-remodeling proteins (Bustamante et al., 2004; Kinosita et al., 2004; Mallik and Gros, 2004). More recently, single-molecule studies described the DNA packaging and ejection methods in greater depth (Chemla et al., 2005; Hugel et al., 2007). Two complementary studies used single-molecule FRET on diffusing molecules and magnetic tweezers to understand mechanistic features of abortive initiation by RNA polymerase (Kapanidis et al., 2006; Revvakin et al., 2006), a process where as an initial transcribing complex, it undergoes several abortive cycles of short RNA synthesis and release. Several single-molecule studies have examined the activity of RNA polymerase, providing information about its binding, activity, and the transcription process in general, and a couple of recent ones are mentioned here (Zlatanova et al., 2006).

Study of molecular nanomachines: Molecular machines in biological cells encompass an enormous range of function, including fuel generation for chemical processes, transport of molecular components within the cell, cellular mobility, signal transduction, and the replication of the genetic code, among many others. Today, researchers are tackling the challenges of studying molecular/nanomachines in their native environments. Splicing is carried out by a cellular nanomachine called the spliceosome that is composed of RNA components and dozens of proteins. Despite decades of study, many fundamentals of the

Introduction on Single-Molecule Science

7

spliceosome function have remained elusive. Single-molecule detection methods enabled investigators to monitor the dynamics of specific splicing components, whole spliceosomes, and even cotranscriptional splicing within living cells (DeHaven et al., 2016).

It is now possible to watch ribosomes in action and their structural rearrangements as they take each step necessary to transform a nucleotide sequence into a protein. Single-molecule methods were successfully applied to study protein synthesis in greater detail. A major class of single-molecule studies of protein synthesis has investigated the kinetics of the ribosome as it progresses through multiple rounds of the elongation cycle; this has been accomplished using both laser tweezers-based, single-molecule manipulation studies and fluorescence-based, single-molecule detection studies. Single-molecule studies of protein synthesis provide the dynamic information necessary to link the structural snapshots of ribosomal complexes provided by X-ray and cryo- electron microscopy structures into a real-time “movie” of how synthesis proceeds (Aitken et al., 2010; Frank and Gonzalaz, 2010). A major class of single-molecule studies of protein synthesis is smFRET studies probing the structural dynamics of the translating ribosome. The primary objective of these studies has been to characterize the conformational dynamics that underlie mRNA-tRNA complex translocation through the ribosome. Cryo-electron microscopy is a powerful technique that could also help to unravel the secrets of ribosomal machinery and its associated protein’s conformational dynamics (Frank and Agarwal, 2000).

1.3 Epigenetic Regulation

Chromatin assembly dynamics and remodeling are involved in virtually all chromosomal processes and have been the subjects of numerous single-molecule, single-fibre studies (Ladoux et al., 2000; Brower-Toland et al., 2002; Zlatanova and Leuba, 2003). Yet another recent study of remodeling enzymes by optical trapping discovered the formation of large intranucleosomal loops, which may be involved both in making nucleosomal DNA accessible and in nucleosomal mobility and eviction (Zhang et al., 2006). Although distinct epigenetic marks correlate with different chromatin states, how they are integrated within single nucleosomes to generate combinatorial signals remains largely unknown. Chen et al. (2012) have reported the successful implementation of single-molecule tools constituting fluorescence correlation spectroscopy (FCS), pulse interleaved excitation-based Förster resonance energy transfer (PIE-FRET), and fluorescence lifetime imaging-based FRET (FLIM-FRET) to elucidate the composition of single nucleosomes containing histone variant H2A.Z. Quantification of epigenetic modifications in nucleosomes has provided a new dimension to epigenetics research and led to a better understanding of how these patterns contribute to the targeting of chromatin-binding proteins and chromatin structure during gene regulation.

Studies of cell surfaces using near-field nanoscopy: Living cells use surface molecules such as receptors and sensors to acquire information

about and to respond to their environments. The cell surface machinery regulates many essential cellular processes, including cell adhesion, tissue development, cellular communication, inflammation, tumor metastasis, and microbial infection. These events often involve multimolecular interactions occurring on a nanometer scale and at very high molecular concentrations. Therefore, understanding how single molecules localize, assemble, and interact on the surface of living cells is an important challenge and a difficult one to address because of the lack of high-resolution single-molecule imaging techniques. However, use of AFM and near-field scanning optical microscopy (NSOM) now make it possible to provide unprecedented possibilities for mapping the distribution of single molecules on the surfaces of cells with nanometer spatial resolution, thereby shedding new light on their highly sophisticated functions.

Single molecule science, which brought a revolution in biology, can be defined as “the ability to observe and track individual molecules and monitor molecular interactions in living cells.” This new field is the resultant of fruitful collaboration among biologists, physicists, and engineers that feeds off the ability to develop new microscopes and other technology to probe single molecules and analyze their behavior in intact cells. It is quite interesting to mention that this field has emerged from developing, using, and combining technologies such as single-molecule fluorescence and super-resolution microscopy, atomic force microscopy, optical and magnetic tweezers, new sensors, and more sophisticated computing and mathematical techniques. The new frontier of single-molecule science presents us with a new way of understanding life processes differs from the bottom-up approaches. Specially, it helps to recognize the behavior and dynamics of molecules in life that give us a handle to better know the cellular architecture and molecular events (such as DNA replication, repair, transcription, translation, signal transduction, protein transport, cell division) in disease biology that ultimately helps to discover new class of drugs.

These groundbreaking discoveries paved the way for innovative techniques such as atomic force microscopy, optical tweezers, tethered particle motion, and magnetic tweezers. Additionally, fluorescence imaging (including confocal microscopy, total-internal-reflection fluorescence [TIRF], and fluorescence imaging with one nanometer accuracy techniques) has helped to visualize molecules in motion in live cells. In recent years, other imaging techniques, including super-resolution microscopies (photoactivated localization microscopy [PALM], stochastic optical reconstruction microscopy [STORM], and stimulated emission depletion [STED] microscopy), have helped biologists achieve spatial resolutions that are better than 50 nanometers (nm). Continued progress in investigating biomolecular systems of increasing complexity will be driven primarily by emerging technological breakthroughs in our abilities to observe and manipulate single molecules.

All of these studies and techniques have enabled us to study biological molecules at the single-molecule level and have led to the launch of the new field of *single-molecule science* or *single-molecule biology*.

We realized that there was the need for a book on this topic for the following reasons:

- (1) In recent years, super-resolution techniques such as PALM, STORM, and STED are being applied to understand biological events at both the single-molecule and/or single-cell level, and this has generated a tremendous interest in academic research laboratories around the world in understanding their applications in molecular biology and disease biology and the potential for developing new instrumentation and diagnostic tools.
- (2) Several hundred research papers have already been published in top journals, including *Nature*, *Nature Biotechnology*, *Nature Methods*, *Nature Physics*, *Biophysical Journal*, *Science*, *Cell*, *PNAS*, and many others, indicating that the use of single-molecule-based techniques in biology is growing at an exponential rate within academia, biopharmaceutical, and photonics industries.
- (3) Based on this technology, many biotech companies, such as Oxford Nanopore Technologies, Ltd. UK, and Pacific Biosciences, USA, have started (using the original investigator's proprietary patent rights) to produce novel tools, reagents, and instruments. Very recently, the large bioanalytical giant Thermo Fisher Scientific has made multimillion dollar agreements with many countries to sequence the DNA sample of human populations using single-molecule biology-based DNA sequencing. Several diagnostic companies (23 and Me, USA; Roche) are also involved in sequencing human samples for diagnostic purposes.

We have carefully selected the chapters, written by experts in the field, and have divided the chapters into appropriate parts to support the theme expressed in the subtitle of this book: *From Super Resolution Microscopy, Molecular Imaging to DNA Mapping and Diagnostics*. Developing *novel molecular imaging tools* is likely to become a prerequisite for *live-image studies* used to understand the inner world of the biological cells. The elegant and revolutionary single-molecule approaches that are covered in this book will undoubtedly have great future commercial promise for the development of innovative molecular image scanning procedures and for the study of biomedicine and human disease biology.

1.4 Scope of This Book

The chapters offered in the present volume discuss insights that have been revealed about their mechanism, structure, or function by single-molecule techniques. Many topics were covered in this text, such as enzymes, motor proteins, membrane channels, DNA, and other key molecules of current interest. An introduction by the editors provides a brief review of key principles with an historical overview. The last part on future perspectives discusses latest applications in relevance to molecular diagnostics and drug discovery.

The text consists of nine chapters, grouped into four parts, starting with the technology leading to descriptions of the applications, as follows:

1.4.1 Part I: Super-Resolution Microscopy and Molecular Imaging Techniques to Probe Biology

This part consists of three chapters.

In the last decade substantial improvements in the area of super-resolution techniques allowed to resolve targets in the 20 nm range, and allowed us to enable precise localization of single molecules. As Herbomel and Patterson have described in the second chapter, that the improvement of super-resolution microscopy in the last decade has led to the development of methods such as STED microscopy and structured illumination microscopy (SIM), which modulate the excitation light to break the diffraction limit, or methods such as PALM or STORM, which utilize the photo-switching properties of fluorescent molecules to enable precise localization of single molecules. Especially, this chapter focuses on PALM/STORM methods, which rely on similar principles and instrumentation for acquiring a series of images of fields of single molecules followed by subsequent image analysis to localize the molecules to much higher precision than their diffraction limited signals.

The third chapter by Gunkel et al. describes high-throughput and high-content screening, single-molecule localization microscopy, and multiscale imaging. Specially, an integration of fully automated targeted single-molecule localization microscopy (SMLM) into a screening platform is presented in order to achieve targeted microscopy and automated super-resolution imaging on tissue microarrays. This chapter also highlights on the application of total internal reflection fluorescence microscopy (TIRF) microscopy to potentiate plasmid-based clustered regularly interspaced short palindromic repeats (CRISPR) usage for screening, which is the most promising high-throughput gene editing technology platform that is directly applied in human or mouse cell culture systems. DNA optical mapping, which has emerged in the 1990s as an alternative approach to optically obtain genetic and epigenetic information from single DNA molecules, allows analyzing large native genomic fragments at single-molecule resolution. Optical mapping has the potential to be an ideal tool for metagenomics studies. The use of fluorescent microscopy allows obtaining several types of information from a single molecule. In the fourth chapter, applicative studies on three different types of optical mapping-based assays, resulting in new types of data with clinical, environmental, and scientific benefits, are well presented in this book by Grunwald et al.

1.4.2 Part II: Protein Folding, Structure, Confirmation, and Dynamics

This part consists of two chapters.

Proteins in biological systems perform a different and highly specialized role within the living cell, determined by its three-dimensional structure, composition, mechanics, and dynamics. Understanding this role is an important task for biophysicists. Very few experimental techniques are available and able to access

information about the structure and dynamics of the individual elements and substructures of protein molecules. Such information is summarized in Part II. Protein nanomachines can be thought of as being built like man-made machines: static parts form a basic scaffold for movable parts whose motion is fuelled by energy source, *e.g.* thermal motion, ligand binding, concentration gradients or chemical reactions.

Using single molecule force spectroscopy, we are now able to identify the basic parts of these “machines”, *i.e.*, functional mechanical elements of proteins. One such technique to understand protein dynamics is single-molecule force spectroscopy by optical trapping, a method for which Arthur Ashkin won the Nobel Prize in Physics in 2018. Using this technique, individual conformational dynamics of protein substructures can be observed in real time, enabling a huge variety of fascinating insights into protein function to be gained. This can be seen from the example of several proteins recently studied using single-molecule force spectroscopy. In the fifth chapter, Zoldák and Tych give the details of how such high-resolution optical tweezers can be used to understand the dynamics of several proteins, including adenylyl cyclase, heat shock protein 70 and 90, and AAA+ proteases.

Protein secretion is a very relevant process because more than 30 percent of synthesized proteins work in organelles or outside the cells. In Chapter 6, Wilson and his group from the University of Chile have described the posttranslational, protein translocation through membranes at the single-molecule level. His group has determined the role of essential accessory protein in translocation such as BiP and Sec 61. Although these studies allowed understanding translocation, the mechanochemistry of such transport mediated by BiP and SecA is still unknown.

1.4.3 Part III: Mapping DNA Molecules at the Single-Molecule Level

This part consists of two chapters.

Biomolecules and biopolymers undergo conformational transitions during many biological processes. For example, some proteins are observed to have multiple intermediate states in the folding/unfolding pathways, intrinsically disordered proteins can form diverse metastable structures, functional proteins can often be switched between active and inactive states through conformational transitions, and nucleosomes are able to regulate DNA unwrapping through their conformational transitions. These dynamic states of DNA and proteins control their biological functions. Since force plays a fundamental role in many, if not all, biological systems, one way to reveal the dynamics of the molecules is to elucidate its intra- and intermolecular force, which can be used as a marker to capture information about their conformational changes.

Single-molecule techniques to measure forces in these biological systems, for example, AFM, optical tweezers, and magnetic tweezers, provide a direct measurement on the force at the single-molecule level. These techniques are versatile for studying various biological processes such as ligand-receptor binding, nucleic

acids unzipping, and protein folding. Therefore, in Chapter 7, Ching-Hwa Kiang's group from Rice University has focused on using AFM in order to understand the dynamic states and pathways of conformational transitions of biological systems, thereby enabling construction of the free energy landscapes of the processes. The atomic force microscope is a key member of a series of scanning probe microscopes that allows the retrieval of local information from a surface by sensing the force operating between a specific point on the surface and a sharp probe. Atomic force microscopy, and AFM-based force spectroscopy in particular, with all the recent advancements in these techniques, is on the verge of addressing the various issues in biological systems. In Chapter 8, Joon Won Park and his colleagues from Pohang University have narrated the use of AFM to show its potential applicability as a diagnostic tool. As an example, they have summarized the use of this tool for the quantification of trace amounts of a DNA biomarker without amplification or labeling.

1.4.4 Part IV: Single-Molecule Biology to Study Gene Expression

This part consists of one chapter.

Determining rules for gene expression regulation is an important step toward predicting how cells are decoding the genome sequence to create a wide variety of phenotypes. Recent advances in imaging technologies revealed the stochastic nature of gene expression, in which different numbers of mRNA and protein molecules can be created in cells that have the same genome sequence. Characterizing and predicting heterogeneity of gene expression in single cells is a key approach to reveal mechanisms involved in stochastic gene expression, but it requires multiple fields of research. To experimentally detect stochastic behaviors of gene expression, advanced fluorescence imaging methods is essential. Especially imaging methods that have sensitivity of single mRNA or protein are crucial to measure stochasticity at any abundance, including Poissonian behaviors that emerge due to the relatively small number of mRNAs and proteins. In Chapter 9, Kumar et al. have nicely reviewed various approaches toward understanding stochastic gene expression from these multiple viewpoints.

Future Perspectives: Over the past decade, new single-cell and single-molecule analysis tools have led to advances that isolate single cells, technologies that can assay each cell's DNA, RNA, proteins, and metabolites, and imaging tools that map cell contents and their molecular interactions. These tools promise new insight on the differences in function between individual cells and molecules; the organization and timing of responses to stimuli; how cells interact as components of a complex system; and how these interactions may change with age, disease, and exposure to environmental stressors. Correspondingly, we will continue to see researchers from a variety of different disciplines make important contributions to the field in the coming years. At the molecular level, the ultimate challenge of a single-molecule experiment would be to acquire global structural time trajectories of single molecules or complexes with atomic and picoseconds resolution, while simultaneously having the ability to perturb